Proceedings of the 27th IEEE International Symposium
on Robot and Human Interactive Communication,
Nanjing, China, August 27-31, 2018

Behavior Explanation as Intention Signaling in Human-Robot Teaming

Ze Gong and Yu Zhang

***Abstract*— Facilitating a shared team understanding is an important task in human-robot teaming. In order to achieve efficient collaboration between the human and robot, it requires not only the robot to understand what the human is doing, but also the robot's behavior be understood by (a.k.a. explainable to) the human. While most prior work has focused on the first aspect, the latter has also begun to draw significant attention. We propose an approach to explaining robot behavior as intention signaling using natural language sentences. In contrast to recent approaches to generating explicable and legible plans, intention signaling does not require the robot to deviate from its optimal plan; neither does it require humans to update their knowledge as generally required for explanation generation.**

The key questions to be answered here for intention signaling are the *what* (content of signaling) and *when* (timing). Based on our prior work, we formulate human interpreting robot actions as a labeling process to be learned. To capture the dependencies between the interpretation of robot actions that are far apart, skip-chain Conditional Random Fields (CRFs) are used. The answers to the *when* and *what* can then be converted to an inference problem in the skip-chain CRFs. Potential timings and content of signaling are explored by fixing the labels of certain actions in the CRF model; the configuration that maximizes the underlying probability of being able to associate a label with the remaining actions, which reflects the human's understanding of the robot's plan, is returned for signaling. For evaluation, we construct a synthetic domain to verify that intention signaling can help achieve better teaming by reducing criticism on robot behavior that may appear undesirable but is otherwise required, e.g., due to information asymmetry that results in misinterpretation. We use Amazon Mechanical Turk (MTurk) to assess robot behavior with two settings (i.e., with and without signaling). Results show that our approach achieves the desired effect of creating more explainable robot behavior.

I. Introduction

With the ever fast development of AI technologies, robots are becoming pervasive in our daily life. We are beginning to see applications of robots in various areas that span household, industries, military, etc. However, making humans and robots to team together remains a challenge. To enable effective human-robot teaming, it is important to maintain a shared team understanding of each other. Towards this goal, it requires not only the robot to understand what the human is doing (so that the robot may choose to assist when necessary), but also the robot's behavior be understood by the human. To address the latter aspect, the robot must be able to behave in a comprehensible way. Otherwise, it would lead to the loss of teaming effectiveness over time, and eventually human trust entirely.

Ze Gong and Yu Zhang are with the Computer Science and Engineering Department at Arizona State University, {zgong11, yzhan442}@asu.edu

There exists research work on making robot behavior comprehensible to humans. Generating legible motions [5], [6], unambiguous natural language sentences [21], [10], and visual projections [1], [22], are a few examples for generating explicit or implicit cues to communicate the robot's intention to the human. For task planning that involves more complex domains, the situation is more complicated since now the differences between the human and robot domain models (i.e., knowledge about the domain dynamics and configurations) must also be taken into account. Subject to such model discrepancies, Zhang *et al.* [24], [25] proposed two metrics that can be used by the robot to generate more explicable plans to the human. When generating such plans is too costly, Chakraborti *et al.* [3] formulated the explanation generation problem for a robot to explain its behavior as a "model reconciliation problem". The goal of an explanation is to make the robot's plan optimal (and hence explicable) according to an updated human model, and explanations are constructed from the updates.

Prior methods on behavior and explanation generation, however, have to either change the robot's plan to make it less optimal but simultaneously more comprehensible, or require modifications to the human knowledge. For situations where such comprehensible plans are too costly, and explanations too complex to understand, they may not always be effective. In such cases, a robot may choose to signal its intention. Intention signaling can be performed by using natural language sentences or visual projections [21], [1]. These prior methods, however, either rely on fixed timings or manually constructed signals and hence fail to address the fundamental questions of *when* and *what* to signal. The simplest method to signal at every action is obviously inefficient, and can significantly increase human cognitive load. In this paper, we propose *intention signaling* as an approach that enables a robot to convey its intention by signaling only when applicable and with what is necessary.

Based as our prior work [25], we formulate human interpreting robot actions as a labeling process. To capture the dependencies between the interpretation of robot actions that are far apart, e.g., when human interpretation of robot actions may be dependent on future context, in contrast to [25], we use skip-chain Conditional Random Fields (CRFs) [19] to learn this process. The two key questions, *when* and *what* to signal during the planning process, can then be formulated as an inference problem with skip-chain CRFs, while maximizing the probability that the human can associate labels to robot actions or in other words interpret the robot's plan, the very information captured by the CRF model.

978-1-5386-7980-7/18/$31.00 ©2018 IEEE

This optimization can be performed by searching through the possible signaling timing and content and fixing the corresponding labels of the robot's actions in the CRF model. Afterwards, natural language sentences can be generated by using predefined templates for intention signaling. We evaluate our approach using a synthetic robot maid domain on Amazon Mechanical Turk (MTurk). Results show that intention signaling effectively reduces human criticism on robot behavior during human-robot teaming.

II. Related Work

It is well known that effective human teaming requires a level of transparency between the teammates [4]. Such transparency enables the team members to estimate and project team status and thus allowing them to maintain situation awareness. This is expected in human-robot teams as well. There are different ways to achieve this. For relatively simple and repetitive tasks, or when sufficient training is allowed, team members can be trained to anticipate others' actions based on a fixed set of interactive scenarios in a given domain [17], [14]. Although this approach is commonly used in human-human teams, it is rather inflexible if not outright infeasible. Another approach is to model the joint behavior altogether [7]. This approach assumes that the joint behavior model can implicitly capture the influence of team member's behaviors on each other and converge to the team behavior model, which is neither always necessary nor possible.

In the more general scenario, this level of transparency between teammates must be maintained in two directions, and each must be explicitly considered. In one direction, the robot must be able to infer information about the human; at the same time, the robot must ensure that the human knows enough information about itself. Plan recognition [15] and human modeling [23] methods are examples of work in the first direction. The other direction has also started to draw significant attention [12], [2], [13]. Several methods have been developed to make the robot's behavior more legible and understandable. For example, robots can generate legible motions [6], [11], [5] and natural language sentences [21], [10], [9], to either implicitly or explicitly explain its behavior during human-robot interaction. Beside natural languages, researchers have also used visual information [20] to improve the interactions, including with Virtual Reality (VR) [16] and Augmented Reality (AR) [1], [22].

More recently, it was pointed out that behavior explanation is influenced by information asymmetry that may lead to misunderstanding between the team members. Zhang *et al.* [24], [25] introduced two metrics to generate explicable plans subject to such asymmetries. In situations where explicable plans are too costly, Chakraborti *et al.* [3] formulated the explanation generation problem as a model reconciliation planning problem. The differences between the models of the robot and human are explored to inform the construction of explanations for the robot's behavior. A balance can be achieved between the two methods [18] to trade off between the costs of explicable plans and making explanations. Our work in this paper provides an alternative method for robot to explain its behavior.

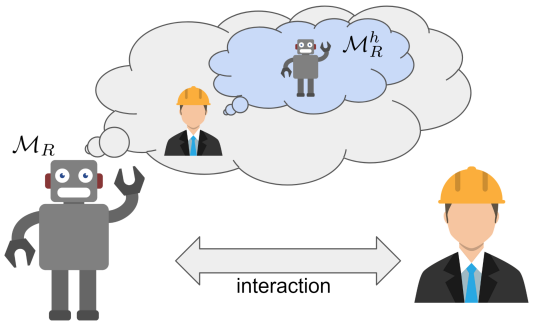


Fig. 1: In human-robot teaming, when the robot generates an optimal plan according to its own model $\mathcal{M}_R$, the plan could be inexplicable to the human if it doesn't match the human's expectation which is based on $\mathcal{M}_R^h$.

III. Intention Signaling

In a human-robot teaming task, given the initial state $\mathcal{I}$ and goal state $\mathcal{G}$, the robot will generate a plan $\pi_{\mathcal{M}_R}$ using its own model $\mathcal{M}_R$. We assume that the robot's planning model is defined in PDDL [8]. The human will have expectations of the robot's behavior based on his understanding of the robot's model, $\mathcal{M}_R^h$ (see Figure 1). During the plan execution, some actions may not be explicable to the human when the robot's behavior doesn't match the human's expectations, even though they contribute to achieving the goal in the robot's model. In order to make its plan explicable, we propose for the robot to signal its intention before executing the inexplicable actions. To achieve this, first of all, the robot must be able to estimate how the plan would be interpreted by the human. Having this estimated human's interpretation of its own plan, the robot can then infer about *when* and *what* to signal its intention to improve this interpretation. This summarizes the two key problems to be solved in this work.

A. Background of Plan Explicability

In our prior work [24], [25], we proposed a metric of plan explicability that captures how the human would interpret the robot's plan. It was assumed that the human understands the robot's plan by assigning each action to a task label. The label for each action is treated as a hidden variable. The sequential labeling process is modeled by a linear-chain CRFs.

Using the learned model of the labeling process, the robot will be able to label the actions for a newly generated plan $\pi_{\mathcal{M}_R}$. Labels are selected from a set of task labels $T = \{T_0, T_1, \cdots, T_m\}$ that are associated with the current domain and zero, one, or multiple labels may be assigned to each action in the plan. Intuitively speaking, the human should be able to associate robot actions with task labels if he can understand its actions. In this paper, we further assume that each robot action can only be assigned to at most one label. If every action gets a task label from T, we say this plan is

explicable to the human. If any action doesn't receive a task label (i.e., get the empty set as the label, which is considered as a special task label in the implementation), we say this action is inexplicable and the entire plan is also (partially or fully) inexplicable.

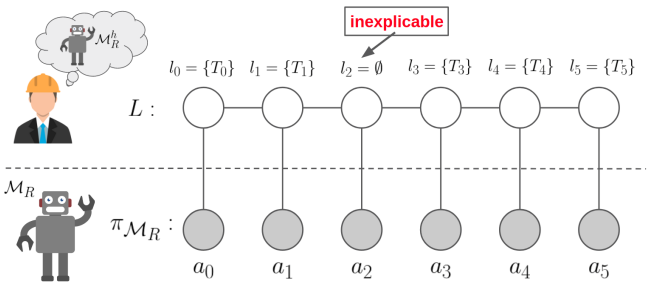
B. Intention Signaling

To illustrate intention signaling, we present an example as shown in Figure 2. Given a task, the robot generates a plan with 6 actions. Before executing the plan, the robot first tries to assign each action to a task label using the learned process that captures the human's interpretation of the robot's plan. Whenever an action gets assigned to the empty set label, it is interpreted as that the labeling process cannot assign it to a meaningful label (i.e., the human may not understand it). As shown in Figure 2a, the sequence of gray circles below the dashed line represents the action sequence of the robot plan and the sequence of white circles above represents the task labels of each action. The gray circles denote the observed variables while the white circles denote the hidden variables. In this example, each action except a_2 gets a label so that a_2 is inexplicable. In order to make a plan explicable, intention signaling allows the robot to signal its intention by providing context information about future actions to maximize the probability that these inexplicable actions can now be labeled. In this example, the robot finds that it should signal its intention right before executing a_2. The content of signaling contains the labels of l_3 and l_4, which are assigned to actions a_3 and a_4 as shown in Figure 2b. By doing this, the robot explains its behavior by providing the information about why a_2 will be taken (in this case, to facilitate the achievement of l_3 and l_4).

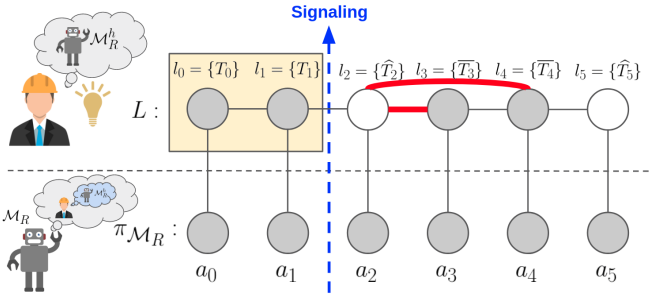
Hence, the search process for intention signaling must determine the timing (assumed to be in between two actions as shown in Figure 2b), and the labels of future actions (the content) to be used in the signaling. During the inference process, the labels before the timing currently being checked remain fixed based on the labeling process (since they will have been observed by the human before the signaling) and we explore all possible task labels (e.g., $\overline{T_3}$ and $\overline{T_4}$) for the selected task variables (e.g., l_3 and l_4), in order to make the labels for the remaining variables (e.g., l_2 and l_5) non-empty, resulting in $\pi_{\mathcal{M}_R}$ being explicable. The inference process is performed on the entire robot plan $\pi_{\mathcal{M}_R}$ before execution. The details of how to find the signaling timing and content will be discussed in the inference section.

C. Modeling the Labeling Process

As in [24], [25], we assume that the human's interpretation of the robot's plan can be formulated as a sequential labeling process. However, in this paper, we need the learning method to capture the influence of future context on the current action label, i.e., humans may look back and change the labels of actions they already assigned after seeing future actions. The labeling of these actions depends on not only the action itself, the plan context in the past, but also future actions. Consider an indoor domain where a human asks a robot to make a cup



(a) Labeling each action of $\pi_{\mathcal{M}_R}$. Labels L are hidden variables. In this case, actions a_0, a_1, a_3, a_4 and a_5 are assigned to task labels T_0, T_1, T_3, T_4 and T_5 respectively. Action a_2 is assigned to $\emptyset$ which means a_2 is inexplicable.



(b) Robot signaling its intention. a_2 was the inexplicable action. In this example, the robot chooses a timing which is right before a_2 for signaling its intention. $\overline{L} = \{l_3, l_4\}$ is the set of task variables the robot chooses to provide more context information for signaling (i.e., "*I perform a_2 in order to achieve $\overline{T_3}$ and then $\overline{T_4}$*"). After incorporating the additional evidence, all the remaining labels including l_2 become non-empty, i.e., the plan becomes explicable.

Fig. 2: An example that illustrates intention signaling. Circles represent variables. In each subfigure, variables below the dashed line are actions in a plan. The variables above are the task labels assigned by the human. The gray circles represent the variables whose values are given while the white circles represent hidden variables.

of coffee for him. His expectation may be for the robot to navigate to the kitchen to make a cup of coffee and then get back and hand it over. The first action of the robot, however, may be to navigate to a locker in the living room since the coffee maker was placed there after it was used last time. The human may be confused at first if the robot doesn't explain its actions. However, if the robot returns later with the coffee maker from the living room, the human will now understand the robot's actions even though they caused confusion in the first place. To capture such dependencies on future context, we use skip-chain CRFs which is a more appropriate model for this problem to capture long-term dependencies as label changes in the training samples.

Features for Learning: Similar to our prior work [25], we use action descriptions and predicates (or state variables) after taking each action as our plan features. For the skip-chain connections, we use the task labels as well as the distance between the connected variables as features since the model should also take the memory span of humans into account – we often will only be able to maintain context

Algorithm 1 Intention signaling: *when* and *what* to signal

Input: π, T_{pre}
Output: best timing and content for signaling
procedure SIGNALING
  $s_{max} \leftarrow 0$ ▷ Value to maximize
  $t_s^* \leftarrow 0$ ▷ Signaling timing
  $\overline{L}^* \leftarrow \{\}$ ▷ Selected task variables
  $T_{\overline{L}^*}^* \leftarrow \{\}$ ▷ Task labels for $\overline{L}^*$
  for $t_s \in \{0, \cdots, i\}$ **do**
    $C_{t_s} \leftarrow$ power set of $\{L_{t_s}, \cdots, L_{i-1}\} \cup \{L_{i+1}, \cdots, L_n\}$, excluding $\emptyset$
    for $\overline{L} \in C_{t_s}$ **do**
      $Perm_{\overline{L}} \leftarrow$ set of possible task labels for $\overline{L}$
      for $T_{\overline{L}} \in Perm_{\overline{L}}$ **do**
        $s \leftarrow s(\overline{L}, T_{\overline{L}}, t_s, T_{pre}, \pi)$
        if $s > s_{max}$ **then**
          $t_s^* \leftarrow t_s$
          $s_{max} \leftarrow s$
          $\overline{L}^* \leftarrow \overline{L}$
          $T_{\overline{L}^*}^* \leftarrow T_{\overline{L}}$
  return $\{t^*, \overline{L}^*, T^*\}$

information within a limited time span. When there are labels changes, both the original label and changed label of an action are recorded and modeled by our skip-chain CRFs. Given a plan $\pi_{\mathcal{M}_R} = \langle a_0, a_1, \cdots, a_n \rangle$, the training sample with a label change will be considered as two instances separately:

Instance 1: $\langle \langle F_0, L_0 \rangle, \langle F_1, L_1 \rangle, \langle F_2, L_2 \rangle, \cdots \rangle$
Instance 2: $\langle \langle F_0, L_0 \rangle, \langle F_1^s, L_1^s \rangle, \langle F_2, L_2 \rangle, \cdots \rangle$

where $\langle F_i, L_i \rangle$ denotes the set of features and task label for the ith action respectively. $\langle F_i^s, L_i^s \rangle$ denotes the features and task label for the ith action after a label change occurred. Instance 1 is the sample without the skip-chain features, and instance 2 is the sample that incorporates the skip-chain features with the label change.

D. Inference

In order to find the timing and content to signal, the robot searches over all the timings, denoted by t_s, before the inexplicable action, denoted by a_i (the ith action in the plan), and all the possible task label assignments (i.e., for actions after t_s excluding a_i) that can make the inexplicable action to be explicable. In this paper, we assume that t_s is always in between two adjacent actions so that $t_s = j$ means that signaling is to happen between the $(j-1)$th and jth actions. At each iteration, the robot picks $t_s \in \{0, \cdots, i\}$ to be the timing it performs signaling and $\overline{L}$ to be the set of task variables it will assign labels to (i.e., these labels determine the content of signaling). We fix all the labels before t_s and only select task variables after t_s for reasons we explained earlier.

For each possible set of task variables selected to use in the signaling, denoted by $\overline{L}$, we search over all possible assignments of task labels to them from T. The optimal timing and content can then be determined by the choice that maximizes a measure (Equation 1) that captures the quality of the signaling which reflects how likely the plan will be made explicable after the signaling. The inference process is shown in Algorithm 1 for plans with a single inexplicable action. When there are two or more inexplicable actions in a plan, every similar process can be used by searching for multiple signalings simultaneously. Or, we can first apply signaling to the first inexplicable action and perform the search process described above from start till some action before the next inexplicable action to make the first part of the plan explicable. This process can then be followed for each remaining inexplicable action in an online fashion.

In order to compare different choices of timing and content to pick the best one, we introduce a new metric to score them, which is denoted as metric s. It measures the probability that the plan will become explicable, or more intuitively, how likely the human would fully understand the robot's plan after the signaling. The inference problem can now be formulated as follows:

$$\underset{\{\overline{L}, T_{\overline{L}}, t_s\}}{\operatorname{argmax}} \; s(\overline{L}, T_{\overline{L}}, t_s, T_{pre}, \pi) \tag{1}$$

where $T_{\overline{L}}$ denotes the task labels that we assign to $\overline{L}$. T_{pre} denotes the labels assigned to plan π which we obtain through the labeling process. The goal is to find a combination of task variables $\overline{L}$ with labels $T_{\overline{L}}$ and signaling timing t_s that maximize the human's understanding of the robot's plan. Given T_{pre} and π, we fix the labels from start to t_s. For $\overline{L}$ and $T_{\overline{L}}$, again, we search over different combinations of task variables and their labels. The metric s above is computed as follows:

$$s(\overline{L}, T_{\overline{L}}, t_s, T_{pre}, \pi) = P(\widehat{L} \in T^{|\widehat{L}|} \mid L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}}, \pi) \tag{2}$$

where $|\widehat{L}|$ is the number of label variables in $\widehat{L}$, $L_{0:t_s-1}$ is the set of task variables from start to $t_s - 1$, $T_{0:t_s-1}$ is the label set assigned to $L_{0:t_s-1}$ based on T_{pre}, and $\widehat{L}$ is the set of remaining variables. Each variable in $\widehat{L}$ must be non-empty since the actions associated with these variables must be explicable (or have a label). We marginalize over all the possible label assignments of $\widehat{L}$:

$$s(\overline{L}, T_{\overline{L}}, t_s, T_{pre}, \pi) = \sum_{\widehat{L} \in T^{|\widehat{L}|}} P(\widehat{L} \mid L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}}, \pi) \tag{3}$$

Based on the definition of conditional probability, we transform Equation 3 as follows:

$$s(\overline{L}, T_{\overline{L}}, t_s, T_{pre}, \pi) = \sum_{\widehat{L} \in T^{|\widehat{L}|}} \frac{P(\widehat{L}, L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}} \mid \pi)}{P(L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}} \mid \pi)} \tag{4}$$

In the denominator, based on the definition of marginal distribution, we have:

$$s(\overline{L}, T_{\overline{L}}, t_s, T_{pre}, \pi) = \sum_{\widehat{L} \in T^{|\widehat{L}|}} \frac{P(\widehat{L}, L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}} \mid \pi)}{\sum_{\widehat{L}'} P(\widehat{L}', L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}} \mid \pi)} \quad (5)$$

where $\widehat{L}' \in \{\{\emptyset\} \cup T\}^{|\widehat{L}|}$. Since $\widehat{L}$ is not in the denominator, we can move the outer summation up to the nominator.

$$s(\overline{L}, T_{\overline{L}}, t_s, T_{pre}, \pi) = \frac{\sum_{\widehat{L} \in T^{|\widehat{L}|}} P(\widehat{L}, L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}} \mid \pi)}{\sum_{\widehat{L}'} P(\widehat{L}', L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}} \mid \pi)} \quad (6)$$

The probability in the nominator and denominator can be rewritten as summation over the label sequence L where $L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}}$.

$$P(\widehat{L}, L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}} \mid \pi) = \sum_{L | L_{0:t_s-1} = T_{0:t_s-1}, \overline{L} = T_{\overline{L}}} P(L \mid \pi) \quad (7)$$

$P(L|\pi)$ measures the probability that L is the correct label sequence for the plan π. It is modeled by a skip-chain CRFs as follows:

$$P(L \mid \pi) = \frac{1}{Z} \prod_{t=1}^{T} \Psi_t(l_t, l_{t-1}, \pi) \prod_{(u,v) \in C} \Psi_{uv}(l_u, l_v, \pi) \quad (8)$$

where Z is the normalization factor. $C = \{(u, v)\}$ is the set of indices of all pairs of task variables where u is the index for the inexplicable action and v is the index for a variable in $\overline{L}$. $\Psi_t(l_t, l_{t-1}, \pi)$ is the factor for label l_t and the previous label l_{t-1}. $\Psi_{uv}(l_u, l_v, \pi)$ is the factor for label l_u and label l_v.

E. "When" to Signal

The timing of intention signaling is determined by t_s which happens after a_{t_s-1} and before a_{t_s}. When the robot searches for $\overline{L}$, we check the task label variables after t_s, not i (i.e., it depends on the selected signal timing, not the index of the inexplicable action). For instance, the robot in the example shown in Figure 2b picks $t_s = 2$ to be the signal timing that is right before the inexplicable action a_2. In this case, The choice of $\overline{L}$ could be any combination of l_3, l_4 and l_5. Similarly, when $t_s = 1$, the choice of $\overline{L}$ would be any combination of l_1, l_3, l_4 and l_5. For example, $\overline{L} = \{l_1, l_3\}$ is a possible choice. In this case, signaling not only uses the future content (i.e., l_3) but also provides a chance of influencing the task labels (i.e., l_1) that are predicted by the labeling process and precede the inexplicable action, in case the preceding labels misguide the human's understanding.

During the inference process, theoretically, all the t_s values should be explored which would lead to a high computational cost due to the large search space. To reduce this computation, we set a constraint on t_s such that $i - 3 \leq t_s < i$ based on the assumption that it may not be helpful to the human if robot signals its intention too early.

F. "What" to Signal

Given t_s, the content of signaling is explored by checking different combinations of the task labels after t_s. After determining the set of label variables and their values for intention signaling, a natural language sentence will be generated. We create a set of templates that can translate the task labels and the (inexplicable) action as follows:

1) "*I will [ACTION].*"
2) "*After [TASK_BEFORE_0] and then [TASK_BEFORE_1] ..., I will [ACTION] in order to achieve [TASK_AFTER_0] and then [TASK_AFTER_1] ....*"
3) "*I will [ACTION] in order to achieve [TASK_0] and then [TASK_1] ...*"

where "*[ACTION]*" refers the action the robot is trying to explain, and "*[TASK]*" refers to the task labels the robot uses to signal to help the human better understand its actions. Template 1 is used when the robot simply signals the inexplicit action before executing it, i.e., when we cannot find a signaling that explains the action. Template 2 is used when the robot chooses to signal before the inexplicable action and uses task labels that are both before and after the inexplicable action for signaling. We use the last template when the robot signals right before the inexplicable action and uses one or more task labels for future actions for signaling. In order to simplify the computation of the inference process, we limit the label variables to be selected from $\{l_{i+1}, l_{i+2}, l_{i+3}\}$, so template 2 is not used. With all the simplifications, our implementation can find the best signaling timing and content in a few seconds.

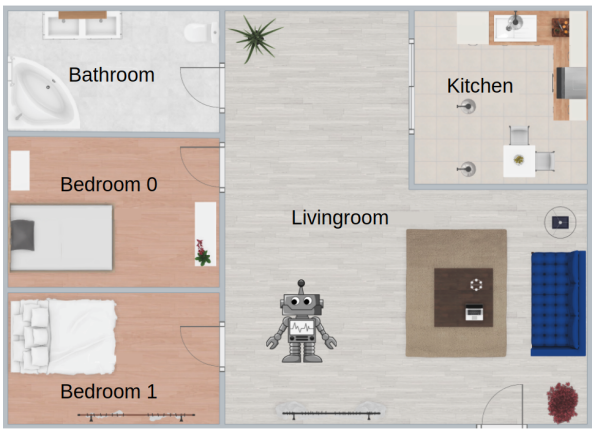
IV. Evaluation

To evaluate our approach, we use a synthetic robot maid domain where the human is unaware of certain information so that he may have a different understanding of the behavior of the robot. Given a task, the robot will generate a plan. It will then use the plan explicability labeling process [25] to decide whether there is any inexplicable action. If so, the robot uses intention signaling to explain its behavior.

A. Synthetic Domain

As shown in Figure 3, the simulated domain has a living room, a kitchen, a bathroom and two bedrooms. Each bedroom and bathroom has a door which may be either open or locked. There is a key for each door in the environment which could be in any of the rooms. In this synthetic domain, we assume that the goal is to make a cup of coffee and clean the bathroom. For making coffee, the robot needs to have a coffee maker and coffee beans. Furthermore, it has to make coffee at kitchen since that is the only place it can get water. The coffee maker can be placed in the kitchen, living room or bedrooms. Similarly, coffee beans can also be placed anywhere. For cleaning the bathroom, the robot needs a vacuum and the vacuum can also be anywhere. In this domain, the human is unaware or unsure about the positions of objects while the robot maid keeps track of about them. [1]

[1]It should not be surprising how forgetful humans are!



(a) Synthetic domain for evaluation.

Fig. 3: Synthetic robot maid domain and four objects (not shown in the environment so that the human is unaware of their locations) in the domain.

Consider an example where the goal for the robot is to clean the bathroom. The initial state of the environment is that the bathroom is locked, the key to the bathroom is in bedroom_0, and the vacuum is in bedroom_1. In this case, the robot's plan is to first navigate to bedroom_0 to get the key to the bathroom and then use the key to open the bathroom door. Next, it goes to bedroom_1 to fetch the vacuum and go back to the bathroom for cleaning. In our domain setting, since the human is unaware of whether the bedrooms and bathroom are open or locked or where the vacuum and keys are, the robot's behavior may seem erratic and inexplicable to the human.

For example, in the situation above, the first action of the robot which is to navigate to bedroom_0, could be inexplicable to the human if the human believes that the bathroom is open and the vacuum is already in the bathroom. Similarly, the human may get confused when he observes that the robot navigates to bedroom_1 later. These inexplicable actions will, most likely, become explicable when the human observes the future robot actions, for example, when the human sees the robot using a key to open the door to the bathroom.

B. Labeling Training Samples

In order to generate our training data set, for each plan we label it systematically using the following method. We assume that, from the human's perspective, if the robot's action does not contribute to the goal, the label will not be assigned to any task labels. Otherwise, it will be assigned a task label based on what the action is doing or where the robot is heading to. In this synthetic domain, we assume only two task labels [MAKE_COFFEE, CLEAN_BATHROOM].

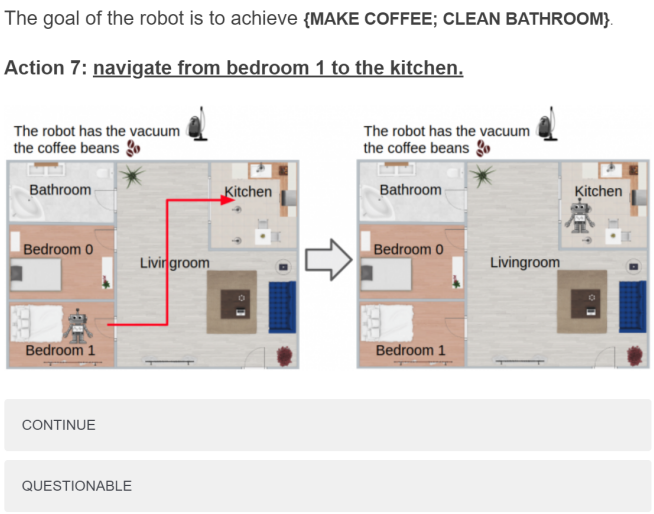


Fig. 4: Actions are displayed to the human subjects on Mturk for evaluation. This example shows the second action of a plan were the robot navigates from bedroom_1 to the kitchen. The subject determines whether this action is "QUESTIONABLE" or not based on the current state.

For making coffee, we assume that the human always considers that the robot should go to the kitchen and get the coffee maker and beans for making coffee. Similarly, for cleaning, we assume that the human always thinks that the robot should get the vacuum and go to the bathroom for cleaning. Depends on the specific situation, any actions satisfying these predefined requirements will be assigned to the corresponding task label. Others are treated as inexplicable and assigned the empty set as their labels.

As discussed in the previous sections, we also need to capture the label changes as well. We will go back and reassign task labels to certain actions. In our synthetic domain, since there is no useless actions in a plan for achieving the goal such that the inexplicable actions always serve as prerequisites for some future actions that contribute to the goal. Thus, we reassign labels to those inexplicable actions the correct task label. For example, when the task is to make a cup of coffee and the inexplicable action is to navigate to the living room, and the next action is to fetch the coffee maker in the living room which we already labeled as MAKE_COFFEE. We will relabel this navigation action as MAKE_COFFEE as well.

C. Experiment Settings and Results

To evaluate our approach, we create surveys using Qualtrics survey system and post them on Amazon Mechanical Turk (MTurk). The participants on MTurk are presented the robot's actions one by one in order as shown in Figure 4. We provide a description of the domain and the goal assigned to the robot at the beginning of the survey. The subjects are tasked to evaluate the robot's action sequence. We provide two choices for them: whenever the participants think the robot's behavior is not explicable, they may choose "QUESTIONABLE". Otherwise, they can select "CONTINUE".

"good task and very interesting"

"It is helpful to understand the robot's behavior. If I did not know why the robot was making some action I might think that it was wrong to do that action."

"It just gave me some insight on the robot's intentions and I could try to help more. "

Fig. 5: Participants' comments on intention signaling.

TABLE I: Comparison of the average number of "QUESTIONABLE" actions per plan between planning with intention signaling and without.

Evaluation Setting	Avg. "QUESTIONABLE" Actions
Number of random actions	1.75
No signaling	3.03
Intention signaling	2.48

We post two evaluation settings with the same set of scenarios. One of the setting simply uses the robot's plan. The other uses the robot's plan along with the natural language sentences generated by our approach for intention signaling. We insert a few random actions on purpose and the participants are evaluated based on how accurately they can tell random actions from valid actions, which they are also aware of. We recruited 20 participants for each setting being evaluated. Each participant assessed 16 scenarios.

We evaluate the performance of our approach by comparing the average number of "QUESTIONABLE" actions with and without signaling. The result is shown in Table I. The average number of random actions inserted is 1.75. For the method without signaling, there are on average 3.03 "QUESTIONABLE" actions and the number is 2.48 with intention signaling which is closer to 1.75. This result shows that intention signaling can largely reduce the criticism on robot actions in teaming settings. We also collect feedbacks from the participants as shown in Figure 5 which also suggested that intention signaling can provide useful information for the participants to better understand the robot's behaviors.

V. Conclusion

To address the problem of how to make the robot's plan explicable to humans, we propose an approach that enables the robot to explain its behavior by signaling its intention. To achieve this, we formulate the human's interpretation of robot actions as a labeling process, similar to [25]. Given a plan, the robot first determines whether it is explicable to the human by checking if every action in the plan can be assigned to a task label. If not, it will search for the timing and content to signal its intention in order to make the human better understand its plan. We develop a new metric for modeling the human's understanding of the robot's plan that takes into account future context. A skip-chain CRFs model is used for capturing the long-term dependencies between action labels. To evaluate our approach, we conduct experiments with a synthetic robot maid domain. We compare the performance of our approach with a baseline approach without signaling. Results show that our approach can largely improve the explicability of the robot's plan and thus benefit teaming.

References

[1] R. S. Andersen, O. Madsen, T. B. Moeslund, and H. B. Amor, "Projecting robot intentions into human environments," in *RO-MAN*. IEEE, 2016, pp. 294–301.

[2] T. Chakraborti, S. Kambhampati, M. Scheutz, and Y. Zhang, "Ai challenges in human-robot cognitive teaming," *arXiv preprint arXiv:1707.04775*, 2017.

[3] T. Chakraborti, S. Sreedharan, Y. Zhang, and S. Kambhampati, "Plan explanations as model reconciliation: Moving beyond explanation as soliloquy," in *IJCAI*, 2017, pp. 156–163.

[4] N. J. Cooke, J. C. Gorman, C. W. Myers, and J. Duran, "Interactive team cognition," *Cognitive Science*, 2013.

[5] A. Dragan and S. Srinivasa, "Generating legible motion," in *Robotics: Science and Systems*, June 2013.

[6] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *HRI*. IEEE, 2013, pp. 301–308.

[7] A. Fern, S. Natarajan, K. Judah, and P. Tadepalli, "A decision-theoretic model of assistance," *J. Artif. Int. Res.*, vol. 50, no. 1, pp. 71–104, May 2014. [Online]. Available: http://dl.acm.org/citation.cfm?id=2693068.2693071

[8] M. Fox and D. Long, "Pddl2. 1: An extension to pddl for expressing temporal planning domains," *Journal of artificial intelligence research*, 2003.

[9] M. Fox, D. Long, and D. Magazzeni, "Explainable planning," *arXiv preprint arXiv:1709.10256*, 2017.

[10] Z. Gong and Y. Zhang, "Temporal spatial inverse semantics for robots communicating with humans," in *ICRA*, 2018.

[11] R. A. Knepper, C. I. Mavrogiannis, J. Proft, and C. Liang, "Implicit communication in a joint action," in *HRI*. ACM, 2017, pp. 283–292.

[12] P. Langley, "Explainable agency in human-robot interaction," in *AAAI Fall Symposium Series*, 2016.

[13] B. Y. Lim, A. K. Dey, and D. Avrahami, "Why and why not explanations improve the intelligibility of context-aware intelligent systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 2119–2128.

[14] 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," *The International Journal of Robotics Research*, vol. 34, no. 14, pp. 1711–1730, 2015. [Online]. Available: http://dx.doi.org/10.1177/0278364915609673

[15] M. Ramírez and H. Geffner, "Plan recognition as planning," in *IJCAI*, 2009, pp. 1778–1783.

[16] E. Rosen, D. Whitney, E. Phillips, G. Chien, J. Tompkin, G. Konidaris, and S. Tellex, "Communicating robot arm motion intent through mixed reality head-mounted displays," *arXiv preprint arXiv:1708.03655*, 2017.

[17] J. Shah, J. Wiken, B. Williams, and C. Breazeal, "Improved human-robot team performance using chaski, a human-inspired plan execution system," in *HRI*. ACM, 2011, pp. 29–36.

[18] S. Sreedharan, T. Chakraborti, and S. Kambhampati, "Balancing explicability and explanation in human-aware planning," *arXiv preprint arXiv:1708.00543*, 2017.

[19] C. Sutton, A. McCallum, *et al.*, "An introduction to conditional random fields," *Foundations and Trends® in Machine Learning*, vol. 4, no. 4, pp. 267–373, 2012.

[20] L. Takayama, D. Dooley, and W. Ju, "Expressing thought: improving robot readability with animation principles," in *HRI*, 2011, pp. 69–76.

[21] S. Tellex, R. A. Knepper, A. Li, D. Rus, and N. Roy, "Asking for Help Using Inverse Semantics," in *Robotics: Science and Systems*, 2014.

[22] M. Walker, H. Hedayati, J. Lee, and D. Szafir, "Communicating robot motion intent with augmented reality," in *HRI*, 2018, pp. 316–324.

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

[24] Y. Zhang, S. Sreedharan, A. Kulkarni, T. Chakraborti, H. H. Zhuo, and S. 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.

[25] ——, "Plan explicability and predictability for robot task planning," in *ICRA*. IEEE, 2017, pp. 1313–1320.