

Improved Inference of Human Intent by Combining Plan Recognition and Language Feedback

Ifrah Idrees¹, Tian Yun¹, Naveen Sharma¹, Yunxin Deng¹, Nakul Gopalan², George Konidaris¹, Stefanie Tellex¹

¹Brown University, USA, ²Arizona State University, USA

Abstract—Conversational assistive robots can aid people, especially those with cognitive impairments, to accomplish various tasks such as cooking meals, performing exercises, or operating machines. However, to interact with people effectively, robots must recognize human plans and goals from noisy observations of human actions, even when the user acts sub-optimally. Previous works on Plan and Goal Recognition (PGR) as planning have used hierarchical task networks (HTN) to model the actor/human. However, these techniques are insufficient as they do not have user engagement via natural modes of interaction such as language. Moreover, they have no mechanisms to let users, especially those with cognitive impairments, know of a deviation from their original plan or about any sub-optimal actions taken towards their goal. We propose a novel framework for plan and goal recognition in partially observable domains—Dialogue for Goal Recognition (D4GR) enabling a robot to rectify its belief in human progress by asking clarification questions about noisy sensor data and sub-optimal human actions. We evaluate the performance of D4GR over two simulated domains—kitchen and blocks domain. With language feedback and the world state information in a hierarchical task model, we show that D4GR framework for the highest sensor noise performs 1% better than HTN in goal accuracy in both domains. For plan accuracy, D4GR outperforms by 4% in the kitchen domain and 2% in the blocks domain in comparison to HTN. The ALWAYS-ASK oracle outperforms our policy by 3% in goal recognition and 7% in plan recognition. D4GR does so by asking 68% fewer questions than an oracle baseline. We also demonstrate a real-world robot scenario in the kitchen domain, validating the improved plan and goal recognition of D4GR in a realistic setting.

I. INTRODUCTION

People with cognitive impairments, such as dementia, often struggle with focusing on everyday tasks and have limited attention spans. Efforts to assist people in tracking task progress have involved various approaches, including modeling tasks as a hierarchical task network (HTN), using a Bayesian Hidden Markov Model, or employing a Partially Observable Markov Decision Process (POMDP) [6, 19]. However, previous approaches focus on observing users rather than engaging them in an interaction. Our work aims to develop a model for a robot capable of assisting people complete tasks with language-based interactions; even when the users perform sub-optimal actions or switch between multiple goals. Our robot tracks the task progress using observations and question-asking using natural language. Such a robot can also benefit an operator building a machine, a child with autism doing their homework, or a child learning to do chores.

Inferring the goals and intents of the human requires plan and goal recognition (PGR) using noisy evidence from action

execution that can be done efficiently using planning techniques [11]. One key challenge in human intent recognition as a PGR problem is that the robot has partial observability of human intentions. This is compounded by noisy sensors that also create partial observability of the environment. Modeling human progress during hierarchical tasks has been done using Hierarchical Task Networks (HTNs) [19, 7]. However, these recognition techniques again do not engage with the users and assume that the user acts optimally. Moreover, incorporating clarification questions and language utterance spoken by humans in PGR is challenging because of the huge space of language observations. The existing solution to this problem is heuristics [16, 5, 10], which are prone to fail as the tasks and environment sensors become complex and noisy.

The main contribution of our paper is a novel formulation that combines expressive and hierarchical task representation of HTNs to represent the human mental states with the sequential decision-making capabilities of a Partially Observable Markov Decision Process, in our Dialogue for Goal Recognition (D4GR). Our method keeps track of the environment, user state and dialogue history internally to perform PGR and guide in successful task completion.

POMDPs can model the uncertainty the robot faces as it performs intent recognitions and enables the robot to ask information-seeking questions. However, POMDP planners traditionally do not decide the relevance of the state using the task network at each sequential time step. Further, the HTNs have no notion of rewards to generate a sequence of actions to maximize the agent’s utility. To solve these challenges, we assume that the user is a planner with goals and subgoals that are represented hierarchically. Moreover, the robot is a POMDP planner performing long-horizon dialogue policy planning. This enables the robot to reason about asking meaningful questions in ambiguous settings, such as the user switching goals during multiple concurrent tasks and also performing sub-optimal actions. Using this information, the robot can better recognize human intents based on the human actions estimated from noisy sensors and their language feedback. This model also explicitly allows for sub-optimal plans by a human user, which D4GR can detect.

We evaluate the usefulness of D4GR by measuring the improved accuracy in PGR and comparing the planning time and number of questions asked to two state-of-the-art baselines in the simulated domain developed by Wang and Hoey [19]. Our system is able to more accurately infer human intents than these baselines using information gathered from a language

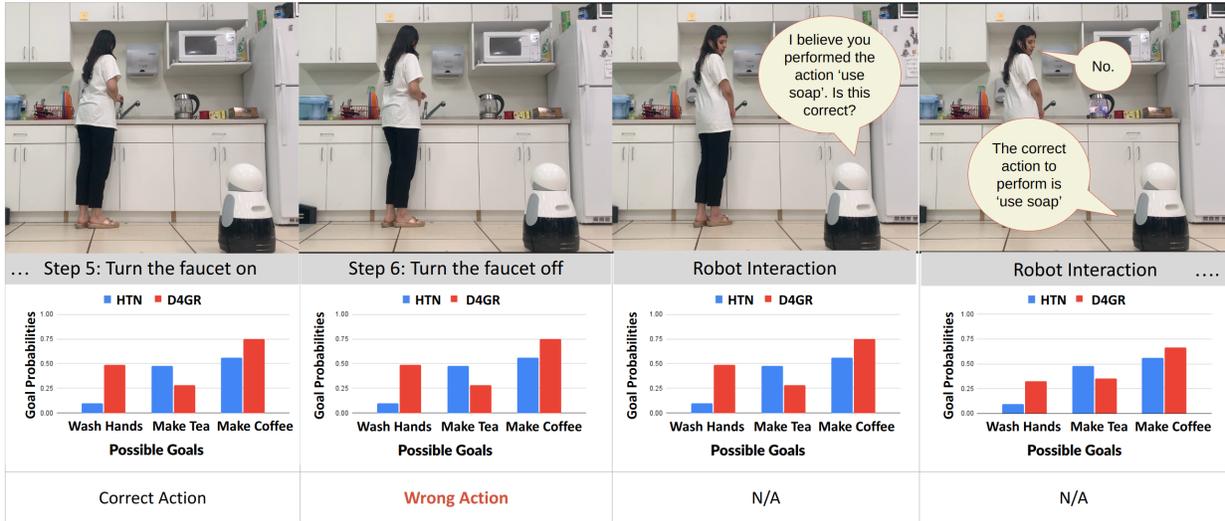


Fig. 1. Kuri robot performing human intent (plan and goal) recognition. The human, while making coffee, starts a new goal of washing hands. However, the human forgets to use soap after turning on the faucet and instead turns off the faucet. The robot (with access to the Hierarchical Task Network representation of goals) observes that the current world observations do not progress the previous goal and also cannot lead to a new goal completion and hence uses D4GR to ask a clarification question. Based on language feedback, the robot reduces its confidence in the goal wash hands and suggests the next action as using soap to complete one of the most plausible goals of washing hands.

without asking unnecessary questions. We run 880 trials for varying sensor noise levels where the simulated human tries to complete a combination of tasks in two domains - the kitchen and blocks domain. In the kitchen domain, there are three tasks - washing hands, making a cup of tea, and making a cup of coffee while in the blocks domain, tasks include an assortment of stacking letters to make 4-7 lettered words - *rote*, *tone*, *tune*, *hawk*, *capstone*. With language feedback and the world state information in a hierarchical task model, we show that our D4GR framework outperforms HTN by 4% on plan accuracy in the kitchen domain and 2% in the blocks domain. In goal accuracy, for the highest sensor noise, our D4GR performs 1% better than HTN in both kitchen and block domains. We also deployed our algorithm on a social robot Kuri as a demonstration of a socially intelligent robot helping confused users complete tasks. In this demonstration, Kuri performs improved PGR by asking clarification questions and reducing uncertainty for a challenging scenario where the user switches between multiple concurrent goals (e.g., washing hands and making coffee) and acts sub-optimally in the kitchen domain. The demo can be found here: <https://youtu.be/Om91zBiDDEY>. An example sequence of the user and robot interaction can be seen in Fig-1.

II. RELATED WORK

Plan and Goal Recognition as Planning: The first work on PGR as planning was introduced by Ramírez and Geffner [15]. This research leverages classical planning systems to solve PGR problems. Wang and Hoey [19] proposed an algorithm for PGR based on hierarchical task network [4] that handles noisy sensors and sub-optimality in human actions. However, their approach detects mistakes heuristically using a manually defined threshold while performing a one-step look ahead without long-horizon planning. Additionally, it lacks dialog

conversation to improve belief in the actor’s progress in the task. Another relevant work that does not engage with the user is by Zhi-Xuan et al. [22], which performs online Bayesian goal inference by modeling the agent as a boundedly rational planning agent but is not designed and evaluated for multiple concurrent hierarchical goals. Mirsky et al. [12] presents favorable results for the hypothesis that feedback from the acting agent can improve plan (goal and step) recognition, but their paper performs goal recognition as reinforcement learning and has a fixed language policy without exploring the observer’s strategy for asking clarification question. Höller et al. [7] employs HTN Planning for PGR but struggles with sub-optimal actions and noisy sensors, focusing on handling missing sensor observations instead.

Context-aware Social Robotics: In the past, research in social robotics has focused on developing non-verbal social behaviors for robots to assist the elderly [5, 10] during task completion. These works place less emphasis on incorporating language feedback/observations for user intent inference. Further, the robot dialog policy (if involved) does not account for the environmental context, dialog context, and user modeling. Research in situated human-robot dialog by Bohus and Horvitz [1], Thomason et al. [18], Idrees et al. [8] grounds speech response in the environment but asks clarification questions heuristically, using rule-based/greedy approaches and without using a decision-theoretic framework. Such heuristics are prone to failure as the tasks get complex and the environment sensors become complex and noisy.

POMDP-based Collaborative Dialog: Partially observable Markov decision processes (POMDPs) provide a rich framework for planning under uncertainty. They excel in optimizing agents’ actions over long horizons in complex environments despite incomplete state information from noisy sensors [6].

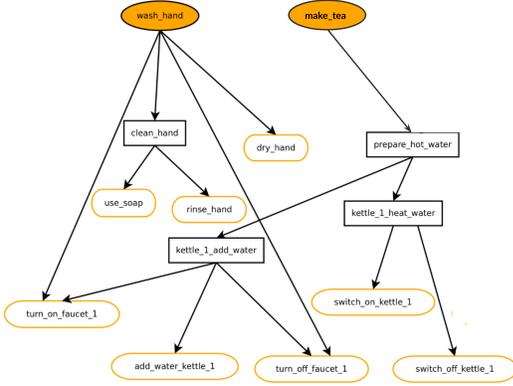


Fig. 2. Hierarchical Task Network for two of our tasks - washing hands, and making tea

Young et al. [21] and Doshi and Roy [3] built POMDP-based dialog systems. However, these only use language as observations, not world observations, for belief updates to choose actions with the highest reward accumulation. Whitney et al. [20] fuses language and world observations for object fetching tasks but does not model the user’s mental state during multi-step task completion. The closest work to ours is by Hoey et al. [6], featuring an engaging assistant that incorporates only world observations and not language observations for multi-step tasks. They can infer human actions and psychological states through hand and towel tracking but can not handle multiple and concurrent tasks and backtracking and are limited to hand washing. Research has recently focused on using reinforcement learning in a collaborative dialog for interactive task learning Chai et al. [2]. However, these works require an existing dataset for offline learning, while our planning approach doesn’t necessitate data collection or learning.

III. TECHNICAL APPROACH

Given the sensor measurements O_w of the world state W , an assistive robot needs to infer the probabilities of the hidden user intent—the person’s current goal G and the current primitive human action α . Partial observability of the latent user intent and noise in the sensor observations impact accurate inference of G , W , and α at every sequence of human actions. We infer the probability distribution of the hidden states G , W , and α using HTN planning to generate feasible plans for the human-given predefined tasks similar to work of Wang and Hoey [19]. With the ability to ask clarification questions, the robot actively improves the inference of the latent states. The user’s response in the form of language observations, o_l , gives the agent additional information about their task. With our D4GR framework, the robot decides when to ask clarification questions and which user action to inquire about, based on information gathered from both the sensors and language. Our framework avoids asking unnecessary questions and balances between information-gathering actions, like asking questions, and goal-inducing actions, such as providing the correct next step.

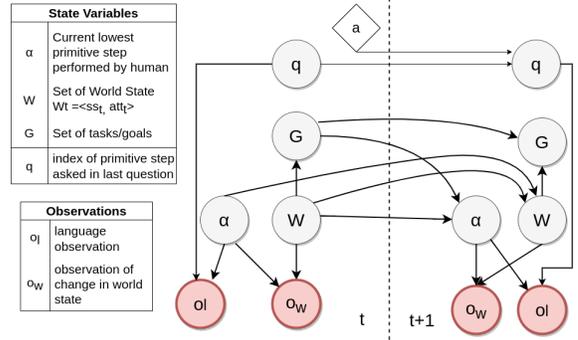


Fig. 3. Influence diagram for D4GR.

A. POMDP Definition

We model our PGR problem as a POMDP [9] planning problem, generating an approximately optimal action policy for the robot. Formally, a POMDP is defined as a tuple $(S, A, T, R, \Omega, O, \gamma, b_0)$ where S is the state space, A is the action space, T is the transition probability, R is the reward function, Ω is a set of observations, O defines an observation probability and γ is the discount factor. Since s_t is not known exactly, the POMDP model updates, at each timestep, the probability distribution over all possible states (belief state b_t). The POMDP agent uses a planner to generate an optimal policy for the robot’s action, which in this case is the communication with the human user.

B. D4GR Formulation

We define a novel model, **Dialogue 4for Goal Recognition**, a Partially Observable Markov Decision Process (D4GR) that combines the goal and plan recognition components as described by Wang and Hoey [19] with the POMDP formalism to allow robots to take action in the environment through dialogue for improved PGR. For efficient human intent recognition and estimation of optimal action policy, our D4GR must handle the large space of the world and language observations. Our formulation leverages the hierarchical task structure of HTN and assumes independent assumptions between state variables set of goals G , human user action α , and world state W for efficient belief update. Our D4GR has the following components: $(S, A, T, R, \Omega, O, \gamma; b_0)$.

State (S): The state, $s_t \in S$, consists of a tuple of the user’s mental state, M_t , and the world state, W_t , along with information needed to track the dialog state. We represent the user mental state using HTNs. We assume access to the HTN’s fixed knowledge graph $TaskNet$ for the tasks, where the root node(s) represent the high-level tasks/goals G that the human can do. The internal nodes are sub-tasks that can be decomposed into leaf nodes depicting primitive human actions. α denotes the current primitive human action. We model the user’s mental state M represented by G and α . The partial $TaskNet$ for our kitchen domain is shown in Fig-2.

The world state W_t combines the states of world smart sensors ss_t and the attributes of objects involved in the task att_t , such as dryness of the hand, state of the faucet, etc. The

dialog state variable q_t stores the latest primitive human action referenced by the robot in its clarification question. Thus, the state s_t can be factored into the following components: $s_t = (M_t, q_t, W_t,)$ where $W_t = (ss_t, att_t)$, $M_t = (G_t, \alpha_t)$. Here $G_t, \alpha_t, att_t, ss_t$ are the hidden variables, while q_t is the known variables, hence making this a Mixed Observable Markov Decision Process (MOMDP) [13]. The influence diagram can be seen in Fig-3.

Action(A) includes the actions of the agent. The robot for this research can perform the following predefined language-based actions - 1) *Wait*: does nothing but advances the time step. 2) *Ask*{ $argmax(\alpha)$ }: The robot asks a clarification question about the primitive action α with the highest belief. The question template used is: "I believe that you just did action (α_i), is this correct?". 3) *inform_next_instruction*: informs the next action that the user should perform at the current timestep based on the current belief. This action is chosen based on a fixed policy and is executed when the user provides a negative language response to the robot's clarification question.

Observations (Ω) encompass both the user's language (o_l) and observations about the world state (o_w). o_w includes discrete observations of the world smart sensor's state ss_t and the attributes of the task-related objects att_t such as *hand_dry == true, faucet_on == false*, etc. These observations are binary for the states in W_t , so the faucet can only be on or off. The language observations are natural language responses.

Observational Model (O): The robot needs a model of $p(o|s) = p(o_l, o_w|s)$ to update its belief. Most of the complexity of our model is captured in this observation model and belief update defined in the sec-III-C.

Transition Model (T) : $T(s, a, s') \equiv p(s_{t+1}|s_t, a_t)$. Our stochastic transition function is factorized as shown in Eq-1, following a similar approach to Wang and Hoey [19]. We factor our mental model M_t into G_t and α_t . Additionally, we assume that the last question asked, q_t , is independent of G , α , and W .

$$p(s_{t+1}|s_t, a_t) = p(G_{t+1}|W_{t+1}, G_t) \times p(W_{t+1}|W_t, \alpha_{i,t+1}) \times p(\alpha_{t+1}|W_t, G_t) \times p(q_{t+1}|q_t, a). \quad (1)$$

In Eq-1, we assume that q_t changes deterministically from null to $max(\alpha_t)$ after the robot asks a clarifying question. Further, G is deterministically carried forward to the next time step.

$$p(q_{i,t+1}|q_{i,t}, a_t) = \begin{cases} 1 & \text{for } max(\alpha) \text{ else } 0, & \text{if } a \neq NULL \\ 1 & \text{for } q_{i,t}, & \text{else } 0 \quad a_t == NULL. \end{cases}$$

Reward ($R(s, a)$) We provide a positive reward (5) for asking a clarifying question when the user is doing the wrong or suboptimal lowest primitive step. A negative reward (-5) for asking a clarifying question when the human user is doing the correct primitive step or when the agent asks a question about a wrong primitive action. Thus, doing nothing accumulates zero rewards until the right question is asked,

while not asking a question or asking a wrong one results in a penalty.

C. Belief Update for Goal Recognition and Planning

Our belief update performs human intent recognition by maintaining a belief over the hidden user's mental state $M_t = (G_t, \alpha_t)$ and the world state W_t . The actions executed by the user produce an observation of the world state o_w indicating the change in the world state W_t . The user can also provide speech/language feedback o_l in response to the clarification question asked. We classify the intent of each sentence into positive or negative feedback using the bag of words approach. Negative responses r_n include { 'no', 'nope', 'other', 'not' } while positive responses r_p include the words { 'yes', 'yeah', 'sure', 'yup' }. Further, our world sensor noise model generates the correct sensor state with probability sr and the incorrect sensor state with $1 - sr$. We adopt the noise model for the sensor described by Wang and Hoey [19].

The observation model can be further expanded and approximated as follows:

$$p(o_t|s_t; a_{t-1}) \propto p(s_t|o_t, a_{t-1}) * p(o_t|a_{t-1}). \quad (2)$$

Overall, the probability of s_t given o_t and a_{t-1} can be factored into the world observation model and language observation model in Eq-3. We assume that the world observation $o_{w,t}$ solely provides information about the W and α . Meanwhile, the language observation is relevant to the human action α , the last question asked q_t , subsequently affecting the goal.

$$p(s_t|o_t, a_{t-1}) \propto \underbrace{p(G_t|W_t) * p(W_t|o_{w,t}) * p(\alpha_t|o_{w,t}) * p(o_{w,t})}_{\text{world observational model}} * \underbrace{p(\alpha_t, q_t|o_{l,t}) * p(o_{l,t}) * p(q_t|a_{t-1})}_{\text{language observational model}}. \quad (3)$$

For both the world and primitive action belief update in eq-3, the components $p(W_t|o_{w,t})$ and $p(\alpha_t|o_{w,t})$ are derived from Wang and Hoey [19]. The Bayesian update is as follows:

$$p(\alpha_t|o_{w,t}) \propto \sum_{w_{t-1} \in W_{t-1}} \sum_{w_t \in W_t} p(\alpha_t, o_{w,t}, w_{t-1}, w_t), \quad (4)$$

$$p(W_t|o_{w,t}) = \sum_{w_{t-1} \in W_{t-1}} \sum_{\alpha_{i,t} \in \alpha_t} p(\alpha_{i,t}, o_{w,t}, w_{t-1}, w_t). \quad (5)$$

We adopt the proposed algorithm for goal recognition, $p(G_t|W_t)$ in Wang and Hoey [19]. The algorithm maintains a goal belief distribution by generating a probabilistic explanation set - *ExplaSet*. Each *expla* \in *Explaset* uses HTN planning to explain the observations so far. The probability of each goal g_i in G given the world state is the sum of probabilities of *expla* \in *ExplaSet* whose *PredictedGoal* $= g_i$. Our algorithm reweighs primitive actions probabilities based on the language observational model described below, influencing the world belief update according to Eq-5 and, consequently, the goal recognition update.

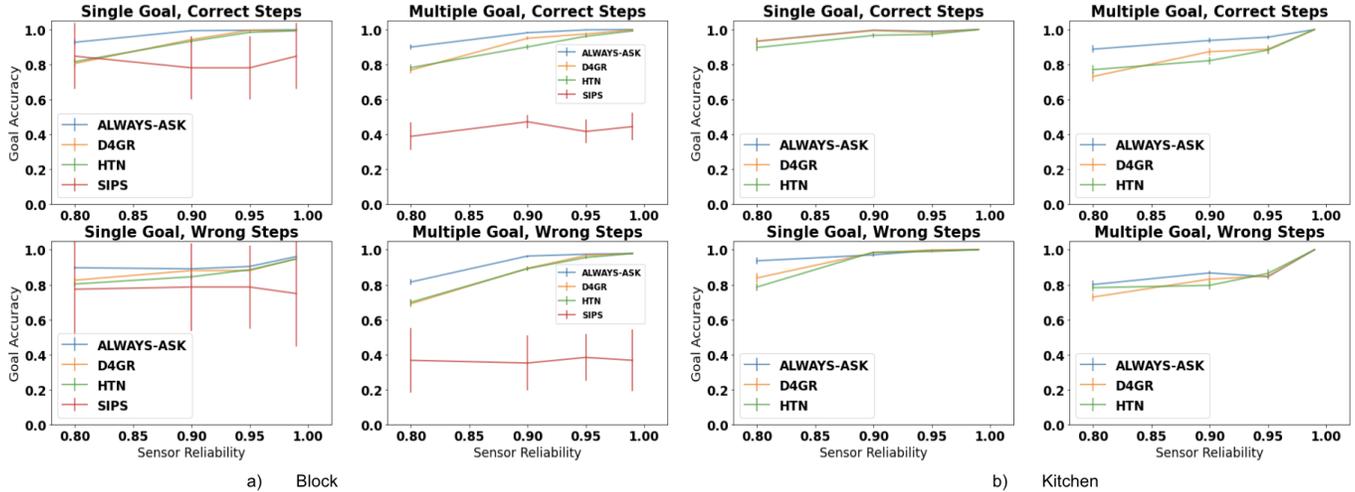


Fig. 4. Results for Top1 Goal Accuracy versus Sensor Reliability for the two domains kitchen and block

The derivation of the language observational model is:

$$p(\alpha_t, q_t | o_{i,t}) * p(o_{i,t}) \propto p(o_{i,t} | \alpha_t, q_t). \quad (6)$$

We adopt a bag-of-words approach as our POMDP’s observational model instead of utilizing a large language model (LLM) like GPT3. LLMs are not inherently grounded. Our model explicitly establishes a connection between sensor information and semantics through a transition model in the POMDP. Although LLMs could be incorporated for intent classification using the right prompt, we did not pursue this direction as it falls outside the focus of our paper.

To estimate the effect of the language observation o_i on α and q , we calculate $p(o_{i,t} | \alpha_t, q_t)$. For this, we consider three possibilities for the state: If the highest belief primitive action $\alpha_{max,t}$ is the same as the question asked, then the user is likely to respond with positive/confirmation feedback. The opposite is true if $\alpha_{max,t} \neq q_t$. If $q_t = Null$, then no question has been asked, so both types of responses are equally likely. The mathematical representation of $p(o_{i,t} | \alpha_t, q_t)$ is governed by the following conditional probability table:

	$p(o_{i,t} = Yes)$	$p(o_{i,t} = No)$
$\alpha_{max,t} = q_t$	0.99	0.01
$\alpha_{max,t} \neq q_t$	0.01	0.99
$q_t = Null$	0.5	0.5

TABLE I
CONDITIONAL PROBABILITY FOR $p(o_{i,t+1} | \alpha_t, q_t)$

At each time step as the human performs an action, we solve the MOMDP using the POUCT solver [17] to approximate the optimal policy for the robot’s communication with the human. The observational model is then employed to update the robot’s belief of user’s mental state $M = \{G, \alpha\}$.

IV. EVALUATION

Our evaluation aims to test the hypothesis that our hierarchical decision-theoretic framework D4GR improves 1) the accuracy of goal recognition and plan recognition of human

activity and 2) the robot’s ability to guide the person towards task success. We evaluate the enhanced performance of our algorithm by measuring the accuracy of goal recognition and prediction of the next human action, also referred to as plan recognition, at every time step. We also measure the planning time, the cumulative expected return, and the number of clarification questions asked for completing the tasks by D4GR and compare it against the three presented methods in simulation. We also perform a robot demonstration of D4GR for the scenario where the human switches between washing hand and making tea tasks.

We use the simulation environment introduced by Wang and Hoey [19] for our experiments. The simulator models real environment state changes that result from the primitive actions specified in the HTN for the virtual human. In the simulator, 44 binary virtual sensors are observing the world state. Some of the examples include sensors for *hand_dry*, *faucet_on*, *block_picked_up*. For our experiments, we vary sensor reliability from 99% to 80%.

A. Domain and Experiment Test Cases

We test our algorithm in two domains: a block domain and a kitchen domain. The Knowledge Base, *TaskNet* of the HTN for the kitchen has three goals: *wash hands*, *make tea*, and *make coffee*. and the block domain has five goals of stacking blocks to make words with varying lengths of 4 to 7. The five goals of our block domain are *rote*, *tone*, *tune*, *hawk*, and *capstone*. The two domains differ in their HTN planning structure, as the block’s domain has higher goals (root nodes) but a shorter tree depth than the kitchen domain. Such a setup allows us to explore the effect of HTN structure on goal and plan recognition performance. We evaluate the performance of our algorithm over four categories of test case scenarios:

Single Goal & Correct Steps: Captures scenarios where the human always executes a correct action sequence for achieving a single goal.

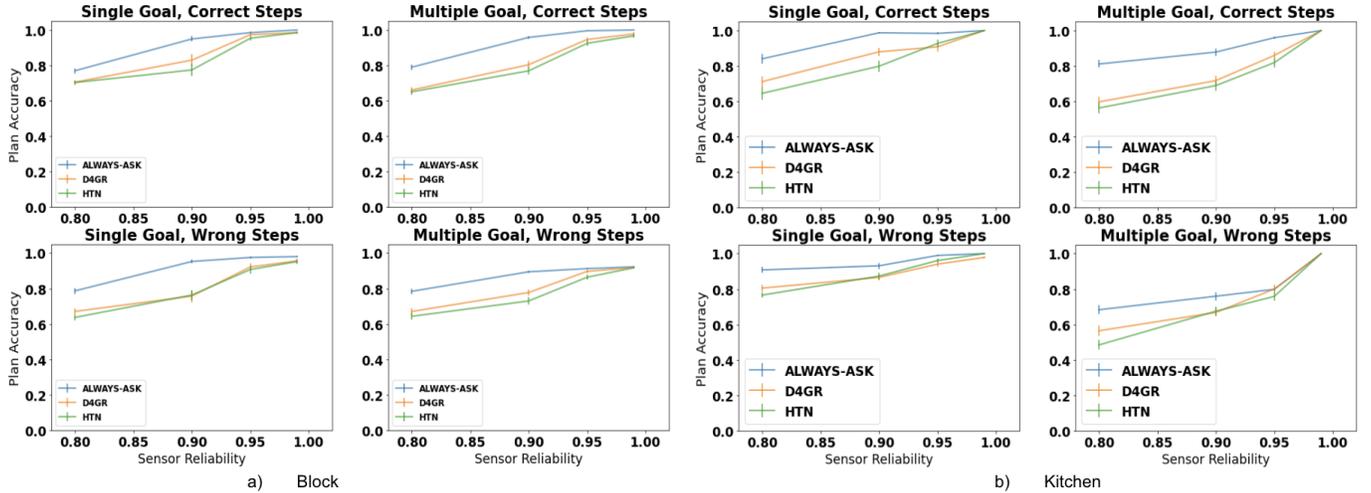


Fig. 5. Results for Top1 Plan Accuracy versus Sensor Reliability for the two domains kitchen and block

Multiple Goals & Correct Steps: The person works on multiple goals simultaneously by switching back and forth,

Single Goal & Wrong Steps: A human has a single goal but can execute wrong steps affecting progress toward the goal.

Multiple Goals & Wrong Steps: The human moves back and forth between goals and executes wrong actions.

The easiest case is Single Goal & Correct Steps, and the hardest is Multiple Goals & Wrong Steps.

B. Baselines:

We compare D4GR’s performance in simulation with three other methods. Our first baseline is **HTN**, a previous method of HTN-based goal recognition introduced by Wang and Hoey [19]. This method passively incorporates partially observable world observations for PGR without engaging with the user. Our second method is **ALWAYS-ASK** which acts as an oracle that always asks the correct clarification question and uses the language feedback for the belief update of D4GR. This baseline always has the highest goal and step recognition accuracy but receives a lower reward because it asks unnecessary questions. Our third baseline is **SIPS**¹ introduced by Zhi-Xuan et al. [22]. This algorithm is not equipped to handle hierarchical nature of goals.

C. Metric Definitions:

We measure 1) Top 1 Accuracy for Goal recognition, 2) Accuracy for Plan Recognition similar to Wang and Hoey [19] averaged over all timesteps. These metrics measure the accuracy of our belief update. To evaluate our POMDP formulation, we also measure the planning time taken: runtime averaged over the steps, the cumulative reward for the whole human action sequence, and the number of questions asked averaged over trials.

¹The SIPS baseline is adopted from the code repository cited in Zhi-Xuan et al. [22] and uses their default setting: static goal transition.

V. RESULTS

Our proposed algorithm aims to improve the capability of the robot for goal and plan recognition. The performance of D4GR depends on how accurately our cognitive assistive robot estimates the belief states for the human mental model M_t : the likelihood of goals G and the human actions α at each simulated step. The ground truth of each human action’s α given the goal G can be obtained from the knowledge base.

A. Exp 1 - Goal Accuracy Performance

In Fig-4, we present results for the average goal accuracies of D4GR and compare them with the baselines over varying sensor reliability and test case categories. The reason for choosing the sensor reliability range from [0.8 to 0.99] is because most of the deep-learned vision and human action detectors have similar average accuracy [14]. Overall, as the sensor reliability decreases, the accuracy performance of HTN-based methods (ALWAYS-ASK, D4GR, HTN) suffers. The oracle baseline, ALWAYS-ASK, always has the highest goal accuracy. Even at lower sensor reliabilities (higher sensor noise), D4GR’s accuracy remains higher than HTN in all experiment categories by 1% on average in both domains. This trend indicates that even when the sensor’s observational model fails, D4GR can better predict the belief states than HTN. The SIPS method did not generate functional plans for our kitchen domain even when the input specification was correct. Their particle filter algorithm could not find feasible plans for the goals. Hence, presenting the results in the blocks domain. Compared to the SIPS baseline, our method is 30 % better in the blocks domain. We significantly outperform the SIPS baseline in the multiple goals scenario by almost 43% because SIPS does not handle the hierarchical nature of goals. The problem categories with single goals (correct steps & wrong steps) have the best performance for the lowest sensor reliability. We see D4GR performance boost by 2.8% in the kitchen domain and 1.4% in the blocks domain as compared to

Domain	Method	Runtime (s) ↓				Cumulative Expected Return ↑				Question Frequency ↓			
		0.8	0.9	0.95	0.99	0.8	0.9	0.95	0.99	0.8	0.9	0.95	0.99
Block	ALWAYS-ASK	2.085	1.869	1.563	1.367	-37.253	-62.769	-69.323	-73.569	1.000	1.000	1.000	1.000
	D4GR	28.750	20.445	18.814	17.084	-26.056	-24.656	-23.014	-23.989	0.297	0.318	0.325	0.358
	HTN	9.056	1.524	1.371	1.040					0.000	0.000	0.000	0.000
	SIPS	43.173	41.615	40.654	40.934								
Kitchen	ALWAYS-ASK	0.484	0.425	0.407	0.343	-41.848	-55.593	-62.342	-73.265	1.000	1.000	1.000	1.000
	D4GR	13.912	13.439	13.572	11.384	-24.527	-26.542	-26.399	-21.437	0.312	0.317	0.316	0.296
	HTN	0.415	0.380	0.391	0.349					0.000	0.000	0.000	0.000

TABLE II
TREND IN QUESTIONS ASKED, REWARDS ACCUMULATED AND RUNTIME (SENSOR RELIABILITY VARIES FROM 80% TO 99%)

HTN. The Oracle, on average, is 6.3% more accurate than the HTN baseline in this category. Our D4GR improves accuracy by inferring when and what to ask a question rather than always asking.

B. Exp 2 - Plan Accuracy Performance

Similar to goal accuracy, we plot the planning accuracy for D4GR and compare it with the baselines for varying sensor noise in Fig-5. Our algorithm overall is 3% more accurate than HTN in both domains. For the lowest sensor reliabilities, D4GR is 4% better than HTN in the kitchen domain and 2% better than HTN in the block domain. For the multiple goal scenarios (correct and wrong steps), D4GR performs the best with an accuracy improvement of 2.7% in the kitchen domain and 2.3% in the block domain.

C. Exp 3 - Trend in Questions Asked, Rewards Accumulated and Runtime:

Our proposed algorithm aims to improve the PGR capabilities of the agent by enabling the robot to engage with the users and ask for language feedback. One performance measure is the number of helpful clarification questions asked with varying sensor reliabilities in Table-II. When the environment and human actions are unambiguous (sensor reliability is high and/or the human is performing correct actions), D4GR enables the robot to intelligently infer that it does not need to ask lots of questions. At sensor reliability 0.99, D4GR still asks questions because users can do sub-optimal actions leading to ambiguity. Overall the agent asks questions 32.4% and 31% of the time in the block and kitchen domains respectively. When compared over varying sensor noise, the change in the number of questions asked is insignificant; the numbers lie within the same standard deviation.

Further, asking a large number of clarification questions, especially if they are not relevant to the current progress of the task, takes more computing resources since planning will have to be done at every timestep. This effect is measured by the runtime and the reward accumulated by D4GR. Our reward function penalizes asking a lot of questions, especially when they reference an irrelevant human action. Our results show that D4GR takes more planning time per step as compared to ALWAYS-ASK and HTN but is 48% faster than SIPS. All these measurements are done on a machine with 31 GB RAM, Intel® Core™ i7-9750H CPU @ 2.60GHz × 12. The code was run single-threaded. D4GR takes more time than

HTN but enables accuracy performance boost as noted in previous sections, Sec-V-A and Section-V-B. Further, D4GR accumulates a higher reward/lower penalty as compared to ALWAYS-ASK by 58% in both domains highlighting that D4GR does not ask unnecessary questions.

D. Robot Demonstration

We performed a robot demonstration to highlight the feasibility of D4GR in the real world. The Kuri robot was used primarily for the demo due to its audio transcription capability. The demonstration consisted of a human performing two interleaving tasks in a kitchen while the Kuri robot observed the actions performed and engaged with the user using the D4GR algorithm. The human begins with making coffee and then moves to wash their hands. Unlike HTN, which struggled to recognize the change in goal, D4GR correctly identified both goals. The demonstration simulated sensor reliability at 0.8 using an oracle and sensor noise model. Kuri used D4GR to intelligently infer when to ask a question and what to ask and is able to perform goal recognition and planning correctly. In the case of negative feedback from humans, Kuri offers its predicted correct step to the human.

VI. DISCUSSION AND FUTURE WORK

Our deployed algorithm D4GR shows improved accuracy for goal and plan recognition than the baseline *HTN* and *SIPS*. It does so while asking fewer questions than the ALWAYS-ASK oracle policy. Our deployed robot with D4GR performs real-time communication, as demonstrated. The time taken for online planning is influenced by two critical parameters of the POMDP solver: 1) d , representing the finite depth of the probabilistic decision tree constructed with state-action pairs, and 2) n , the finite number of observations sampled from each node. Increasing d and n enhances the solver's accuracy and increases runtime. To achieve real-time communication, we conducted empirical experiments and determined that setting $d = 19$ and $n = 6$ provided appropriate action choices within a reasonable time.

Our algorithm shows promising improvements in PGR accuracy, although it comes with increased runtime compared to HTN. Our algorithm is designed to assist users with cognitive impairment in their daily tasks, focusing on non-time-critical activities. By providing delayed feedback, our social assistive robot increases the likelihood of users learning from mistakes

and avoiding continuous repetition of errors compared to using HTN. We can reduce runtime further by retaining only the highest probable explanation sets, denoted as *ExplaSet* in HTN planning. However, this impacts the PGR accuracy since *ExplaSet* with multiple goals during initial steps can get pruned due to their lower probabilities. In the multi-goal and low sensor reliability setting, D4GR shows slightly lower PGR accuracy than HTN. This is due to D4GR’s reliance on noisy beliefs and user switching goals, leading to a higher probability of asking questions about irrelevant actions. The rational language feedback also adversely affects the update of the explaset, potentially diminishing its utility in later timesteps of the episode.

Our work is also limited by the type of clarification questions the robot can ask. We have a fixed template for the question. It will be interesting to see how humans respond to various clarification strategies and how the robot can plan over a space of such categories. This will increase the action space requiring more exploration by the POMDP solver. Further, our language observational model is a bag of words model. It can be more expressive by incorporating inference from LLMs.

Further, our work assumes access to a pre-defined knowledge base for the tasks. One thing that we will be exploring in the future is how to make the knowledge base adaptive to a layman user’s needs and preferences as the task progresses through interactive dialogue. Our research opens venues for language grounding and human intent recognition in other collaborative tasks like building machines/complex furniture together by humans and robots. This is an encouraging step toward enhancing the sensory capabilities of home-service robots that can assist people in completing tasks with language-based interactions.

VII. CONCLUSION

We propose a novel algorithm for robots to interactively keep track of people’s ongoing progress in a task using questions. Moreover, our D4GR framework can suggest plan improvements to users in solving a task if required. Our work shows that: 1) modeling the user as an HTN and incorporating language feedback improves robots’ belief of human’s progress in simulation; 2) POMDPs are effective methods for tracking a task’s progress and asking clarification questions. Our D4GR formulation has a similar goal and step recognition accuracy as the best baseline ALWAYS-ASK method while asking 68% fewer questions. In future work, we aim to conduct a user study with the targeted population to measure our approach’s usefulness during the interaction. D4GR’s ability to intelligently balance between clarifying uncertainty with a lesser number of questions allows for realistic interactions between a social robot and a human. This ability in the future can allow for realistic interactions with human users during collaborations over tasks between humans and robots.

VIII. ACKNOWLEDGEMENT

We thank our labmates at Brown and Rutgers for their valuable insights. This work was supported by NSF under award

number IIS-1652561, ONR under award numbers N00014-21-1-2584 and N00014-22-1-2592, and with funding from Echo Labs.

REFERENCES

- [1] Dan Bohus and Eric Horvitz. Facilitating multiparty dialog with gaze, gesture, and speech. In *ICMI-MLMI*, ICMI-MLMI ’10, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781450304146. doi: 10.1145/1891903.1891910.
- [2] Joyce Y Chai, Qiaozi Gao, Lanbo She, Shaohua Yang, Sari Saba-Sadiya, and Guangyue Xu. Language to action: Towards interactive task learning with physical agents. In *IJCAI*, pages 2–9, 2018.
- [3] Finale Doshi and Nicholas Roy. Efficient model learning for dialog management. In *HRI*, pages 65–72, 2007.
- [4] Kutluhan et al Erol. Htn planning: Complexity and expressivity. In *AAAI*, volume 94, pages 1123–1128, 1994.
- [5] Juan Fasola and Maja J Mataric. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, 2(2):3–32, 2013.
- [6] Jesse Hoey, Pascal Poupart, Axel von Bertoldi, Tammy Craig, Craig Boutlier, and Alex Mihailidis. Automated handwashing assistance for persons with dementia using video and a partially observable markov decision process. *Computer Vision and Image Understanding*, 114(5): 503–519, 2010.
- [7] Daniel Höller, Gregor Behnke, Pascal Bercher, and Susanne Biundo. Plan and goal recognition as htn planning. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 466–473. IEEE, 2018.
- [8] Ifrah Idrees, Stefanie Tellex, and Momotaz Begum. Grounding human-robot dialogue in environment sensors to help elderly in activities of daily living. *Robodial Workshop - SigDial*, 2020.
- [9] 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.
- [10] Cory D Kidd and Cynthia Breazeal. Robots at home: Understanding long-term human-robot interaction. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3230–3235. IEEE, 2008.
- [11] Felipe Rech Meneguzzi and Ramon Fraga Pereira. A survey on goal recognition as planning. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), 2021, Canada.*, 2021.
- [12] Reuth Mirsky, Roni Stern, Meir Kalech, et al. Sequential plan recognition. *arXiv preprint arXiv:1703.01083*, 2017.
- [13] Sylvie CW Ong, Shao Wei Png, David Hsu, and Wee Sun Lee. Pomdps for robotic tasks with mixed observability. In *Robotics: Science and Systems*, volume 5, page 4, 2009.
- [14] Preksha Pareek and Ankit Thakkar. A survey on video-based human action recognition: recent updates, datasets, challenges, and applications. *Artificial Intelligence Review*, 54(3):2259–2322, 2021.
- [15] Miquel Ramirez and Hector Geffner. Plan recognition as planning. In *Twenty-First IJCAI*, 2009.
- [16] Muhammad Asif Razzaq, Wajahat Ali Khan, and Sungyoung Lee. Intent-context fusioning in healthcare dialogue-based systems using jdl model. In *ICOST*, pages 61–72. Springer, 2017. doi: 10.1007/978-3-319-66188-9_6.
- [17] David Silver and Joel Veness. Monte-carlo planning in large pomdps. *Advances in neural information processing systems*, 23, 2010.
- [18] Jesse Thomason, Aishwarya Padmakumar, and et al. Improving grounded natural language understanding through human-robot dialog. In *ICRA*, pages 6934–6941. IEEE, 2019.
- [19] Dan Wang and Jesse Hoey. Hierarchical task recognition and planning in smart homes with partial observability. In *International Conference on Ubiquitous Computing and Ambient Intelligence*, pages 439–452. Springer, 2017.
- [20] David Whitney, Eric Rosen, James MacGlashan, Lawson LS Wong, and Stefanie Tellex. Reducing errors in object-fetching interactions through social feedback. In *2017 (ICRA)*, pages 1006–1013. IEEE, 2017.
- [21] Steve Young, Milica Gašić, Blaise Thomson, and Jason D Williams. Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179, 2013.
- [22] Tan Zhi-Xuan, Jordyn Mann, Tom Silver, Josh Tenenbaum, and Vikash Mansinghka. Online bayesian goal inference for boundedly rational planning agents. *NEURIPS*, 33:19238–19250, 2020.