On Trust-aware Assistance-seeking in Human-Supervised Autonomy

Dong Hae Mangalindan, Ericka Rovira, and Vaibhav Srivastava

***Abstract*— Using the context of human-supervised object collection tasks, we explore policies for a robot to seek assistance from a human supervisor and avoid loss of human trust in the robot. We consider a human-robot interaction scenario in which a mobile manipulator chooses to collect objects either autonomously or through human assistance; while the human supervisor monitors the robot's operation, assists when asked, or intervenes if the human perceives that the robot may not accomplish its goal. We design an optimal assistance-seeking policy for the robot using a Partially Observable Markov Decision Process (POMDP) setting in which human trust is a hidden state and the objective is to maximize collaborative performance. We conduct two sets of human-robot interaction experiments. The data from the first set of experiments is used to estimate POMDP parameters, which are used to compute an optimal assistance-seeking policy that is used in the second experiment. For most participants, the estimated POMDP reveals that humans are more likely to intervene when their trust is low and the robot is performing a high-complexity task; and that robot asking for assistance in high-complexity tasks can increase human trust in the robot. Our experimental results show that the proposed trust-aware policy yields superior performance compared with an optimal trust-agnostic policy.**

I. Introduction

While autonomous systems are becoming pervasive across a range of domains including, healthcare, agriculture, and transportation, they often require human assistance to deal with complex and uncertain environments [1]–[4]. One of the key factors in the successful design of such human-robot teams is the calibration of the human partner's trust in the robot's capabilities. Trust in robot/automation is defined as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability [5]". The disuse, misuse, and abuse of automation and its underpinnings in terms of under- and over-trust in automation is well-documented [6]–[8]. Therefore, trust must be accounted for not just in the physical design of robots, but also in the design of its policies in complex and uncertain environments.

In this paper, using the context of human-supervised object collection tasks, we study trust-aware policies for an autonomous system to seek assistance from a human supervisor. We adopt a Partially Observable Markov Decision Process (POMDP) framework to design optimal trust-aware policies and compare them with trust-agnostic optimal policies which as opposed to the trust-aware policy, do not take into account the state of human trust.

This work was supported by NSF award ECCS 2024649.

D. Mangalindan and V. Srivastava are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824 e-mail: {mangalin, vaibhav}@egr.msu.edu

E. Rovira is with the Department of Engineering Psychology, U.S. Military Academy, West Point, NY 10996 e-mail: ericka.rovira@westpoint.edu

The influence of factors such as reliability, transparency, and workload on humans has been extensively explored [6]. However, these traditional studies have treated trust as a static parameter. Lately, a significant focus has been devoted to understanding the dynamic evolution of human-automation trust [2], [9]–[17]. These works focus on the dynamic evolution of trust and the influence of factors such as the quality of robot performance, transparency of the robot's operation, and forgetfulness of the human agent about prior robot performance. From a modeling perspective, broadly speaking, the models can be divided into deterministic linear models and probabilistic models.

The deterministic linear models [13]–[15] use a linear time-invariant dynamical system in which the states are variables such as current trust level, cumulative trust level, and expectation bias, and input includes human experience in terms of performance and reliability. These models have also been extended to capture mutual trust between humans and robots [18], [19], and have been used to study the influence of information transparency in reconnaissance missions [14], human reliance on driving assistance systems [13], subtask allocation in collaborative assembly tasks [18], and scheduling teleoperation in underwater robotic navigation [19].

The probabilistic models treat the trust as a hidden state and estimate its distribution conditioned on variables such as robot performance and human actions in a Bayesian setting [20]. The POMDP models [2], [16], [17], [21]–[23] are a class of probabilistic models that treat human mental states, such as trust and workload, as discrete (usually, binary) hidden variables and infer their transition matrices conditioned on human and robot actions, environmental state, and task outcome. These models have been used to design task schedules in human-supervised robotic operations such that human interventions [16] and monitoring rate [21] are minimized. These models have also been used to design optimal recommendations and associated levels of explanation based on human trust and workload [2], [22]. In contrast to these works, we focus on a scenario where the robot may seek assistance from a human operator to avoid loss of trust and design an optimal assistance-seeking policy.

Using the context of human-supervised object collection tasks, we study trust-aware policies for an autonomous system to seek assistance from a human supervisor. We adopt a POMDP framework in which the future trust state depends on the current trust state, robot action, environmental complexity, and prior experience of the human supervisor. We characterize experience by the success/failure of the robot in accomplishing its goals. The observations in the POMDP model are human actions, which depend on the current trust, environmental complexity, and robot actions. We conduct two sets of human subject experiments. In the first set, we adopt a randomized assistance-seeking policy and adopt an Input-Output Hidden Markov Model (IOHMM) [24] framework to estimate the parameters of the POMDP model. In the second set, we leverage the learned POMDP to design a trust-aware optimal assistance-seeking policy and show that it achieves superior team performance than a trust-agnostic optimal policy. The major contributions of this work are twofold. First, we design an optimal trust-aware assistance-seeking policy for the robot such that the team performance is maximized. Second, we validate the efficacy of the proposed policy through human subject experiments.

The remainder of the paper is organized as follows. We review key ideas of IOHMM and POMDP in Section II. We describe the collaborative object collection task in Section III and design the optimal trust-aware assistance-seeking policy in Section IV. We discuss the estimated human behavioral model and its implications in Section V. We compare trust-aware and trust-agnostic policies through human-in-the-loop experiments in Section VI and we conclude in Section VII.

II. Background: IOHMM and POMDP

Let $\{\bar{S}_t \in \bar{\mathcal{S}}\}_{t\in\mathbb{N}}$ be a Markov chain with an input $\{u_t \in \mathcal{U}\}_{t\in\mathbb{N}}$, i.e., the probability distribution of $\bar{S}_{t+1}$ is completely determined by $\bar{S}_t$ and u_{t+1}. Here, $\bar{\mathcal{S}}$ and $\mathcal{U}$ are some finite sets. In an IOHMM [24], the state $\bar{S}_t$ is not available for measurement (it is hidden), and only an output $\{y_t \in \mathcal{Y}\}_{t\in\mathbb{N}}$ is measured, which is a realization of an unknown probability distribution $\mathbb{P}(y_t|\bar{S}_t, u_t)$. Here, $\mathcal{Y}$ is some finite set.

Given $\{y_t, u_t\}_{t\in\mathbb{N}}$, $\bar{\mathcal{S}}$, and $\mathcal{Y}$, a key inference goal for IOHMM is to estimate the state transition probabilities $\mathbb{P}(\bar{S}_{t+1}|\bar{S}_t, u_{t+1})$, emission probabilities $\mathbb{P}(y_t|\bar{S}_t, u_t)$, and the prior distribution of S_1. An extended Baum-Welch algorithm [24] has been developed to estimate these probabilities.

A POMDP is a generalization of the IOHMM in which some of the inputs are controllable and other inputs are assumed Markovian and included in the state space. Let $\{S_t \in \mathcal{S}\}_{t\in\mathbb{N}}$ be the augmented state Markov chain, $\{a_t \in \mathcal{A}\}_{t\in\mathbb{N}}$ be the action sequence, i.e., the sequence of controllable inputs, and $\{o_t \in \mathcal{O}\}_{t\in\mathbb{N}}$ be the observation sequence. Here, $\mathcal{S}, \mathcal{A}$, and $\mathcal{O}$ are finite sets. The POMDP is defined by state transition probabilities $\mathbb{P}(S_{t+1}|S_t, a_t)$, observation probabilities $\mathbb{P}(o_t|S_t, a_t)$, and a reward function $\mathcal{R}(S_t, a_t)$. Note that state S_t includes hidden state $\bar{S}_t$ and non-controlled inputs in the IOHMM and o_t includes perfect measurements of the non-controlled inputs and output y_t in the IOHMM.

III. Human-Supervised Robotic Object Collection Experiment

We consider a human-supervised object collection task in which a human and a mobile manipulator collaborate to collect incorrectly placed items on grocery shelves and deposit them in a bin attached to the manipulator. The human supervisor has access to the live feeds of three views: the world view, a local view, and an end-effector view; see Fig. 1 for a snapshot of the experiment interface. The human supervisor can teleoperate the manipulator with a joystick.

We refer to each attempt by the manipulator to collect an object as a trial. At each trial, the mobile manipulator can either collect an object autonomously or can ask for human assistance, which allows the human to teleoperate the manipulator and collect the item. We refer to the former robot action by a^- and the latter by a^+. The human supervisor monitors the robot's operation for safety and efficiency. She teleoperates the manipulator to collect objects either to assist the robot when asked or to voluntarily intervene when she perceives that the robot may fail. We refer to the observed human action, when not asked to assist, as *intervene* o^- when she intervenes and *rely* o^+, otherwise.

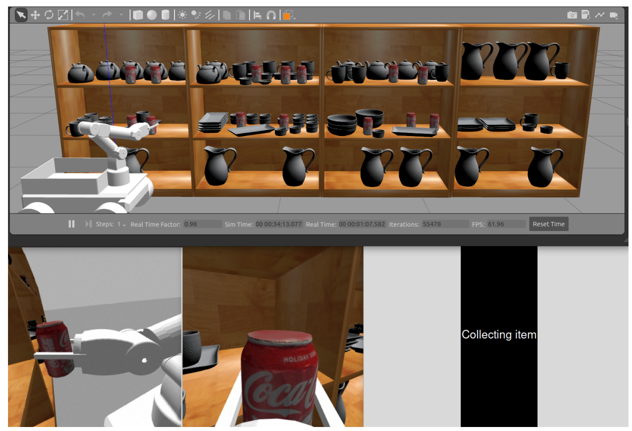


Fig. 1: Experimental Setup. The top figure shows the world view of the environment available to the human supervisor. The bottom left and middle figures show the local view and end-effector view used by the human during teleoperation. The bottom right figure shows the operation status. The setup uses ROS-Gazebo [25] and resources available in [26], [27].

The complexity C of a trial is determined by the presence of an obstacle in the direct path between the manipulator and the object to be collected, and it can either be low C^L or high C^H. We assume that the complexity of trials is an i.i.d. Bernoulli random variable with p_{C^H} being the probability of the complexity being high.

The outcome of each trial can be a success or a failure. An autonomous object collection is considered a failure if the robot does not safely collect the object, i.e., if the robot collides with its surroundings and/or fails to deposit the object in the bin. Otherwise, autonomous object collection is considered a success. We assume that autonomous object collection is successful with probabilities p_L^{suc} and p_H^{suc} in environments with low and high complexities, respectively. We also assume that, when teleoperating, the human always successfully collects the item and deposits it in the bin.

We define the human experience E_{t+1} at the end of trial t by the trial's outcome (robot performance). For an autonomous operation, the experience is reliable E^+, when the robot succeeds; and faulty E^-, when the robot fails. For a voluntary intervention by the supervisor, the experience is faulty E^-. Since the robot is expected to operate without assistance in low-complexity trials, assistance seeking in high complexity is labeled as reliable E^+, and assistance seeking in low complexity is labeled as faulty E^-.

The human-robot team objective is to maximize the discounted cumulative team reward across all trials. The reward in each trial is defined by

$$\mathcal{R}^a_{o,E} = \begin{cases} +2, & \text{if } (a, o, E) = (a^-, o^+, E^+), \\ +1, & \text{if } a = a^+, \\ 0, & \text{if } (a, o) = (a^-, o^-), \\ -3, & \text{if } (a, o, E) = (a^-, o^+, E^-). \end{cases} \tag{1}$$

The reward in (1) is designed to ensure that the number of collected objects is maximized while minimizing human effort from assistance and intervention.

IV. Trust-aware Optimal Assistance-seeking

The human-in-the-loop system studied in this paper is shown in Fig. 2, and in this section, we describe its different subsystems.

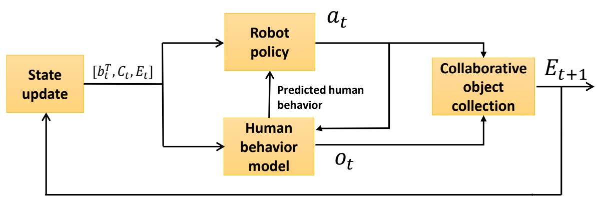


Fig. 2: Collaborative object collection setup. The human behavioral model with trust as a hidden state yields the likelihood of human intervention and is used by the robot to design an optimal policy in a POMDP setting.

A. Trust-modulated Human Behavioral Model

The human behavioral model determines the probability of the human taking action o^- when the robot takes action a^-. We assume that this probability depends on the complexity of the trial and a hidden state that captures the human perception of the expected robot performance (success or failure). We refer to this hidden state as the *trust* T_t and assume it takes binary values: high T^H and low T^L.

Building upon existing models of the dynamic evolution of trust [13], we assume that trust T_{t+1} at the end of trial t is influenced by trust T_t at the end of trial $t-1$, complexity C_t of trial t, experience E_{t+1} at the end of trial t, and action a_t of the robot in trial t. Fig. 3 shows the IOHMM-based dynamic human behavioral model that is modulated by the hidden trust dynamics. Once this IOHMM has been trained, it yields the probabilities $\mathbb{P}(T_{t+1}|T_t, E_{t+1}, C_t, a_t)$ and $\mathbb{P}(o_t|T_t, a_t, C_t)$.

B. POMDP-based Optimal Assistance-Seeking Policy

We pose the design of optimal assistance-seeking policy as a POMDP. The elements of this POMDP are

States, actions, and observations. We take the POMDP states as $S_t = (T_t, E_t, C_t) \in \{T^L, T^H\} \times \{E^-, E^+\} \times \{C^L, C^H\}$, where T_t is a hidden state and (E_t, C_t) are observed states; the actions $a_t \in \{a^-, a^+\}$; and observations $o_t \in \{o^-, o^+\}$.

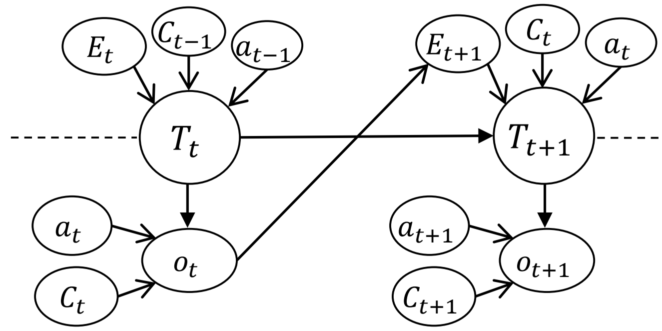


Fig. 3: IOHMM-based dynamic human behavioral model. The hidden trust dynamics T_t modulate the human behavioral outcome o_t.

State transition matrices and observation probabilities. Using the IOHMM structure in Fig. 3 and the fact that C_t's are i.i.d., we can write

$$\begin{aligned}\mathbb{P}(S_{t+1}|S_t) = \mathbb{P}(T_{t+1}|E_{t+1}, T_t, C_t, a_t) \\ \times \mathbb{P}(C_{t+1})\mathbb{P}(E_{t+1}|T_t, C_t, a_t),\end{aligned}$$

where $\mathbb{P}(T_{t+1}|E_{t+1}, T_t, C_t, a_t)$ is known from the human behavioral model, $\mathbb{P}(C_{t+1} = C^H) = p_{C^H}$ by design, and $\mathbb{P}(E_{t+1}|T_t, C_t, a_t)$ is computed as

$$\mathbb{P}(E_{t+1}|T_t, C_t, a_t) = \sum_{o_t \in \mathcal{O}} \mathbb{P}(E_{t+1}|o_t, C_t, a_t)\mathbb{P}(o_t|T_t, C_t, a_t)$$

where $\mathbb{P}(o_t|T_t, C_t, a_t)$ is known from the trained human behavior model, and

$$\begin{aligned}&\mathbb{P}(E_{t+1} = E^+|o_t, C_t, a_t) \\ &\quad = \begin{cases} p^{\text{suc}}_L, & \text{if } (a_t, C_t, o_t) = (a^-, C^L, o^+), \\ p^{\text{suc}}_H, & \text{if } (a_t, C_t, o_t) = (a^-, C^H, o^+), \\ 0, & \text{if } (a_t, C_t, o_t) = (a^-, *, o^-), \\ 0, & \text{if } (a_t, C_t, o_t) = (a^+, C^L, *), \\ 1, & \text{if } (a_t, C_t, o_t) = (a^+, C^H, *). \end{cases}\end{aligned}$$

Reward function and discount factor. Using the reward defined in (1), the reward function for the POMDP is

$$\begin{aligned} R(S, a^-) &= \mathcal{R}^{a^-}_{o^+,E^+}\mathbb{P}(E^+|o^+, C, a^-)\mathbb{P}(o^+|T, C, a^-) \\ &\quad + \mathcal{R}^{a^-}_{o^+,E^-}\mathbb{P}(E^-|o^+, C, a^-)\mathbb{P}(o^+|T, C, a^-) \\ &\quad + \mathcal{R}^{a^-}_{o^-,E^+}\mathbb{P}(o^-|T, C, a^-), \\ R(S, a^+) &= \mathcal{R}^{a^+}_{*,*}, \end{aligned} \tag{2}$$

where, for brevity, we have represented events by their outcomes; for example, we have represented the event $\{E = E^+\}$ by E^+. The discount factor γ is chosen as 0.987.

C. Belief MDP Reformulation of the POMDP

To solve the above POMDP, we reformulate it as an equivalent MDP with belief states $S^b = (b^T_t, E_t, C_t)$, where $b^T_t = \mathbb{P}(T_t = T^H|E_t, C_{t-1}, a_{t-1}, b^T_{t-1})$. The action set and the discount factor will be the same as the POMDP formulation. The belief update can be written as

$$b_{t+1}^T = \mathbb{P}(T_{t+1} = T^H | T_t = T^H, E_{t+1}, C_t, a_t) b_t^T + \mathbb{P}(T_{t+1} = T^H | T_t = T^L, E_{t+1}, C_t, a_t)(1 - b_t^T). \quad (3)$$

Likewise, $\mathbb{P}(E_{t+1} | b_t^T, C_t, a_t)$ is computed by computing the expected value of $\mathbb{P}(E_{t+1} | T_t, C_t, a_t)$ over T_t. The reward function for the belief MDP is computed by computing the expected value of $R(S_t, a)$ in (2) over T_t. For computational purposes, we discretize b_t^T and use update (3) to compute its state transition matrix.

V. Estimated Human Behavioral Model

To estimate the human behavioral model, we conducted a human subject experiment[1]. We recruited a total of 9 participants for this experiment. Each participant performed 71 trials out of which 41 involved a low complexity environment, while the remaining 30 involved high complexity environment. When the robot is operating autonomously, it has a success probability $p_H^{\text{suc}} = 0.83$ in the high complexity environments and $p_L^{\text{suc}} = 0.97$ in low complexity. In each trial, the robot asks for human assistance with a probability of 0.10, if the environment complexity is low; and with a probability of 0.4, if the environment complexity is high. For each trial, the data collected are human action o, robot action a, complexity C, and experience E.

The data collected from all participants are pooled together and are used to estimate the parameters of the human behavioral model using an extended version of the Baum-Welch algorithm. We now present and discuss the estimated parameters in the human behavioral model. For brevity, we have represented events by their outcomes.

A. Initial Trust and Observation Probabilities.

The initial probabilities of trust being high are estimated to be $\mathbb{P}(T^H) = 1$. This indicates that humans are estimated to start the session with high trust. The estimated observation probabilities $\mathbb{P}(o_t | T_t, C_t, a_t)$ are shown in Fig. 4.

Fig. 4a shows that, the estimated $\mathbb{P}(o_t^- | T_t, C_t, a_t^+) = 1$ for any C and T. This is consistent with the fact that when the robot seeks assistance, then the human teleoperates irrespective of the trust state and environment complexity.

Fig. 4b shows the estimated $\mathbb{P}(o_t^+ | T_t, C_t^L, a_t^-)$. If the robot does not ask for assistance in low complexity, then the human relies on the robot with probability 1 and 0.98 when the trust is high and low, respectively. This is consistent with the high success probability for low environment complexity.

Fig. 4c shows the estimated $\mathbb{P}(o_t^+ | T_t, C_t^H, a_t^-)$. If the robot does not ask for assistance in high complexity, then the human relies on the robot with a probability of 0.95 and 0.46 when the trust is high and low, respectively. This is also consistent with the lower success probability for high environment complexity.

[1]The human behavioral experiments were approved under Michigan State University Institutional Review Board Study ID 391.

B. State Transition Matrices.

Fig. 5 shows the estimates of state transition matrices for low environment complexity. Fig. 5a shows the estimate of $\mathbb{P}(T_{t+1} | T_t, E_{t+1}^+, C_t^L, a_t^-)$. Here, if the current trust is high, then it remains high with probability 1; while if the current trust is low, then it remains low with probability 1. Thus, success in an easy task does not repair trust.

Fig. 5b shows the estimate of $\mathbb{P}(T_{t+1} | T_t, E_{t+1}^-, C_t^L, a_t^-)$. Here, if the current trust is high, then it transitions to low with a probability of 0.39; while if the current trust is low, then it remains low with a probability of 0.99. This shows that failure in an easy task has a negative effect on trust.

Fig. 5c shows the estimate of $\mathbb{P}(T_{t+1} | T_t, E_{t+1}^-, C_t^L, a_t^+)$. Here, the trust state does not change after the trial with probability 1. Thus, seeking assistance in low complexity has minimal effect on trust.

Fig. 6 shows the estimates of state transition matrices for high environment complexity. Fig. 6a shows the estimate of $\mathbb{P}(T_{t+1} | T_t, E_{t+1}^+, C_t^H, a_t^-)$. Here, if the current trust is high, then it remains high with probability 1; while if the current trust is low, then it transitions to high with probability 1. Thus, success in complex tasks has more impact on increasing trust compared with easier tasks.

Fig. 6b shows the estimate of $\mathbb{P}(T_{t+1} | T_t, E_{t+1}^-, C_t^H, a_t^-)$. Here, if the current trust is high, then it transitions to low with probability 1; while if the current trust is low, then it remains low with probability 0.7. Thus, high-complexity tasks are high-risk, high-reward: the success of the autonomous operation can increase trust significantly; while a failure decreases trust significantly.

Fig. 6c shows the estimate of $\mathbb{P}(T_{t+1} | T_t, E_{t+1}^+, C_t^H, a_t^+)$. Here, if the current trust is high, then it remains high with a probability of 0.94; while if the current trust is low, then it transitions to high with a probability of 0.44. Thus, seeking assistance in high complexity may increase trust, possibly due to the human perceiving the system as "being careful" by "not risking" to attempt high-complexity tasks. Thus, seeking assistance in high complexity can help build/repair trust and prevent interventions by the human supervisor.

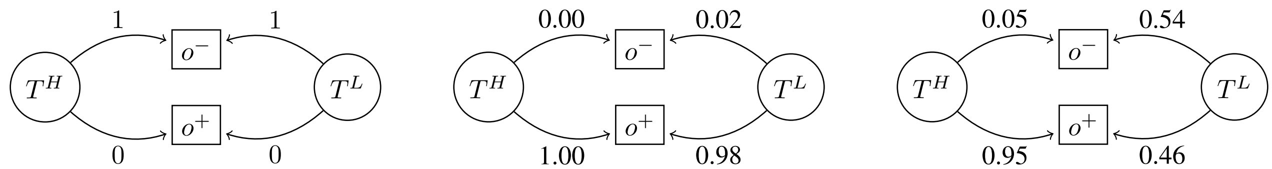
VI. Validation of Trust-aware Assistance-Seeking Policy

In this section, we present the optimal assistance-seeking policy based on the estimated POMDP in Section V and validate its efficacy with data from human subject experiments.

A. Computed Assistance-seeking Policy

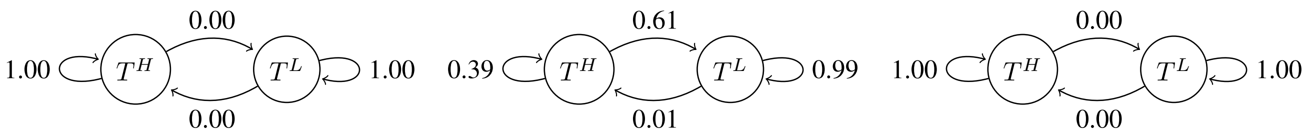
Using the POMDP parameters estimated in Section V, we computed the optimal state-action value function $Q(b_t, E_t, C_t, a)$ for the belief-MDP in Section IV-C using the value iteration algorithm.

The state-action value functions $Q(b_t, E_t, C_t, a)$ for $C_t = C^L$ and $C_t = C^H$ are shown in Fig. 7. Note that experience E_t has no influence on the state-action value functions. This may be attributed to the fact that E_t only affects human behavior through T_t (see Fig. 3).



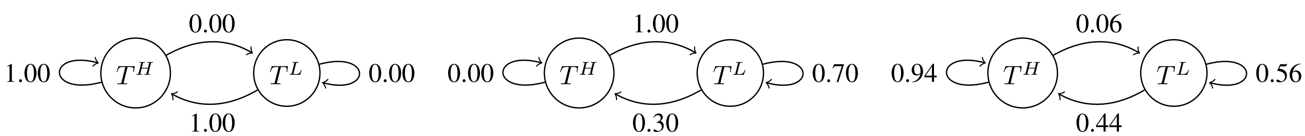
(a) $a = a^+, \forall C$ (b) $a = a^-, C = C^L$ (c) $a = a^-, C = C^H$

Fig. 4: Estimated observation probabilities $\mathbb{P}(o_t|T, C, a)$



(a) $[E^+, C^L, a^-]$ (b) $[E^-, C^L, a^-]$ (c) $[E^-, C^L, a^+]$

Fig. 5: Estimated trust state transition probabilities $\mathbb{P}(T_{t+1}|T_t, E_{t+1}, C_t, a_t)$ for low environment complexity $C_t = C^L$



(a) $[E^+, C^H, a^-]$ (b) $[E^-, C^H, a^-]$ (c) $[E^+, C^H, a^+]$

Fig. 6: Estimated trust state transition probabilities $\mathbb{P}(T_{t+1}|T_t, E_{t+1}, C_t, a_t)$ for high environment complexity $C_t = C^H$

Based on the state-action value functions in Fig. 7, it can be deduced that the optimal action is to never seek assistance in a low-complexity trial. For high-complexity trials, the optimal policy switches from seeking assistance to not seeking assistance as the belief b^T crosses a threshold of ≈ 0.75. Thus, the optimal policy attempts to gather higher reward by autonomously collecting the object in a high complexity trial, if human trust is sufficiently high; otherwise, it works towards increasing human trust by seeking assistance and avoiding any loss in reward due to unnecessary interruption by the human supervisor.

We also computed an optimal trust-agnostic assistance-seeking policy. For a trust-agnostic policy, the only attribute of a trial is the environmental complexity. We used the data from the experiment in Section V to estimate the probability $\mathbb{P}(o_t = o^-|C_t)$ and used it with p_L^{suc} and p_H^{suc} to compute the expected reward for actions a^- and a^+ for both low and high complexities. These calculations suggested that the optimal trust-agnostic policy is to never seek assistance in low-complexity trials and to always seek assistance in high-complexity trials.

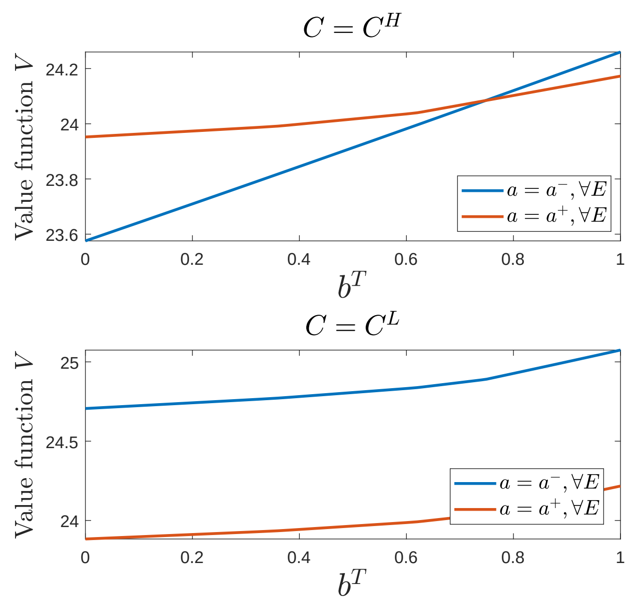


Fig. 7: Estimated action-value functions for the belief-MDP

B. Evaluation of the proposed policy

We performed a second set of human subject experiments to evaluate the proposed optimal trust-aware assistance-seeking policy and compared it with the optimal trust-agnostic policy. Five participants were recruited for these experiments. Every participant performed two blocks of experiments in which the robot adopted trust-aware and trust-agnostic policies, respectively. The order of the blocks was randomly selected. 3 participants started with the trust-aware policy block and 2 participants started with the trust-agnostic policy block. In each block, the participants performed 17 low-complexity and 17 high-complexity trials.

The cumulative reward for all the participants under both policies is shown in Fig. 8. Using the trust-agnostic policy, the median score is 51, while the median score for the trust-aware policy is 57. The trust-aware policy outperformed the trust-agnostic policy for most participants. With the trust-aware policy, the manipulator collected more items autonomously as compared to the trust-agnostic policy.

*Remark 1 (**Timescale of trust repair**):* During the experiments, we noticed that low trust is easily triggered for some participants compared to others. These participants incurred low cumulative rewards in both trust-aware and trust-agnostic settings because they often intervened in low-complexity trials. Our pooled estimates did not specialize to individuals and failed to capture such conservative behavior. This also highlights how certain individuals may require a longer number of trials for their trust to be repaired.

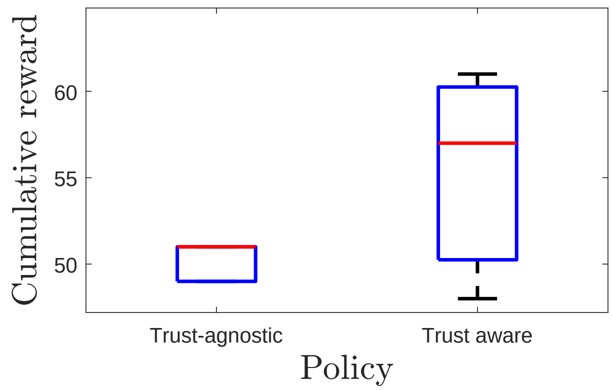


Fig. 8: Cumulative reward statistics for the trust-agnostic and trust-aware policies. The trust-aware policy outperformed the trust-agnostic policy for most participants.

VII. Conclusions and Future Directions

In this paper, using the context of human-supervised object collection tasks, we designed optimal trust-aware policies for a mobile manipulator to seek assistance from a human supervisor. We adopted a POMDP framework and showed that the optimal policy admits a threshold structure: it only seeks assistance when the human trust is below a threshold, and this threshold varies with the complexity of the task. We conducted human subject experiments and showed that the proposed optimal trust-aware policy outperforms the optimal trust-agnostic policy.

An interesting direction of future research is to schedule the low and high-complexity trials such that overall team performance is maximized. Other directions of future research include designing trust-aware policies for non-supervisory human-robot collaboration tasks.

References

[1] S. Musić and S. Hirche, "Control sharing in human-robot team interaction," *Annual Reviews in Control*, vol. 44, pp. 342–354, 2017.

[2] K. Akash, G. McMahon, T. Reid, and N. Jain, "Human trust-based feedback control: Dynamically varying automation transparency to optimize human-machine interactions," *IEEE Control Systems Magazine*, vol. 40, no. 6, pp. 98–116, 2020.

[3] J. R. Peters, V. Srivastava, G. S. Taylor, A. Surana, M. P. Eckstein, and F. Bullo, "Human supervisory control of robotic teams: Integrating cognitive modeling with engineering design," *IEEE Control Systems Magazine*, vol. 35, no. 6, pp. 57–80, 2015.

[4] P. Gupta and V. Srivastava, "Structural properties of optimal fidelity selection policies for human-in-the-loop queues," *ArXiv*, vol. abs/2201.09990, 2022.

[5] J. Lee and K. See, "Trust in technology: Designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, 2004.

[6] K. A. Hoff and M. Bashir, "Trust in automation: Integrating empirical evidence on factors that influence trust," *Human Factors*, vol. 57, no. 3, pp. 407–434, 2015.

[7] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011.

[8] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors*, vol. 39, no. 2, pp. 230–253, 1997.

[9] J. Lee and N. Moray, "Trust, control strategies and allocation of function in human-machine systems," *Ergonomics*, vol. 35, no. 10, pp. 1243–1270, 1992.

[10] J. D. Lee and N. Moray, "Trust, self-confidence, and operators' adaptation to automation," *International Journal of Human-Computer Studies*, vol. 40, no. 1, pp. 153–184, 1994.

[11] D. Manzey, J. Reichenbach, and L. Onnasch, "Human performance consequences of automated decision aids: The impact of degree of automation and system experience," *Journal of Cognitive Engineering and Decision Making*, vol. 6, no. 1, pp. 57–87, 2012.

[12] M. Desai, P. Kaniarasu, M. Medvedev, A. Steinfeld, and H. Yanco, "Impact of robot failures and feedback on real-time trust," in *ACM/IEEE International Conference on Human-Robot Interaction*, 2013, pp. 251–258.

[13] K. Akash, W.-L. Hu, T. Reid, and N. Jain, "Dynamic modeling of trust in human-machine interactions," in *2017 American Control Conference (ACC)*, 2017, pp. 1542–1548.

[14] X. J. Yang, V. V. Unhelkar, K. Li, and J. A. Shah, "Evaluating effects of user experience and system transparency on trust in automation," in *ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 408–416.

[15] X. Wang, Z. Shi, F. Zhang, and Y. Wang, "Dynamic real-time scheduling for human-agent collaboration systems based on mutual trust," *Cyber-Physical Systems*, vol. 1, pp. 76 – 90, 2015.

[16] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa, "Planning with trust for human-robot collaboration," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 307–315.

[17] S. Nikolaidis and J. Shah, "Human-robot teaming using shared mental models," in *Workshop on Human-Agent-Robot Teamwork, ACM/IEEE International Conference on Human Robot Interaction*, 2012.

[18] S. M. Rahman and Y. Wang, "Mutual trust-based subtask allocation for human–robot collaboration in flexible lightweight assembly in manufacturing," *Mechatronics*, vol. 54, pp. 94–109, 2018.

[19] Y. Wang, Z. Shi, C. Wang, and F. Zhang, *Human-Robot Mutual Trust in (Semi)autonomous Underwater Robots*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 115–137.

[20] A. Xu and G. Dudek, "Optimo: Online probabilistic trust inference model for asymmetric human-robot collaborations," in *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2015, pp. 221–228.

[21] Z. Zahedi, M. Verma, S. Sreedharan, and S. Kambhampati, "Trust-aware planning: Modeling trust evolution in longitudinal human-robot interaction," *CoRR*, vol. abs/2105.01220, 2021.

[22] N. Wang, D. V. Pynadath, and S. G. Hill, "Trust calibration within a human-robot team: Comparing automatically generated explanations," in *ACM/IEEE International Conference on Human Robot Interaction*, 2016, p. 109–116.

[23] K. Akash, N. Jain, and T. Misu, "Toward adaptive trust calibration for level 2 driving automation," in *Proceedings of International Conference on Multimodal Interaction*, 2020, pp. 538–547.

[24] Y. Bengio and P. Frasconi, "Input-output HMMs for sequence processing," *IEEE Transactions on Neural Networks*, vol. 7, no. 5, pp. 1231–1249, 1996.

[25] N. Koenig and A. Howard, "Design and use paradigms for Gazebo, an open-source multi-robot simulator," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2004, pp. 2149–2154.

[26] L. Downs, A. Francis, N. Koenig, B. Kinman, R. Hickman, K. Reymann, T. B. McHugh, and V. Vanhoucke, "Google scanned objects: A high-quality dataset of 3D scanned household items," in *International Conference on Robotics and Automation*, 2022, pp. 2553–2560.

[27] G. Ramkumar and J. Lentin, *ROS Robotics Projects : Build and Control Robots Powered by the Robot Operating System, Machine Learning, and Virtual Reality*. Packt Publishing, 2019, vol. Second edition.