Trust-Aware Decision Making for Human-Robot Collaboration: Model Learning and Planning

MIN CHEN*, National University of Singapore
STEFANOS NIKOLAIDIS, University of Southern California
HAROLD SOH, National University of Singapore
DAVID HSU, National University of Singapore
SIDDHARTHA SRINIVASA, University of Washington

Trust in autonomy is essential for effective human-robot collaboration and user adoption of autonomous systems such as robot assistants. This paper introduces a computational model which integrates trust into robot decision-making. Specifically, we learn from data a partially observable Markov decision process (POMDP) with human trust as a latent variable. The trust-POMDP model provides a principled approach for the robot to (i) infer the trust of a human teammate through interaction, (ii) reason about the effect of its own actions on human trust, and (iii) choose actions that maximize team performance over the long term. We validated the model through human subject experiments on a table-clearing task in simulation (201 participants) and with a real robot (20 participants). In our studies, the robot builds human trust by manipulating low-risk objects first. Interestingly, the robot sometimes fails intentionally in order to modulate human trust and achieve the best team performance. These results show that the trust-POMDP calibrates trust to improve human-robot team performance over the long term. Further, they highlight that maximizing trust alone does not always lead to the best performance.

CCS Concepts: • **Human-centered computing** → **Collaborative interaction**; • **Computing methodologies** → *Planning under uncertainty*;

Additional Key Words and Phrases: Trust models, Human-robot collaboration, Partially observable Markov decision process (POMDP)

1 INTRODUCTION

Trust is essential for seamless human-robot collaboration and user adoption of autonomous systems, such as robot assistants. Over-trusting robot autonomy may lead to misuse of such systems, where people rely excessively on automation, failing to intervene in the case of critical failures [26]. On the other hand, lack of trust leads to disuse of autonomous systems: users ignore the systems' capabilities, with negative effects on overall performance.

We witnessed an example of users' distrust in the system in one of our studies, where a human participant and a robot collaborated to clear a table (Figure 1). Although the robot was fully capable

*Chen and Nikolaidis contributed equally to the work. The user studies in this work were approved by the NUS-IRB: A-16-356.

Authors' addresses: Min Chen, National University of Singapore, chenmin@comp.nus.edu.sg; Stefanos Nikolaidis, University of Southern California, snikolai@alumni.cmu.edu; Harold Soh, National University of Singapore, harold@comp.nus.edu.sg; David Hsu, National University of Singapore, dyhsu@comp.nus.edu.sg; Siddhartha Srinivasa, University of Washington, siddh@cs.uw.edu.

2573-9522/2019/4-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Fig. 1. A robot and a human collaborate to clear a table. The human, with low initial trust in the robot, intervenes to stop the robot from moving the wine glass.

of handling all objects on the table, inexperienced participants did not trust that the robot was able to succeed and stopped the robot from moving the wine glass, since they were afraid that the glass may fall and break. It was clear that their trust was poorly calibrated with respect to the robot's true capabilities. This, in turn, had a significant effect on the interaction.

This study revealed that, in order to achieve fluent human-robot collaboration, the robot should *monitor* human trust and *influence* it so that it matches the system's capabilities. In our study, for instance, the robot should build human trust first by acting in a trustworthy manner, before going for the wine glass.

We propose a trust-based computational model of robot decision making: Since trust is not fully observable, we model it as a latent variable in a partially observable Markov decision process (POMDP) [22]. Our trust-POMDP model contains two key components: (i) a trust dynamics model, which captures the evolution of human trust in the robot, and (ii) a human decision model, which connects trust with human actions. Our POMDP formulation can accommodate a variety of trust dynamics and human decision models. Here, we adopt a data-driven approach and learn these models from data.

Although prior work has studied human trust elicitation and modeling [14, 25, 44, 46], we close the loop between trust modeling and robot decision-making. The trust-POMDP enables the robot to systematically infer and influence the human collaborator's trust, and leverage trust for improved human-robot collaboration and long-term task performance.

Consider again the table clearing example (Figure 2). The trust-POMDP strategy first removes the three plastic water bottles to build up trust and only attempts to remove the wine glass afterwards. In contrast, a baseline myopic strategy maximizes short-term task performance and does not account for human trust in choosing the robot actions. It first removes the wine glass, which offers the highest reward, resulting in unnecessary interventions by human collaborators with low initial trust.

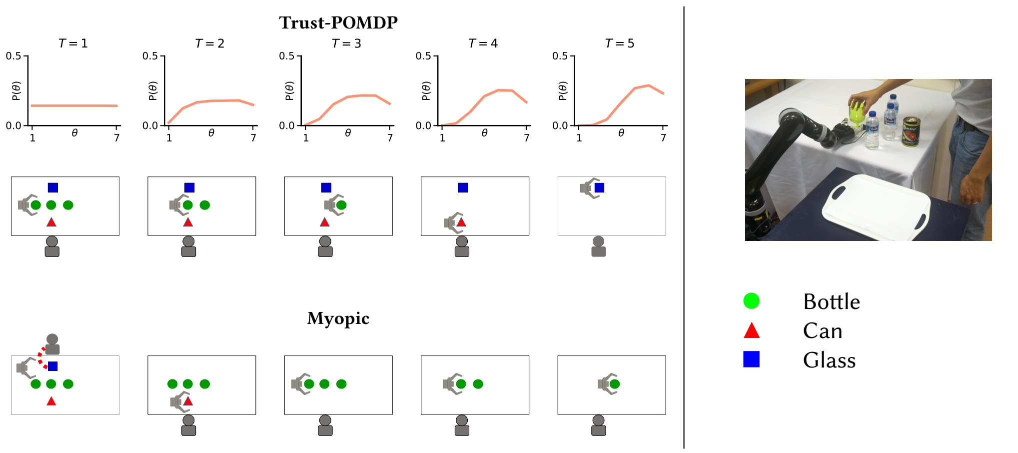


Fig. 2. Sample runs of the trust-POMDP strategy and the myopic strategy on a collaborative table-clearing task. The top row shows the probabilistic estimates of human trust over time on a 7-point Likert scale. The trust-POMDP strategy starts by moving the plastic bottles to build trust ($T = 1, 2, 3$) and moves the wine glass only when the estimated trust is high enough ($T = 5$). The myopic strategy does not account for trust and starts with the wine glass, causing the human with low initial trust to intervene ($T = 1$).

We validated the trust-POMDP model through human subject experiments on the collaborative table-clearing task, both online in simulation (201 participants) and with a real robot (20 participants). Compared with the myopic strategy, the trust-POMDP strategy significantly reduced participants' intervention rate, indicating improved team collaboration and task performance.

In these experiments the robot always succeeded. Robots, however, fail frequently. What if the robot is likely to *fail* when picking up the wine glass? The robot should then assess human trust in the beginning of the task; if trust is too high, the robot should effectively *communicate* this to the human, in order to calibrate human trust to the appropriate level. While human teammates are able to use natural language to communicate expectations [28], our assistive robotic arm does not have verbal communication capabilities. The trust-POMDP strategy in this case enables the robot to modulate human trust by *intentionally failing* when picking up the bottles, before attempting to grasp the wine glass. This prompts the human to intervene when the robot attempts to pick up the wine glass, preventing failure.

This paper builds upon our previous work [7] by introducing robot failures into the computational framework. In particular, (i) we augment the dynamics model with robot failures, add a new session of data collection to learn the model and discuss the effect of failures on different levels of trust; (ii) we simulate and visualize robot policies with the learned model; (iii) we provide an analysis of the results in the case of an adaptive policy that enables the robot to assess participants' initial trust and intentionally fail.

Integrating trust modeling and robot decision making enables robot behaviors that leverage human trust and actively *modulate* it for seamless human-robot collaboration. Under the trust-POMDP model, the robot deliberately chooses to fail in order to reduce the trust of an overly trusting user and achieve better task performance over the long term. Further, embedding trust in a reward-based POMDP framework makes our robot task-driven: when the human collaboration is unnecessary, the robot may set aside trust building and act to maximize the team task performance directly. All these diverse behaviors emerge automatically from the trust-POMDP model, without explicit manual robot programming.

2 RELATED WORK

Trust has been studied extensively in the social science research literature [16, 23], with Mayer et al., suggesting that three general levels summarize the bases of trust: ability, integrity, and benevolence [29]. Trust in automation differs from trust between people in that (i) humans are more collaborative and trusting of human partners even when observed agent actions are identical [45], and (ii) automation lacks intentionality [26]. Additionally, in a human-robot collaboration task, human and robot share the same objective metric of task performance. Therefore, similar to previous work [9, 35, 36, 44, 47], we assume that human teammates will not expect the robot to deceive them on purpose, and their trust will depend mainly on the *perceived robot ability* to complete the task successfully.

Binary measures of trust [17], as well as continuous measures [9, 25, 47], and ordinal scales [18, 20, 30] have been proposed. For real-time measurement, Desai [9] proposed the Area Under Trust Curve (AUTC) measure, which was recently used to account for one's entire interactive experience with the robot [48].

Researchers have also studied the temporal dynamics of trust conditioned on the task performance: Lee and Moray [25] proposed an autoregressive moving average vector form of time series analysis; Floyd et al. [14] used case-based reasoning; Xu and Dudek [46] proposed an online probabilistic trust inference model to estimate a robot's trustworthiness; Wang et al. [44] showed that adding transparency in the robot model by generating explanations improved trust and performance in human teams; Desai et al. [10, 11] showed that robot failures had a negative impact on human trust, and early robot failures led to dramatically lower trust than later robot failures. While previous works have focused on either quantifying trust or studying the dynamics of trust in human-robot interaction, our work enables the robot to leverage upon a model of human trust and choose actions to maximize task performance.

In human-robot collaborative tasks, the robot often needs to reason over the human's hidden mental state in its decision-making. The POMDP provides a principled general framework for such reasoning. It has enabled robotic teammates to coordinate through communication [5] and software agents to infer the intention of human players in game AI applications [27]. The model has been successfully applied to real-world tasks, such as autonomous driving where the robot car interacts with pedestrians and human drivers [2, 3, 15]. When the state and action space of the POMDP model become continuous, one can use hindsight optimization [19], or value of information heuristics [39], which generate approximate solutions but are computationally more efficient.

Nikolaidis et al. [33] proposed to infer the human type or preference online using models learned from joint-action demonstrations. This formalism recently extended from one-way adaptation (from robot to human) to human-robot *mutual* adaptation [31, 32], where the human may choose to change their preference and follow a policy demonstrated by the robot in the recent history. In this work, we provide a general way to link the whole interaction history with the human policy, by incorporating human trust dynamics into the planning framework.

3 TRUST-POMDP

3.1 Human-robot team model

We formalize the human-robot team as a Markov Decision Process (MDP), with world state $x \in X$, robot action $a^{\mathrm{R}} \in A^{\mathrm{R}}$, and human action $a^{\mathrm{H}} \in A^{\mathrm{H}}$. The system evolves according to a probabilistic state transition function $p(x'|x, a^{\mathrm{R}}, a^{\mathrm{H}})$ which specifies the probability of transitioning from state x to state x' when actions a^{R} and a^{H} are applied in state x. After transitioning, the team receives a real-valued reward $r(x, a^{\mathrm{R}}, a^{\mathrm{H}})$, which is constructed to elicit the desirable team behaviors.

We denote by $h_t = \{x_0, a_0^{\mathrm{R}}, a_0^{\mathrm{H}}, x_1, r_1, \ldots, x_{t-1}, a_{t-1}^{\mathrm{R}}, a_{t-1}^{\mathrm{H}}, x_t, r_t\} \in H_t$ as the history of interaction between robot and human until time step t. In this paper, we assume that the human observes the robot's current action and then decides their own action. In the most general setting, the human uses the entire interaction history h_t to decide the action. Thus, we can write the human's (possibly stochastic) policy as $\pi^{\mathrm{H}}(a_t^{\mathrm{H}} | x_t, a_t^{\mathrm{R}}, h_t)$ which outputs the probability of each human action a_t^{H}.

Given a robot policy π^{R}, the *value*, *i.e.*, the expected total discounted reward of starting at a state x_0 and following the robot and human policies is

$$v(x_0 | \pi^{\mathrm{R}}, \pi^{\mathrm{H}}) = \mathop{\mathbb{E}}_{a_t^{\mathrm{R}} \sim \pi^{\mathrm{R}}, a_t^{\mathrm{H}} \sim \pi^{\mathrm{H}}} \sum_{t=0}^{\infty} \gamma^t r(x_t, a_t^{\mathrm{R}}, a_t^{\mathrm{H}}), \tag{1}$$

where x_t is the state at time step t, a_t^{R} and a_t^{H} are the robot action and human action at time step t. γ is the discount factor that favors immediate rewards over future ones.

The robot's optimal policy π_*^{R} can be computed as

$$\pi_*^{\mathrm{R}} = \mathop{\arg\max}_{\pi^{\mathrm{R}}} v(x_0 | \pi^{\mathrm{R}}, \pi^{\mathrm{H}}). \tag{2}$$

In our case, however, the robot does not know the human policy in advance. It computes the optimal policy under expectation over the human policy:

$$\pi_*^{\mathrm{R}} = \mathop{\arg\max}_{\pi^{\mathrm{R}}} \mathop{\mathbb{E}}_{\pi^{\mathrm{H}}} v(x_0 | \pi^{\mathrm{R}}, \pi^{\mathrm{H}}). \tag{3}$$

Key to solving Eq. 3 is for the robot to model the human policy, which potentially depends on the entire history h_t. The history h_t may grow arbitrarily long and make the optimization extremely difficult.

3.2 Trust-dependent human behaviors

Our insight is that in a number of human-robot collaboration scenarios, *trust is a compact approximation of the interaction history* h_t. This allows us to condition human behavior on the inferred trust level and in turn find the optimal policy that maximizes team performance.

Following previous work on trust modeling [46], we assume that trust can be represented as a single scalar random variable θ. Thus, the human policy is rewritten as

$$\pi^{\mathrm{H}}(a_t^{\mathrm{H}} | x_t, a_t^{\mathrm{R}}, \theta_t) = \pi^{\mathrm{H}}(a_t^{\mathrm{H}} | x_t, a_t^{\mathrm{R}}, h_t). \tag{4}$$

3.3 Trust dynamics

Human trust changes over time. We adopt a common assumption on the trust dynamics: trust evolves based on the robot's performance e_t [25, 46]. Performance can depend not just on the current and transitioned world state but also the human and robot's actions

$$e_{t+1} = \texttt{performance}(x_{t+1}, x_t, a_t^{\mathrm{R}}, a_t^{\mathrm{H}}). \tag{5}$$

For example, `performance` may indicate success or failure of the robot to accomplish a task. This allows us to write our trust dynamics equation as

$$\theta_{t+1} \sim p(\theta_{t+1} | \theta_t, e_{t+1}). \tag{6}$$

We detail in Section 4 how trust dynamics is learned via interaction.

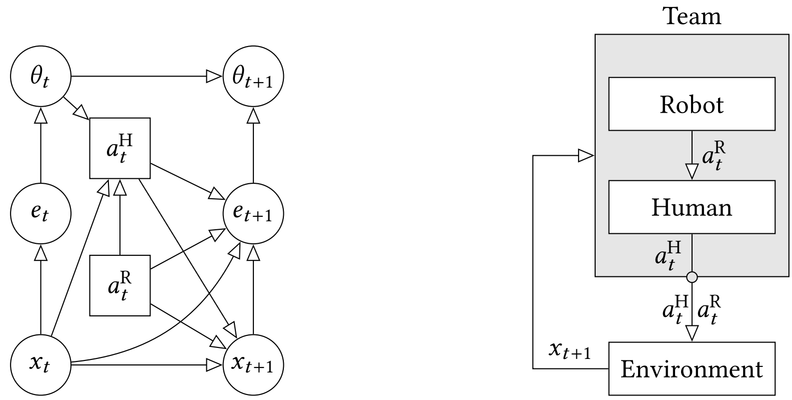


Fig. 3. The trust-POMDP graphical model (left) and the team interaction flowchart (right). The robot action a_t^R depends on the world state x_t and its belief over trust θ_t.

3.4 Maximizing team performance

Trust cannot be directly observed by the robot and therefore must be inferred from the human's actions. In addition, armed with a model, the robot may actively modulate the human's trust for the team's long-term reward.

We achieve this behavior by modeling the interaction as a partially observable Markov decision process (POMDP), which provides a principled general framework for sequential decision making under uncertainty. A graphical model of the Trust-POMDP and a flowchart of the interaction are shown in Figure 3.

To build trust-POMDP, we create an augmented state space with the augmented state $s = (x, \theta)$ composed of the fully-observed world state x and the *partially-observed* human trust θ. We maintain a belief b over the human's trust. The trust dynamics and human behavioral policy are embedded in the transition dynamics of trust-POMDP. We describe in Section 4 how we learn the trust dynamics and the human behavioral policy.

The robot now has two distinct objectives through its actions:

- *Exploitation.* Maximize the team's reward
- *Exploration.* Reveal and change the human's trust so that future actions are rewarded better.

The solution to a Trust-POMDP is a policy that maps belief states to robot actions, *i.e.*, $a^{\mathrm{R}} = \pi^{\mathrm{R}}(b_t, x_t)$. To compute the optimal policy, we use the SARSOP algorithm [24], which is computationally efficient and has been previously used in various robotic tasks [3].

4 LEARNING TRUST DYNAMICS AND HUMAN BEHAVIORAL POLICIES

Nested within the trust-POMDP is a model of human trust dynamics $p(\theta_{t+1}|\theta_t, e_{t+1})$, and behavioral policy $\pi^{\mathrm{H}}(a_t^{\mathrm{H}}|x_t, a_t^{\mathrm{R}}, \theta_t)$. We adopted a data-driven approach and built the two models for the table clearing task from data collected in Amazon's Mechanical Turk (AMT) study. Suitable probabilistic models derived via alternative approaches can be substituted for these learned models (*e.g.*, for other tasks and domains).

4.1 Data Collection

Table clearing task. A human and a robot collaborate to clear objects off a table. The objects include three water bottles, one fish can, and one wine glass. At each time step, the robot picks

Table 1. The reward function R for the table-clearing task.

	Bottle	Fish Can	Wine Glass
SP-success	1	2	3
SP-fail	0	-4	-9
IT	0	0	0

up one of the remaining objects. Once the robot starts moving towards the intended object, the human can choose between two actions: {intervene and pick up the object that the robot is moving towards, stay put and let the robot pick the object by itself}. This process is repeated until all the objects are cleared from the table.

Each object is associated with a different reward, based on whether the robot successfully clears it from the table (which we call SP-success), the robot fails in clearing it (SP-fail), or the human intervenes and puts it on the tray (IT). Table 1 shows the rewards for each object and outcome. We assume that a robot success is always better than a human intervention, since it reduces human effort. Additionally, there is no penalty if the robot fails by dropping one of the sealed water bottles, since the human can pick it up. On the other hand, dropping the fish can result in some penalty, since its contents will be spilled on the floor. Breaking the glass results in the highest penalty. We see that staying put when the robot attempts to pick up the bottle has the lowest risk, since there is no penalty if the robot fails. On the other hand, staying put in the case of the glass object has the largest risk-return trade off. We expect the human to let the robot pick up the bottle even if their trust is low, since there is no penalty if the robot fails. On the other hand, if the human does not trust the robot, we expect them to likely intervene on glass or can, rather than risking a high penalty in case of robot failure.

Participants. For the data collection, we recruited in total 231 participants through AMT [1]. The participants were all from United States, aged 18-65 and with approval rate higher than 95%. Each participant was compensated $1 for completing the study. To ensure the quality of the recorded data, we asked all participants an attention check question that tested their attention to the task. We removed 9 data points either because the participants failed on the attention check question or the their data were incomplete. This left us 222 valid data points for model learning.

Procedure. Each participant was asked to perform an online table clearing task together with a robot. Before the task started, the participant was informed of the reward function in Table 1. We first collected the participant's initial trust in the robot. We used Muir's questionnaire [30], with a seven-point Likert scale as a human trust metric, *i.e.*, trust ranges from 1 to 7. Muir's questionnaire is a well established measure of trust in the literature [9]. It integrates Barber's and Rempel *et al.*'s model to include the dimensions of persistence, technical competence, reliability and predictability of the robot [4, 37], which are important when measuring performance-centric trust (*i.e.*, the type of trust studied in this work). The set of questions we used is listed in Table 2.

At each time step, the participant watched a video of the robot attempting to pick up an object, and were asked to choose to intervene or stay put. They then watched a video of either the robot picking up the object, or them intervening based on their action selection. Then, they reported their updated trust in the robot.

We were interested in learning the trust dynamics and the human behavioral policies for any state and robot action. However, the number of open-loop [2] robot policies is $O(K!)$, where K is the

[1]We conducted two sessions of data collection, one where the robot always succeeded and one when the robot failed with high probability. Our previous work [7] presented the results of the first session only.

[2]When collecting data from AMT, the robot follows an open-loop policy, *i.e.*, it does not adapt to the human behavior.

Table 2. Muir's questionnaire.

1. To what extent can the robot's behavior be predicted from moment to moment?
2. To what extent can you count on the robot to do its job?
3. What degree of faith do you have that the robot will be able to cope with similar situations in the future?
4. Overall how much do you trust the robot?

number of objects on the table. In order to focus the learning on a few interesting robot policies (i.e. picking up the glass in the beginning vs in the end), while still covering a large space of policies, we split the data collection process, so that in one half of the trials the robot randomly chose a policy out of a set of pre-specified policies, while in the other half the robot followed a random policy.

Data Format. The data we collected from each participant has the following format:

$$d_i = \{\theta_0^{\mathrm{M}}, a_0^{\mathrm{R}}, a_0^{\mathrm{H}}, e_1, \theta_1^{\mathrm{M}}, \ldots, a_{K-1}^{\mathrm{R}}, a_{K-1}^{\mathrm{H}}, e_K, \theta_K^{\mathrm{M}}\}$$

where K is the number of objects on the table. θ_t^{M} is the estimated human trust at time t by averaging the participants' responses to the Muir's questionnaire to a single rating between 1 and 7. a_t^{R} is the action taken by the robot at time step t. a_t^{H} is the action taken by the human at time step t. e_{t+1} is the performance of the robot that indicates whether the robot succeeded at picking up the object, the robot failed, or the human intervened.

4.2 Trust dynamics model

We model human trust evolution as a linear Gaussian system, which is similar to [46]. Our trust dynamics model relates the human trust causally to the robot task performance e_{t+1}.

$$\begin{aligned} &P(\theta_{t+1}|\theta_t, e_{t+1}) = \mathcal{N}(\alpha_{e_{t+1}}\theta_t + \beta_{e_{t+1}}, \sigma^2_{e_{t+1}}) \\ &\theta_t^M \sim \mathcal{N}(\theta_t, \sigma^2)\ ,\ \theta_{t+1}^M \sim \mathcal{N}(\theta_{t+1}, \sigma^2) \end{aligned} \tag{7}$$

where $\mathcal{N}(\mu, \sigma)$ denotes a Gaussian distribution with mean μ and standard deviation σ. $\alpha_{e_{t+1}}$ and $\beta_{e_{t+1}}$ are linear coefficients for the trust dynamics, given the robot task performance e_{t+1}. In the table clearing task, e_{t+1} indicates whether the robot succeeded at picking up an object, the robot failed, or the human intervened, *e.g.*, e_{t+1} can represent that the robot succeeded at picking a water bottle, or that the human intervened at the wine glass. θ_t^M and θ_{t+1}^M are the observed human trust (Muir's questionnaire) at time step t and time step $t+1$.

The unknown parameters in the trust dynamics model include $\alpha_{e_{t+1}}$, $\beta_{e_{t+1}}$, $\sigma_{e_{t+1}}$ and σ. We performed full Bayesian inference on the model through Hamiltonian Monte Carlo sampling using the Stan probabilistic programming platform [6].

Trust in our model is a discrete scalar that ranges from 1 to 7. We computed the trust transition dynamics by sampling from the learned linear Gaussian model. For example, suppose the current trust is $\theta_t = 1$, the next trust θ_{t+1} follows a Gaussian distribution $\mathcal{N}(\alpha_{e_{t+1}}\theta_t + \beta_{e_{t+1}}, \sigma^2_{e_{t+1}})$, where we can compute the probability of $\theta_{t+1} = 1, \theta_{t+1} = 2, \ldots, \theta_{t+1} = 7$, and we normalize the probability afterwards. Figure 4 shows the trust transition matrices for all trust levels and all possible robot performance in the table clearing task. As we can see, human trust in the robot gradually increased with observations of successful robot actions (as indicated by transitions to higher trust levels when the participants stayed put and robot succeeded), while it decreased with observations of robot failures. Trust tended to remain constant or decrease slightly when interventions occurred. It

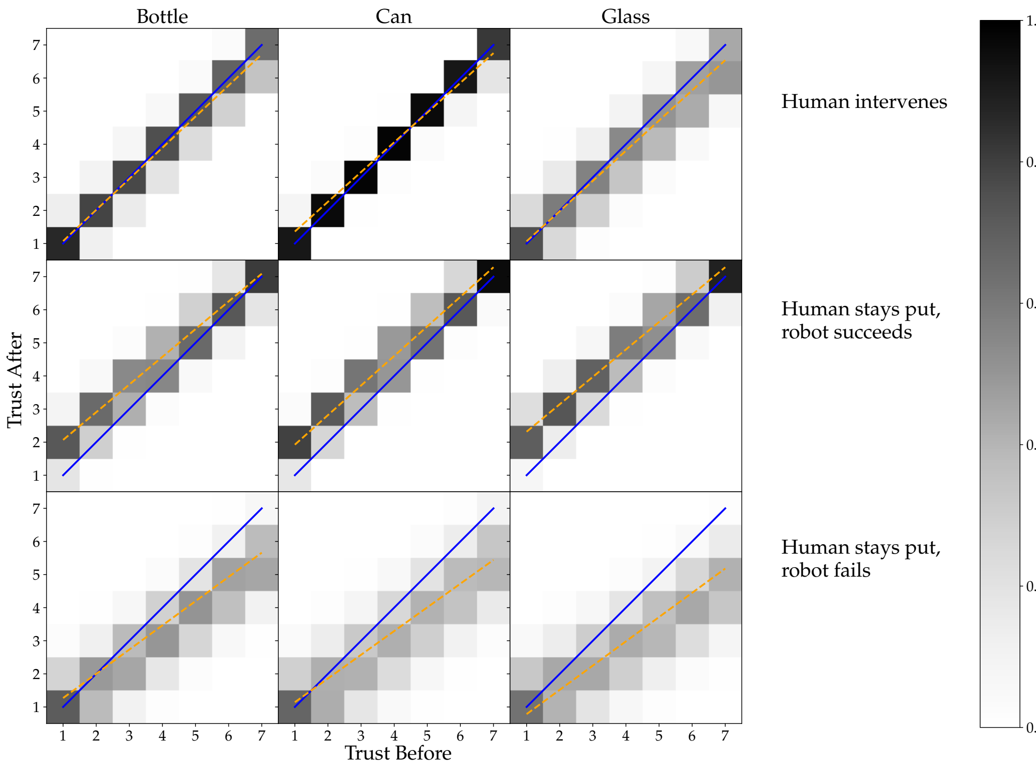


Fig. 4. Trust transition matrices, which represent the change of trust given the robot performance, shown by the linearly regressed line (yellow) contrasted with the X-Y line (blue). In general, trust stays constant or decreases slightly when the human intervenes (top row). It increases when the human stays put and the robot succeeds (middle row), while it decreases when the robot fails (bottom row).

also appears that the higher the trust, the greater the loss upon failure, and vice versa upon success. These results matched our expectations that successful robot performance positively influenced trust, while robot failures negatively affected trust.

4.3 Human behavioral policies

Our key intuition in the human model is that human's behavior depends on the trust in the robot. To support our intuition, we consider two types of human behavioral models. The first model is a trust-free human behavioral model that ignores human trust, while the second is a trust-based human behavioral model that explicitly models human trust. In both human models, we assume humans follow the *softmax rule* [3] when they make decisions in an uncertain environment [8]. More explicitly,

- Trust-free human behavioral model: At each time step, the human selects an action probabilistically based on the actions' relative expected values. The expected value of an action depends on the human's belief on the robot to succeed and the risk of letting robot to do the

[3]According to the *softmax rule* , the human's decision of which action to take is determined probabilistically on the actions' relative expected values.

task. In the trust-free human model, the human's belief on the robot success on a particular task does not change over time.

- Trust-based human behavioral model: Similar to the model above, the human follows the *softmax rule* at each time step. However, the trust-based human model assumes that human's belief on the robot success changes over time, and it depends on human's trust in the robot.

Before we introduce the models, we start with some notations. Let j denote the object that the robot tries to pick at time step t. Let r_j^{S} be the reward if the human stays put and the robot succeeds, and r_j^{F} be the reward if the human stays put and the robot fails. Let θ_t be the human trust in the robot at time step t. $\mathcal{S}(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function, which is equivalent to the softmax function in the case of binary human actions. $\mathcal{B}(p)$ is the Bernoulli distribution that takes action stay put with probability p.

The trust-free human behavioral model is as follows,

$$\begin{aligned} P_t &= \mathcal{S}(b_j r_j^{\mathrm{S}} + (1 - b_j) r_j^{\mathrm{F}}) \\ a_t^{\mathrm{H}} &\sim \mathcal{B}(P_t) \end{aligned} \tag{8}$$

where, b_j is the human's belief on the robot successfully picking up object j, and it remains constant. $0 < P_t < 1$ is the probability that human stays put at time step t. a_t^{H} is the action human took at time step t.

Next, we introduce the trust-based human behavioral model:

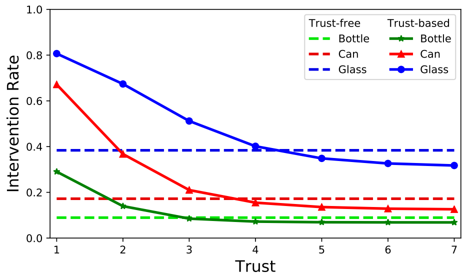
$$\begin{aligned} b_j^t &= \mathcal{S}(\gamma_j \theta_t + \eta_j) \\ P_t &= \mathcal{S}(b_j^t r_j^{\mathrm{S}} + (1 - b_j^t) r_j^{\mathrm{F}}) \\ \theta_t^M &\sim \mathcal{N}(\theta_t, \sigma^2)\ ,\ a_t^{\mathrm{H}} \sim \mathcal{B}(P_t) \end{aligned} \tag{9}$$

where b_j^t is the human's belief on robot success on object j at time step t, and it depends on the human's trust in the robot. γ_j and η_j are the linear coefficients for object j. $0 < P_t < 1$ is the probability that the human stays put at time step t. θ_t^M is the observed human trust from Muir's questionnaire at time step t, and we assume it follow a Gaussian distribution with mean θ_t and standard deviation σ. a_t^{H} is the action human took at time step t.

The unknown parameters here include b_j in the trust-free human model, and γ_j, η_j, σ in the trust-based human model. We performed Bayesian inference on the two models above using Hamiltonian Monte Carlo sampling [6]. The trust-based human model (log-likelihood = −153.37) fit the collected AMT data better than the trust-free human model (log-likelihood = −156.40). The log-likelihood values are relatively low in both two models due to the large variance among different users. Nevertheless, this result supports our notion that the prediction on human behavior is improved when we explicitly model human trust.

Figure 5 shows the model prediction on the mean probability of human interventions with respect to trust (left column) and observed human behaviors during the study (right column). Overall, the prediction of the trust-based model is closer to the observed human behaviors during the study. For both the trust-based model and trust-free model, the human tends to intervene more on objects with higher risk, *i.e.*, the human intervention rate on glass is higher than can, which is again higher than bottle. The trust-free human behavioral model ignores human trust, thus the human intervention rate does not change. On the other hand, the trust-based human behavioral model has a general falling trend, which indicates that participants are less likely to intervene when their trust in the robot is high. This is observed particularly for the highest-risk object (glass), where the object intervention rate drops significantly when human trust score is maximum.

To summarize, the results of Sec. 4.2 and Section 4.3 indicate that



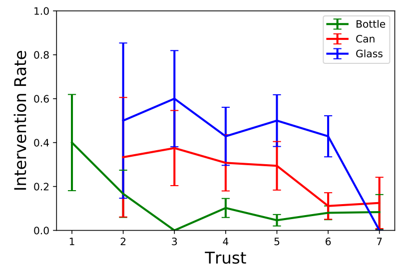


Fig. 5. The model prediction on the mean of human intervention rate with respect to trust (left, column), and the observed human intervention rate during the study (right column). Under the trust-free human behavioral model, the human intervention rate stays constant. Under the trust-based human behavioral model, the intervention rate decreases with increasing trust, which is similar to the observed human intervention rate, and the rate of decrease depends on the object (it is more sensitive to the risker objects).

- Human trust is affected by robot performance: human trust can be built up by successfully picking up objects (Figure 4). In addition, it is a good strategy for the robot to start with low risk objects (bottle), since the human is less likely to intervene even if the trust in the robot is low (Figure 5).
- Human trust affects human behaviors: the intervention rate on the high risk objects could be reduced by building up human trust (Figure 5).

5 EXPERIMENTS

We conducted two human subjects experiments, one on AMT with human participants interacting with recorded videos and one in our lab with human participants interacting with a real robot. The purpose of our study was to test whether the trust-POMDP robot policy would result in better **team performance** than a policy that did not account for human trust. To simplify the analysis of the different behaviors in these experiments, we had the robot always succeed when attempting to pick up the objects.

We had two experimental conditions, which we refer to as “trust-POMDP” and “myopic”.

- In the trust-POMDP condition, the robot uses human trust as a means to optimize the long term team performance. It follows the policy computed from the trust-POMDP described in Section 3.4, where the robot’s perceived human policy is modeled via the trust-based human behavioral model described in Section 4.3.
- In the myopic condition, the robot ignores human trust. It follows a myopic policy by optimizing Eq. 3, where the robot’s perceived human policy is modeled via the trust-free human behavioral model described in Section 4.3.

5.1 Online AMT experiment

Hypothesis 1. *In the online experiment, the performance of teams in the trust-POMDP condition will be better than of the teams in the myopic condition.*

We evaluated team performance by the accumulated reward over the task. We expected the trust-POMDP robot to reason over the probability of human interventions, and act so as to minimize the intervention rate for the highest reward objects. The robot would do so by actively building up human trust before it goes for high risk objects. On the contrary, the myopic robot policy was agnostic to how the human policy may change from the robot and human actions.

Procedure. The procedure is similar to the one for data collection (Sec. 4.1), with the difference that, rather than executing random sequences, the robot executes the policy associated with each condition. While we kept the Muir's questionnaire in the experiment as a groundtruth measure of trust, the robot *did not use the score*, but estimated trust solely from the trust dynamics model as described in Sec. 4.2.
Model parameters. In the formulation of Section 3.4, the observable state variable x represents the state of each object (on the table or removed). We assume a discrete set of values of trust θ : $\{1, 2, 3, 4, 5, 6, 7\}$. The transition function incorporates the learned trust dynamics and human behavioral policies, as described in Sec. 4. The reward function R is given by Table 1. We used a discount factor of $\gamma = 0.99$, which favors immediate rewards over future rewards.
Subject Allocation We chose a between-subjects design in order to not bias the users with policies from previous conditions. We recruited 208 participants through Amazon Mechanical Turk, aged $18-65$ and with approval rate higher than 95%. Each participant was compensated \$1 for completing the study. We removed 7 wrong (participants failed on the attention check question) or incomplete data points. In the end, we had 101 data points for the trust-POMDP condition, and 100 data points for the myopic condition.

5.2 Real-robot experiment

In the real-robot experiment we followed the same robot policies, model parameters and procedures as the online AMT experiment, with that the participants interacted with a real-robot in person.
Hypothesis 2. *In the real-robot experiment, the performance of teams in the trust-POMDP condition will be better than of the teams in the myopic condition.*
Subject Allocation. We recruited 20 participants from our university, aged 21-65. Each participant was compensated \$10 for completing the study. All data points were kept for analysis, *i.e.*, 10 data points for the trust-POMDP condition and 10 data points for the myopic condition.

5.3 Team performance

We used the accumulated rewards as a measure of team performance. We performed a one-way ANOVA test of the accumulated rewards in the trust-based condition and the myopic condition, In addition, we computed the effect size (Cohen's d) for comparison between the means of accumulated rewards in the two conditions.

In the online AMT experiment, the accumulated rewards of trust-based condition was significantly larger than the myopic condition ($F(1, 199) = 7.81, p = 0.006$). The effect size $d = 0.394$, indicating a medium difference between the two means. This result supports Hypothesis 1.

Similarly, in the real-robot experiment, the accumulated rewards of the trust-based condition was significantly larger than the myopic condition ($F(1, 18) = 11.22, p = 0.004$). The effect size $d = 1.498$, indicating a huge difference between the two means. This result supports Hypothesis 2. The effect size showed that the difference in performance was more significant in the real-robot experiment, compared with online AMT experiment. This was mainly because the data from AMT was noisy, and some online participants did not pay full attention to the experiment.

Overall, for both the online AMT experiment and the real-robot experiment, the trust-POMDP robot outperformed the myopic robot. The difference in performance occurred because participants' intervention rate in the trust-POMDP condition was significantly lower than myopic condition (Figure 6 - left column). In the online AMT experiment, the intervention rate in the trust-POMDP condition was 54% and 31% lower in the can and glass object. In the real-robot experiment, the intervention rate in the trust-POMDP condition dropped to zero (100% lower) in the can object and 71% lower in the glass object.

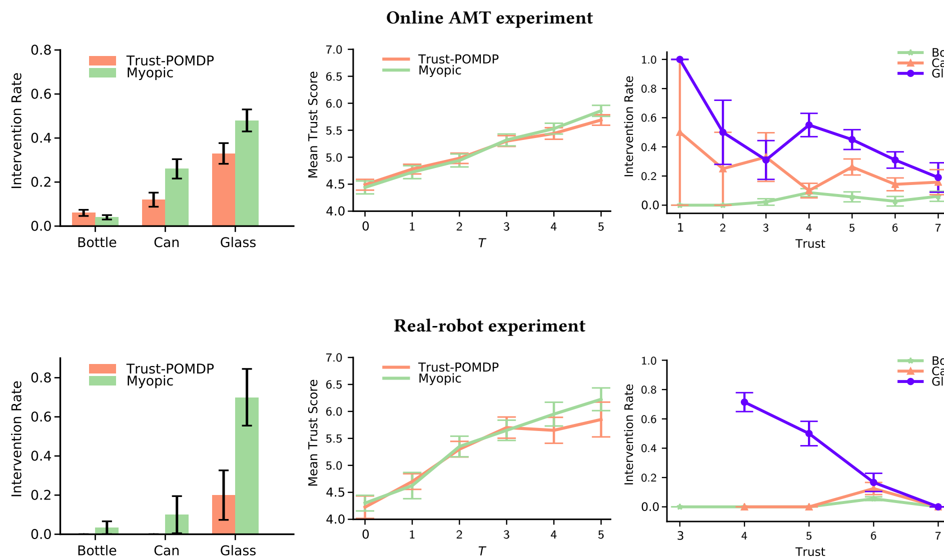


Fig. 6. Comparison of the Trust-POMDP and the myopic policies in the AMT experiment and the real-robot experiment.

In the myopic condition, the robot picked the objects in the order of highest to lowest reward (Glass, Can, Bottle, Bottle, Bottle). In contrast, the trust-based human behavior model influenced the trust-POMDP robot policy by capturing the fact that interventions on high-risk objects were more likely if trust in the robot was insufficient. Therefore, the trust-POMDP robot reasoned that it was better to start with the low risk objects (bottles), build human trust (Figure 6 - center column) and go for high risk object (glass) last. In this way, the trust-POMDP robot minimized the human intervention ratio on the glass and can object, which significantly improved the team performance.

In summary, the trust-POMDP robot was able to make good decisions on whether to pick up the low risk object to increase human trust, or to go directly to the high risk object when trust is high enough. This is one main advantage that trust-POMDP robot has over the myopic robot.

5.4 Trust evolution

Figure 6 (center column) shows the participants' trust evolution in the online AMT experiment and the real-robot experiment. We make two key observations. First, successfully completing a task increased participants' trust in the robot. This is consistent with the human trust dynamics model we learned in Section 4.2. Second, there is a lack of significant difference in the *average* trust evolution between the two conditions (Figure 6, center column), especially given that fewer human interventions occurred under the trust-POMDP policy. This can be partially explained by a combination of averaging and nonlinear trust dynamics, specifically that robot performance in the earlier part of the task has a more pronounced impact on trust [9]. This is a specific manifestation of the "primacy effect", a cognitive bias that results in a subject crediting a performer more if the performer succeeds earlier in time [21]. Figure 7 shows this time-varying aspect of trust dynamics in our experiment; the change in the mean of trust was larger if the robot succeeded earlier, most clearly seen for the Can and Glass objects in the real-robot experiment. As such, in the myopic condition, although there were more interventions on the glass/can at the beginning, this was averaged out by a larger increase in the human trust.

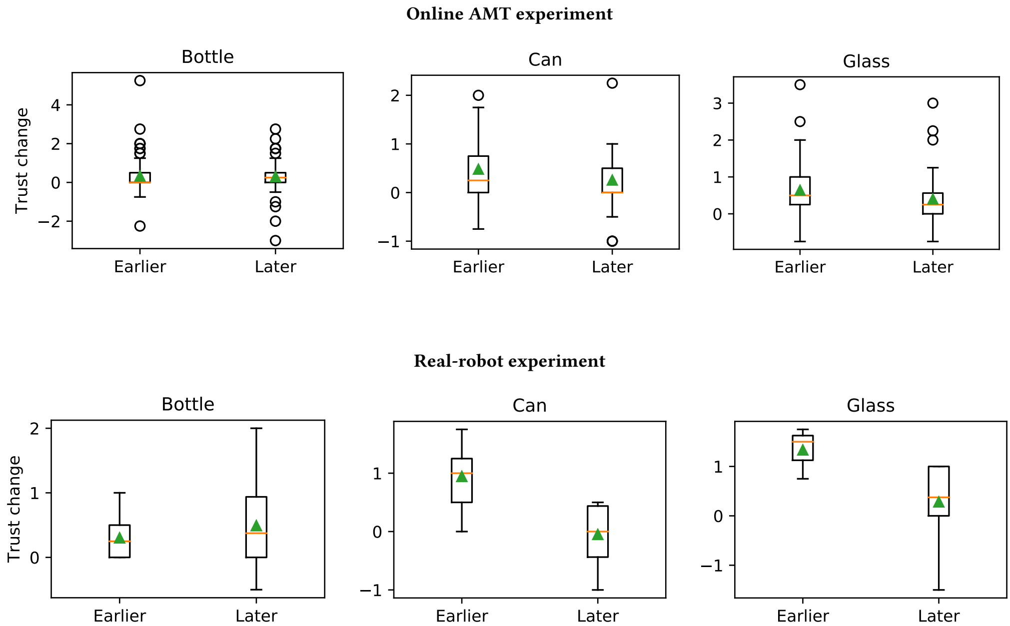


Fig. 7. Time-varying trust dynamics. The same outcome has greater effect on trust when it occurs earlier than later.

Importantly, the trust-POMDP uses trust as the means to maximize task performance and builds up human trust only when necessary, *i.e.*, the trust-POMDP robot does not maximize human trust. In addition, it may choose to ignore or even reduce human trust for maximized performance (see Section 7 for examples).

5.5 Human behavioral policy

Figure 6 (right column) shows the observed human behaviors given different trust levels. Consistent with the trust-based human behavioral model (Section 4.3), participants were less likely to intervene as their trust in the robot increased. The human's action also depended on the type of object. For low risk objects (bottles), participants allowed the robot's attempt to complete the task even if their trust in the robot was low. However, for a high risk object (glass), participants intervened unless they trusted the robot more.

6 TIME-VARYING TRUST DYNAMICS

We have observed in the experiment that robot performance in earlier part of the task has a more pronounced impact on human trust. In this section, we learned a trust dynamics model that explicitly conditioned on the time. This is different from the trust dynamics model in Section 4.2, which is time-invariant. We incorporated this time-varying trust dynamics model into a trust-POMDP for robot decision making. The resulting robot policy is more aggressive at building human trust, especially at the beginning of a task.

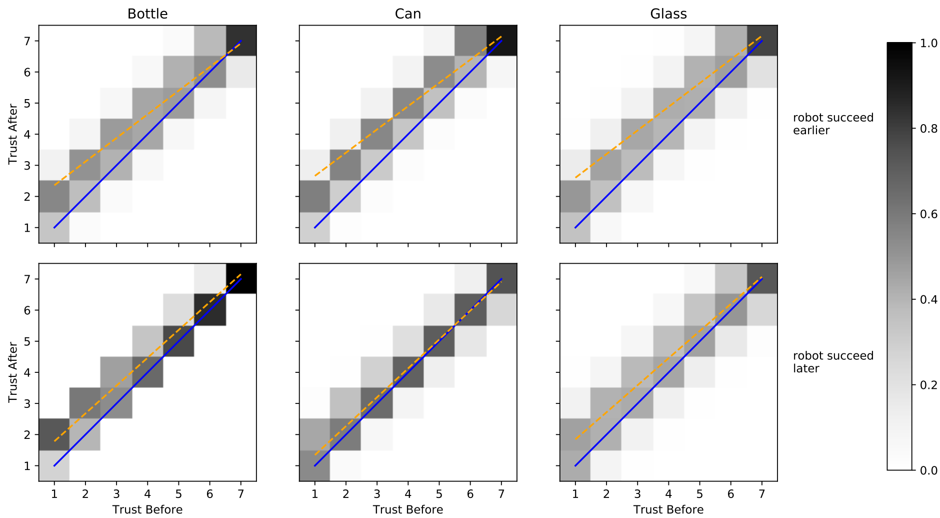


Fig. 8. Time-varying trust transition matrices, which represent the change of trust at different time of the task, shown by the linearly regressed line (yellow) contrasted with the X-Y line (blue). Overall, human trust increased more when robot succeeded earlier in time.

6.1 Time-varying trust dynamics model

Similar to Section 4.2, we model trust evolution as a linear Gaussian system that relates trust causally to the robot task performance e_{t+1}. The key difference here is that we further condition trust evolution on the time. For the table clearing task, we manually divided the time into *earlier* (1th-2th step), and *later* (3th-5th step). We formally write our time-varying trust dynamics model as follows:

$$
\begin{aligned}
&P(\theta_{t+1}|\theta_t, m_t, e_{t+1}) = \mathcal{N}(\alpha_{e_{t+1},m_t}\theta_t + \beta_{e_{t+1},m_t}, \sigma^2_{e_{t+1},m_t}) \\
&\theta_t^M \sim \mathcal{N}(\theta_t, \sigma^2)\ ,\ \theta_{t+1}^M \sim \mathcal{N}(\theta_{t+1}, \sigma^2)
\end{aligned}
\tag{10}
$$

where m_t indicates the current progress of the task, *i.e.*, m_t = *earlier* when $t \leq 2$, and m_t = *later* when $t > 2$. $\mathcal{N}(\mu, \sigma)$ denotes a Gaussian distribution with mean μ and standard deviation σ. α_{e_{t+1},m_t} and β_{e_{t+1},m_t} are the linear coefficients for the trust dynamics that explicitly conditioned on the robot performance e_{t+1} and current progress m_t. θ_t^M and θ_{t+1}^M are the observed human trust (Muir's questionnaire) at time step t and time step $t+1$.

The unknown parameters in the time-varying trust dynamics model include α_{e_{t+1},m_t}, β_{e_{t+1},m_t}, σ_{e_{t+1},m_t}, and σ. We used the trust data collected from the real-robot experiment, and we performed full Bayesian inference on the model through Hamiltonian Monte Carlo sampling using the Stan probabilistic programming platform [6].

Figure 8 shows the trust transition matrices at different time during the task (*i.e.*, earlier or later), given that the robot succeeded at that step. As we can see, human trust in the robot increased more significantly when the robot succeeded earlier (top row), compared with robot succeeded later (bottom row).

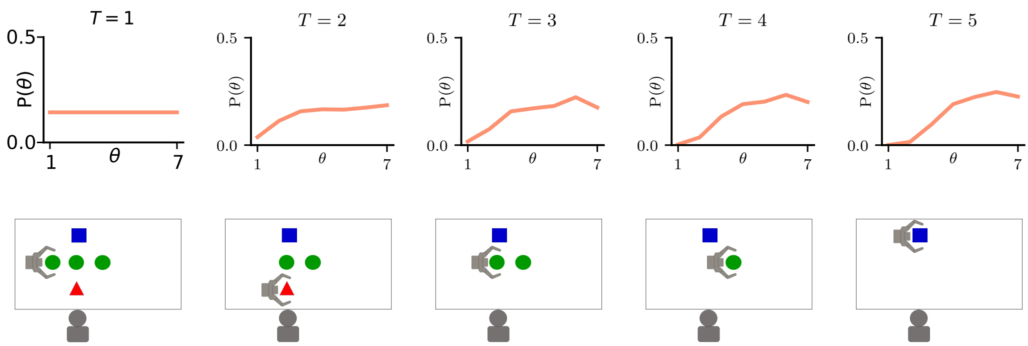


Fig. 9. Sample run of the trust-POMDP strategy with time-varying trust dynamics. The robot goes for the can object at the second step, because succeeding at the can object earlier increases trust significantly more compared with later success.

6.2 Robot policy with time-varying trust dynamics

We incorporate the time-varying trust dynamics model into a trust-POMDP, where we simply replacing the trust dynamics model in Section 4.2 with the time-varying trust dynamics model. The resulting robot policy is shown in Figure 9, which is different from the robot policy with time-invariant trust dynamics model (Figure 2, top row). In particular, the robot moves the Can object at the second step, instead of the fourth step (Figure 2, top row). This is primary due to the fact that robot moving the Can object earlier increases trust more significantly compared with moving the Can object later (Figure 8, central column).

7 ROBOT FAILURES

The previous experimental results show that the trust-POMDP policy significantly outperforms the myopic policy that ignores trust in robot decision-making. The trust-POMDP robot was able to make good decisions on whether to pick up the low risk object to increase human trust, or to go directly to the high risk object when trust is high enough. This is one main advantage that trust-POMDP robot has over the myopic robot.

In these experiments the robot always succeeded. However, in real world, the robot is also likely to fail, and we want to explore the behavior of the trust-POMDP when the robot may fail in its attempt to pick up an object with some known probability.

To incorporate robot failure into our trust-POMDP, we manually set the parameters of our robot to make it fail on the glass cup with high probability (0.9). Contrary to when the robot always succeeds, in this case it is actually beneficial for the human to intervene and pick up the glass themselves, in order to avoid the large penalty from a likely robot failure.

7.1 Robot policy with robot failures

We learned the trust dynamics model when the robot may fail on the glass with high probability, and the results are shown in Figure 4, bottom row. In general, human trust tends to decrease if the robot fails on a task. We recompute the trust-POMDP policy using the trust dynamics model with robot failures. Fig. 10 shows the computed trust-POMDP policy and belief updates: the robot starts with the glass cup, since the beginning of the task is when the human is most likely to intervene and not let the robot pick up the glass, which prevents robot failures.

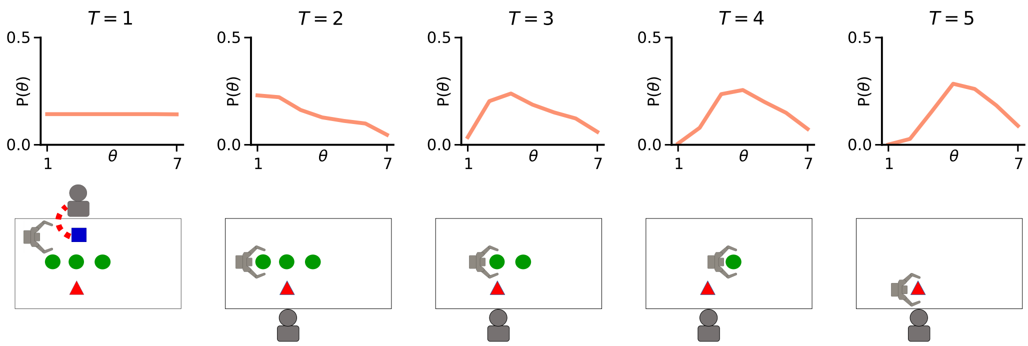


Fig. 10. Sample run of the trust-POMDP strategy when the robot may fail on the glass cup with high probability.

While this shows that the robot can reason over human intervention rate to reduce failure, intuitively the robot should also be able to *actively reduce* trust to affect human behavior. While there is a range of behaviors that can reduce human trust [43, 44], we focused on active trust reduction through failures. Therefore, we expanded the robot's action space, so that it can intentionally fail on any object. Keeping the failure probability for glass at 0.9 and reducing the reward for robot success when picking up the bottles to 0.3 results in the exciting behavior demonstrated at Fig. 11.

When following the trust-POMDP policy (Fig. 11 top and middle row) the robot attempts to pick up the can first; This is an *information seeking* action, that the robot uses to estimate the initial human trust. If the human stays put, the robot infers that human trust is high, and it will then *fail intentionally* at the bottles to reduce trust, before going for the glass cup. By the time the robot goes for the glass cup, human trust has been reduced sufficiently so that the human is likely to intervene, avoiding failure. On the other hand, if the human intervenes, the robot infers that the human trust is already low. The robot then does not need to fail intentionally, since it does not need to reduce human trust any further, so it subsequently goes for the glass cup.

The resulting policy contrasts the policy that the robot follows, if it maximizes human trust instead (Fig. 11, bottom row). When following the trust-maximizing policy, the robot starts with the glass. Human trust at the beginning is relatively low, therefore, the human is most likely to intervene and stop the robot from picking the glass. This prevents significant reduction of human trust, as the robot will fail on the glass with high probability.

7.2 Simulation results

We further illustrate the difference between the two policies by simulating policy runs and showing the evolution of the expected trust and mean accumulated reward over time (Fig. 12, 13).

The plots show that the performance-maximizing policy is significantly different from the trust-maximizing policy, *i.e.*, the performance-maximizing policy actively reduces human trust to maximize reward. The mean accumulated reward over 10^4 policy runs for the performance-maximizing policy is -1.33, compared to -1.67 for the trust-maximizing policy, with a statistically significant difference ($F(1, 19998) = 29.42, p < 0.001$). This evaluation indicates that maximizing trust can be suboptimal in the presence of robotic failures.

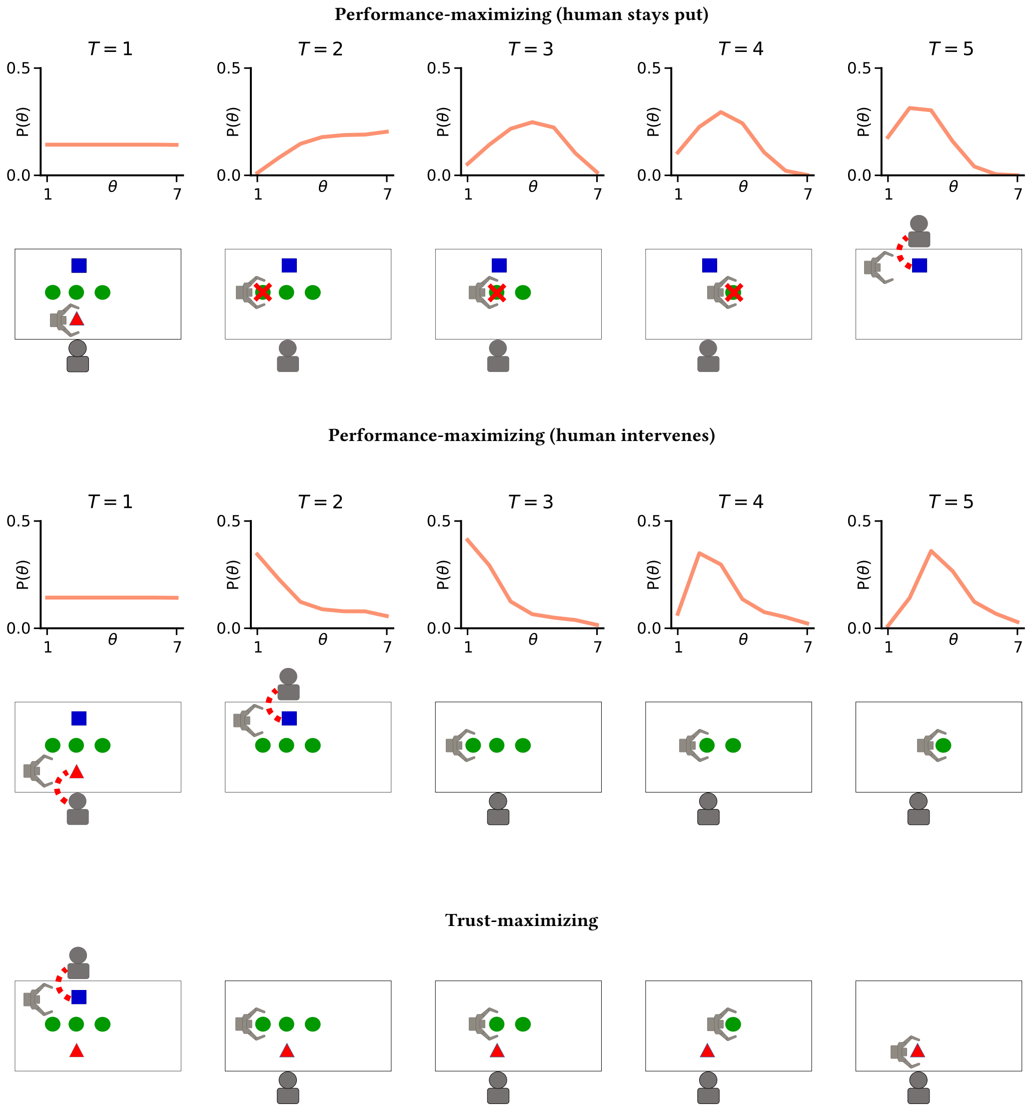


Fig. 11. Sample runs of the performance-maximizing policy (*i.e.*, trust-POMDP, top and middle row) and the trust-maximizing policy (bottom row) when the robot may fail on the glass cup with high probability, and the robot can fail intentionally in any object. The adaptive trust-POMDP policy starts with can object and branches out at $T = 1$. If the human stays put (top row), the robot intentionally fails in the bottles to reduce human trust and maximize the probability of the human intervening when it goes for the glass at $T = 5$. If the human intervenes (middle row), the robot realizes that the human trust is already low, and it goes for the glass cup at $T = 3$. On the other hand, the trust-maximizing policy starts with the glass, since the human is most likely to intervene at the beginning as the human trust is relatively low.

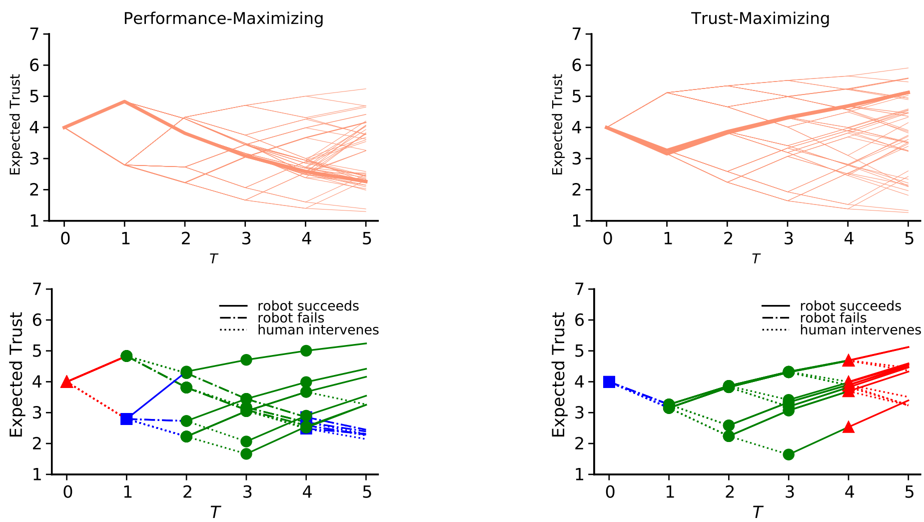


Fig. 12. (Top) Expected trust for all possible human action sequences for the performance-maximizing and trust-maximizing policy. Each sequence is represented with a line of width proportional to the likelihood of that sequence, based on the learned model. (Bottom) Annotated robot actions for the 16 most likely sequences.

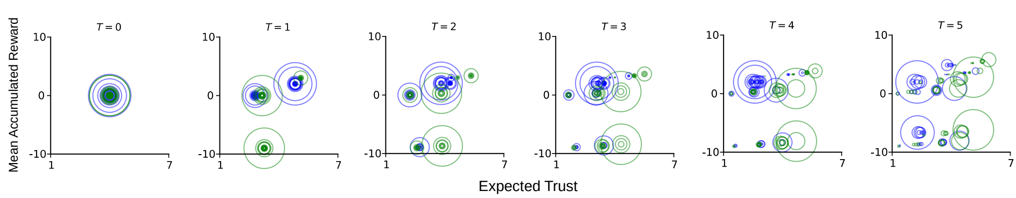


Fig. 13. Scatterplot of mean accumulated reward as a function of human trust over time for all human action sequences of the performance-maximizing policy (blue) and the trust-maximizing policy (green). The radius of each circle is proportional to the likelihood of the corresponding sequence, based on the learned model. The performance-maximizing policy (blue) gradually reduces human trust to maximize the accumulated reward, while the trust-maximizing policy (green) focuses on increasing trust.

8 DISCUSSION

Summary. This paper presents the trust-POMDP, a computational model for integrating human trust into robot decision making. The trust-POMDP closes the loop between trust models and robot decision making. It enables the robot to infer and influence human trust systematically and to leverage trust for fluid collaboration.

Our experimental results in a table-clearing task show that the trust-POMDP policy calibrates human trust to match it to the robot's manipulation capabilities: If trust is overly low, the robot prioritizes picking up the low risk objects to increase trust. This results in better performance, compared to the myopic robot that ignores trust. On the other hand, if trust is overly high, the robot fails intentionally in the low risk objects. Our results show that always maximizing trust can be in fact detrimental to performance in the presence of robotic failures.

Scaling up. Solving a trust-POMDP is essentially solving a standard POMDP model [22], which is computationally expensive in general. The SARSOP algorithm [24, 34] used in this paper is the state of the art offline POMDP solver, which computes an optimal action offline for every possible belief state. SARSOP is able to scale up to problems with $O(10^5)$ number of states [34], which roughly equals to the table clearing task with 9 objects on the table (the number of states in the table clearing task is 4^N where N is the number of objects).

To scale up trust-POMDP to large robotic tasks, one can apply the online POMDP algorithms. Unlike offline POMDP algorithms, the online POMDP algorithm performs a look-ahead search and computes a best action for the current belief only [38]. After the robot executes the action, the algorithm updates the belief based on the observation received. The process then repeats at the new belief for the next time step. Online search algorithms scale up by focusing on the current belief only, rather than all possible beliefs that the robot may encounter. POMCP [40] and DESPOT [42] are the fastest online POMDP algorithms available today. Both employ the idea of sampling future contingencies. In particular, DESPOT has been successfully implemented for real-time ($3hz$) autonomous driving in a crowd [1], where the problem has continuous state and a planning horizon of 90 steps.

Limitations. There are several limitations in our current work. Similar to previous works [9, 46], we modeled trust as a single real-valued latent variable that reflected the capabilities of the entire system. However, a multi-dimensional parameterization of trust that captured the different functions and modes of automation could be a more accurate representation. In addition, the evolution of trust might also depend on the type of motion executed by the robot (e.g., for expressive or deceptive motions [12, 13]). The current trust-POMDP model also assumes static robot capabilities, but a robot's true capabilities may change over time. In fact, the trust-POMDP can be extended to model robot capabilities via additional state variables that affect the state transition dynamics. Furthermore, the reward function is manually specified in this work. However, the reward function may be difficult to specify in practice. One possible way to resolve this is to learn the reward function from human demonstrations (e.g., [33]).

Future works. This work points to several directions for further investigation. In addition to addressing the limitations aforementioned, we consider two particularly promising topics.

First, the trust-POMDP is applicable in many real-world robotic tasks. In the current work, we have chosen a simple table clearing task to test the trust-POMDP model, because it allows us to analyze experimentally the core technical issues on human trust without interference from confounding factors. However, we believe that the overall technical approach in our work is general and not restricted to this particular simplified task. What we learned here on the trust-POMDP for a simplified task will be a stepstone towards more complex, large-scale applications. For example, consider a rescue task, where the robot and the human collaborate to rescue survivors in a disaster scene. The robot can quickly search over the disaster scene, but the robot sensors have noises and might fail at certain scenarios. To complete the task successfully, the human and the robot have to collaborate and take advantage of their own strengths. Modeling human trust is important in the rescue task. If the human trust is too low (under-reliance), the rescue will be inefficient as the robot is not well utilized. On the other hand, if the human trust is too high (over-reliance), they may miss some survivors as the robot can fail. With trust-POMDP, the robot can actively modulate human trust to avoid under-reliance and over-reliance.

Second, the trust-POMDP is a generative decision model conditioned explicitly on trust. Explicit trust modeling provides several advantages: it fits better to experimental data (see Section 4.3), and potentially improves the efficiency of learning by reducing sample complexity. Most importantly, the trust model learned on one task may transfer to a related task [41]. This last aspect is another interesting direction for future work.

9 ACKNOWLEDGEMENTS

This work was funded in part by the Singapore Ministry of Education (grant MOE2016-T2-2-068), the National University of Singapore (grant R-252-000-587-112), US National Institute of Health R01 (grant R01EB019335), US National Science Foundation CPS (grant 1544797), US National Science Foundation NRI (grant 1637748), and the Office of Naval Research.

REFERENCES

[1] H.Y. Bai, S.J. Cai, D. Hsu, and W.S. Lee. 2015. Intention-Aware Online POMDP Planning for Autonomous Driving in a Crowd. In *Proc. IEEE Int. Conf. on Robotics & Automation.*

[2] Haoyu Bai, Shaojun Cai, Nan Ye, David Hsu, and Wee Sun Lee. 2015. Intention-aware online POMDP planning for autonomous driving in a crowd. In *2015 IEEE International Conference on Robotics and Automation (ICRA).* IEEE, 454–460.

[3] Tirthankar Bandyopadhyay, Kok Sung Won, Emilio Frazzoli, David Hsu, Wee Sun Lee, and Daniela Rus. 2013. Intention-aware motion planning. In *Algorithmic Foundations of Robotics X.* Springer, 475–491.

[4] Bernard. Barber. 1983. *The logic and limits of trust / Bernard Barber.* Rutgers University Press New Brunswick, N.J. 189 p. ; pages.

[5] Samuel Barrett, Noa Agmon, Noam Hazon, Sarit Kraus, and Peter Stone. 2014. Communicating with unknown teammates. In *Proceedings of the twenty-first european conference on artificial intelligence.* IOS Press, 45–50.

[6] Bob Carpenter, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2016. Stan: A probabilistic programming language. *Journal of Statistical Software* 20 (2016), 1–37.

[7] Min Chen, Stefanos Nikolaidis, Harold Soh, David Hsu, and Siddhartha Srinivasa. 2018. Planning with trust for human-robot collaboration. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 307–315.

[8] Nathaniel D Daw, John P O'doherty, Peter Dayan, Ben Seymour, and Raymond J Dolan. 2006. Cortical substrates for exploratory decisions in humans. *Nature* 441, 7095 (2006), 876–879.

[9] Munjal Desai. 2012. Modeling trust to improve human-robot interaction. (2012).

[10] Munjal Desai, Poornima Kaniarasu, Mikhail Medvedev, Aaron Steinfeld, and Holly Yanco. 2013. Impact of robot failures and feedback on real-time trust. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction.* IEEE Press, 251–258.

[11] Munjal Desai, Mikhail Medvedev, Marynel Vázquez, Sean McSheehy, Sofia Gadea-Omelchenko, Christian Bruggeman, Aaron Steinfeld, and Holly Yanco. 2012. Effects of changing reliability on trust of robot systems. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction.* ACM, 73–80.

[12] Anca D Dragan, Rachel M Holladay, and Siddhartha S Srinivasa. 2014. An Analysis of Deceptive Robot Motion.. In *Robotics: science and systems.* 10.

[13] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on.* IEEE, 301–308.

[14] Michael W Floyd, Michael Drinkwater, and David W Aha. 2015. Trust-Guided Behavior Adaptation Using Case-Based Reasoning. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence.* 4261–4267.

[15] Enric Galceran, Alexander G Cunningham, Ryan M Eustice, and Edwin Olson. 2015. Multipolicy decision-making for autonomous driving via changepoint-based behavior prediction. In *Proc. Robot.: Sci. & Syst. Conf.* 2.

[16] Robert T Golembiewski and Mark McConkie. 1975. The centrality of interpersonal trust in group processes. *Theories of group processes* 131 (1975), 185.

[17] Robert J Hall. 1996. Trusting your assistant. In *Knowledge-Based Software Engineering Conference, 1996., Proceedings of the 11th.* IEEE, 42–51.

[18] Guy Hoffman. 2013. Evaluating fluency in human-robot collaboration. In *International conference on human-robot interaction (HRI), workshop on human robot collaboration*, Vol. 381. 1–8.

[19] Shervin Javdani, Siddhartha S Srinivasa, and J Andrew Bagnell. 2015. Shared autonomy via hindsight optimization. *arXiv preprint arXiv:1503.07619* (2015).

[20] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. 2000. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics* 4, 1 (2000), 53–71.

[21] Edward E Jones, Leslie Rock, Kelly G Shaver, George R Goethals, and Lawrence M Ward. 1968. Pattern of performance and ability attribution: An unexpected primacy effect. *Journal of Personality and Social Psychology* 10, 4 (1968), 317.

[22] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra. 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101, 1–2 (1998), 99–134.

[23] Roderick M Kramer and Tom R Tyler. 1995. *Trust in organizations: Frontiers of theory and research.* Sage Publications.

[24] Hanna Kurniawati, David Hsu, and Wee Sun Lee. 2008. SARSOP: Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces.. In *Robotics: Science and Systems*, Vol. 2008. Zurich, Switzerland.

[25] John Lee and Neville Moray. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35, 10 (1992), 1243–1270.

[26] John D Lee and Katrina A See. 2004. Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46, 1 (2004), 50–80.

[27] Owen Macindoe, Leslie Pack Kaelbling, and Tomás Lozano-Pérez. 2012. Pomcop: Belief space planning for sidekicks in cooperative games. (2012).

[28] John E Mathieu, Tonia S Heffner, Gerald F Goodwin, Eduardo Salas, and Janis A Cannon-Bowers. 2000. The influence of shared mental models on team process and performance. *Journal of applied psychology* 85, 2 (2000), 273.

[29] Roger C Mayer, James H Davis, and F David Schoorman. 1995. An integrative model of organizational trust. *Academy of management review* 20, 3 (1995), 709–734.

[30] Bonnie Marlene Muir. 1990. *Operators' trust in and use of automatic controllers in a supervisory process control task.* University of Toronto.

[31] Stefanos Nikolaidis, David Hsu, and Siddhartha Srinivasa. 2017. Human-robot mutual adaptation in collaborative tasks: Models and experiments. *International Journal of Robotics Research* 36, 5-7 (2017), 618–634.

[32] Stefanos Nikolaidis, Anton Kuznetsov, David Hsu, and Siddhartha Srinivasa. 2016. Formalizing Human-Robot Mutual Adaptation: A Bounded Memory Model. In *HRI*. IEEE Press, 75–82.

[33] Stefanos Nikolaidis, Ramya Ramakrishnan, Keren Gu, and Julie Shah. 2015. Efficient model learning from joint-action demonstrations for human-robot collaborative tasks. In *HRI*. ACM, 189–196.

[34] Sylvie CW Ong, Shao Wei Png, David Hsu, and Wee Sun Lee. 2010. Planning under uncertainty for robotic tasks with mixed observability. *The International Journal of Robotics Research* 29, 8 (2010), 1053–1068.

[35] Alyssa Pierson and Mac Schwager. 2016. Adaptive inter-robot trust for robust multi-robot sensor coverage. In *Robotics Research*. Springer, 167–183.

[36] Charles Pippin and Henrik Christensen. 2014. Trust modeling in multi-robot patrolling. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 59–66.

[37] John K Rempel, John G Holmes, and Mark P Zanna. 1985. Trust in close relationships. *Journal of personality and social psychology* 49, 1 (1985), 95.

[38] Stephane Ross, Joelle Pineau, Sebastien Paquet, and Brahim Chaib-Draa. 2008. Online Planning Algorithms for POMDPs. *J. Artif. Intell. Res.(JAIR)* 32 (2008), 663–704.

[39] Dorsa Sadigh, Shankar Sastry, Sanjit A Seshia, and Anca D Dragan. 2016. Planning for autonomous cars that leverages effects on human actions. In *Proceedings of the Robotics: Science and Systems Conference (RSS)*.

[40] D. Silver and J. Veness. 2010. Monte-Carlo planning in large POMDPs. In *Advances in Neural Information Processing Systems*.

[41] Harold Soh, Pan Shu, Min Chen, and David Hsu. 2018. The Transfer of Human Trust in Robot Capabilities across Tasks. *arXiv preprint arXiv:1807.01866* (2018).

[42] A. Somani, N. Ye, D. Hsu, and W.S. Lee. 2013. DESPOT: Online POMDP Planning with Regularization. In *Advances in Neural Information Processing Systems*.

[43] Rik van den Brule, Ron Dotsch, Gijsbert Bijlstra, Daniel HJ Wigboldus, and Pim Haselager. 2014. Do robot performance and behavioral style affect human trust? *International journal of social robotics* 6, 4 (2014), 519–531.

[44] Ning Wang, David V Pynadath, and Susan G Hill. 2016. Trust calibration within a human-robot team: Comparing automatically generated explanations. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 109–116.

[45] Yaqi Xie, Indu Prasad, Desmond Ong, David Hsu, and Harold Soh. 2019. Robot Capability and Intention in Trust-based Decisions across Tasks. In *HRI*. ACM.

[46] Anqi Xu and Gregory Dudek. 2015. Optimo: Online probabilistic trust inference model for asymmetric human-robot collaborations. In *HRI*. ACM, 221–228.

[47] Anqi Xu and Gregory Dudek. 2016. Towards Modeling Real-Time Trust in Asymmetric Human–Robot Collaborations. In *Robotics Research*. Springer, 113–129.

[48] Jessie Yang, Vaibhav Unhelkar, Kevin Li, and Julie Shah. 2017. Evaluating Effects of User Experience and System Transparency on Trust in Automation. In *HRI*.