Design of Human-Aware Robotic Decision Support Systems

Manisha Natarajan[1], Chunyue Xue[1], Karen Feigh[2] and Matthew Gombolay[1]

***Abstract*— Advances in robotics and artificial intelligence (AI) have enabled the possibility of human-robot teaming. One potential avenue for collaborative robots is to provide decision-support for human partners in complex decision-making tasks. However, such agents are imperfect in real-world scenarios and may provide incorrect or suboptimal recommendations. Thus, it is imperative for human collaborators to understand when to trust the robot's suggestions for maximizing task performance. Explainable AI (xAI) attempts to improve user understanding by providing explanations or rationales for agent recommendations. However, constantly providing explanations is unnecessary and can induce cognitive overload among users. In this work, we propose a POMDP framework that allows the robot to infer the users' latent trust and preferences to provide appropriate and timely explanations for maximizing human-robot team performance in a sequential decision-making game.**

I. Introduction

Robots and Artificial Intelligence (AI) agents have the potential to improve human decision-making by providing recommendations or relevant information. Currently, AI systems are aiding humans in decision-making for safety-critical scenarios such as clinical diagnosis [10], as well as everyday activities such as navigation. Robots and other automated agents are designed to increase efficiency and reduce cognitive demands on humans. However, such agents can be imperfect in complex settings, and dependence on imperfect automation can lead to severe consequences, especially in safety-critical tasks [1], [14]. In this work, we explore how to help humans discern when to trust and depend on suboptimal robotic decision support systems (DSS) by providing appropriate and timely explanations.

Trust is defined as an "attitude that an agent (automation or another person) will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" [8], and dependence is a behavioral measure that indicates the extent to which the user accepts recommendations from an agent. Users' dependence upon robots and other AI agents is often correlated with their trust in the agent [3], [11]. To avoid negative outcomes, we need to mitigate inappropriate dependence (i.e., Type I and Type II errors). *Type I* error or *over-reliance* refers to the extent to which users accept poor recommendations from an agent, and *Type II* error or *under-reliance* refers to the extent to which users reject good recommendations from an agent [2].

Our current research is centered on robotic DSS as previous studies have demonstrated that robot embodiment can foster trust and appropriate reliance among both experts [5] and novice users [12]. Additionally, robots deployed as DSS are significantly more capable than their software counterparts, as they can perform physical tasks such as navigation, manipulation and acquire real-time data from sensors for providing decision support. Some real-world applications of robotic DSS include manufacturing – where robots can help with quality control and inventory management [13] and healthcare – for clinical diagnosis assistance [10] and resource allocation [5]. Some recent works have also investigated using robots to provide assistance in making judicial decisions [4].

Our proposed research has three goals. First, we develop a computational cognitive model for human decision-making with robotic DSS, which infers the latent trust dynamics and individual capability based on the user's interaction history with the robot. Second, we utilize the cognitive model to learn a robot policy for providing effective assistance in human decision-making. We model the human-robot interaction (HRI) as a Partially Observable Markov Decision Process (POMDP), where the robot agent must decide (1) when to provide assistance, (2) whether it is beneficial to provide explanations, and (3) the type of explanation based on individual preferences and the task context. Lastly, we will evaluate our proposed approach by comparing the performance of the human-robot team employing the POMDP framework against other baselines that do not utilize the cognitive decision-making model. Our proposed work is novel since it addresses modeling multiple latent human states in a sequential decision-making task, where users cannot immediately evaluate the optimality of the robot.

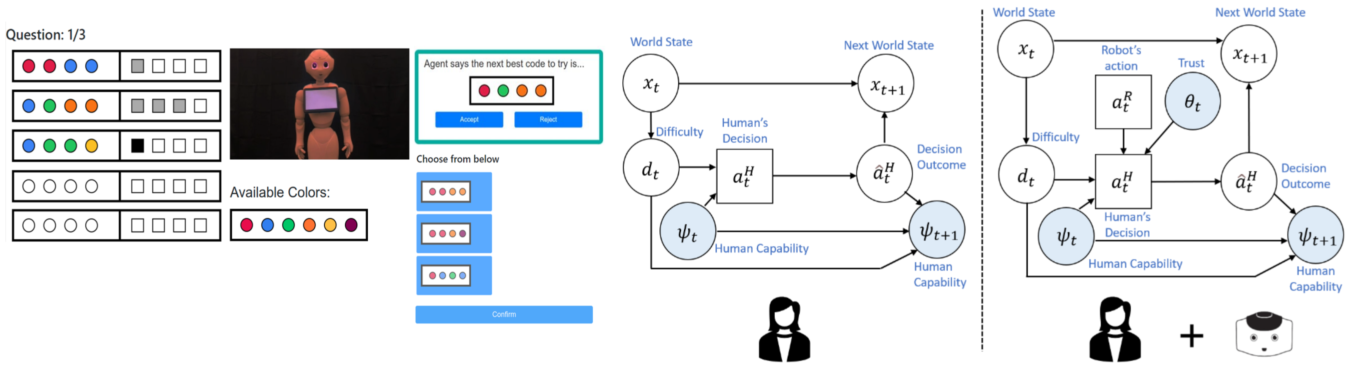
II. Methodology

The objective of this work is to develop an adaptive robot policy that can determine when and how to assist humans to mitigate *Type I* and *Type II* errors and optimize human-robot team performance in sequential decision-making tasks.

A. Task Domain

In this work, we will use Mastermind [7], a sequential decision-making game, for modeling humans. In Mastermind, users must identify a four-color secret code in the fewest turns possible using feedback from the game (colored squares) as shown in Figure 1a. At every turn, the users must choose from a list of choices, and their decision will influence the choices available in subsequent turns. Users are scored based on the number of turns and time taken to identify the code. The game ends when the users identify the correct code or exceed the maximum turns. The users solve

[1]College of Computing, Georgia Institute of Technology, USA.
[2]Dept. of Aerospace Engineering, Georgia Institute of Technology, USA.
Corresponding Author: Manisha Natarajan, mnatarajan30@gatech.edu



(a) Robot providing decision-support in Mastermind. The user can accept or reject the robot's suggestion at each turn.

(b) Cognitive models of human decision-making with and without robotic decision-support. Shaded regions imply latent states unknown to the robot.

Fig. 1: Design framework for human-aware robotic DSS. Figure 1a depicts the task domain with robot assistance and Figure 1b depicts the cognitive model that will be used by the robot agent to determine when and how to assist humans.

the task independently or with robotic assistance, where they can either accept or reject the robot's suggestion.

B. *Cognitive Model for Human Decision Making*

We model the human-robot interaction for suboptimal robotic DSS as a POMDP, with world state $x \in \mathcal{X}$, robot action $a^R \in \mathcal{A}^R$, and human action $a^H \in \mathcal{A}^H$. The world state x is an abstract representation of the current game state (the current turn, a boolean indicator for identifying the code, and an information gain metric to assess how far away the user is from guessing the correct code). At each turn t, both the robot and the user perform an action — a_t^R and a_t^H, respectively. Since the robot only acts as a DSS, its actions will only influence the human's actions and not directly change the world state, i.e., world state dynamics is only impacted by the human's action – $p(x_{t+1}|x_t, a_t^R, a_t^H) = p(x_{t+1}|x_t, a_t^H)$. We assume that the user's decision in each turn is dependent on their trust in the robot θ_t, their ability to solve the task independently ψ_t, and the task difficulty d_t. The task difficulty d_t reduces as the user moves closer to finding the code by eliminating various possibilities. The user's trust θ_t and capability ψ_t are latent or unknown to the robot *apriori*. Thus, the robot agent maintains a belief over the user's latent states and updates them by observing the user's actions a_t^H (accept or reject) and outcome $\hat{a}_t^H$.

1) Modeling Human Decisions without Robot Assistance: In the case of no robotic assistance, the human's decision-making policy will only be dependent on their capability ψ_t and the task difficulty d_t $(= f(x_t))$, i.e., $\pi^H(a_t^H|\hat{s}_t) = p(a_t^H|d_t, \psi_t)$, where $\hat{s}_t \equiv (x_t, \psi_t)$ is the augmented state. Then the overall transition dynamics is given by:

$$p(\hat{s}_{t+1}|a_t^H, \hat{s}_t) = p(x_{t+1}|x_t, a_t^H) \times p(\psi_{t+1}|\psi_t, a_t^H, d_t) \quad (1)$$

2) Modeling Human Decisions with Robot Assistance: In the case of robotic assistance, the human's decision-making policy will additionally be dependent on their trust in the robot θ_t, and the action taken by the robot (e.g., the type of explanation provided by the robot) a_t^R. Thus, the human's decision-making policy can be modeled as $\pi^H(a_t^H|s_t, a_t^R) = p(a_t^H|d_t, \psi_t, \theta_t, a_t^R)$, where $s_t \equiv (x_t, \psi_t, \theta_t)$ is the augmented state. Then the overall transition dynamics becomes:

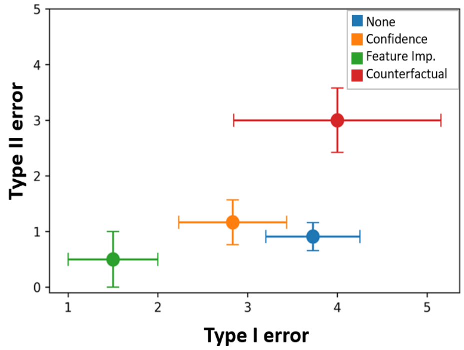
$$p(s_{t+1}|a_t^H, a_t^R, s_t) = p(x_{t+1}|x_t, a_t^H) \times p(\psi_{t+1}|\psi_t, a_t^H, d_t) \times p(\theta_{t+1}|\theta_t, a_t^H, a_t^R, d_t) \quad (2)$$

C. *Learning an Adaptive Policy for Effective Assistance*

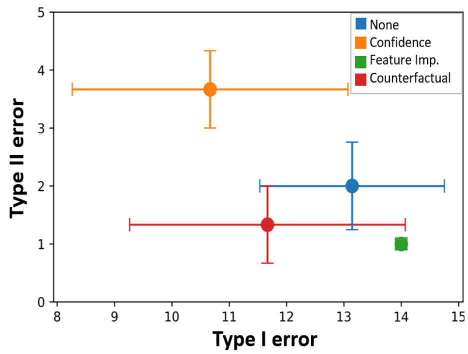
Since the user's trust and capability are unknown to the robot, we model the HRI as a POMDP. We propose to use the Bayes-Adaptive POMDP framework (BA-POMDP) [15] to simultaneously learn the parameters of the task domain (i.e., the transition and observation probabilities) and learn a policy for the robot that maximizes a reward function. The BA-POMDP framework utilizes Dirichlet distributions to represent unknown distributions. Maximizing the reward function $R = \alpha \times \Delta\text{info gain} - \beta \times \Phi(a_t^R)$ enables the robot's policy to balance when to provide recommendations based on its estimation of the human's ability to solve the task. The Δinfo gain is a pseudo-measure for the difference in task progress between human actions with and without robot assistance, and the cost Φ of robot actions is proportional to the cognitive effort required by the human to evaluate. We propose to use a variant of POMCP [17], an online solver, for learning the robot policy, as shown in prior work [9].

D. *Human Subjects Experiments*

We propose a two-phase user study design to (1) learn the parameters for the proposed cognitive model and train the robot policy and (2) evaluate the trained policy on new users. Our first user study (**data collection phase**) is designed to assess user performance in our task domain with robots providing various explanations common in xAI literature, such as counterfactual, feature importance, and confidence-based explanations [16]. Preliminary results from Phase I are discussed in Section III. In the second user study (**evaluation phase**), we will compare the performance of human-robot teams employing the proposed adaptive BA-POMCP policy against other baselines discussed in Section II-E.



(a) Type I / Type II errors for high performers



(b) Type I / Type II errors for low performers

Fig. 2: Type I and Type II error rates with respect to different explanation types from the data collection study.

E. Evaluation Baselines

We propose to compare our approach against prior work by Hong et al. [6] that uses offline reinforcement learning to influence suboptimal humans for maximizing human-AI team performance. Our approach is different as it is aimed at collaborative settings where both the human and the robot agents are suboptimal.

III. Data Collection Study: Preliminary Results

We conducted a 1×3 (explanation type) between-subjects experiment to assess user performance with and without robotic assistance for various explanation types. Each user played three rounds of Mastermind with the robot providing either no assistance, only suggestions, or suggestions with an explanation (confidence / counterfactual / feature importance) in random order. We collected data from 36 subjects on the Prolific platform upon IRB approval. We segregated users as high or low performers based on their performance without robotic assistance. We find that feature importance explanations reduced Type I and Type II errors in high performers and Type II errors in low performers. Overall, high performers tended to rely more appropriately on the robot (had fewer errors) as shown in Figures 2a and 2b.

IV. Conclusion

In this work, we propose a novel, adaptive robotic DSS framework that seeks to maximize human-robot team performance by providing appropriate and timely explanations in a sequential decision-making game. Our preliminary investigation reveals that only certain explanations are useful in establishing appropriate dependence. We need to further investigate which robot actions can influence appropriate dependence across different user populations. Upon collecting further data, we seek to train the adaptive robot policy using a variant of the BA-POMCP algorithm. We expect that the performance of human-robot teams employing the BA-POMCP framework will be significantly higher than our baselines from prior work.

References

[1] Key Dismukes, Benjamin A Berman, and Loukia D Loukopoulos. *The limits of expertise: Rethinking pilot error and the causes of airline accidents*. Ashgate Publishing, Ltd., 2007.

[2] Stephen Dixon and Christopher Wickens. Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human factors*, 48:474–86, 02 2006.

[3] Mary T Dzindolet, Scott A Peterson, Regina A Pomranky, Linda G Pierce, and Hall P Beck. The role of trust in automation reliance. *International journal of human-computer studies*, 58(6):697–718, 2003.

[4] Pedro Rubim Borges Fortes. Paths to digital justice: Judicial robots, algorithmic decision-making, and due process. *Asian Journal of Law and Society*, 7(3):453–469, 2020.

[5] Matthew Craig Gombolay, Xi Jessie Yang, Bradley Hayes, Nicole Seo, Zixi Liu, Samir Wadhwania, Tania Yu, Neel Shah, Toni Golen, and Julie A. Shah. Robotic assistance in coordination of patient care. In *Robotics: Science and Systems*, 2016.

[6] Joey Hong, Anca Dragan, and Sergey Levine. Learning to influence human behavior with offline reinforcement learning. *arXiv preprint arXiv:2303.02265*, 2023.

[7] Donald E Knuth. The computer as master mind. *Journal of Recreational Mathematics*, 9(1):1–6, 1976.

[8] John D. Lee and Katrina A. See. Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1):50–80, 2004.

[9] Joshua Lee, Jeffrey Fong, Bing Cai Kok, and Harold Soh. Getting to know one another: Calibrating intent, capabilities and trust for human-robot collaboration. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6296–6303. IEEE, 2020.

[10] Izidor Mlakar, Urška Smrke, Vojko Flis, Andrej Bergauer, Nina Kobilica, Tadej Kampič, Samo Horvat, Damjan Vidovič, Bojan Musil, and Nejc Plohl. A randomized controlled trial for evaluating the impact of integrating a computerized clinical decision support system and a socially assistive humanoid robot into grand rounds during pre/post-operative care. *Digital Health*, 8:20552076221129068, 2022.

[11] Manisha Natarajan and Matthew Gombolay. Effects of anthropomorphism and accountability on trust in human robot interaction. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 33–42, 2020.

[12] Richard Pak, Nicole Fink, Margaux Price, Brock Bass, and Lindsay Sturre. Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics*, 55(9):1059–1072, 2012.

[13] Ahmed Rhiat, Lamine Chalal, and Allal Saadane. A smart warehouse using robots and drone to optimize inventory management. In *Proceedings of the Future Technologies Conference (FTC) 2021, Volume 1*, pages 475–483. Springer, 2022.

[14] Paul Robinette, Ayanna Howard, and Alan Wagner. Effect of robot performance on human–robot trust in time-critical situations. *IEEE Transactions on Human-Machine Systems*, PP:1–12, 01 2017.

[15] Stephane Ross, Brahim Chaib-draa, and Joelle Pineau. Bayes-adaptive pomdps. *Advances in neural information processing systems*, 20, 2007.

[16] Andrew Silva, Mariah Schrum, Erin Hedlund-Botti, Nakul Gopalan, and Matthew Gombolay. Explainable artificial intelligence: Evaluating the objective and subjective impacts of xai on human-agent interaction. *International Journal of Human–Computer Interaction*, 2022.

[17] David Silver and Joel Veness. Monte-carlo planning in large pomdps. *Advances in neural information processing systems*, 23, 2010.