Learning and Communicating the Latent States of Human-Machine Collaboration

Vaibhav V. Unhelkar and **Julie A. Shah**
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
{unhelkar, julie_a_shah}@csail.mit.edu

1 Introduction

Artificial agents (both embodied robots and software agents) that interact with humans are increasing at an exceptional rate. Yet, achieving seamless collaboration between artificial agents and humans in the real world remains an active problem [Thomaz *et al.*, 2016]. A key challenge is that the agents need to make decisions without complete information about their shared environment and collaborators. For instance, a human-robot team performing a rescue operation after a disaster may not have an accurate map of their surroundings. Even in structured domains, such as manufacturing, a robot might not know the goals or preferences of its human collaborators [Unhelkar *et al.*, 2018].

Algorithmically, this challenge manifests itself as a problem of decision-making under uncertainty in which the agent has to reason about the latent states of its environment and human collaborator. However, in practice, quantifying this uncertainty (i.e., the state transition function) and even specifying the features (i.e., the relevant states) of human-machine collaboration is difficult. Thus, the objective of this thesis research is to develop novel algorithms that enable artificial agents to learn and reason about the latent states of human-machine collaboration and achieve fluent interaction.

2 Research Problems

Consider a human-machine team (depicted in Fig. 1) performing a sequential collaborative task with a known and shared objective (reward). Both the (human and artificial) agents have autonomy over their actions, receive separate observations and can communicate with each other. During task execution, however, there is information asymmetry between the agents — either due to the inherently decentralized nature of the multi-agent task, incomplete knowledge (due to the latent states) of the environment, an inaccurate model (due to the latent states) of the teammate's behavior, or differences in reasoning approaches of the agents.

Prior studies of teamwork [Jonker *et al.*, 2011; Butchibabu *et al.*, 2016] indicate that *effective* information sharing between teammates – both prior to and during task execution – is critical to the success of human-machine collaboration. Hence, our solutions leverage communication between human and artificial agents for effective learning and utilization of the latent states of human-machine collaboration. Specifically, we consider the following problems,

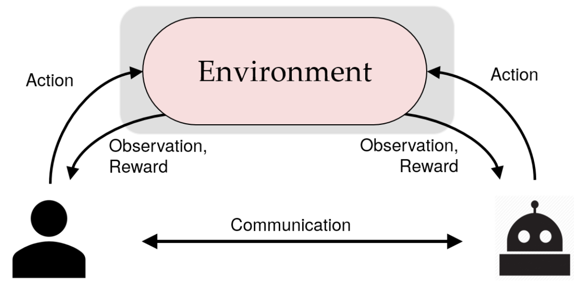

Figure 1: An abstraction for a human-machine team performing a collaborative task in a shared environment.

- learning a model of the human collaborator's sequential decision-making without complete specification of the features that impact the human's decisions, and
- enabling artificial agents to effectively share information during execution of human-machine collaborative tasks in unknown (or partially known) environments.

3 Learning Latent States of Decision-Making

Models of human collaborator's decision-making are useful for anticipating their actions, gauging their information needs and coordinating shared plans. Indeed, several representations and algorithms have been developed to model and learn another agent's decision-making behavior [Albrecht and Stone, 2018; Ziebart *et al.*, 2008]. Irrespective of the algorithms used for learning, a human's decision-making behavior can be summarized by their policy (π), i.e., the probability of choosing an action (a) in a given situation or state (f). However, implicit to describing the policy is the knowledge of available actions (i.e., actions $\{a \in A\}$) and features that impact the decisions (i.e., states $\{f \in F\}$).

Based on the sequential task, it is usually possible to specify the action space (A) of the human. However, the factors (f) that impact human's decisions – in addition to the known task-specific features (s) – often include internal variables (e.g., fatigue and trust) that are latent, dynamic and difficult to specify. Learning another agent's decision-making model when the decision factors are unknown is an open problem [Albrecht and Stone, 2018], and requires algorithms that can identify the unmodeled states (x) of decision-making. The fact that humans may not behave optimally makes this problem further challenging [Kahneman, 2003].

Towards this problem, we have developed a novel approach that poses the problem of learning decision-making models as one of Bayesian inference [Unhelkar and Shah, 2018]. Briefly, we provide iAMM, a Bayesian nonparametric generative model, to describe the human's sequential decision-making behavior. By utilizing Bayesian nonparametric priors, iAMM models the number of unmodeled states x (i.e., state space), their dynamics (i.e., transition function) and the human's policy (π) as unknown random variables. By separately modeling known (s) and unknown (x) state features, the model can incorporate task-specific knowledge as probabilistic priors. In order to learn the unknown parameters of the decision-making model, we have designed an integrated sampling and querying approach that includes

- an MCMC sampler inspired by the beam sampling algorithm for HDP-HMM [Van Gael *et al.*, 2008] to generate candidate decision-making models, and
- an active querying approach to identify the true model from the candidate models generated by the sampler.

Our approach uses both observations of human decisions (in the form of time series of observable states s and actions x) and human input (in the form of queries posed to the decision maker) to learn the decision-making model. In evaluations using simulated domains, we observe that our integrated sampling and querying approach can identify the number of latent states and estimate model variables (transition function and policy). As the next steps, we intend to evaluate our approach with a human in the loop, improve scalability using function approximation techniques, and explore the development of human-in-the-loop inference algorithms that can speed up inference by judiciously utilizing human input.

4 Decision-Making for Communication

In several applications, such as disaster response, collaborative agents have to execute the task without an accurate model of their environment (transition function). Thus, in addition to learning latent states of their collaborators, agents need algorithms to reason about latent states of their environment. Effective communication between teammates is critical for coordinating actions and accomplishing shared goals, especially in unknown environments [Butchibabu *et al.*, 2016]. However, enabling effective communication in human-machine teams – despite the presence of a communication modality – is challenging, since communication incurs cost and communication decisions about unmodeled phenomena have to be made during execution [Unhelkar *et al.*, 2017]. Thus, to enable effective communication for human-machine teams operating in unknown environments, we have developed a novel algorithm called ConTaCT detailed in [Unhelkar and Shah, 2016]. By representing belief of their teammate and gauging the value of information, the algorithm enables agents to make execution-time communication decisions. Our evaluations show that ConTaCT can achieve comparable task performance despite making fewer communications as compared to a baseline motivated by prior art [Roth *et al.*, 2005]. This behavior is indicative of anticipatory information sharing in teams and desirable for human-machine teams, where excessive communication can be detrimental.

In conclusion, to become successful collaborators, agents operating in real environments require the capability to learn, reason about and communicate latent states of human-machine collaboration. This abstract summarizes our ongoing research to provide collaborative agents this capability.

References

[Albrecht and Stone, 2018] Stefano Albrecht and Peter Stone. Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artificial Intelligence*, 258:66 – 95, 2018.

[Butchibabu *et al.*, 2016] Abhizna Butchibabu, Christopher Sparano-Huiban, Liz Sonenberg, and Julie Shah. Implicit coordination strategies for effective team communication. *Human Factors*, 58(4):595–610, 2016.

[Jonker *et al.*, 2011] Catholijn M Jonker, M Birna Van Riemsdijk, and Bas Vermeulen. Shared mental models. In *COIN*, 2011.

[Kahneman, 2003] Daniel Kahneman. Maps of bounded rationality. *The American Economic Review*, 93(5):1449–1475, 2003.

[Roth *et al.*, 2005] Maayan Roth, Reid Simmons, and Manuela Veloso. Reasoning about joint beliefs for execution-time communication decisions. In *AAMAS*, 2005.

[Thomaz *et al.*, 2016] Andrea Thomaz, Guy Hoffman, and Maya Cakmak. Computational human-robot interaction. *Foundations and Trends in Robotics*, 2016.

[Unhelkar and Shah, 2016] Vaibhav V Unhelkar and Julie A Shah. ConTaCT: deciding to communicate during time-critical collaborative tasks in unknown deterministic domains. In *AAAI*, 2016.

[Unhelkar *et al.*, 2017] Vaibhav V Unhelkar, X Jessie Yang, and Julie A Shah. Challenges for Communication Decision-Making in Sequential Human-Robot Collaborative Tasks. In *R:SS workshop on HRI*, 2017.

[Unhelkar *et al.*, 2018] Vaibhav V Unhelkar*, Przemyslaw Lasota* and others. Human-Aware Robotic Assistant for Collaborative Assembly: Integrating Human Motion Prediction with Planning in Time. *IEEE Robotics and Automation Letters*, 2018.

[Unhelkar and Shah, 2018] Vaibhav V Unhelkar and Julie A Shah. Learning Models of Sequential Decision-Making without Complete State Specification using Bayesian Nonparametric Inference and Active Querying. CSAIL Technical report, MIT, 2018.

[Van Gael *et al.*, 2008] Jurgen Van Gael, Yunus Saatci, Yee Whye Teh, and Zoubin Ghahramani. Beam sampling for the infinite hidden Markov model. In *ICML*, 2008.

[Ziebart *et al.*, 2008] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum Entropy Inverse Reinforcement Learning. In *AAAI*, 2008.