Coordinating and Incorporating Trust in Teams of Humans and Robots with Multi-Robot Reinforcement Learning

Christopher Amato & Stacy Marsella

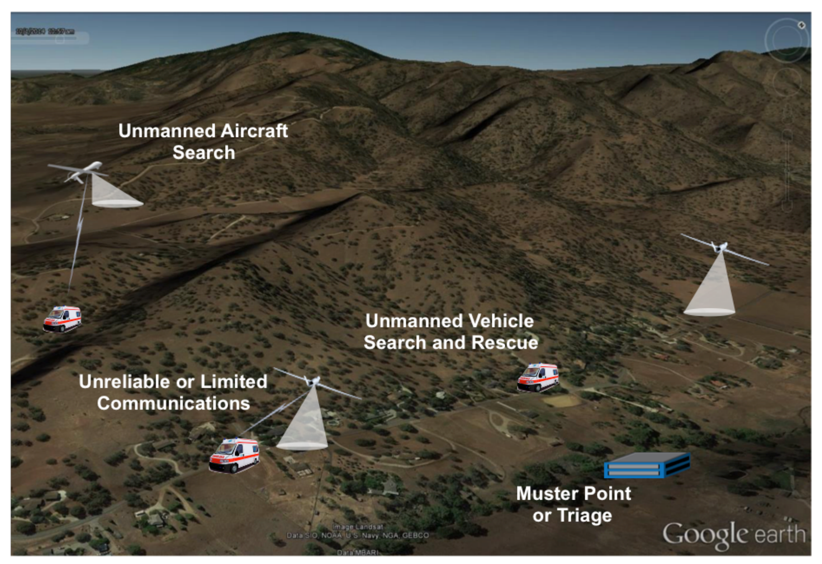
Northeastern University

Project overview

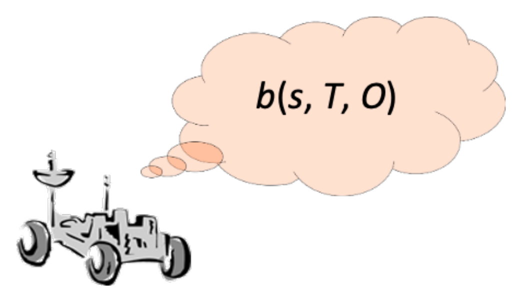
Goal: How can teams of robots learn to collaborate with humans given the partial observability and uncertainty in human interactions as well as the vast differences in reasoning between robots and humans?

We plan to develop:

1. Teams of robots **learning** to assist humans even with **incorrect and incomplete human models**
2. Teams of robots learning to **coordinate and interact with humans using shared mental models**
3. Teams of robots learning to coordinate and interact with humans by **incorporating trust**
4. **Test in simulation and hardware** in (simulated) search and rescue and other scenarios



Modeling and Learning Human Models

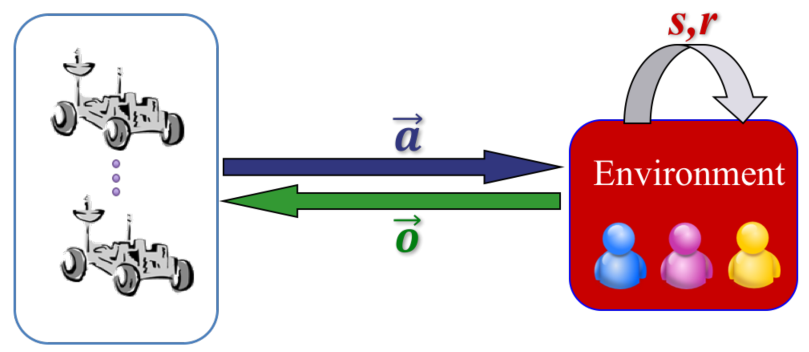


Overview

- Generate initial POMDP-based human models from simulation data
- Use these models in model-based (partially observable) reinforcement learning (RL)
- Improve models and solutions using Bayesian RL

POMDPs

- S, a set of states
- A, a set of actions
- T, the state transition model: $\Pr(s'|s,a)$
- R, the reward model: $R(s,a)$
- O, a set of observations
- Z, the observation model: $\Pr(o|s',a)$

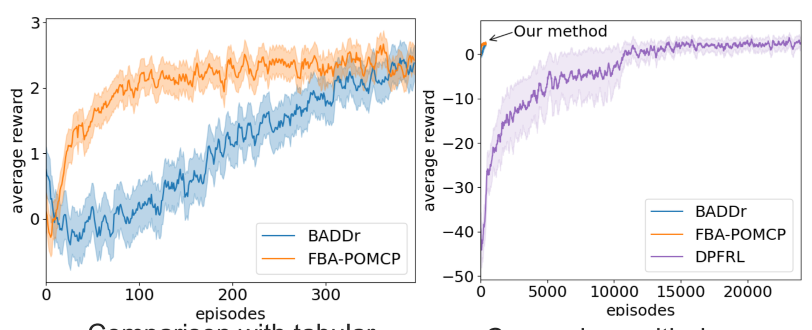


Generating initial model

- Will explore methods such as Automated Cognitive Behavior Analysis for generating from data with humans as part of the POMDP environment (based on [4])

Efficient Bayesian RL for POMDPs [1]

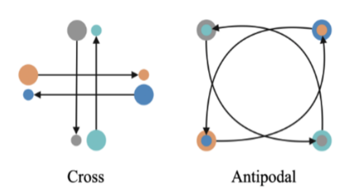
- Bayesian RL can optimally balance exploration and exploitation
- Ideal for online learning—optimally sample efficient so can learn from very few interactions with humans
- Can be computationally challenging, but building on our work combining deep RL with Bayesian RL to improve scalability [1] and exploiting structure in our human interaction domains
- Our approach, Bayes-adaptive deep dropout RL (**BADDr**), approximates the possible POMDP models with neural networks and approximates the Bayesian update with dropout in an ensemble of models
  - Slower than tabular Bayesian RL in small domains
  - But tabular Bayesian RL doesn't scale
  - Much faster than deep RL

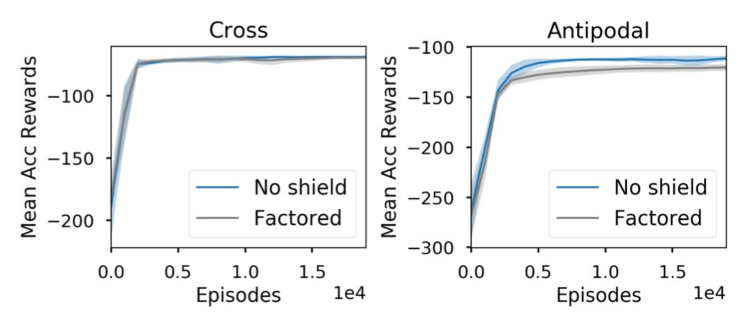


Comparison with tabular Comparison with deep

Safe multi-agent reinforcement learning [3]

- (Multi-agent) RL is popular, but there are no safety guarantees
- We guarantee safety properties (e.g., no collisions) during training & execution with a shield
- Formalize safety using LTL
- Scalable by developing factored shields over problem
- Can be included into any MARL method
- Results show we can generate near-optimal solutions without unsafe actions (collisions)



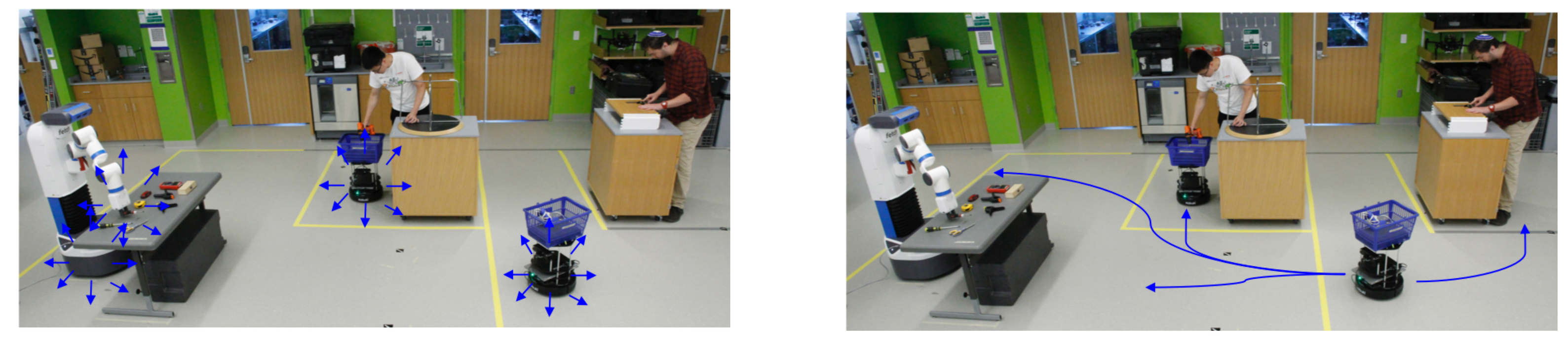


	MADDPG	MADDPG with Shield
Cross	207.20	**0.00**
Antipodal	14,419.20	**0.00**

Number of collisions

Shared Mental Models

- Provide humans and robots with a shared model of the task as well as models of each other [6] by which to interpret and coordinate behavior
- Uses a hierarchical model over shared key coordination points
- Allows flexibility the low level but coordination and communication at the shared high level
- Extend Bayesian RL methods above to incorporate hierarchies [from 2]
- Will be more scalable since they learn over fewer, higher-level actions and enable consistent mental models across the robots and people along with natural communication about those mental models.



Incorporating Trust and Interpretability

- True cooperation requires humans and robots to see each other as trusted teammates
- Therefore, we will incorporate trust into our POMDP models of human interaction
- And develop methods for improving trust by generating and sharing more interpretable robot solutions

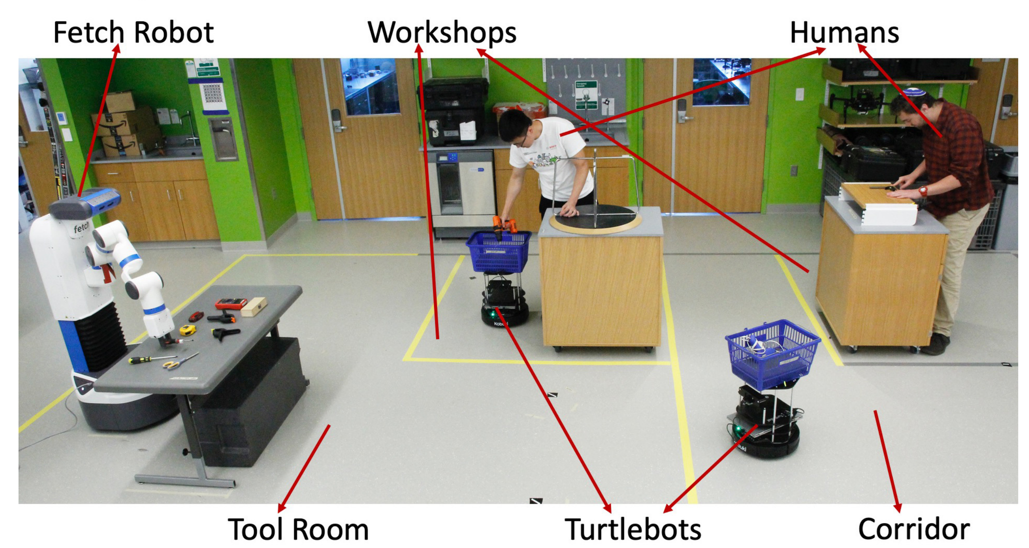
Trust: Want to allow human teammates to appropriately judge the trustworthiness of robots using approaches based on the human response data and using Appraisal Theory to inform the design of trustworthy interaction and communication [5]

Interpretability: Interpretability in RL for HRI is important but general approaches don't currently exist. We will develop high-level interpretable representations of our Bayesian RL methods and other solutions along with uncertainty estimates.

Experimental domains

Evaluate our methods in simulated scenarios and then hardware

- Simulated scenarios to test our ideas
  - Search and rescue scenarios in Minecraft and coordinated resource gathering task
  - Much easier to gather data and test methods in simulation
- Will also test ideas in standard partially observable benchmarks
- Test in hardware using simple manufacturing setting (below) as well as more complex (simulated) search and rescue domain with humans teaming with aerial and ground robots



Current domains that will be used for testing (Minecraft and manufacturing)

[1] **BADDr: Bayes-adaptive Deep Dropout RL for POMDPs.** Sammie Katt, Frans A. Oliehoek and Christopher Amato. *In the Proceedings of the Eighteenth International Conference on Autonomous Agents and Multi-Agent System* (AAMAS-22), May 2022.
[2] **Macro-Action-Based Deep Multi-Agent Reinforcement Learning.** Yuchen Xiao, Joshua Hoffman and Christopher Amato. In *Proceedings of the 2019 Conference on Robot Learning* (CoRL-19), October 2019.
[3] **Safe Multi-Agent Reinforcement Learning via Shielding.** Ingy Elsayed-Aly, Suda Bharadwaj, Christopher Amato, Rudiger Ehlers, Ufuk Topcu and Lu Feng. In the *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems* (AAMAS-21), May 2021.
[4] **Analyzing Human Negotiation using Automated Cognitive Behavior Analysis: The Effect of Personality.** Pedro Sequeira and Stacy Marsella. In *Proceedings of Cognitive Science*, 2018.
[5] **Contextually-Based Utility: An Appraisal-Based Approach at Modeling Framing and Decisions.** Jonathan Ito and Stacy Marsella. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence* (AAAI), San Francisco, 2011.
[6] **Reinforcement Learning for Adaptive Theory of Mind in the Sigma Cognitive Architecture.** David V. Pynadath, Paul S. Rosenbloom, and Stacy C. Marsella. In *Artificial General Intelligence*, AGI, 2014.