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

Christopher Amato & Stacy Marsella

Northeastern University

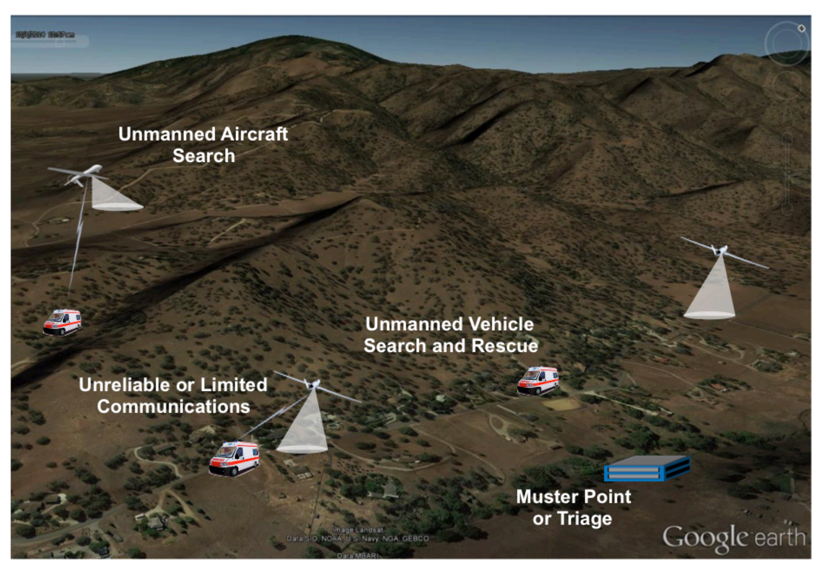
Project overview

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?

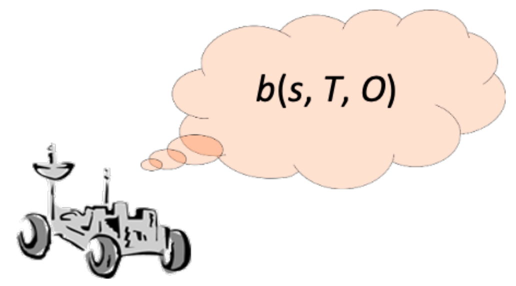
We plan to develop:

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

Project just began so this is our plan for the next three years.



Modeling and Learning Human Models

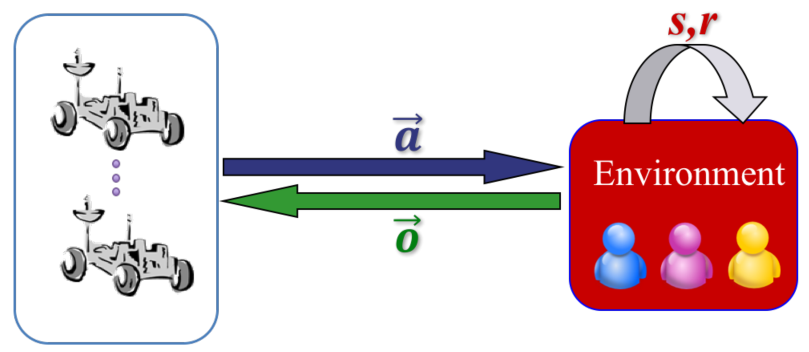


Overview

- 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])

Bayesian RL for POMDPs

- Bayesian RL can optimally balance exploration and exploitation
- Ideal for online learning—optimally sample efficient so can learn from very few interactions with humans

Scalable solution methods

- Can be computationally challenging, but building on our work combining deep RL with Bayesian RL to improve scalability [1,2] and exploiting structure in our human interaction domains

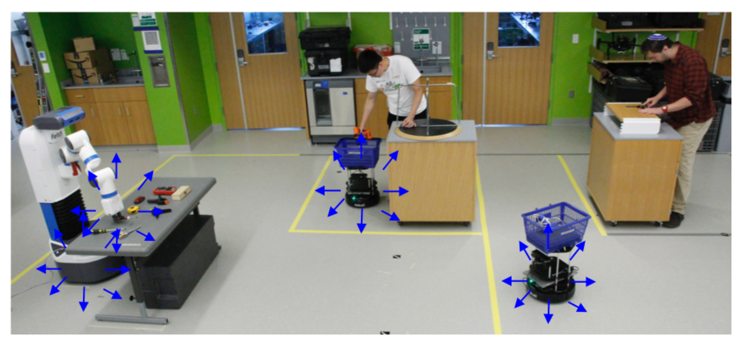
Shared Mental Models

Overview

- 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
- 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.

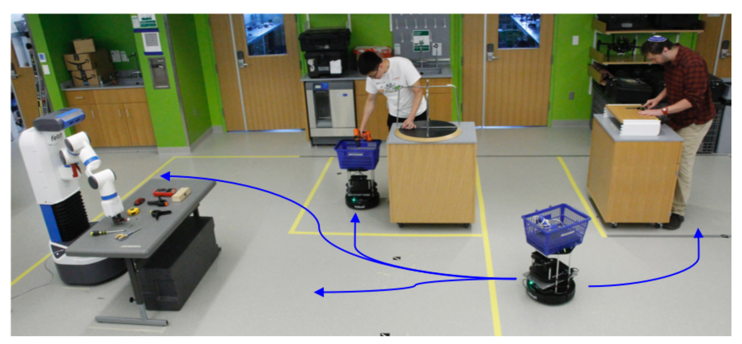
Hierarchies of macro-actions

- Robots will execute at different task levels (e.g., high-level *macro-actions* and low-level controls)
- This is result in asynchronous execution for the robots
- Macro-action goals will be defined with humans
- Robots (and humans) will communicate progress toward goals



Learning with macro-actions

- Some of our previous work defined macro-actions and developed deep RL methods for learning over them [3]
- But current methods are model-free and non-Bayesian
- We'll develop the first hierarchical Bayesian deep reinforcement learning methods for POMDPs



Incorporating Trust and Interpretability

Overview

- 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
- We intend to explore multiple approaches to incorporating explicit trust goals into the design
- One is largely driven by analysis of the human response data from previous experiments,
- One takes a theory driven perspective using Appraisal Theory to inform the design of trustworthy interaction and communication [5]
- Finally, a learning approach that incorporates the achievement and maintenance of trust between humans and robots as part of the robots' objective functions, so that in essence the robot learns to become more trustworthy

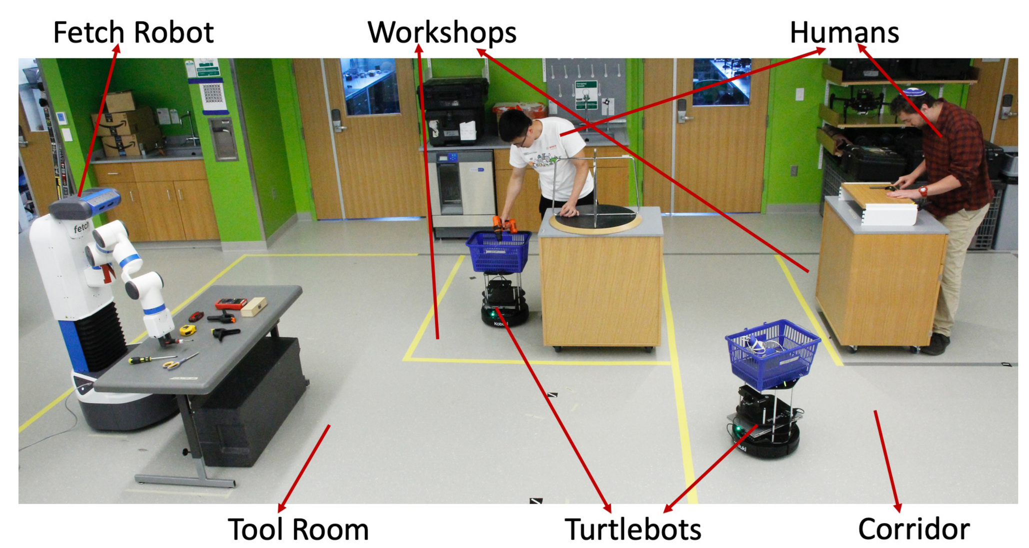
Interpretability

- Interpretability in RL for HRI is important but general approaches don't currently exist
- We will start by generating high-level finite-state controller solutions from our Bayesian RL methods that are updated with the currently executing macro-action
- Generate improved representations that are more helpful to humans
- Update solutions (shared with appropriate communication) as learning progresses
- Also will provide uncertainty estimates and robot beliefs to humans given by Bayesian RL

Experimental domains

Evaluate our methods in Minecraft search and rescue scenarios and then hardware

- Develop search and rescue scenarios in Minecraft to test our ideas
- 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] **Bayesian Reinforcement Learning in Factored 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-19), May 2019

[2] **BADDr: Bayes-adaptive Deep Dropout RL for POMDPs.** Sammie Katt, Frans A. Oliehoek and Christopher Amato. *In review*

[3] **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.

[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.