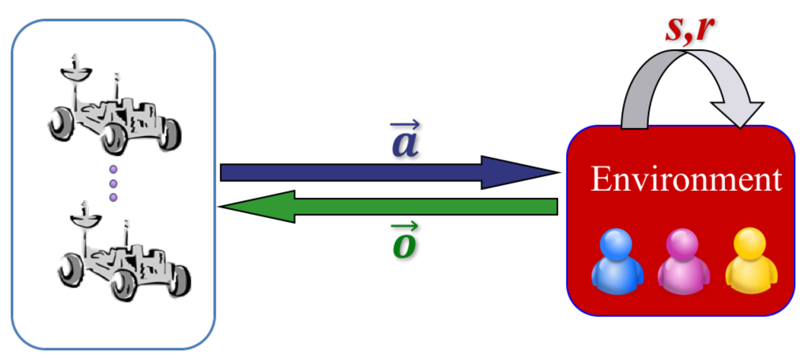
Coordinating Human-Robot Teams in Uncertain Environments

Christopher Amato, *Northeastern University* & **Laurel D. Riek**, *UC San Diego*

Project overview

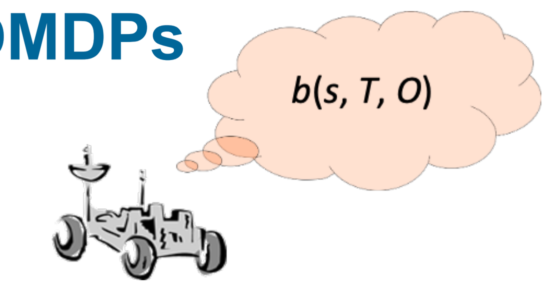
Goal: create and solve realistic models for coordinating teams of humans and robots in uncertain environments

1. Re-conceptualize multi-human teamwork that include dynamic, stochastic environments
   - See **Intention Modelling for Teaming under Uncertainty** [4][5]
2. Develop realistic (POMDP) models of human-robot teamwork with uncertainty and partial observability
   - In progress



3. Create scalable techniques for planning and learning in these models
   - See **Bayesian Reinforcement Learning (RL) for POMDPs** [2] and **Belief-Grounded Networks for Accelerated Robot Learning under Partial Observability** [3]
4. Test in simulation and emergency department (ED) settings
   - See **Situating Robots in the Emergency Department** [6][7]

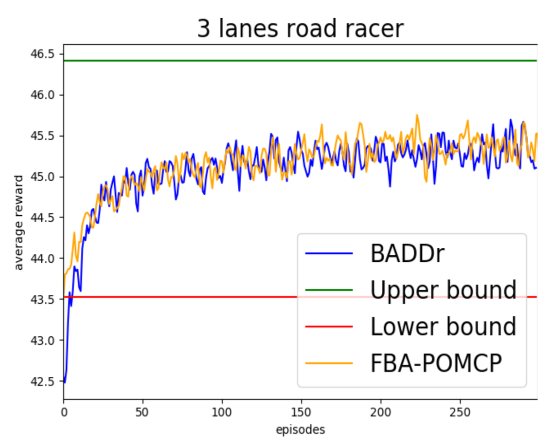
Bayesian Reinforcement Learning (RL) for POMDPs



Developed scalable deep Bayesian RL methods for POMDPs [2]

- Bayesian RL can optimally balance exploration and exploitation
- Ideal for online learning—optimally sample efficient!
- Can be computationally challenging, but combined with deep reinforcement learning to improve scalability
- These methods outperform previous methods, allowing learning in large POMDPs

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: $\Pr(o|s',a)$



Small problem/similar performance

Bayesian RL for POMDPs (e.g., [Ross et al. JMLR 11])

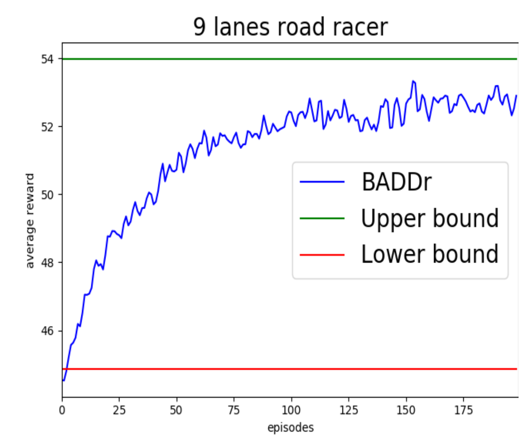
- Explicitly consider uncertainty over possible trans. and obs. models
- Can start with prior over models and update based on observations
- Can now have belief over state *and models*

Scalable solution methods

- Combined scalability of deep RL with sample efficiency of Bayesian RL [2]
- Sampling method for scalable particle filtering and planning [2]

Results

- Can learn similar to tabular methods [1] on small problems
- Can now solve significantly larger problems

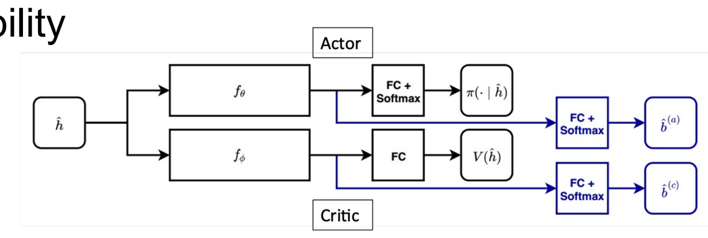


Big problem/previous methods can't solve

Belief-Grounded Networks for Accelerated Robot Learning under Partial Observability [3]

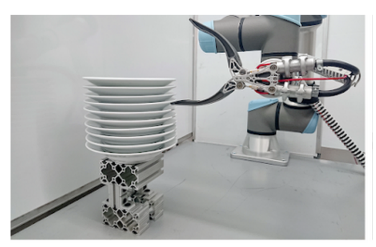
Problem

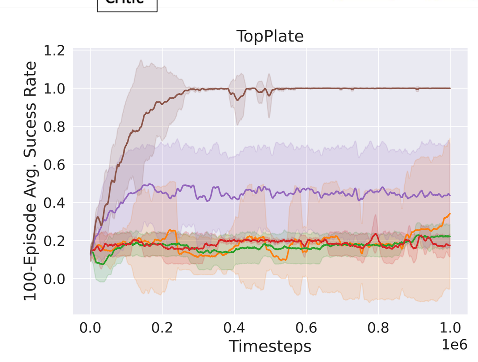
- Many tasks can be simulated and have privileged information (true states, models) during training
- However, when being deployed the agent only has partial observability
- How to leverage this privileged information during training?



Method

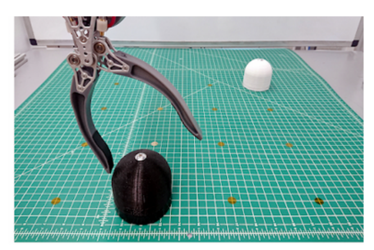
- Use ground-truth beliefs to improve the history feature extractors --> better features --> better action selection
- Applied to an advantage actor-critic (A2C) agent
- During training, add two branches (blue) to a standard A2C agent to reconstruct the beliefs from the history
- During deployment, only the actor network is used

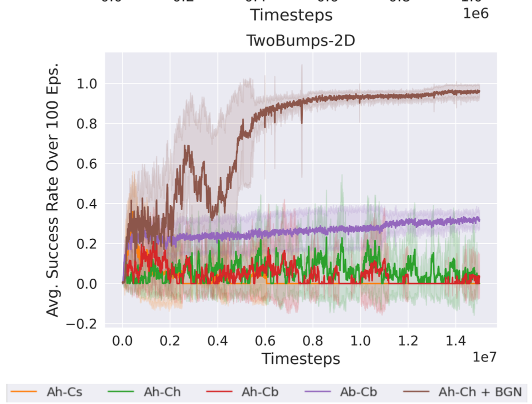




Results

- The proposed method (brown line, +BGN) outperforms all baselines (even the one using ground-truth beliefs)
- Policies learned in simulation can be transferred zero-shot to work w/ real hardware

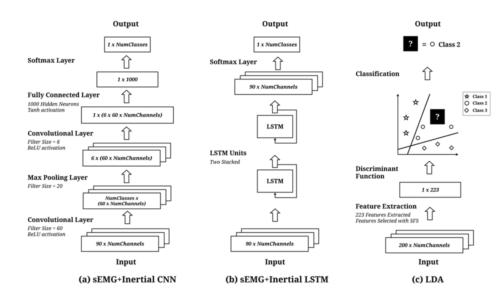


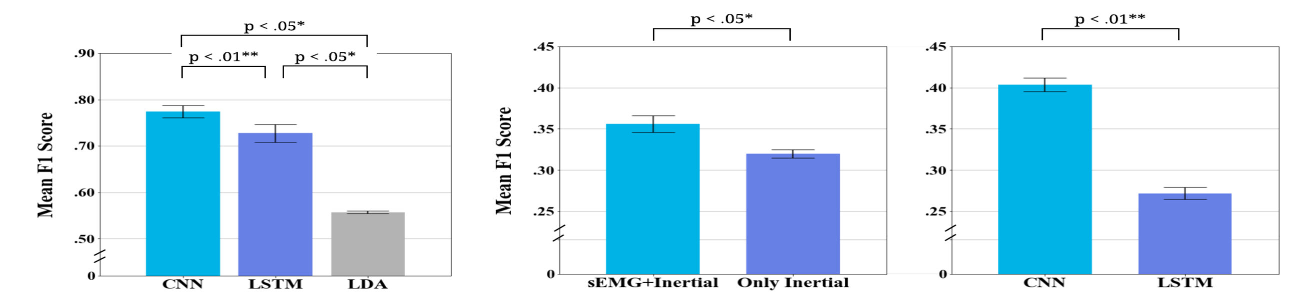


Intention Modelling for Teaming under Uncertainty

Created new deep learning methods for non-visual activity modelling [4]

- Can detect both fine and gross motor movements, is immune to occlusion and avoids privacy concerns of visual sensing, augments linear and angular velocity w/ muscle activity data from wearable (Myo)
- Our methods outperformed the state-of-the-art classifiers by 28%, sEMG+Inertial yielded significantly higher classification accuracy than inertial alone
- Wearables are well-suited to activity recognition in uncertain environments



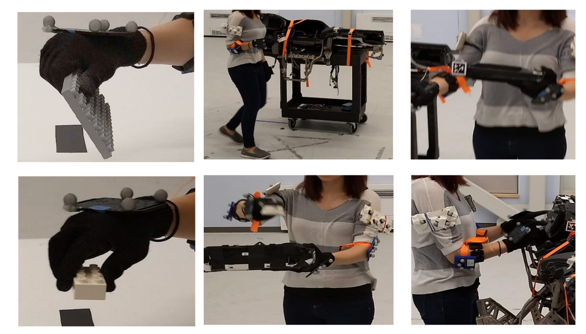


Designed new approaches for multimodal contextualized activity recognition [5]

- Created new multimodal dataset of gross and fine motor tasks (EMG/Inertial/MoCap), compared multiple activity recognition approaches for recognition suitability, employed early fusion
- Results suggest complementary strengths of each sensor type – task type should be taken into account when engaging in sensor selection

MEAN F1 SCORES OBTAINED FOR EACH DATA MODALITY ON EACH DATASET USING DIFFERENT CLASSIFIERS. ACROSS THE DATASETS AND SENSORS, WE AVERAGED THE F1 SCORES FROM EVERY TRIAL. A HIGHER F1 SCORE IS BETTER.

	SVM			LDA			KNN		
	Vicon	Myo	Vicon+Myo	Vicon	Myo	Vicon+Myo	Vicon	Myo	Vicon+Myo
Automotive (Gross motion)	**.79**	.42	.43	**.76**	.48	.49	**.88**	.58	.59
Block (Fine-grained motion)	.09	**.37**	.36	.23	**.39**	.36	.32	**.43**	.43

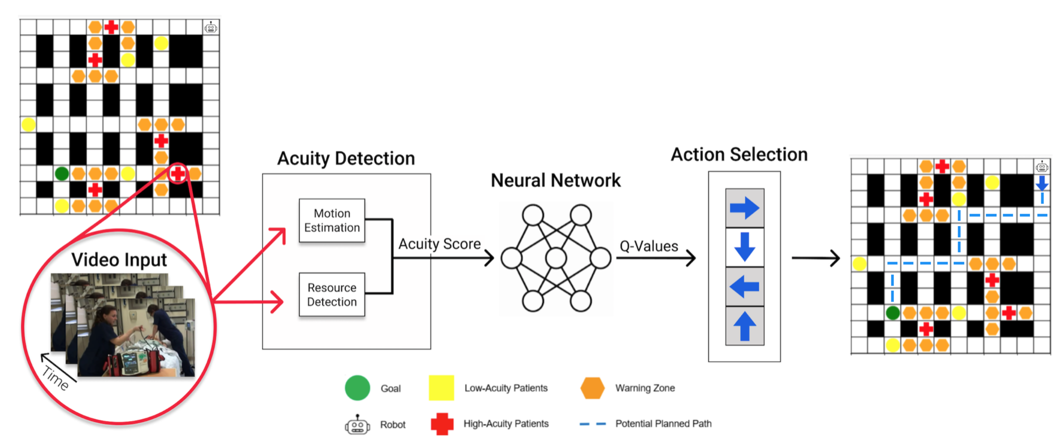


Situating Robots in the Emergency Department [6,7]

Problem: The ED is an uncertain environment in which mistakes can be deadly and providers are over burdened.; well-designed & contextualized robots could relieve providers of non-value added tasks and enable them to spend more time on patient care. e.g., delivery robots.

Approach: We used domain knowledge to characterize staff workflow and patient experience, identify key considerations for robots in the ED, inc.: safety, physical and behavioural attributes, usability, and training. We introduced a task representation [5] and new acuity-aware social navigation algorithm [7] which incorporates both patient criticality and staff workflow.

Results: We introduced the Safety-Critical Deep Q-Network (SafeDQN) system [7], a new acuity-aware navigation system for mobile robots. SafeDQN is based on two insights about care in EDs: high-acuity patients tend to have more HCWs in attendance and those HCWs tend to move more quickly. We compared SafeDQN to three classic navigation methods, and show that it generates the safest, quickest path for mobile robots when navigating in a simulated ED environment.



Method	Avg. Path Length ↓		Avg. Reward ↑		Avg. HA Penalties ↓		Avg. LA Penalties ↓	
	OF	KD	OF	KD	OF	KD	OF	KD
Random Walk	243.6	231.0	-54.8	-53.3	5.9	5.6	15.4	25.0
A*	12.6	11.7	-2.4	-2.5	0.1	0.2	0.4	0.9
Dijkstra	11.6	10.4	-2.4	-2.1	0.1	0.3	**0.2**	**0.4**
SafeDQN	**11.3**	**9.4**	**-0.6**	**-0.6**	**0**	**0.1**	0.7	0.7

[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] **Belief-Grounded Networks for Accelerated Robot Learning under Partial Observability.** Hai Nguyen, Brett Daley, Xinchao Song, Christopher Amato and Robert Platt. In the *Proceedings of the Conference on Robot Learning* (CoRL-20), November 2020.
[4] **Frank, A., Kubota, A., and Riek, L.D. (2019). "Wearable activity recognition for robust human-robot teaming in safety-critical environments via hybrid neural networks".** Andrea Frank, Alyssa Kubota, Laurel D. Riek. *IEEE International Conference Intelligent Robots and Systems* (IROS-19), 2019
[5] **Activity recognition in manufacturing: The roles of motion capture and sEMG+inertial wearables in detecting fine vs. gross motion.** Alyssa Kubota, Tariq Iqbal, Julie A. Shah, and Laurel D. Riek. *IEEE International Conference on Robotics and Automation* (ICRA-19), 2019.
[6] **Situating Robots in the Emergency Department.** Angelique Taylor, Sachiko Matsumoto, and Laurel D. Riek. *AAAI Spring Symposium on Applied AI in Healthcare: Safety, Community, and the Environment* (AAAI-20) , 2020
[7] **Social Navigation for Mobile Robots in the Emergency Department.** Angelique Taylor, Sachiko Matsumoto, Wesley Xiao, and Laurel D. Riek. *IEEE International Conference on Robotics and Automatino* (ICRA-21), 2021