Online Inverse Reinforcement Learning under Occlusion

NRI:FND:Robust Inverse Learning for Human-Robot Collaboration

Award: 1830421, 08/18/2018

Prashant Doshi & Yi Hong – University of Georgia, Kenneth Bogert – UNC Asheville

Challenge

- Expert demonstrations are continuous, not arriving in batches

Solution

- A framework for online IRL called **I2RL**
- New I2RL method: Online Latent MaxEnt IRL using gradient descent
- Theory: Bound feature expectations to ensure convergence

A learner robot observes a continuous patrol of a hallway by two other robots. By using an online-inverse learning approach, the learner can attempt to penetrate the patrol as soon as it is confident in its learned model instead of after a pre-defined amount of observations

Scientific Impact

- Convergence and Regret guarantees in an *online* setting

Broader Impact

- Expand IRL applications to those with continuous and cyclical demonstrations
- A critical step towards the creation of autonomous robot learners

IRL in Partially Controlled Settings

NRI:FND:Robust Inverse Learning for Human-Robot Collaboration

Award: 1830421, 08/18/2018

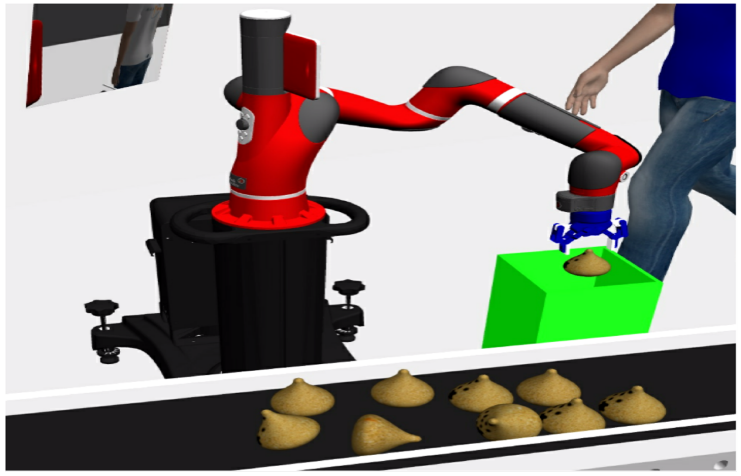
Prashant Doshi & Yi Hong – University of Georgia, Kenneth Bogert – UNC Asheville

Challenge

- IRL outside the lab must contend with persistent sources of noise and confounding elements

Solution

- New method: Hierarchical Bayesian model of percepts during demo – solve with LME-IRL
- Learn mixture model of percept sources – from expert or confounding elements



A robot arm sorts onions while a person walks in the background. The learning agent is confused by the person's shirt being the same color as the gripper it is trying to track. Our approach is to treat all such percepts received as a mixture of the expert and confounding elements present in the environment.

Scientific Impact

- Relax unrealistic assumption of knowing the expert's true trajectory
- Exploit indirect data sources toward IRL

Broader Impacts

- Reduce the deployment time of IRL & apprenticeship applications
- Expand IRL applicability to many real-world settings where the environment can't be completely controlled

Maximum Entropy Multi-task IRL

NRI:FND:Robust Inverse Learning for Human-Robot Collaboration

Award: 1830421, 08/18/2018

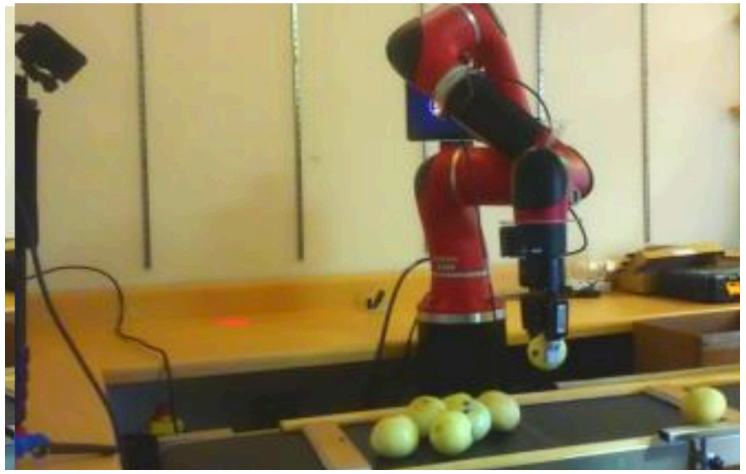
Prashant Doshi & Yi Hong – University of Georgia, Kenneth Bogert – UNC Asheville

Challenge

- Expert may demonstrate multiple interleaved approaches to the same problem or perform multiple tasks

Solution

- Assume each trajectory comes from a specific but unknown cluster – each cluster represents a reward function
- New method: Combine Dirichlet process based minimum entropy clustering with MaxEnt IRL



A robot arm sorts onions, however, it may also roll onions with its gripper to change the visible portion. Both these tasks relate to sorting out blemished onions and are interleaved during the overall demonstrations. They are modeled with separate MDPs with unknown rewards to be learned with IRL

Scientific Impact

- Generalizes well-known MaxEnt IRL to multi-task environments with unknown task labels

Broader Impact

- Reduce the demonstration burden on the expert as multi-modal or multiple technique demonstrations are learnable
- Expands applicability of IRL to more natural task performance