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

Prashant Doshi (PI), Computer Science, University of Georgia
Yi Hong (Co-PI), Computer Science, University of Georgia
Kenneth Bogert (Co-PI), Computer Science, University of North Carolina at Asheville

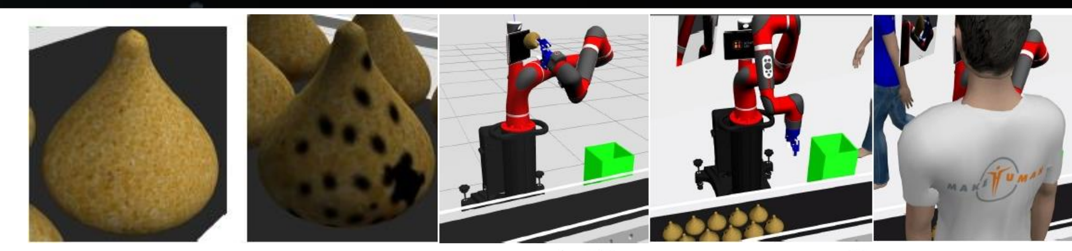
Project URL: http://thinc.cs.uga.edu

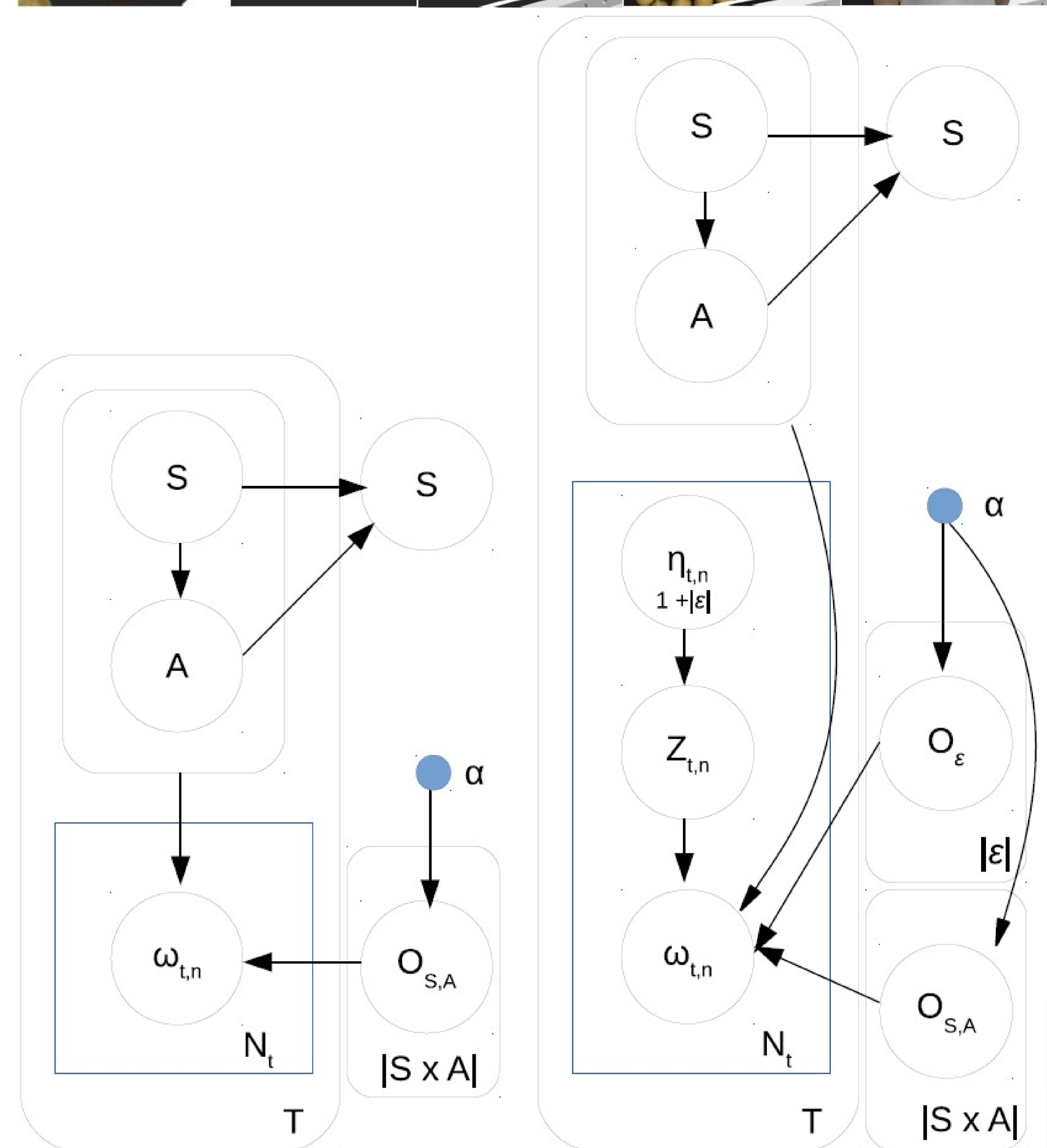
Challenges
- Co-bots working with humans trained using passive demonstrations in real-world scenarios
- The demonstrations by human experts have confounding elements, occlusion, and occur in uncontrolled settings
- The human experts may interleave different ways of solving a given task exhibiting differing preferences

Approach
- Hierarchical Bayes model for IRL in partially-controlled settings
- Develop online learning, incremental learning IRL algorithms
- Maximum entropy multi-task IRL algorithm

Scientific Impacts
- Relax knowing the expert's true trajectory due to perception noise. Exploit indirect data sources toward IRL
- A framework and new method for online IRL
- Generalizes well-known MaxEnt IRL to multi-task environments with unknown task labels

Broader Impacts
- Reduce the deployment time of apprentice robots
- Expand IRL applicability to more natural and continuous task demonstrations in environments that cannot be controlled



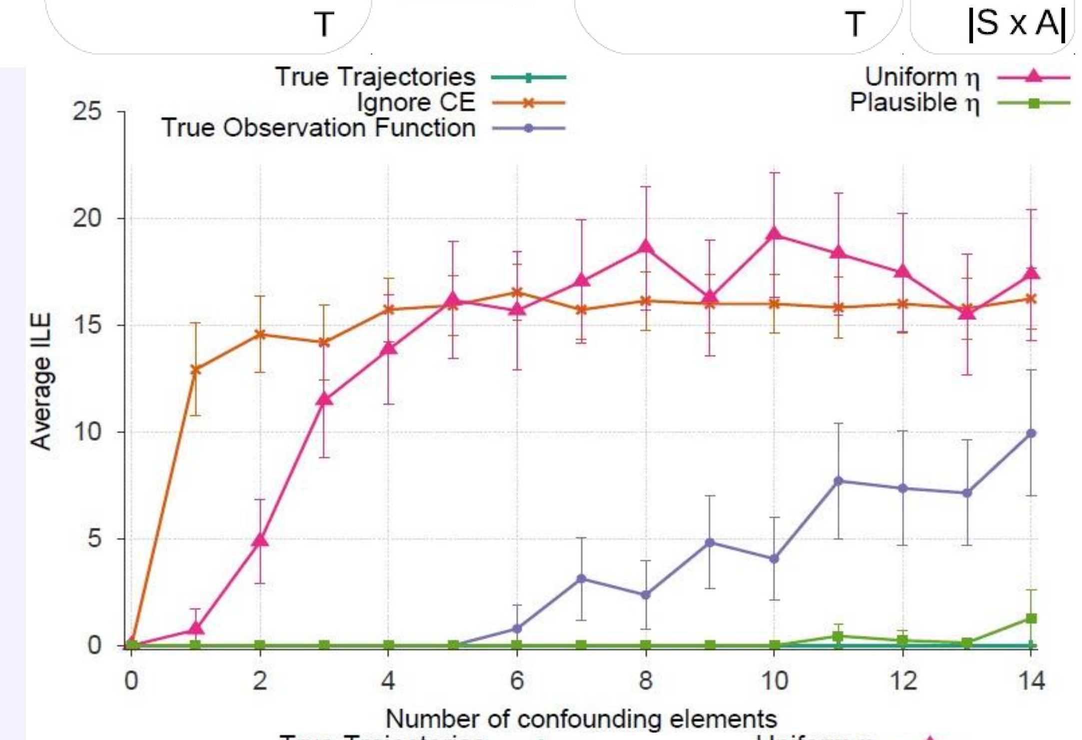


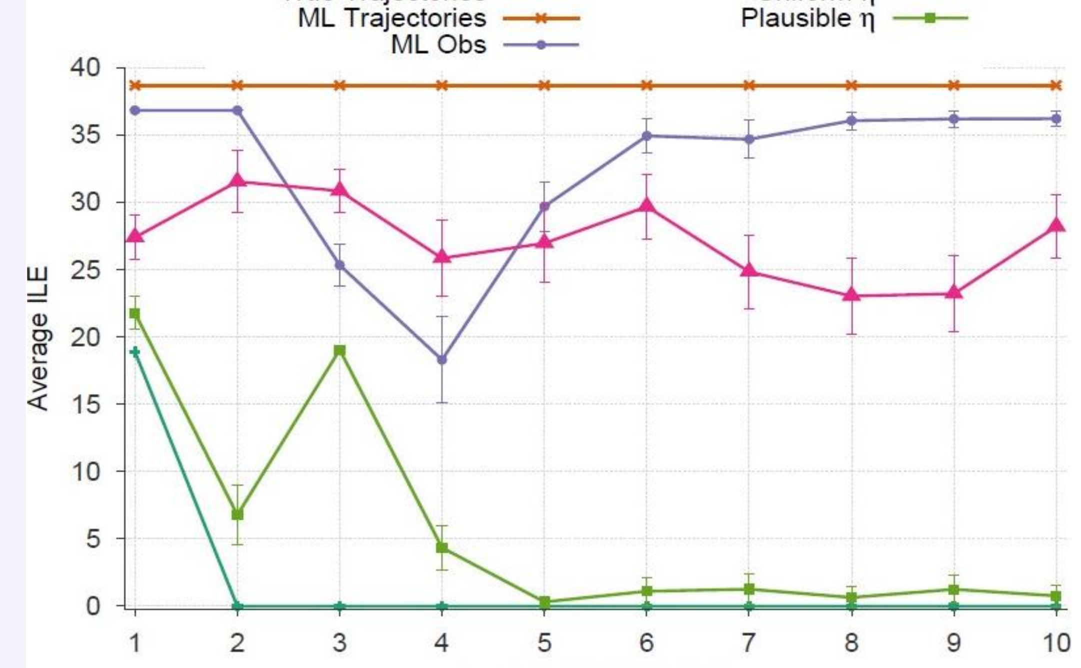
Inverse RL in Partially-Controlled Settings

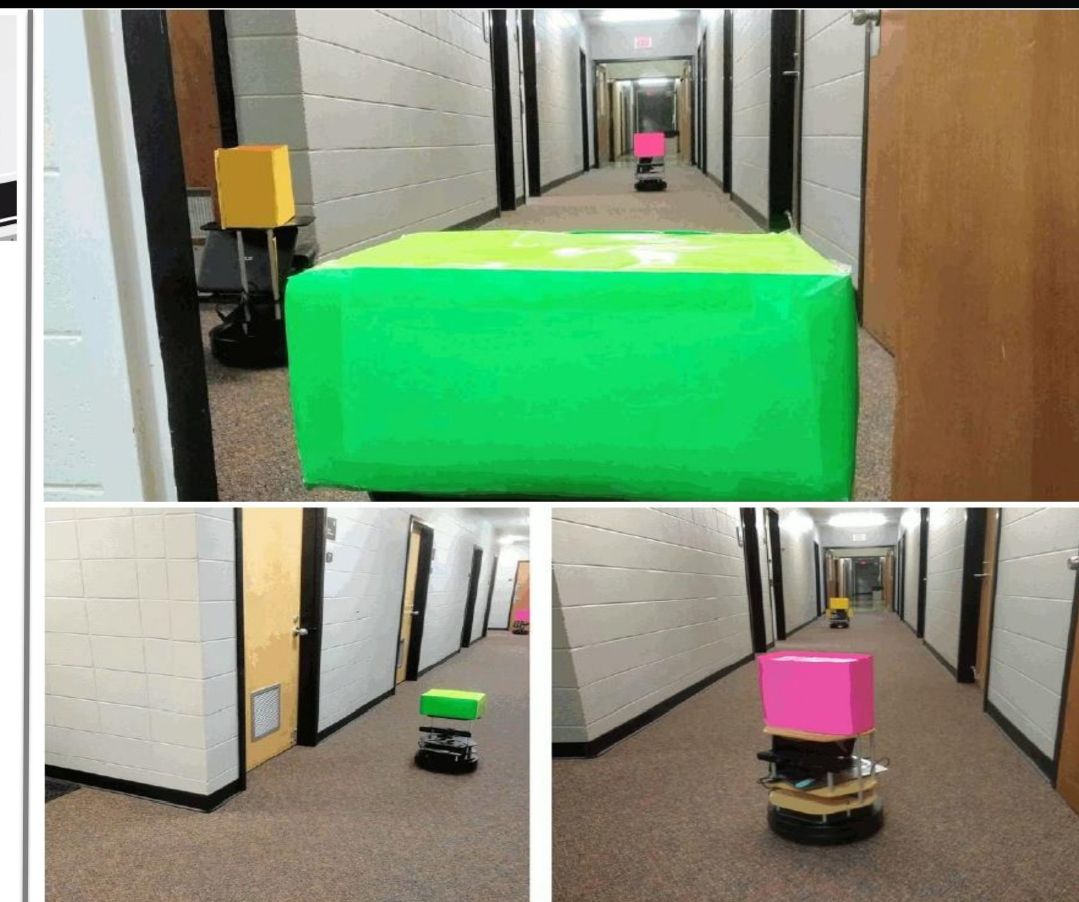
- A general hierarchical Bayesian model that captures the uncertainty present in observations during expert agent demonstrations in partially controlled, real-world situations to perform apprenticeship learning.
- Generalize the E-step of HiddenDataEM to environmental and sensing noise to compute a distribution over possible demonstrations X given the observations.
- Deal with partially controlled environments by finding the expert's trajectory distribution and exploiting indirect observations.

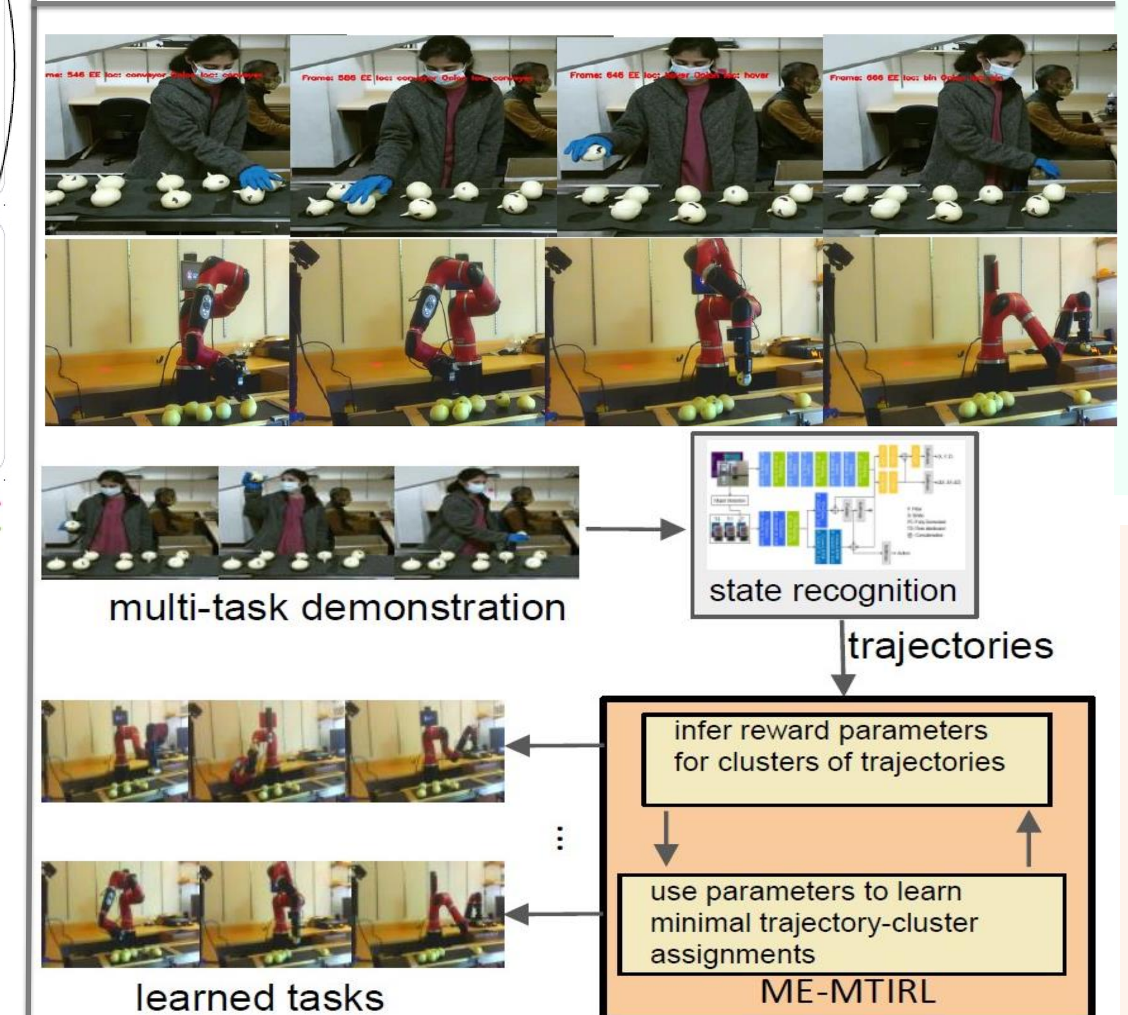
Experiment 1:
- Formative evaluation on Gridworld with added confounding elements.

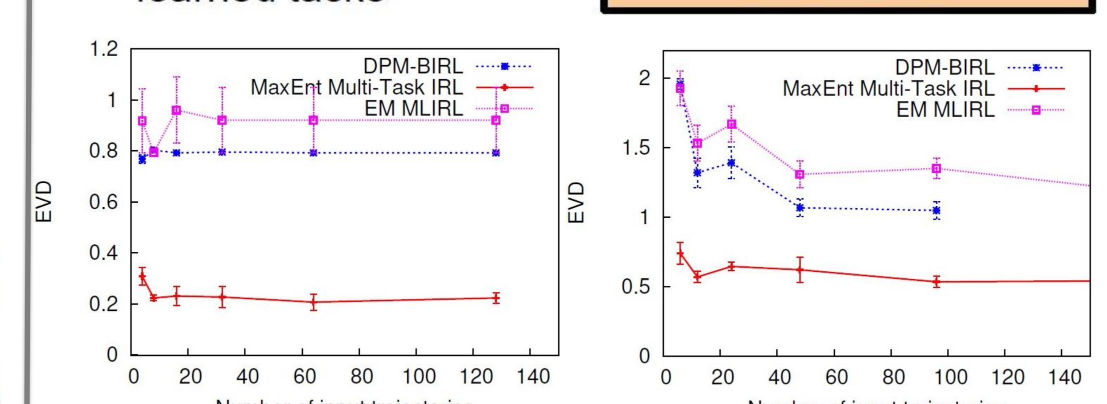
Experiment 2:
- Summative assessment using Robotic Onion Sorting
- The Sawyer robot arm is tasked with inspecting onions moving down a conveyor belt and sorting good onions from blemished ones.











		(TP,FP,FN,TN)	P%,R%
Expert	P-I-P R-P-P	(7,0,12,5) (9,4,8,3)	100.00, 58.33 69.23, 75.00
Learned (ME-MTIRL)	P-I-P R-P-P	(7,1,11,5) (9,4,8,3)	87.50, 58.33 69.23, 75.00
Learned (DPM-BIRL)	P-I-P R-P-P	(7,3,9,5) (9,4,8,3)	70.00, 58.33 69.23, 75.00
Learned (EM-MLIRL)	P-I-P R-P-P	(6,3,9,6) (6,4,8,6)	66.67, 50.00 60.00, 50.00

I2RL: Online IRL Under Occlusion

- Introduce I2RL , framework for online IRL, which establishes the key components and offers candidate stopping criteria.
- Establish Key theoretical method called online latent maximum entropy IRL.
- It offers the capability to perform online IRL in contexts where portions of the observed trajectory may be occluded.

Task:
- Two independent mobile robots patrolling corridors of varying configurations. A third robot observes the patrollers and uses learned reward functions to predict patrolling trajectories and identify a path to the goal location.

Experiments:
- Done using physical robots (Turtlebots) to test how well I2RL can be implemented and extended in real world.
- The learner robot can only accesses to less than 30% observability. Compared to online GAIL, LME I2RL performs significantly better and achieves a reasonable result.

Maximum Entropy Multi-Task Inverse RL

- An expert may solve a problem in multiple distinct ways each of which optimizes a different reward function while still sharing the features.
- Combine MaxEnt IRL with the Dirichlet process based minimum entropy clustering of the observed data.
- Yield a single nonlinear optimization problem called MaxEnt Multi-task IRL, which can be solved using Lagrangian relaxation and gradient descent methods.

Task:
- Human experts will teach robot learner about how to distinguish blemished and unblemished onion and pick up blemished onion by its gripper.

Experiment:
- Use expected value difference averaged over the trajectories and Precision and Recall as metrics.
- Use YOLO as objects recognition tools.
- Two distinct sorting methods:
  - pick-inspect-place and roll-pick-place

2021 NRI & FRR Principal Investigators' Meeting
March 10-12, 2021

Award ID#: : IIS 1830421