CSE 571
Inverse Optimal Control
(Inverse Reinforcement Learning)

Many slides by Drew Bagnell
Carnegie Mellon University

Autonomous Navigation

UPI
Somerset Woods
May, 2007

Optimal Control Solution

Cost Map

2-D Planner

Y
(Path to goal)
Mode 1: Training example

Mode 1: Learned behavior

Mode 1: Training example

Mode 1: Learned behavior
Mode 1: Learned cost map

Mode 2: Training example

Mode 2: Training example

Mode 2: Learned behavior
Mode 2: Learned behavior

Ratliff, Bagnell, Zinkevich 2005
Ratliff, Bradley, Bagnell, Chestnutt, NIPS 2006
Silver, Bagnell, Stentz, RSS 2008

Mode 2: Learned cost map

Ratliff, Bagnell, Zinkevich, ICML 2006
Ratliff, Bradley, Bagnell, Chestnutt, NIPS 2006
Silver, Bagnell, Stentz, RSS 2008
Learned Cost Function Examples

Ratliff, Bradley, Chesnutt, Bagnell 06
Zucker, Ratliff, Stolle, Chesnutt, Bagnell, Adkeson, Kuffner 09
21. **Learned Cost Function Examples**

22. **Pedestrian Trajectory Prediction**

23. **Staying out of People’s Path**

24. **Learning Manipulation Preferences**

- **Input:** Human demonstrations of preferred behavior (e.g., moving a cup of water upright without spilling)
- **Output:** Learned cost function that results in trajectories satisfying user preferences
Demonstration(s)

Graph

Projection
Setup

- **Binary** state-dependent features (~95)
  - Histograms of distances to objects
  - Histograms of end-effector orientation
  - Object specific features (electronic vs non-electronic)
  - Approach direction w.r.t goal

- **Task**
  - Hold cup upright while not moving above electronics

Laptop task: Demonstration

(Not part of training set)
Laptop task: LTO + Discrete graph path

Laptop task: LTO + Smooth random path

Readings

- Imitation learning (Ermon): [https://cs.stanford.edu/~ermon/](https://cs.stanford.edu/~ermon/)
- Human/manipulation (Dragan): [https://people.eecs.berkeley.edu/~anca/research.html](https://people.eecs.berkeley.edu/~anca/research.html)