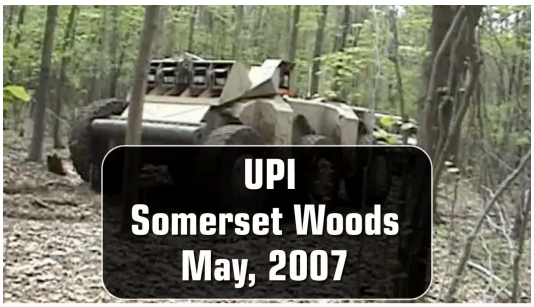
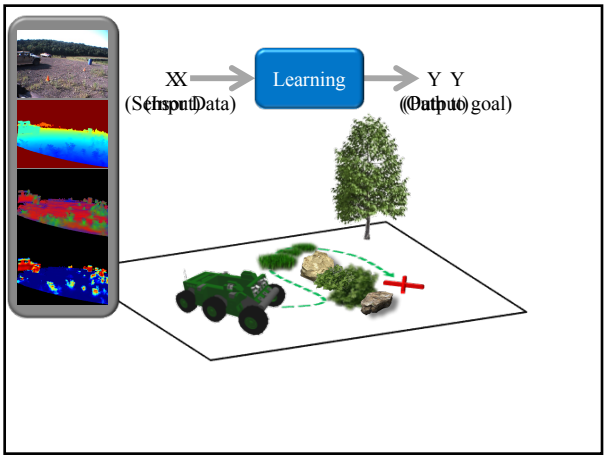
CS 473: Intro to Artificial Intelligence
Dieter Fox

Inverse Optimal Control
(Inverse Reinforcement Learning)

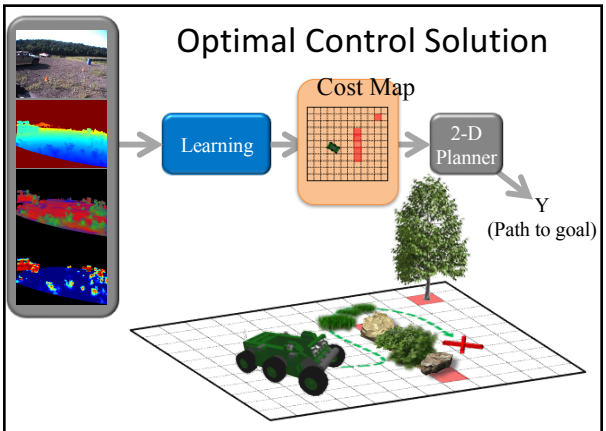
Most slides by
Drew Bagnell / Brian Ziebart
Carnegie Mellon University

Autonomous Navigation





Optimal Control Solution



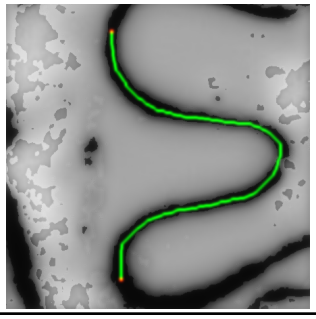
Mode 1: Training example

Mode 1: Training example

Mode 1: Learned behavior

Mode 1: Learned behavior

Mode 1: Learned cost map



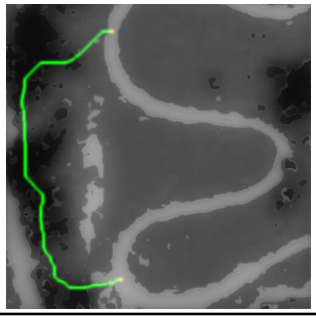
Mode 2: Training example

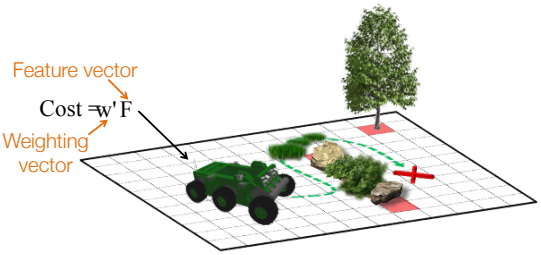
Mode 2: Training example

Mode 2: Learned behavior

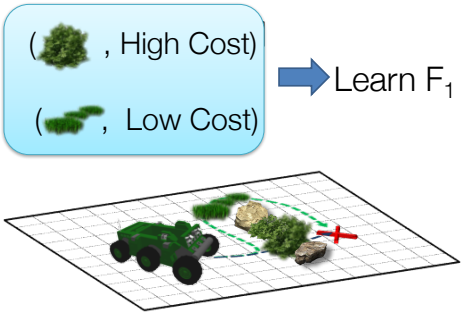
Mode 2: Learned behavior

Mode 2: Learned cost map

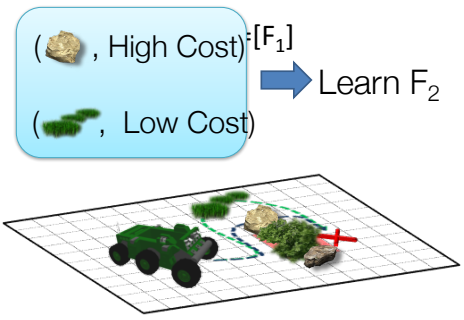




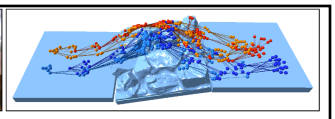
Ratliff, Bagnell, Zinkevich 2005
Ratliff, Bradley, Bagnell, Chestnutt, NIPS 2006
Silver, Bagnell, Stentz, RSS 2008



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



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

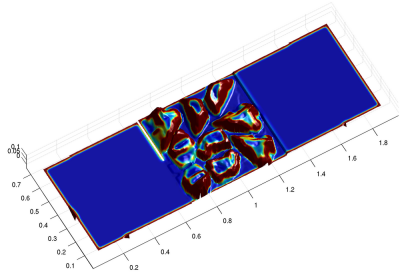



Learning Locomotion Phase III
Matt Zucker - The Robotics Institute

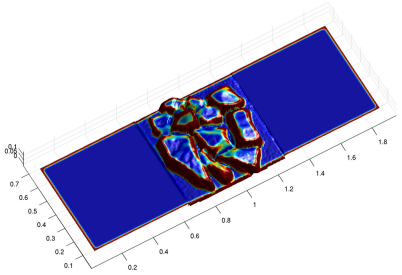
Ratliff, Bradley, Chesnutt, Bagnell 06

Zucker, Ratliff, Stolle, Chesnutt, Bagnell, Atkeson, Kuffner 09

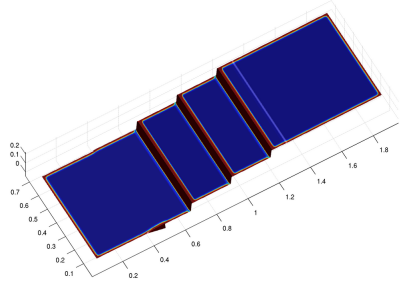
Learned Cost Function Examples



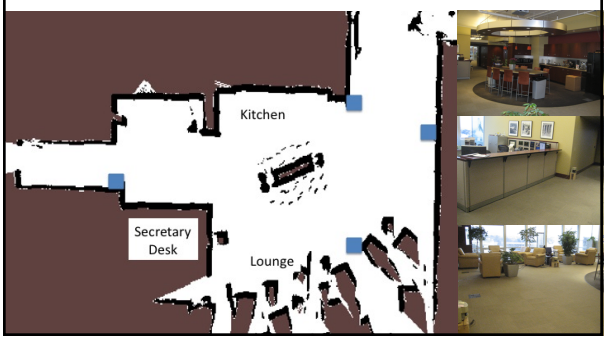
Learned Cost Function Examples



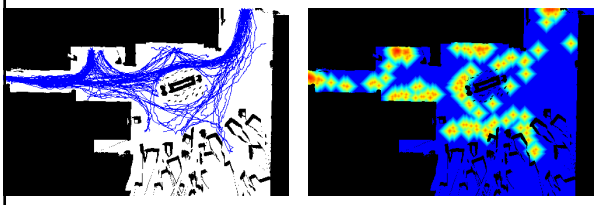
Learned Cost Function Examples



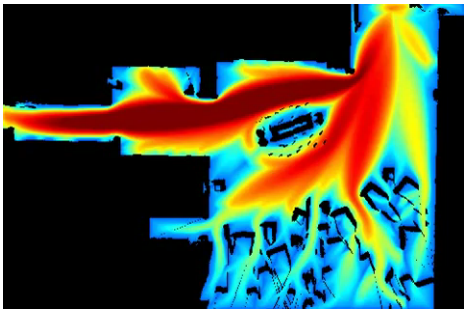
Pedestrian Trajectory Prediction



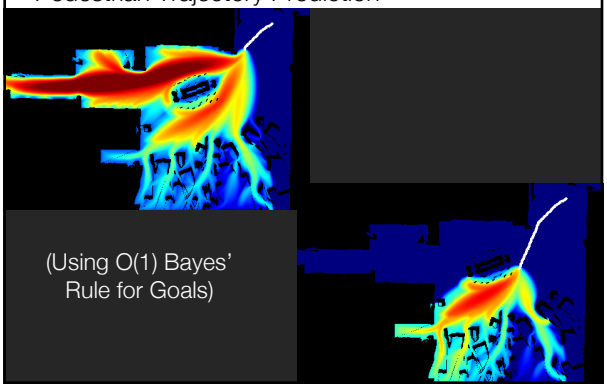
Pedestrian Trajectory Prediction



Pedestrian Trajectory Prediction

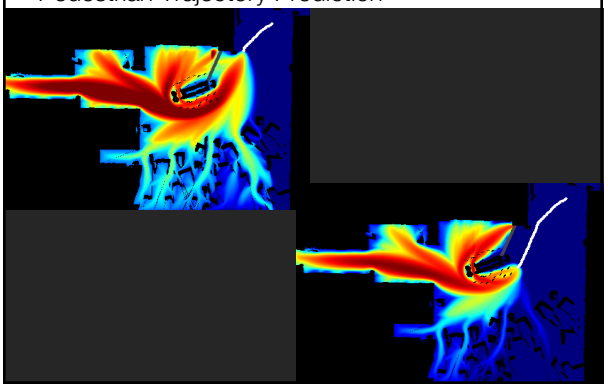


Pedestrian Trajectory Prediction



(Using O(1) Bayes' Rule for Goals)

Pedestrian Trajectory Prediction

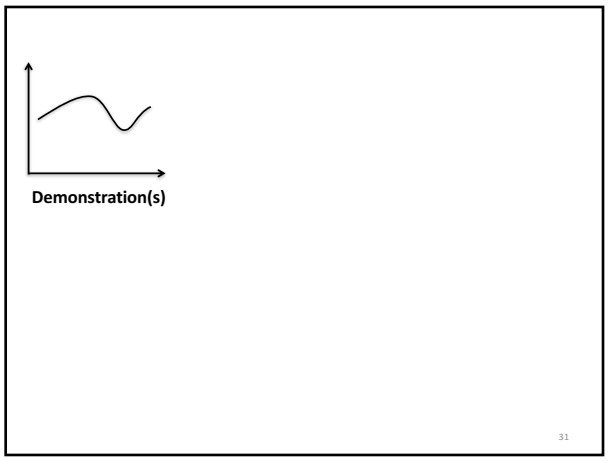


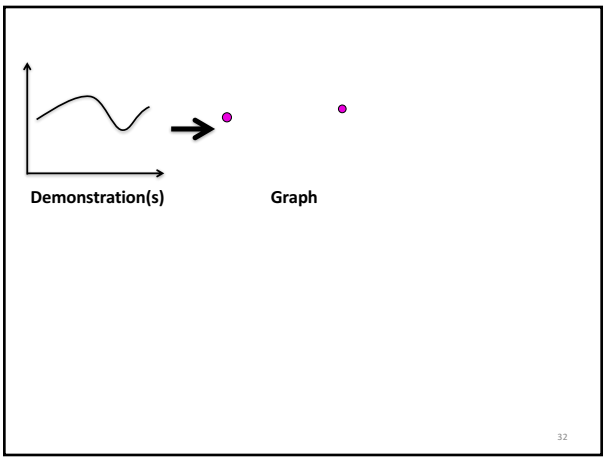
Staying out of People's Path

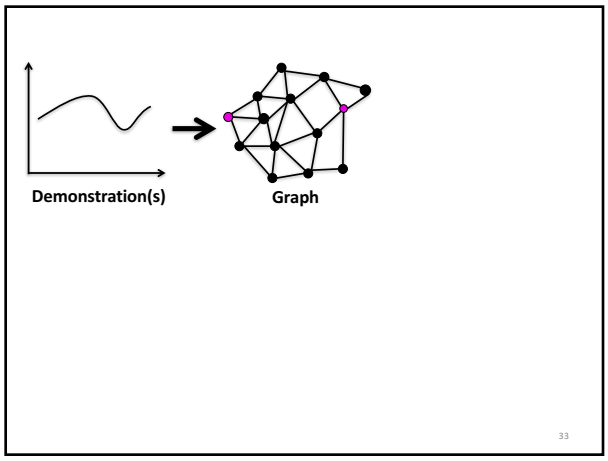
Learning Manipulation Preferences

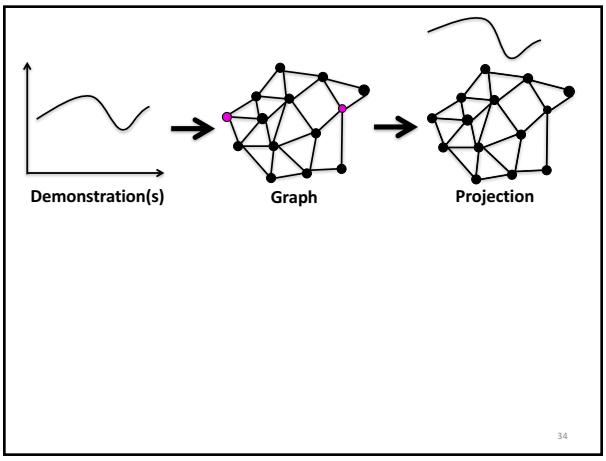
- ***Input:*** Human demonstrations of preferred behavior (e.g., moving a cup of water upright without spilling)
- ***Output:*** Learned cost function that results in trajectories satisfying user preferences

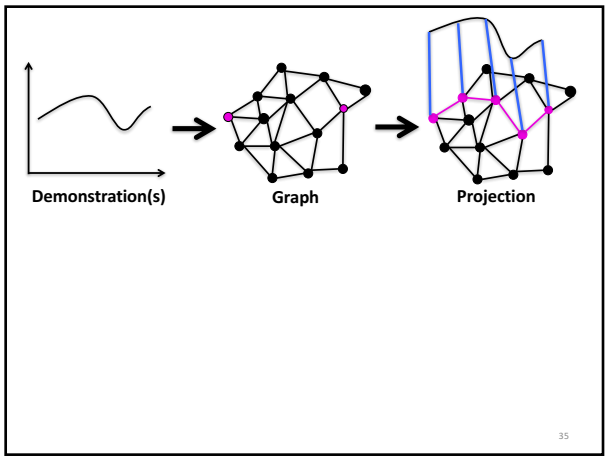
30

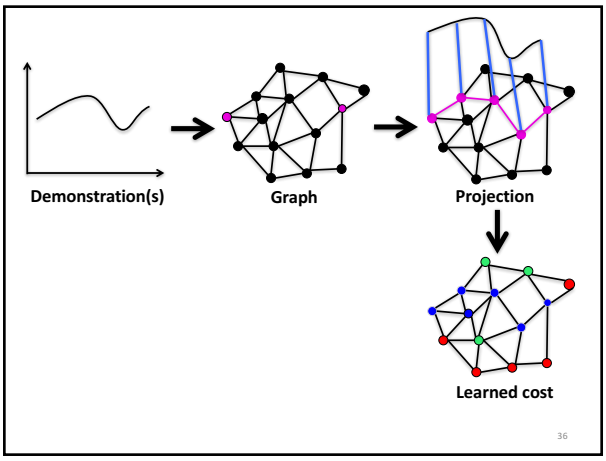

Demonstration(s)
31

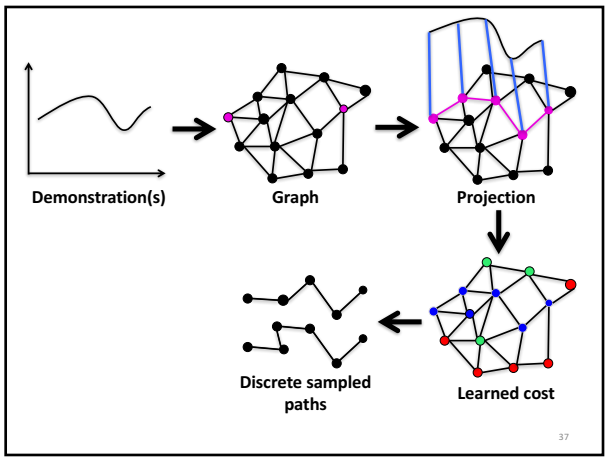

Demonstration(s)
Graph
32

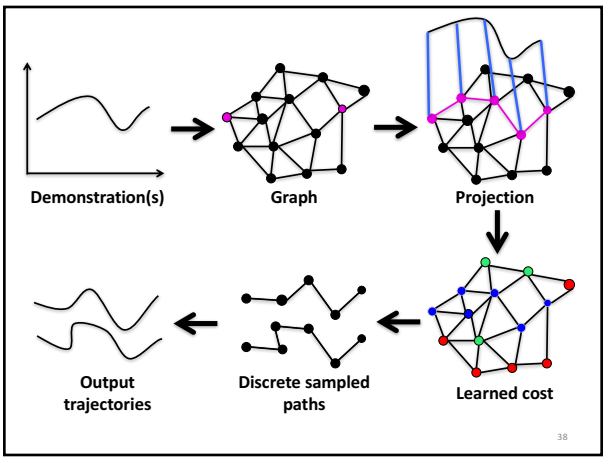

Demonstration(s)
Graph
33

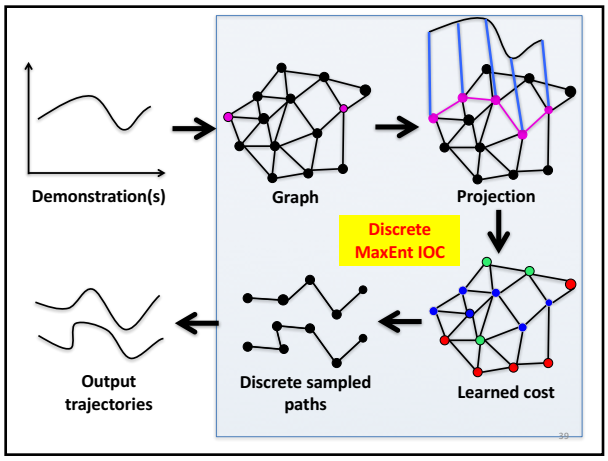

Demonstration(s)
Graph
Projection
34

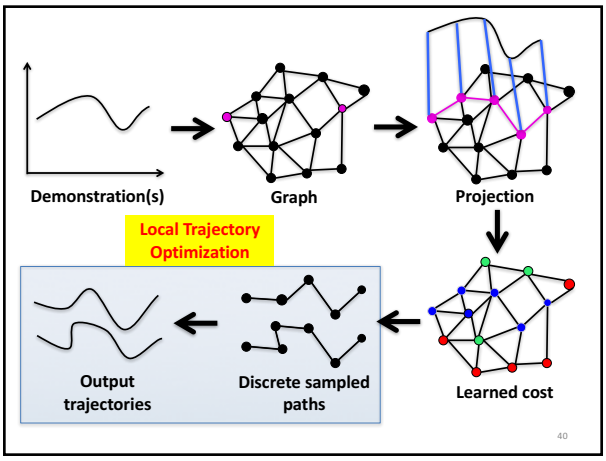

Demonstration(s)
Graph
Projection
35


Demonstration(s)
Graph
Projection
Learned cost
36









Laptop task: Demonstration

(Not part of training set)

Laptop task: LTO + Discrete graph path

That's all for Reinforcement Learning!

- Very tough problem: How to perform any task well in an unknown, noisy environment!
- Traditionally used mostly for robotics, but becoming more widely used
- Lots of open research areas:
  - How to best balance exploration and exploitation?
  - How to deal with cases where we don't know a good state/feature representation?
  - How to best bootstrap the learning process from demonstrations?

Conclusion

- We're done with Part I: Search and Planning!
- We've seen how AI methods can solve problems in:
  - Search
  - Constraint Satisfaction Problems
  - Games
  - Markov Decision Problems
  - Reinforcement Learning
- Next up: Part II: Uncertainty and Learning!