

# Learning from Limited Demonstrations

**Beomjoon Kim**

Amir-massoud Farahmand

Joelle Pineau

Doina Precup



# Learning from Demonstrations



Successes:

- Helicopter Flight [Abbeel 07]
- Robot Surgical Task [Berg 10]
- UAV Control [Ross 13]
- Navigation in Human Environments [Kim 13]

# What if we have limited demonstrations?



Suboptimal Expert

and/or



Expensive Expert

**Key Idea:** Learn from both **trial-and-error** and **demonstrations**

**Reinforcement Learning**

+

**Learning from Demonstrations**

# Approximate Policy Iteration with Demonstrations (APID)

Approximate Policy Iteration

Large Margin Classification

$$\hat{Q} \leftarrow \operatorname{argmin}_{Q \in \mathcal{F}|\mathcal{A}|, \xi \in \mathbb{R}_+^m} \left( \|Q - T^\pi Q\|_n^2 + \lambda J^2(Q) \right) + \frac{\alpha}{m} \sum_{i=1}^m \xi_i$$

$$\text{s.t. } Q(X_i, \pi_E(X_i)) - \max_{a \in \mathcal{A} \setminus \pi_E(X_i)} Q(X_i, a) \geq 1 - \xi_i$$

- PAC guarantee on **Bellman Error**
- **Real Robot** Path Finding Experiment

