High-level Learning

Jane Li

Assistant Professor
Mechanical Engineering Department, Robotic Engineering Program
Worcester Polytechnic Institute
Quiz 10 (pts)

• (2 pts) How to avoid obstacles when reproducing a trajectory using a learned DMP?

• (2 pts) How to synchronize the motions reproduced using DMP across multiple DOFs?

• (3 pts) Describe the phase matching problem in human-robot interaction learning

• (3 pts) Compared to DTW and GMM/GMR, what are the advantages of ProMP?
Phase estimation problem

- Observes human partner’s motion
  - Sparse
- Predict end user’s motion
  - Prediction may vary by fitting sparse data to variants of a model that differ by temporal scaling
- Generate motions that match
  - Wrong prediction leads to mismatch between human and robot motions
Related work - Trajectory alignment

- Dynamic time warping (DTW)
Related work – GMM/GMR [4,6]
Limitation of DTW

• Represent averaged behavior?
  • No. Pairwise matching

• How to represent variability?
  • No. Pairwise matching

• How to align trajectory?
  • Yes. Align fast and slow trajectory (deterministic, need to choose a unique reference)

• How to match phase in the learning of human-robot interaction?
  • Yes.
Limitation of GMM/GMR

• Represent averaged behavior?
  • Yes

• How to represent variability?
  • Yes

• How to align trajectory?
  • No. Cannot align fast or slow motions

• How to match phase in the learning of human-robot interaction?
  • No.
Pro-MP

• Represent averaged behavior?
  • Yes

• How to represent variability?
  • Yes

• How to align trajectory?
  • Yes. Align motions by phase matching

• How to match phase in the learning of human-robot interaction?
  • Yes.
Selected student presentations
Best work for Assignment 10

- Heramb Nemlekar (selected for presentation)
- Gunner Hover
- Aishwary Jagetia
- Sanjuksha Nirgude (selected for presentation)
Best work for Assignment 11

- Heramb Nemlekar
- Tess Meier (selected for presentation)
- Sihui Li
- Aishwary Jagetia (selected for presentation)
Rewards

• Student presentation – Spotlight talk
  • Each one gives a 5-min talk
  • Interactive session for Q&A after all the presentation

• Publish good examples of presentation and review on canvas

• Grade boosting for presentation
  • You can choose a low-grade assignment/quiz to replace with full score
  • Let our TA know which you prefer to replace, by Wednesday this week
High-Level Learning
How actions derived from \textit{low-level learning} can be used to learn \textit{higher level tasks}
Overview

• What can be learned at high level?
  • State-action mapping function, i.e., policy
  • Task plan, objectives, features
  • Reference frame, affordance

• How to learn?
  • Supervised/unsupervised learning
  • Reinforcement learning
• Action Primitives = MP or sequence of MP
  • E.g., Reach-To, Pick-Up, Move-Forward, etc.

• Can be hand-coded, developed using planner, learned from demos

• Often parameterized
Define desired behavior using action primitives

• Explicit task goals
  • Pre- and post-condition of action primitives
  • A particular configuration of an object

• Implicit task goals = Reward function
  • Sparse/dense reward
  • Learn reward function? → IOC, IRL
Learning mapping functions
Learning mapping function

• Input state $\rightarrow$ output action

• Demonstration = state-action pairs

• Learning policy

• Objectives underlying the policy? $\rightarrow$ don’t care
Decision Tree [1,2]

**Decision Tree: Should I accept a new job offer?**

- **Salary at least $50,000?**
  - Yes
  - Commute more than 1 hour?
    - Yes
    - Offers free coffee?
      - Yes
      - Accept offer
      - No
      - Decline offer
    - No
    - Decline offer
  - No
  - Offers free coffee?
    - Yes
    - Accept offer
    - No
    - Decline offer
  - No
  - Decline offer

**Root node**

**Decision nodes**

**Leaf nodes**
Training process

Learn state transition function as hierarchical classifier over features

\[ T_S(f_t) \rightarrow S' \]
Decision Tree

Goal is out of view
Example of a Leaned Policy
Other behaviors (policies)

Wander

Start -> Fowards
Front Clear -> FrontLeft Blocked
Front Clear -> FrontRight Blocked

Right -> Left

Disperse(Color)

Start -> Wander
No(Color) -> Left(Color)
No(Color) -> Right(Color)

Right -> Left
More complicated behaviors (policies)

**Attack(Color)**

- Start → Servo(\text{Color}) → Close(\text{Color}) → Stop, Signal \textit{Done}

**Patrol**

- Start → Disperse(T) → See(I) → Attack(I) → Done
- Attack(H) → "Go Home" → Done
Memory-based Techniques

• Lazy learning algorithms
  • Memorize demonstrated action in a robot state
  • In a new states, search for similar old states and apply the corresponding action
  • Used for learning navigation from demonstrations

• Method for measuring similarity?
  • KNN [4]
  • Case-based reasoning [5,6]
Bayesian Method [7]

- Observe and memorize a sequence of states (sub-goals)
- Pick up the action that maximizes the chance of taking the agent from current state to the memorized next state

$$a^* = \arg \max_a p(S_t \rightarrow S_{t+1})$$
Essentially, a classification/regression problem

Estimate classification confidence
- Integrate a measure of confidence in classification/regression
- Address the uncertainty in action

Methods for estimating classification confidence
- Bayesian methods [8]
- Confidence-Based Autonomy algorithm [9]
- Locally Weighted Projection Regression [10]
Learning a task plan
Learning a task plan

• Represent desired robot behavior as a plan

• Generalized state-action mapping
  • State: Pre-condition, post-condition
  • Action: Sequence of action between initial and goal states

• Underlying objectives = goal state
How to represent a plan?

- Bayesian models
- Finite-state Automaton

Finite-state Automaton
Compared to learning policy

- More sparse states
  - Initial states
  - Goal states
- What if something goes wrong in-between?
  - Provide additional demonstration as correction
  - Iterative and incremental learning
Interactive correction

- Re-do parts of the demo
- Re-segment old data
- Add new corrections and rebuild FSA
Assignment 14 (30 pts) – Due Dec 4

- Read
  - Section 5.5 Learning task features

- Prepare 7-10 presentation slides
  - Digest over multiple papers

- To reflect your understanding
  - Add notes to your presentation slides, or
  - Submit 2-page review

Reference


Reference


