Good Afternoon Colleagues

- Are there any questions?
Good Afternoon Colleagues

- Are there any questions?

- Pending questions:
  - Helicopter: MDP or POMDP?
  - Bagnell Theorem 2 correct?
Logistics

- Tom Dietterich visiting tomorrow:
  “Three Challenges for Machine Learning Reserch”
  3pm, ACES 2.302
Learning from Primitives

• The main reason I chose this paper...
Learning from Primitives

• The main reason I chose this paper...
• Another reason: Q-learning works! (Fig. 7)
Learning from Primitives

- The main reason I chose this paper...
- Another reason: Q-learning works! (Fig. 7)
- Start with 5 hand-coded primitives:
  ball goal $\mapsto$ motor actions
Learning from Primitives

- The main reason I chose this paper...
- Another reason: Q-learning works! (Fig. 7)
- Start with 5 hand-coded primitives: ball goal $\mapsto$ motor actions
- **Memory-based** learning for when to pick primitives and which **Subgoals** to use
Learning from Primitives

- The main reason I chose this paper...
- Another reason: Q-learning works! (Fig. 7)
- Start with 5 hand-coded primitives:
  ball goal $\mapsto$ motor actions
- **Memory-based** learning for when to pick primitives and which **Subgoals** to use
- Learning from practice amounts to adjusting the **kernel** function: (state, experience) $\mapsto$ value
Learning from Primitives

• The main reason I chose this paper...

• Another reason: Q-learning works! (Fig. 7)

• Start with 5 hand-coded primitives:
  ball goal $\mapsto$ motor actions

• **Memory-based** learning for when to pick primitives and which **Subgoals** to use

• Learning from practice amounts to adjusting the **kernel** function: (state, experience) $\mapsto$ value
  – State can be discretized, stored in a table
  – Or approximated continuously with **LWPR**: locally weighted polynomial regression
Learning from Primitives

- The main reason I chose this paper...
- Another reason: Q-learning works! (Fig. 7)
- Start with 5 hand-coded primitives: ball goal $\mapsto$ motor actions
- **Memory-based** learning for when to pick primitives and which **Subgoals** to use
- Learning from practice amounts to adjusting the **kernel** function: (state, experience) $\mapsto$ value
  - State can be discretized, stored in a table
  - Or approximated continuously with **LWPR**: locally weighted polynomial regression
- Any active exploration?
Learning from Primitives

• The main reason I chose this paper...
• Another reason: Q-learning works! (Fig. 7)
• Start with 5 hand-coded primitives: ball goal $\rightarrow$ motor actions
• **Memory-based** learning for when to pick primitives and which **Subgoals** to use
• Learning from practice amounts to adjusting the kernel function: $(state, experience) \rightarrow value$
  – State can be discretized, stored in a table
  – Or approximated continuously with **LWPR**: locally weighted polynomial regression
• Any active exploration?
Student-led Discussion

- Michael on biological primitives
Discussion Points

• Where does the reward function come from? What’s up with corners?
Discussion Points

• Where does the reward function come from? What’s up with corners?

• Can we use the primitives without observations? or is the action space too large?
Discussion Points

• Where does the reward function come from? What’s up with corners?

• Can we use the primitives without observations? or is the action space too large?

• Could you learn from bad trials instead of good ones?
Discussion Points

• Where does the reward function come from? What’s up with corners?

• Can we use the primitives without observations? or is the action space too large?

• Could you learn from bad trials instead of good ones?

• How could they make this model-based?
Discussion Points

- Where does the reward function come from? What’s up with corners?
- Can we use the primitives without observations? or is the action space too large?
- Could you learn from bad trials instead of good ones?
- How could they make this model-based?
- How do primitives compare to options, MAXQ?
Discussion Points

- Where does the reward function come from? What’s up with corners?

- Can we use the primitives without observations? or is the action space too large?

- Could you learn from bad trials instead of good ones?

- How could they make this model-based?

- How do primitives compare to options, MAXQ?

- Should we be impressed?
Discussion Points

- Where does the reward function come from? What’s up with corners?

- Can we use the primitives without observations? or is the action space too large?

- Could you learn from bad trials instead of good ones?

- How could they make this model-based?

- How do primitives compare to options, MAXQ?

- Should we be impressed?