

CS394R
Reinforcement Learning:
Theory and Practice

Scott Niekum and Peter Stone

Department of Computer Science
The University of Texas at Austin

Good Morning Colleagues

- Are there any questions?

Logistics

- Start thinking about final project

Logistics

- Start thinking about final project
- Next week's readings

Logistics

- Start thinking about final project
- Next week's readings
 - Eligibility traces: blending different n-step returns
 - Some sections can be skimmed

Logistics

- Start thinking about final project
- Next week's readings
 - Eligibility traces: blending different n-step returns
 - Some sections can be skimmed

Chapter 10

- On-policy control with approximation

Chapter 10

- On-policy control with approximation
- Policy improvement with complex state (and action) spaces

Chapter 10

- On-policy control with approximation
- Policy improvement with complex state (and action) spaces
 - Semi-gradient Control (SARSA)
 - Average reward RL

Chapter 10

- On-policy control with approximation
- Policy improvement with complex state (and action) spaces
 - Semi-gradient Control (SARSA)
 - Average reward RL
- Later: policy gradient methods

Mountain Car

- TAMER webpage: <http://www.bradknox.net/human-reward/tamer/tamer-in-action/mountain-car/>
- 3D MC slides

Common Questions

- Why not use average reward RL in the first place?

Common Questions

- Why not use average reward RL in the first place?
- Is ergodicity a common assumption in RL?

Common Questions

- Why not use average reward RL in the first place?
- Is ergodicity a common assumption in RL?
- Why does the policy improvement theorem fail with function approximation?

Common Questions

- Why not use average reward RL in the first place?
- Is ergodicity a common assumption in RL?
- Why does the policy improvement theorem fail with function approximation?
 - Proof on p.78 relies on q and v being exact (correct)

Other Interesting Questions

- Chima Ezeilo: Can the assumption of ergodicity be used in the episodic case or is it strictly for the continuous case?

Other Interesting Questions

- Chima Ezeilo: Can the assumption of ergodicity be used in the episodic case or is it strictly for the continuous case?
- Jiaxun Cui: Can we get at least asymptotic improvement for the on-policy control under function approximation?

Other Interesting Questions

- Chima Ezeilo: Can the assumption of ergodicity be used in the episodic case or is it strictly for the continuous case?
- Jiaxun Cui: Can we get at least asymptotic improvement for the on-policy control under function approximation?
- Garrett Gu: Does it really matter that we cannot guarantee convergence in corner cases if the algorithm still converges in most real-world use cases?