

CS394R

Reinforcement Learning: Theory and Practice

Amy Zhang and Peter Stone

Departments of ECE and CS
The University of Texas at Austin

Good Morning Colleagues

- Are there any (logistics) questions?

Logistics

- Programming assignment 3 due Sunday

Logistics

- Programming assignment 3 due Sunday
- Continue towards final project proposal

Logistics

- Programming assignment 3 due Sunday
- Continue towards final project proposal
- Grades so far visible in canvas

Logistics

- Programming assignment 3 due Sunday
- Continue towards final project proposal
- Grades so far visible in canvas
- Midterm after spring break

Logistics

- Programming assignment 3 due Sunday
- Continue towards final project proposal
- Grades so far visible in canvas
- Midterm after spring break
- Next week's readings

Logistics

- Programming assignment 3 due Sunday
- Continue towards final project proposal
- Grades so far visible in canvas
- Midterm after spring break
- Next week's readings
 - Policy gradient methods

Chapter 12 - Eligibility Traces

- Another way to blend TD \rightarrow MC (other than n-step returns)

Chapter 12 - Eligibility Traces

- Another way to blend TD \rightarrow MC (other than n-step returns)
- Equally applicable in continuous and discrete settings

Common Questions

- When do we use online vs offline TD?
- Please discuss true online TD lambda further
- Please explain the relationship between the forward and backward views
- Why is $T D(\lambda)$ an approximation of the off-line λ -return algorithm? Where is the approximation?

Other Common Questions

- How do we set the proper lambda (and other hyperparameters) for TD Learning?

Other Common Questions

- How do we set the proper lambda (and other hyperparameters) for TD Learning?
- Is it possible to combine the λ and γ parameters?

Other Common Questions

- How do we set the proper lambda (and other hyperparameters) for TD Learning?
- Is it possible to combine the λ and γ parameters?
- Please discuss pseudo-termination further

Other Common Questions

- How do we set the proper lambda (and other hyperparameters) for TD Learning?
- Is it possible to combine the λ and γ parameters?
- Please discuss pseudo-termination further
 - Predict quantities that aren't part of the problem

Other Common Questions

- How do we set the proper lambda (and other hyperparameters) for TD Learning?
- Is it possible to combine the λ and γ parameters?
- Please discuss pseudo-termination further
 - Predict quantities that aren't part of the problem
 - e.g. reward in 4 steps
 - number of steps to landmark