Reinforcement Learning, Tutorial 05

Philipp Kratzer

Machine Learning and Robotics Lab

University of Stuttgart
Germany

May 27th, 2020

Outline

Announcements

- Next week no lecture/tutorials
- Next lecture on June 9th
- Next exercise sheet will be published on June 9th
- Register for the exam (campus)

Outline

1a

Task: Random Walk Example: What happened on the first episode? Why was only the estimate for this one state changed? By exactly how much was it changed (assuming $\alpha = 0.1$)?

1a

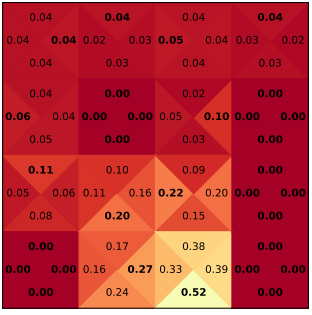
Task: Random Walk Example: What happened on the first episode? Why was only the estimate for this one state changed? By exactly how much was it changed (assuming $\alpha = 0.1$)?

1. Must have ended left
2. For other states td error was 0
3. TD-update after the last transition:

$$\begin{aligned} V(A) &\leftarrow V(A) + \alpha\big[R_{t+1} + \gamma V(S_{t+1}) - V(S_t)\big] \\ &= 0.5 + 0.1\big[0 + \gamma \cdot 0 - 0.5\big] \\ &= 0.5 + 0.1(-0.5) \\ &= 0.45 \end{aligned}$$

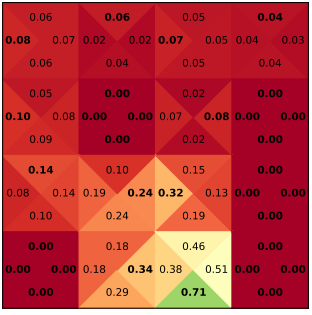
2a

Task: Implement sarsa



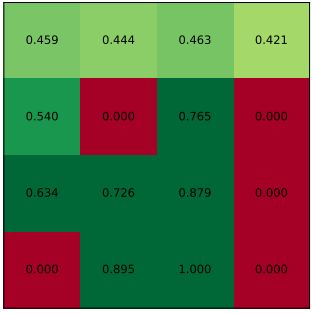
2b

Task: Implement Q-learning

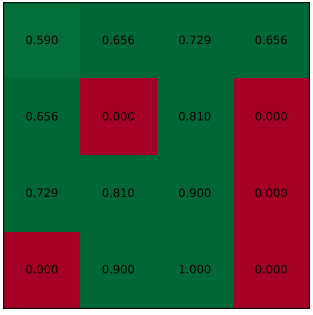


2c

Task: non-slippery



sarsa



q-learning

- sarsa prefers the safer way

Announcements

- Next week no lecture/tutorials
- Next lecture on June 9th
- Next exercise sheet will be published on June 9th