

Reinforcement Learning, Tutorial 05

Philipp Kratzer

Machine Learning and Robotics Lab



University of Stuttgart
Germany

May 27th, 2020

Outline

1. Announcements

2. Solutions Discussion

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

1. Announcements

2. Solutions Discussion

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:

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

2a

Task: Implement sarsa

0.045	0.041	0.054	0.036
0.061	0.000	0.101	0.000
0.114	0.204	0.224	0.000
0.000	0.270	0.520	0.000

0.04	0.04	0.04	0.04				
0.04	0.04	0.02	0.03	0.05	0.04	0.03	0.02
0.04		0.03		0.04		0.03	
0.04		0.00		0.02		0.00	
0.06	0.04	0.00	0.00	0.05	0.10	0.00	0.00
0.05		0.00		0.03		0.00	
0.11		0.10		0.09		0.00	
0.05	0.06	0.11	0.16	0.22	0.20	0.00	0.00
0.08		0.20		0.15		0.00	
0.00		0.17		0.38		0.00	
0.00	0.00	0.16	0.27	0.33	0.39	0.00	0.00
0.00		0.24		0.52		0.00	

2b

Task: Implement Q-learning

0.078	0.062	0.071	0.042
0.099	0.000	0.078	0.000
0.145	0.244	0.317	0.000
0.000	0.342	0.713	0.000

0.06	0.07	0.02	0.02	0.05	0.04	0.03
0.08	0.06	0.04	0.05	0.04		
0.05	0.08	0.00	0.02	0.08	0.00	0.00
0.10	0.09	0.00	0.02		0.00	
0.14	0.10	0.15		0.00		
0.08	0.14	0.19	0.24	0.32	0.13	0.00
0.10	0.24	0.19				0.00
0.00	0.18	0.46				0.00
0.00	0.00	0.18	0.34	0.38	0.51	0.00
0.00	0.29	0.71				0.00

2c

Task: non-slippery

0.459	0.444	0.463	0.421
0.540	0.000	0.765	0.000
0.634	0.726	0.879	0.000
0.000	0.895	1.000	0.000

sarsa

- sarsa prefers the safer way

0.590	0.656	0.729	0.656
0.656	0.000	0.810	0.000
0.729	0.810	0.900	0.000
0.000	0.900	1.000	0.000

q-learning

Announcements

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