Reinforcement Learning, Tutorial 04

Philipp Kratzer

Machine Learning and Robotics Lab

University of Stuttgart
Germany

May 20th, 2020

Outline

Announcements

- Exercise sheet not graded yet
- Next exercise sheet is available

Announcements

- We want this session to be more interactive, so I will ask for volunteers to present their solutions
- Solutions presented do not have to be perfect, use the chance to get feedback directly on what you did!

Outline

1a

Task: Advantages of Monte Carlo methods over dynamic programming?

1a

Task: Advantages of Monte Carlo methods over dynamic programming?

1. No model required
2. Can be used with simulations
3. Can focus on small subset of the state space
4. Less harmed to violations of the Markov property

1b

Task: Example environment where it could perform better?

2

Task: Implement Monte Carlo for Blackjack
a) First-visit Monte Carlo prediction
b) Monte Carlo ES

2b

Optimal policy:

	A	2	3	4	5	6	7	8	9	10
21	S	S	S	S	S	S	S	S	S	S
20	S	S	S	S	S	S	S	S	S	S
19	S	S	S	S	S	S	S	S	S	S
18	H	S	S	S	S	S	S	S	H	H
17	H	H	H	H	H	H	H	H	H	H
16	H	H	H	H	H	H	H	H	H	H
15	H	H	H	H	H	H	H	H	H	H
14	H	H	H	H	H	H	H	H	H	H
13	H	H	H	H	H	H	H	H	H	H
12	H	H	H	H	H	H	H	H	H	H

usable ace

	A	2	3	4	5	6	7	8	9	10
21	S	S	S	S	S	S	S	S	S	S
20	S	S	S	S	S	S	S	S	S	S
19	S	S	S	S	S	S	S	S	S	S
18	S	S	S	S	S	S	S	S	S	S
17	S	S	S	S	S	S	S	S	S	S
16	H	S	S	S	S	S	H	H	H	H
15	H	S	S	S	S	S	H	H	H	H
14	H	S	S	S	S	S	H	H	H	H
13	H	S	S	S	S	S	H	H	H	H
12	H	H	H	S	S	S	H	H	H	H

no usable ace

Outline

Next exercise sheet

- Next exercise sheet available
- It is about TD methods
- Programming part is again on the FrozenLake environment
- Sourcecode on github
  `https://github.com/humans-to-robots-motion/rl-course`