Kahoot!

Lecture 05: Planning and Learning

#cs #rl

74 plays • 119 players

A public kahoot

Questions (7)

1 - Quiz

In model-based reinforcement learning,

60 sec

▲	the MDP of the environment is given.	✗
◆	we learn a model of the underlying environment.	✓
●	we learn a value function from samples from the environment.	✗
■	we can use planning on a model to obtain a value function.	✓

2 - True or false

It is always easier to learn a dynamics model than a policy.

60 sec

▲	True	✗
◆	False	✓

3 - Quiz

It can be a good choice to learn the state difference rather than the transition to a global state -- why?

60 sec

▲	The numbers are usually smaller.	✗
◆	The model suffers less from accumulating errors.	✗
●	There can be local similarities w.r.t. the state differences.	✓

4 - Quiz

In sample-based planning, we ... 60 sec

▲	we solve the MDP directly.	✗
◆	apply planning to the MDP.	✗
●	we apply model-free RL to sampled experience.	✓
■	suffer less from the curse of dimensionality.	✓

5 - Quiz

In Dyna, we learn the value function/ policy from 60 sec

▲	samples from the learned model.	✗
◆	samples of the real environment.	✗
●	imaginations of the real world.	✗
■	samples from the learned model and the real environment.	✓

6 - Quiz

In Prioritized Sweeping, we update (s,a)-pairs according to ... 60 sec

▲	their absolute Q-value.	✗
◆	their absolute TD-error.	✓
●	the number of states that lead to them.	✗
■	their negative distance to the goal.	✗

7 - Quiz

In monte carlo tree search (MCTS), we ... 60 sec

▲	combine in-tree policies and out-of-tree policies.	✓
◆	traverse the tree randomly to obtain MC simulations.	✗
●	use a greedy policy as an in-tree policy.	✗
■	use a greedy policy as an out-of-tree policy.	✗