8. Q-value Iteration

Q-value Iteration

and do exactly the same computation

and directly get Qs.

And that's exactly what you will do in your exercise.

You will reformulate this algorithm

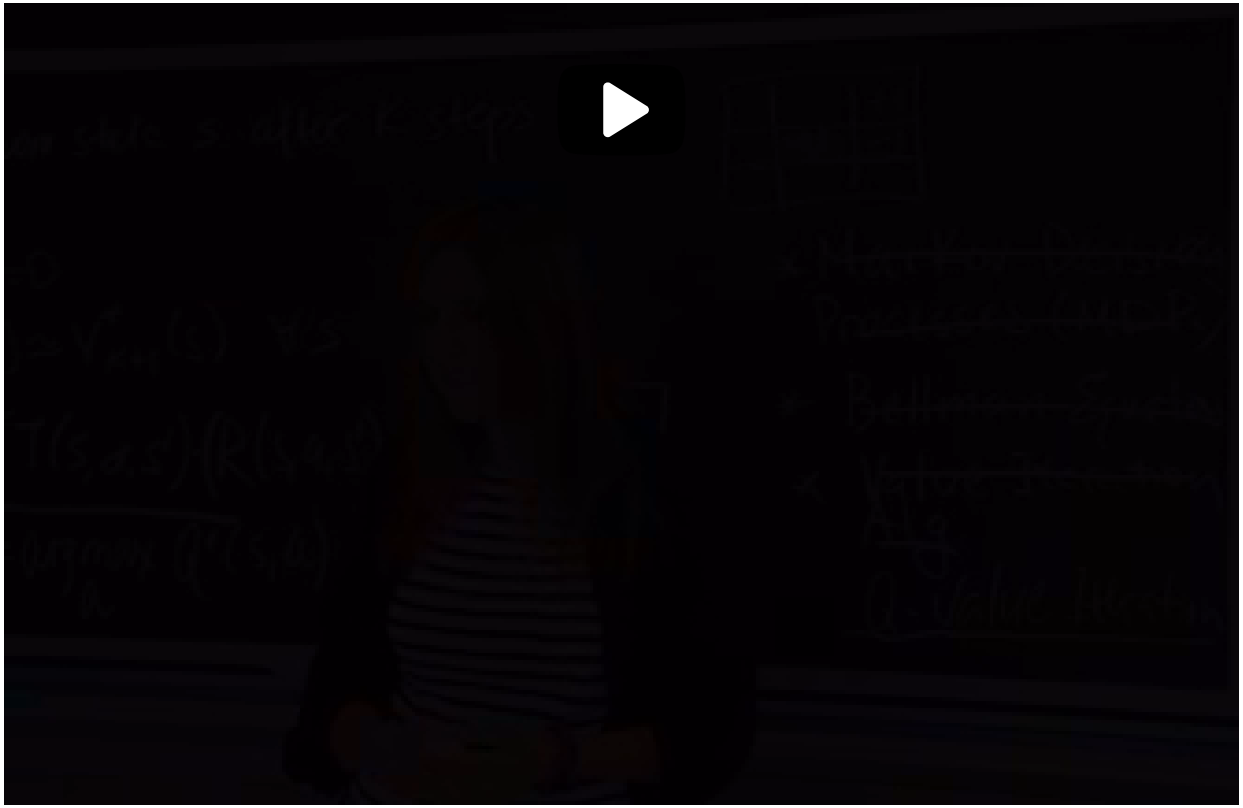
and get a new algorithm which is called Q value iterations.

But the idea-- exactly the same.

And we will use this algorithm in our discussion

about reinforcement learning next time.

Thank you.



0:59 / 0:59 1.50x

Video

Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

In this problem, we compute the optimal policy using Q-value iteration.

First, let's recall some of the bellman equations:

$$Q^*(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma V^*(s')\right)$$

We also know that, $V^*(s) = \max_a Q^*(s,a)$. Therefore,

$$Q^*(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma \max_{a'} Q^*(s',a')\right)$$

Q value iteration update rule

1/1 point (graded)

Q-value iteration update rule from the above equation could be derived similar to the value iteration update rule.

Pick the most suitable Q-value iteration update rule from the options below:

- $Q^*_{k+1}(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma \max_{s'} Q^*_k(s',a)\right)$
- $Q^*_{k+1}(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma V^*(s')\right)$
- **$Q^*_{k+1}(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma \max_{a'} Q^*_k(s',a')\right)$** ✔
- $Q^*_{k+1}(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma Q^*_k(s',a)\right)$

Solution:

First let's recall the following bellman equations:

$$Q^*(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma V^*(s')\right)$$

We also know that, $V^*(s) = \max_a Q^*(s,a)$. Therefore,

$$Q^*(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma \max_{a'} Q^*(s',a')\right)$$

Q value iteration would use previous iteration of the Q-value on the right hand side of the above equation to update the Q value estimate of the current step. Hence, the Q value update for k^{th} step would look like:

$$Q^*_{k+1}(s,a) = \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma \max_{a'} Q^*_k(s',a')\right)$$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 5 Reinforcement Learning (2 weeks) :Lecture 17. Reinforcement Learning 1 / 8. Q-value Iteration

Add a Post

Show all posts ▼ by recent activity ▼

There are no posts in this topic yet.

✖