4. Utility Function

Bookmark this page

Exercises due May 3, 2023 08:59 -03 Completed

Video note: In the video below at 1:25, Prof Barzilay miswrote “final horizon" on the boa should be meant **finite horizon**.

Utility Function

0:00 / 0:00 1.0x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Finite Horizon vs Discounted Reward

1/1 point (graded)

The main problem for MDPs is to optimize the agent's behavior. To do so, we first need that we are trying to maximize in terms of accumulated rewards. We will define a **utility** its expectation.

We consider two different types of utility functions:

1. **Finite horizon based utility** : The utility function is the sum of rewards after acti steps. For example, in the case when the rewards depend only on the states, th

How do these two types of utility function depend on the time steps?
(Choose all that apply.)

- [x] The action at state that maximizes a finite horizon based utility can depend on been taken.
- [ ] The action at state that maximizes a finite horizon based utility does **not** deper have been taken.
- [x] The action at state that maximizes a discount reward based utility does **not** de steps have been taken.
- [ ] The action at state that maximizes a discount reward based utility can depend have been taken.

✔

Submit You have used 1 of 3 attempts

Discounte

1/1 point (grade

Recall that the discounted reward in the case when is given by:

‹ Previous Next ›

Utility Function

R(s0) expression

What does Gamma represent in this expression? Why is the reward bounded at R_max / (1-gamma)?