2. Estimating Inputs for RL algorithm

Estimating Inputs for RL algorithm

0:00 / 0:00 1.0x

Start of transcript. Skip to the end.

So let me start showing to you the first approach,

which I will not necessarily take.

But I will propose you this approach.

And let's see together why it is not a good approach.

So the first idea you can say, OK, what I want to do,

I want to end up in this state.

Why?

Because they already have algorithms for

Video
Download video file

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

MDP Specifications in the Real World

1/1 point (graded)

We would need to provide the following tuple $< S, A, T, R >$ for our RL algorithms in an MDP setting.

However, in the real world, not all of these might be directly available.

From the options below select, which one(s) would not be readily available in most real-world scenarios.

- [ ] S: State Space
- [ ] A: Action Space
- [x] T: Transition Probabilities ✔
- [x] R: Reward Structure ✔

✔

Solution:

We would need to have some idea of the state space for our problem before we could even define $A, T \text{or} R$.
Since, the goal of Reinforcement Learning is to train an intelligent agent that can perform interesting tasks, it is possible to identify the set of actions that this agent is allowed to take.
T, R might not be so readily available in the non-deterministic noisy real world.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Estimating Transition Probabilities and Rewards

1/1 point (graded)

What might be some issue(s) with trying to estimate $\hat{T}, \hat{R}$ in the following manner?

$$\hat{T} = \frac{\text{count}(s,a,s')}{\sum_{s'} \text{count}(s,a,s')}$$

$$\hat{R} = \frac{\sum_{t=1}^{\text{count}(s,a,s')} R_t(s,a,s')}{\text{count}(s,a,s')}$$

Select one or more options from below that are correct:

- [ ] All states are guaranteed to be visited while collecting these statistics
- [x] Certain states might not be visited at all while collecting the statistics for $\hat{T}, \hat{R}$ ✔
- [x] Certain states might be visited much less often than others leading to very noisy estimates of $\hat{T}, \hat{R}$ ✔
- [ ] There are no issues with estimating $\hat{T}, \hat{R}$ in the above manner

✔

Solution:

Statistics for $\hat{T}, \hat{R}$ for a given state cannot be collected unless the agent visits this state during the estimation process. For tasks with a large state space it is possible that the agent only circles around a subset of the state space leaving the rest unexplored.

For the estimates, $\hat{T}, \hat{R}$ to be reliable, we would need to collect multiple samples of them for each state. If the state space is large, then during the exploration phase, some states would be significantly less explored than others resulting in very noisy estimates for these states.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Model Free vs Model Based Approaches

1/1 point (graded)

Say we want to estimate the expectation of a function $f(x)$.

$$\mathbb{E}[f(x)] = \sum_x p(x) \cdot f(x)$$

We have access K samples in order to compute our estimation. Select one or more statement(s) from the options below that are correct about model free or model based approaches

- [ ] Model Free Approach first tries to estimates the probability distribution before estimating the expectation
- [x] Model Based Approach first tries to estimates the probability distribution before estimating the expectation ✔

- [ ] Model Based Approach estimate would be given by $\frac{\sum_{i=1}^{K} f(X_i)}{K}$

- [x] Model Free Approach estimate would be given by $\frac{\sum_{i=1}^{K} f(X_i)}{K}$ ✔

✔

Solution:

- Model Based Approach works by sampling K points from the distribution $P(x)$ and estimating,

$$\hat{p}(X) = \text{count}(x)/K$$

before estimating the expectation of $f(X)$ as follows:

$$E[f(X)] \approx \sum p\hat{(x)} f(x)$$

- Model Free Approach would sample K points from $P(X)$ and directly estimate the expectation of $f(X)$ as follows:

$$E[f(X)] \approx \frac{\sum_{i=1}^{K} f(X_i)}{K}$$

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 5 Reinforcement Learning (2 weeks) :Lecture 18. Reinforcement Learning 2 / 2. Estimating Inputs for RL algorithm