

Reinforcement Learning Tutorial 1, Week 2

Introduction / MDPs

Sanjay Rakshit, Pavlos Andreadis

January 2020

Overview: The following tutorial questions relate to material taught in weeks 1 and 2 of the 2019-20 Reinforcement Learning course. They aim at encouraging engagement with the course material and facilitating a deeper understanding.

Problem 1 - Discussion

”How is Reinforcement learning different from Supervised learning?”

Open ended question for discussion. Attempt to write a short answer before the tutorial. It could be interesting to start from the comparison at this link [Shaikh \[2017\]](#), at the end of section 2, as it is flawed. Can you tell how? Consider also Section 1.1 in [Sutton and Barto \[2018\]](#).

Problem 2 - Discussion

Provide two different examples of applications of Reinforcement Learning in the industry. What aspects of these problems make them solvable by Reinforcement Learning?

You may find this link [Gary \[2018\]](#) to be useful, as it has a few examples and relevant discussion in section II. Can you find other examples? Consider also the first paragraph of Section 1.1 in [Sutton and Barto \[2018\]](#).

Problem 3 - Modelling: Frog on a Rock

A friendly frog, Hop-along was stranded on a rock surrounded by water. It needs to get to land without falling in. The only way to safety is for it to jump on to neighbouring rocks till it arrives on land.

There are 2 rocks, and Hop-along can jump from a rock to another or from the final rock to land. However, sometimes it misses and ends up in the water, having to climb back onto the same rock it tried to jump from. The rocks are arrayed in a row leading from its starting rock, which we will name $rock_0$, to $rock_1$, and then finally land which can only be reached from $rock_1$.

Consider the control problem where the current state is specified by Hop-along's location and the actions Hop-along can take is to attempt to jump towards another reachable rock or land.

Assume that Hop-along's jumps always have a 90% chance of reaching the intended destination, while the rest of the time Hop-along falls in the water.

Formulate a Markov Decision Process (MDP) for the problem of deciding on Hop-along's actions in order to help it reach land. Why did you formulate it as you did? What additional assumptions did you have to make?

References

- Gary. Applications of reinforcement learning in real world. <https://towardsdatascience.com/applications-of-reinforcement-learning-in-real-world-1a94955bcd12/>, 2018.
- F. Shaikh. Simple beginner's guide to reinforcement learning & its implementation. <https://www.analyticsvidhya.com/blog/2017/01/introduction-to-reinforcement-learning-implementation/>, 2017.
- Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.