

Reinforcement Learning Tutorial 1, Week 2

Introduction

Sanjay Rakshit, Pavlos Andreadis, Michael Herrmann

January 2025

Overview: The following tutorial questions relate to material taught at the start of week 1 of the 2024-25 Reinforcement Learning course. They aim at encouraging engagement with the course material and facilitating a deeper understanding.

This first tutorial starts with a couple of discussions on the concept of *Reinforcement Learning*. Use this tutorial as an opportunity to engage with your tutor and classmates, and to get a broader sense of the course and the subject area. This is also a good time to start thinking of how to use the resources provided by the course in setting up your personal learning plan.

As weeks progress, tutorials will also involve problem modeling exercises and algorithm applications, among other things. Answers provided to your tutorial questions and exercises will often pose further, “1-step ahead” questions. Consider these when studying and ask your tutor for advice, if you are stuck.

Reinforcement Learning is built around a few central concepts, and many algorithms are merely variations of the same basic approach. Trying to understand how one approach is similar to or different from another can be a good way to build an understanding of the fundamentals and how individual algorithms operate. You can also refer to the course algorithms chart (see the course resources page) to see how different approaches fit into the bigger picture of reinforcement learning. But all that in good time.

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, which is questionable. Can you give a better characterisation of the difference between these two types of learning? Consider also Section 1.1 in [Sutton and Barto, 2018](#). Also, discuss the reward hypothesis [Silver e.a., 2021](#) that was mentioned in the first lecture. Is there an analogous hypothesis underlying to supervised learning?

Problem 2 — Discussion

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

It may help to consider the possible state, action and reward spaces for your examples. You may find this link [Chan, 2018](#) useful, as it has a few examples and relevant discussion in section II, although it is now a little outdated. Can you find other examples? Consider also the first paragraph of Section 1.1 in [Sutton and Barto, 2018](#).

References

- Gary Chan. Applications of reinforcement learning in real world. <https://towardsdatascience.com/applications-of-reinforcement-learning-in-real-world-1a94955bcd12/>, 2018 (updated 2023).
- JalFaizy 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, <http://incompleteideas.net/book/the-book.html>, 2018.
- Yann LeCun Cake Analogy, <https://medium.com/syncedreview/yann-lecun-cake-analogy-2-0-a361da560dae>, 2016.
- Boutilier, Craig. A POMDP formulation of preference elicitation problems, 2002.
- David Silver, Satinder Singh, Doina Precup and Richard S Sutton. Reward is enough. *Artificial Intelligence* **299**, p. 103535, 2021.