Assignment 4: Free project
Course: Reinforcement Learning, Leiden University
Written by: Daniël Pelt

In this assignment, you can freely choose a topic to study and write a 6-page report about. The topic should include writing new code and running experiments to investigate behavior. The report should include a detailed explanation of what you did and what results you obtained (including relevant figures with captions). Make sure to choose a topic that is rich enough to write a full report about. Hand in both the report and the code that was used to run experiments with. For this assignment, being creative and challenging yourself will be rewarded. Furthermore, it is not required that the algorithms you try succeed in the task you are testing: observing that certain approaches do not work is interesting as well! However, make sure that you clearly describe what you did in the report, and try to interpret the results, giving possible explanations for observed behavior.

You can choose any topic related to reinforcement learning that you like. If you would like some guidance, some possible directions for topics are:

- Implementing approximation RL methods and applying them to either a tabular environment with a very large state-action space or a continuous environment, testing which approximation methods work well and which do not. Some possible continuous environments can be found here: `https://gym.openai.com/envs/#classic_control`.

- Implementing policy-based RL methods and applying them to a new environment or one of the environments we used before, observing differences with other algorithms we have studied in this course.

- Implementing n-step TD methods and applying them to a new environment or one of the environments we used before, observing differences with other algorithms we have studied in this course.

- Investigating other (potentially better!) ways of exploring within the RL methods we have studied in this course. As an example, Example 8.3 in Sutton & Barto describes an exploration heuristic that takes into account how many time steps ago a certain state-action pair has been visited in the real environment. Such exploration approaches could be implemented and tested on interesting environments (e.g., ones that change).

If you are unsure about whether your chosen topic is suitable, or if you have trouble picking a suitable topic, please contact us.