

SIT796 Reinforcement Learning

Distinction Task 3.1D: Exact policy iteration implementation for MDPs

Overview

During week 3, you have learnt about Markov Decision Processes. You have also been provided with an example of the implementation of Value Iteration in the workshop for the week.

In this task you need to implement the exact policy iteration method and test it in the environment you scoped in Tasks 1.1P. Your implementation should employ a linear system to find the exact solution for a policy that can achieve the goal. Your report, in the other hand, should refer to the relevant pseudocode in the your Jupyter Notebook. You are also expected to show a graphical model similar to that used in Week 3 Practical corresponding to your state and transition probabilities. Also include the average reward result over 1000 steps.

To complete this assignment, you need to refer back to Week 1 lecture, Task 1.1P and Practical 3

Submission Details

For this task you need to provide the code and a report on the implementation of the exact policy iteration that allows you to achieve the goal. You may use Practical 3 as the basis for the demonstration of your policy, but the environment itself has to be that you scoped in Task 1.1P. The implementation of the policy iteration method should be from scratch. To solve the actual linear system you may use NumPy.

The policy iteration should be demonstrated using a Jupiter notebook. The report and the Jupiter notebook need to be submitted to OnTrack. To do so, Convert the Jupyter Notebook to a **PDF** and append it to your report so as to submit that document. Also include a URL link to a GitHub repo containing the Notebook file and the environment implementation.

You may have to install pandoc to convert a Jupyter Notebook to a PDF document:

<http://pandoc.org/installing.html>

Constraints

The code used for the environment should be attributable only to the author of the report. The report should have a **high-quality writing style** not exceeding **600 words** in length.