Lund University
Automatic Control

FRTN70 Project in Systems, Control and Learning
Spring 2021

Reinforcement Learning for Double Pendulum

Project Plan Group H

Cem Alpturk[1] Pukashawar Pannu[2]

Project Advisor: Emil Vladu

[1] ce5368al-s@student.lu.se
[2] pu6218pa-s@student.lu.se

1 Project Purpose

The goal of this project is to teach a model to control a double pendulum on a cart to be standing upright, in it's equilibrium point, using reinforcement learning.

2 Equipments and Material

This project does not require any physical equipment. It will be fully simulated.

3 Modelling and System Design

Focus of the project is to train a controller for the dobule pendulum and not to model the actual double pendulum on a cart.

4 Division of Labour

We have decided to split our project into three parts: Setting up Environment, Train the Agents, Presentation.

4.1 Setting up Environment

In order to teach an agent to control a double pendulum we need to simulate it. This section is about setting up the requirements for the training phase. Equations for the double pendulum on a cart will be researched online and from other available sources and implemented using Python. Model will be verified by comparing with online models and an attempt will be made to animate the results to see if it looks as expected.

4.2 Training the Agents

This will be the main focus of the project, which is to create models and train them to control the double pendulum. The training will utilize reinforcement learning where the agent will learn from experience by trial and error. If the agent manages to keep the pendulum upright for a long enough time, the time has not been decided as of yet, it is considered to have solved the problem.

4.2.1 Deciding Network Architecture

Before we start training the models we need to decide on what type of architecture to use. We need to investigate the advantages and disadvantages of different types of networks such as feed forward, recurrent networks etc.

4.2.2 Deciding Simulation Conditions

We need to define a reward system for the training process, together with a completion and a failure condition.

4.3 Demo and Report

Create a basic demo showing our results and write a detailed report.

5 Time Plan

5.1 Subtasks

- Search online sources for the equations of motion for the double pendulum and implement it in Python. Compare pendulum behavior to other available simulations in order to make sure that the behavior is accurate. **Estimated Deadline**: **Apr 5**
- Animate the model using available Python libraries. **Estimated Deadline**: **Apr 5**

- Set up Gym Environment and documentation. **Estimated Deadline**: **Apr 9**
- Practice Reinforcement Learning on a more basic problem. **Estimated Deadline**: **Apr 9**
- Decide on the network architecture for the model. **Estimated Deadline**: **Apr 16**
- Decide on a reward system for the model. **Estimated Deadline**: **Apr 21**
- Decide on a completion condition. **Estimated Deadline**: **Apr 16**
- Decide on a fail condition. **Estimated Deadline**: **Apr 16**
- Train a model that is able to keep both pendulum in upright position without any disturbances. **Estimated Deadline**: **Apr 23**
- Train a model that is able to keep both pendulums in upright position in the presence of disturbances. **Estimated Deadline**: **Apr 27**
- Finalize documentation. **Estimated Deadline**: **May 5**
- If there is enough time, train different models to balance the pendulums in the other equilibrium points.
- If there is enough time, train a model that can utilize swing-up and stay in balance.
- If there is enough time, write a demo in Jupyter Notebook so that other students can run it easily.

5.2 Important dates

- Mar 29 - Hand in project plan.
- Apr 22 - Feedback seminar 1 on the modeling and design.
- May 5 - Report should be pushed to git to allow peer review by other groups.
- May 11 - Peer review done
- May 12 - Feedback seminar 2 on the design and implementation.
- May 20 - Project done and demonstrated and final report handed in
- May 27 - Demo film upload and peer review of final report done
- June 3 - Final presentation and demonstration and revised final report handed in

5.3 Gantt Chart

The project plan can be seen in estimated duration in days, plotted in Figure 1

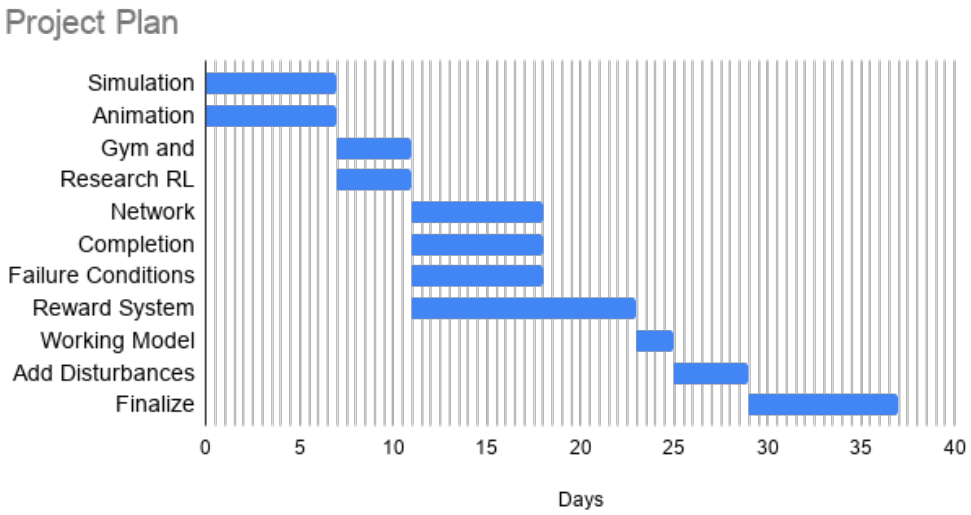

Figure 1: The Gantt chart showing the work flow.