Reinforcement Learning

Exercise 0

Submitted by:
Mostafa Hussein
Miray Yuce
Arlette Perez

1) Introduction to RL

A Reinforcement Learning agent may include one or more of these components: Policy, Value function and/or Model.

Model

An agent is connected to its environment via perception and action. The agent mimics the behaviour of the environment (not the environment itself). On each step of interaction, the agent receives as input the current state of the environment. Then, the agent chooses an action by prediction. This action modifies the state of the environment. Therefore, a model predicts what the environment will do next. The agent learns the behaviour of the environment; thus, maps a plan.

There are two types:
- Transitions model: the agent learns to predict the next state.
- Rewards model: it learns to predict the next reward.

Policy

The policy represents the agent's behaviour function. The action selection is represented as a map, which goes from state to action. The policy is learned from experience.
There are two types:
- Deterministic policy ☐ maps from state to action:
- Stochastic policy ☐ which is the probability of taking an action given a state

Value-function

It is a prediction of future reward. It is important how much reward the agent will obtain for evaluating how well is the agent doing in each state and/or action. The agent chooses the next movement between states by selecting based on the corrected bases. For example: the agent will move to the next state, if the reward is better than any other state.

Two types:
- State value function: It expresses the expected value following a policy in a certain state.
- Action value function: It expresses the expected value following a policy using an action in a certain state.