# IT21479632 - Wijekoon W.M.K.G.S.S.B

# Deep Learning - SE4050

# Lab 08

## Question 01

## Markov\_Decision\_Process

A screen shot of a computer

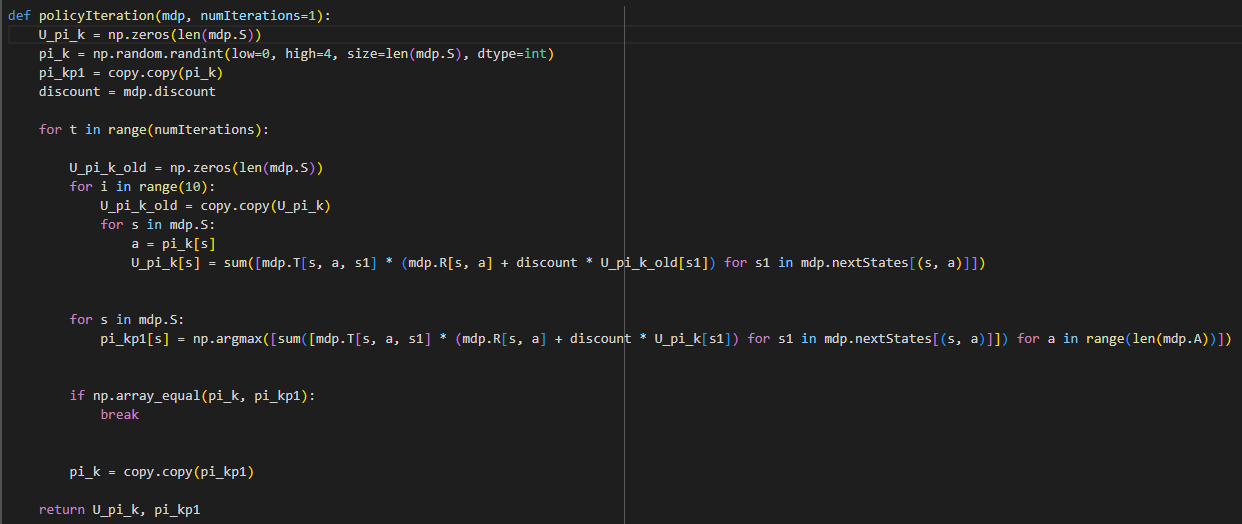
Description automatically generated

**A computer screen shot of a black screen

Description automatically generated**

**A black background with colorful lines

Description automatically generated**

****

**GridWorld**

**A screen shot of a computer program

Description automatically generated**

Increasing The Grid size and comparing the execution and time to converge

**A screen shot of a computer program

Description automatically generated**

Initial Execution Graph  
A screen shot of a graph

Description automatically generated

For 12 x 12 Grid

A screen shot of a graph

Description automatically generated

8 x 8

A graph with blue lines

Description automatically generated

12 x 12

A screen shot of a graph

Description automatically generated

## Question 02

2)

In reinforcement learning, Model-Based and Model-Free algorithms represent two different approaches to solving decision-making problems. **Model-Based algorithms** rely on having access to a model of the environment, which includes knowing the transition probabilities (i.e., how likely it is to move from one state to another given a specific action) and the reward function for each state-action pair. With this model, algorithms like **Policy Iteration** and **Value Iteration** simulate the effects of different actions to compute the expected utility of states and optimize the policy. These methods essentially **"think before they act"** by using the environment's dynamics to make informed decisions. The advantage of Model-Based approaches is that they often converge more quickly to an optimal solution, as they can efficiently plan ahead. However, their limitation lies in the fact that they require an accurate model of the environment, which may not always be available or feasible to construct in complex real-world scenarios.

On the other hand, **Model-Free algorithms** do not rely on an explicit model of the environment. Instead, they learn directly from experience by interacting with the environment and observing the outcomes of their actions. Methods like **Q-Learning** and **SARSA** update their policy based on the rewards received during these interactions, without needing to know the transition probabilities or reward functions beforehand. Model-Free algorithms **"learn by doing"**, making them more flexible for environments where the model is unknown or too complex to define. However, this flexibility comes at the cost of slower learning, as these algorithms require many iterations of trial and error to converge to an optimal policy.

3)

A screenshot of a graph

Description automatically generated

Model-Free Approach

A screenshot of a computer

Description automatically generated

## Question 03

DQN:

A screen shot of a graph

Description automatically generated

A graph of a chart

Description automatically generated with medium confidence