

MSc in Artificial Intelligence

Machine Learning

**Q-Learning**

Instructor: G. Vouros

|  |
| --- |
| Goulas Thodoris |
| Stavrianos Konstantinos |

# Introduction

In this exercise, an auction game between two players-agents will be implemented. Each agent is able to choose between two actions A, D. Agents actions are completely independent. After choosing their actions, the environment will return the agents the respective rewards based on the following table:

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Column Agent | |
| Α | D |
| Row Agent | Α | a,b | d,c |
| D | c,d | b,a |

where a,b,c,d are randomly selected so that a,b > c,d.

Row agent will receive the right reward and column agent will receive the left one. Agents will be trained to act based on ε-greedy Q-learning algorithm, in a predefined number of episodes. Each episode will be composed by a total of 20 moves per agent.

# Methodology

Two separate environments were created per agent. Each environment will be consisted by 4 states according to the above-mentioned table. Environment corresponding to row agent will return the right rewards to the respective agent, while column agent environment will return the left ones. Each environment will choose next state based on agents’ actions.

A total number of 150 episodes will be conducted. Since, ε-greedy q-learning algorithm will be implemented, exploration rate (ε) will be initialized to 1.0 and will be reduced by 0.01 per episode, meaning that in the final 50 episodes no exploration will occur, studying agents’ behavior in a stable equilibrium. Finally, a discount factor gamma = 0.9 will be used.

# Results

To do

# GitHub repository

[auction-qlearning](https://github.com/teogoulas/auction-qlearning.git)