U.PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

IART

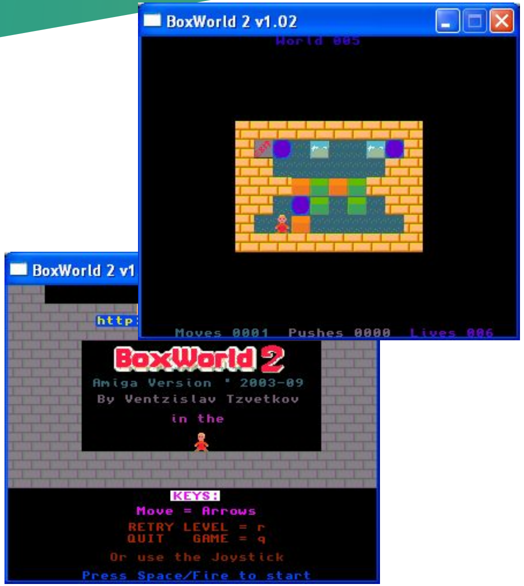
Class 5 - 1A
Box World 2

António Dantas - up201703878@fe.up.pt

Eduardo Macedo - up201703658@fe.up.pt

Gonçalo Pereira - up201705971@fe.up.pt

Specification



- Our project consists in replicating the famous *Box Word 2* game, as well as implementing an *AI* that can easily beat it. The game consists in a series of puzzles, where the player has to reach the exit of the stage in order to advance to the next level. Various obstacles are placed around the arena, making the goal harder. At some stages, there are holes, which can be filled with the various boxes placed around.

- Our AI will consist in an agent solving the puzzles using **Reinforcement Learning**. The two algorithms we will focus on will be both **SARSA** and **Q-Learning.**

Algorithm implementation

Reason	Reward
Reach the exit	10.0
Block the exit	-10.0
Take a step	-0.1

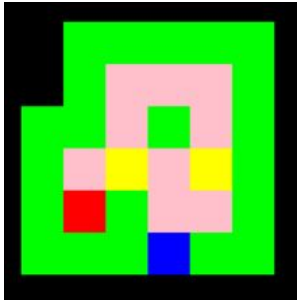
In this project we implemented both **SARSA** and **Q-Learning**. We chose these two algorithms because of the difference in how they will navigate the environment. SARSA, being an on-policy algorithm, will use the same policy all the time, whilst Q-Learning may evaluate the outcome of different one's to choose the best movement decision. It was interesting to analyze both types of results. To evaluate the probability distribution over different actions given our game states we developed the following reward system (table on the left).

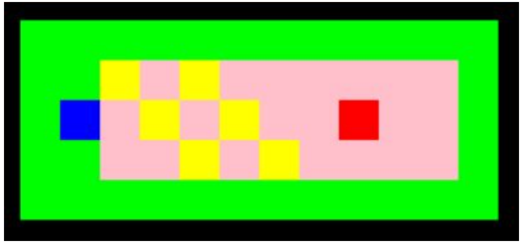
Regarding the libraries we used, we included **matplotlib**(that will help us plot the results of each algorithm and analyze them) and **random**(to solve exploration and exploitation)

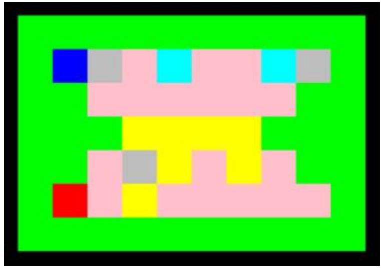
Our Implementation

We have decided use our previous implementation of *Box World 2* of Project 1 which is simplistic representation of the game from the ground up using Python.

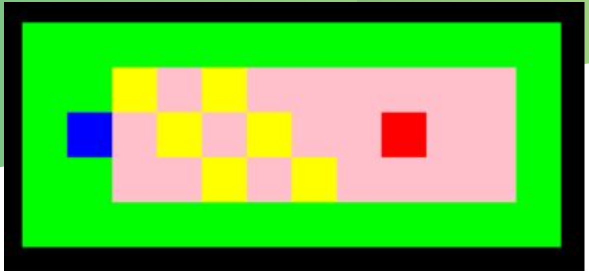
We have eight working levels, some of them are displayed in the images. The red square represents the player that is trying to reach the blue square (exit). The yellow squares represent boxes that can be pushed around to achieve our goal. The light blue squares represent the Ice Boxes and the grey squares represent holes. These holes can be filled up with any type of box making it possible for the player to move to this position.



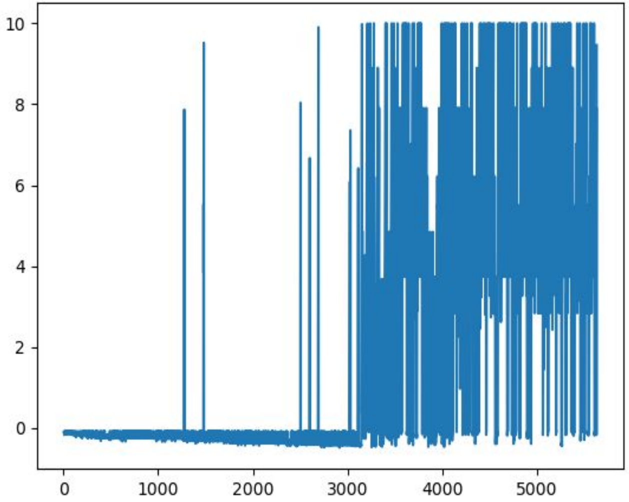


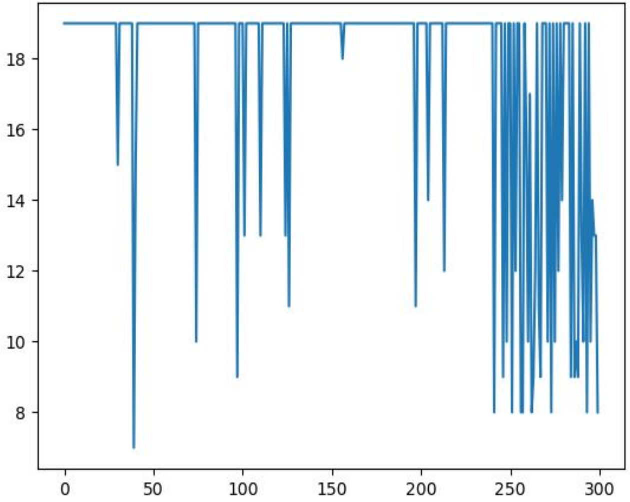


Experimental Evaluation

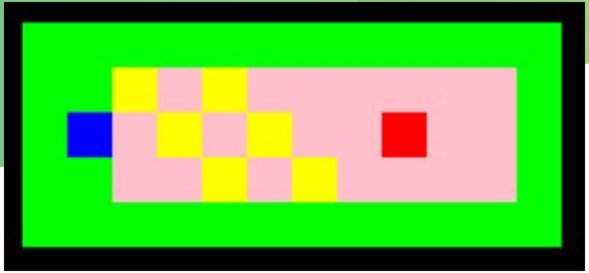


Level 2

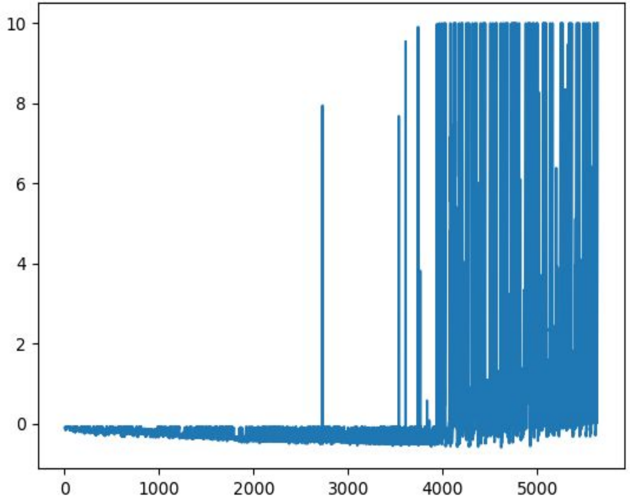


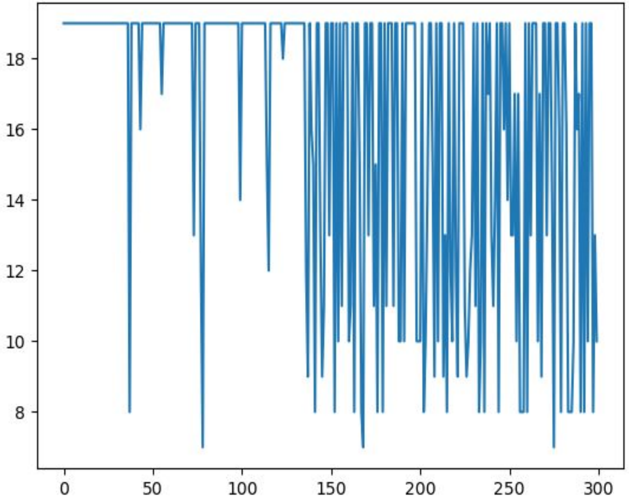


Experimental Evaluation (cont.)

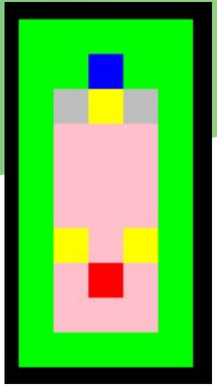


Level 2

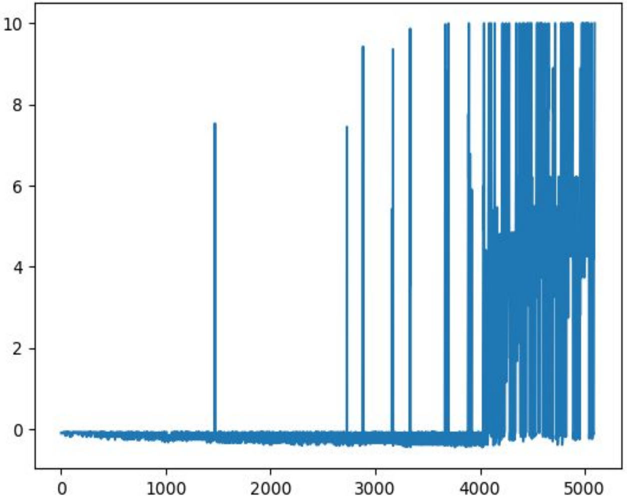


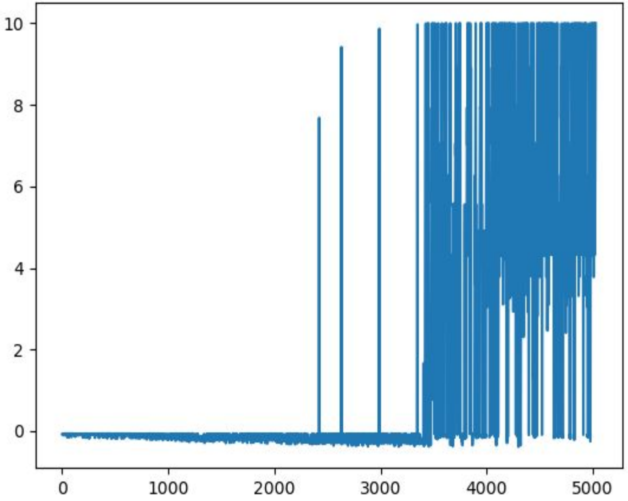


Experimental Evaluation (cont.)



Level 3





References

- Code bullet (Q learning) - https://www.youtube.com/watch?v=-NJ9frfAWRo
- Code bullet (Q learning) - https://www.youtube.com/watch?v=r428O_CMcpI
- QLearning demo - https://github.com/llSourcell/q_learning_demo
- StackOverflow - https://stackoverflow.com/questions/6848828/what-is-the-difference-between-q-learning-and-sarsa
- BoxWorld 2 - http://hirudov.com/others/BoxWorld2.php
- Wikipedia Q-Learning - https://en.wikipedia.org/wiki/Q-learning