

# REINFORCEMENT LEARNING - Bonus

deboshre - 50291550

December 5, 2018

## 1 Implementation

In the bonus part, we have implemented the DQN based reinforcement learning technique in a different environment: CARTPOLE. We will see how the system behaves in the environment.

### 1.1 Cartpole Environment: bonus1.ipynb

CartPole is one of the simplest environments in OpenAI gym (a game simulator). The goal of CartPole is to balance a pole connected with one joint on top of a moving cart. Instead of pixel information, there are 4 kinds of information given by the state, such as angle of the pole and position of the cart. An agent can move the cart by performing a series of actions of 0 or 1 to the cart, pushing it left or right.

The implementation of cartpole is similar to the tom and jerry grid environment given to us though there are some significant differences. As the environment is pre-implemented, there is no need to define the moves of the agent and how it earns rewards. We have trained our model over 1000 episodes keeping the hyper-parameter values as follows:

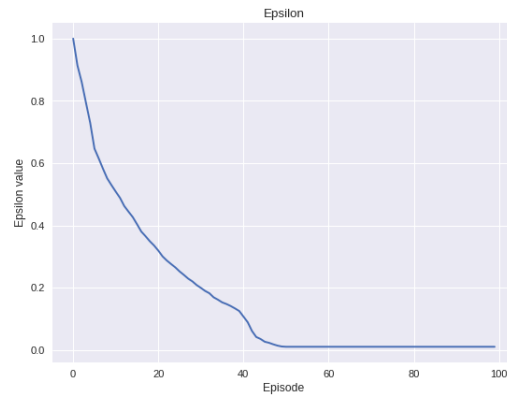
```
self.gamma = 0.95
self.epsilon = 1.0
self.epsilon-min = 0.01
self.epsilon-decay = 0.995
self.learning-rate = 0.001
```

The DQN model implemented has 3-layers with 2 hidden-layers and the activation as RELU-RELU-LINEAR, keeping Adam as the optimizer. We will have one input layer that receives 4 information. But we are going to have 2 nodes in the output layer since there are two buttons (0 and 1) for the game. The model goes random for some 50 iterations.

```

episode: 0/500, score: 40, e: 0.96
episode: 1/500, score: 18, e: 0.88
episode: 2/500, score: 12, e: 0.83
episode: 3/500, score: 12, e: 0.78
episode: 4/500, score: 23, e: 0.69
episode: 5/500, score: 13, e: 0.65
episode: 6/500, score: 47, e: 0.51
episode: 7/500, score: 18, e: 0.47
episode: 8/500, score: 9, e: 0.45
episode: 9/500, score: 10, e: 0.43
episode: 10/500, score: 13, e: 0.4
episode: 11/500, score: 12, e: 0.38
episode: 12/500, score: 12, e: 0.35
episode: 13/500, score: 9, e: 0.34
episode: 14/500, score: 9, e: 0.32
episode: 15/500, score: 9, e: 0.31
episode: 16/500, score: 8, e: 0.3
episode: 17/500, score: 13, e: 0.28
episode: 18/500, score: 8, e: 0.27
episode: 19/500, score: 9, e: 0.26
episode: 20/500, score: 9, e: 0.24
episode: 21/500, score: 9, e: 0.23

```



## 1.2 Epsilon decay with number of episodes

## 2 References

<https://www.safaribooksonline.com/library/view/reinforcement-learning-with/9781788835725/15cab92f-50a8-4114-8942-906c80ef93f1.xhtml>