ÉCOLE
D'INGÉNIEURS
DE L'AIR ET DE L'ESPACE

Introduction to Machine Learning

From a childhood game to reinforced learning

Authors :
Matthieu **Thomas**
Julien **Huynh**

Supervised by :
Mme **Frontera-Pons**

Version 0.1
April 9, 2020

Contents

Chapter 1

The legendary PinBall 3D game

Back in the old days when multiplayer gaming and unlimited internet connection was not as widespread, as kids, some of us eventually turned to a Windows pre-installed game, PinBall 3D.

1.1 Objective

Score as many points as possible

Chapter 2

Reinforcement learning

2.1 Why reinforcement learning ?

While thinking about the way to use Machine Learning to play PinBall, we thought of various possibilities but the one that stood out from the rest was reinforcement learning. The advantage of this method is that it works well when there are various states through time which is the case in a game of Pinball. Moreover, we can fairly easily say that the reward involved in this method will be related to the score and the lives.

2.2 Basics of reinforcement learning

2.2.1 Concepts

Simply put, reinforcement learning is a method that takes advantage of the interactions between :

- Agent
- Actions
- Rewarding system
- Environment
- States
- Policy
- Value

Other concepts are also used but the ones mentioned above are the most essential ones.

2.2.2 Main idea

In reinforcement learning methods, the agent does a set of actions in order to maximize the reward it will get. At each given time, the agent will compute an action being given the state and this will then have a reaction from the environment surrounding it as a feedback which is the reward and will put it in a new state.

One representation of some of those interactions is the following loop :

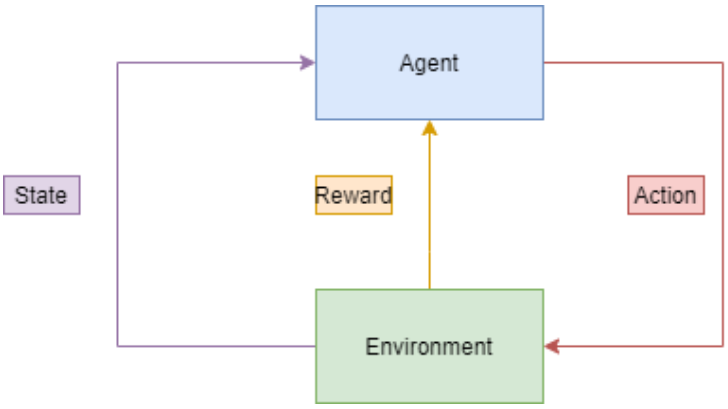

Figure 2.1: Reinforcement learning concept interactions

On this loop, at a given time t, the agent knows the state s_t and the reward r_t at time t and then computes an action a_t with respect to s_t and r_t. Then, the environment will give a new state s_{t+1} and a new reward r_{t+1} as a response to a_t. Obviously, the better the action taken was with respect to the objective, the higher the reward will be.

2.2.3 Policy, Value and Trajectory

While they did not appear in the loop, the trajectory, policy and value are essential in reinforcement learning.

Policy

The policy is the strategy the agent will use to reach its objective. If we refer to path-finding, it could be the different ways to go from a starting point to the objective point using different algorithms (Dijkstra, A*, RRT, potential field, ...). Simply put, it defines how the agent will choose its action to have an impact on the environment.

The learning part is done at the **Agent** part which, given the state and the previous reward, will learn the best policy to maximize the value.

Value

In reinforcement learning, the value is basically the long term reward, it is what we would expect as an overall reward while being in a certain state and following a certain policy. The difference with the reward is that the latter is instantaneous, meaning that it is simply a response at a given time while the value is a much more global concept. Mathematically, it is defined as :

$$V(s) = \mathbb{E}\left[\sum_n \gamma^n r_n\right]$$

With n the states, r_n the reward for the state and $\gamma \in [0, 1]$ a kind of damping factor for the reward, called discount factor. This factor helps giving a balance in the importance between the current reward and the ones to be expected in the "future". This is important in some situations when for example, an action is taken, giving a high reward at the given time but will also greatly reduce the expected future rewards thus, the value.

Trajectory

Roughly speaking, the trajectory τ is the set of states the agent went through after interacting with t

2.3 Reinforcement learning methods

Chapter 3

Reinforcement learning - Application to PinBall 3D

blabla applicatif + git