DDPG + HER in FetchPush-v1 Gym Environment with Python3 and Tensorflow 2.0

Simone De Angelis 1760464
Veronica Romano 1580844

January 5, 2021

Chapter 1

Introduction and Work Purposes

Reinforcement Learning is learning what to do and how to map situations to actions. The end result is to maximize the numerical reward signal. The learner is not told which action to take, but instead must discover which action will yield the maximum reward. As illustrated in Figure ?, an agent, which is who have to reach a goal, stars from a state s_t and through actions a_t, manipulates the environment. From this it takes a reward r_t+1, if goes more close to the goal and it's intended to encourage good agent behavior, and discovers the next state s_t+1. These two are the reward and the state that permits the agent to re-implement the cycle. This proceeds until the goal is reached. Reinforcement Learning also differs from the other machine learning techniques. In fact in supervised learning there is a "supervisor" which has the knowledge of the environment and shares it with the agent to complete the task. while in reinforcement learning is the agent that learns from its own experience. Moreover there is a reward function which acts as a feedback to the agent. With respect to unsupervised learning where there is no mapping from input to output, in reinforcement learning the mapping is present. This type of learning is usually modeled through Markov Decision Process. If the environment is non-deterministic we need a transition model $P(s'|s, a)$ and the utility function is the sum of the received rewards.

This work consists in implementing the FetchPush-v1 environment provided by Gym OpenAI, and using a specific algorithm, which in our specific case is DDPG+HER, experimenting on this environment. Also a specific framework will be assigned for the implementing the code. All the result reached will be illustrated in Chapter 3. While in the following sections we will exploit in details the algorithm used for our task, the environment, and the implementation framework.

Chapter 2

Implementation Assignment

2.1 DDPG + HER Algorithm

2.2 Gym Environment

The release provided by OpenAI Gym, contains four environments using the Fetch research platform. These environments use the MuJoCo physics simulator. Our environment is, as previously mentioned, FetchPush-v1. Fetch goal is to move a box by pushing it until it reaches a desired goal position. This tasks has the concept of a "goal". By default it uses a sparse reward of -1 if the desired goal was not yet achieved and 0 if it was achieved (within some tolerance).

FIGURE

2.3 Implementation Framework

Chapter 3

Experiments and Results