ARTIFICIAL INTELLIGENCE

FALL 2018-2019 ASSIGNMENT #3

Onat Şahin - 150150129

As a problem, I chose the fetch, pick and place problem. In this problem, the robotic arm is required to pick up a block and carry it to a randomly selected goal point in 3d space.

a) LITERATURE REVIEW

As I was searching for papers that tackle the problem of fetch, pick and place, or a similar problem, I realized that people who does research in this area are taking fetching robotic arm problems as a whole. Since fetch pick and place, fetch push, fetch reach and fetch slide problems in the OpenAI Gym are similar, researchers are trying to develop models that can be adapted to any of these problems.

The first paper that caught my eye was the paper by Andrychowicz et al. from OpenAI called "Hindsight Experience Replay"[1]. This paper first goes through classic reinforcement techniques like trying to obtain the Q*, which is the optimal Q-function. It follows this with deep Q-networks, which are neural networks used to approximate the Q*, since Q* is really costly to calculate. After this, the paper argues that sparse and binary rewards should be used since complicated reward functions are costly to come up with and are domain specific. However, models with binary rewards are very hard to train since they do not give any information other than whether the agent has reached the goal. To be able to achieve a good reinforcement learning model with sparse and binary rewards, the paper proposes the approach called "Hindsight Experience Replay", or HER. This approach replays each trajectory of the agent during the learning process while changing its goal from the true goal state to the final state reached in that trajectory. This way, the agent gets rewards and learns to visit some states even though those states may not be the true goal. Learning to reach different states eventually helps the agent learn to achieve the true goal way quicker than the classic way.

In the paper "Multi-Goal Reinforcement Learning: Challenging Robotics Environments and Request for Research" by Plappert et al.[2], Hindsight Experience Replay is used to solve every problem in OpenAI Gym, instead of just fetch problems. For fetch problems, sparse and binary rewards are used. The actions are 4 dimensional: x,y,z coordinates for movement and a bit for gripper being open or closed. The goal is three dimensional, being the desired coordinates of the object. This paper also experiments with dense rewards and models that do not use HER. As a result, they state that using HER with sparse rewards gives the best results.

[1] M. Andrychowicz et al., "Hindsight Experience Replay", 2018

[2]M. Plappert et al., "Multi-Goal Reinforcement Learning: Challenging Robotics Environments and Request for Research", 2018.

b) FETCH, PICK AND PLACE MODEL

The objective is to carry the fetched block to the target location. I would model my problem similar to how it is modeled in papers I have cited.

STATES: Possible locations of the arm.

ACTIONS: x,y,z coordinates and gripper being open or close.

GOAL: Target location where the block will be dropped.

REWARDS: 1 if the goal is reached, 0 if the goal is not reached.

c) ALGORITHMS EMPLOYED

I would implement a Hindsight Experience Replay algorithm to establish intermediary goals so that the agent can learn even with binary rewards. I would also use a neural network to estimate Q* values that satisfy the Bellman equation.

d) ADVANTAGES AND DISADVANTAGES

Binary rewards makes the reinforcement learning model really adaptable to similar different problems. Also, it makes the model more scalable. HER enables the model to learn with only binary rewards.

Even though this approach is more successful compared to the classic reinforcement learning approaches, it is harder to implement. In addition to the classic algorithms, Hindsight Experience Replay and a neural network must be implemented. Furthermore, the neural network must be trained, which requires datasets for training and testing.