Chapter 70
Deep Reinforcement Learning for Optimization

Md Mahmudul Hasan
Anglia Ruskin University, UK

Md Shahinur Rahman
https://orcid.org/0000-0002-9913-2454
Daffodil International University, Bangladesh

Adrian Bell
Anglia Ruskin University, UK

ABSTRACT

Deep reinforcement learning (DRL) has transformed the field of artificial intelligence (AI) especially after the success of Google DeepMind. This branch of machine learning epitomizes a step toward building autonomous systems by understanding of the visual world. Deep reinforcement learning (RL) is currently applied to different sorts of problems that were previously obstinate. In this chapter, at first, the authors started with an introduction of the general field of RL and Markov decision process (MDP). Then, they clarified the common DRL framework and the necessary components RL settings. Moreover, they analyzed the stochastic gradient descent (SGD)-based optimizers such as ADAM and a non-specific multi-policy selection mechanism in a multi-objective Markov decision process. In this chapter, the authors also included the comparison for different Deep Q networks. In conclusion, they describe several challenges and trends in research within the deep reinforcement learning field.

INTRODUCTION

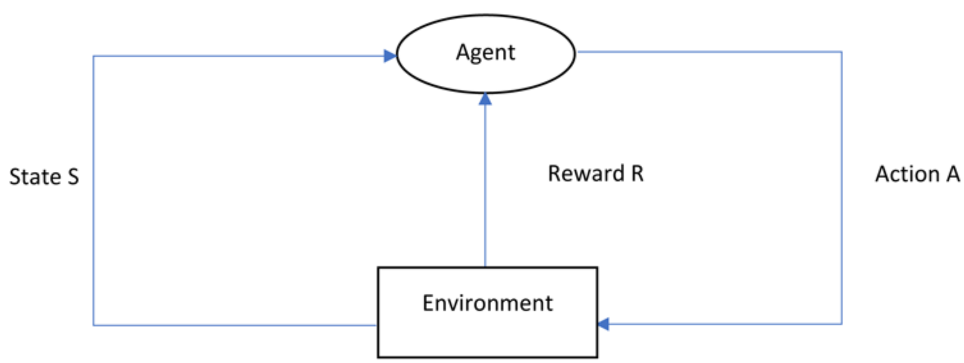
Nowadays it is more important to train up a machine and interact with environment to determine potential behaviours. Deep reinforcement learning is a powerful and most usable technique for communicate between an agent and an environment. Reinforcement Learning is a technique to understand how an agent can communicate with the environment and find out which action is best based on every step by trial

DOI: 10.4018/978-1-7998-7705-9.ch070

and error (H. Li, Wei, Ren, Zhu, & Wang, 2017). In machine learning there are three main categories which are supervised learning, Unsupervised Learning and Reinforcement Learning. In this chapter, we are given an overview about one of the most exciting topic of Machine learning is Reinforcement Learning. It is more important to find out the best solution of a problem but most of this process it's so much difficult to find the exact solution without any reaction. RL can take a decision which action is best and how can an agent learn behaviour in environment by action and seeing result. To overcome this problem one of most important step is using optimization with RL. Optimization is movement process to take the best compromising solution based on a set of all possible solution reduce leftover. Let think about a robots movement. A robot may take a long step in front and it can fall. Again the robot may take a short step and can hold balanced ("Reinforcement learning explained - O'Reilly Media," 2016). So here using RL we can get a possible solution set based on environment and from the possible solution set using optimization we can extract the best compromising solution. If we go through the definition of Reinforcement Learning then we can say that in initial position a robot doesn't know anything but when train up the robot how to walk, take action, keep balanced based on Environment this is called the Reinforcement Learning.

The following Figure 1 shows a reinforcement learning model where an agent takes action for an environment for each states and earned some rewarding points (Zoltán Gábor, Zsolt Kalmár, & Csaba Szepesvári, 1998).

Figure 1. Reinforcement learning model



There are several ways to solve the control or sequential decision making problems using reinforcement learning techniques. They are as follows (Watkins & Dayan, 1992):

1. Markov Decision Process
2. Dynamic Programming
3. Temporal difference learning
4. Q Learning
5. Deep Learning
6. Monte-Carlo Tree Search (MCTS)

Related Content