

Experience Replay Makes Others Become Flatterers

Kunfeng Li
College of Intelligence and Computing
Tianjin University
TJ, 300350 PRC
a1070033289@qq.com

Abstract

ChuiPeng is a Chinese word meaning to flattery, originating from "The Will of the Chief." Some people also call it Guli, Chuibi, etc. Getting others' flattery is good for our mood, and we all want to maximize the flattery we get. I proposed a more-flattery approach to obtaining more flattery. I have found that with the help of the experience replay mechanism in deep reinforcement learning, we can get more flattery and significantly reduce the training time.

1. Introduction

We have found that once people get others' flattery, they will feel more happy and work harder. Therefore, some leaders with bad intentions will further squeeze the value of employees through appropriate praise, and the efficiency of employees will be improved at the cost of almost zero. But we don't consider this situation here. Our goal is to get as much flattery as possible, in addition to its help to physical and mental health, the pleasure can help us alleviate hair loss. So we all want to maximize the flattery we got.

Experience replay is the main practice of DQN, which stores the data obtained by the system exploration environment, and then update the parameters of the deep neural network in virtue of random sampling. The advantages are as following:

- 1) High data utilization because one sample is used multiple times.
- 2) The correlation of consecutive samples will make the variance of the parameter update larger, and this mechanism can reduce this correlation.

2. Model

This section will penetrate into my more-flattery approach.

2.1. Problem Destination

The process of getting flattery is defined as Markov Decision Process (MDP). The environment is scenarios where the agent is revealing itself, such as group presentations, evaluation of scholarship presentations, etc. State S represents the silhouette of the environment at each moment. The action set includes the intonation, posture and expression of the agent. Reward is others' flattery. Each person's single flattery is equivalent. Because in general people won't speak together at the same time, the training goal is to maximize the number of pieces of flattery the agent got at the end of a particular scene.

2.2. More-Flattery Approach

The way to add experience replay is to record your own performance process and then store it on the hard disk. Just like dancers record their own dance process and chase down their own shortcomings through video.

The Agent finds strategies for gathering more flattery by randomly observing the video. In this way, you won't agonize about unlearning the performance process and details that you can't engrave, which improves the data utilization. In addition, when you vividly explain a concept, some people boast that you are too humorous, some people boast that you understand too well, then you may fall into chaos, which is more important. But if you have experience replay, you can observe past records and find the apposite priorities. Some conceptual humor is more important, while others are more important to speak clearly.

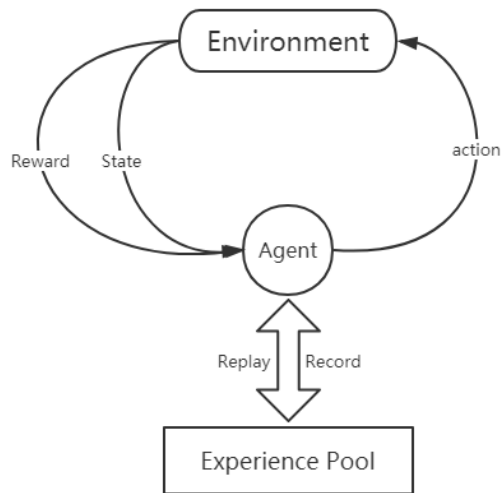


Figure 1. More-Flattery Approach - model architecture.

3. Conclusions

Record life.