

عنوان مقاله:

Actor Double Critic Architecture for Dialogue System

محل انتشار:

مجله نوآوری های مهندسی برق و کامپیوتر، دوره 11، شماره 2 (سال: 1402)

تعداد صفحات اصل مقاله: 10

نویسندگان:

.Y. Saffari - *Department of Electrical and Computer Engineering, University of Kashan, Kashan, Iran*

.J. Salimi Sartakhti - *Department of Electrical and Computer Engineering, University of Kashan, Kashan, Iran*

خلاصه مقاله:

Background and Objectives: Most of the recent dialogue policy learning methods are based on reinforcement learning (RL). However, the basic RL algorithms like deep Q-network, have drawbacks in environments with large state and action spaces such as dialogue systems. Most of the policy-based methods are slow, cause of the estimating of the action value using the computation of the sum of the discounted rewards for each action. In value-based RL methods, function approximation errors lead to overestimation in value estimation and finally suboptimal policies. There are works that try to resolve the mentioned problems using combining RL methods, but most of them were applied in the game environments, or they just focused on combining DQN variants. This paper for the first time presents a new method that combines actor-critic and double DQN named Double Actor-Critic (DAC), in the dialogue system, which significantly improves the stability, speed, and performance of dialogue policy learning. **Methods:** In the actor critic to overcome the slow learning of normal DQN, the critic unit approximates the value function and evaluates the quality of the policy used by the actor, which means that the actor can learn the policy faster. Moreover, to overcome the overestimation issue of DQN, double DQN is employed. Finally, to have a smoother update, a heuristic loss is introduced that chooses the minimum loss of actor-critic and double DQN. **Results:** Experiments in a movie ticket booking task show that the proposed method has more stable learning without drop after overestimation and can reach the threshold of learning in fewer episodes of learning. **Conclusion:** Unlike previous works that mostly focused on just proposing a combination of DQN variants, this study combines DQN variants with actor-critic to benefit from both policy-based and value-based RL methods and overcome two main issues of both of them, slow learning and overestimation. Experimental results show that the proposed method can make a more accurate conversation with a user as a dialogue policy learner.

کلمات کلیدی:

Actor-Critic, Dialogue system, DQN, Actor Double Critic

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1681168>



