Research Reports on Computer Science
https://ojs.wiserpub.com/index.php/RRCS/

Editorial

A Note on Reinforcement Learning

Ying Tan

Department of Machine Intelligence, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China
E-mail: ytan@pku.edu.cn

Received: 10 November 2022; **Accepted:** 10 November 2022

In the past decade, deep reinforcement learning (DRL) has drawn much attention in theoretical research, meanwhile, it has seen huge success across multiple application areas, such as combinatorial optimization, recommender systems, autonomous driving, intelligent healthcare system and robotics. As one of three basic machine learning paradigms, reinforcement learning concerns with how intelligent agents learn in an interactive environment through trial and error to maximize the total cumulative reward of the agents. Even though many progresses of reinforcement learning have been presented, there are still many challenging research topics due to the complexity of the problems.

Intrinsic reward designing. As one of the essential problems of reinforcement learning, the efficiency of exploration greatly affects the performance of RL algorithms. Curiosity-driven intrinsic reward encourages an agent to explore spaces that are less visited. Curiosity-based reward schemes provide powerful exploration mechanisms which enable the discovery of solutions for complex, sparse or long-horizon tasks [1]. Nonetheless, while the agent learns and evolves to unfamiliar spaces, the learning algorithm autonomously shifted the objective to reward new areas, many behaviors emerge and later diminish as it is being overruled by the continual change of objective. An intrinsic reward with higher shift adaptability is one of the most challenging topics.

Offline reinforcement learning. One of the major hindrances to the adoption of reinforcement learning algorithms is due to its nature which implements a fundamentally online learning paradigm. The process of reinforcement learning involves iteratively collecting experience by interacting with the environment, normally with the latest learned policy, and then using that experience to improve the policy [2]. This online interaction is often impractical in many real-world tasks. In those feasible settings, pre-collected human-level data can gratefully help policy and avoid invalid exploration. Offline reinforcement learning studies how an agent learns from human behaviors. Developing a data-driven reinforcement learning paradigm could inspire a new era of reinforcement learning.

Multi-agent reinforcement learning. A multi-agent system is closer to reality as many real-world problems can be abstracted as a multi-agent system. Meanwhile, there are many multi-agent system-specific problems, such as non-stationarity, communication, coordination, scalability, partial observability, credit assignment and curse of dimensionality. Therefore, multi-agent reinforcement learning is a topic worth studying. There are some trends targets in state-of-the-art approaches, curriculum learning, memory, communication, and training paradigm.

Sequence modeling. This is a new perspective of offline reinforcement learning. Recent work has shown that Transformers can model large-scale distributions of semantic concepts. Sequence modeling utilizes transformer architecture to model human-level behaviors and solves long-term credit assignment tasks. Decision Transformer (DT) shows great potential for solving complex tasks which Temporal Difference (TD) learning is not good at, and has

DOI: https://doi.org/10.37256/rrcs.1220222153

attracted a lot of researchers' interest.

Research Reports on Computer Science (RRCS) mainly reports on innovative research results that cover novel theories, technologies, and engineering applications in the fields of computer science and engineering. DRL is one of the top hot-spots of artificial intelligence, it has great potential to solve a bunch of real-world problems and even may become one of the keys to solving artificial general intelligence. In the near future, RRCS will increase attention in researches on reinforcement learning. You are welcome to present recent advances in your research on DRL and related researches on RRCS.

References

[1] Groth O, Wulfmeier M, Vezzani G, Dasagi V, Hertweck T, Hafner R, et al. Is curiosity all you need? On the utility of emergent behaviours from curious exploration. *ArXiv* [Preprint] 2021. https://doi.org/10.48550/arXiv.2109.08603

[2] Sutton RS, Barto AG. *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press; 1998.