Bai. *Complex Eng Syst* 2024;4:8
DOI: 10.20517/ces.2024.18

Complex Engineering Systems

Editorial

Introduction to discrete-time reinforcement learning control in Complex Engineering Systems

Weiwei Bai

Navigatioin College, Dalian Maritime University, Dalian 116026, Liaoning, China.

Correspondence to: Prof. Weiwei Bai, Navigation College, Dalian Maritime University, No. 1 Linghai Road, Dalian 116026, Liaoning, China. E-mail: baiweiweidl@dlmu.edu.cn

How to cite this article: Bai W. Introduction to discrete-time reinforcement learning control in Complex Engineering Systems. *Complex Eng Syst* 2024;4:8. https://dx.doi.org/10.20517/ces.2024.18

Received: 22 Apr 2024 **Accepted:** 25 Apr 2024 **Published:** 28 Apr 2024

Academic Editor: Hamid Reza Karimi **Copy Editor:** Fangling Lan **Production Editor:** Fangling Lan

Within the context of the Complex Engineering Systems (CES), this editorial aims to describe the recent progress of discrete-time reinforcement learning (RL) control.

Considering the widespread application of the digital computer in the CES, which process the data in discrete-time form, and the nonlinearity inherent in the engineering, the nonlinear discrete-time control design has garnered growing attention in the modern control engineering. For instance, when the backstepping method is employed to design the controller for a nonlinear discrete-time system, it may suffer from a noncausal problem. As the authors claimed, the causality contradiction problem may arise[1], where the future signal is embedded in the current control signal, leading to controller design failure. To solve this problem, various system transformations were developed, making them one of the perennial topics for nonlinear discrete-time systems.

The control design is one of the most important topics of CES; as an optimal control strategy, RL has received increasing attention. It can not only make a compromise between the control cost and performance but also decrease the impact of external environment through continuous exploration[2]. Many RL techniques exist, such as Q-learning, adaptive dynamic programming (ADP), policy iteration, *etc*. Among them, actor-critic is one of the classical RL techniques, which is simple and easy to apply. However, it should be noted that the gradient descent method is employed to learn the weight vector that searches for the optimal solution from a single point, and it may easily fall into the local optimal. Therefore, solving the local optimal problem is one of the hot topics in the RL control design. In the following section, the above control topics of research progress will be introduced.

1. SYSTEM TRANSFORMATION

The noncausal problem was first pointed out by Yeh[3] and it was solved using a time-varying mapping technique for the parameter-strict-feedback and parameter-pure-feedback systems. Based on this result, it was further extended to the time-varying parameters and nonparametric uncertainty systems[4]. However, this transformation is inapplicable to a class of more general nonlinear strict feedback systems. To address this issue, Ge *et al.* transformed the nonlinear strict feedback system into a novel sequential decrease cascade form, and the noncausal problem was solved[5]. Notably, the nonlinear function $f_i(\bar{x}_i(k))$ only includes partial states $\bar{x}_i(k) = [x_1(k), \dots, x_i(k)]^{\mathrm{T}}$ (refer to[6] for more details). However, it may incorporate all the states of the control system $\bar{x}_n(k) = [x_1(k), \dots, x_n(k)]^{\mathrm{T}}$, which is the so-called non-strict feedback system that is more general than the strict one. The system transformation in[5] is no longer applicable again. In[7], the discrete-time non-strict feedback system was transformed into a time instant recursive form, requiring all the past information at the current time, which will lead to great difficulty for the application. Subsequently, a universal system transformation was recently devised[8], and the noncausal problem was solved thoroughly. This is the main research progress about the system transformation up to now, and this is a perennial topic.

2. LOCAL OPTIMAL PROBLEM

Another hot topic within the control field is the issue of local optimal. As previously stated, the gradient descent method searches for the optimal solution from a single point, which may easily fall into the local optimal. Therefore, the local optimal problem should be considered in the controller design. Genetic (GA) and evolutionary algorithms can effectively tackle this problem by exploring optimal solutions from multiple points of view. However, the evolutionary algorithms have a heavy computation burden when it has a large population rendering them unsuitable for the online learning. After that, the experience replay technique is developed, involving repeatedly learning the past information, allowing the parameters to converge to the true value[9]. Nevertheless, this technique is a traditional adaptive technique that cannot realize the optimization. Based on the key idea of the experience replay technique, a multi-gradient recursive approach is developed to learn the weight vector and solve the local optimal problem[10]. Consequently, this issue has received much attention in the adaptive control field recently, emerging as a prominent subject in adaptive RL control.

DECLARATIONS

Authors' contributions

The author contributed solely to the article.

Availability of data and materials

Not applicable.

Financial support and sponsorship

This work was supported by the National Natural Science Foundation of China (No. 52271360), the Dalian Outstanding Young Scientific and Technological Talents Project (No. 2023RY031), and the Basic Scientific Research Project of Liaoning Education Department (Grant No. JYTMS20230164).

Conflicts of interest
The author declared that there are no conflicts of interest.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Copyright

REFERENCES
1. Bai W, Li T, Long Y, Chen CLP. Event-triggered multigradient recursive reinforcement learning tracking control for multiagent systems. *IEEE Trans Neural Netw Learn Syst* 2023;34:366-79. DOI PubMed
2. Yang Q, Cao W, Meng W, Si J. Reinforcement-learning-based tracking control of waste water treatment process under realistic system conditions and control performance requirements. *IEEE Trans Syst Man Cybern Syst* 2022;52:5284-94. DOI
3. Yeh PC, Kokotović PV. Adaptive control of a class of nonlinear discrete-time systems. *Int J Control* 1995;62:303-24. DOI
4. Zhang Y, Wen C, Soh YC. Discrete-time robust backstepping adaptive control for nonlinear time-varying systems. *IEEE Trans Automat Control* 2000;45:1749-55. DOI
5. Ge SS, Li GY, Lee TH. Adaptive NN control for a class of strict-feedback discrete-time nonlinear systems. *Automatica* 2003;39:807-19. DOI
6. Li YM, Min X, Tong S. Adaptive fuzzy inverse optimal control for uncertain strict-feedback nonlinear systems. *IEEE Trans Fuzzy Syst* 2020;28:2363-74. DOI
7. Bai W, Li T, Tong S. NN reinforcement learning adaptive control for a class of nonstrict-feedback discrete-time systems. *IEEE Trans Cybern* 2020;50:4573-84. DOI PubMed
8. Bai W, Li T, Long Y, et al. A novel adaptive control design for a class of nonstrict-feedback discrete-time systems via reinforcement learning. *IEEE Trans Syst Man Cybern Syst* 2024;54:1250-62. DOI
9. Modares H, Lewis FL, Naghibi-Sistani MB. Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks. *IEEE Trans Neural Netw Learn Syst* 2013;24:1513-25. DOI PubMed
10. Bai W, Zhou Q, Li T, Li H. Adaptive reinforcement learning neural network control for uncertain nonlinear system with input saturation. *IEEE Trans Cybern* 2020;50:3433-43. DOI