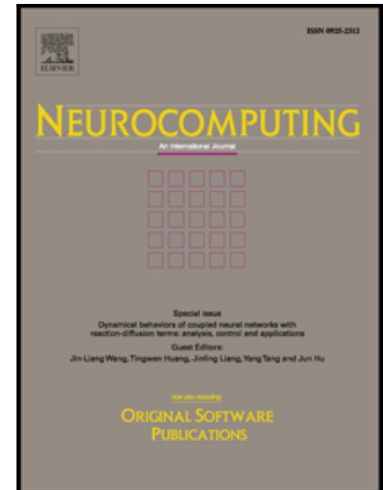


Accepted Manuscript

A New History Experience Replay Design for Model-free Adaptive Dynamic Programming

Naresh Malla, Zhen Ni

PII: S0925-2312(17)30862-7
DOI: [10.1016/j.neucom.2017.04.069](https://doi.org/10.1016/j.neucom.2017.04.069)
Reference: NEUCOM 18443



To appear in: *Neurocomputing*

Received date: 14 February 2017
Revised date: 16 April 2017
Accepted date: 29 April 2017

Please cite this article as: Naresh Malla, Zhen Ni, A New History Experience Replay Design for Model-free Adaptive Dynamic Programming, *Neurocomputing* (2017), doi: [10.1016/j.neucom.2017.04.069](https://doi.org/10.1016/j.neucom.2017.04.069)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A New History Experience Replay Design for Model-free Adaptive Dynamic Programming

Naresh Malla, Zhen Ni

Department of Electrical Engineering and Computer Science

South Dakota State University

Brookings, SD, 57007, USA

Email: (naresh.malla, zhen.ni)@sdstate.edu

Abstract

An adaptive dynamic programming (ADP) controller is a powerful control technique that has been investigated, designed and tested in a wide range of applications for solving optimal control problems in complex systems. The performance of the ADP controller is usually obtained by long training periods because the data usage efficiency is low as it discards the samples once used. History experience, also known as experience replay, is a powerful technique showing potential to accelerate the training process of learning and control. However, the existing design of history experience cannot be directly used for the model-free ADP design, because the existing work focuses on the forward temporal difference (TD) information (e.g., state-action pair). This information is between the current time step and the future time step and will need a model network for future information prediction. This paper proposes a new history experience replay design to avoid the usage of the model network or identifier of the system/environment. Specifically, we designed the experience tuple with one step backward state-action information and the TD can be achieved by a previous time step and a current time step. In addition, a systematic approach is proposed to integrate history experience in both the critic and action networks of the ADP controller design. The proposed approach is tested for two case studies: a cart-pole balancing task and a triple-link pendulum balancing task.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات