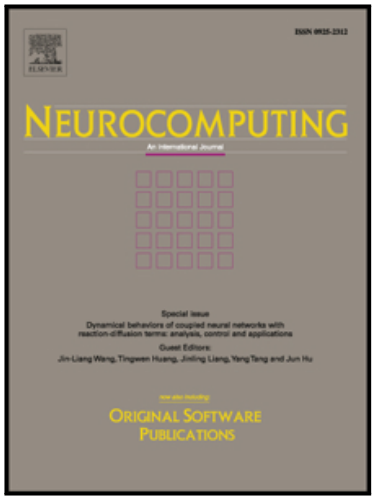
Communicated by Shaocheng Tong

Accepted Manuscript



General value iteration based reinforcement learning for solving optimal tracking control problem of continuous-time affine nonlinear systems

Geyang Xiao, Huaguang Zhang, Yanhong Luo, Qiuxia Qu

PII: S0925-2312(17)30553-2
DOI: 10.1016/j.neucom.2017.03.038
Reference: NEUCOM 18266

To appear in: *Neurocomputing*

Received date: 19 September 2016
Revised date: 4 January 2017
Accepted date: 14 March 2017

Please cite this article as: Geyang Xiao, Huaguang Zhang, Yanhong Luo, Qiuxia Qu, General value iteration based reinforcement learning for solving optimal tracking control problem of continuous-time affine nonlinear systems, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.03.038

General value iteration based reinforcement learning for solving optimal tracking control problem of continuous-time affine nonlinear systems

Geyang Xiao, Huaguang Zhang*, Yanhong Luo, Qiuxia Qu

College of Information Science and Engineering, Northeastern University, Box 134, 110819, Shenyang, P.R. China
The Key Laboratory of Integrated Automation of Process Industry (Northeastern University) of the National Education Ministry, 110004, Shenyang, P.R. China

Abstract

In this paper, a novel reinforcement learning (RL) based approach is proposed to solve the optimal tracking control problem (OTCP) for continuous-time (CT) affine nonlinear systems using general value iteration (VI). First, the tracking performance criterion is described in a total-cost manner without a discount term which can ensure the asymptotic stability of the tracking error. Then, some mild assumptions are assumed to relax the restriction of the initial admissible control in most existing references. Based on the proposed assumptions, the general VI method is proposed and three situations are considered to show the convergence with any initial positive performance function. To validate the theoretical results, the proposed general VI method is implemented by two neural networks on a nonlinear spring-mass-damper system and two situations are considered to show the effectiveness.

Keywords: Adaptive dynamic programming, optimal control, reinforcement learning, continuous-time systems

*Corresponding author
Email addresses: xgyalan@outlook.com (Geyang Xiao), hgzhang@ieee.org (Huaguang Zhang*), neuluo@gmail.com (Yanhong Luo), quqiuxia2010@163.com (Qiuxia Qu)