INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING
Int. J. Adapt. Control Signal Process. (2014)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/acs.2485

Online concurrent reinforcement learning algorithm to solve two-player zero-sum games for partially unknown nonlinear continuous-time systems

Sholeh Yasini*,†, Ali Karimpour, Mohammad-Bagher Naghibi Sistani and Hamidreza Modares

Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad 91775-1111, Iran

SUMMARY

Online adaptive optimal control methods based on reinforcement learning algorithms typically need to check for the persistence of excitation condition, which is necessary to be known *a priori* for convergence of the algorithm. However, this condition is often infeasible to implement or monitor online. This paper proposes an online concurrent reinforcement learning algorithm (CRLA) based on neural networks (NNs) to solve the H_∞ control problem of partially unknown continuous-time systems, in which the need for persistence of excitation condition is relaxed by using the idea of concurrent learning. First, H_∞ control problem is formulated as a two-player zero-sum game, and then, online CRLA is employed to obtain the approximation of the optimal value and the Nash equilibrium of the game. The proposed algorithm is implemented on actor–critic–disturbance NN approximator structure to obtain the solution of the Hamilton–Jacobi–Isaacs equation online forward in time. During the implementation of the algorithm, the control input that acts as one player attempts to make the optimal control while the other player, that is, disturbance, tries to make the worst-case possible disturbance. Novel update laws are derived for adaptation of the critic and actor NN weights. The stability of the closed-loop system is guaranteed using Lyapunov technique, and the convergence to the Nash solution of the game is obtained. Simulation results show the effectiveness of the proposed method.

Received 21 July 2013; Revised 11 December 2013; Accepted 21 February 2014

KEY WORDS: H_∞ control; two-player zero-sum games; neural networks; online concurrent reinforcement learning algorithm

1. INTRODUCTION

The H_∞ control theory gives an answer to a major control problem, which is to conceive controllers not designed for a single plant under known inputs but for a class of plants under unknown inputs, which can be disturbances. The H_∞ optimal control framework was initiated by Zames [1], designed in frequency domain for sensitivity reduction and disturbance rejection. Many theoretical studies have been carried out on H_∞ control theory during few past decades [2–5]. In its equivalent time domain formulation, the H_∞ control problem is in fact a *minmax* optimization problem and, hence, a zero-sum differential game where the game is designed with the aim of obtaining a controller that minimizes a cost functional in the presence of worst-case disturbances [5, 6]. In order to solve zero-sum game problem arising in H_∞ optimal control, one needs to obtain the closed-loop solution of the Hamilton–Jacobi–Isaacs (HJI) equation, which is very difficult or even impossible to solve in case of nonlinear systems. This has inspired researchers to present some methods for

*Correspondence to: Sholeh Yasini, Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad 91775-1111, Iran.
†E-mail: sh.yasini@stumail.um.ac.ir