Neural Networks 105 (2018) 142–153

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Adaptive critic designs for optimal control of uncertain nonlinear systems with unmatched interconnections☆

Xiong Yang [a,b], Haibo He [b,*]

[a] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China
[b] Department of Electrical, Computer and Biomedical Engineering, University of Rhode Island, Kingston, RI 02881, USA

ARTICLE INFO

Article history:
Received 4 January 2018
Received in revised form 13 April 2018
Accepted 4 May 2018

Keywords:
Adaptive critic designs
Adaptive dynamic programming
Neural networks
Optimal control
Unmatched interconnection

ABSTRACT

In this paper, we develop a novel optimal control strategy for a class of uncertain nonlinear systems with unmatched interconnections. To begin with, we present a stabilizing feedback controller for the interconnected nonlinear systems by modifying an array of optimal control laws of auxiliary subsystems. We also prove that this feedback controller ensures a specified cost function to achieve optimality. Then, under the framework of adaptive critic designs, we use critic networks to solve the Hamilton–Jacobi–Bellman equations associated with auxiliary subsystem optimal control laws. The critic network weights are tuned through the gradient descent method combined with an additional stabilizing term. By using the newly established weight tuning rules, we no longer need the initial admissible control condition. In addition, we demonstrate that all signals in the closed-loop auxiliary subsystems are stable in the sense of uniform ultimate boundedness by using classic Lyapunov techniques. Finally, we provide an interconnected nonlinear plant to validate the present control scheme.

1. Introduction

The design of stabilizing controllers for interconnected nonlinear dynamical systems has drawn intensive attention over the past several decades. This is mainly because interconnected nonlinear dynamical systems have widely emerged in real world applications, such as electric power systems, computer networks, transportation systems, and aerospace systems. A core challenge of designing stabilizing controllers for such systems is that it often suffers from computational expensiveness when using a centralized control scheme (Jain & Khorrami, 1997). To overcome the difficulty, the decentralized control method was introduced. The decentralized control approach aims at partitioning the control problem of the overall plant into several manageable subproblems (Bakule, 2008). Then, the overall system can be controlled by a set of independent controllers. In other words, the decentralized controller is comprised of the independent controllers. A significant characteristic of the decentralized control method is that it only uses locally available subsystem states rather than the whole system states. Due to this property, plenty of studies on decentralized control has been reported in the literature (Hou, Cheng, & Tan, 2009; Li & Tong, 2017; Liu, Jiang, & Hill, 2012).

The optimal control methodology introduced to solve decentralized control problems can be dated back to at least 1980s. In Saberi (1988), a decentralized control of interconnected nonlinear systems was derived from an optimal control point of view. As pointed out by Saberi (1988), the decentralized controller for the overall system was able to be obtained by solving the optimal control problems of isolated subsystems. It is well-known that solving nonlinear optimal control problems often boils down to solving a class of nonlinear partial differential equations, namely, the Hamilton–Jacobi–Bellman equations (HJBEs). However, the HJBEs generally cannot be solved analytically (Aliyu, 2018). Thus, many researchers tend to find the approximate solutions of nonlinear optimal control problems. In 1970s, adaptive critic designs (ACDs) were first introduced as effective tools to approximately solve the optimal control problems (Werbos, 1974; Widrow, Gupta, & Maitra, 1973). The typical structure used in ACDs is the actor-critic architecture which consists of two networks: The actor network performs an action to the controlled system, and the critic network evaluates the value of that action and provides feedback information to the actor network. Because adaptive dynamic programming (ADP) (Liu, Wei, Wang, Yang, & Li, 2017) and reinforcement learning (RL) (Vrabie, Vamvoudakis, & Lewis, 2013) are almost in the same spirits as ACDs, they are usually viewed as synonyms for ACDs. In this paper, we take ADP and RL as a kind of ACDs.

☆ This work was supported in part by the National Natural Science Foundation of China under Grant 61503379, and in part by China Scholarship Council under the State Scholarship Fund, and in part by the National Science Foundation under Grant CMMI 1526835.

* Corresponding author.
E-mail addresses: xiong.yang@tju.edu.cn (X. Yang), he@ele.uri.edu (H. He).

https://doi.org/10.1016/j.neunet.2018.05.005
0893-6080/

Over the past three decades, many kinds of ADP and RL have been developed, such as goal representation ADP (Zhong, Ni, & He, 2017), robust ADP (Gao, Jiang, Jiang, & Chai, 2016; Jiang & Jiang, 2017), policy iteration ADP (Luo, Liu, Wu, Wang, & Lewis, 2017; Zhang, Jiang, Luo, & Xiao, 2017), value iteration ADP (Bertsekas, 2017; Wei, Liu, & Lin, 2016), data-based ACD (Wang, Liu, Zhang, & Zhao, 2016), integral RL (Lee, Park, & Choi, 2015; Yang, Liu, Luo, & Li, 2016), online RL (Kiumarsi, Vamvoudakis, Modares, & Lewis, 2018; Zhao & Zhu, 2015), and off-policy RL (Modares, Lewis, & Jiang, 2016; Zuo, Song, Lewis, & Davoudi, 2017).

In recent years, applications of ACDs to the design of optimal controller for interconnected systems have been extensively studied. In Mehraeen and Jagannathan (2011), an online decentralized optimal control scheme was developed for partially unknown discrete-time nonlinear systems via direct neural dynamic programming (NDP) (note: according to Liu et al. (2017), NDP is a synonym for ADP). After that, in Bian, Jiang, and Jiang (2015), a robust ADP was introduced to obtain the decentralized adaptive optimal control of continuous-time large-scale interconnected systems with completely unknown dynamics. Both Bian et al. (2015) and Mehraeen and Jagannathan (2011) employed the actor-critic architecture, where the actor aimed at approximating the optimal control and the critic tended to evaluate the cost of the overall system. Later, in Mu, Sun, Wang, Song, and Qian (2018), an ADP-based decentralized optimal control strategy was presented for continuous-time nonlinear systems with matched interconnections. Unlike Bian et al. (2015) and Mehraeen and Jagannathan (2011), Mu et al. (2018) only used critic networks to derive the decentralized optimal controller. In this sense, it had a simpler architecture. It is worth emphasizing here that all the above mentioned decentralized control approaches require the initial admissible control while implementing them. However, as stated in Liu, Yang, Wang, and Wei (2015), the initial admissible control is nothing else but the sub-optimal control, which is generally hard to obtain. Recently, Tong, Sun, and Sui (2018) developed a fuzzy adaptive decentralized optimal control scheme for continuous-time strict-feedback interconnected nonlinear systems via ADP. To implement this proposed ADP, there was no need to provide an initial admissible control, which was an advantage. Nevertheless, in comparison with Bian et al. (2015) and Mu et al. (2018), the performance for the overall system was not taken into account.

Inspired by aforementioned works, in this paper, a new optimal control strategy is developed for a class of uncertain nonlinear systems with unmatched interconnections. Owing to the present scheme having much in common with decentralized control approaches, this optimal control strategy can be regarded as a kind of decentralized control methods. In the beginning, a stabilizing feedback controller for the interconnected nonlinear systems is designed through modifying an array of optimal control laws of auxiliary subsystems. Meanwhile, this feedback controller is demonstrated to be able to optimize a prescribed cost function. Then, under the framework of ACDs, the critic networks are utilized to solve the HJBEs associated with auxiliary subsystem optimal control laws. The critic network weights are updated via the gradient descent method combined with an additional stabilizing term. Based on the newly established weight update rules, the initial admissible control is no longer indispensable. In addition, all signals in the closed-loop auxiliary subsystems are proved to be uniformly ultimately bounded (UUB) by using Lyapunov method.

The remainder of the paper is arranged as follows. Section 2 presents preliminaries and problem formulations Section 3 describes the optimal control strategy for interconnected nonlinear systems. Section 4 illustrates that the approximation solutions of HJBEs can be obtained via ACDs. Meanwhile, the stability of closed-loop auxiliary subsystems is discussed. Section 5 provides an example to validate the developed theoretical results. Finally, Section 6 presents the discussion and concluding remarks.

Notation: $\mathbb{R}$ and $\mathbb{Z}^+$ represent the set of all real numbers and the set of all positive integers, respectively. $\mathbb{R}^{m_i}$ and $\mathbb{R}^{n_i \times m_i}$ represent the spaces of all real m_i-vectors and all $n_i \times m_i$ real matrices, respectively. I_{n_i} is the identity matrix of dimension $n_i \times n_i$. $\triangleq$ means 'equal by definition'. T denotes the transpose. C^1 denotes the class of functions with continuous derivative. When $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^{\mathsf{T}} \in \mathbb{R}^{n_i}$, $\|x_i\| = \sqrt{\sum_{s=1}^{n_i} |x_{is}|^2}$ represents the Euclidean norm of x_i. When $A_i \in \mathbb{R}^{n_i \times m_i}$, $\|A_i\|$ denotes the Frobenius-norm of A_i. $\nabla V_i(x_i)$ denotes the partial derivative of $V_i(x_i)$ with respect to x_i, i.e., $\nabla V_i(x_i) = \partial V_i(x_i)/\partial x_i$.

2. Preliminaries and problem formulations

Consider the continuous-time interconnected nonlinear systems formulated as

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(t) + \Delta f_i(x(t)) \quad i = 1, 2, \ldots, N \tag{1}$$

where $x_i \in \mathbb{R}^{n_i}$ and $u_i \in \mathbb{R}^{m_i}$ are the measurable state and control input of the ith subsystem, respectively, $x = [x_1^{\mathsf{T}}, x_2^{\mathsf{T}}, \ldots, x_N^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{R}^n$ ($n = \sum_{i=1}^N n_i$) is the overall state, $f_i(x_i) \in \mathbb{R}^{n_i}$, $g_i(x_i) \in \mathbb{R}^{n_i \times m_i}$, and $\Delta f_i(x) \in \mathbb{R}^{n_i}$ are the known internal dynamics, the known input matrix, and the *uncertain* interconnection of the ith subsystem, respectively.

The following assumptions are provided to facilitate our later discussion. These assumptions were employed in Tripathy, Kar, and Paul (2018), Wang, Liu, Mu, and Ma (2016) and Yang and He (2018).

Assumption 1. $f_i(x_i)$ and $g_i(x_i)$, $i = 1, 2, \ldots, N$, are continuously differentiable in their arguments. $x_i = 0$ is the equilibrium point of the ith subsystem when $u_i(t) = 0$ and $\Delta f_i(x(t)) = 0$ for all $t \geq 0$. Furthermore, $x_{i0} = x_i(0)$ is the initial state of the ith subsystem, where $i = 1, 2, \ldots, N$.

Assumption 2. The interconnected term $\Delta f_i(x)$ satisfies the unmatched condition, i.e.,

$$\Delta f_i(x) = k_i(x_i)\omega_i(x)\,(k_i(x_i) \neq g_i(x_i)), \; i = 1, 2, \ldots, N$$

where $k_i(x_i) \in \mathbb{R}^{n_i \times p_i}$ is a known smooth function, and $\omega_i(x) \in \mathbb{R}^{p_i}$ is an *uncertain* function bounded as

$$\|\omega_i(x)\| \leq \sum_{j=1}^N a_{ij} P_{ij}(x_j), \quad i = 1, 2, \ldots, N \tag{2}$$

where $P_{ij}(x_j)$, $j = 1, 2, \ldots, N$, are positive-definite functions (Khalil, 2002) and a_{ij}, $j = 1, 2, \ldots, N$, are nonnegative constants. Meanwhile, $\omega_i(0) = 0$ and $P_{ij}(0) = 0, j = 1, 2, \ldots, N$.

Let

$$P_i(x_i) = \max\{P_{1i}(x_i), P_{2i}(x_i), \ldots, P_{Ni}(x_i)\}. \tag{3}$$

Then, we can further develop (2) as

$$\|\omega_i(x)\| \leq \sum_{j=1}^N b_{ij} P_j(x_j), \quad i = 1, 2, \ldots, N \tag{4}$$

where $b_{ij} \geq a_{ij} P_{ij}(x_j)/P_j(x_j)$, $j = 1, 2, \ldots, N$, are nonnegative constants.

Assumption 3. For the ith subsystem, the control matrix $g_i(x_i)$ has the full column rank, and $g_i^{\mathsf{T}}(x_i)k_i(x_i) = 0$.

The cost function for interconnected system (1) is given by

$$J(x(t), u(t)) = \int_t^\infty r(x(s), u(s))\mathrm{d}s \tag{5}$$