Neurocomputing 73 (2010) 3020–3027

Contents lists available at ScienceDirect

Neurocomputing

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

# Optimal control laws for time-delay systems with saturating actuators based on heuristic dynamic programming ☆

Ruizhuo Song [a], Huaguang Zhang [a,*], Yanhong Luo [a], Qinglai Wei [b]

[a] School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110004, China
[b] The Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

*Article history:*
Received 19 November 2009
Received in revised form
24 March 2010
Accepted 12 July 2010
Communicated by W. Yu
Available online 27 August 2010

*Keywords:*
Adaptive dynamic programming
Approximate dynamic programming
Nonlinear systems
Time delays
Optimal control
Saturating actuators

ABSTRACT

In this paper a new iterative heuristic dynamic programming (HDP) algorithm is proposed to solve the optimal control problem for a class of nonlinear discrete time-delay systems with saturating actuators. Considering the saturation nonlinearity in the actuators, a nonquadratic performance index function is introduced. In the meantime, a state modification is used to deal with the obstacle induced by time delays. In the new iterative HDP algorithm the local and global optimization searching processes are developed to solve the optimal feedback control problem with convergence analysis. In the presented iterative HDP algorithm, two neural networks are used to facilitate the implementation of the iterative algorithm. Finally, two simulation examples are given to demonstrate the convergence and feasibility of the proposed optimal control scheme.



## 1. Introduction

Saturation nonlinearity is unavoidable in most of the actuators. Due to the nonanalytic nature of the actuator nonlinear dynamics and the fact that the exact actuator nonlinear functions are unknown, such systems present a challenge to control engineers. So the control of systems with saturating actuators has been the focus of many researchers for many years. Several methods for deriving control laws considering the saturation phenomena can be found in [4,13]. On the other hand, time delays often occur in the transmission or material between different parts of systems [16]. So, the research on the optimal control for nonlinear time-delay systems with saturating actuators is beneficial to practical application.

It is well known that there are many methods for designing stable control for nonlinear systems [10,21,25]. However, stability is only a bare minimum requirement in a system design. Ensuring optimality guarantees the stability of the nonlinear systems. Dynamic programming is a very useful tool in solving optimization and optimal control problems by employing the principle of optimality. However, solving the associated Hamilton–Jacobi–Bellman (HJB) equation demands a large (rather infeasible) number of computations and storage space dedicated to this purpose [22]. An innovative idea was proposed in [18] to get around this numerical complexity by using an adaptive/approximate dynamic programming (ADP) formulation. And heuristic dynamic programming (HDP) is one of the main schemes of ADP. In recent years, ADP algorithms were further developed by Lewis [1,7,8], Powell [24], Jagannathan [6,17], Murray [14], and Si et al. [5]. Moreover, in [3], a greedy iterative HDP scheme with convergence proof was proposed for solving the optimal control problem for nonlinear discrete-time systems. In [20], a forward-in-time optimal control method for a class of discrete-time nonlinear systems with general multi-objective performance indices was proposed with unknown system dynamics. In [11], the near-optimal control problem for a class of nonlinear discrete-time systems with control constraints was solved by iterative adaptive dynamic programming algorithm. In [19], an optimal control scheme for a class of nonlinear systems with time delays in state and control variables with respect to a quadratic performance index function was proposed using a new iterative ADP algorithm.

Though ADP algorithms have made great progress in the optimal control field [9,23,26], to the best of our knowledge, it is still an open problem how to solve the optimal control problem for time-delay systems with saturating actuators based on ADP algorithms. This motivates our research. First, the HJB equation for time-delay systems with saturating actuators is derived using nonquadratic function.

☆ This work was supported by the National Natural Science Foundation of China (50977008, 60774048, 60821063, 61034005), the Program for Cheung Kong Scholars, National Basic Research Program of China (2009CB320601).

* Corresponding author. Tel.: +86 24 83687762; fax: +86 24 83671498.
*E-mail addresses:* ruizhuosong@163.com (R. Song), zhanghuaguang@ise.neu.edu.cn, zhg516516@gmail.com, hgzhang@ieee.org (H. Zhang).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.07.005

In order to solve this HJB equation, two optimization searching processes are proposed. One is local optimization searching process which is used to modify the states for every iteration, and the other is global optimization searching process which is used to get the optimal control law. We also give the convergence proof for the new iterative HDP algorithm. At last, two networks are used to implement the iterative HDP algorithm. They are critic neural network and action neural network which approximate the performance index function and compute the corresponding control law, respectively.

This paper is organized as follows. In Section 2, we present the problem formulation. In Section 3, the optimal control scheme is developed based on iterative HDP algorithm and the convergence proof is given. In Section 4, the neural network implementation for the optimal control scheme is discussed. In Section 5, two examples are given to demonstrate the effectiveness of the proposed control scheme. In Section 6, the conclusion is given.

## 2. Problem formulation

Consider a class of affine nonlinear discrete time-delay systems with saturating actuators as follows:

$$\begin{cases} x_{k+1}=f(x_{k-\sigma_0},\ldots,x_{k-\sigma_m})+g(x_{k-\sigma_0},\ldots,x_{k-\sigma_m})u_k, & k\geq 0,\\ x_k=\lambda_k, & k=-\sigma_0,-\sigma_1,\ldots,-\sigma_m, \end{cases} \tag{1}$$

where $x_{k-\sigma_0},x_{k-\sigma_1},\ldots,x_{k-\sigma_m}\in\Re^n$, $u_k\in\Re^m$, $f(x_{k-\sigma_0},x_{k-\sigma_1},\ldots,x_{k-\sigma_m})\in\Re^n$, $g(x_{k-\sigma_0},x_{k-\sigma_1},\ldots,x_{k-\sigma_m})\in\Re^{n\times m}$, $\lambda_k$ describes the initial condition. Set $0=\sigma_0<\sigma_1<\cdots<\sigma_m,\sigma_i$, $i=1,\ldots,m$ is positive integer number. Here assume that the system is controllable on $\Omega\subset\Re^n$. Assume that $f,g$ are all Lipschitz continuous functions. We denote $\Omega_u=\{u_k|u_k=[u_k(1)\quad u_k(2)\ldots u_k(m)]^{\mathrm{T}}\in\Re^m,|u_k(i)|\leq\overline{u}(i),i=1,\ldots,m\}$, where $\overline{u}(i)$ is the saturating bound for the $i$th actuator. Let $\overline{U}=diag[\overline{u}(1)\overline{u}(2)\ldots\overline{u}(m)]$.

This paper is desired to find an optimal control law for the system (1), which minimizes a generalized nonquadratic performance index function

$$J(x_k,u_k)=\sum_{i=k}^{\infty}\{Q(x_i)+W(u_i)\}, \tag{2}$$

where $Q(x_i)$, $W(u_i)$ are positive definite, and $U(x_i,u_i)=Q(x_i)+W(u_i)$ is the utility function denoting the cost incurred in going from $k$ to $k+1$ using control $u_k$.

Note that, for optimal control problems, the state feedback control laws $u_k$ not only stabilize the system (1) on $\Omega$, but also guarantee (2) is finite. Such controls are defined to be admissible [11].

**Definition 1** (*Admissible control*). A control law $u(x_k)$ is defined to be admissible with respect to (2) on $\Re^n$ if $u(x_k)$ is continuous on $\Re^n$, $u(0)=0$, and for $\forall x_k\in\Re^n$, $u(x_k)$ stabilizes (1) on $\Re^n$, and $\forall x_0\in\Re^n\, J(x_0,u(\cdot))$ is finite, where $u(\cdot)=(u_0,u_1,\ldots)$ and $u_k=u(x_k)$, $k=0,1,\ldots$ .

In this paper, we mainly discuss the optimal control for discrete time-delay system with saturating actuators. Inspired by [11,15], we let

$$Q(x_i)=X_{i-\sigma}^{\mathrm{T}}QX_{i-\sigma} \tag{3}$$

and

$$W(u_i)=2\int_0^{u_i}\phi^{-\mathrm{T}}(\overline{U}^{-1}s)\overline{U}R\,ds, \tag{4}$$

where $X_{i-\sigma}=[x_{i-\sigma_0};x_{i-\sigma_1};\ldots;x_{i-\sigma_m}]$, $Q,R$ are positive define, and we assume that $R$ is diagonal for simplicity of analysis, $s\in\Re^m$, $\phi\in\Re^m$, $\phi^{-1}(u_i)=[\psi^{-1}(u_i(1))\psi^{-1}(u_i(2))\cdots\psi^{-1}(u_i(m))]^{\mathrm{T}}$, $\psi(\cdot)$ is a bounded single mapping function satisfying $|\psi(\cdot)|\leq 1$ and belonging to $C^p$ $(p\geq 1)$ and $L_2$. Moreover, it is a monotonic increasing odd function with its first derivative bounded by a constant $M$. We know that the hyperbolic tangent function $\psi(\cdot)=\tanh(\cdot)$ is one example of such function. Noting that $W(u_i)$ is assured to be positive define by the definition above, because $\psi^{-1}(\cdot)$ is a monotonic odd function and $R$ is positive definite.

Let $J^*(x_k)=\min_{u_k}J(x_k,u_k)$ denote the optimal performance index function, and let $u^*{}_k$ denote the corresponding optimal control law. According to Bellman's principle of optimality, the optimal performance index function $J^*(x_k)$ should satisfy the following HJB equation:

$$J^*(x_k)=\min_{u_k}\sum_{i=k}^{\infty}\{Q(x_i)+W(u_i)\}=\min_{u_k}\{Q(x_k)+W(u_k)+J^*(x_{k+1})\}, \tag{5}$$

and the optimal controller $u_k^*$ should satisfy

$$u_k^*=\arg\min_{u_k}\sum_{i=k}^{\infty}\{Q(x_i)+W(u_i)\}=\arg\min_{u_k}\{Q(x_k)+W(u_k)+J^*(x_{k+1})\}. \tag{6}$$

The optimal control problem for the nonlinear discrete time-delay system with saturating actuators can be solved if the optimal performance index function $J^*(x_k)$ can be obtained from (5). However, there is currently no quite effective method for solving this performance index function for the nonlinear discrete time-delay system with saturating actuators. Therefore, in the following part we will discuss how to utilize the iterative HDP algorithm to seek the approximate optimal control solution.

## 3. The optimal control based on iterative HDP algorithm

### 3.1. Derivation of the iterative HDP algorithm

First, for any given initial state $\lambda_k$ and initial control policy $u_k$ we start with initial iteration performance index function $J^0(\cdot)=0$. Then we find the control vector $u_k^0$ as follows:

$$u_k^0=\arg\min_{u_k}\left\{(X_{k-\sigma}^{\mathrm{T}})QX_{k-\sigma}+2\int_0^{u_k}\phi^{-\mathrm{T}}(\overline{U}^{-1}s)\overline{U}R\,ds+J^0(x_{k+1})\right\}, \tag{7}$$

and the performance index function is updated as

$$J^1(x_k)=\min_{u_k}\left\{(X_{k-\sigma}^{\mathrm{T}})QX_{k-\sigma}+2\int_0^{u_k}\phi^{-\mathrm{T}}(\overline{U}^{-1}s)\overline{U}R\,ds+J^0(x_{k+1})\right\}, \tag{8}$$

where $x_1,\ldots,x_k$ are obtained under the action $u_k$, and

$$x_{k+1}=f(x_{k-\sigma_0},\ldots,x_{k-\sigma_m})+g(x_{k-\sigma_0},\ldots,x_{k-\sigma_m})u_k^0. \tag{9}$$

Moreover, for $i=1,2,\ldots$ the iterative HDP algorithm iterates between

$$u_k^i=\arg\min_{u_k}\left\{(X_{k-\sigma})^{\mathrm{T}}QX_{k-\sigma}+2\int_0^{u_k}\phi^{-\mathrm{T}}(\overline{U}^{-1}s)\overline{U}R\,ds+J^i(x_{k+1})\right\} \tag{10}$$

and

$$J^{i+1}(x_k)=\min_{u_k}\left\{(X_{k-\sigma})^{\mathrm{T}}QX_{k-\sigma}+2\int_0^{u_k}\phi^{-\mathrm{T}}(\overline{U}^{-1}s)\overline{U}R\,ds+J^i(x_{k+1})\right\}, \tag{11}$$

where $x_1,\ldots,x_k$ are obtained under the action $u_k^{i-1}$, and

$$x_{k+1}=f(x_{k-\sigma_0},\ldots,x_{k-\sigma_m})+g(x_{k-\sigma_0},\ldots,x_{k-\sigma_m})u_k^i. \tag{12}$$

We further compute the control law $u_k^i$ from Eq. (10):

$$u_k^i=\overline{U}\phi\left(-\frac{1}{2}(\overline{U}R)^{-1}g^{\mathrm{T}}(x_{k-\sigma_0},\ldots,x_{k-\sigma_m})\frac{\partial J^i(x_{k+1})}{\partial x_{k+1}}\right) \tag{13}$$