ACADEMIC
World Academic Union

ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 10, No. 2, 2015, *pp*. 106-113

Robust Reinforcement Learning Decoupling Control Based on Integral Quadratic Constraints

Wang Teng *
School of Computer, South China Normal University, Guangzhou 510631, China
(*Received August 07, 2013, accepted January 06, 2015)*

Abstract. In order to keep stable in reinforcement learning process, a novel robust reinforcement learning decoupling control (RRLDC) based on integral quadratic constraints（IQC）is presented in this paper. It composes of a linear model to approximate the nonlinear plant, a state feedback K controller to generate the basic control law, and an adaptive critic unit to evaluate decoupling performance, which tunes an actor unit to compensate decoupling action and model uncertainty as well as system nonlinearity. By replacing nonlinear and time-varying aspects of a neural network and model uncertainty with IQC, the stability of the control loop is analyzed. As a result, the range of the adjusted parameters is found within which the stability is guaranteed, the control system performance is improved through learning and the algorithm convergence speed is accelerated. The proposed RRLDC is applied to gas collector pressure control of coke ovens. The simulation results show the proposed control strategy can not only obtain the good performance, but also avoid unstable behavior in learning process. It is an effective multivariable decoupling control method for a class of strong coupling systems such as the gas collector pressure control of coke ovens.the effectiveness of proposed control strategy for the collector gas pressure of coke ovens

Keywords: reinforcement learning, decoupling control, integral

1. Introduction

Reinforcement learning (RL) is a kind of machine learning method which can be applied to solve optimal control problems[1].

Although classical reinforcement learning algorithm does not need environment models such as Sutten's TD (temporal difference) algorithm, Watkin's Q-learning algorithm and AHC (adaptive heuristic critic), there exist problems of slow convergence speed and low convergence precision. Environment models added can increase convergence speed. The nature of reinforcement learning is exploration and exploitation. Successful applications often allow systems to accumulate experience, learn from failures, and eventually succeed. However, it is difficult to ensure system stability and tracking performance in reinforcement learning process, which may greatly limit its applications in complex industrial process control.

Robust control theory is applied to uncertain systems to analyze their stability and design a controller with certain performance[2]. It has a unified framework including gain margin, phase margin, tracking, noise disturbance poise and calm concept. Most control system design methods based on robust control theory provide strong stability guarantees and certain performances. Hence, Reinforcement learning combining with robust control theory has become a hotspot in the past decades. Kretchmar et al combined robust control theory with reinforcement learning[3]. Morimoto and Doya[4] applied robust control theory to improve the performance of a reinforcement learning system under certain disturbances. Perkins and Barto[5] used Lyapunov principles to design stable controllers and achieved a level of desired performance. Jose[6] presented a series of reinforcement learning algorithms which could learn quickly, generalize properly over continuous state spaces and be robust to a high degree of environmental noise. Dong[7] proposed a novel quantum-inspired reinforcement learning algorithm for navigation control of autonomous mobile robots. This approach was then applied to navigation control of a real mobile robot. The simulation and experimental results show its effectiveness.

In the coking process, the stability of gas collector pressure has influence on coke ovens' quality, life-time and their production environment. Also, its control has a direct impact on the operating conditions of the whole coke oven system. However, the control system of gas collector pressure is a complicated multivariable, nonlinear, time-varying and big time-delay control object, which has many disturbance factors and the

* Corresponding author. Tel.: +86-15360060605.
E-mail address of corresponding author Wang Teng: towangteng@263.net

coupling phenomenon is serious. Therefore, it is difficult to establish a mathematical model to reflect collector gas pressure accurately and obtain desired control effect using traditional control methods[8].

In this paper, a novel intelligent decoupling control method based on IQC reinforcement learning is proposed and an actor-critic architecture is adopted in the RRLC, which uses a neural network to generate the decoupling control compensation and a reinforcement learning automata (RLA) to learn the neural network parameters. Meanwhile, reinforcement learning to find the optimal solution is integrated with domain knowledge available to analyze the system stability. The strategy is used for decoupling control of coke ovens plant. The simulation results show its effectiveness for gas pressure control of coke ovens.

2. The methods

A reinforcement learning algorithm

The essence of reinforcement learning is how an agent takes actions in environment so as to maximize long-term reward. The Agent consists in critic unit and actor unit. In the process of information interaction, the Agent receives environment state s and reward signal r. At the same time, the critic unit evaluates the control performance and guides the actor unit to find a policy that maps state s to action a. This action output will lead to environment change to create a new state. Every time, the choice behavior principle for the Agent is to maximize long-term reward, determine whether a new state meets the learning goals and generate a corresponding TD error signal to adjust actor unit strategies. After repeated trial and error, the Agent can finally obtain optimal behavioral strategies and complete the learning task.

B RL-based Decoupling control system architecture

RL-based decoupling control system is an actor-critic architecture, which has five components: adaptive critic evaluation element (ACE), associative search element (ASE), system model (SM), state feedback controller K, IQC stability analysis and optimal element. The system model approximates the linear part of plant. The state feedback K controller is used in the control loop giving the system a guaranteed initial performance. The ACE learns to evaluate the system performance and tune the ASE to learn nonlinear decoupling compensation mapping. The input of actual control is generated by summing the output of controller K and ASE.

Through replacing nonlinear and time-varying aspects of a neural network and model uncertainty with IQC, the stability is guaranteed in robust reinforcement learning process even as the neural network is being trained.

C. STATE FEEDBACK K

In fact, the coke oven system is a nonlinear and time-varying system. It is sure that there exist error and residuals in statistics, if a linear system, as shown in Eq(1), is used to model it.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{1}$$

Where *y* is controlled system output, *X* is state of system, A, B, C, D is matrix of system.

The quadratic constraint *J*, as performance index, is introduced to evaluate the LTI system.

$$J = \int_0^\infty [x^T(t)Q(t)x(t) + U^T(t)R(t)U(t)]dt \tag{2}$$

H function is got when minimize *J* [2]:

$$\frac{\partial H}{\partial u} = 0 \Rightarrow u^*(t) = -R^{-1}B^T Px(t) \tag{3}$$

Where *P* is got from Riccati matrix equations. Suppose $K = R^{-1}B^T P$, then the basic control law $-Kx(t)$ is obtained.The overall output, as mentioned above, is given by

$$u(t) = -Kx(t) + \tau(t) \tag{4}$$

Where decoupling compensation τ is the ASE output adding random exploration.

D. ACE AND ASE