J Control Theory Appl 2011 9 (3) 309
DOI 10.1007/s11768-011-1104-1

Editorial

Special issue on approximate dynamic programming and reinforcement learning

We are extremely pleased to present this special issue of the Journal of Control Theory and Applications. Approximate dynamic programming (ADP) is a general and effective approach for solving optimal control and estimation problems by adapting to uncertain environments over time. ADP optimizes the sensing objectives accrued over a future time interval with respect to an adaptive control law, conditioned on prior knowledge of the system, its state, and uncertainties. A numerical search over the present value of the control minimizes a Hamilton-Jacobi-Bellman (HJB) equation providing a basis for real-time, approximate optimal control.

In the past few decades, various types of adaptive controllers provide implementation strategies that employ online observations of system performance either to drive the system to an equilibrium state, or track prescribed trajectories. Many techniques have been developed for online controller tuning without requiring the system dynamics. However, the control engineer is often constrained in the choice of performance measure or cost employed for the optimization.

A key feature of ADP is that it can approximate the optimal control law (or policy) with only either partial or minimal knowledge of the value function or of the system dynamics, in an online fashion. In an effort to reduce computational complexity, the control law and value function have been represented by nonlinear function approximators, such as adaptive neural networks, fuzzy logic and other online approximators, giving rise to Neurodynamic Programming (NDP). Much work is being done in this area toward the development of adaptive control systems that are based on value or policy iteration, or both (i.e., adaptive critics).

It is becoming clear that ADP techniques, on the other hand, do allow the design of optimal controllers online in real time in terms of a (freely) prescribed performance measure. The key lies effectively in solving Bellman's optimality condition forward in time through repeated iterations that involve: 1) computing the cost or value of using a current control, then 2) based on that value performing a control policy update, or control improvement. However, in the most recent literature, it was shown that iterating the value and control policies within a sampling interval may not be feasible in many control implementations and merely updating the value and policy at each sampling interval is sufficient for guaranteed convergence and performance.

The ADP methodology can be viewed as a type of reinforcement learning comprising two components, a critic and an action network. The former evaluates currently instantiated control via policy or value iteration, and the latter improves controller design based on the latest evaluation. Usually, online approximators are utilized for the critic and action networks.

It is important to note that ADP extends Riccati equation-like linear optimal controller design methods to unknown nonlinear systems by using universal function approximators. In addition, in the recent works, ADP controllers have been derived without needing the measured system states via software observers and by transforming the system into input-output form usually referred to as data-based control. A number of ADP applications to formation control, engine control, smart grid, stock trend prediction have appeared in the literature.

This special issue aims to showcase recent findings by the automatic control and computational intelligence communities in order to highlight emerging trends and frontiers in ADP methods and applications. Many of the ADP practitioners are being represented in this special issue.

Silvia Ferrari
Laboratory for Intelligent Systems and Control (LISC)
Department of Mechanical Engineering & Materials Science
Box 90300, Duke University, Durham, NC 27708-0005
E-mail: sferrari@duke.edu
Tel: +1 919-660-5484; Cell: +1 919-597-9586; Fax: +1 919-660-8963

Jagannathan Sarangapani
Department of Electrical & Computer Engineering
University of Missouri-Rolla, MO, 65401, U.S.A.
E-mail: sarangap@mst.edu

Frank L. Lewis
Automation and Robotics Research Institute
The University of Texas at Arlington
7300 Jack Newell Blvd. S, Ft. Worth, Texas, 76118-7115, U.S.A.
E-mail: lewis@uta.edu

Control Theory and Technology

Publishing model: **Hybrid**

Volume 9, Issue 3

August 2011

14 articles in this issue