
Electrical & Computer Engineering
S E M I N A R
Louisiana State University

**Reinforcement Learning Feedback Control
Using Reduced Output Measurements**

F. L. Lewis

University of Texas at Arlington Riverbend

Abstract—In this talk we present new results in H-infinity control using reinforcement learning (RL) techniques, which use observed system responses to update the control policy in real-time in an optimal fashion. We show how to implement RL feedback controllers for continuous-time systems using newly developed techniques. We present a new method for RL control that requires only output measurements. Traditional RL methods require full state variable feedback.

Optimal Control design techniques have provided very effective feedback controllers for modern systems in aerospace, vehicle systems, industrial process control, robotics, mobile robots, wireless sensor networks, and elsewhere. Optimal control design is fundamentally a backwards-in-time procedure based on dynamic programming, specifically on Bellmans Optimality Principle. This means that most existing optimal control design methods must be carried out off-line. Moreover, the full system dynamical description must generally be known to compute optimal controllers using well-known techniques, such as Riccati equation design.

In this talk we show how to implement optimal controllers on-line forward in time for systems whose dynamical description is not known or is partially known. A family of on-line Optimal Adaptive Controllers is provided, whereby adaptive learning techniques are used to learn the optimal control strategy in real time using system measured data along the system trajectories. In the linear time-invariant case, this amounts to solving the Riccati equation on-line in real-time without knowing the system plant matrix.

These Optimal Adaptive Controllers are based on Approximate Dynamic Programming (ADP) and Q learning. Reinforcement Learning is a method for on-line learning of control policies based on stimuli from the environment in response to current control policies. Such methods were used by I.M. Pavlov for learning in canines. Particularly interesting are the actor-critic structures, including those based on policy iteration those based on value iteration. A special case of value iteration is the ADP structures. Q-learning is a method of actor critic reinforcement learning that does not require any knowledge of the system

dynamics, yet finds optimal control policies on-line in real time. ADP and Q learning have been well developed by the Computational Intelligence Community, primarily for Markov Decision processes, and have not been fully explored for feedback control purposes within the Control systems Community

When: Friday, 1 October 2010, 13:30 - 14:30

Where: 117 Electrical Engineering Building

Info: <http://www.ece.lsu.edu/seminar>