



The Center for Control, Dynamical Systems, and Computation
University of California at Santa Barbara
Spring 2009 Seminar Series
Presents

Q-learning and Pontryagin's Minimum Principle

Sean Meyn

University of Illinois at Urbana-Champaign

Friday, May 15, 2009 3:00 - 4:00pm WEBB 1100

Abstract: Q-learning is a technique used to compute an optimal policy for a controlled Markov chain based on observations of the system controlled using a non-optimal policy. It has proven to be effective for models with finite state and action space. In this talk we demonstrate how the construction of the algorithm is identical to concepts more classical nonlinear control theory - in particular, Jacobson & Mayne's differential dynamic programming introduced in the 1960's. I will show how Q-learning can be extended to deterministic and Markovian systems in continuous time, with general state and action space. The main ideas are summarized as follows. (i) Watkin's "Q-function" is an extension of the Hamiltonian that appears in the Minimum Principle. Based on this observation we obtain extensions of Watkin's algorithm to approximate the Hamiltonian within a prescribed finite-dimensional function class. (ii) A transformation of the optimality equations is performed based on the adjoint of a resolvent operator. This is used to construct a consistent algorithm based on stochastic approximation that requires only causal filtering of the time-series data. (iii) Examples are presented to illustrate the application of these techniques, including application to distributed control of multi-agent systems. Reference: P. Mehta and S. Meyn. Q-learning and Pontryagin's Minimum Principle. Submitted to the 48th IEEE Conference on Decision and Control, December 16-18 2009.

About the Speaker: Sean P. Meyn received the B.A. degree in Mathematics Summa Cum Laude from UCLA in 1982, and the PhD degree in Electrical Engineering from McGill University in 1987 (with Prof. P. Caines, McGill University). After a two year postdoctoral fellowship at the Australian National University in Canberra, Dr. Meyn and his family moved to the Midwest. He is now a Professor in the Department of Electrical and Computer Engineering, and a Research Professor in the Coordinated Science Laboratory at the University of Illinois. He is also an IEEE fellow. He is coauthor with Richard Tweedie of the monograph Markov Chains and Stochastic Stability, Springer-Verlag, London, 1993, and received jointly with Tweedie the 1994 ORSA/TIMS Best Publication In Applied Probability Award. The 2009 edition is published in the Cambridge Mathematical Library. His new book, Control Techniques for Complex Networks is published by Cambridge University Press. He has held visiting positions at universities all over the world, including the Indian Institute of Science, Bangalore during 1997-1998 where he was a Fulbright Research Scholar. During his latest sabbatical during the 2006-2007 academic year he was a visiting professor at MIT and United Technologies Research Center (UTRC). His research interests include stochastic processes, optimization, complex networks, and information theory. Current funding is provided by NSF, Motorola, DARPA, and UTRC.