MARKOV GAME BASED HYBRID CONTROLLERS FOR NONLINEAR SYSTEMS

by

Rajneesh Sharma

Electrical Engineering Department

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the

Indian Institute of Technology, Delhi

February 2007

Certificate

This is to certify that the thesis entitled *"Markov Game based Hybrid Controllers for Nonlinear Systems"* being submitted by **Rajneesh Sharma** for the award of the degree of **Doctor of Philosophy** to the **Indian Institute of Technology, Delhi** is a record of the bonafide research work he has carried out under my supervision. The results contained in this thesis have not been submitted to any other University or Institute for the award of a degree or a diploma.

MGopal

(Prof. M. Gopal)
Supervisor

Department of Electrical Engineering

Indian Institute of Technology, Delhi

New Delhi – 110 016

Acknowledgements

My deepest gratitude goes to my supervisor **Prof. M. Gopal**. My demands on his valuable time, over the past six years for research directions, manuscript writing, and discussions have been quite phenomenal. Sir provided me with 'right from scratch' support in research problem formulation, and in tackling problems that came on, both motivational and logistic wise. He is the one who provided breakthroughs whenever I found myself in real soup, and at times handled my stubbornness to pursue wrong way. He taught me the nitty-gritty of research and encouraged (at times pressurized) me to perform beyond my own capabilities. Sir bailed me out from many a difficult situations through his enormous experience and knowledge. I can only summarize his contribution by saying that he will always be a true role-model for me in pursuing my academic goals.

I would also like to thank my Student Research Committee (SRC) members, especially, Prof. Ranjan K. Mallik, and Prof. I. N. Kar. Both of them provided critical comments, suggested references, were willing to discuss anything, and patiently endured my SRC presentations on research progress.

I owe my thanks to Dr. Ranjeet Singh, Director, NSIT, for sponsoring my Ph.D. program at IIT-Delhi, and Prof. A. P. Mittal, HOD, ICE division, NSIT, for providing a congenial environment at my workplace.

I have no words to say thanks to my spouse, **Pratibha**, who stood by me during difficult times. She encouraged me to go on and believe in myself. My little daughter, **Titiksha**, made me happy through her angelic smile and small enquiries. Both of them have been patient enough to allow me concentrate on my work.

There is no way in which I can express my gratitude to my parents for they taught me everything that today I stand for. I would also like to express my appreciation to my Mother-in-law for relieving me to concentrate on my work by taking care of my wife and daughter.

I would also like to thank the lab staff at IIT-Delhi, Mr. Jaipal Singh in particular, for the light moments he shared with me, and for all the logistic help he provided.

My thanks go to other research students, past and presents, who have contributed in one way or other in my efforts.

Finally, I pay my tributes to the **ALMIGHTY LORD SHIVA** for blessing me in all my endeavors.

Rajneesh S.

Rajneesh Sharma

ABSTRACT

This thesis proposes novel reinforcement learning (RL) based game-theoretic formulations for adaptive-optimal control of nonlinear systems, affected by external disturbances and/or modeling errors. In control literature, one of the most widely used adaptive control approach to deal with disturbances is the well-known H_∞ control methodology. However, conventional H_∞ based approaches depend heavily on the analytical solvability of a corresponding Hamilton-Jacobi-Isaac (HJI) equation, and assume full knowledge of the system dynamics. For majority of practical systems, these requirements are hard to satisfy. RL based approaches, on the other hand, frame controller optimization problem as finding optimal control policy, for the environment modeled as a Markov decision process (MDP). This assumption of modeling the environment as an MDP severely limits the scope of application of RL methods, as it places a strong constraint on the structure of environment. Further, MDP formalism is quite restrictive in the sense that it allows only a single agent operating in a stationary environment.

Work in this thesis is an attempt to overcome these limitations, by providing an alternative framework based on the theory of Markov games, for adaptive-optimal control of *unknown* nonlinear systems, in presence of external disturbances and uncertainties. In the proposed Markov game based formulation of controllers, controller-disturber tussle is viewed as a two player zero-sum Markov game, in which a 'disturbing' agent tries to make worst possible disturbance while a 'control' agent tries to make best control input. The problem is framed as finding 'minimax' solution of a value function. We propose an

online procedure for learning optimal value function, and calculating optimal control policy.

Comprehensive empirical evaluations of the proposed Markov game based control methodology on several standard tasks, such as control of two-link robot manipulators, inverted pendulum swing-up, and steering control of a mobile robot, illustrate effectives of the approach in handling external disturbances and parameter variations. Further, performance comparison of the proposed approach against other control approaches, e.g., H_∞ theory based robust game controller, RL based robust controller, and Q-learning based controller, indicates viability of Markov game based control. A formal convergence analysis shows that the approach converges to optimal function values. We extend the proposed approach to continuous state-action space control problems, by implementing a continuous-action variation of Markov games in conjunction with generic function approximators such as neural networks, and fuzzy inference systems. Fuzzy inference system based approach has been termed fuzzy Markov games as it adapts fuzzy Q-learning approach to Markov games.

Next, we investigate sufficiency of using a minimax based learning in Markov games. We propose two novel hybrid Markov game based control approaches that incorporate 'fictitious play' and 'cautious fictitious play' algorithms from the game theory literature for improving performance of Markov game based controllers. Proposed fictitious play approach to Markov game based control attempts at 'safe and consistent' control, i.e., obtaining a performance superior to the pure Markov game approach. Objective of Markov game based hybrid control using cautious fictitious play is to strive for 'safety and universal consistency', i.e., reasonable control performance against all environments. Empirical evidence shows superiority of proposed hybrid methodologies over baseline Markov game based and other non-Markov game based approaches. We give convergence analysis of the hybrid Markov game based control.

In addition to the work on Markov game based control, our contributions include work on hybridizing the model based and model free RL approaches for the non-game or MDP based control. We propose a novel approach termed Trajectory based Modified Policy Iteration (TMPI) that uses finite-length trajectories through the state-space, to iteratively build and improve an MDP model along with cost-to-go value estimates. Empirical comparison of TMPI, against other RL algorithms illustrates effectiveness of the TMPI approach.

Table of Contents