



Analysis Seminar

Near Optimality of Finite Memory Policies for Partially Observed Stochastic Control Problems under Filter Stability

By

Ali Devran Kara
(University of Michigan)

Abstract: The talk focuses on partially observed Markov Decision processes (POMDPs). In POMDPs, the existence of optimal policies has in general been established via converting the original partially observed stochastic control problem to a fully observed one on the belief space, leading to a belief-MDP. However, computing an optimal policy for this fully observed model using classical methods is challenging even if the original system has a finite state and action spaces, since the state space of the fully observed belief-MDP model is always uncountable. We provide an approximation technique for POMDPs that use a finite window history of past information variables. We establish near optimality of finite window control policies in POMDPs under filter stability conditions and the assumption that the measurement and action sets are finite (and the state space is real vector valued). We also establish a rate of convergence result which relates the finite window memory size and the approximation error bound, where the rate of convergence is exponential under the filter stability conditions, where filter stability refers to the correction of an incorrectly initialized filter for a partially observed stochastic dynamical system (controlled or control-free) with increasing measurements.

Finally, we establish the convergence of the associated Q learning algorithm for control policies using such a finite history of past observations and control actions (by viewing the finite window as a 'state') and we show near optimality of such limit Q functions under the filter stability condition.

While there exist many experimental results for POMDPs, (i) the near optimality with an explicit rate of convergence (in the memory size) and relations to filter stability, and (ii) the asymptotic convergence (to the approximate MDP value function) for such finite-memory Q-learning algorithms are results that are new to the literature, to our knowledge.

-Joint work with Serdar Yuksel (Queen's University).

Date: Friday, April 29, 2022

Time: 18:00-19:00, GMT+3

Place: ZOOM

To request the event link, please send a message to goncha@fen.bilkent.edu.tr