

Bounded Rationality Methods in Reinforcement Learning-driven Autonomy



Kyriakos Vamvoudakis

The Daniel Guggenheim School of Aerospace
Engineering at Georgia Tech

Thursday, February 17, 2022
10:30 AM • POTR 234 (Fu Room)

Zoom info: <https://purdue-edu.zoom.us/j/93806327446> ~ Meeting ID: 938 0632 7446

Please note the confidentiality of this seminar and do not post publicly

Abstract

Autonomous systems will be tasked with operating in complex, human-centric environments both in cooperation and in competition with humans and other agents. In this talk, issues of prediction in the context of game theory and multi-agent cyber-physical systems will be addressed. To account for the cognitive limitations of human and machine decision-makers, we introduce ideas and principles of bounded rationality for autonomy using tools from control theory and reinforcement learning. Specifically, we will formulate level-k thinking and cognitive hierarchy models in nonlinear and linear noncooperative differential games, where each agent is assigned an intelligence level corresponding to a number of thinking iterations. The applicability of this approach will be highlighted via the example of a pursuit evasion game between Unmanned Aerial Vehicles. The versatility of the proposed methods will be shown via results of level-k thinking in discrete stochastic games. Finally, in order to design more advanced decision-making algorithms that explicitly exploit the learning abilities of the other agents in the environment, a meta-learning framework in games will be presented, via which an autonomous agent can achieve learning manipulation and deception.

Bio

Kyriakos G. Vamvoudakis was born in Athens, Greece. He received the Diploma (a 5-year degree, equivalent to a Master of Science) in Electronic and Computer Engineering from the Technical University of Crete, Greece in 2006 with highest honors. After moving to the United States of America, he studied at The University of Texas and he received his M.S. and Ph.D. in Electrical Engineering in 2008 and 2011 respectively. From May 2011 to January 2012, he was working as an Adjunct Professor and Faculty Research Associate at the University of Texas at Arlington and at the Automation and Robotics Research Institute. During the period from 2012 to 2016 he was a project research scientist at the Center for Control, Dynamical Systems and Computation at the University of California, Santa Barbara. He was an assistant professor at the Kevin T. Crofton Department of Aerospace and Ocean Engineering at Virginia Tech until 2018. He currently serves as an Assistant Professor at The Daniel Guggenheim School of Aerospace Engineering at Georgia Tech. He holds a secondary appointment in the School of Electrical and Computer Engineering. His research interests include reinforcement learning, control theory, cyber-physical security, bounded rationality, and safe/assured autonomy. Dr. Vamvoudakis is the recipient of a 2019 ARO YIP award, a 2018 NSF CAREER award, a 2021 GT Chapter Sigma Xi Young Faculty Award, and of several international awards including the 2016 International Neural Network Society Young Investigator (INNS) Award, the Best Paper Award for Autonomous/Unmanned Vehicles at the 27th Army Science Conference in 2010, the Best Presentation Award at the World Congress of Computational Intelligence in 2010, and the Best Researcher Award from the Automation and Robotics Research Institute in 2011. He currently is a member of the IEEE Control Systems Society Conference Editorial Board, an Associate Editor of: Automatica; IEEE Computational Intelligence Magazine; IEEE Transactions on Systems, Man, and Cybernetics: Systems; IEEE Transactions on Artificial Intelligence; Neurocomputing; Journal of Optimization Theory and Applications; IEEE Control Systems Letters; and of Frontiers in Control Engineering-Adaptive, Robust and Fault Tolerant Control. He is also a registered Electrical/Computer engineer (PE), a member of the Technical Chamber of Greece, and a Senior Member of both IEEE and AIAA.

Host: Professor Shreyas Sundaram, sundara2@purdue.edu, 765-496-0406