

Deceptive Planning for Supervised Autonomous Agents

Mustafa Karabag
University of Texas at Austin



Wednesday, October 5, 2022
10:30AM • EE 317

Zoom link: <https://purdue-edu.zoom.us/j/94546264748?pwd=MLF5ampZRzdwV2lNS3JCM3RJSmxYZz09>

Abstract

Deception naturally emerges in situations that involve two competing parties. In an adversarial environment, an autonomous agent needs to employ deceptive strategies to hide its intentions and maintain its continued operation. In this talk, I will discuss the intention deception problem of supervised agents. I will first formulate deception as a hypothesis testing problem and discuss the synthesis of optimal deceptive policies for fully observable Markov decision processes. As a dual problem, I will discuss the computational complexity of synthesizing optimal reference policies that prevent deception and provide a decomposable minimax optimization method for synthesizing optimal reference policies. Finally, I will address the deception problem under partial observability and in the context of two-player stochastic games.

Bio

Mustafa O. Karabag is a Ph.D. candidate at the University of Texas at Austin in Electrical & Computer Engineering. Before joining Ph.D., he earned his B.S. from Bogazici University in Electrical & Electronics Engineering. His research focuses on sequential decision-making for autonomous systems in adversarial or information-scarce environments with a particular interest in deception in autonomy.

Host: Abolfazl Hashemi (abolfazl@purdue.edu) and Mahsa Ghasemi (mahsa@purdue.edu)