HUMANITY CENTERED ROBOTICS INITIATIVE

BROWN

Stefanos Nikolaidis
Carnegie Mellon University

"Formalizing Human-Robot Collaboration"

Friday, April 14, 2017
2-3pm
CIT 477 Lubrano Conference Room

Abstract: The goal of my research is to improve human-robot collaboration by integrating mathematical models of human behavior into robot decision making. I develop game-theoretic algorithms and probabilistic planning techniques that reason over the uncertainty in the human internal state and its dynamics, enabling robots to act optimally in a variety of real-world collaborative settings. I first focus on robots that follow the human preference, and I propose cross-training, an algorithm inspired by human team training practices. Cross-training is an effective way for humans to communicate their preference to their robotic counterparts in human-robot collaboration settings, allowing robots to integrate seamlessly into human teams.

What if the communicated preference is suboptimal, however? Using insights from behavioral economics, I propose models of human behavior that allow autonomous systems to reason over the effects of their actions to future human actions, guiding humans towards better ways of doing their task. I start with a bounded-memory model of human adaptation parameterized by the human adaptability - the probability of the human switching towards a strategy newly demonstrated by the robot. I then examine more subtle forms of adaptation, where the human teammate adapts to the robot, without replicating the robot's policy. I model the interaction as a repeated game, and present an optimal policy computation algorithm that has complexity linear to the number of robot actions. Integrating these models into robot action selection allows for human-robot mutual-adaptation: the robot adapts its own actions, based on the probability of the human adapting to the robot. I show that mutual adaptation significantly improves human-robot team performance, compared to one-way robot adaptation to the human, in a variety of collaboration and shared-autonomy settings..

Stefanos Nikolaidis is a PhD Candidate at the Personal Robotics Lab in Carnegie Mellon's Robotics Institute, working with Prof. Siddhartha Srinivasa. His research lies at the intersection of human-robot interaction, algorithmic game-theory and planning under uncertainty. He develops decision-making algorithms that leverage mathematical models of human behavior to support deployed robotic systems in real-world collaborative settings. Stefanos has a MS from MIT, a MEng from the University of Tokyo and a BS from the National Technical University of Athens. He has additionally worked as a research specialist at MIT and as a researcher at Square-Enix in Tokyo. He has received a Best Enabling Technologies Award from the IEEE/ACM International Conference on Human-Robot Interaction and was a Best Paper Award Finalist in the International Symposium on Robotics.

Host: George Konidaris/HCRI