Decision-Theoretic User Modeling for Human-Robot Mutual Adaptation

Tuesday, September 6, 2016
GHC 4405
9:00 a.m.

Abstract

In many application domains, robots co-exist in the same physical space with humans and aim to become trustworthy partners. We particularly envision personal robots arranging furniture with a human partner, manufacturing robots performing spar assembly with human co-workers, or rehabilitation robots assisting spinal cord injury patients. In such collaborative settings, humans often have inaccurate models of the robot capabilities, which leads the team towards suboptimal strategies. On the other hand, the robot frequently knows the optimal way of executing the task based on some objective performance metric. This thesis proposes a set of decision-theoretic models of human teammates, that allow the robot to reason in a principled way over the effects of its actions on the future human behavior, and guide the human towards new, optimal strategies, unknown to them in advance. We formalize human adaptability, that is their willingness to adapt to a robot strategy, and propose a human-robot mutual adaptation formalism based on a bounded-memory model. We evaluate the impact of adaptability on collaboration paradigms: a shared-location collaborative task, and a shared-autonomy setting. We show that the formalism significantly improves team performance when the starting human preference of executing the task is suboptimal. We expect that the proposed models will increase task performance, human trust in the robot and perceived collaboration on a variety of joint-action collaborative tasks.