Direct Predictive Collaborative Control of a Prosthetic Arm

Craig Sherstan
Reinforcement Learning & Artificial Intelligence Lab
Department of Computer Science
University of Alberta
Edmonton, AB
sherstan@ualberta.ca

Joseph Modayil
Reinforcement Learning & Artificial Intelligence Lab
Department of Computing Science
University of Alberta
Edmonton, AB
jmodayil@cs.ualberta.ca

Patrick M. Pilarski
Reinforcement Learning & Artificial Intelligence Lab &
Division of Physical Medicine and Rehabilitation
University of Alberta
Edmonton, AB
pilarski@ualberta.ca

Abstract

We have developed an online learning system for the collaborative control of an assistive device. Collaborative control is a complex setting requiring a human user and a learning system (automation) to co-operate towards achieving the user's goals. There are many control domains where the number of controllable functions available to a user surpass what a user can attend to at a given moment. Such domains may benefit from having automation assist the user by controlling those unattended functions. How exactly this interaction between user decision making and automated decision making should occur is not clear, nor is it clear to what degree automation is beneficial or desired. We should expect such answers to vary from domain to domain and possibly from moment to moment. One domain of interest is the control of powered prosthetic arms by amputees. Upper-limb amputees are extremely limited in the number of inputs they can provide to a prosthetic device and typically control only one joint at a time with the ability to toggle between joints. Control of modern prostheses is often considered by users to be laborious and non-intuitive. To address these difficulties, we have developed a collaborative control framework called Direct Predictive Collaborative Control (DPCC), which uses a reinforcement learning technique known as general value functions to make temporal predictions about user behavior. These predictions are directly mapped to the control of unattended actuators to produce movement synergies. We evaluate DPCC during the human control of a powered multi-joint arm. We show that DPCC improves a user's ability to perform coordinated movement tasks. Additionally, we demonstrate that this method can be used without the need for a specific training environment, learning only from user's behavior. To our knowledge this is also the first demonstration of the combined use of the new True Online TD(λ) algorithm with general value functions for online control.

Keywords: collaborative control, prediction, temporal-difference learning, prosthetics, assistive robotics

Acknowledgements

The authors would like to thank Harm van Seijen and Richard Sutton for a number of helpful discussions, and acknowledge support from the Alberta Innovates Centre for Machine Learning (AICML), Alberta Innovates – Technology Futures (AITF), and the Natural Sciences and Engineering Research Council of Canada (NSERC).