Real-time Machine Learning in Rehabilitation Robotics for Adaptable Artificial Limbs

Patrick M. Pilarski, Michael R. Dawson, Thomas Degris, Farbod Fahimi, Jason P. Carey, and Richard S. Sutton

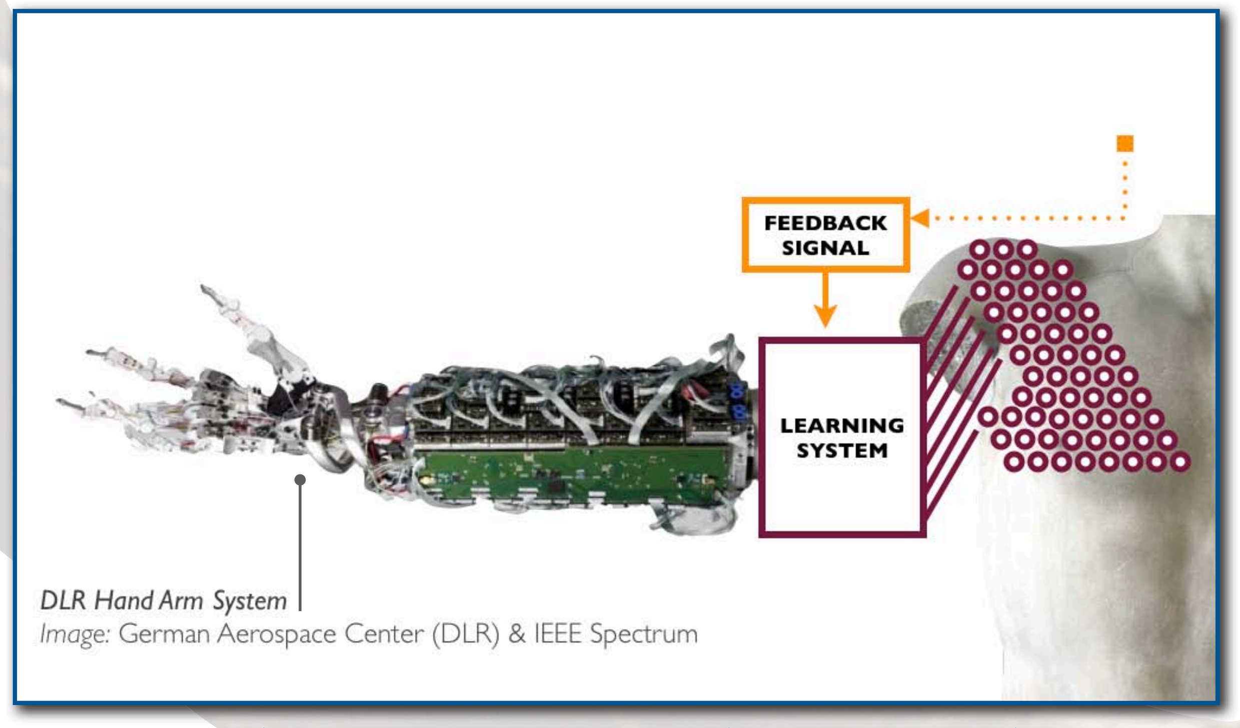
Reinforcement Learning and Artificial Intelligence Group & Alberta Innovates Centre for Machine Learning, Dept. of Computing Science, University of Alberta

Alberta Innovates Technology Futures

MYOELECTRIC CONTROL

Myoelectric prostheses are assistive rehabilitation devices that interpret muscle signals from an amputee's body to actuate a multiple-joint robotic appendage. While advances in biosensor and robotics technology promise to greatly improve the utility of these devices, new electromyographic (EMG) control methods are required to deal with the corresponding increase in available sensor information and actuation capability. In addition, current myoelectric control methods lack the ability to adapt online to changes in amputee use patterns or physical condition; controller calibration and improvement is largely impossible outside of a clinical setting.

LEVERAGING COMPLEXITY FOR ADVANCED CONTROL



Above: machine learning approach to complex human-machine interfaces.

In this work, we present a real-time machine learning approach for generalized myoelectric control that promises to flexibly scale to large and diverse sensorimotor spaces.

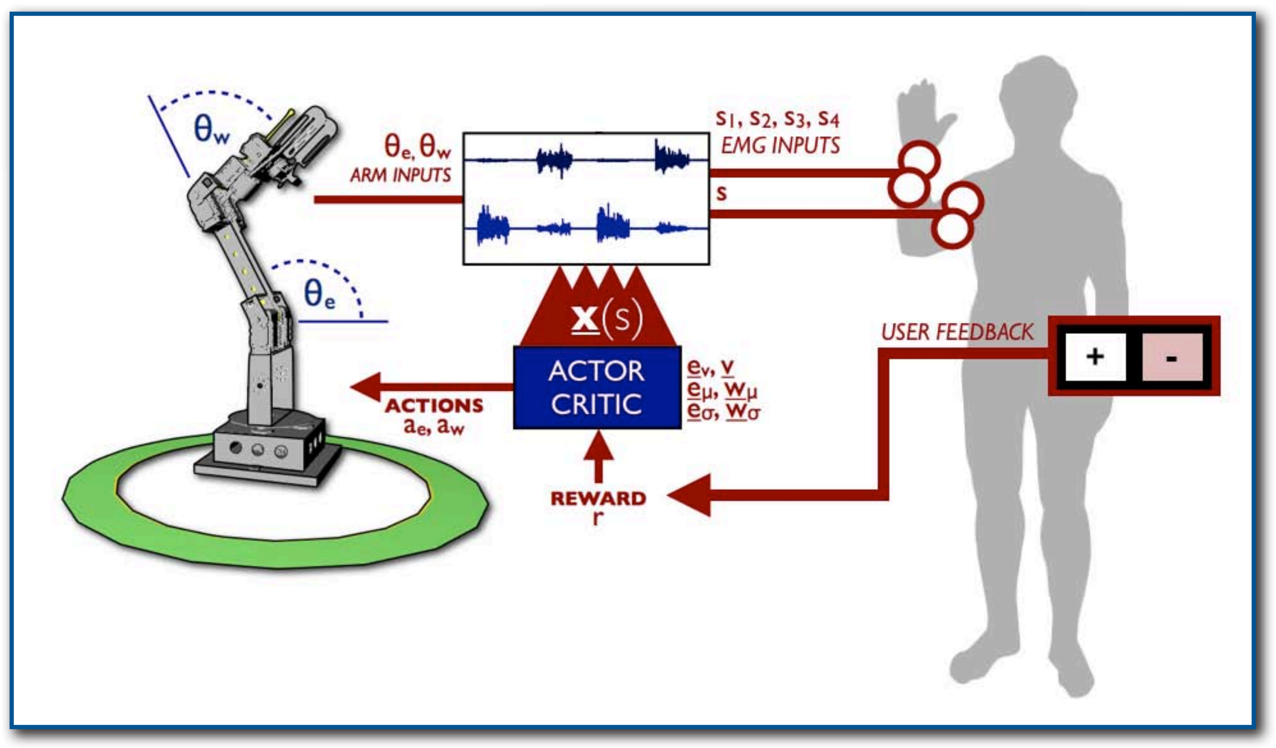
This learning controller: (1) can be trained in an online, ongoing fashion by a human user, (2) does not require expert knowledge to adjust for different users, sensors, and actuators, and (3) requires no *a priori* data.

ONLINE HUMAN TRAINING OF A REINFORCEMENT LEARNING CONTROL SYSTEM

Reinforcement learning (RL) is a machine learning approach for solving optimal control problems. In RL, a control policy is learned through repeated trial and error interactions between a learning system and its environment. A system aims to maximize the expected sum of a scalar feedback signal, termed *reward*, even in cases where an *a priori* model of the problem domain is unavailable.

Algorithm 1 Continuous Actor-Critic Algorithm

initialize: $\mathbf{w}_\mu, \mathbf{w}_\sigma, \mathbf{v}, \mathbf{e}_\mu, \mathbf{e}_\sigma, \mathbf{e}_\mathbf{v}, s$
repeat:
$\mu \leftarrow \mathbf{w}_\mu^T \mathbf{x}(s)$
$\sigma \leftarrow \exp[\mathbf{w}_\sigma^T \mathbf{x}(s) + \log(\sigma_c)]$
$a \leftarrow \mathcal{N}(\mu, \sigma^2)$
take action a, **observe** r, s'
$\delta \leftarrow r + \gamma \mathbf{v}^T \mathbf{x}(s') - \mathbf{v}^T \mathbf{x}(s)$
$\mathbf{e}_\mathbf{v} \leftarrow \lambda \mathbf{e}_\mathbf{v} + \mathbf{x}(s)$
$\mathbf{v} \leftarrow \mathbf{v} + \alpha_\mathbf{v} \delta \mathbf{e}_\mathbf{v}$
$\mathbf{e}_\mu \leftarrow \lambda \mathbf{e}_\mu + (a - \mu)\mathbf{x}(s)$
$\mathbf{w}_\mu \leftarrow \mathbf{w}_\mu + \alpha_\mathbf{w} \delta \mathbf{e}_\mu$
$\mathbf{e}_\sigma \leftarrow \lambda \mathbf{e}_\sigma + [(a-\mu)^2/\sigma^2 - 1]\mathbf{x}(s)$
$\mathbf{w}_\sigma \leftarrow \mathbf{w}_\sigma + \alpha_\mathbf{w} \delta \mathbf{e}_\sigma$
$s \leftarrow s'$



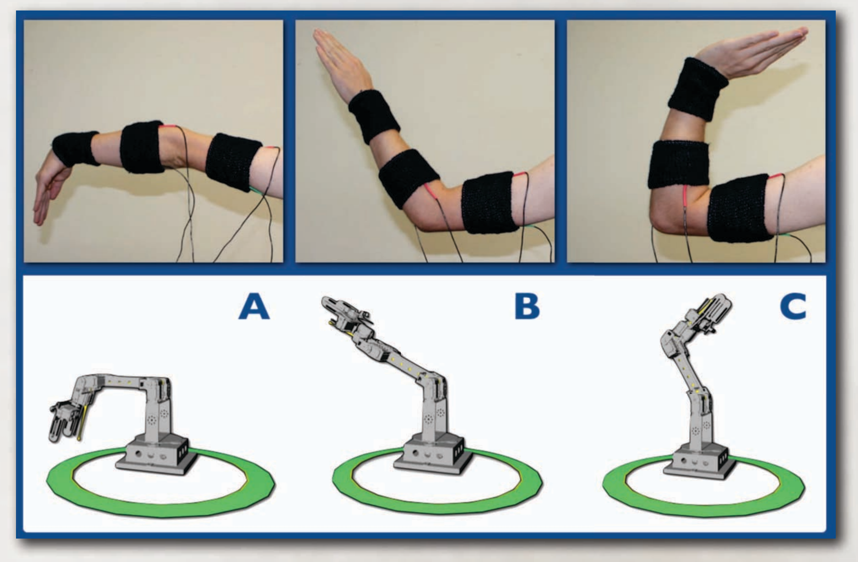
Above: Algorithm and schematic diagram of actor-critic reinforcement learning applied to a two-joint robotic arm. At each time step, a state approximation ***x(s)*** and a scalar user-provided reward signal ***r*** are given to the learning system. Based on this, the system updates its control policy and generates two continuous joint velocity actions that are given as input to the robotic arm.

P.M. Pilarski, T. Degris, and R.S. Sutton are with the Department of Computing Science, University of Alberta, Edmonton, AB, Canada; J.P. Carey is with the Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada; M.R. Dawson is with the Glenrose Rehabilitation Hospital, Edmonton, AB, Canada; F. Fahimi is with the Department of Mechanical and Aerospace Engineering, University of Alabama in Huntsville, Huntsville, AL, USA. **Please direct correspondence to:** pilarski@ualberta.ca

EXPERIMENTS

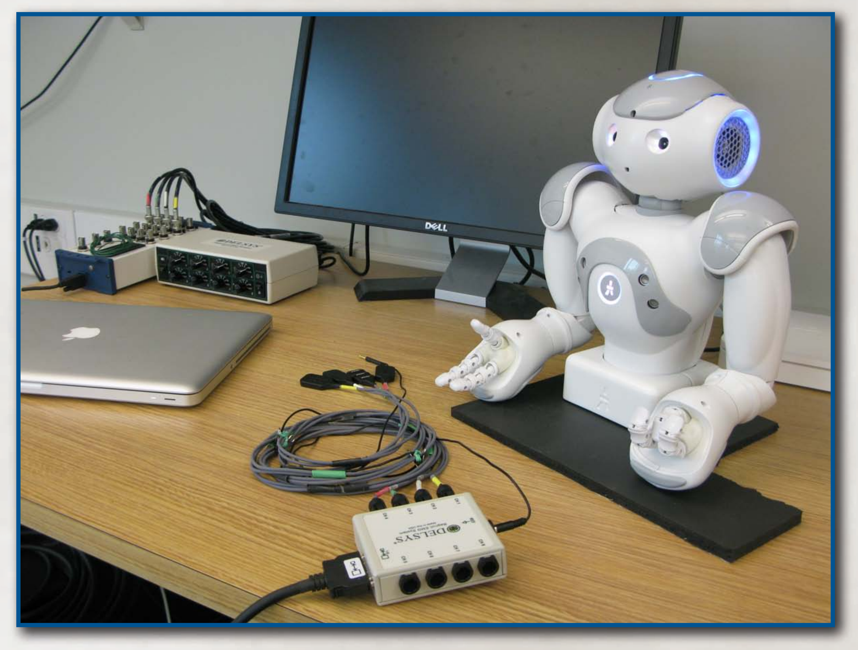
We assessed this control method using able-bodied subjects. While subjects engaged in a series of movement tasks, EMG signals were recorded using electrodes affixed to the skin of their arms and torso. The learning system was given interactive positive or negative reward feedback according to how well it transformed these EMG signals into control commands for an artificial limb (below).

A simulated upper-arm prosthesis:



Right: Simulated AX-12 Smart Arm and motion targets. EMG signals were recorded with four BL-AE-N pre-amplified surface electrodes via a National Instruments PCI-6259 data acquisition card at a sampling rate of 200Hz.

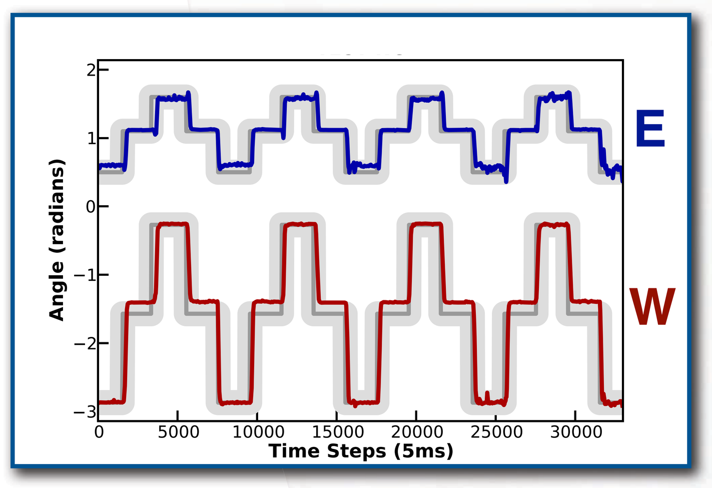
A physical robotic limb:



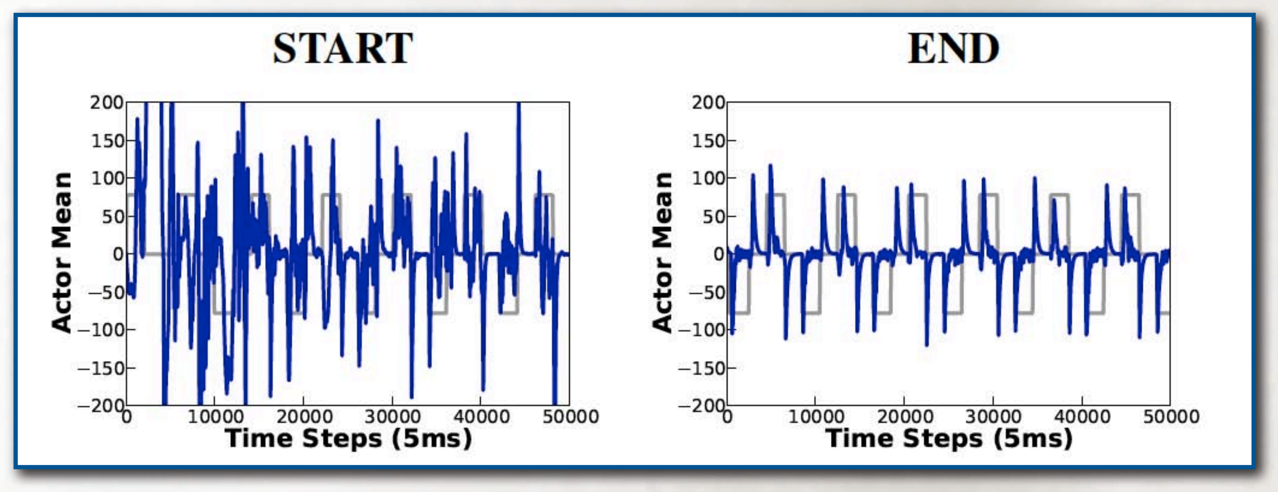
Right: Nao T14 robot torso (Aldebaran Robotics, France) and EMG system. EMG signals used in device control and learning were obtained using a Bagnoli-8 (DS-B03) EMG System with four DE-3.1 Double Differential Detection EMG sensors (Delsys, Boston, USA), and a National Instruments USB-6216 BNC analog-to-digital signal converter. Electrodes were sampled at 40Hz.

RESULTS & DISCUSSION

These experiments showed that the learning system was able to interpret online myoelectric input and human guidance to form user-specific control policies in real time. In addition, the system demonstrated the ability to adapt online to changes in user preferences and the experimental environment. *Below:* AX-12 motion profile results.



ANGLE TRACKING AFTER HUMAN TRAINING



VELOCITY PROFILE AFTER HUMAN TRAINING

This type of adaptable, online learning controller represents an important step forward for artificial limb technology. In the long term, we expect this approach to increase the autonomy of amputees, while decreasing clinical costs and removing commercialization barriers for the next generation of truly intelligent prosthetic devices.

FOR MORE DETAIL AND REFERENCES: P.M. Pilarski, M.R. Dawson, T. Degris, F. Fahimi, J.P. Carey, and R.S. Sutton, *Proc. of the 2011 IEEE International Conference on Rehabilitation Robotics*, June 29–July 1, 2011, Zurich, Switzerland, pp. 134-140.

This work was made possible through support from **Alberta Innovates—Technology Futures,** the **Alberta Innovates Centre for Machine Learning,** the **Natural Sciences and Engineering Research Council,** and the **Glenrose Rehabilitation Hospital.** We greatly appreciate the contributions of these partners and funders.

UNIVERSITY OF ALBERTA