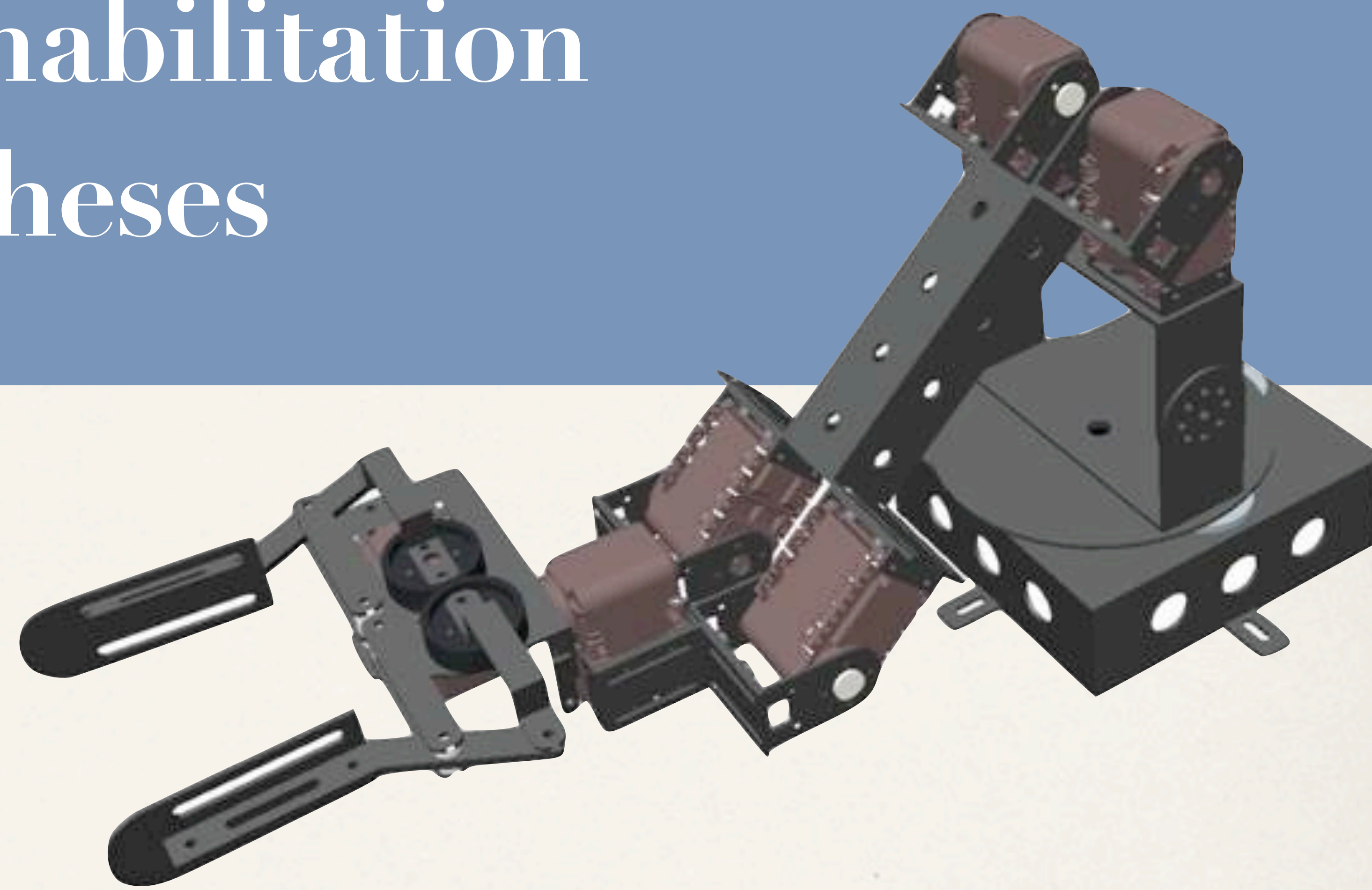
Reinforcement Learning Artificial Intelligence in Rehabilitation for Adaptive Prostheses



Patrick M. Pilarski

Postdoctoral Fellow, RLAI & AICML
University of Alberta

Joint work with Michael Rory Dawson, Thomas Degris, Farbod Fahimi, Jason Carey, *and* Richard S. Sutton

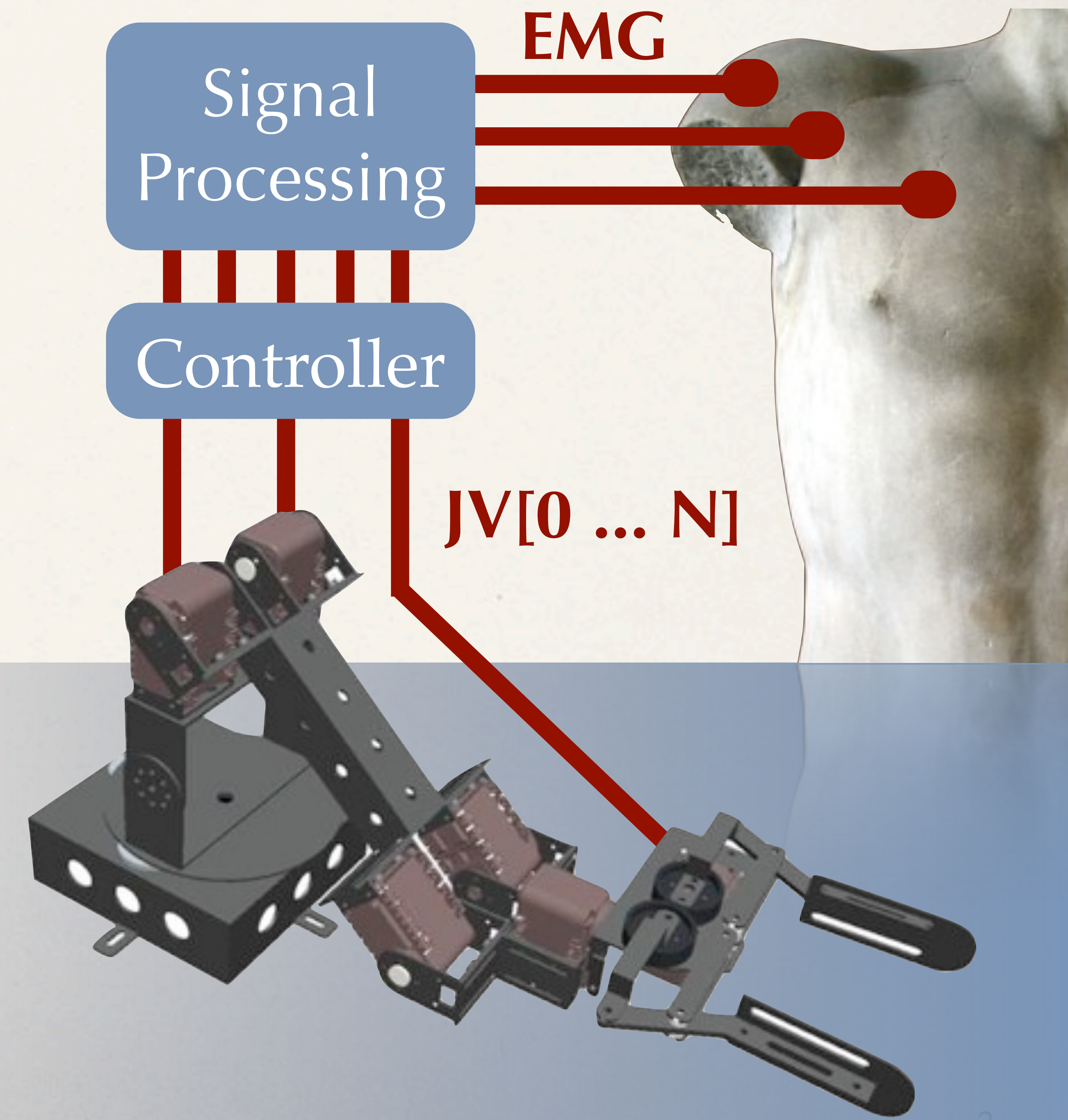
November 3rd, 2010

Overview

- **Prostheses: Upper Arm Myoelectric Control**
  - Clinical relevance.
  - The challenges of multi-function myoelectric control.
  - Changing data & the need for adaptation.
- **Reinforcement Learning Artificial Intelligence**
- **Applications of RL to Prosthetics and Myoelectric control.**
  - Results from an EMG-based control task.
- **Conclusions and Thoughts to Leave With**

Multi-Function Prosthetics

- Devices that monitor electrical signals produced by muscle tissue in limb-deficient patients (EMG signals).
- Use these signals to control the movement of a multiple-actuator robotic appendage.
- Can be from physiologically-mapped muscle sources, or from other muscle areas (Targeted Muscle Reinnervation, TMR).



Clinical Motivation

- Recent amputees find the transition to their new prosthetic device challenging & frustrating, often due to the complex control scheme.
- Use patterns and patient physiology change, often requiring ongoing calibration of the artificial limb by patients and physiotherapists.
- Adaptive, intuitive prostheses could help increase controllability and learning rates for patients, but there are no examples in clinical use.
- Current work at the U of A focuses on an inexpensive training tool for use by new TMR patients at the GRH (Dawson, Carey, Fahimi: MTT).

The Myoelectric Training Tool

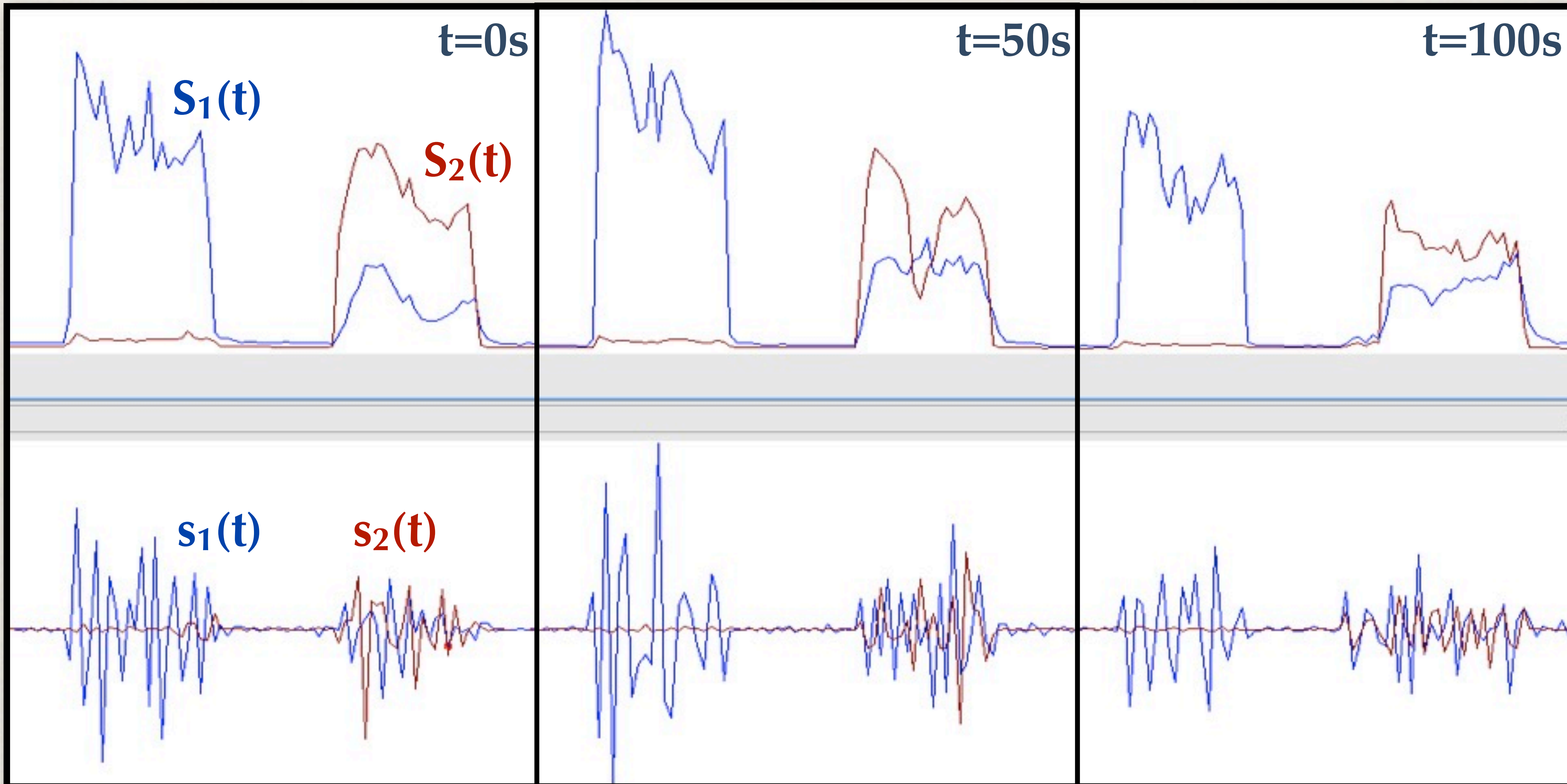
Video by Michael Rory Dawson

Complex Time-Varying Data

- Patient intent is not mapped directly to EMG data; the same intended command may generate widely varying muscle activation patterns.
- EMG signals from different muscle groups may overlap in unpredictable and/or detrimental ways.
- Signal amplitude and frequency components may change as the body, sensors, and environmental conditions change.
- Signal drift can happen over a period of minutes, days, or weeks.

Parker et al., Journal of Electromyography and Kinesiology (2006)

Complex Time-Varying Data



Open Questions for Research

1. How best to translate multiple and possibly overlapping muscle signals into usable control commands for a mechanical limb.

2. How to automatically tailor the system to needs and specific physical conditions of individual patients, without constant manual intervention and periods of frustration and/or reduced function.

3. How to improve limb control based on (sparse) patient feedback.

These directly relate to fundamental problems for BCI/HCIs.

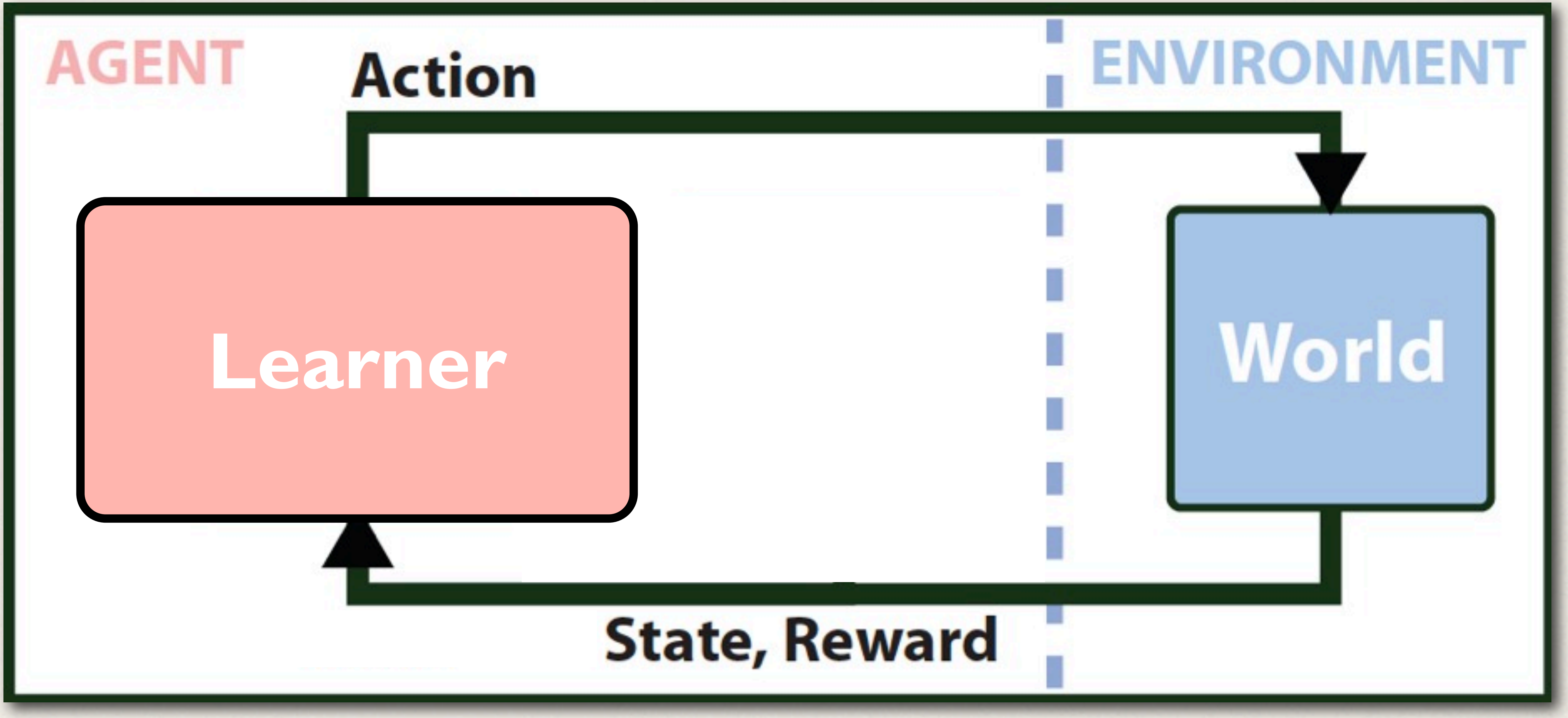
Parker et al., Journal of Electromyography and Kinesiology (2006)

Reinforcement Learning
Artificial Intelligence

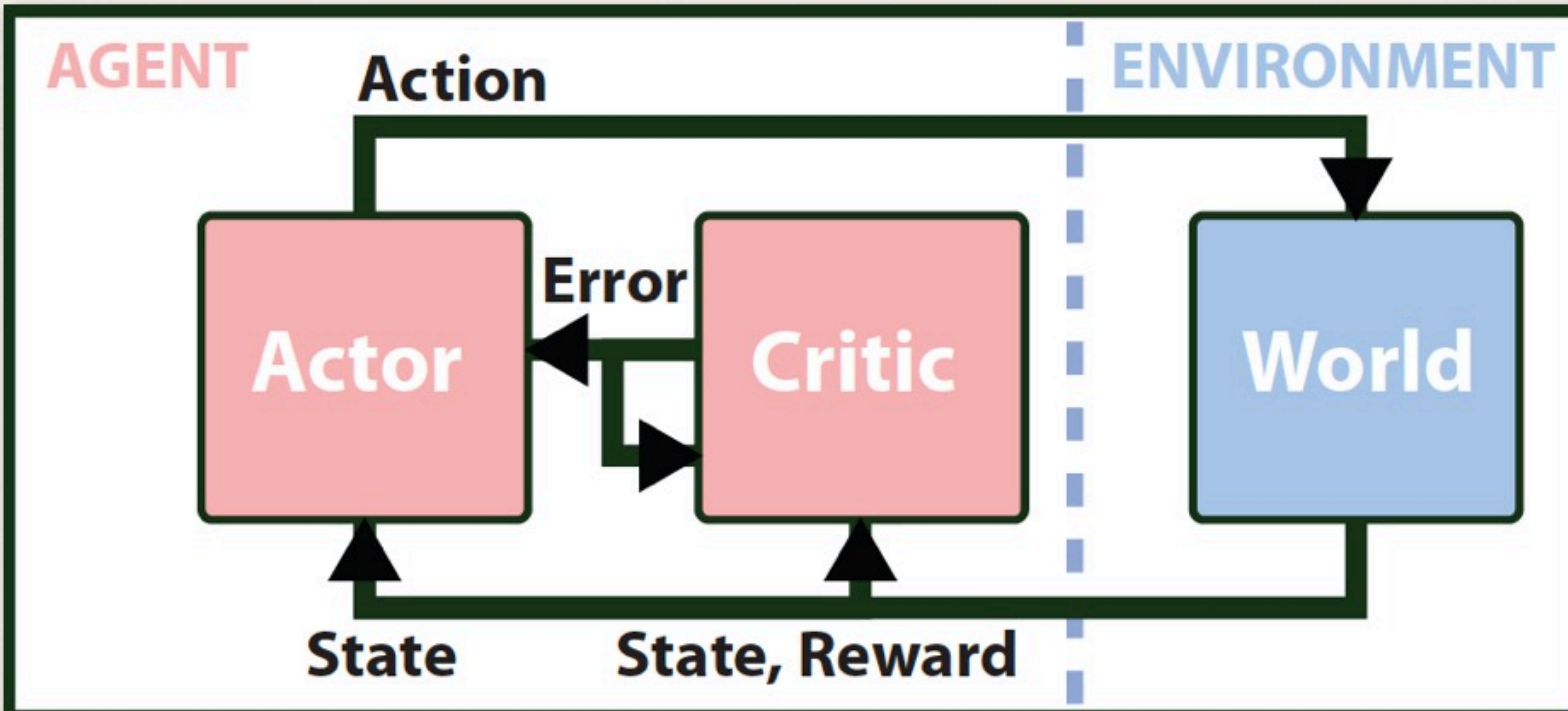
- Reinforcement learning (RL) involves an **agent** and an **environment**.
- The agent perceives the state of the environment via a set of **observations** and takes **actions**.
- It then receives a new set of observations and a **reward** from the environment.
- These observations and rewards are used to predict *future* rewards, and to change the agent's **policy** (how it selects actions).
- **Key point:** RL methods involve **semi-supervised learning**. A single, scalar reward signal drives learning.

Sutton and Barto, MIT Press (1998)

Continuous Actor-Critic Reinforcement Learning



Continuous Actor-Critic Reinforcement Learning



This has roots in *Bhatnagar et al., Automatica (2009); Williams, Machine Learning (1992)*

What does this mean for real-world applications?

- RL systems can learn well when an end goal or desired behaviour is known but it is difficult (or impossible) to model the problem domain.
- Fast computation and low memory requirements allow for realtime deployment, especially on embedded or distributed systems.
- This also permits online adaptation: the learner can change in response to user needs and variation in the environment. This increases the robustness and versatility of systems.
- Very little hand tuning is required, and automatic tuning further reduces the need for ongoing maintenance. This saves human labour.

… and Specifically for Amputees?

- These artificial intelligence techniques can be used by health professionals *and patients* to increase the power of existing methods *without* the requirement for detailed technical knowledge (human training with no computer programming skills needed).

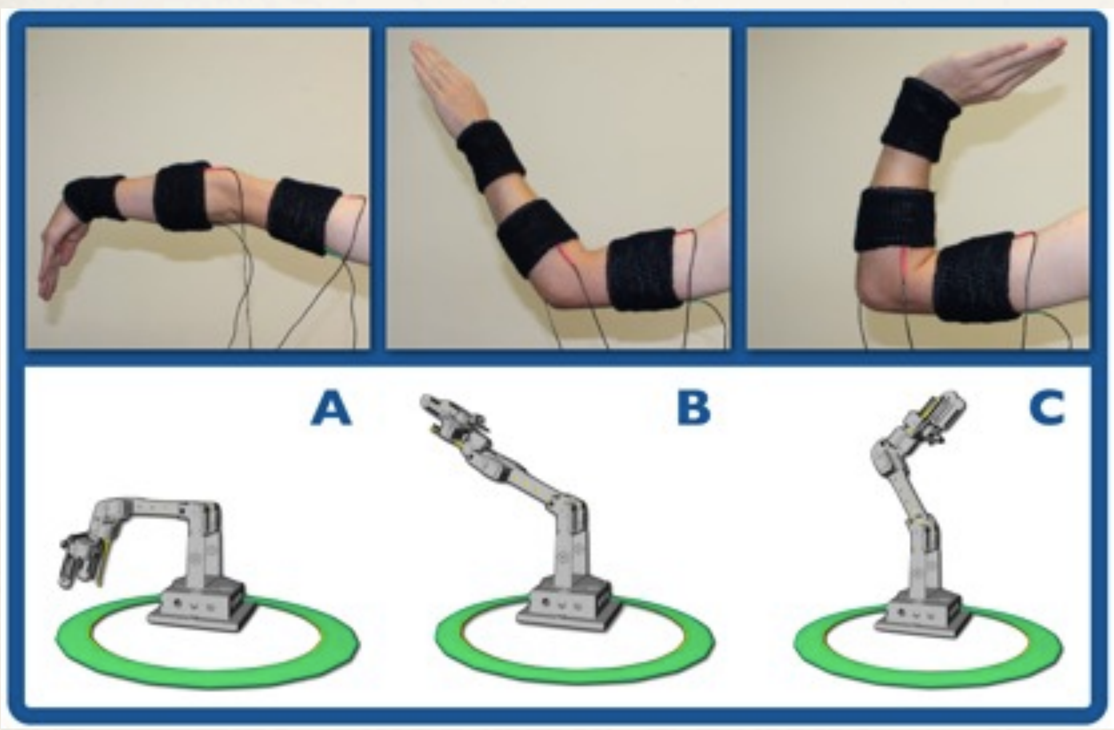
- Methods can flexibly adapt to the needs of individual patients and are not dependent on a fixed set of calibrations or sensor positions.

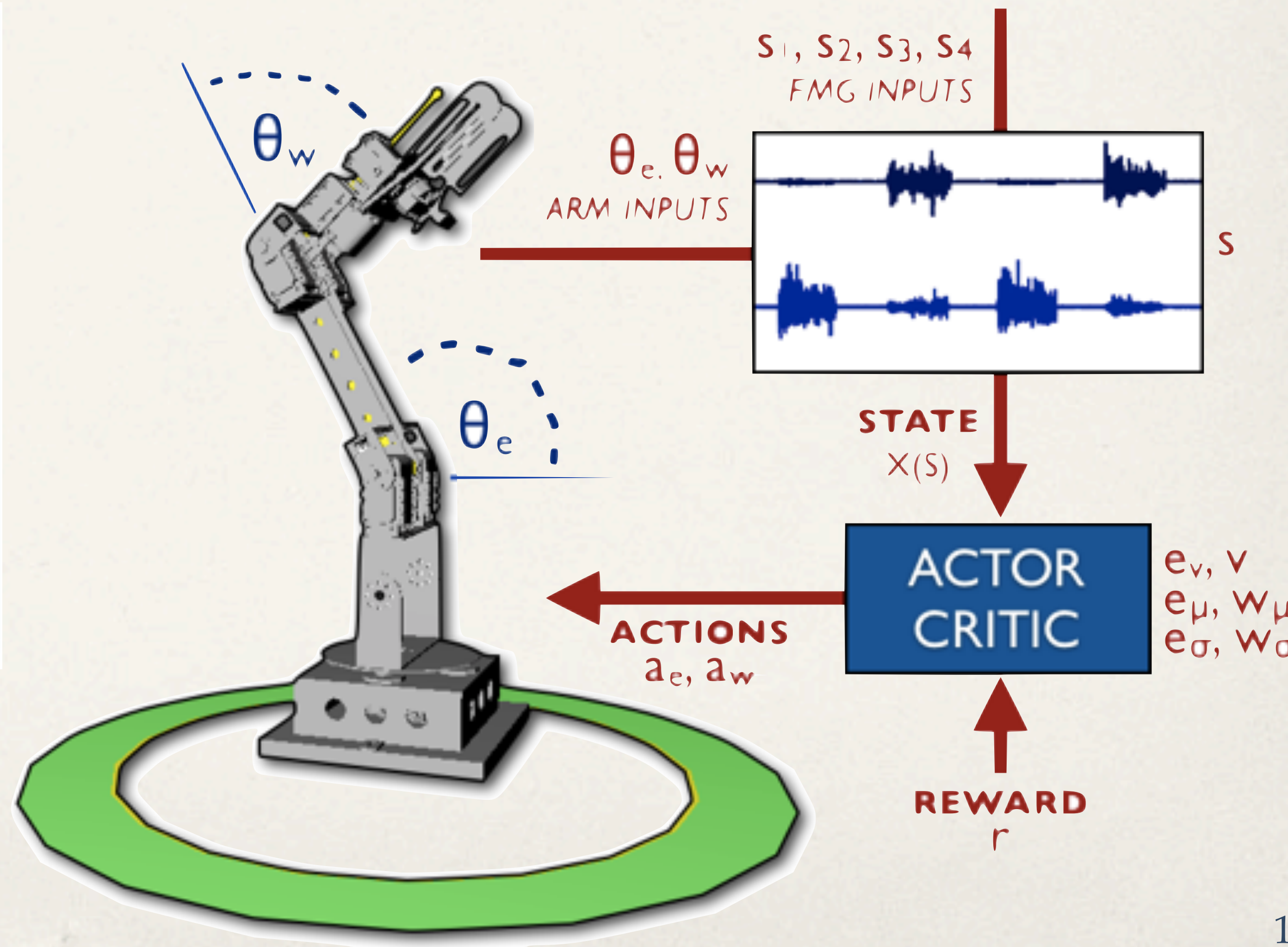
- Because these RL methods operate and learn in real-time, they can improve with time and training, and change with the patient (both in terms of biology and use patterns).

- Ability to perform fluid, multi-joint actions, not just staged motion.

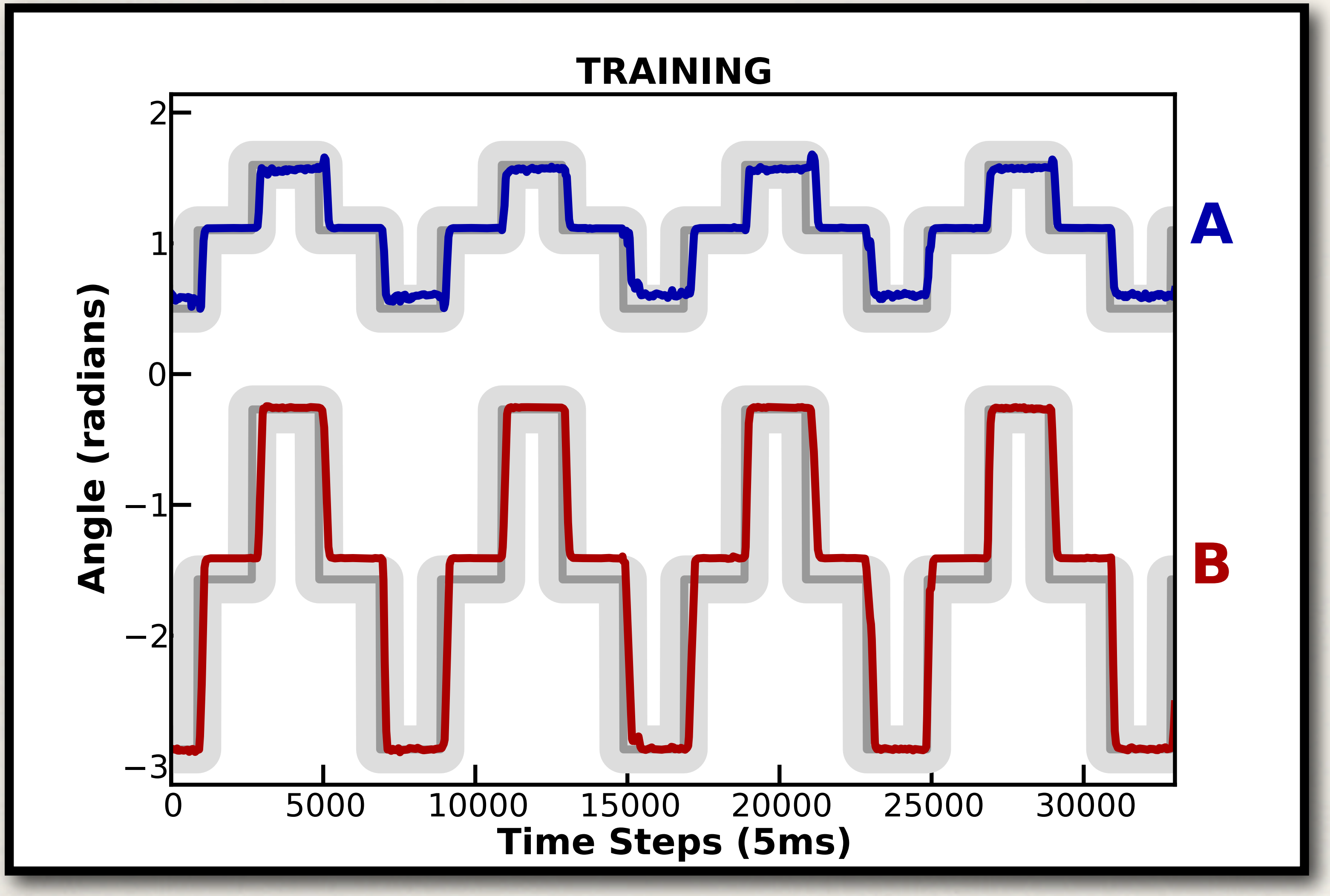
Example: EMG-based Control

- **Learning a robotic arm control policy with input from an able-bodied subject:** human performs a reaching task, and rewards the robotic arm when it performs the desired (correct) movements.



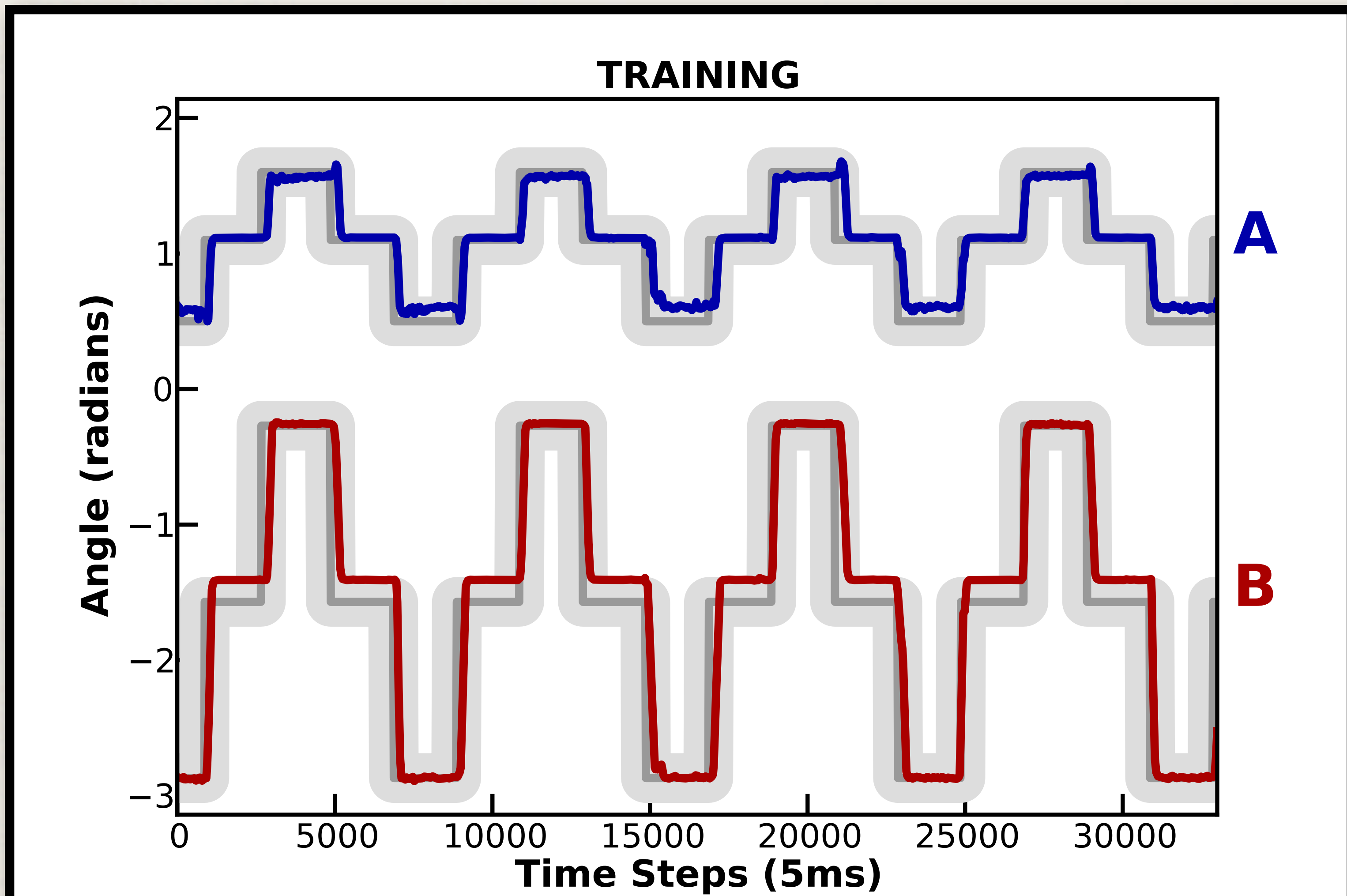


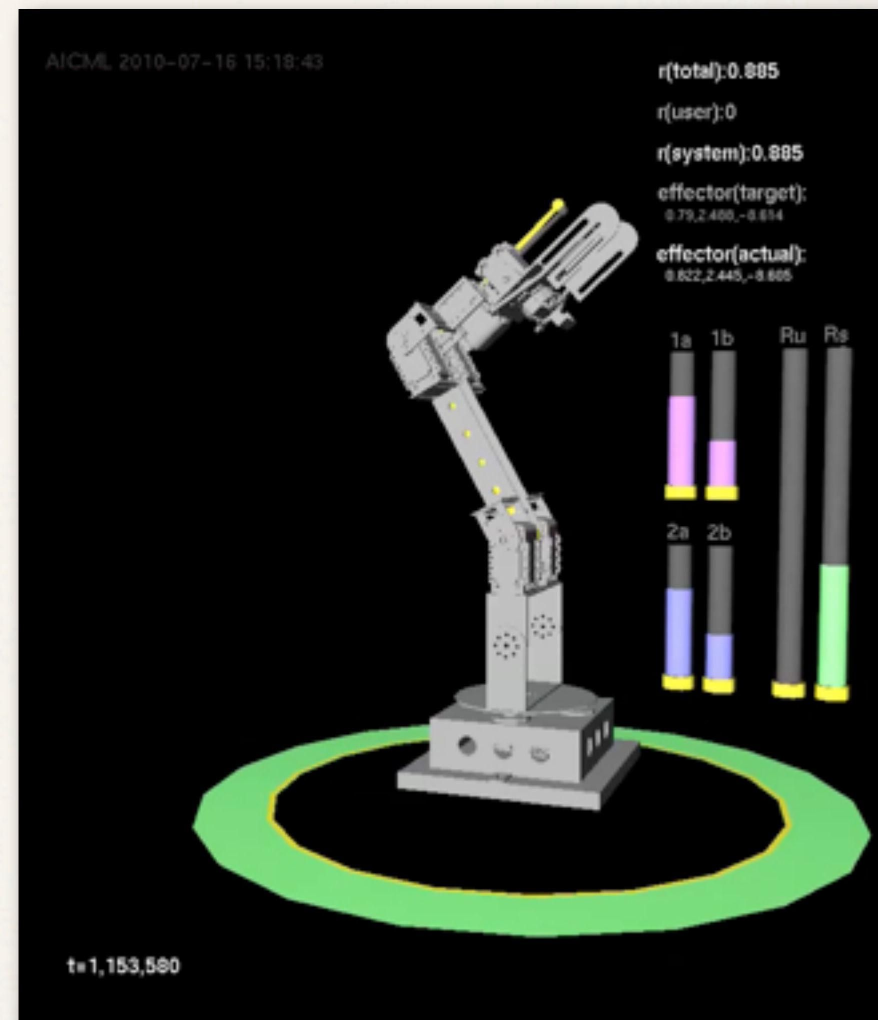
Results



TRAINING
Angle (radians)
2
1
0
−1
−2
−3
0
5000
10000
15000
20000
25000
30000
Time Steps (5ms)
A
B

Results


TRAINING
Angle (radians)
2
1
0
−1
−2
−3
0
5000
10000
15000
20000
25000
30000
Time Steps (5ms)
A
B


AICML 2010-07-16 15:18:43
r(total):0.885
r(user):0
r(system):0.885
effector(target):
effector(actual):
1a
1b
Ru
Rs
2a
2b
t=1,153,580

Key Messages to Leave With

- Reinforcement learning artificial intelligence methods are well suited to use in a biomedical problem domain (semi-supervised & flexible).
- Adaptive control methods of this type will increase the speed and success with which amputees can learn to use their powered prostheses, and improve patient artificial limb function.
- Facilitates devices that adapt to daily use patterns and changes in the patient, without the need for constant intervention by specialists.
- This points to more customized treatment, increased patient engagement, and reduced load on the medical system.

Acknowledgements

- **Collaborators:**

Michael Rory Dawson	*(MECE, U of Alberta)*
Thomas Degris	*(CS, U of Alberta)*
Farbod Fahimi	*(MECE, U of Alberta & U of Alabama)*
Jason Carey	*(MECE, U of Alberta)*
Richard S. Sutton	*(CS, U of Alberta)*

- **Funding and Support:**

Alberta Ingenuity Centre for Machine Learning ▪ iCORE
Glenrose Rehabilitation Hospital ▪ Alberta Innovates – TF
Natural Sciences and Engineering Research Council of Canada