Applied RL on Robots

Patrick M. Pilarski

Canada Research Chair in Machine Intelligence for Rehabilitation
Division of Physical Medicine and Rehabilitation, Dept. of Medicine

Fellow, Alberta Machine Intelligence Institute (Amii)
Applied RL on Robots

Patrick M. Pilarski

Canada Research Chair in Machine Intelligence for Rehabilitation
Division of Physical Medicine and Rehabilitation, Dept. of Medicine
Fellow, Alberta Machine Intelligence Institute (Amii)
IA
(Intelligence Amplification)

I get by AI with a little help from my friends’ brains

John Lennon
SPINNING WHEEL

{animal fur, plant matter}  
>>  
{thread, yarn}
... and, in short order ...
Marvin the paranoid android from THHGTTG.
Whole point of this talk:

*communicate one research design pattern that can help you tackle ambitious real-world problems*
Never storm a castle through the front gates...
Direct brain-computer interfaces: study participant Jan Scheuermann feeding herself with a robotic limb (University of Pittsburgh); [http://www.upmc.com/media/media-kit/bci/Pages/default.aspx](http://www.upmc.com/media/media-kit/bci/Pages/default.aspx)
Brain-body-machine interfaces: “Amputee Makes History with APL’s Modular Prosthetic Limb” (JHU Applied Physics Laboratory); https://youtu.be/9NOncx2jU0Q
Brain-body-machine interfaces: “APL’s Modular Prosthetic Limb Reaches New Levels of Operability” (JHU Applied Physics Laboratory); https://youtu.be/-0srXvOQlu0
Brain-body-machine interfaces: “Brain-Machine Interface @ EPFL- Wheelchair” (École polytechnique fédérale de Lausanne); https://youtu.be/0-1sdtnugcE
Commercially Deployed Pattern Recognition for Prostheses
Consumer-Available BCI and BMI

Muse

Myo (Thalmic Labs)
These examples all involve machine intelligence or machine learning.
Human Body

Machine Intelligence

Assistive Machine
Would you do trial-and-error control learning on all those I/O channels?
Whole point of this talk:

highlight prediction learning
Whole point of this talk:

highlight prediction learning as a foundation for more advanced control solutions
PREDICTIONS

Momentary.
(e.g., classification decision)

PREDICTIONS

Momentary. (e.g., classification decision)

Temporally extended. (e.g., expected return)


PREDICTIONS

Momentary. (e.g., classification decision)  Temporally extended. (e.g., expected return)

Can be acquired or updated in batches or in real time.


PREDICTION PRECEDES CONTROL

Flanagan et al., *Current Biology* 13(2), 2003: “Prediction precedes control in motor learning”
Desmurget et al., *Science* 324(5928), 2009: “Movement intention after parietal cortex stimulation in humans”
Hallmarks of Intelligence:
Artificial, Machine (and Human)
Hallmarks of Intelligence: Artificial, Machine (and Human)
human embodiment of a robot body part is really tricky...
University of Alberta: http://blinclab.ca, https://www.smartnetworkcentre.ca/
University of Alberta: http://blinclab.ca
### Planning and Meta-learning

<table>
<thead>
<tr>
<th>Control</th>
<th>2009</th>
</tr>
</thead>
</table>

#### Prediction

#### Representation
Online Human Training of a Myoelectric Prosthesis Controller via Actor-Critic Reinforcement Learning

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

Abstract—As a contribution toward the goal of adaptable, intelligent artificial limbs, this work introduces a continuous actor-critic reinforcement learning method for optimizing the control of multi-function myoelectric devices. Using a simulated upper-arm robotic prosthesis, we demonstrate how it is possible to derive successful limb controllers from myoelectric data using only a sparse human-delivered training signal, without requiring detailed knowledge about the task domain. This reinforcement-based machine learning framework is well suited for use by both patients and clinical staff, and may be easily adapted to different application domains and the needs of individual amputees. To our knowledge, this is the first myoelectric control approach that facilitates the online learning of new amputee-specific motions based only on a one-dimensional (scalar) feedback signal provided by the user of the prosthesis.

I. INTRODUCTION

Fig. 1. A schematic diagram of a continuous reinforcement learning application.
Planning and Meta-learning

Control

Prediction

Representation

2010
EXAMPLE 0

making predictions

(Nexting in real time)

~18k predictions about bits learned and made in real time


Whole point of this talk:

say the phrase “Pavlovian control” enough times that you remember it next week
Pavlovian control involves a
fixed mapping between learned
predictions and control actions.

EXAMPLE 1

adaptive switching

(predictions change an interface)
Adaptive Switching

Edwards et al., *MEC*, 2014
Edwards et al., *Prosthetics Orthotics Int.*, 2015
Predicting the Future

Pilarski et al., 2012, BioRob
Autonomous Switching
(learning and unlearning automatic control actions)

Edwards et al., BioRob, 2016
EXAMPLE 2

motor synergies

(predictions as actions)

EXAMPLE 3

communication

(predictions as feedback)

Exoskeletons: UC Berkeley spin-off suitX exoskeleton technology; https://www.youtube.com/watch?v=l3roYI3CB2Y
#ConstructivistAGI
Whole point of this talk:

was to keep you awake
with cool videos long enough
to hear Rich’s talk
Whole point of this talk:

or think about Intelligence Amplification
as the grand challenge
for DL and RL
Whole point of this talk:

or maybe just that “prediction then control” research pattern we talked about earlier. either way, all good.
Start with prediction.