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)

UNIVERSITY OF ALBERTA
EDMONTON·ALBERTA·CANADA

amii

~~Applied RL on Robots~~

Boring title! (remember to send organizers a better one.)

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)

this is kind of stuffy and boring too...

UNIVERSITY OF ALBERTA
EDMONTON·ALBERTA·CANADA

amii

AI

IA

(Intelligence Amplification)

Ashby, W. R. (1956). *An introduction to cybernetics*. London: Chapman & Hall.

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

John Lennon


OBARA
KAWASAKI
OBARA


UNIMATE
Kawasaki
Kawasaki

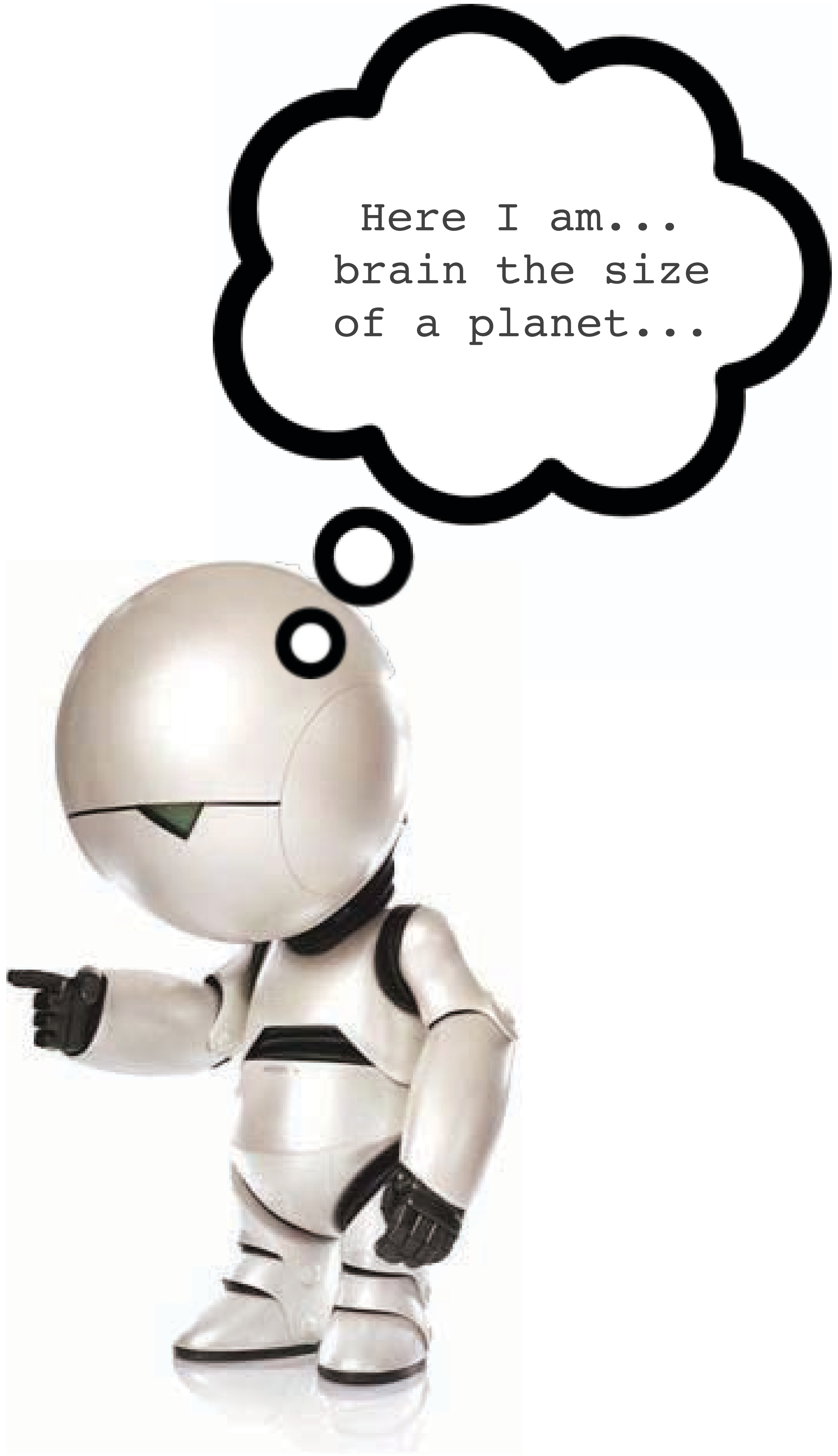

ショックアブソーバガイド
ナット締付装置
ナットランナ
エンジン・シャシ搭載リフタ

SPINNING WHEEL

{animal fur,
plant matter}
>>
{thread,
yarn}


muratec

… and, in short order …

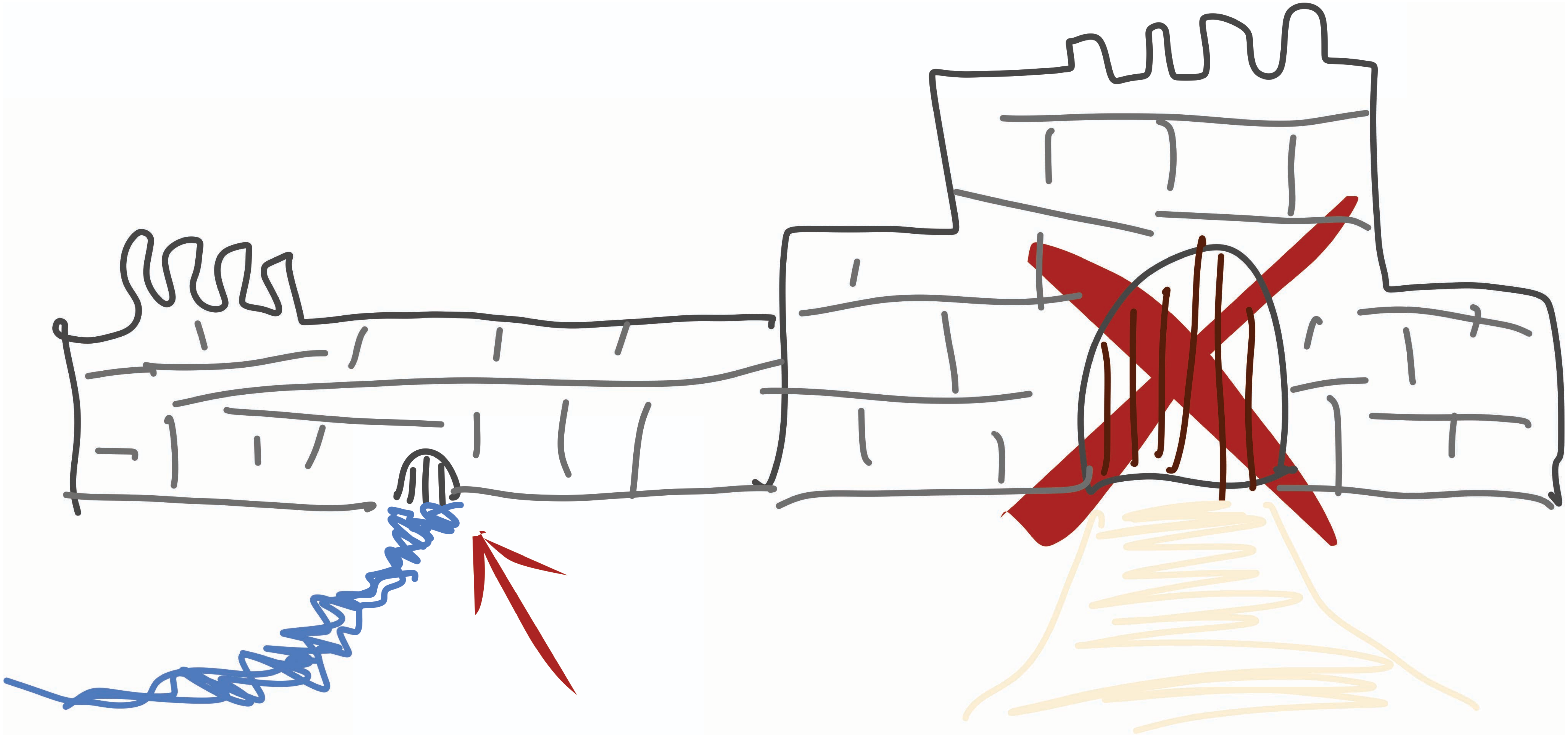


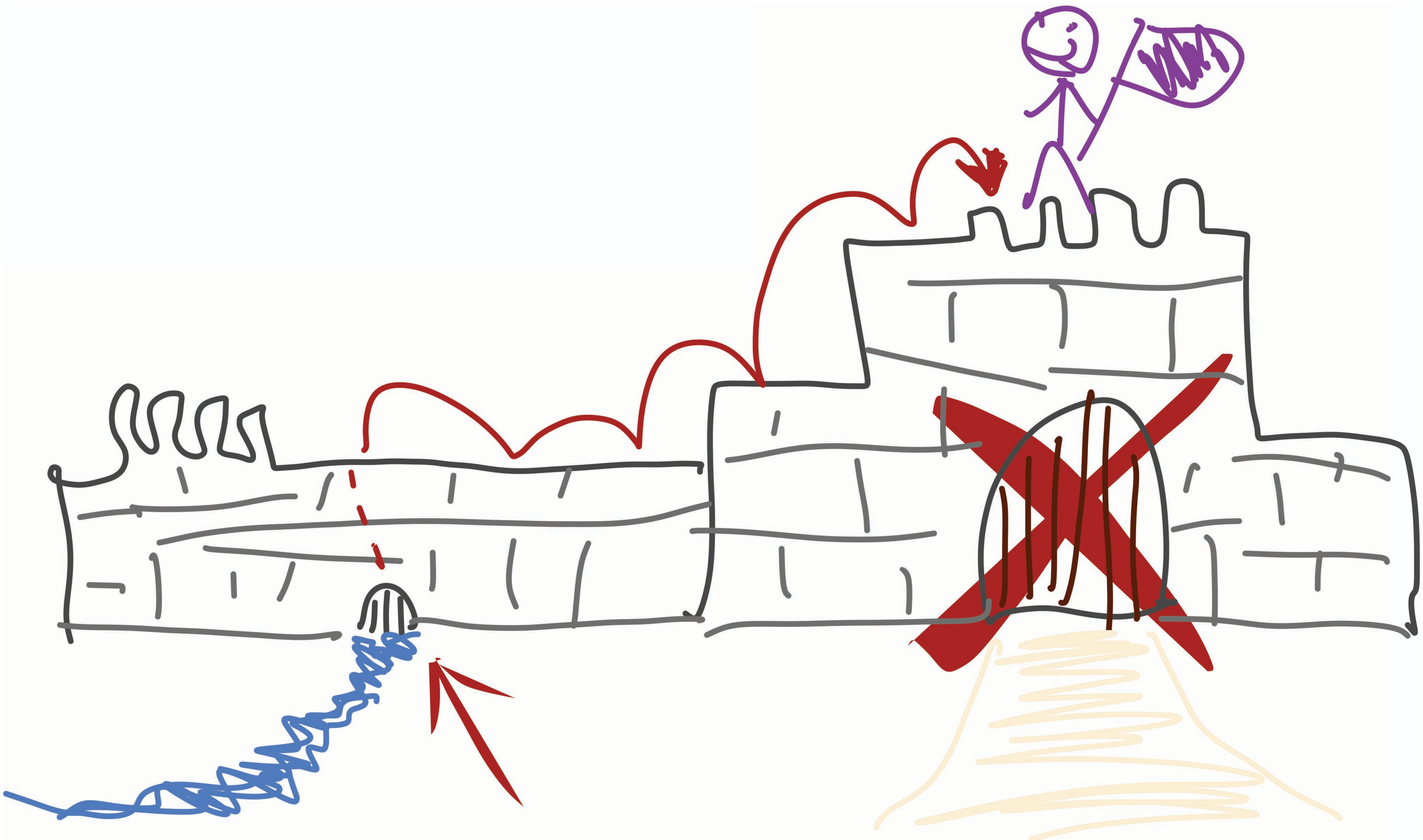
Marvin the paranoid android from THHGTTG.

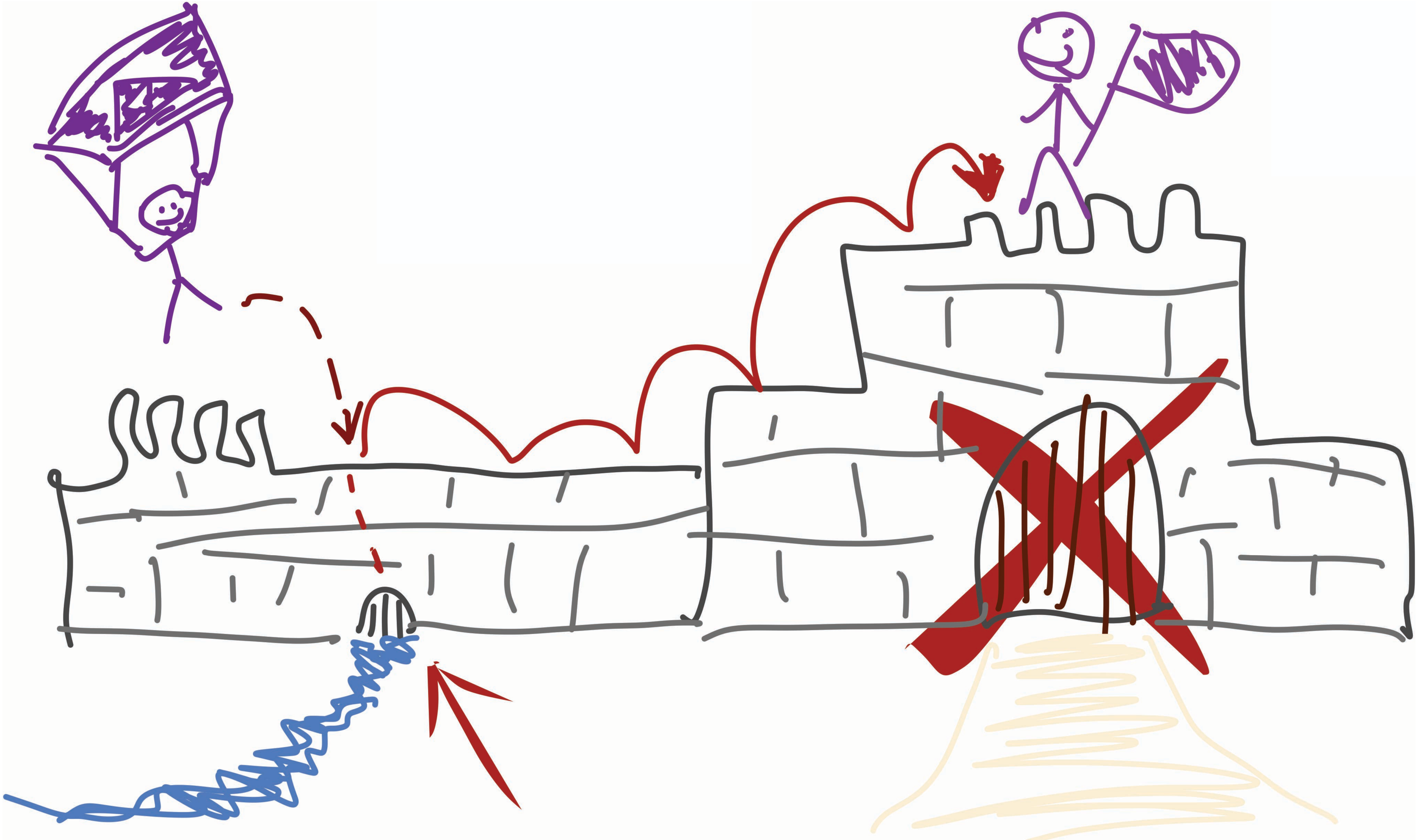
Whole point of this talk:

communicate one research design pattern
that can help you tackle
ambitious real-world problems

Never storm a castle through the front gates...

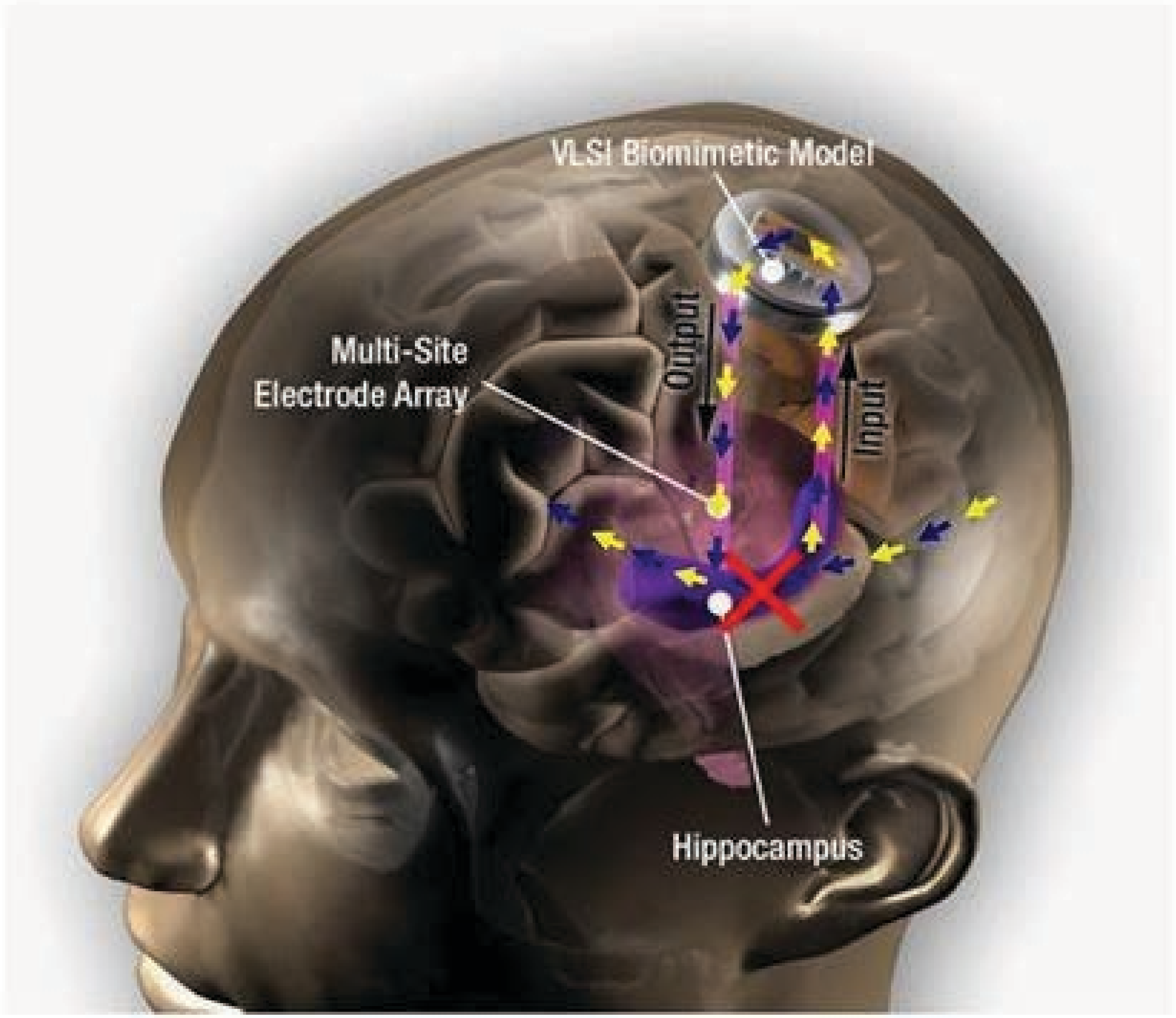


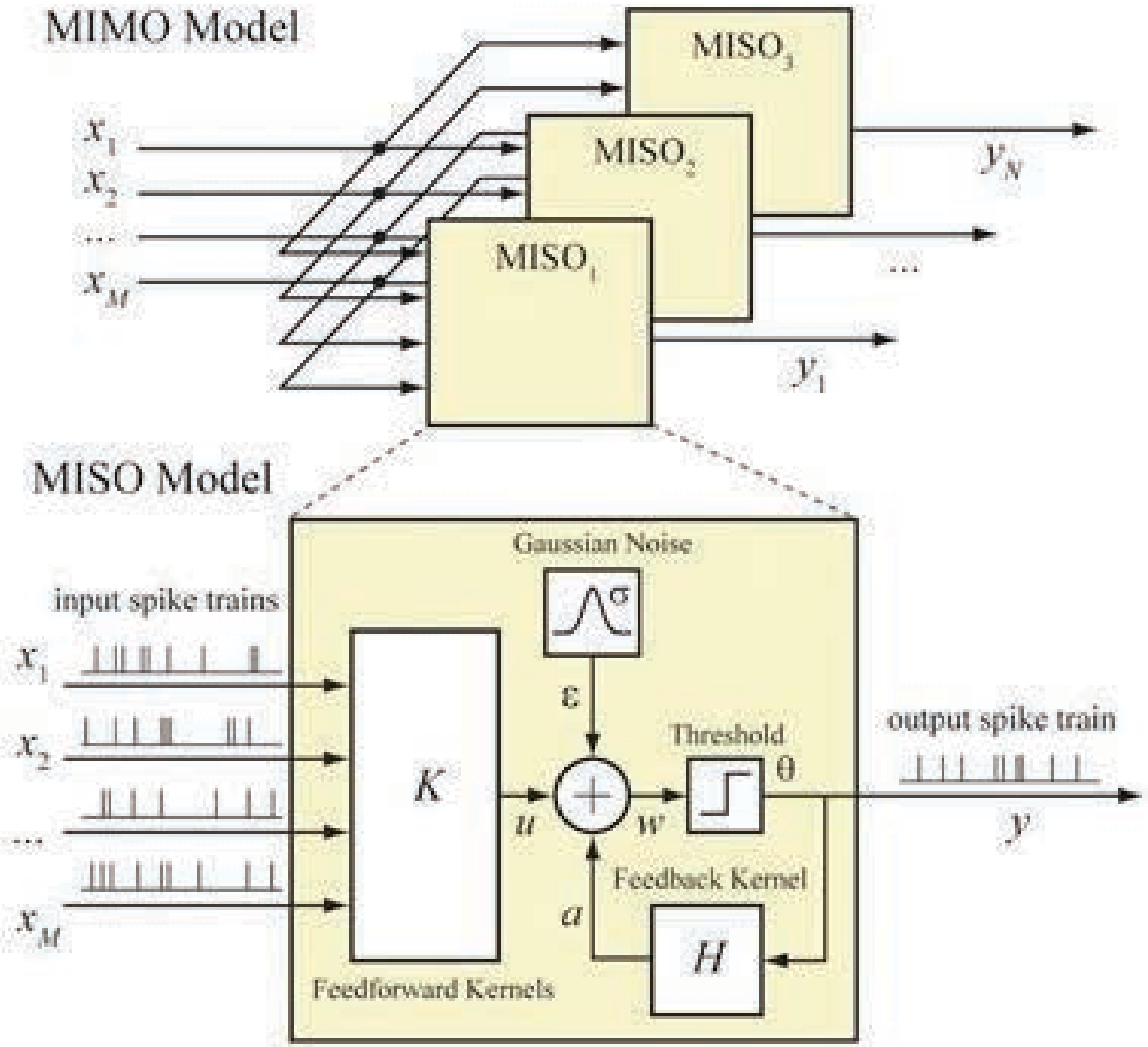




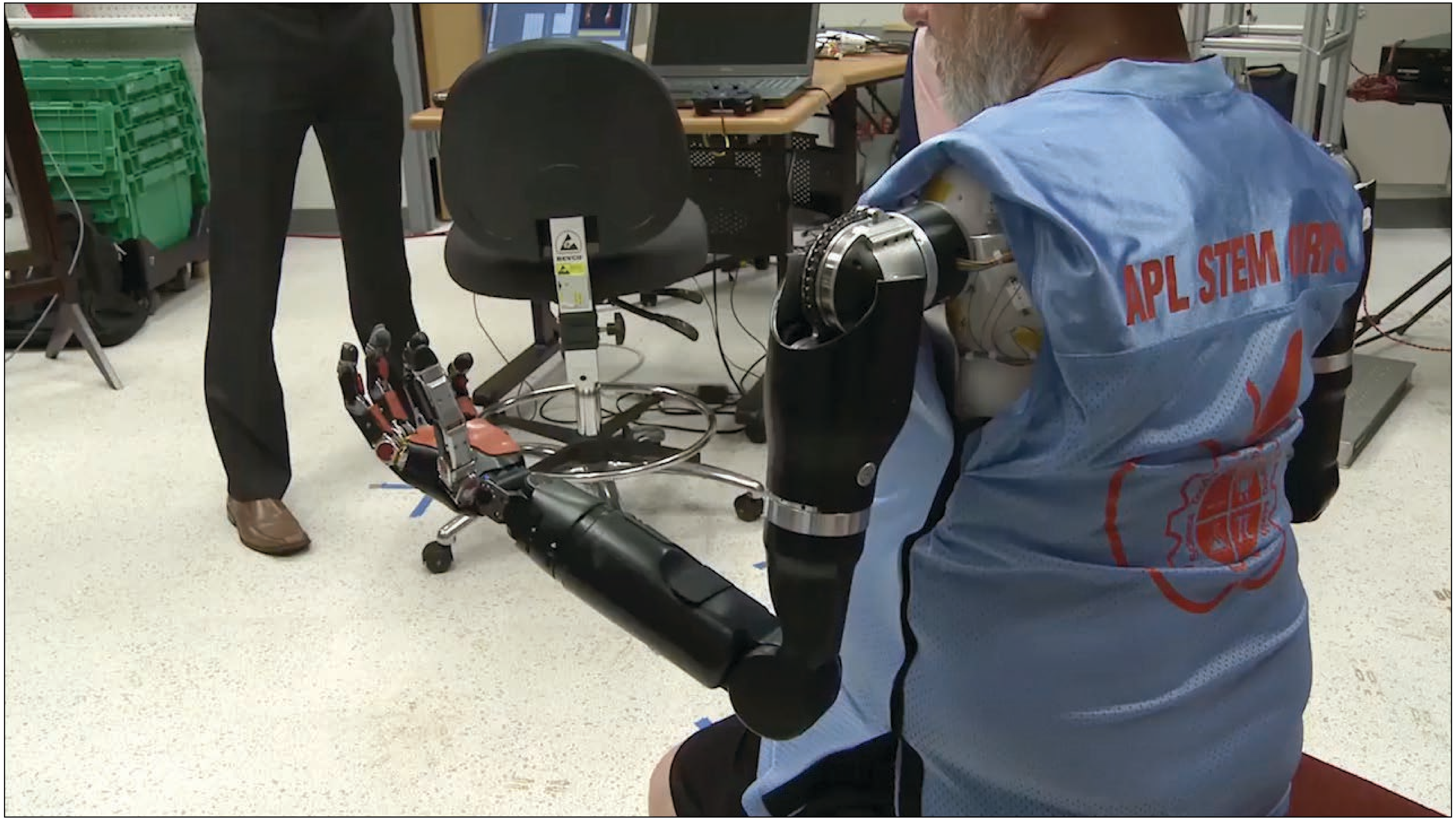
IA

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



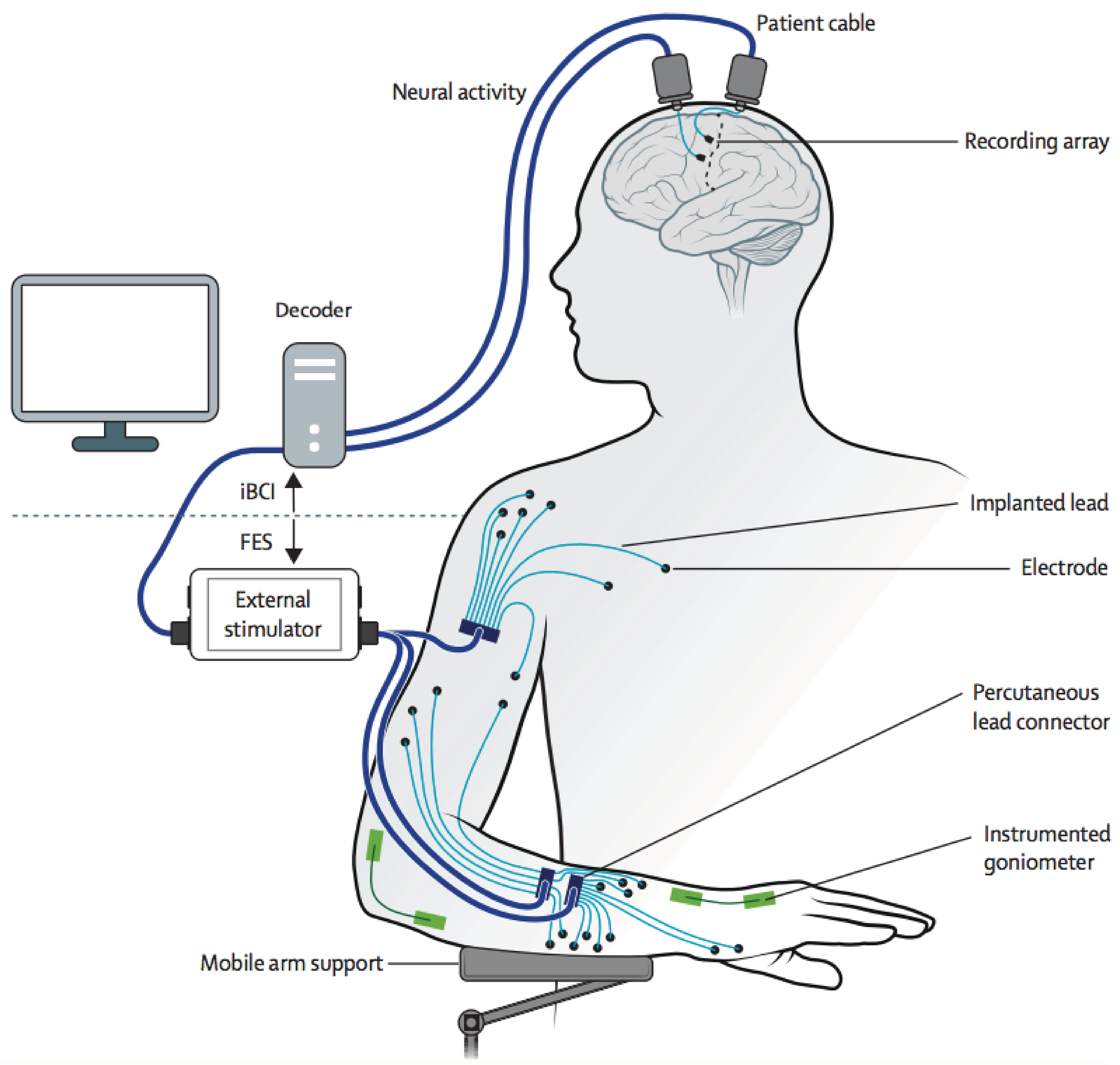


Direct brain-computer interfaces: *memory protheses* from the Center for Neural Engineering, Viterbi School of Engineering. https://cne.usc.edu/neural-prosthesis-for-hippocampal-memory-function/ and IEEE Trans Neural Syst Rehabil Eng. 2018, 26(2):272-280.



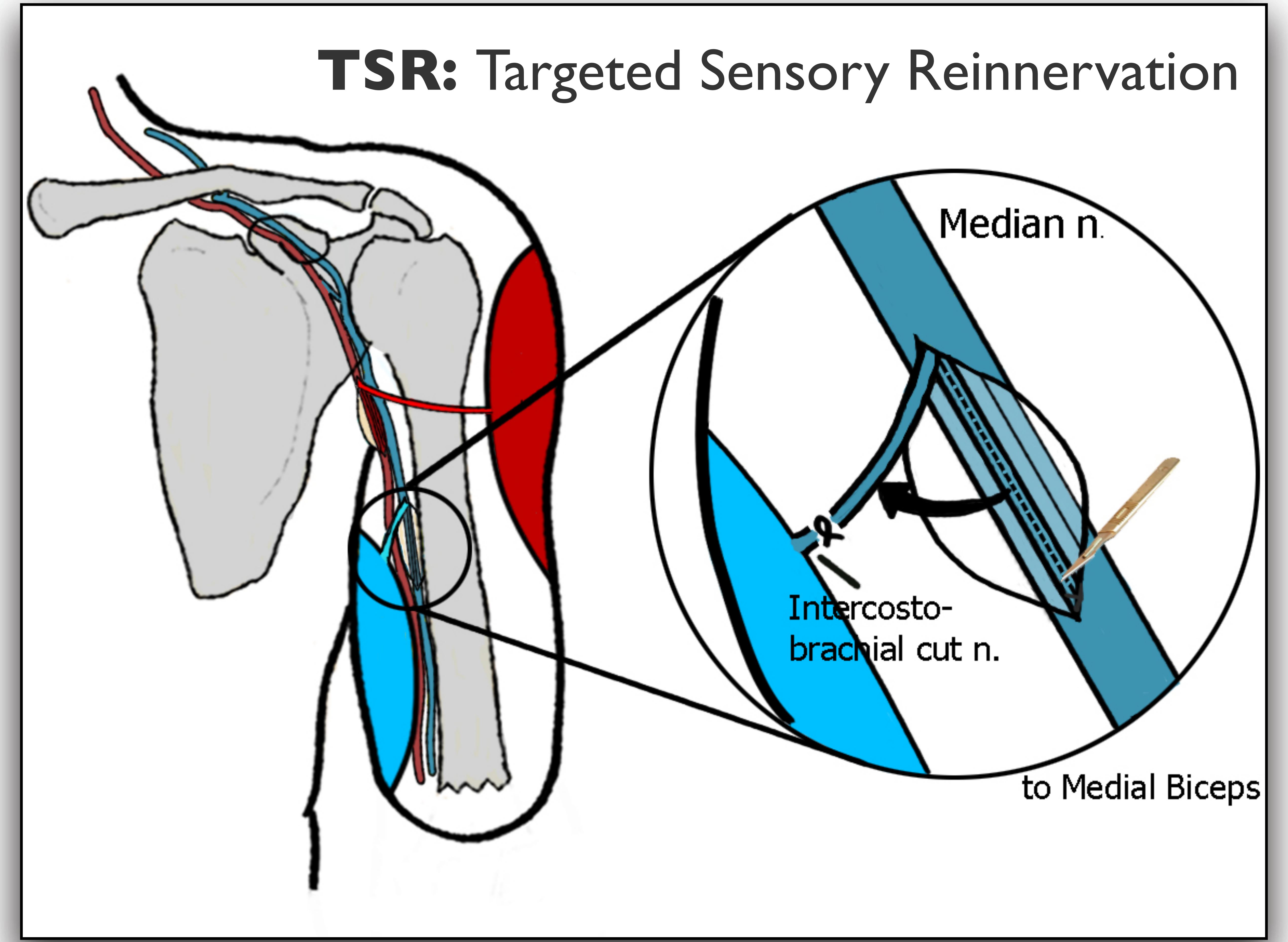
Brain-body-machine interfaces: “Amputee Makes History with APL’s Modular Prosthetic Limb” (JHU Applied Physics Laboratory); https://youtu.be/9NOncx2jU0Q

Brain-body-machine interfaces: “APL’s Modular Prosthetic Limb Reaches New Levels of Operability” (JHU Applied Physics Laboratory); https://youtu.be/-0srXvOQIu0



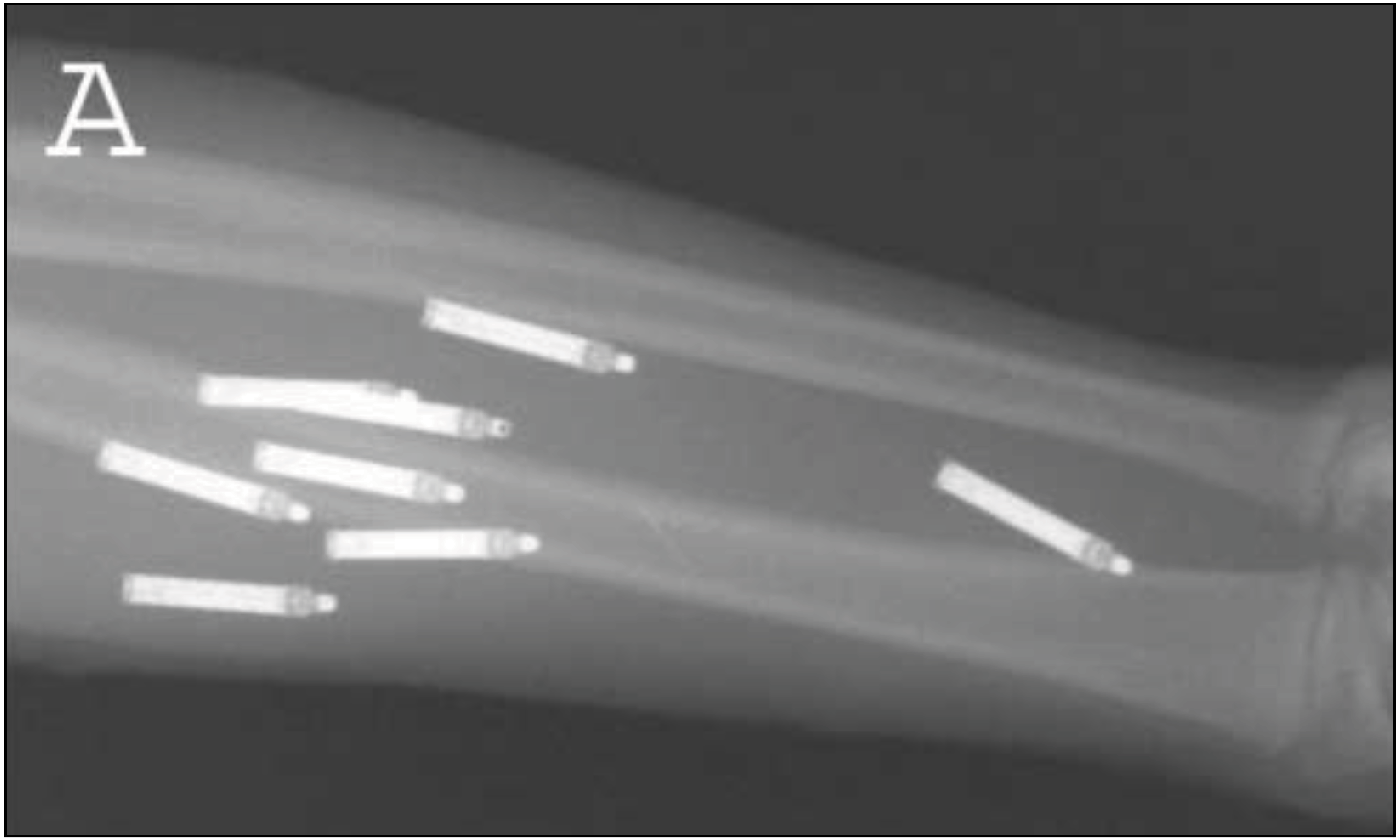
Brain-body-machine interfaces: "Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration" Ajiboye, A Bolu et al., *The Lancet*, Volume 389 , Issue 10081, 1821-1830, 2017.

Re-wiring the Nerves



Hebert JS, Olson JL, Morhart MJ, Dawson MR, Marasco PD, Kuiken TA, Chan KM, "Novel Targeted Sensory Reinnervation Technique To Restore Functional Hand Sensation After Transhumeral Amputation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* Vol 22, No 4, pages 765-773, 2014.

Hebert et al. 2014, IEEE-TNSRE



Brain-body-machine interfaces: Baker et al., “Continuous Detection and Decoding of Dexterous Finger Flexions With Implantable MyoElectric Sensors,” IEEE TNSRE 18(4):424-32, 2010.

Brain-body-machine interfaces: “Brain-Machine Interface @ EPFL- Wheelchair” (École polytechnique fédérale de Lausanne); https://youtu.be/0-1sdtnuqcE

Commercially Deployed
Pattern Recognition for Prostheses

Muse

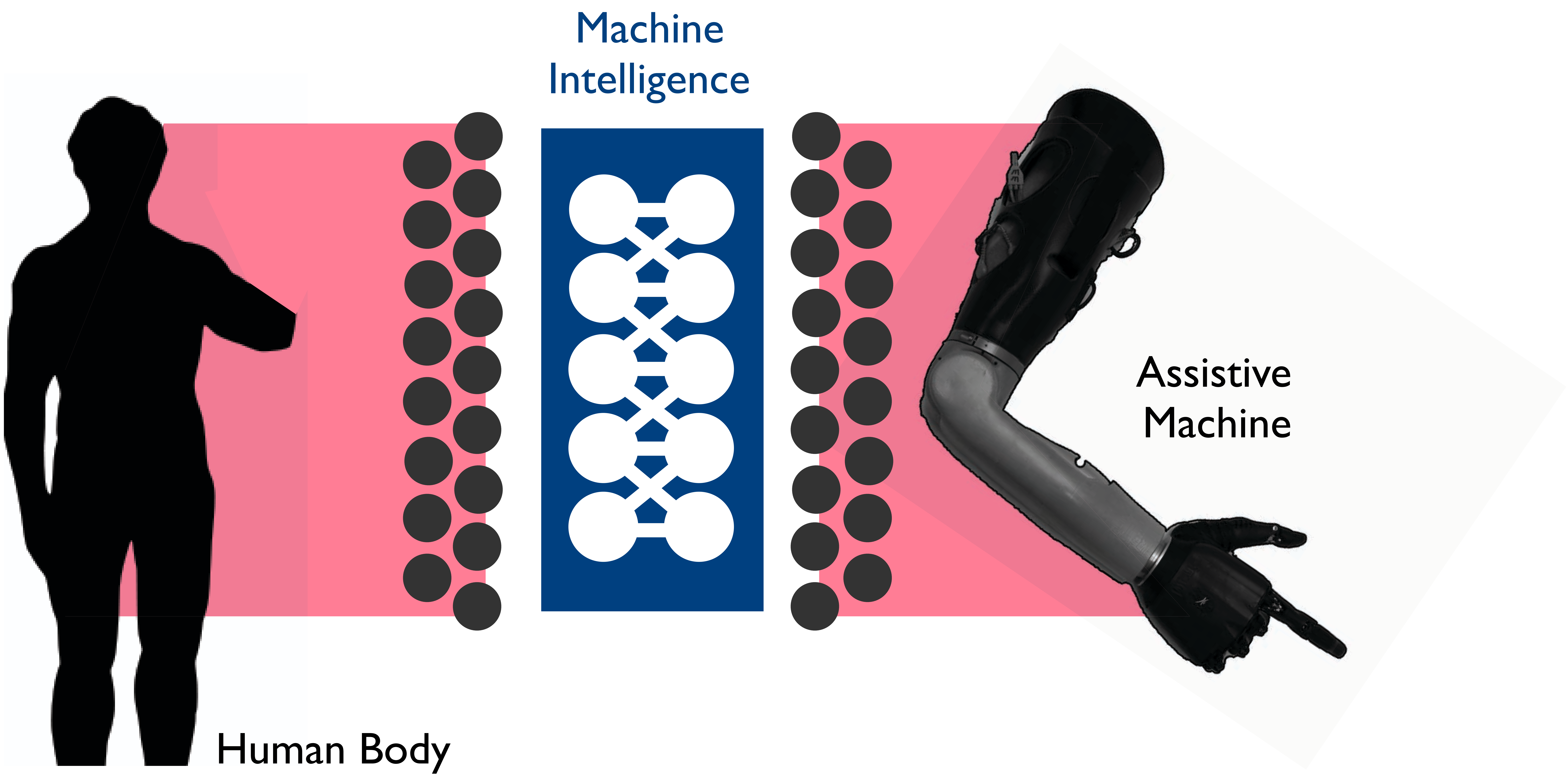
Myo (Thalmic Labs)

Consumer-Available BCI and BMI

BLINCdev.ca

File photo by *The Canadian Press*/Amber Bracken, 2019

These examples
all involve machine intelligence
or machine learning


Machine
Intelligence
Assistive
Machine
Human Body

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)

S. Micera, J. Carpaneto, and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Rev. Biomed. Eng.*, 2010.

PREDICTIONS

Momentary.
(e.g., classification decision)

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

S. Micera, J. Carpaneto, and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Rev. Biomed. Eng.*, 2010.

Sutton et al., "Horde: A Scalable Real-time Architecture for Learning Knowledge from Unsupervised Sensorimotor Interaction," *Proc. of 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2011.

PREDICTIONS

Momentary.
(e.g., classification decision)

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

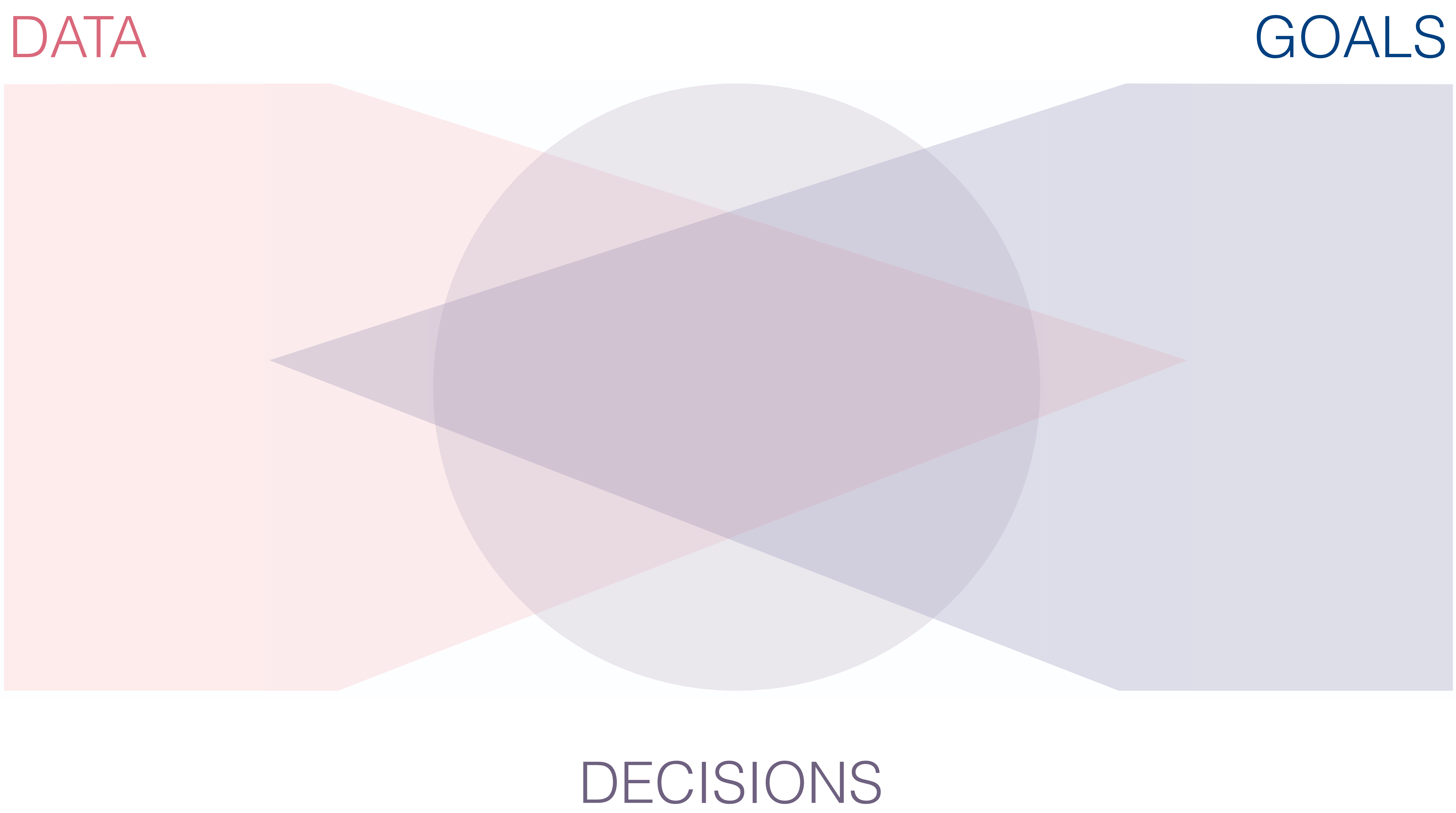
Can be acquired or updated in batches or in real time.

S. Micera, J. Carpaneto, and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Rev. Biomed. Eng.*, 2010.

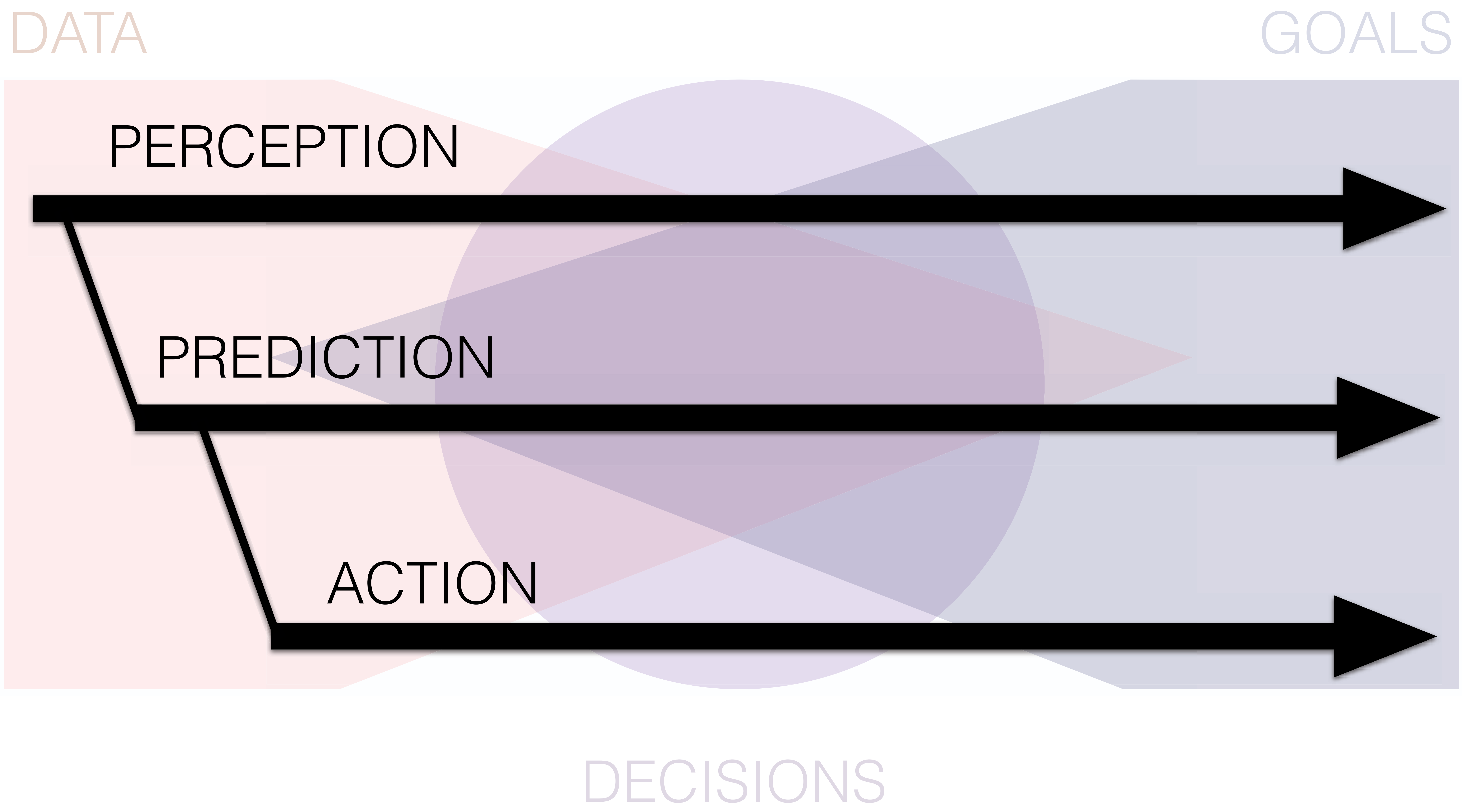
Sutton et al., "Horde: A Scalable Real-time Architecture for Learning Knowledge from Unsupervised Sensorimotor Interaction," *Proc. of 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2011.

PREDICTION PRECEDES CONTROL

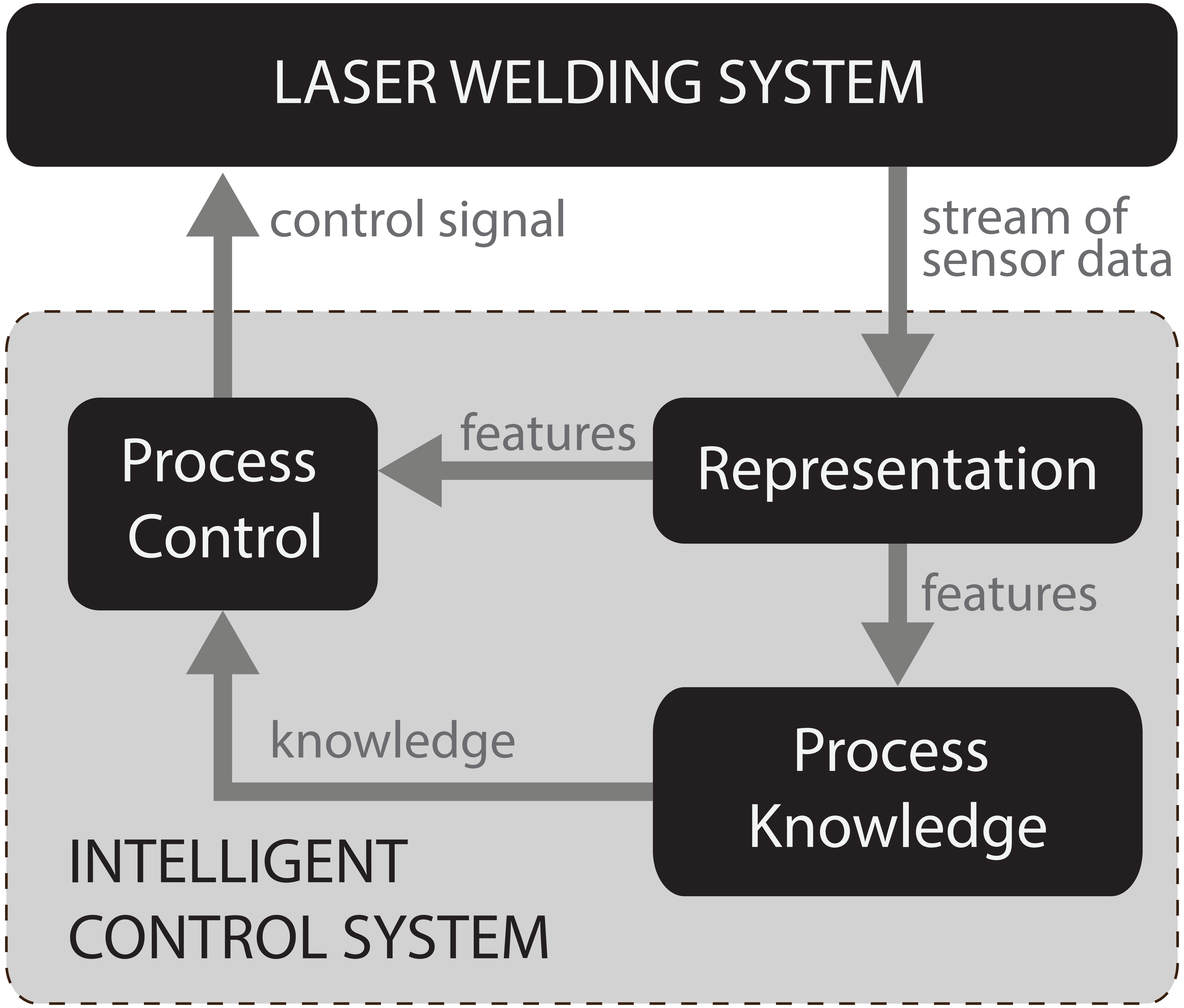
Wolpert et al., *Trends Cog Sci* 5(11), 2001: “Perspectives and problems in motor learning”
Flanagan et al., *Current Biology* 13(2), 2003: “Prediction precedes control in motor learning”
Desmurget et al., *Science* 324(5928), 2009: “Movement intention after parietal cortex stimulation in humans”



Hallmarks of Intelligence:
Artificial, Machine (and Human)



Hallmarks of Intelligence:
Artificial, Machine (and Human)



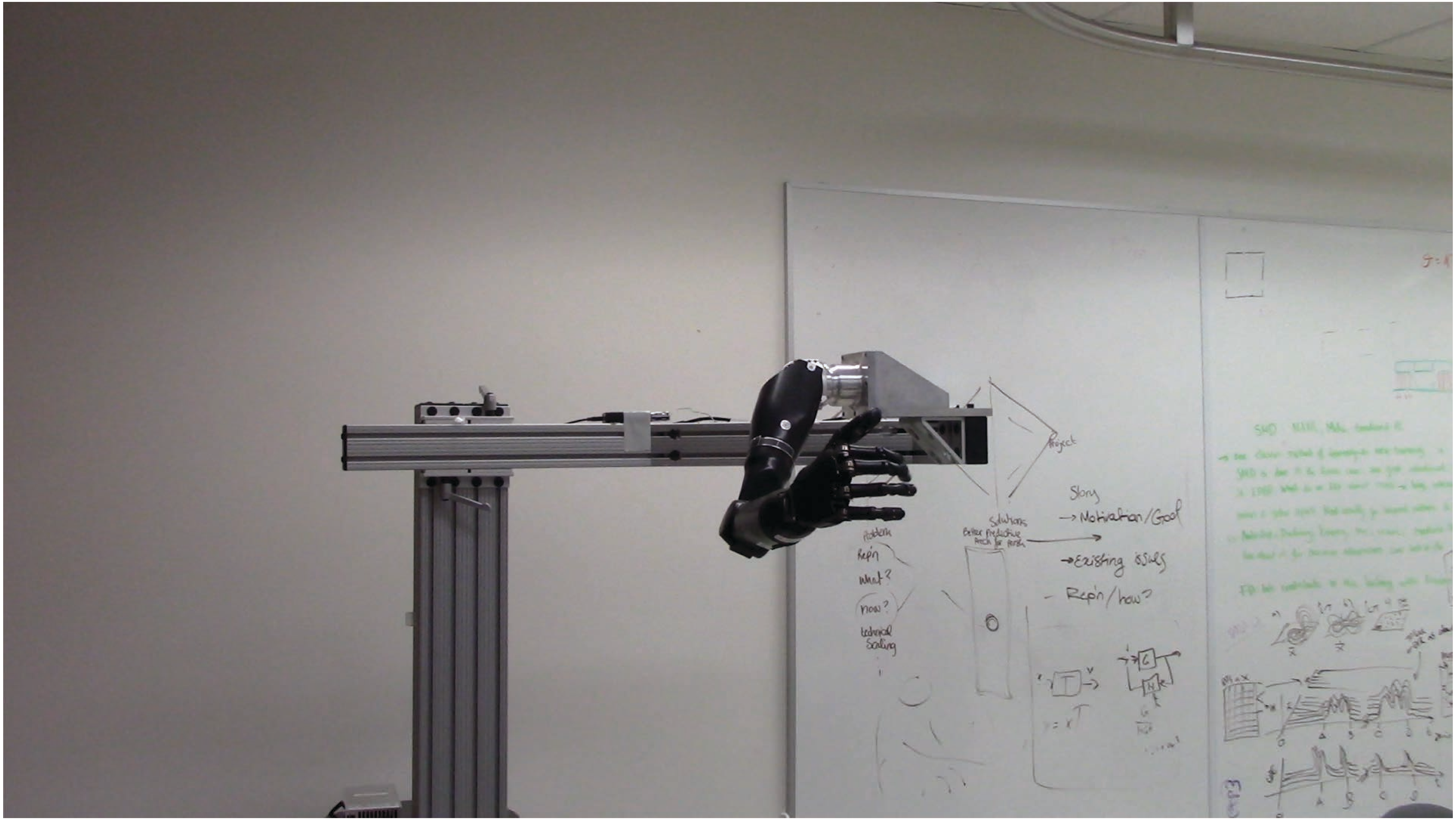
J. Gunther, P. M. Pilarski, G. Helfrich, H. Shen, K. Diepold, “Intelligent Laser Welding through Representation, Prediction, and Control Learning: An Architecture with Deep Neural Networks and Reinforcement Learning,” *Mechatronics*, vol. 34, pp. 1–11, March 2016.

CASE STUDY

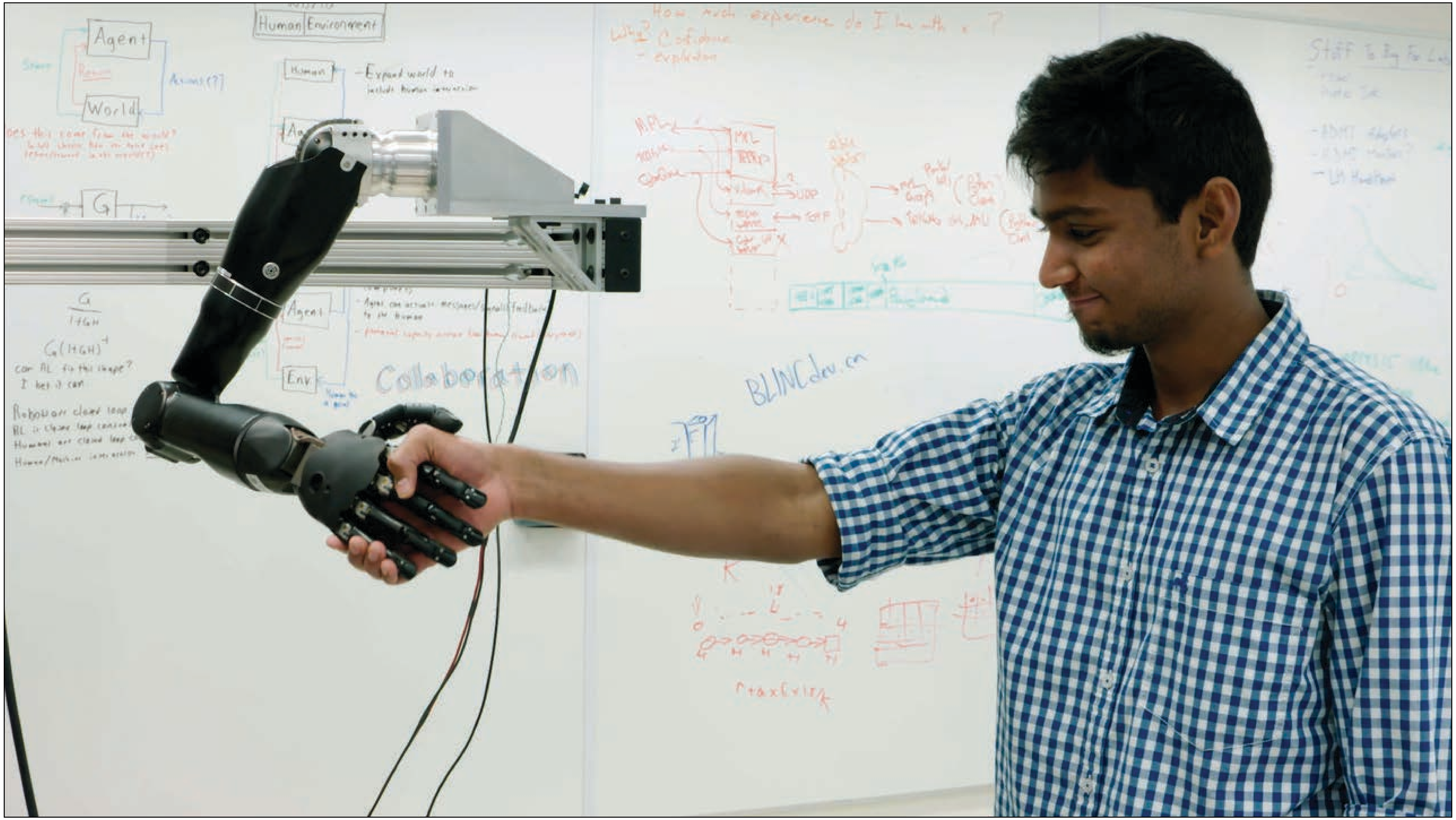
human embodiment of a robot body part

is really tricky...

P. M. Pilarski, R. S. Sutton, K. W. Mathewson, C. Sherstan, A. S. R. Parker, A. L. Edwards, "Communicative Capital for Prosthetic Agents," arXiv:1711.03676 [cs.AI]: 33 pages, 2017.



University of Alberta: http://blinclab.ca, https://www.smartnetworkcentre.ca/

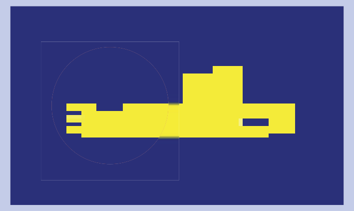


Pilarski Lab
August 2016

University of Alberta: http://blinclab.ca

Planning and Meta-learning

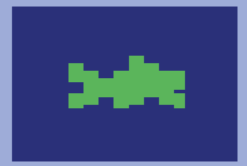
Control

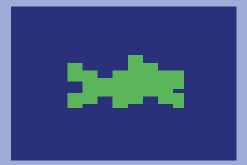


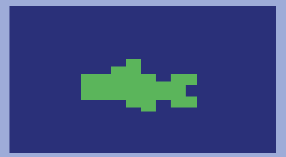
2009

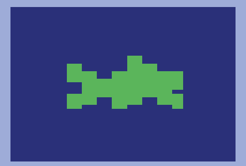
Prediction

Representation





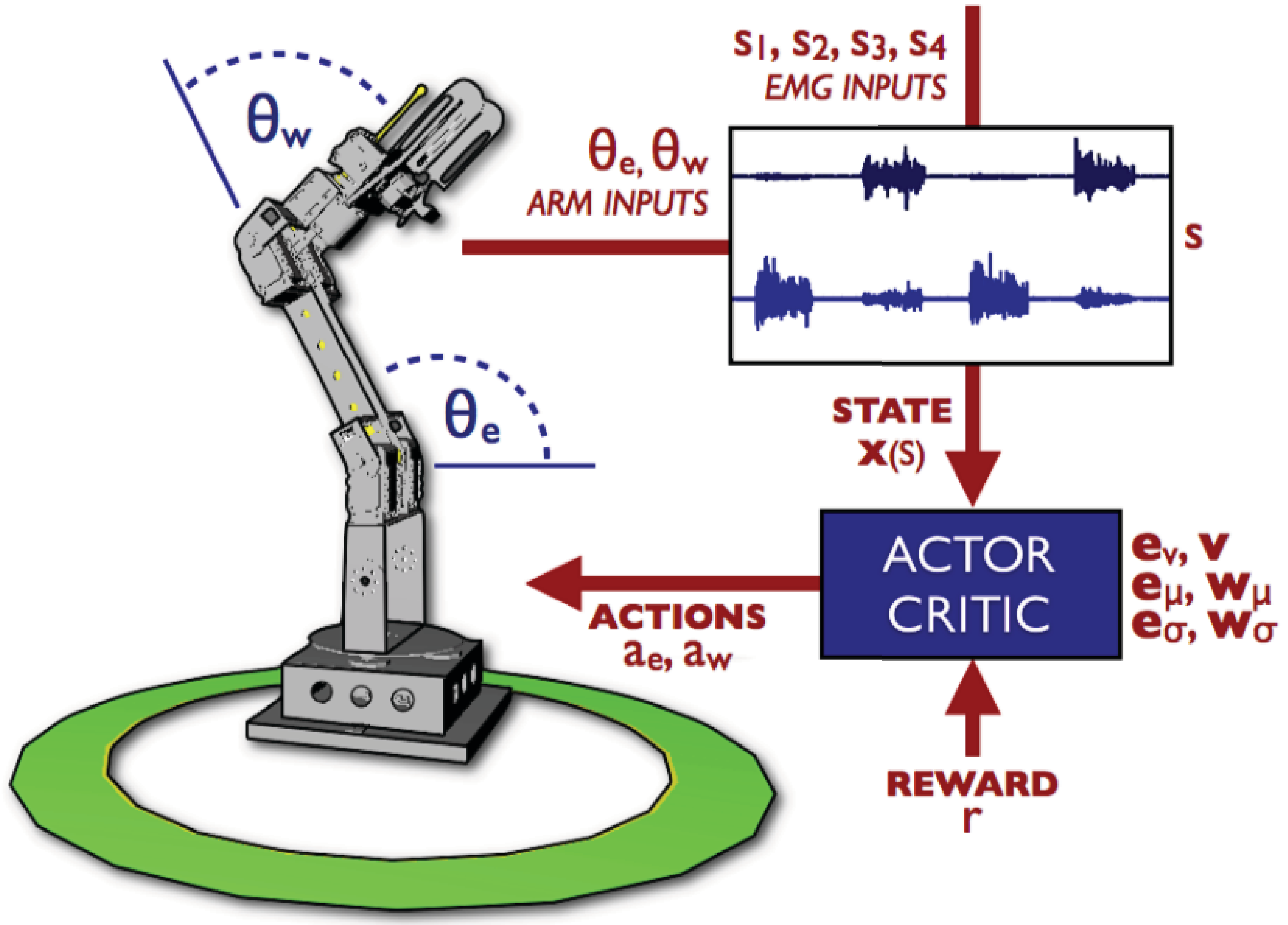




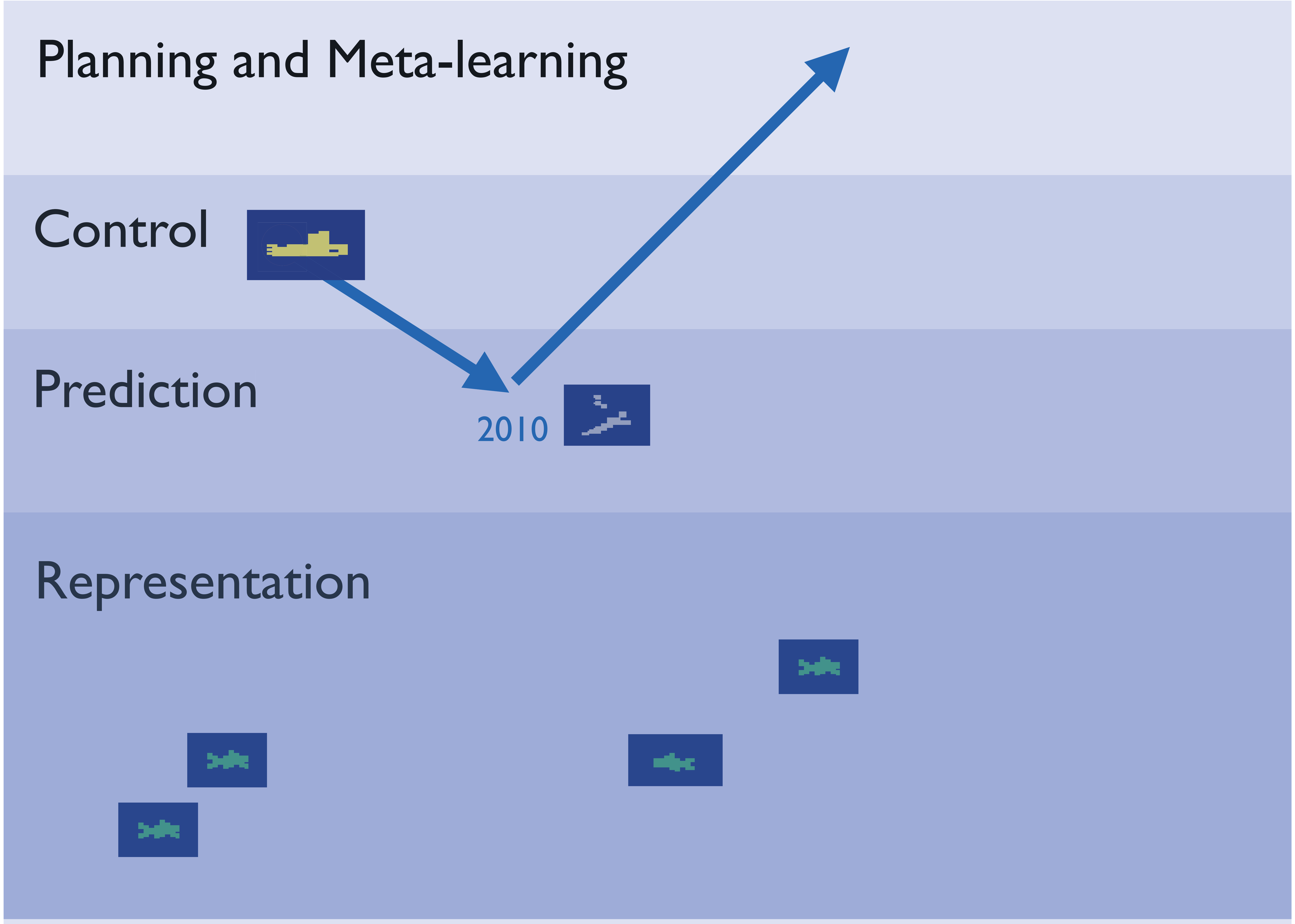
Online Human Training of a Myoelectric Prosthesis Controller via Actor-Critic Reinforcement Learning

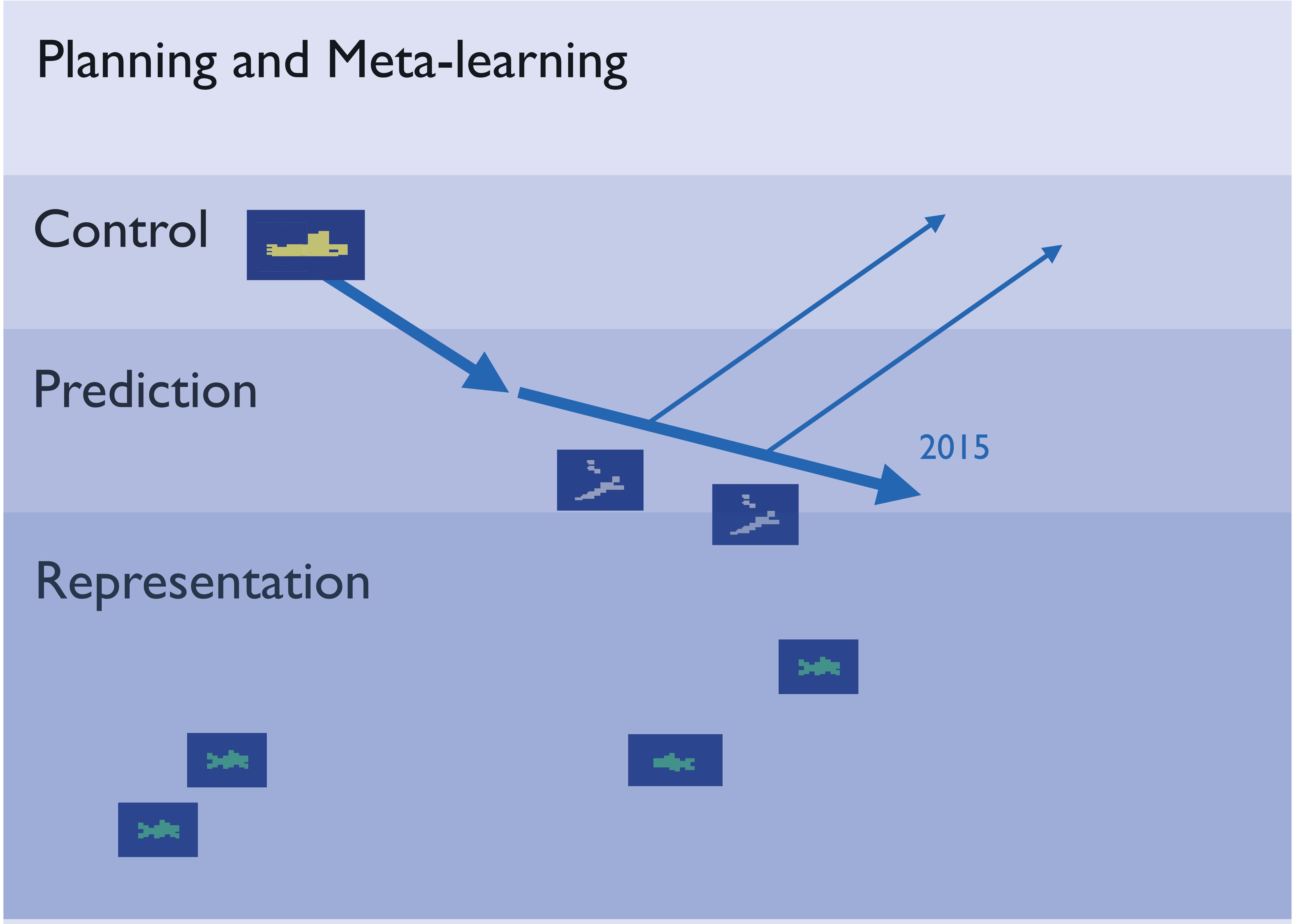
Patrick M. Pilarski, Michael R. Dawson, Thomas Degris, Farbod Fahimi, Jason P. Carey, and Richard S. Sutton

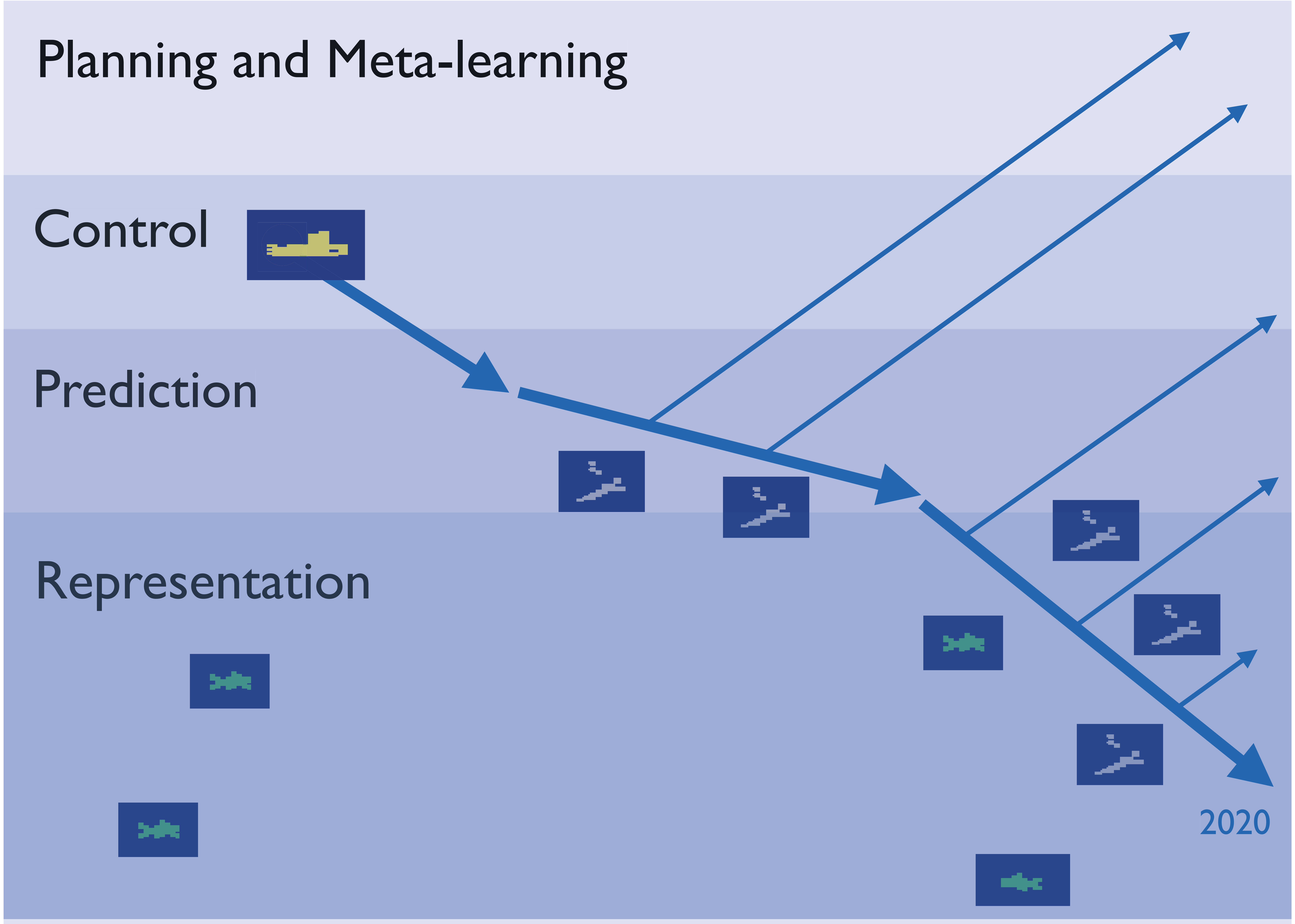
Abstract— **As a contribution toward the goal of adaptable, intelligent artificial limbs, this work introduces a continuous actor-critic reinforcement learning method for optimizing the control of multi-function myoelectric devices. Using a simulated upper-arm robotic prosthesis, we demonstrate how it 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


Planning and Meta-learning
Control
Prediction
2010
Representation


Planning and Meta-learning
Control
Prediction
Representation
2015

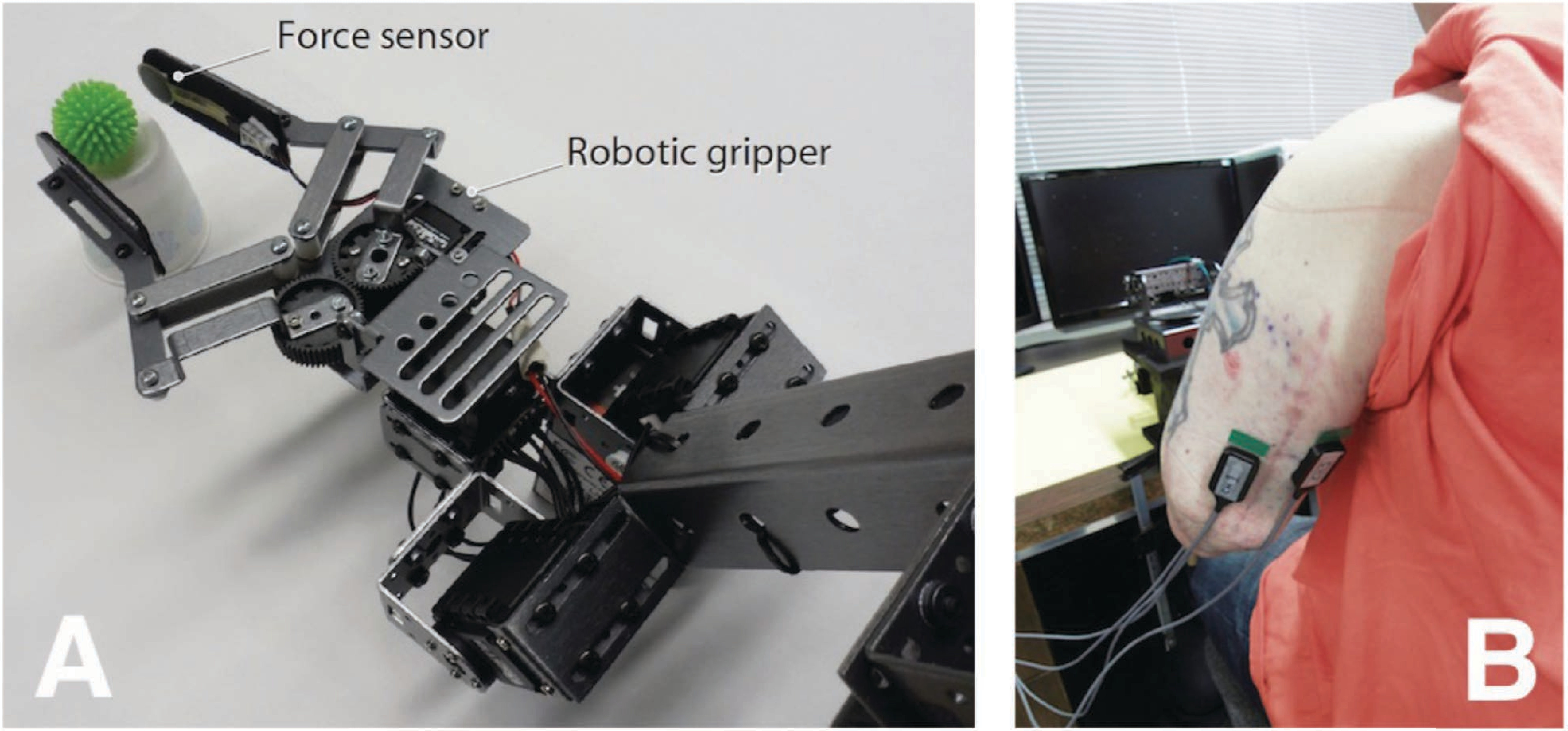

Planning and Meta-learning
Control
Prediction
Representation
2020

EXAMPLE 0

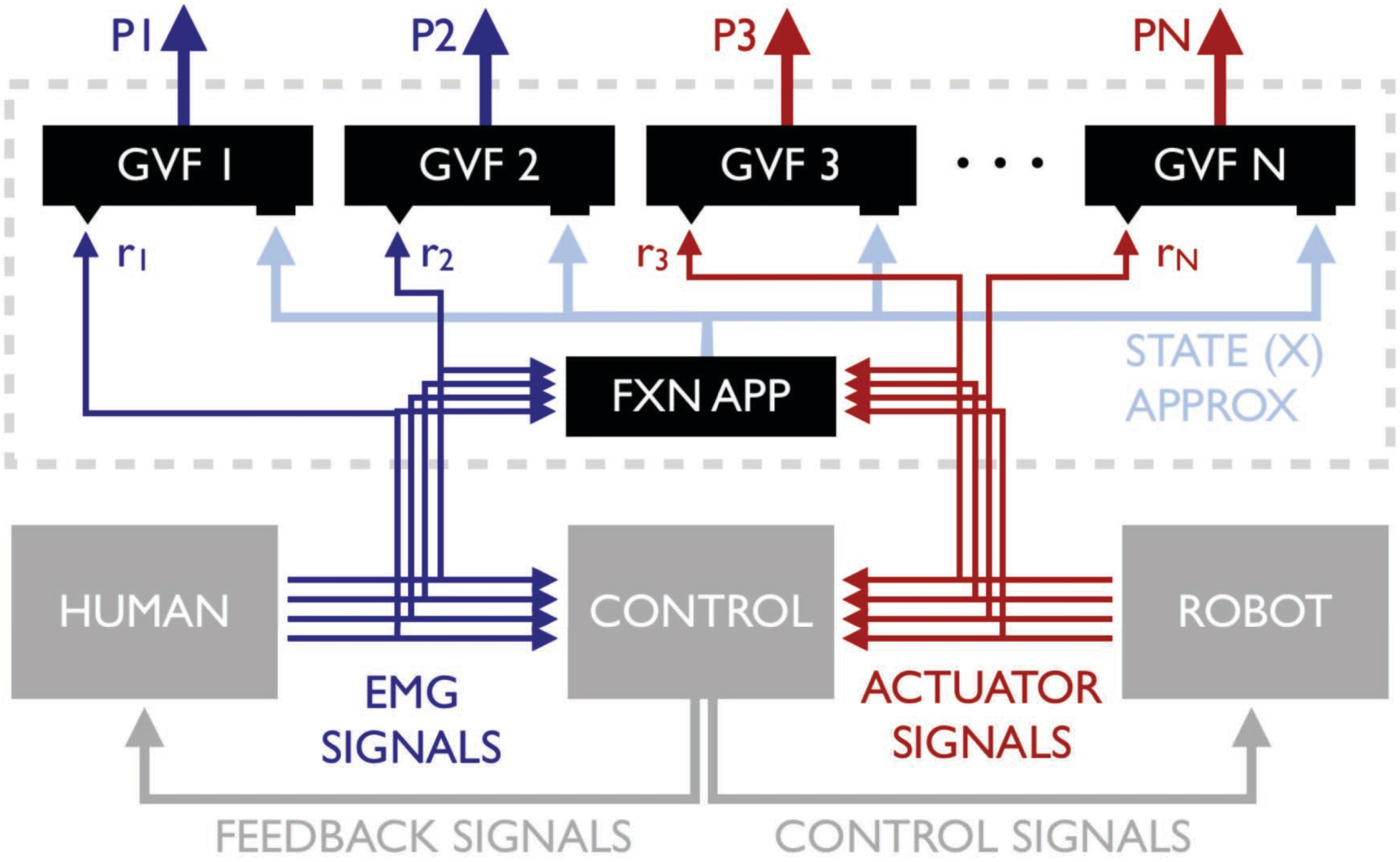
making predictions

(Nexting in real time)

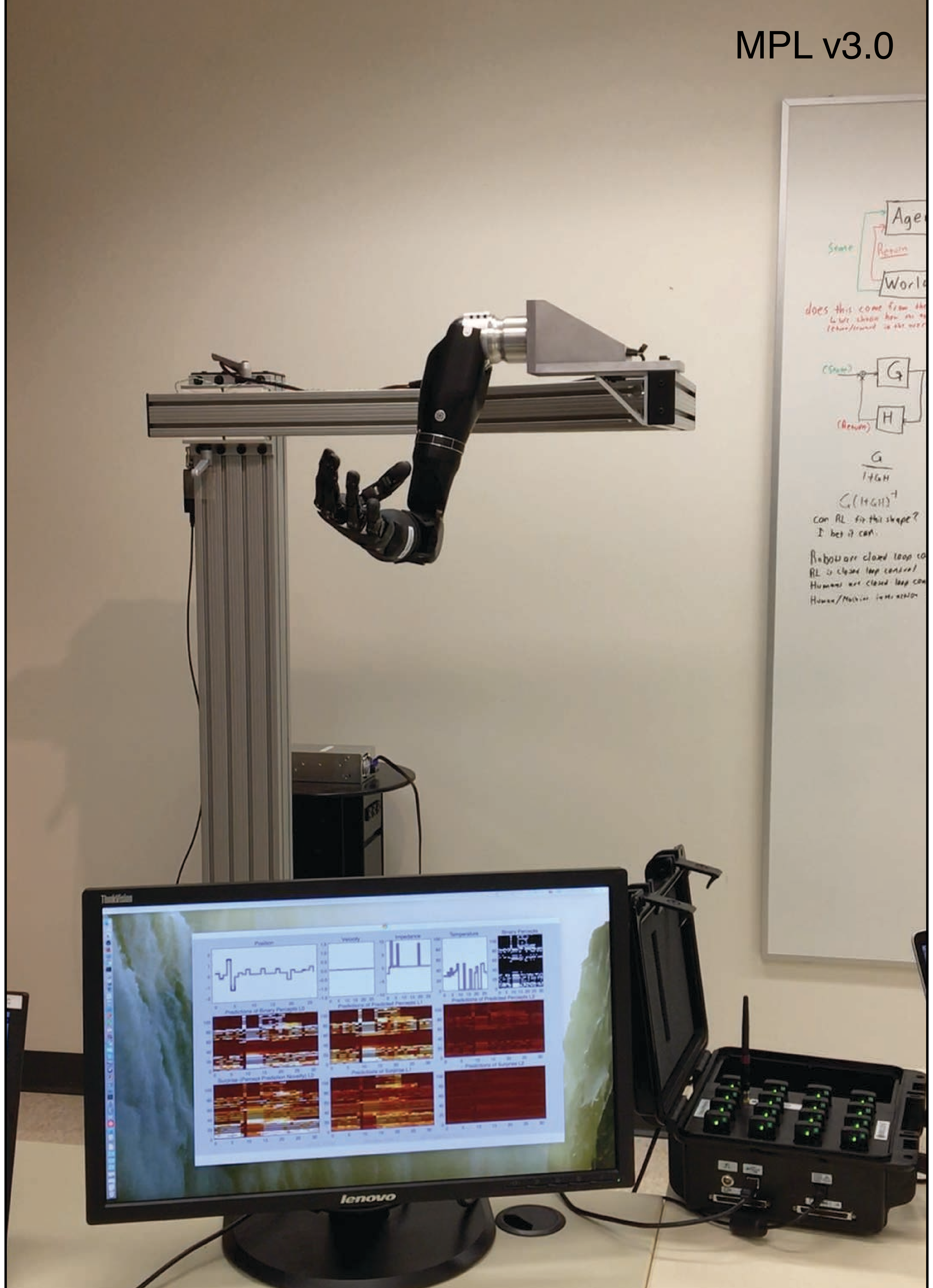
J. Modayil, A. White, R. S. Sutton, “Multi-timescale Nexting in a Reinforcement Learning Robot, *Adaptive Behavior* 22(2), April 2014. pages 146-160.



P.M. Pilarski, M.R. Dawson, T. Degris, J.P. Carey, K.M. Chan, J.S. Hebert, and R.S. Sutton, "Adaptive Artificial Limbs: A Real-time Approach to Prediction and Anticipation," *IEEE Robotics & Automation Magazine*, Vol. 20(1): 53–64, March 2013.

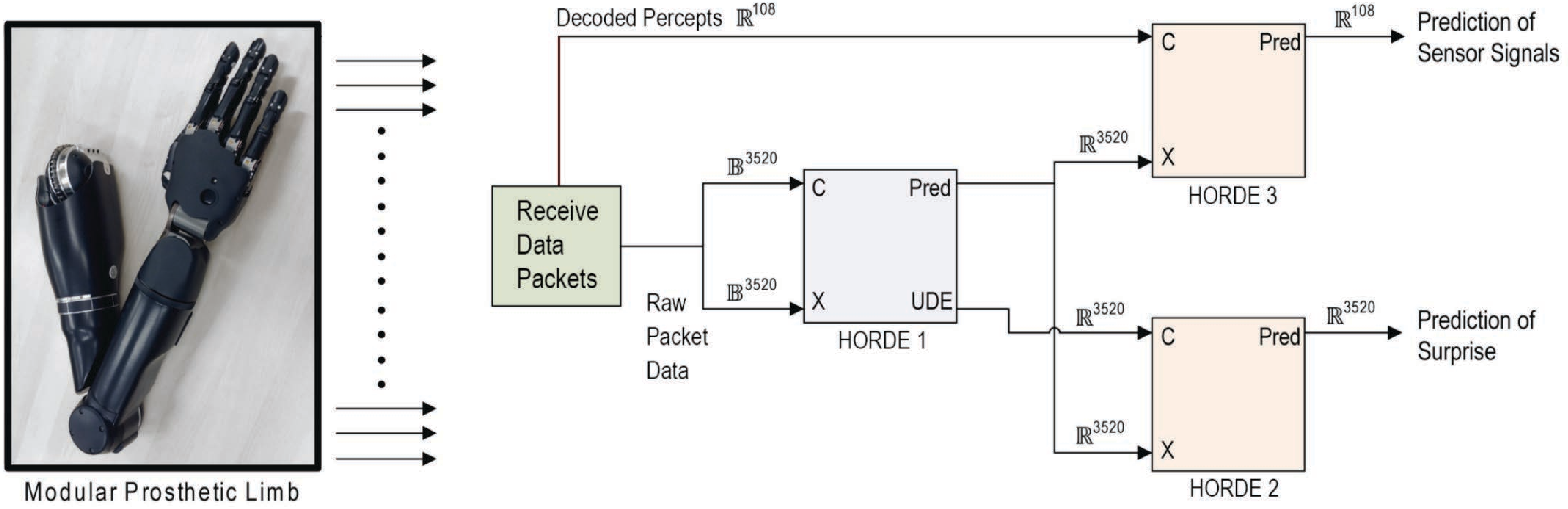


P.M. Pilarski, M.R. Dawson, T. Degris, J.P. Carey, K.M. Chan, J.S. Hebert, and R.S. Sutton, "Adaptive Artificial Limbs: A Real-time Approach to Prediction and Anticipation," *IEEE Robotics & Automation Magazine*, Vol. 20(1): 53–64, March 2013.

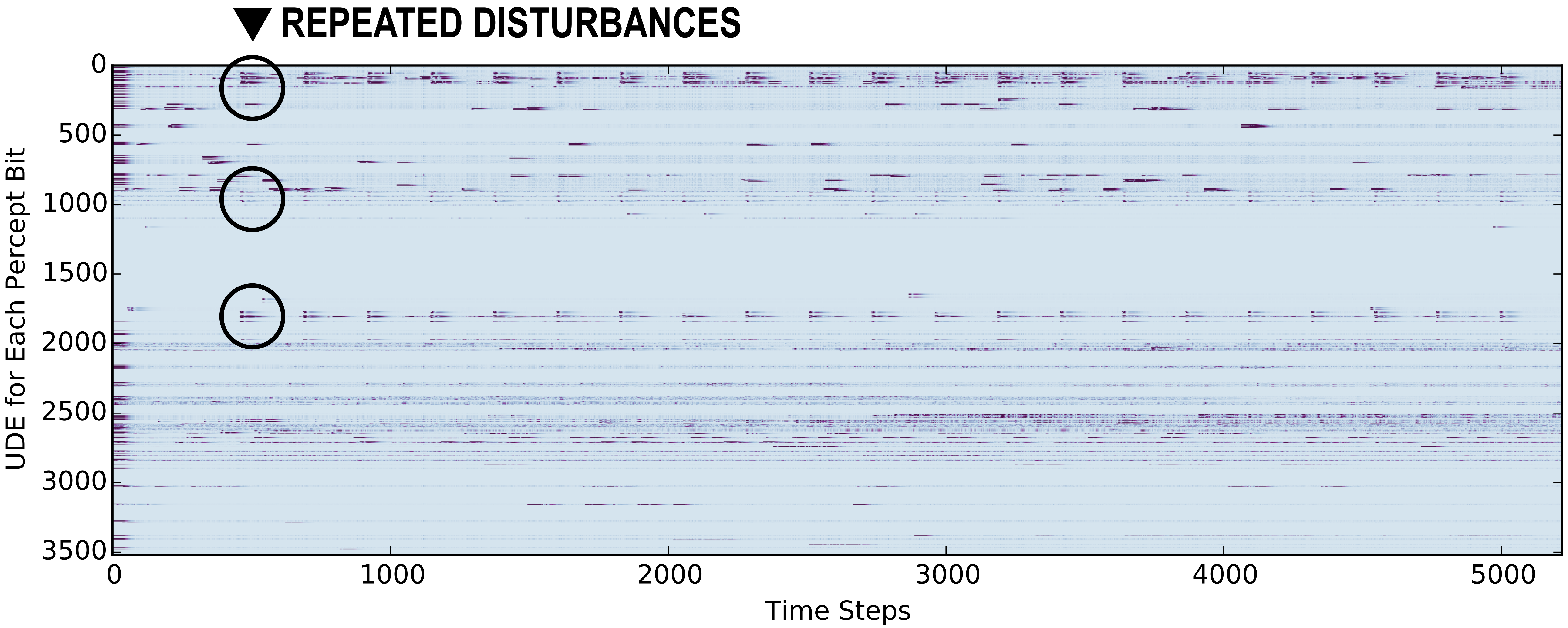


P. M. Pilarski, C. Sherstan, “Steps Toward Knowledgeable Neuroprostheses,” *Proceedings of the 6th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob2016)*, June 26-29, 2016, Singapore, pp. 220.

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



J. Gunther, A. Kearney, M. R. Dawson, C. Sherstan, P. M. Pilarski, "Predictions, Surprise, and Predictions of Surprise in General Value Function Architectures," *Proc. AAAI 2018 Fall Symposium on Reasoning and Learning in Real-World Systems for Long-Term Autonomy,* Arlington, USA, October 18-20, 2018, pp. 22–29.



J. Gunther, A. Kearney, N. M. Ady, M. R. Dawson, P. M. Pilarski, "Meta-learning for Predictive Knowledge Architectures: A Case Study Using TIDBD on a Sensor-rich Robotic Arm," *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, Montreal, Canada, May 13–17, 2019, pp. 1967–1969.

J. Gunther, A. Kearney, M. R. Dawson, C. Sherstan, P. M. Pilarski, "Predictions, Surprise, and Predictions of Surprise in General Value Function Architectures," *Proc. AAAI 2018 Fall Symposium on Reasoning and Learning in Real-World Systems for Long-Term Autonomy,* Arlington, USA, October 18-20, 2018, pp. 22–29.

Whole point of this talk:

say the phrase "Pavlovian control"
enough times that you
remember it next week

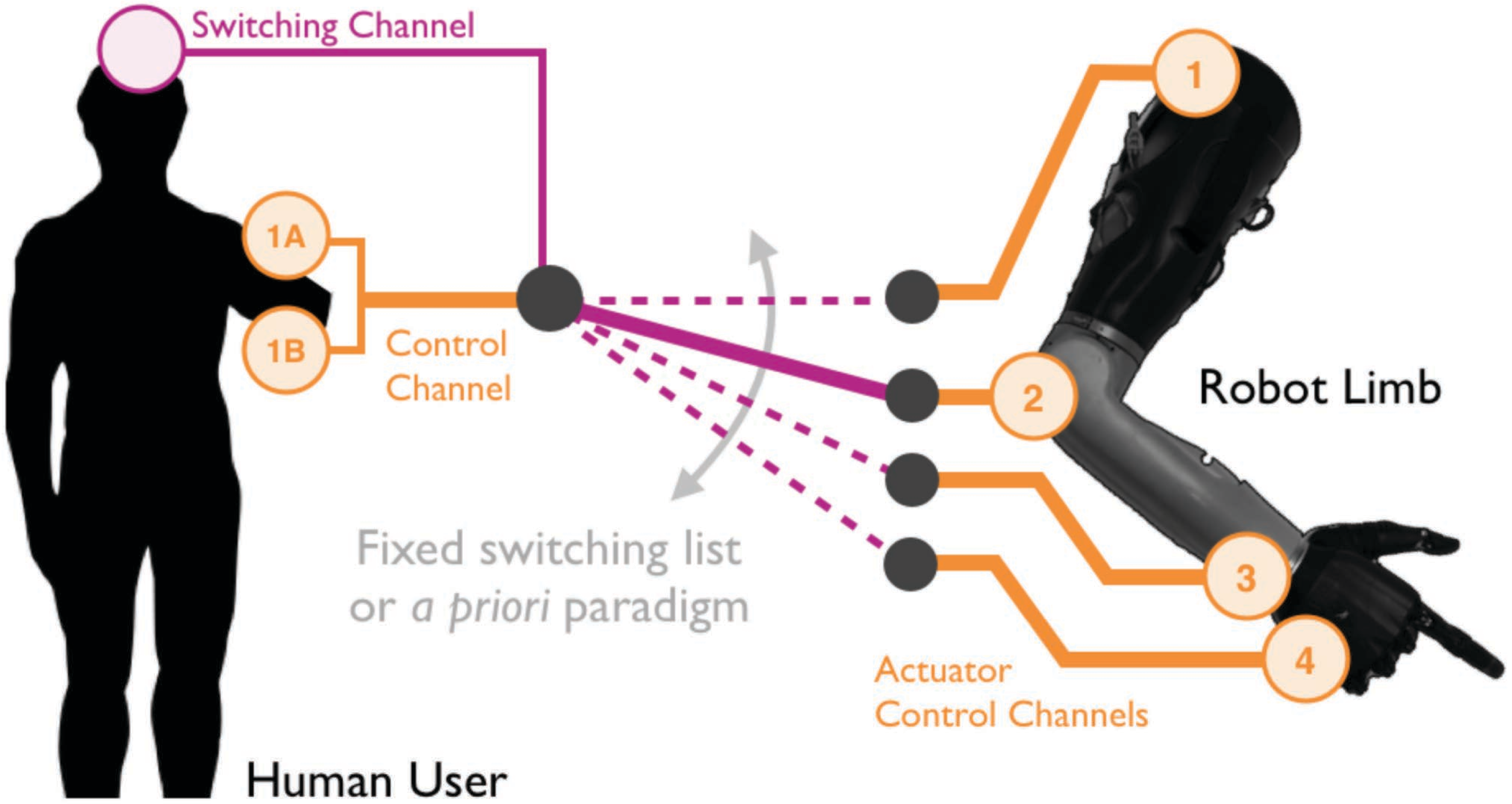
Pavlovian control involves a **fixed mapping between learned predictions** and control actions

J. Modayil and R. S. Sutton, "Prediction Driven Behavior: Learning Predictions that Drive Fixed Responses," *AAAI Workshop on AI and Robotics,* 2014.

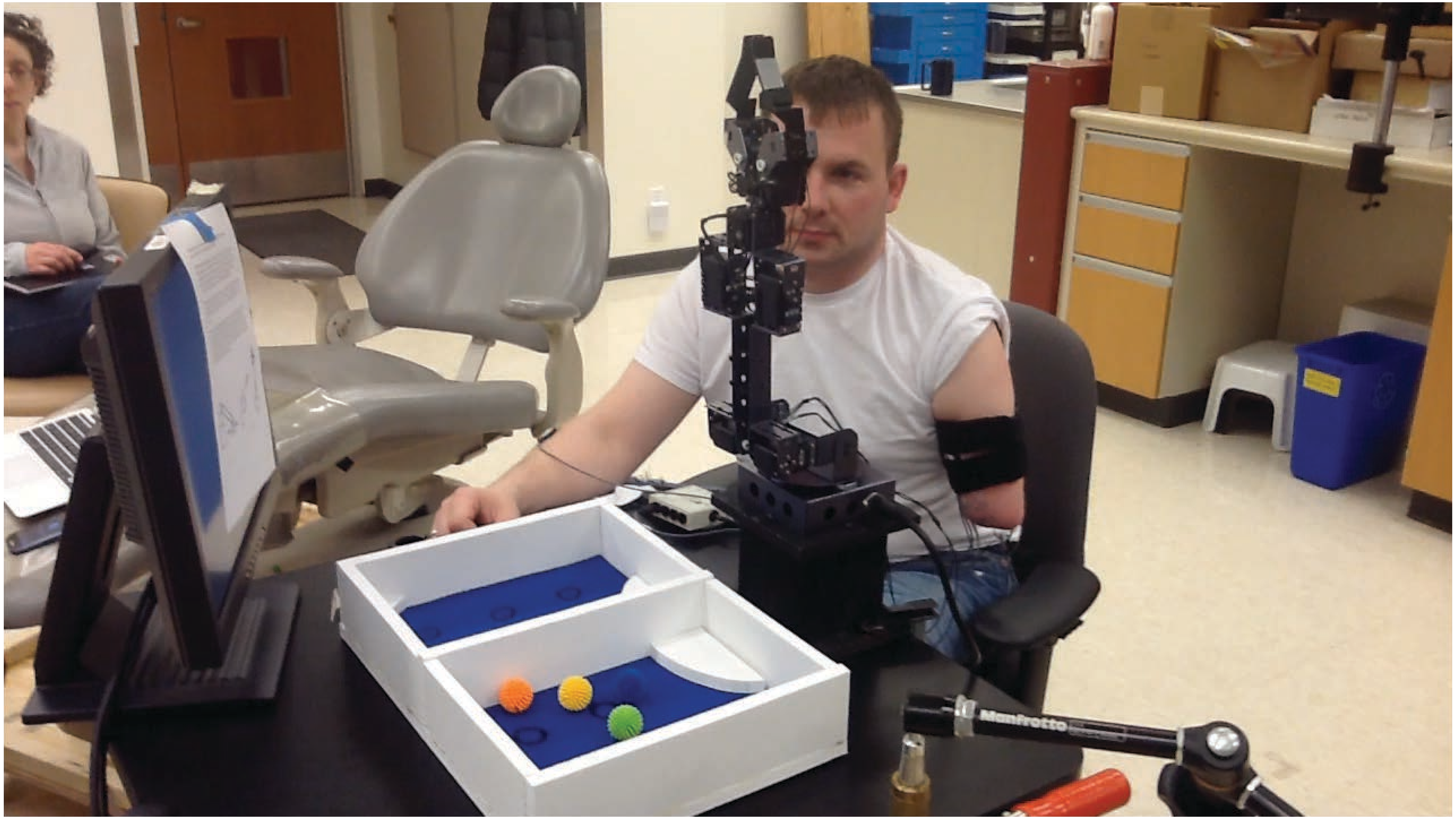
EXAMPLE 1

adaptive switching

(predictions change an interface)



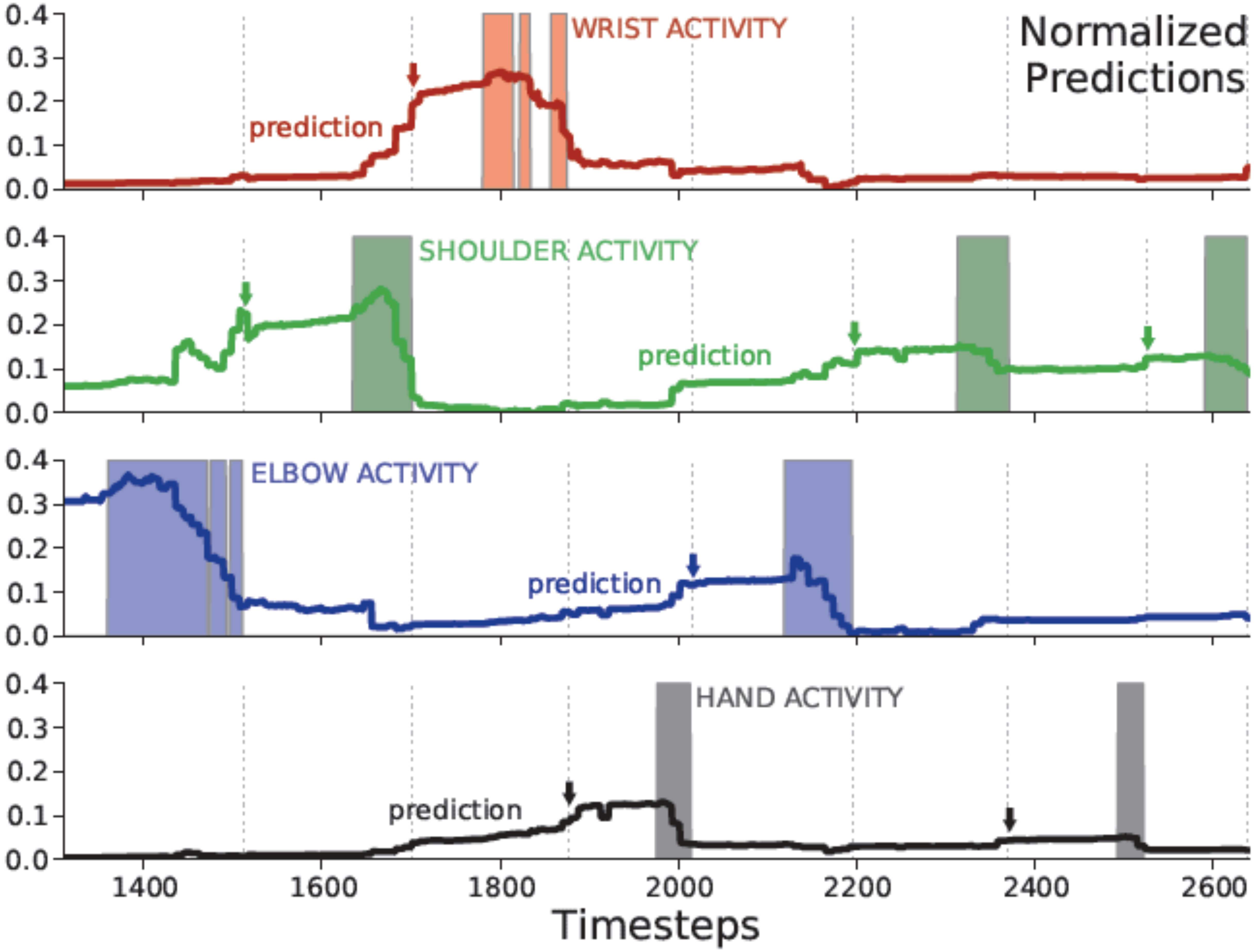
A. L. Edwards, “Adaptive and Autonomous Switching: Shared Control of Powered Prosthetic Arms Using Reinforcement Learning,” MScRS Thesis, Faculty of Rehabilitation Medicine, University of Alberta, 2016.



Adaptive Switching

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

Predicting the Future



Pilarski et al., 2012, BioRob

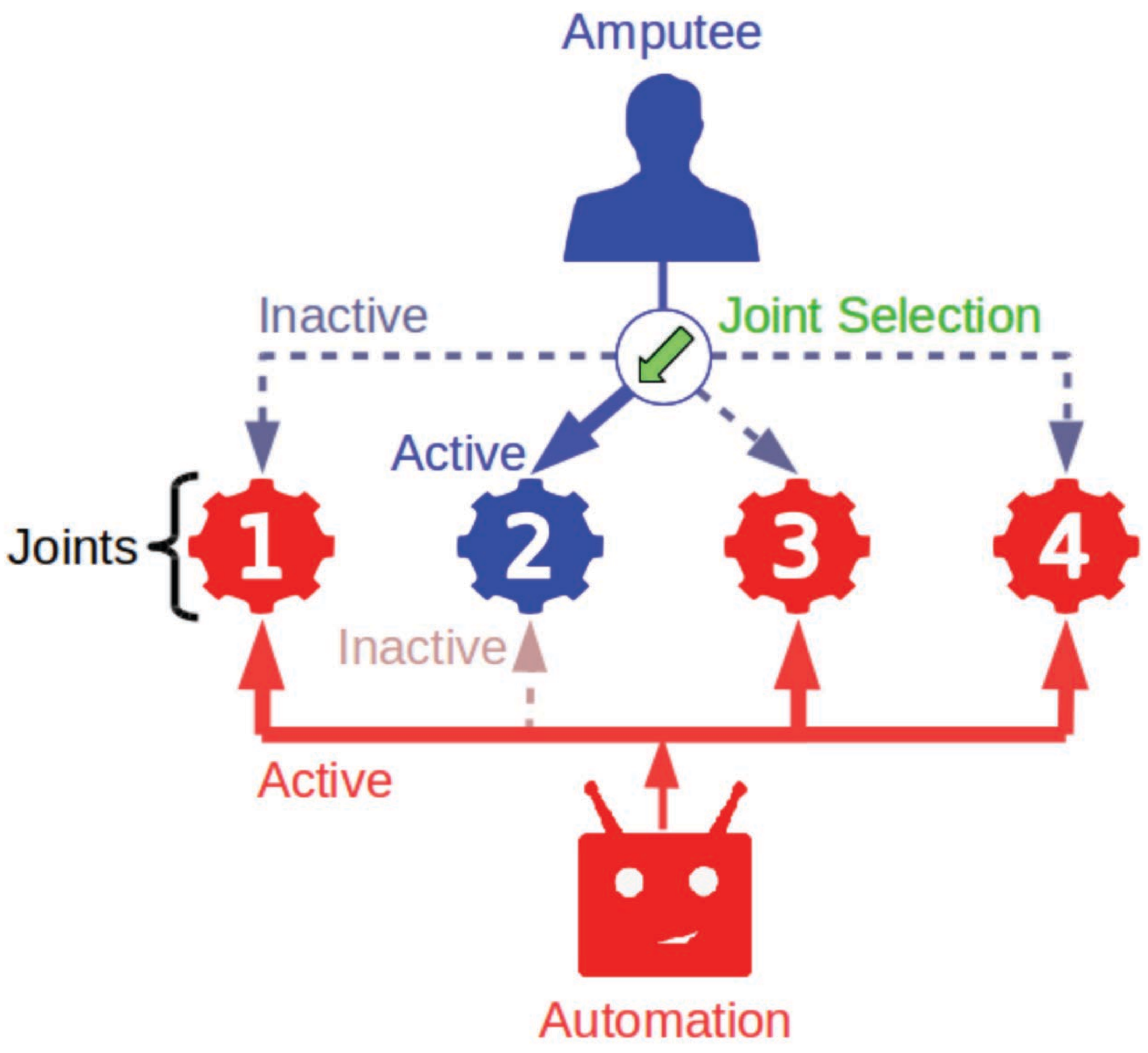
Autonomous Switching Edwards et al., *BioRob*, 2016

(learning and unlearning automatic control actions)

EXAMPLE 2

motor synergies

(predictions as actions)



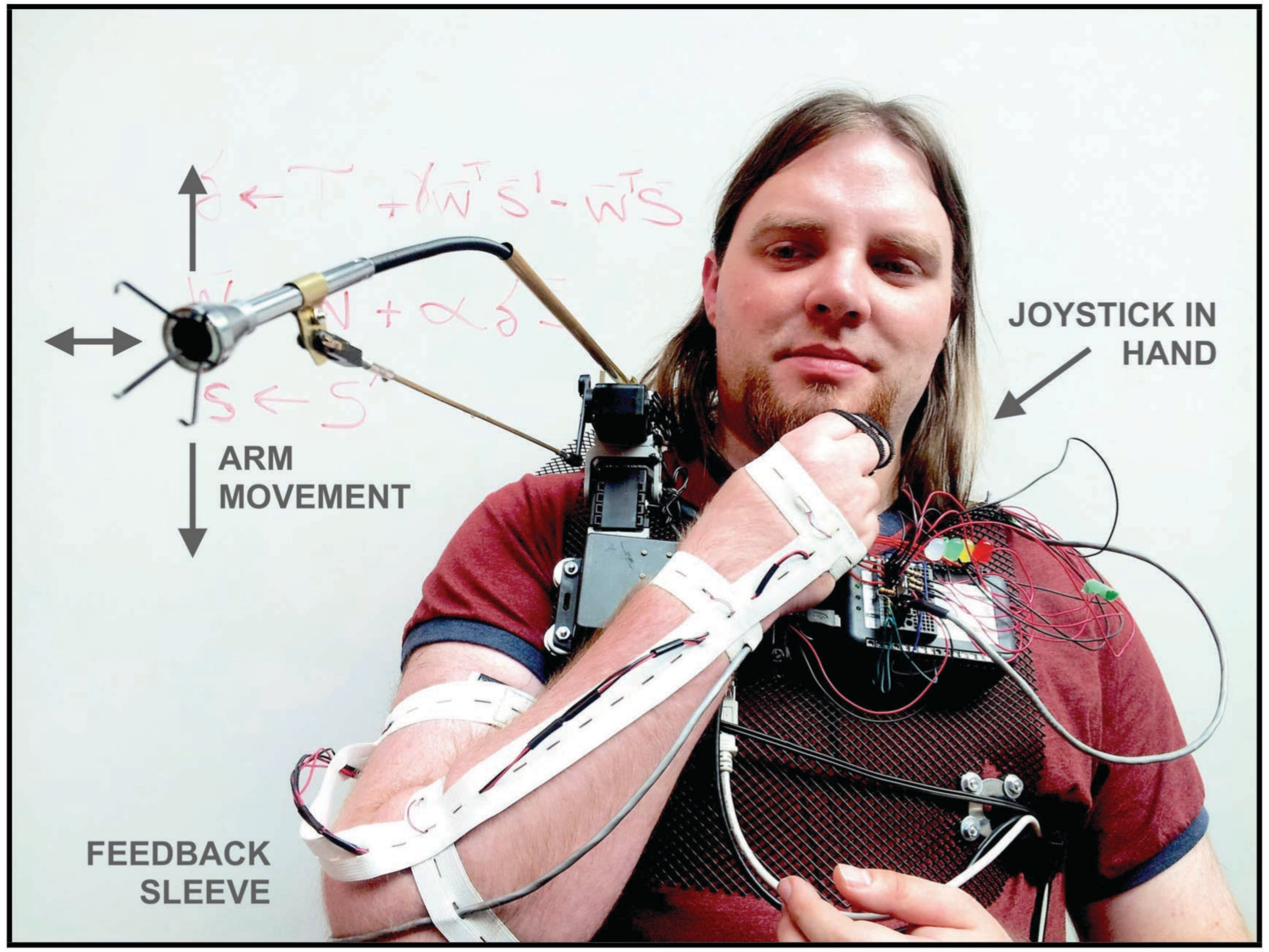
C. Sherstan, J. Modayil, P.M. Pilarski, "A Collaborative Approach to the Simultaneous Multi-joint Control of a Prosthetic Arm," *Proc. of the 14th IEEE/RAS-EMBS International Conference on Rehabilitation Robotics (ICORR)*, August 11–14, Singapore, 2015, pp. 13–18.

P.M. Pilarski, T.B. Dick, and R.S. Sutton, "Real-time Prediction Learning for the Simultaneous Actuation of Multiple Prosthetic Joints," *Proc. of the 2013 IEEE International Conference on Rehabilitation Robotics (ICORR)*, Seattle, USA, June 24–26, 2013. 8 pages.

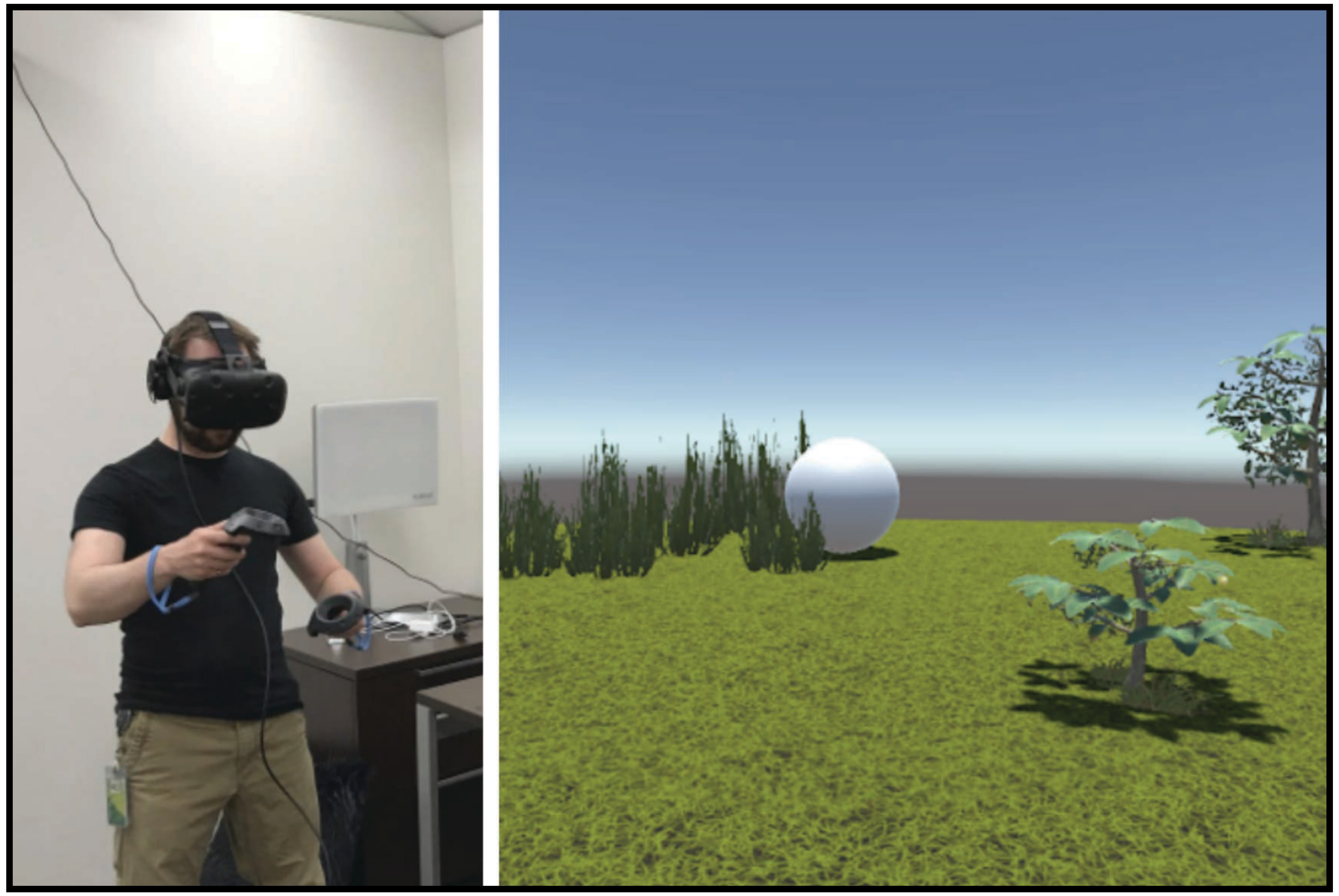
EXAMPLE 3

communication

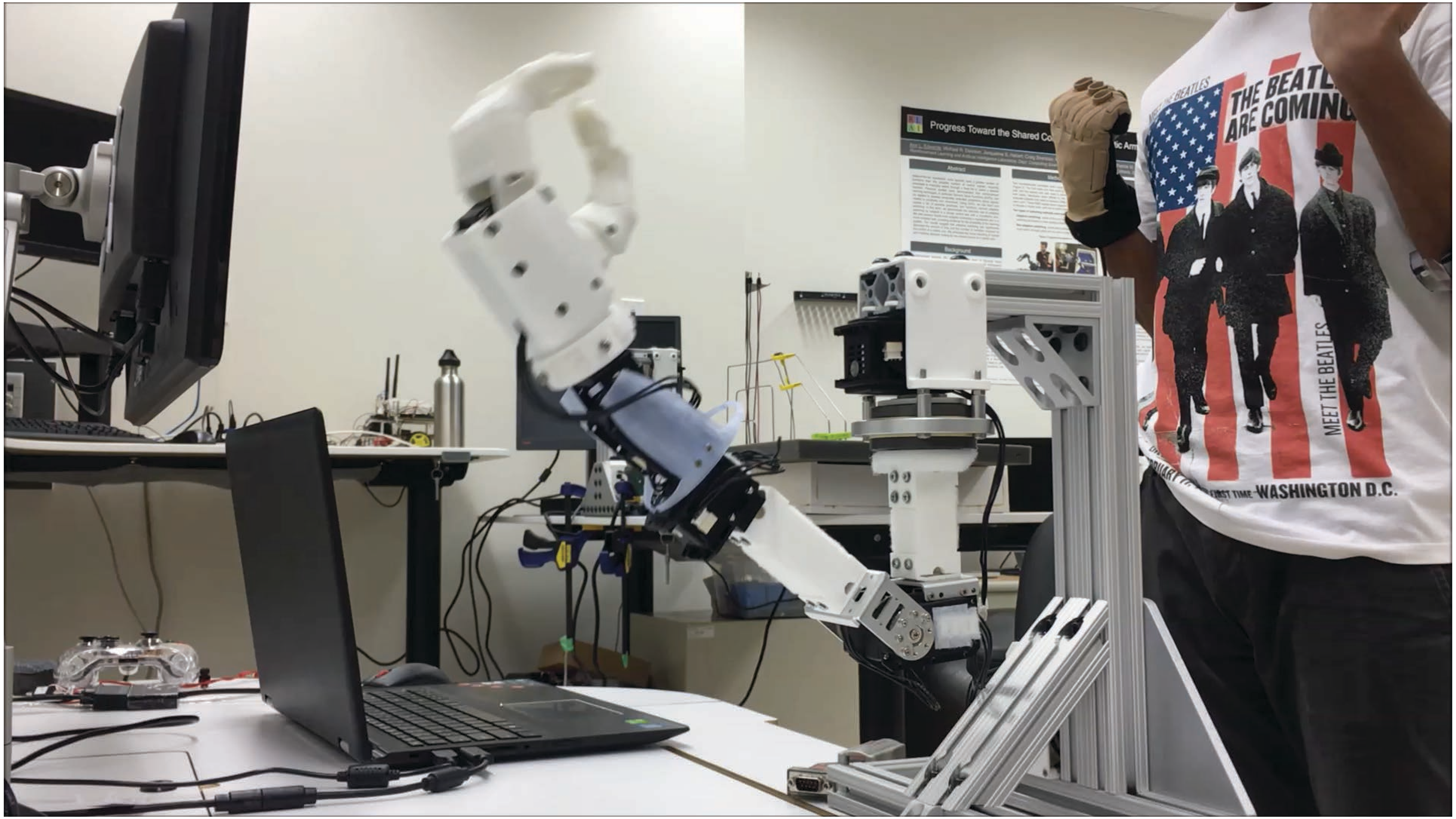
(predictions as feedback)



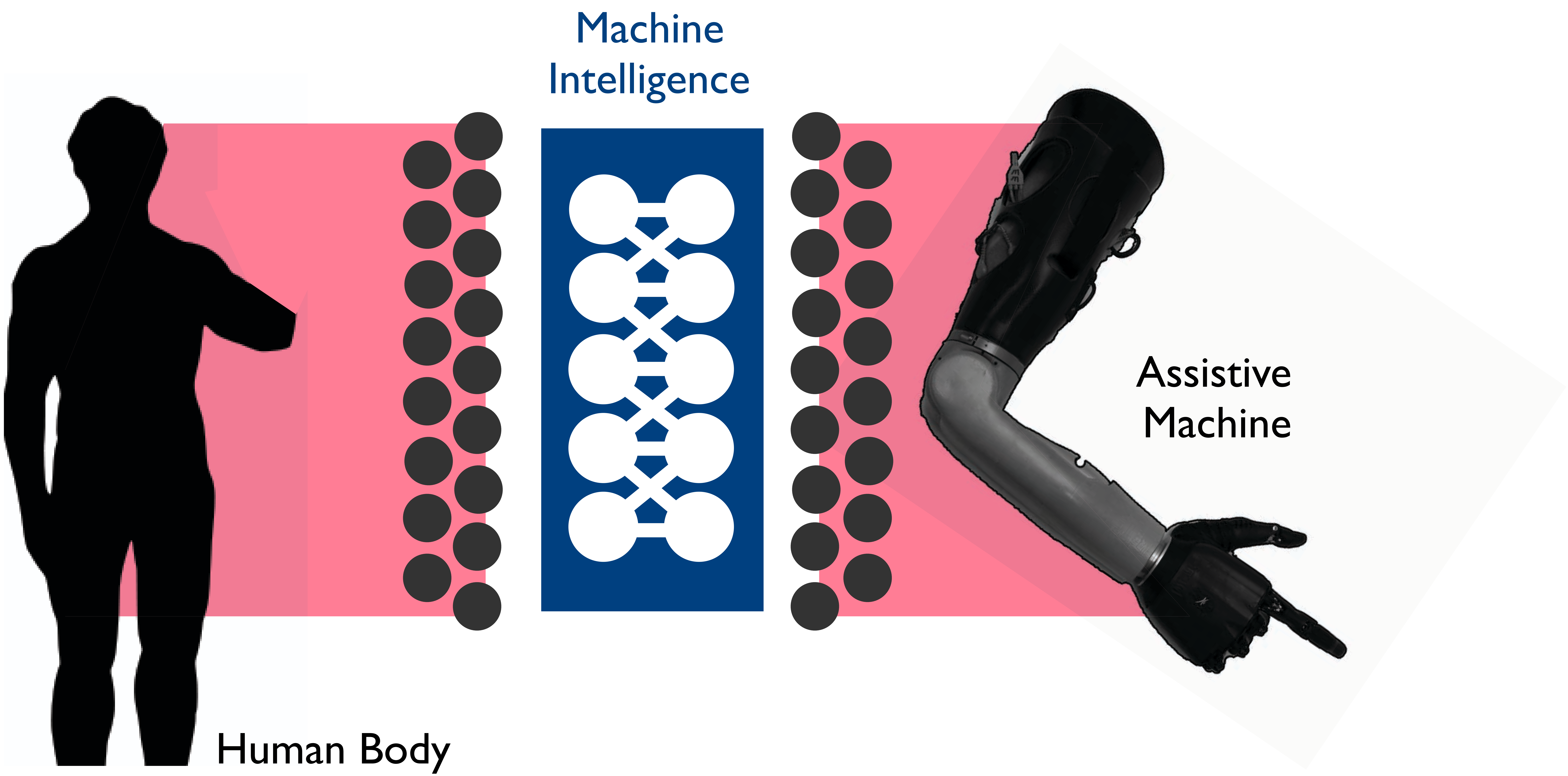
A. S. R. Parker, A. L. Edwards, P. M. Pilarski, "Exploring the Impact of Machine-Learned Predictions on Feedback from an Artificial Limb," *2019 IEEE-RAS-EMBS International Conference on Rehabilitation Robotics (ICORR)*, 24-28 June, 2019, Toronto, 8 pages.



P. M. Pilarski, A. Butcher, M. Johanson, M. M. Botvinick, A. Bolt, A. S. R. Parker, "Learned human-agent decision-making, communication and joint action in a virtual reality environment," *RLDM 2019 / arXiv:1905.02691* [cs.AI], 5 pages, 2019.



G. Vasan, “Teaching a Powered Prosthetic Arm with an Intact Arm Using Reinforcement Learning,” MSc Thesis, Dept. Computing Science, University of Alberta, 2017.


Machine
Intelligence
Human Body
Assistive
Machine

Exoskeletons: UC Berkeley spin-off suitX exoskeleton technology; https://www.youtube.com/watch?v=I3roYI3CB2Y


Machine
Intelligence
Human Mind
Perception
Action
Cognition

#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.

hello@amii.ca
www.amii.ca

Amii is proudly supported by:

UNIVERSITY OF ALBERTA
ALBERTA INNOVATES
Alberta Economic Development and Trade
Canada