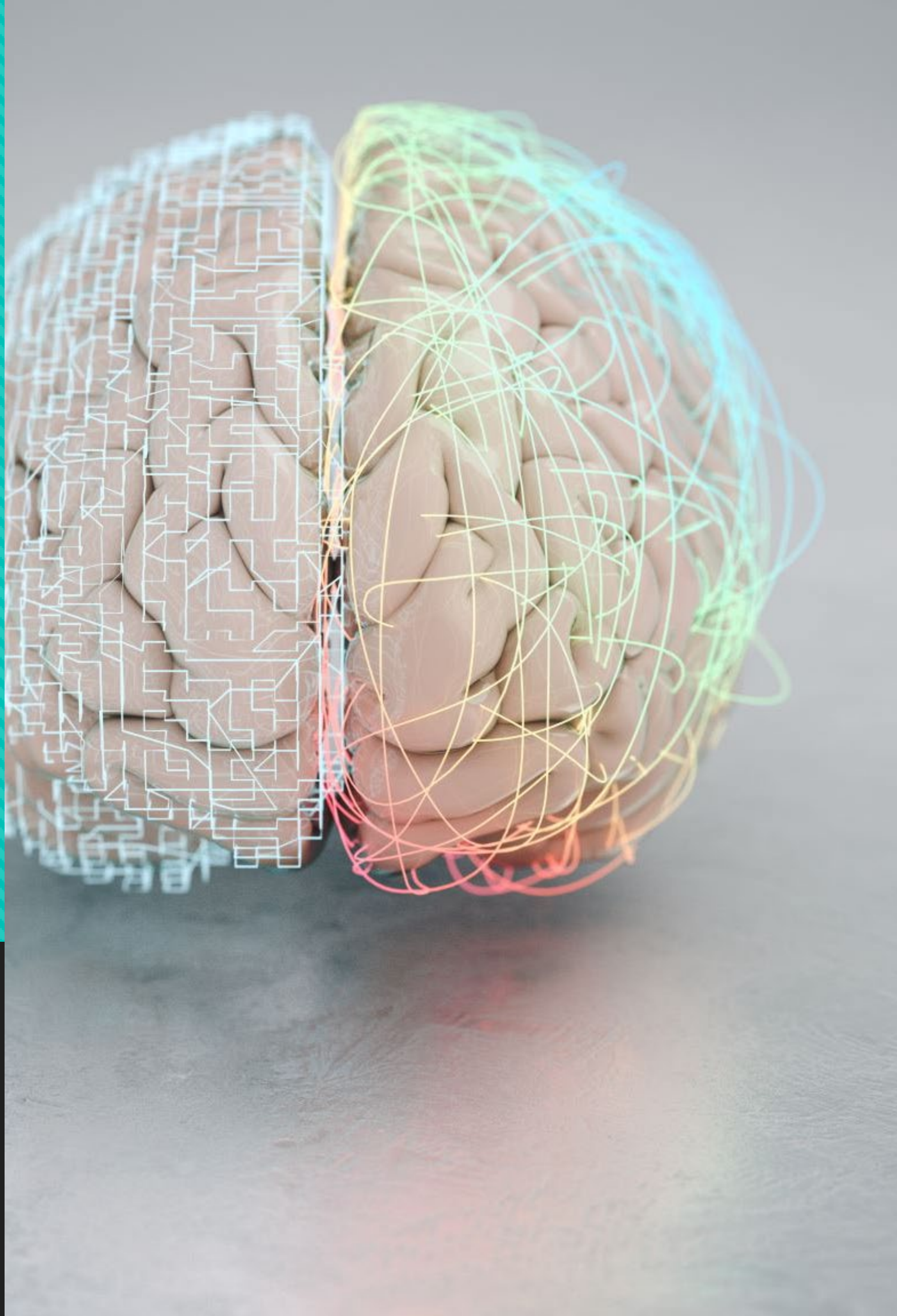
Neural Decoders in Reinforcement Learning Brain Machine Interfaces

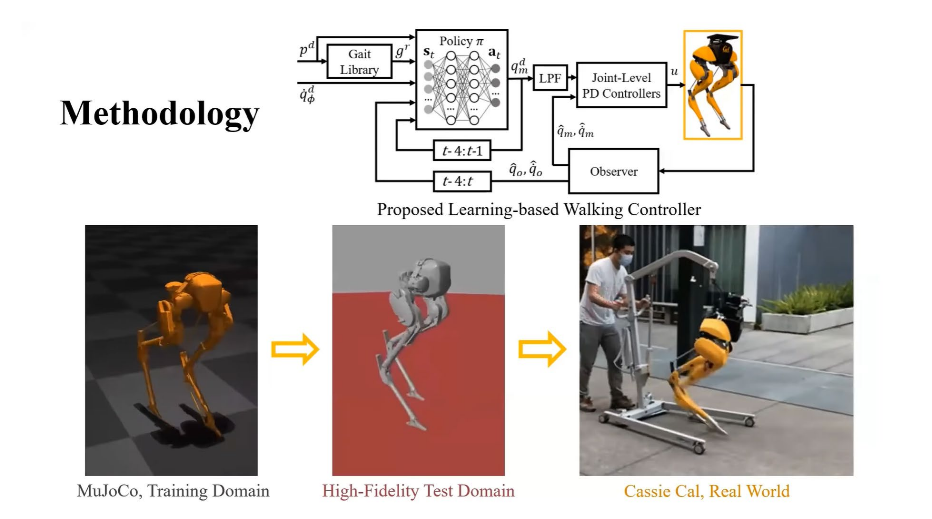
Jihye Bae, PhD
Electrical and Computer Engineering
University of Kentucky



Contents

Powerful Reinforcement Learning (RL)

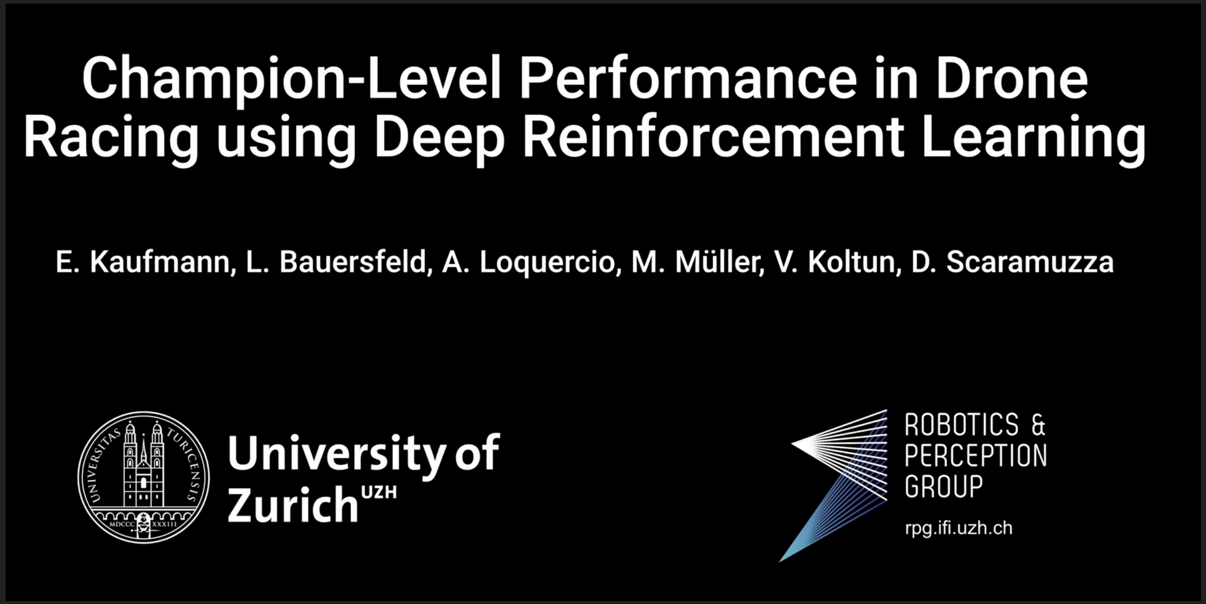
- AlphaGo by Google DeepMind
- Automated Robotics



https://www.technologyreview.com/2021/04/08/1022176/boston-dynamics-cassie-robot-walk-reinforcement-learning-ai/

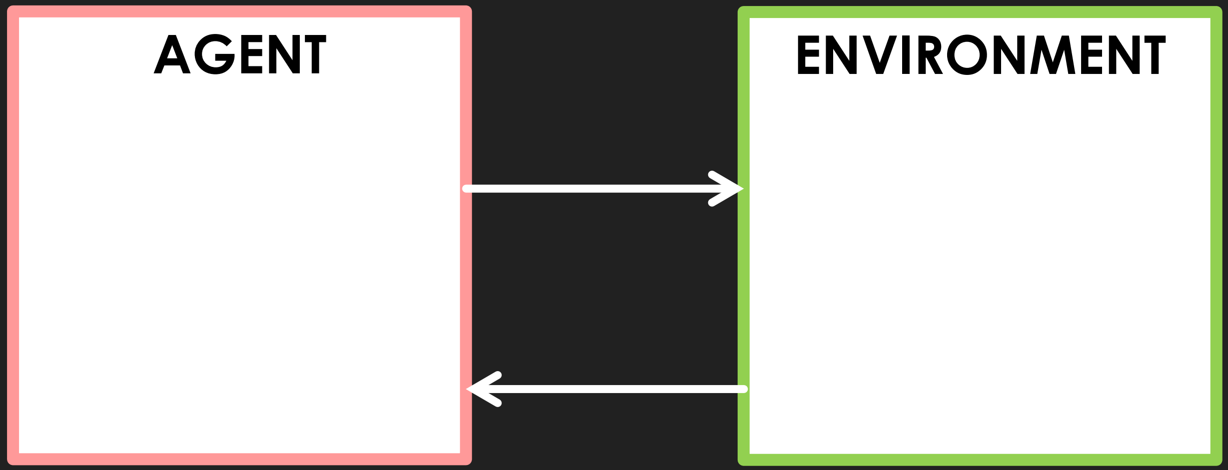
Powerful Reinforcement Learning (RL)

- Recently Reported Drone Racing



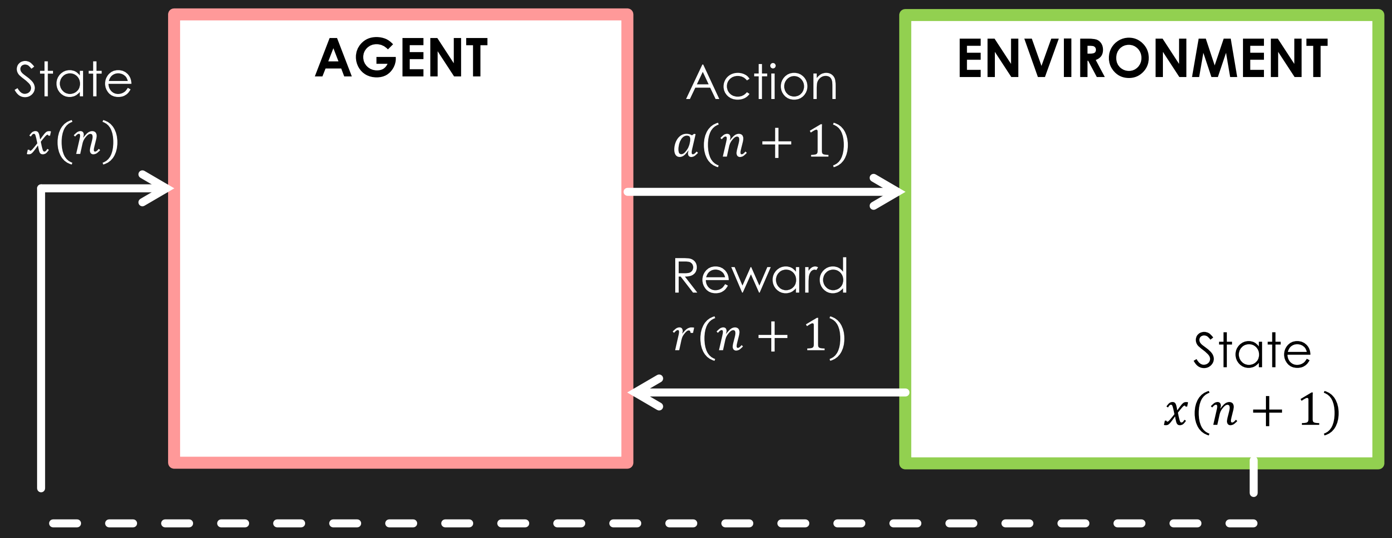
https://www.youtube.com/watch?v=fBiataDpGIo

Principle of RL



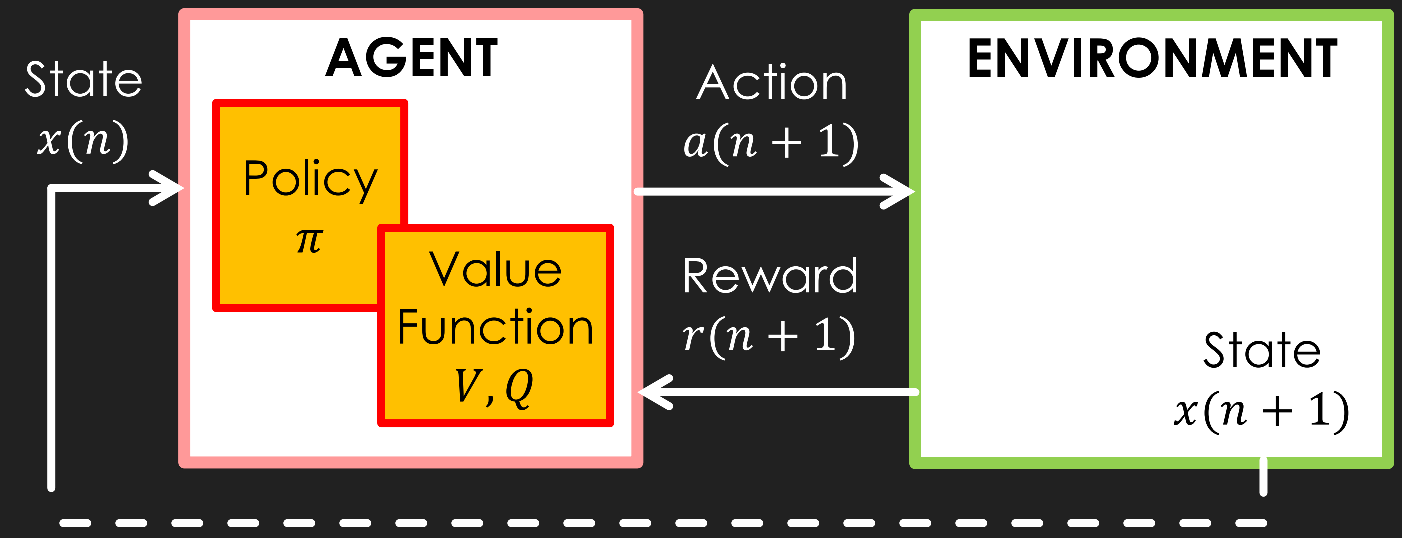
- Learning from the interaction between **Agent** and **Environment**

Key Components in RL



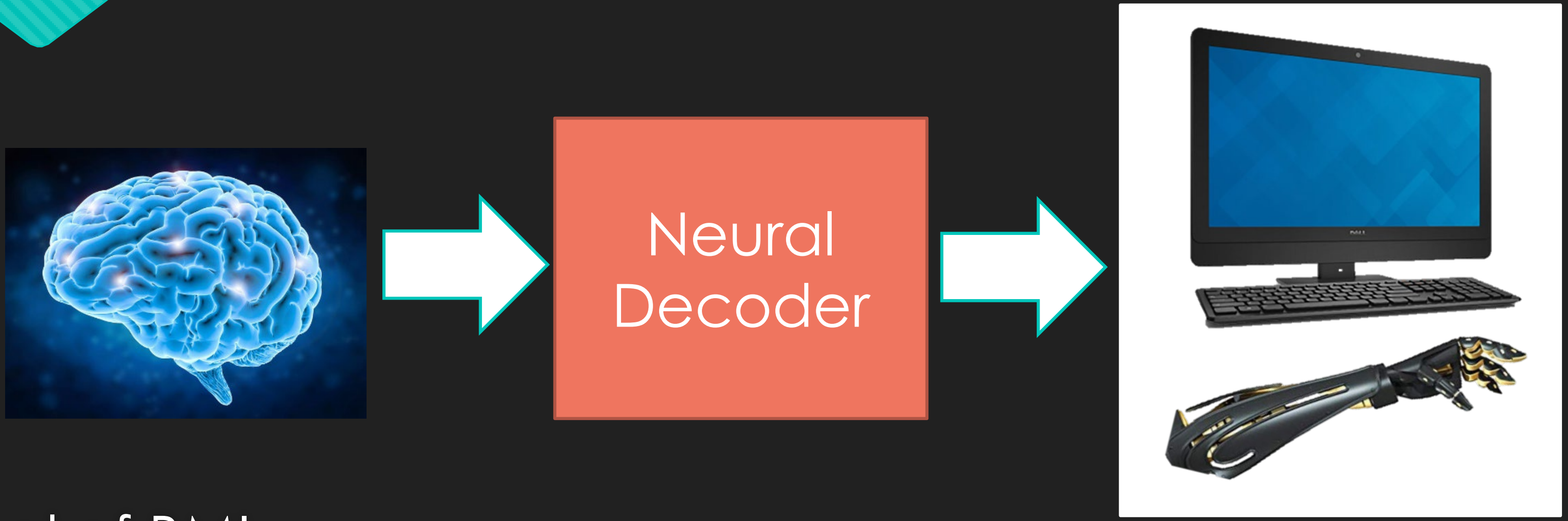
- **Agent** and **Environment** communicate in terms of **State**, **Action**, and **Reward**
- Goal of Agent is to maximize cumulative Reward
- RL is well-suited for multistep prediction problems (Markov Decision Process)

Key Components in Agent



- **Policy:** mapping function from state to action, $\pi{:}X \rightarrow A$
- **Value Function:** measure of long-term performance
  - **State value function:** $V\big(x(n)\big) = E[R(n)|x(n)]$
  - **Action value function:** $Q^{\pi}(x(n), a(n)) = E[R(n)|x(n), a(n)]$

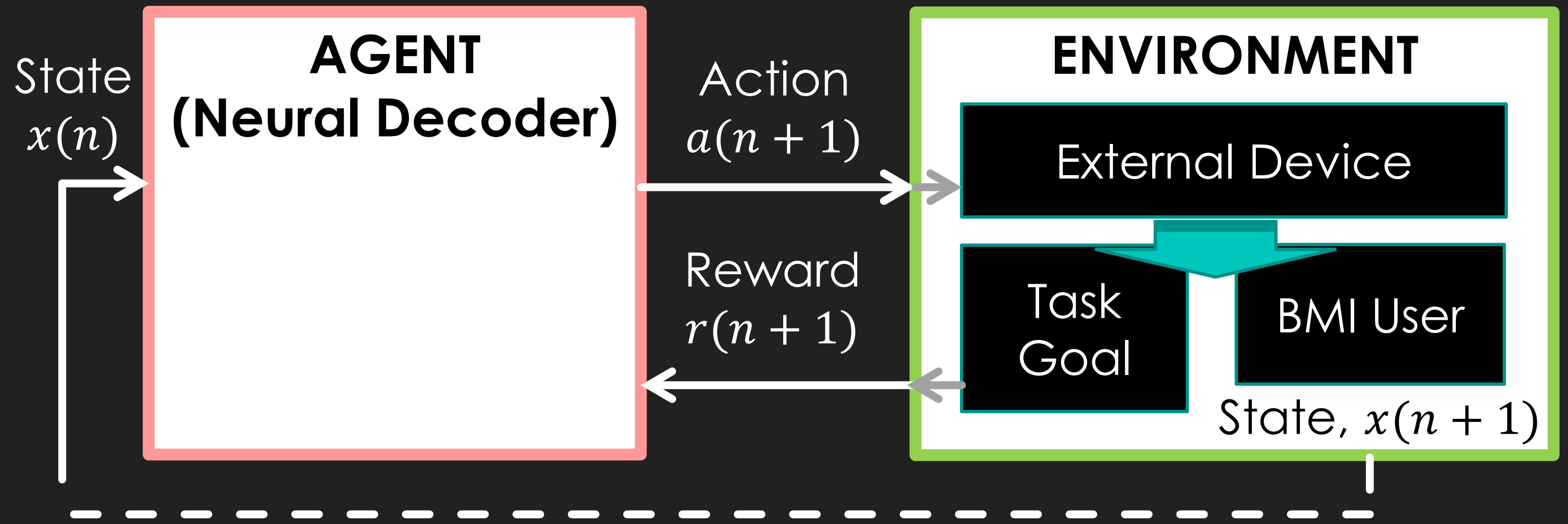
Key Element in BMIs



Images from Google Image

- Goal of BMI
  - Direct communication between user and external device
- **A neural decoder is a key element in BMIs.**
  - Neural decoder: a system that maps neural signals into control commands for external devices

Structure of Reinforcement Learning in Brain Machine Interfaces

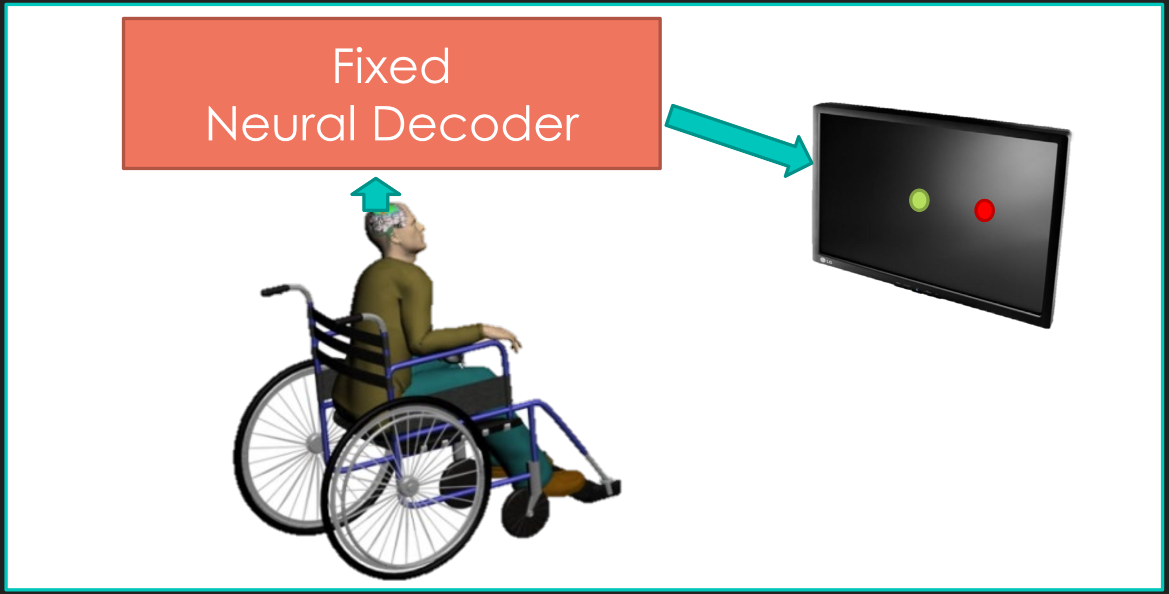


- **Neural Decoder can be considered as the Agent.**

J. DiGiovanna, B. Mahmoudi, J. Fortes, J. C. Principe, J. C. Sanchez. "Coadaptive Brain–Machine Interface via Reinforcement Learning", IEEE Trans. On Biomedical Engineering, 15(6), 2009.

Limitations of Conventional Approach

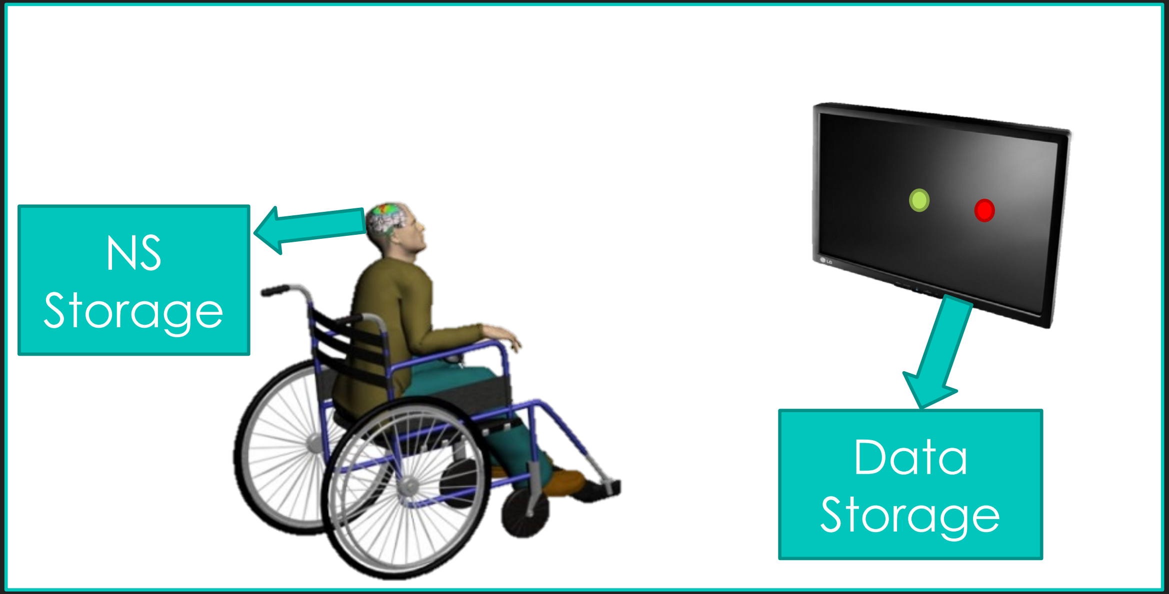
- **Supervised learning**: learning based on desired signal



- Requires training and testing modes
- **No instantaneous updates; No system adaptation**

Supervised Learning Paradigm in BMIs

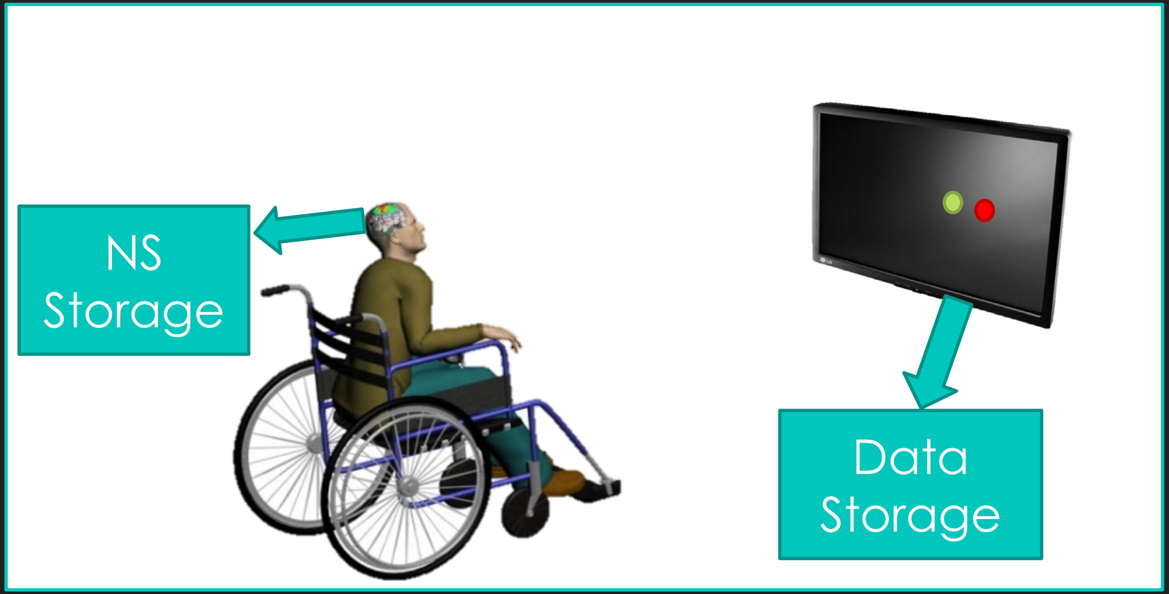
- **Training Mode**



- Patient imagines controlling the cursor while observing the cursor's predefined movement
- Both NS and Cursor information are stored

Supervised Learning Paradigm in BMIs

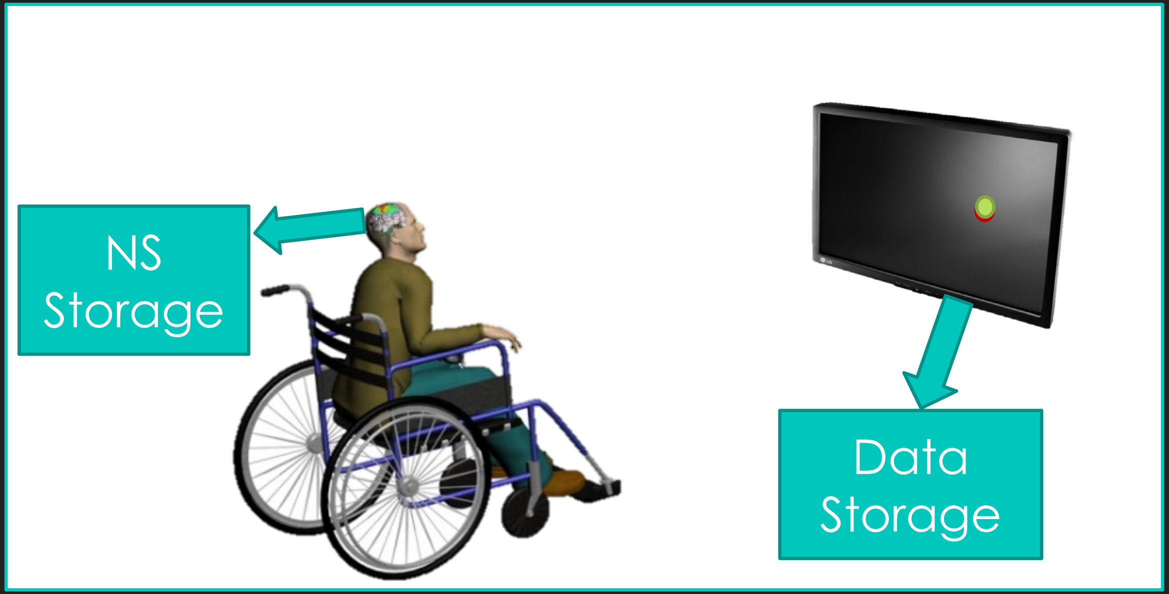
- Training Mode



- Patient imagines controlling the cursor while observing the cursor's predefined movement
- Both NS and Cursor information are stored

Supervised Learning Paradigm in BMIs

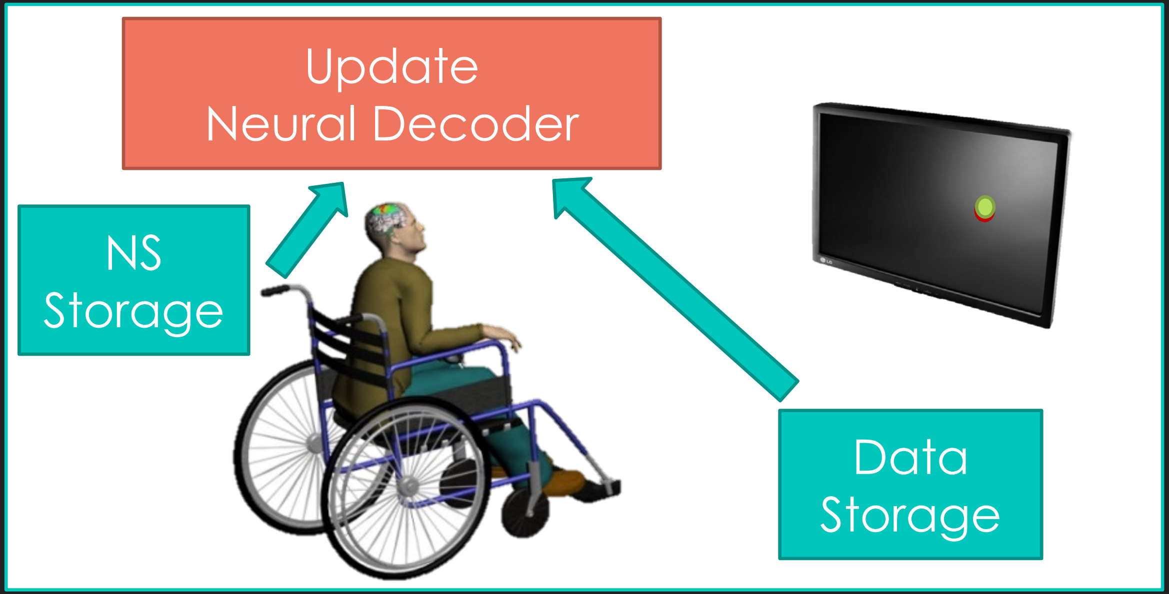
- Training Mode



- Patient imagines controlling the cursor while observing the cursor's predefined movement
- Both NS and Cursor information are stored

Supervised Learning Paradigm in BMIs

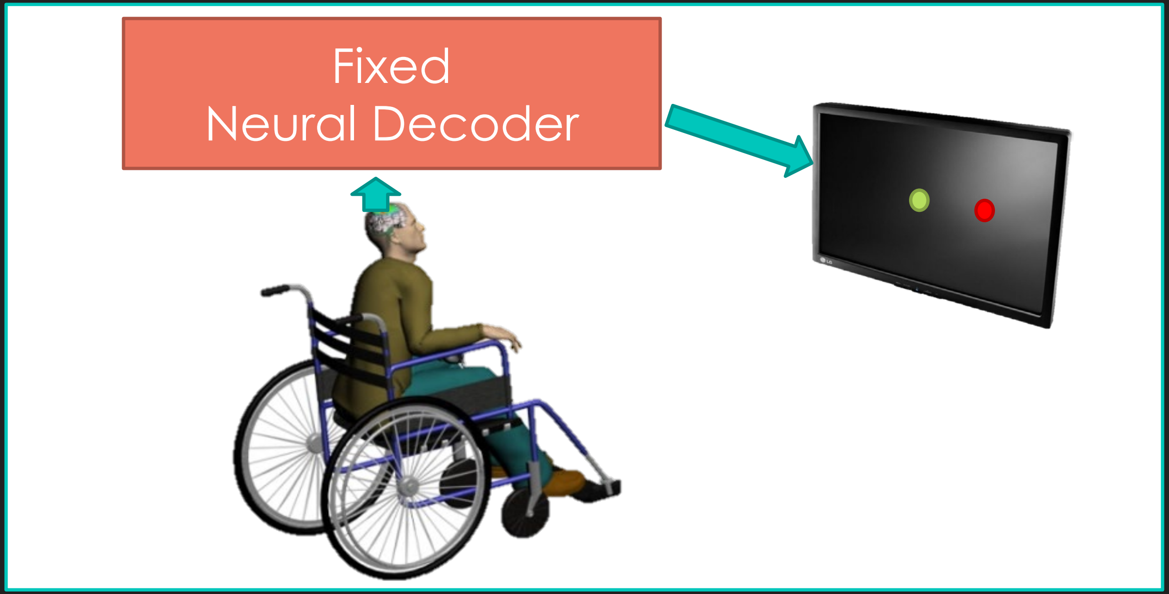
- **Neural Decoder Update**



- Neural Decoder parameters are updated using stored NS and cursor information
- Goal: minimize difference with desired signal

Supervised Learning Paradigm in BMIs

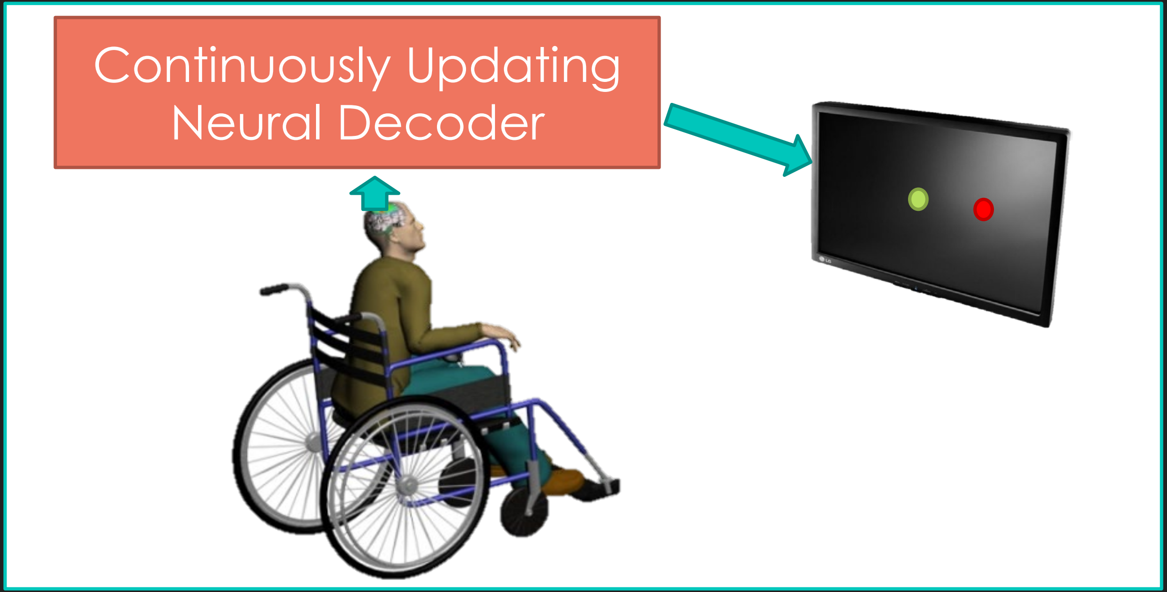
- **Testing Mode**: use fixed Neural Decoder



- **No instantaneous updates; No system adaptation**
- Reinforcement Learning can overcome these issues!

Basic Idea of Reinforcement Learning in BMIs

- **Reinforcement learning: learning based on reward**

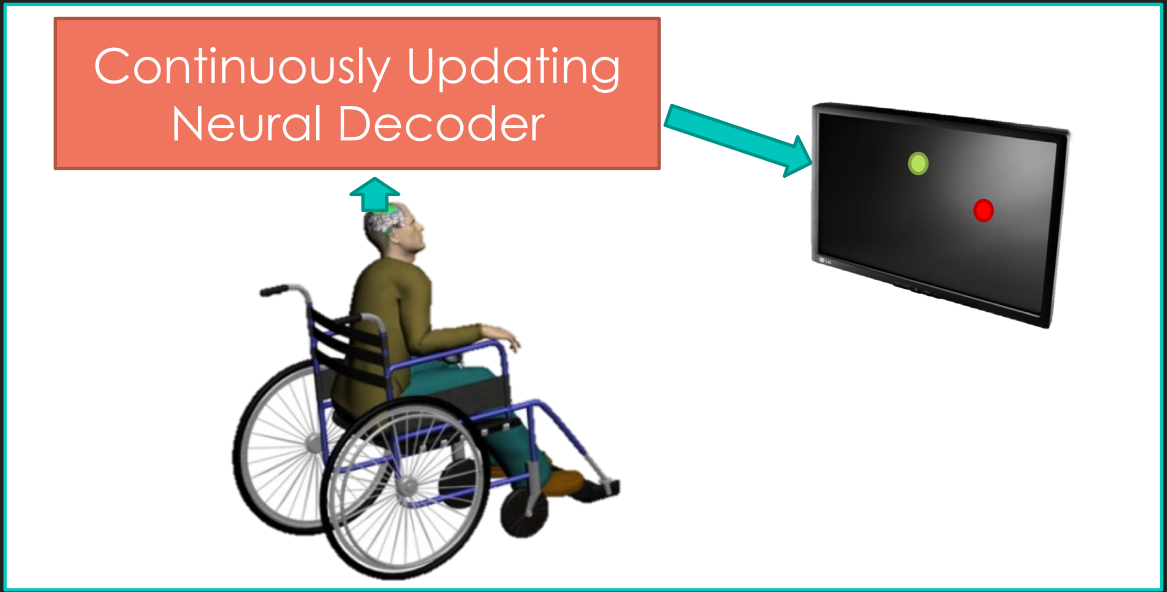


- Neural Decoder can be continuously updated
- Goal: maximize cumulative reward

$$R(n) = \sum_{i=0}^{\infty} \gamma^i r(n+i),\ 0 < \gamma < 1$$

Basic Idea of Reinforcement Learning in BMIs

- **Reinforcement learning: learning based on reward**

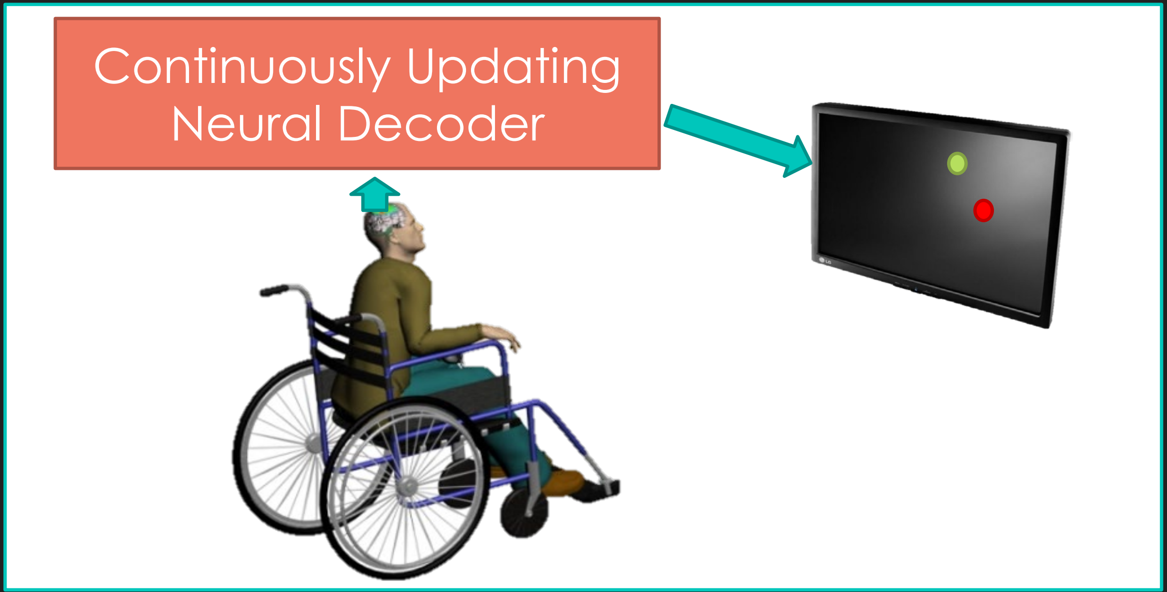


- Neural Decoder can be continuously updated
- Goal: maximize cumulative reward

$$R(n) = \sum_{i=0}^{\infty} \gamma^i r(n+i), \; 0 < \gamma < 1$$

Basic Idea of Reinforcement Learning in BMIs

- **Reinforcement learning: learning based on reward**

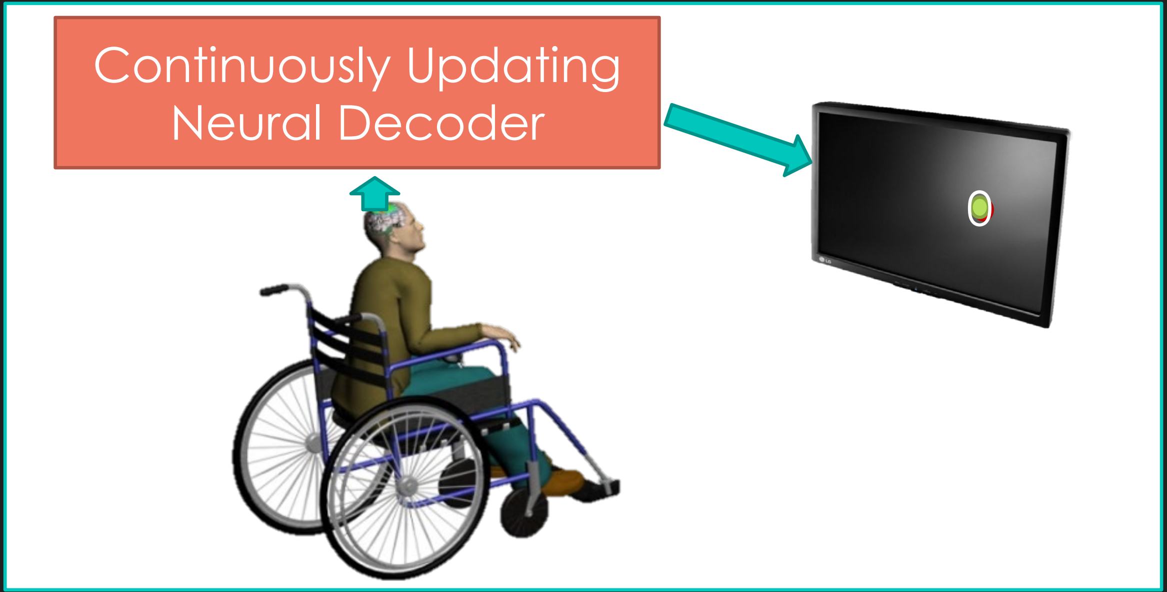


- Neural Decoder can be continuously updated
- Goal: maximize cumulative reward

$$R(n) = \sum_{i=0}^{\infty} \gamma^i r(n+i), \; 0 < \gamma < 1$$

Basic Idea of Reinforcement Learning in BMIs

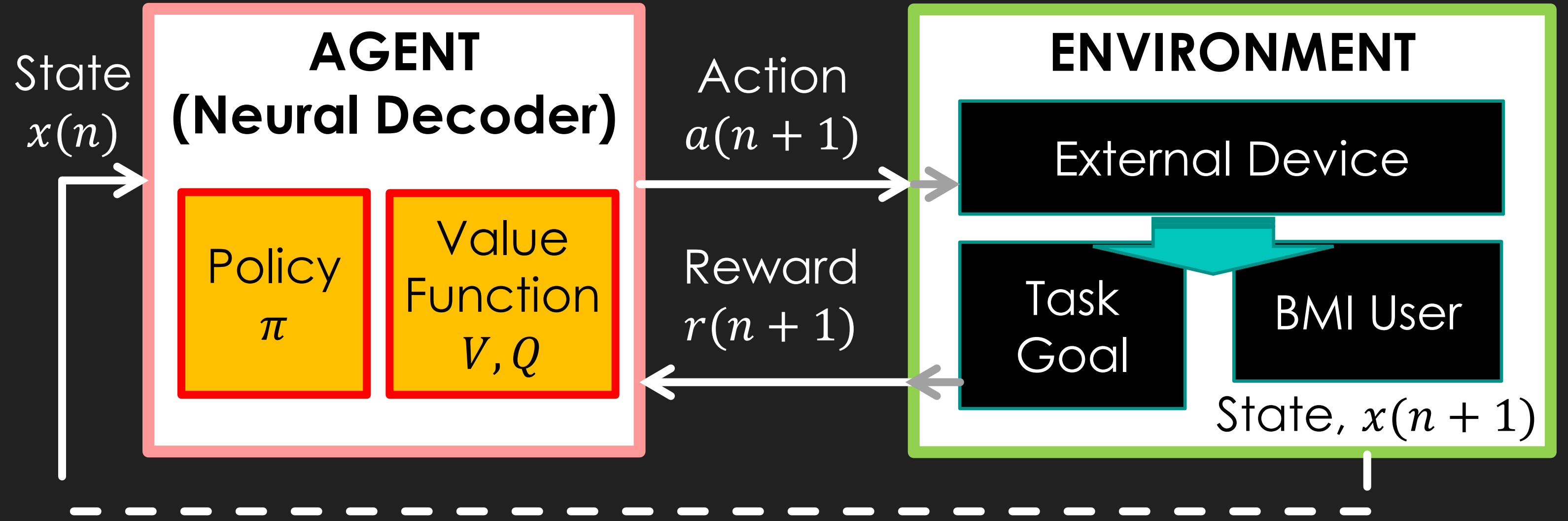
- **Reinforcement learning: learning based on reward**



- Neural Decoder can be continuously updated
- Goal: maximize cumulative reward

$$R(n) = \sum_{i=0}^{\infty} \gamma^i r(n+i),\ 0 < \gamma < 1$$

Neural Decoders in RLBMIs



- **Functional approximations** of the value function and policy are the key approaches for neural decoders in RLBMIs

B. Girdler, W. Caldbeck, and J. Bae. "Neural Decoders Using Reinforcement Learning in Brain Machine Interfaces: A Technical Review." Frontiers in Systems Neuroscience, 16, 2022, doi=10.3389/fnsys.2022.836778.

Functional approximations of the value function and policy

- Value Function Approximations
  - $\tilde{V}^{\pi}(x_t) = f_v(x_t; \theta_{f_v})$
  - $\tilde{Q}^{\pi}(x_t, a_t) = f_q(x_t, a_t; \theta_{f_q})$
- Policy Approximations
  - $\pi: a_t \approx f_\pi(x_t; \theta_{f_\pi})$

2 Mainly Considered RL Models in Neural Decoders

- Q-learning
- Actor-Critic

Q-learning

- Goal: find an optimal action sequence

$$A(n) = \underset{a}{\operatorname{argmax}}\, Q^*(x, a)$$

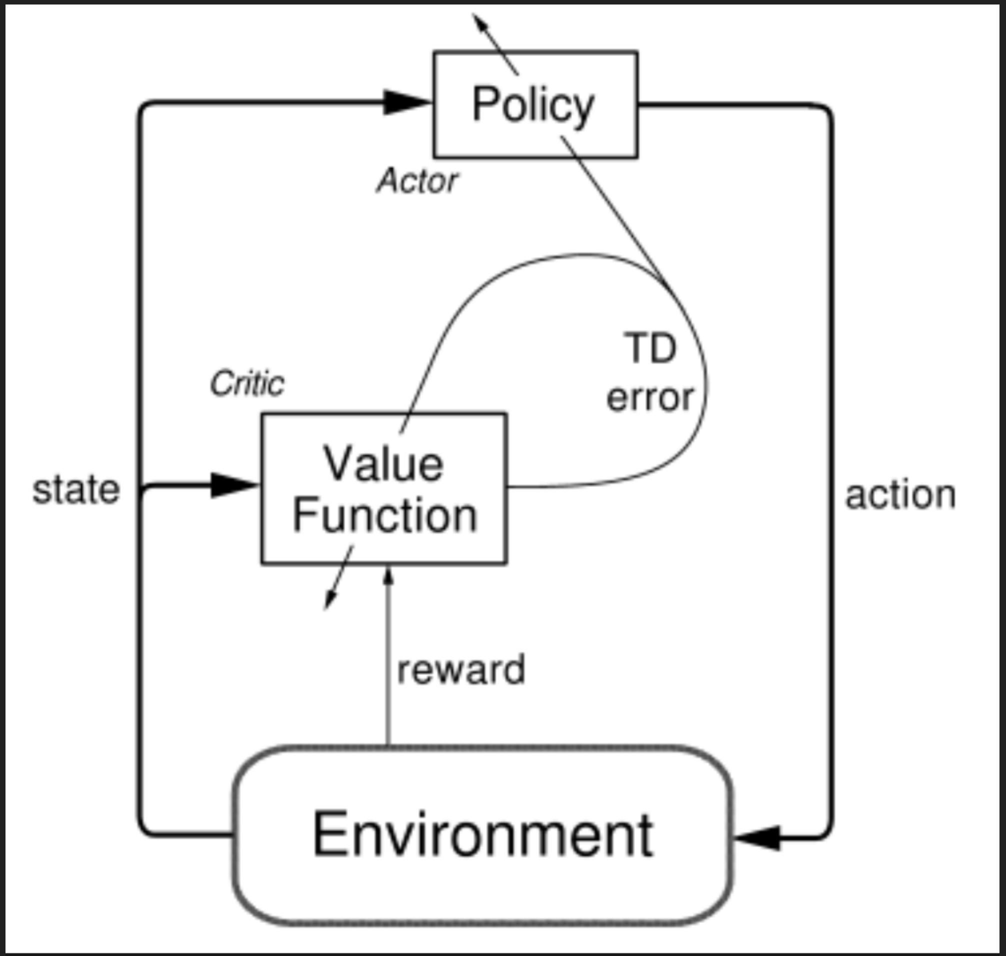
- Q-learning finds a policy using the action value function Q
  - Basic idea: $\pi \geq \pi' \;\; iff \;\; Q^{\pi}(x, a) \geq Q^{\pi'}(x, a)$
- **Function approximation of the action value function**

$$\boldsymbol{A(n) = \underset{a}{argmax}\, \tilde{Q}(x, a)}$$

B. Girdler, W. Caldbeck, and J. Bae. "Neural Decoders Using Reinforcement Learning in Brain Machine Interfaces: A Technical Review." Frontiers in Systems Neuroscience, 16, 2022, doi=10.3389/fnsys.2022.836778.

Actor-Critic

- Actor: Decide which action to take (Policy)
- Critic: Tells the Actor how good its action was and how it should adjust (Value Function, V)



R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 2018.

2 Commonly Used Functional Approximators in Neural Decoders

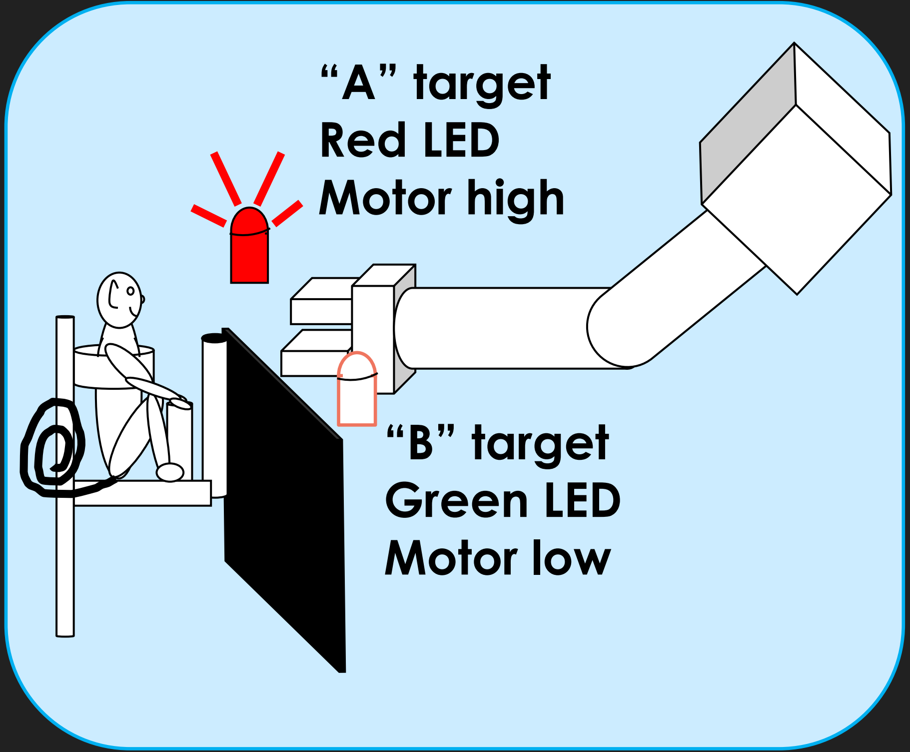
- Kernel Expansions: nonlinearly map the input data to a high-dimensional feature space of vectors

- Artificial Neural Networks (Feedforward Neural Networks and Convolutional Neural Networks)

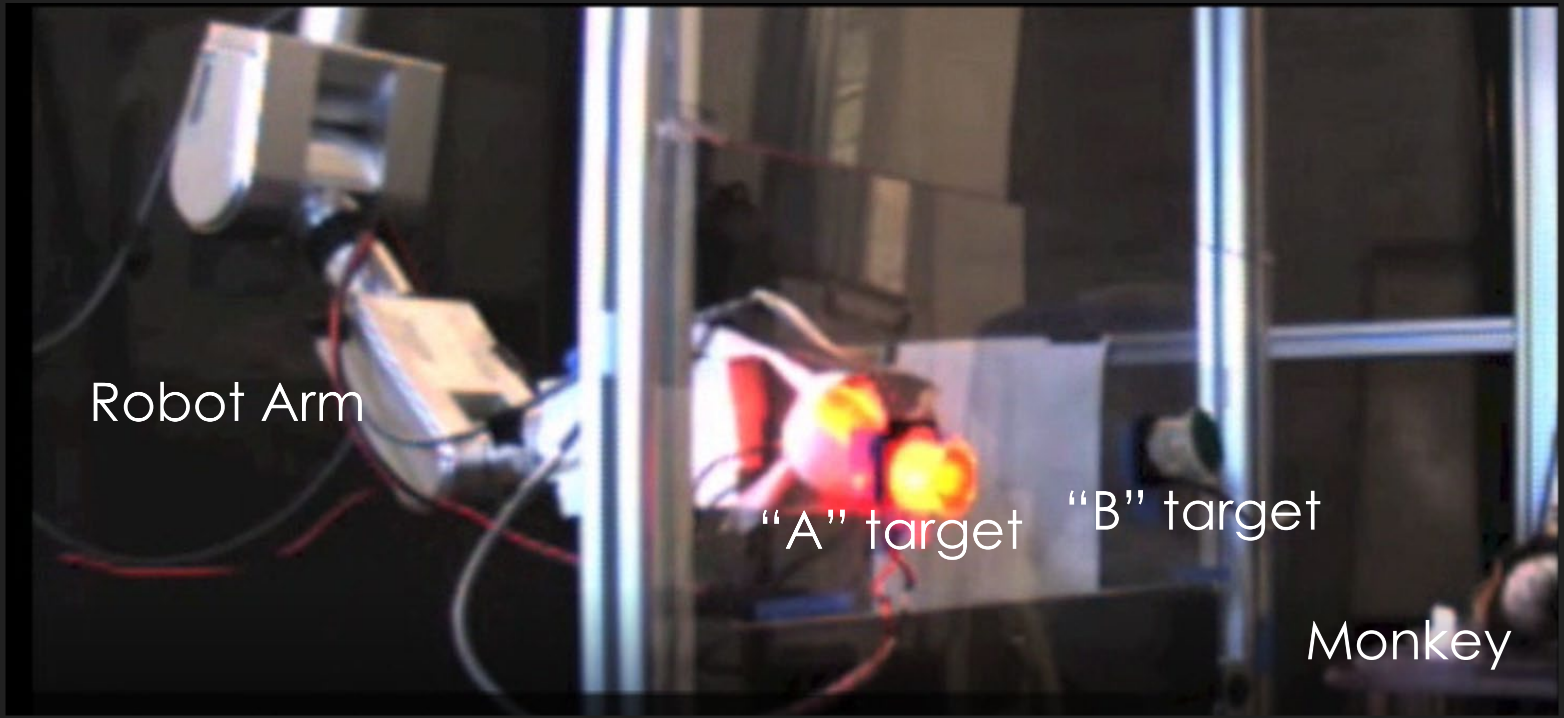
B. Girdler, W. Caldbeck, and J. Bae. "Neural Decoders Using Reinforcement Learning in Brain Machine Interfaces: A Technical Review." Frontiers in Systems Neuroscience, 16, 2022, doi=10.3389/fnsys.2022.836778.

Example: Q-KTD (λ) in Closed-Loop RLBMI

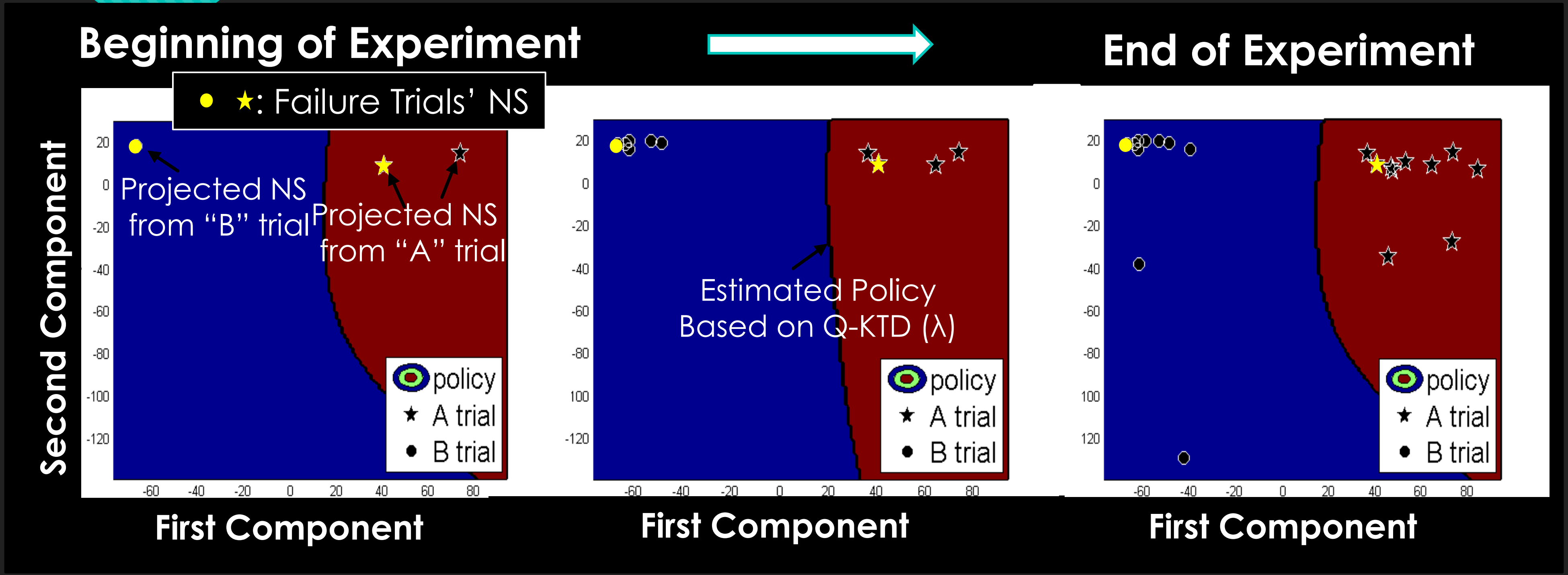
- Q-KTD(λ): Integration of Q-learning and Kernel methods
- Task evaluated: Go/No Go task
- Marmoset monkey implanted in motor cortex (M1) with 10 electrodes (14 units)
- Monkey trained to associate motor high and low states with the targets



Closed-Loop RLBMI Implementation Video



Adaptive Behavior of Q-KTD (λ)



- The decoder can estimate a **nonlinear policy**
- The decoder **adapts to the environment on-line**

J.Bae, L. G. Sanchez Giraldo, E. A. Pohlmeyer, J. C. Sanchez, J. Principe ,A New Method of Concurrently Visualizing States, Values, and Actions in Reinforcement Based Brain Machine Interfaces, *EMBC*, 2013.

High Success Rates of Q-KTD (λ)

- No use of pre-trained system
- **The decoder can find an appropriate state to action map.**

Success rates for 4 experiments

	Total Number of Trials (A, B Trial)	Success Rates (%)
Exp. 1	20(10,10)	90.00
Exp. 2	32(16,16)	84.38
Exp. 3	53(27, 26)	77.36
Exp. 4	52(27, 25)	78.85

J. Bae, L. G. Sanchez Giraldo, E. A. Pohlmeyer, J. T. Francis, J. C. Sanchez, and J. C. Principe. "Kernel Temporal Differences for Neural Decoding." Computational Intelligence and Neuroscience, 2015.

Limitations of RLBMIs

- Slow learning in the beginning due to the nature of trial and error-based learning paradigm
- Most reported works use intracortical signals in animals
- No study provided to compare different decoders

Future Directions

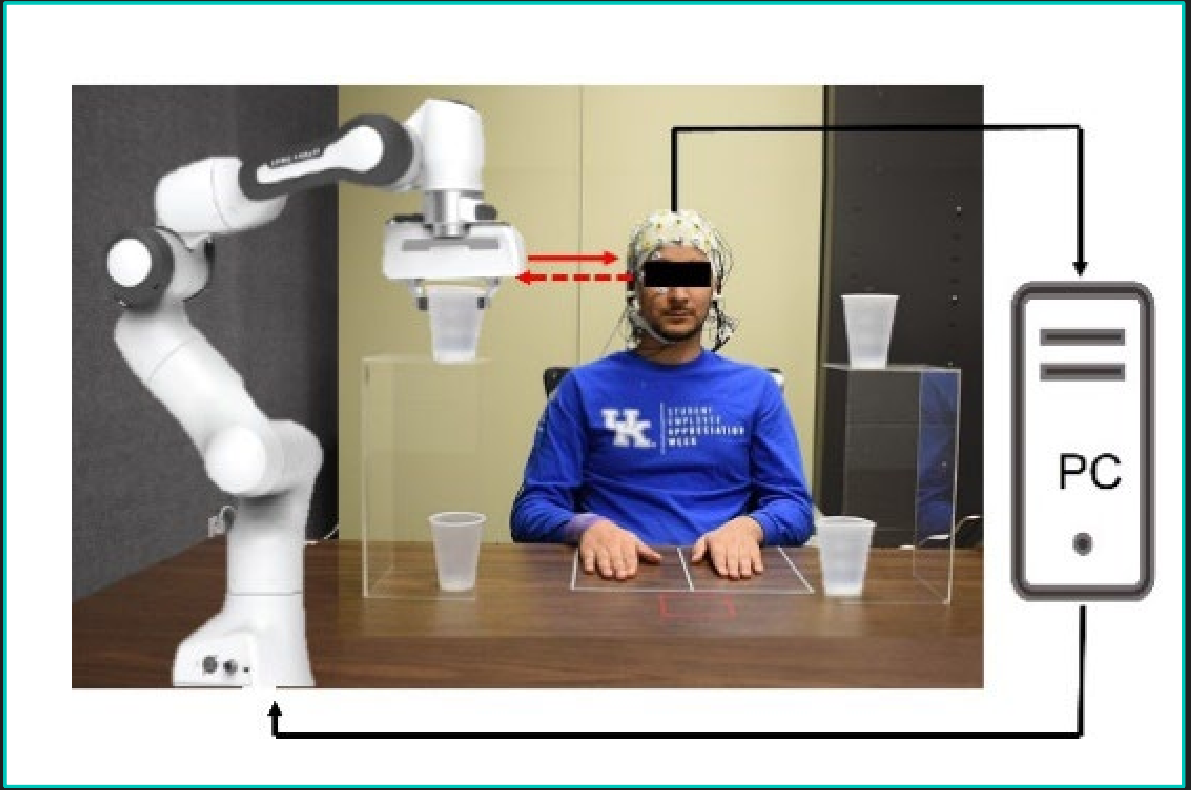
- Possible use of transfer learning approach has been investigated by transferring learned model to different days
- Evidence for the potential use of human scalp EEG in RLBMIs have been reported (2 open loop and 1 closed loop experiments)
- Further evaluation of different RL models and functional approximation approaches need to be conducted

Future Directions
in the Neural Interfaces Lab at UKY

- Confirmed applicability of Q-KTD (λ) on EEG-based RLBMI (open loop)
- On-going recording of human scalp EEG
- Future Work:

  Integration of robotic arm

  for closed loop evaluation



B. R. Thapa, D. Restrepo Tangarife, and J. Bae. "Kernel Temporal Differences for EEG-based Reinforcement Learning Brain Machine Interfaces." Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), 2022, pp.3327-3333, doi: 10.1109/EMBC48229.2022.9871862.

Thank You!

Questions?