Is Cerebellum a Model-Based Reinforcement Learning Agent?

Bharath Masetty[1,2], Reuth Mirsky[1], Ashish D. Deshpande[2]
Michael Mauk[3] and Peter Stone[1,4]

ReNeu Robotics Lab
Rehabilitation and Neuromuscular Robotics

LARG
Learning Agents Research Group
The University of Texas at Austin

[1] Learning Agents Research Group, Computer Science Department, The University of Texas at Austin
[2] Rehabilitation and Neuromuscular Robotics Lab, Mechanical Engineering Department, The University of Texas at Austin
[3] Mauklab, Center for Learning and Memory, Department of Neuroscience, The University of Texas at Austin
[4] Sony AI

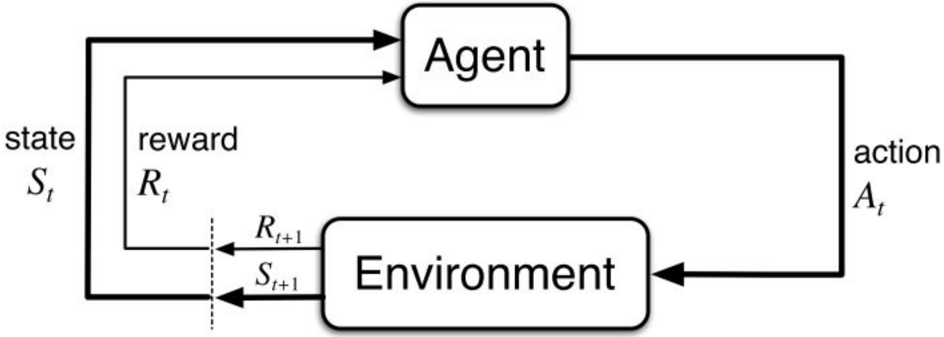
Motivation

- Review cerebellum's functionality from a reinforcement learning perspective
- Propose novel experiments using cerebellum simulations and ideas in RL to understand more about cognitive and motor learning in humans.

Background: Cerebellum

- Cerebellum is a major structures of the brain located near brainstem
- 10 % of brain's volume but has more neurons than the rest of the brain
- Neural substrate responsible for movement coordination and motor control
- Consists of functional subdivisions called *microzones* which modulate activity in specific muscle groups

Background: Reinforcement Learning



Markov Decision Process (**S**,**A**, ***p***, ***r***, **γ**):

- **S :** State space
- **A** : Action space
- Transition Function ***P*(*s'* | *s*, *a*)** : Probability of being in state *s'* when taken action *a* in state *s*.
- Reward Function ***r*(*s*, *a*)**: Determines reward *r* when taken action *a* in state *s*.
- **γ :** Discount factor

Model-Based Methods: Uses models of the environment to optimize the policy.
Model-Free Methods: Do not use models of the environment to optimize the policy.

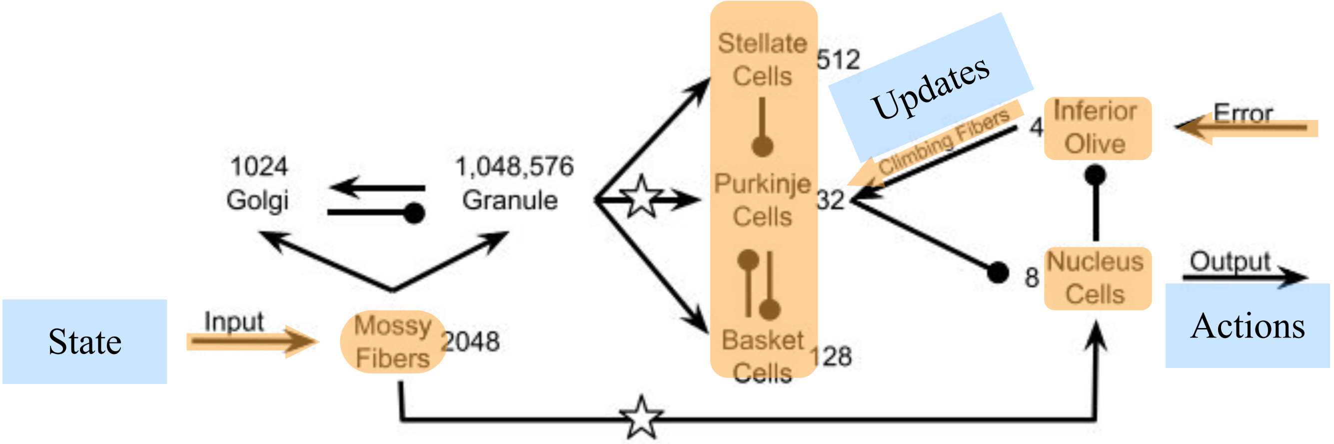
Hypothesis

Cerebellum Functionality		Process in RL
Ability to modulate motor commands and control movement	≡	Modulation of a Control Policy
Ability to predict future sensory states	≡	Learning a forward dynamics model
Ability to encode external rewards	≡	Learning a reward function

Model-based reinforcement learning ***can*** be one of the functionalities of cerebellum

- Literature supporting the hypothesis
- Propose a way to test the hypothesis using a cerebellum simulation

Cerebellum: Topology

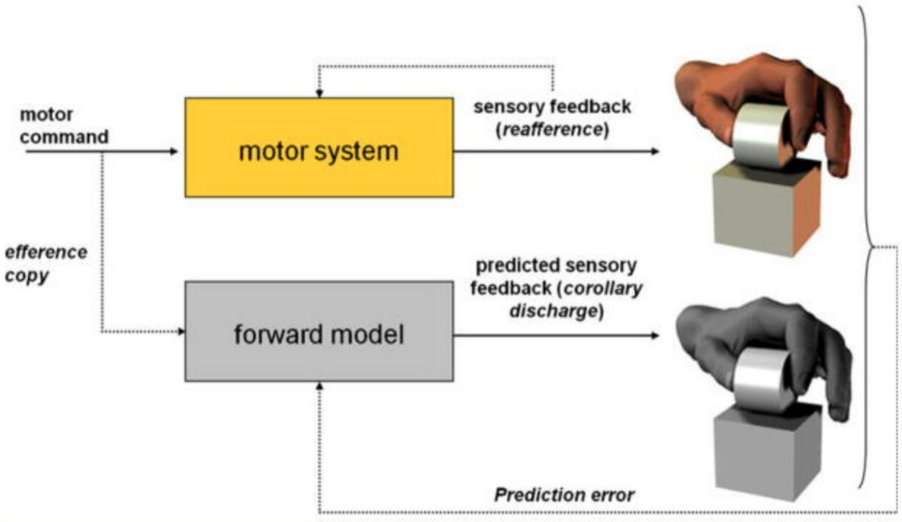


- *Mossy Fibers* are the primary inputs to the cerebellum (CB)
- *Nucleus Cells* provide the primary output
- *Climbing Fibers* delivers error signals to modulate the synaptic plasticity of intermediate layers
- Divided into functional subdivisions called *microzones* controlling specific muscle groups

Cerebellum: Forward Models

- The cerebellum controls motor commands using prediction of future sensory states via its *internal forward models*
- Sensory prediction error acts as a training signal to learn the internal forward models

Internal Models in Cerebellum[7]



[7] Mario Manto, James M Bower, Adriana Bastos Conforto, Jos´e M Delgado-Garc´ıa, Suzete Nasci-mento Farias Da Guarda, Marcus Gerwig, Christophe Habas, Nobuhiro Hagura, Richard B Ivry, PeterMari¨en, et al. Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement.The Cerebellum, 11(2):457–487, 2012

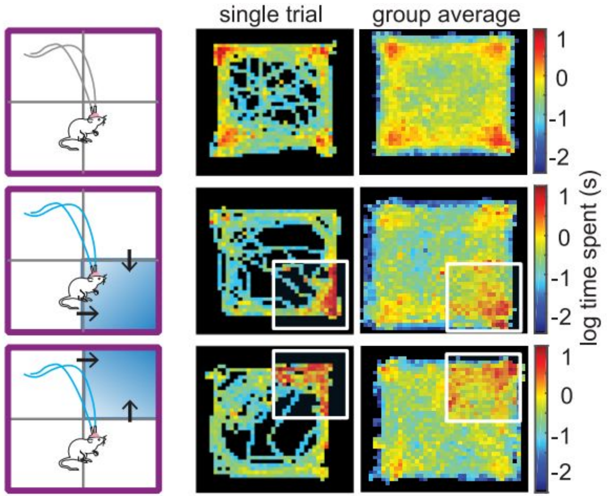
Hypothesis

Cerebellum Functionality		Process in RL
Ability to modulate motor commands and control movement	≡	Modulation of a Control Policy
Ability to predict future sensory states	≡	Learning a forward dynamics model
Ability to encode external rewards	≡	Learning a reward function

Model-based reinforcement learning *can* be one of the functionalities of cerebellum

Cerebellum: Reward Learning

- Recent experiments[8] revealed that the cerebellum has direct excitatory projections to the Ventral Tegmental Area (VTA)
- VTA also known as brain's rewarding center
- Experiments on rodents showed that CB-VTA projects can encode external reward functions



[8] Ilaria Carta, Christopher H Chen, Amanda L Schott, Schnaude Dorizan, and Kamran Khodakhah. 2019. Cerebellar modulation of the reward circuitry and social behavior.Science 363, 6424 (2019).

Hypothesis

- Ability to modulate motor commands and control movement ≡ Modulation of a Control Policy
- Ability to predict future sensory states ≡ Forward dynamics of environments
- Ability to learn external reward functions ≡ Reward function of environments

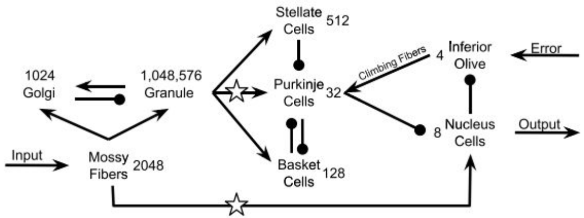
Model-based reinforcement learning *can* be one of the functionalities of cerebellum

How to test this hypothesis?

- We propose to use a simulated cerebellum for this purpose.

Simulated Cerebellum: Related Work

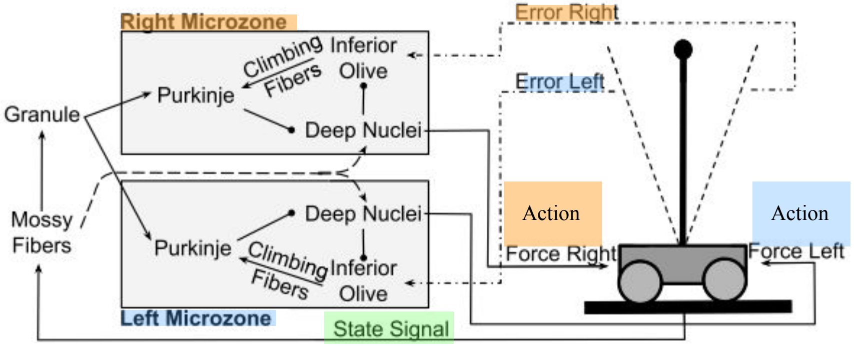
- Cerebellum's well understood topology makes it a good candidate for simulation neuroscience



- Adaptive cerebellar spiking model[9] to control robotic arm
- Cerebellum-inspired neural network[10] for state estimation and control
- Biologically constrained cerebellum[11] simulation was used to perform:
  - Cartpole balancing
  - PID Control
  - Robot Balancing using RL
  - Classification
  - Pattern Recognition

[9] Niceto R Luque, Jesús Alberto Garrido, Richard R Carrillo, Silvia Tolu, and Eduardo Ros. 2011. Adaptive cerebellar spiking model embedded in the control loop: Context switching and robustness against noise.International Journal of Neural Systems 21, 05 (2011), 385–401
[10] Christopher Assad, Sanjay Dastoor, Salomon Trujillo, and Ling Xu. 2005. Cerebellar dynamic state estimation for a biomorphic robot arm. In 2005 IEEE International Conference on Systems, Man and Cybernetics, Vol. 1. IEEE, 877–882
[11] Matthew Hausknecht, Wen-Ke Li, Michael Mauk, and Peter Stone. 2016. Machine learning capabilities of a simulated cerebellum.IEEE transactions on neural networks and learning systems 28, 3 (2016), 510–522.

Cartpole Interface



- The state signal is encoded into the ***mossy fibers***
- Each output comes from a distinct ***microzone***
- Every microzone is associated with its own error signal
- All microzones share a common input signal
- Outputs are inferred from the firing rates of nucleus neurons
- A ***pair*** of microzones for each class of output
  - One of them will increase the output for that class and the other will decrease

[11] Matthew Hausknecht, Wen-Ke Li, Michael Mauk, and Peter Stone. 2016. Machine learning capabilities of a simulated cerebellum.IEEE transactions on neural networks and learning systems 28, 3 (2016), 510–522.

Testing the Hypothesis

Model-based reinforcement learning could be one of the functionalities of cerebellum

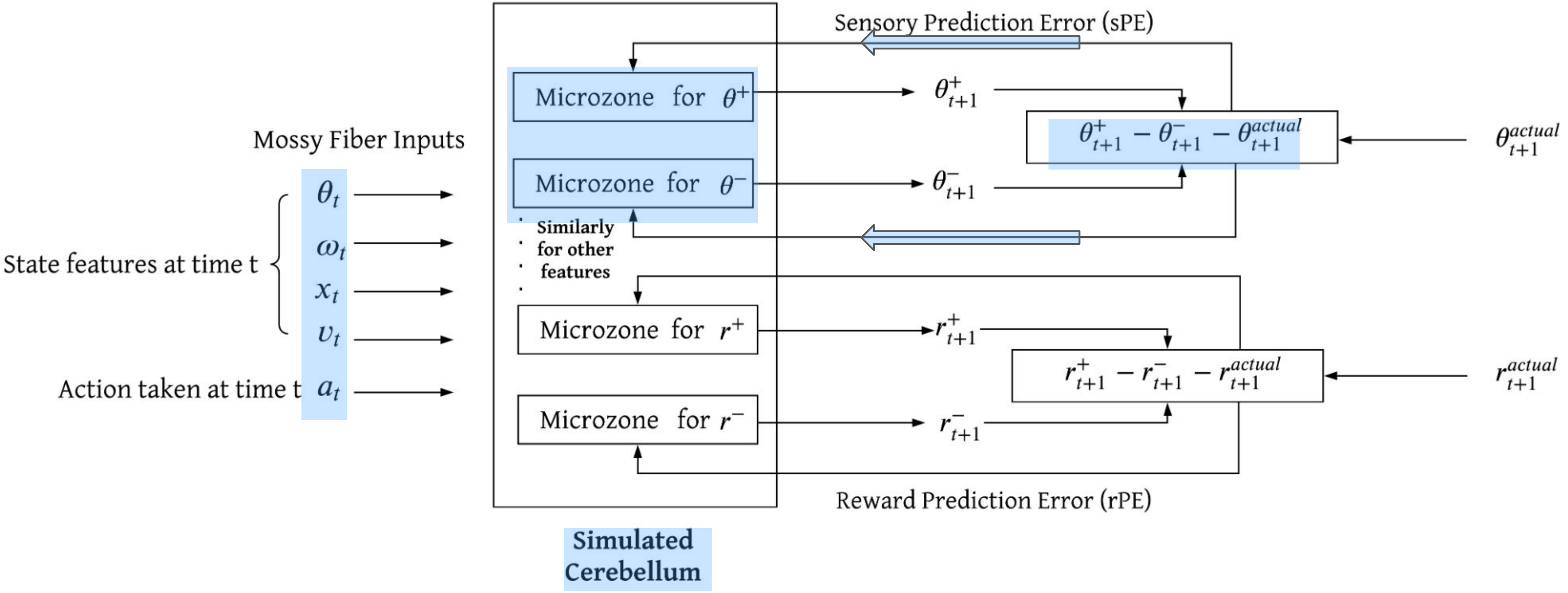
- The goal is to test if the simulated cerebellum ***can*** act as a reinforcement learning model.

We propose a two step method:

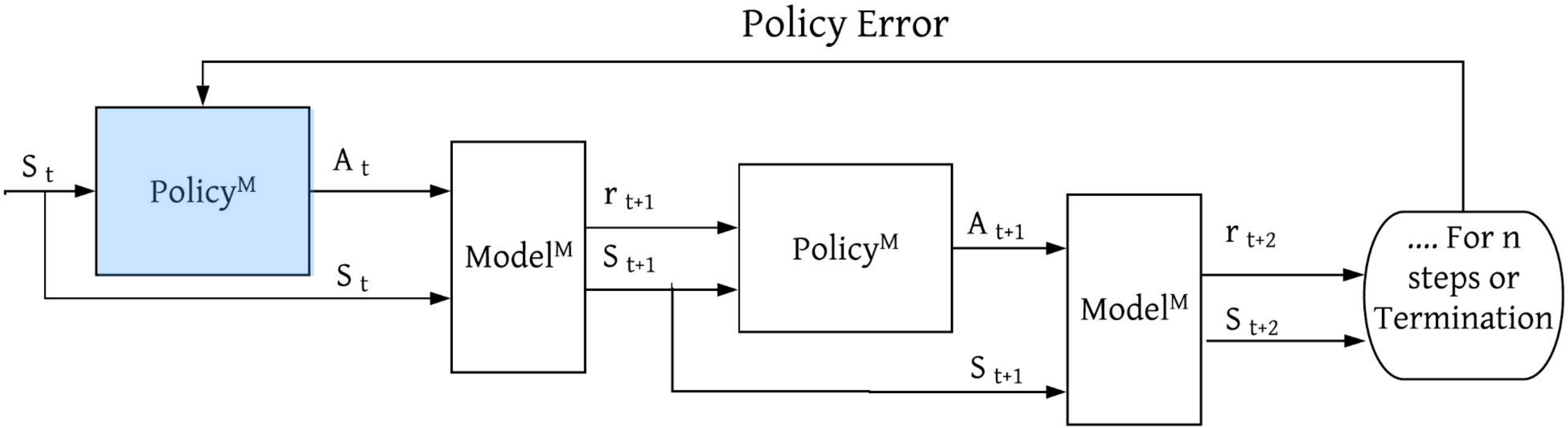
Step 1: Learn the forward dynamics model and reward function of the environment

Step 2: Perform an n-step look ahead using the learnt model to test for policy optimization.

Step 1: Model Learning



Step 2: Policy Learning



Potential Outcomes

- In step 2, if the simulation is able to show improvement in its policy
    - Cerebellum simulation ***can*** perform model-based RL
    - Supports our hypothesis

- If the simulation does not show any improvement in policy
    - Hypothesis is not true, or
    - Cannot be tested within the scope of simulated cerebellum or
    - Could be due to limitations of the simulation model

Summary

- We combine popular consensus with recent evidence to hypothesize that the cerebellum can perform model based reinforcement learning

- We propose a two-stage method to test this hypothesis using a simulated cerebellum
  - Learn the **forward dynamics** and **reward function** of the environment
  - Perform an **n-step look ahead** on policy microzones using model microzones to evaluate policy optimization

- **Potential challenges:**
  - Biological accuracy and level of abstraction in the cerebellum simulation
  - Hyperparameter tuning

- **Potential outcomes**:
  - Algorithmic understanding of reinforcement learning in cerebellum
  - Inspiration for new sample-efficient RL methods.

Is Cerebellum a Model-Based Reinforcement Learning Agent?

ReNeu Robotics Lab
Rehabilitation and Neuromuscular Robotics

LARG
Learning Agents Research Group
The University of Texas at Austin


Bharath Masetty
bmasetty@utexas.edu


Reuth Mirsky


Ashish D. Deshpande


Michael Mauk


Peter Stone