









# Neural circuit policies enabling auditable autonomy

Mathias Lechner <sup>1,4</sup> , Ramin Hasani <sup>2,3,4</sup> , Alexander Amini<sup>3</sup>, Thomas A. Henzinger <sup>1</sup>,  
Daniela Rus<sup>3</sup> and Radu Grosu <sup>2</sup>

2021. 06. 22.

Hyunsoo, Yu

# INDEX

## 1. Introduction

## 2. Methods

## 3. Results

1. Learning a compact neural representation
2. Avoiding crashes under increasing input perturbations
3. Robustness of the output decisions in the presence of input
4. Driving with smooth neural activity
5. Enhance interpretability

# INTRODUCTION

- This study designed a **brain-inspired intelligent agent** that learns to **control an autonomous vehicle** directly from its camera inputs.
- C.elegans have mastered such an ability to perform locomotion, motor control, and navigation with **near-optimal nervous system** structure (302 neural cells).
- This approach can lead to more expressive artificial intelligence agents with models that are **simultaneously accurate and explainable**.
- **Safety-critical domain** like controlling vehicles demands **interpretable dynamics**.
- **Neural Circuit Policies(NCP)** can be the key for these problem.

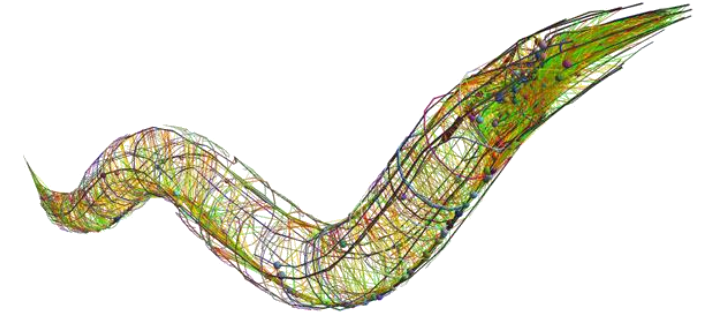


Lane tracking



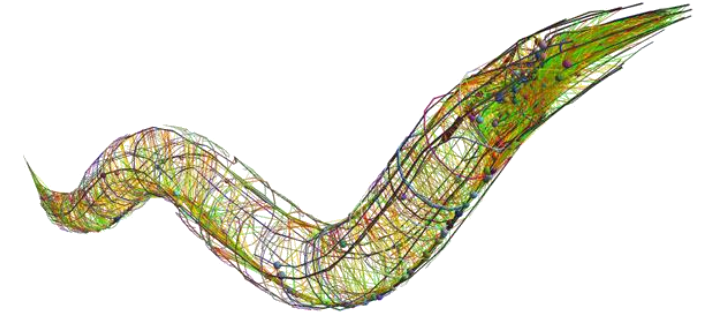
C.Elegans – fully analyzed its neural system<sup>3</sup>

# METHODS

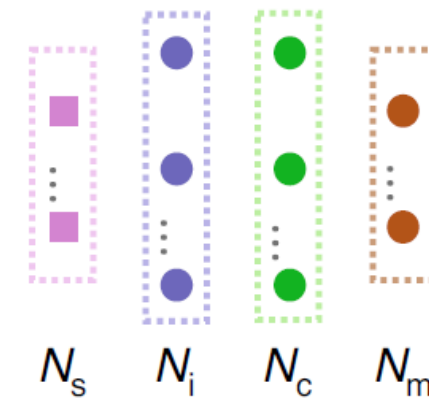


- NCP (Neural Circuit Policies) is inspired by C.elegans nematode
- Nematode's nervous system are constructed by a distinct 4-layer hierarchical network topology.
  - Sensory neurons – receive environmental observation
  - Inter-neurons & Command neurons – generate an output decision
  - Motor neurons – actuate muscles

# METHODS

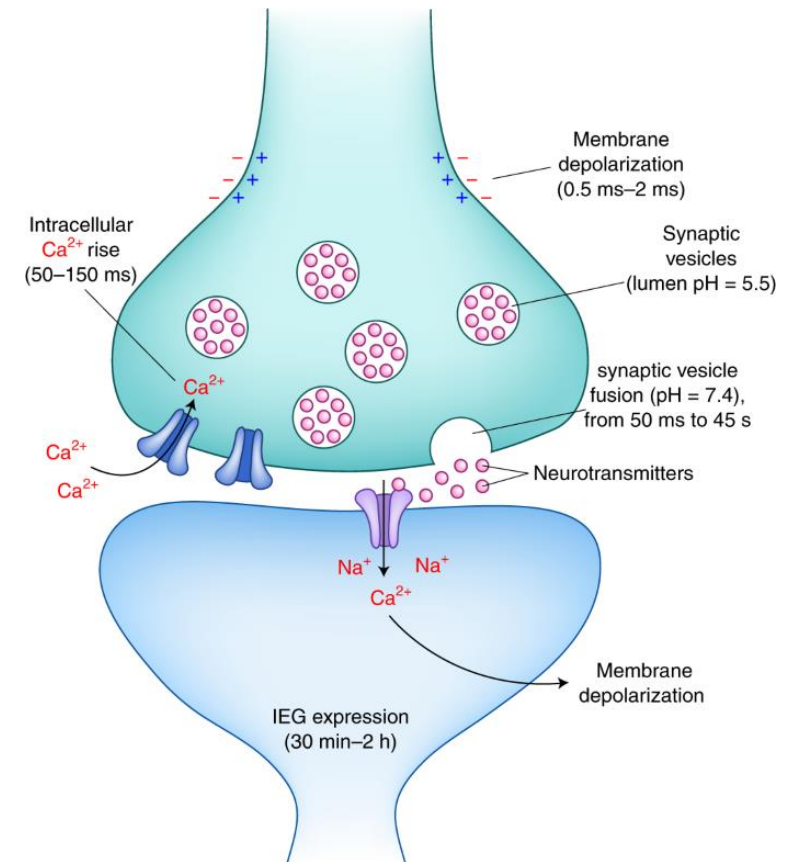
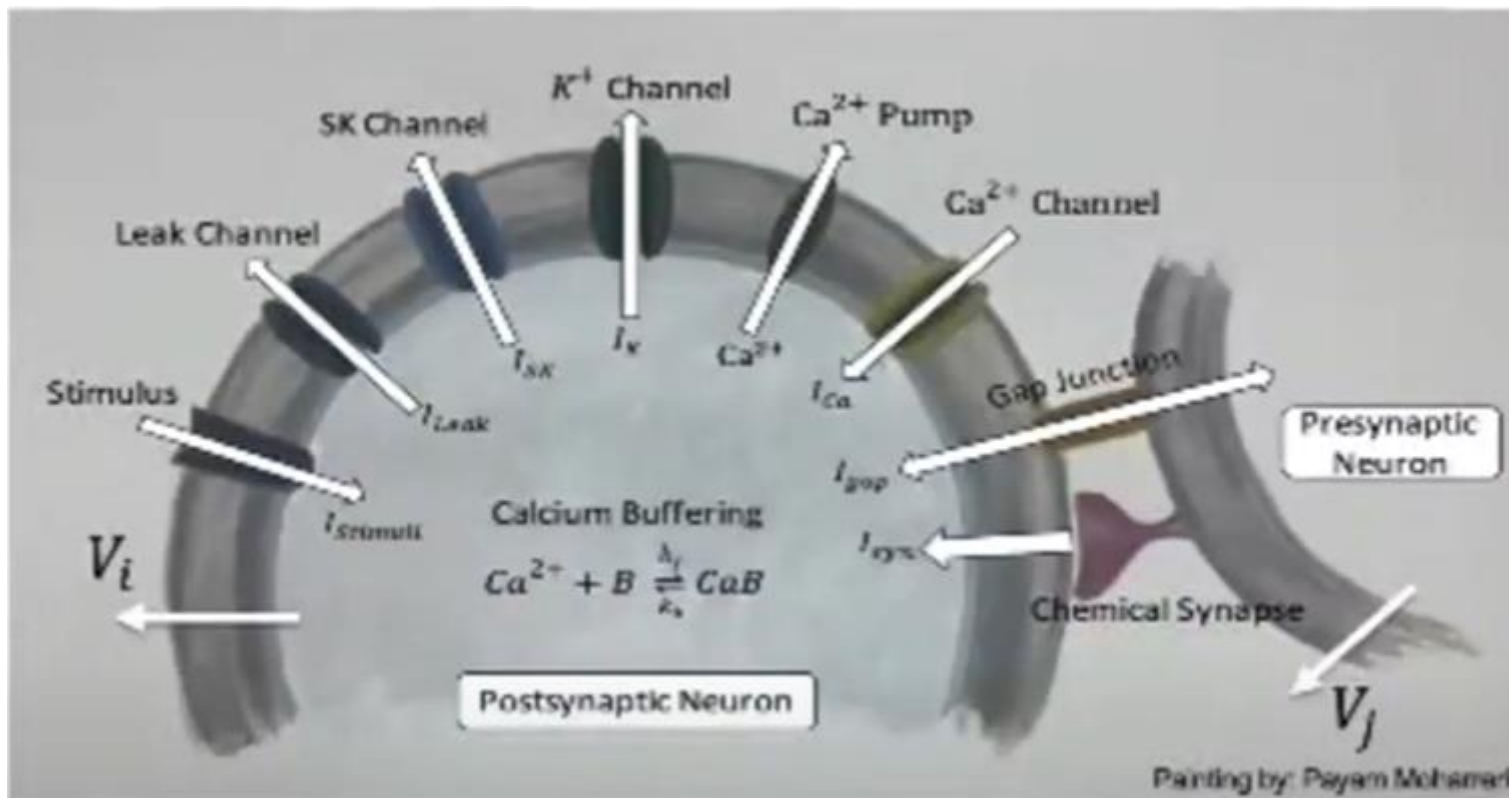


- NCP (Neural Circuit Policies) is inspired by C.elegans nematode
- Nematode's nervous system are constructed by a distinct 4-layer hierarchical network topology.
  - Sensory neurons – receive environmental observation
  - Inter-neurons & Command neurons – generate an output decision
  - Motor neurons – actuate muscles



■ Sensory      ■ Inter-neuron  
■ Command    ■ Motor

# METHODS



# METHODS

