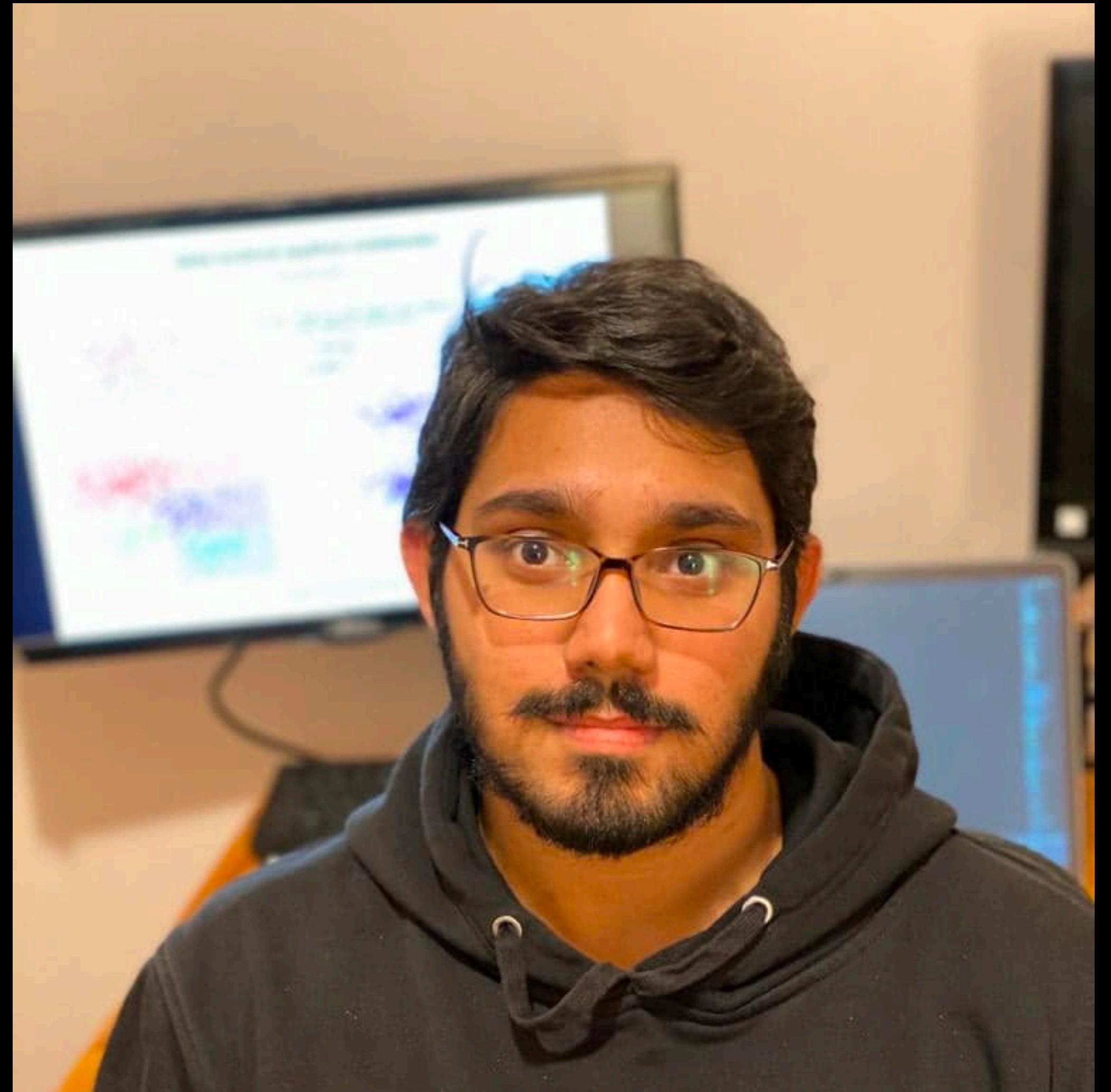


Who am I?

- Computer Scientist
 - Machine Learning Enthusiast
 - Blockchain Enthusiast
 - Summer Student
-
- CSC376, CSC413, CSC411

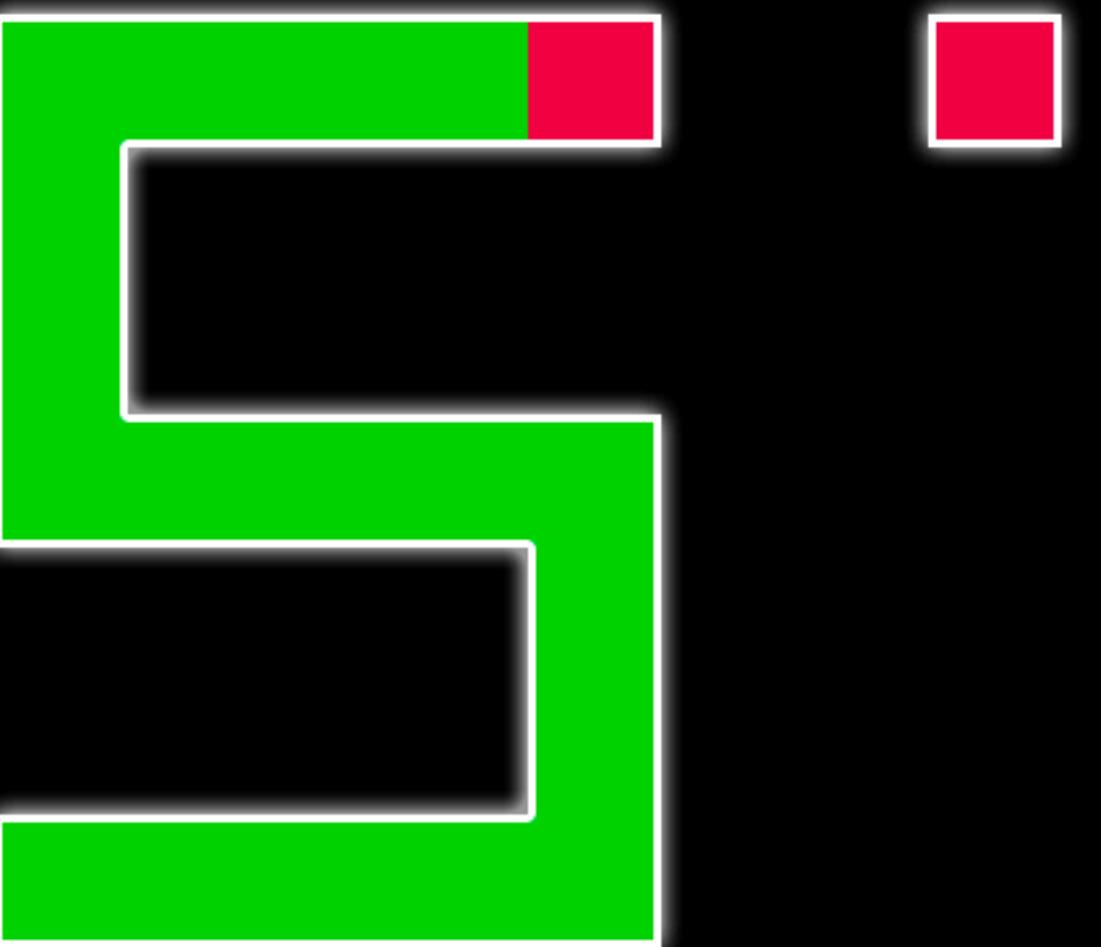


METABot Learning to Play

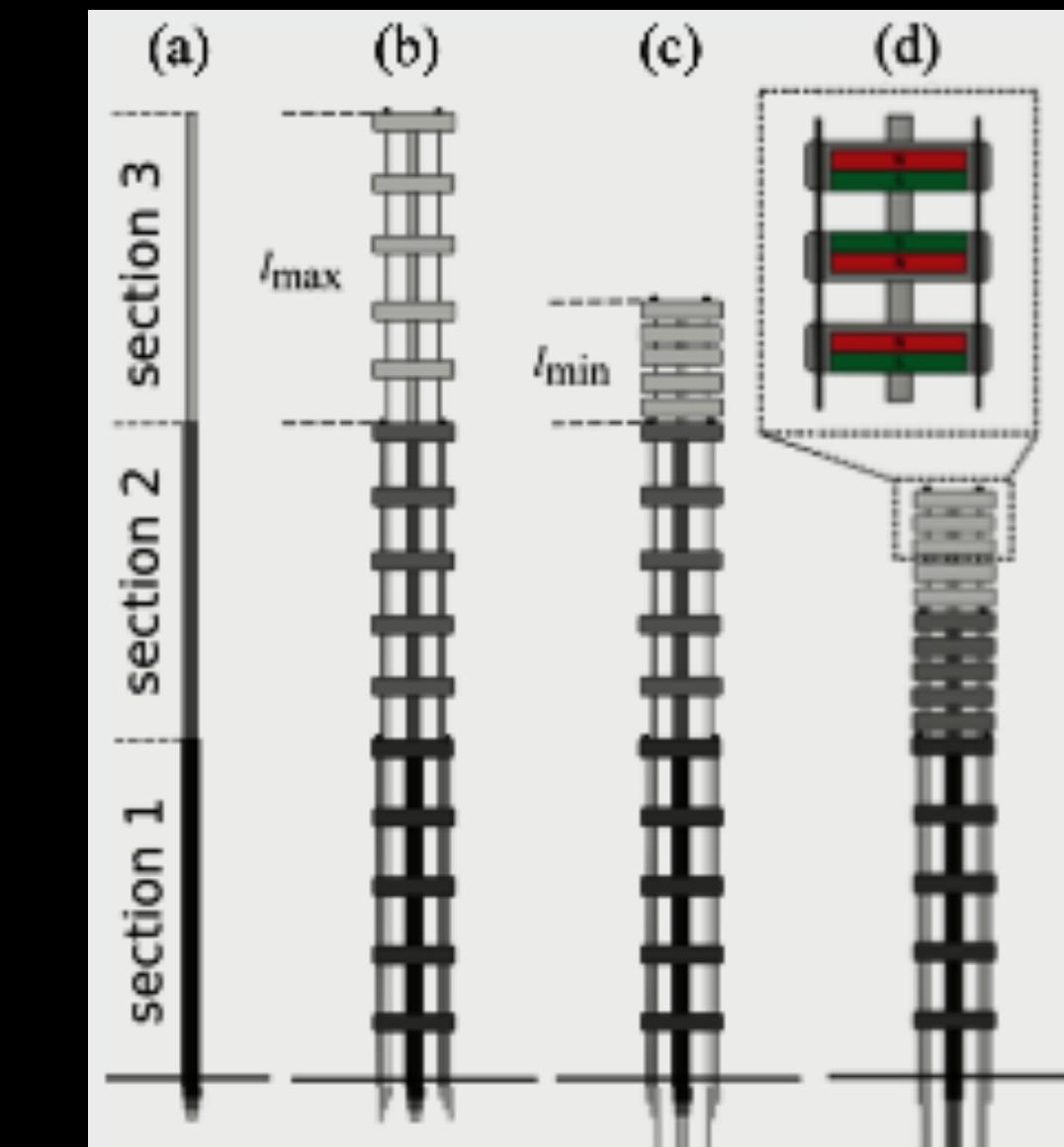
Reinforcement Learning Approach

Abdul Wasay Mehar

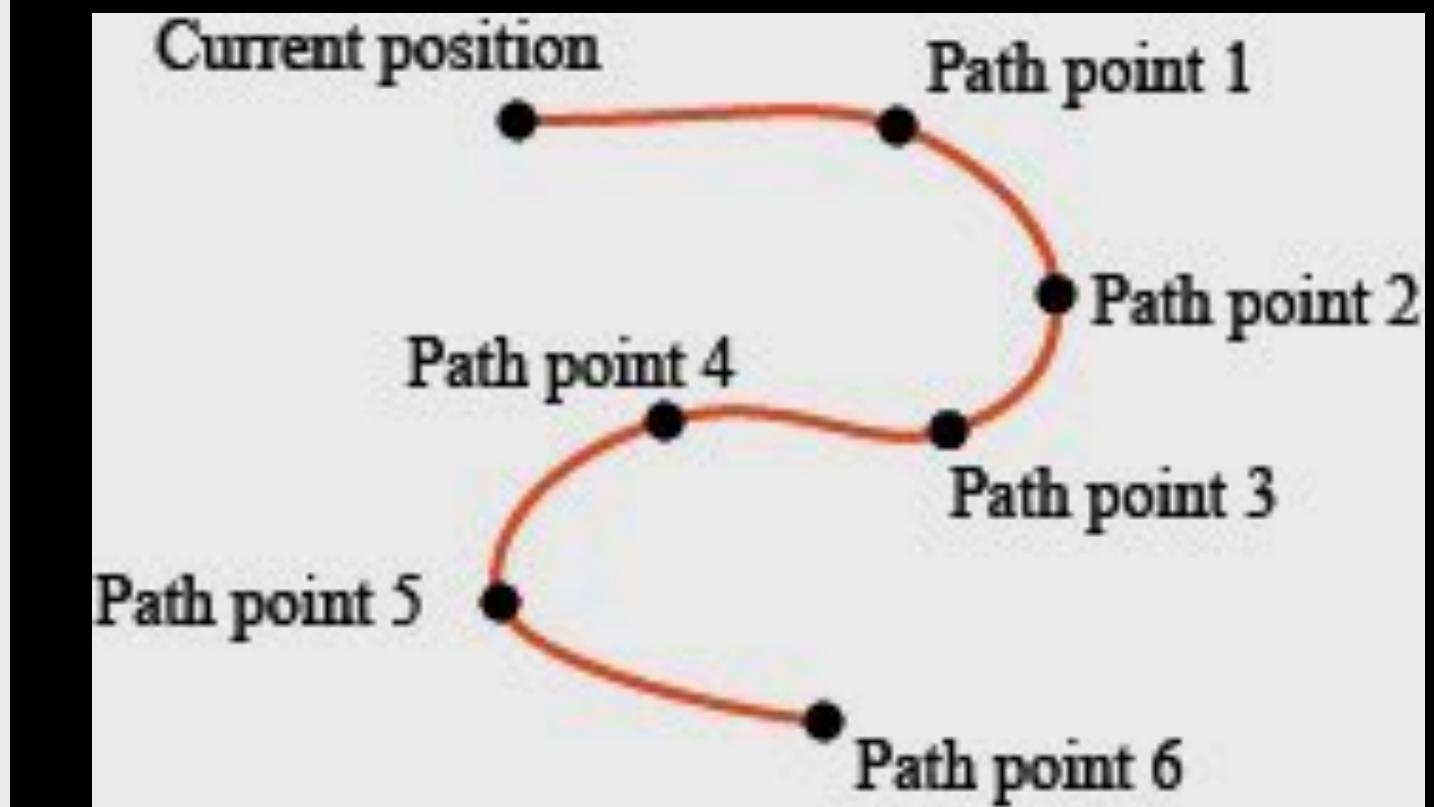
Supervision: Jessica Burgner-Kahrs & Reinhard Grassmann



METABot Learning to Play



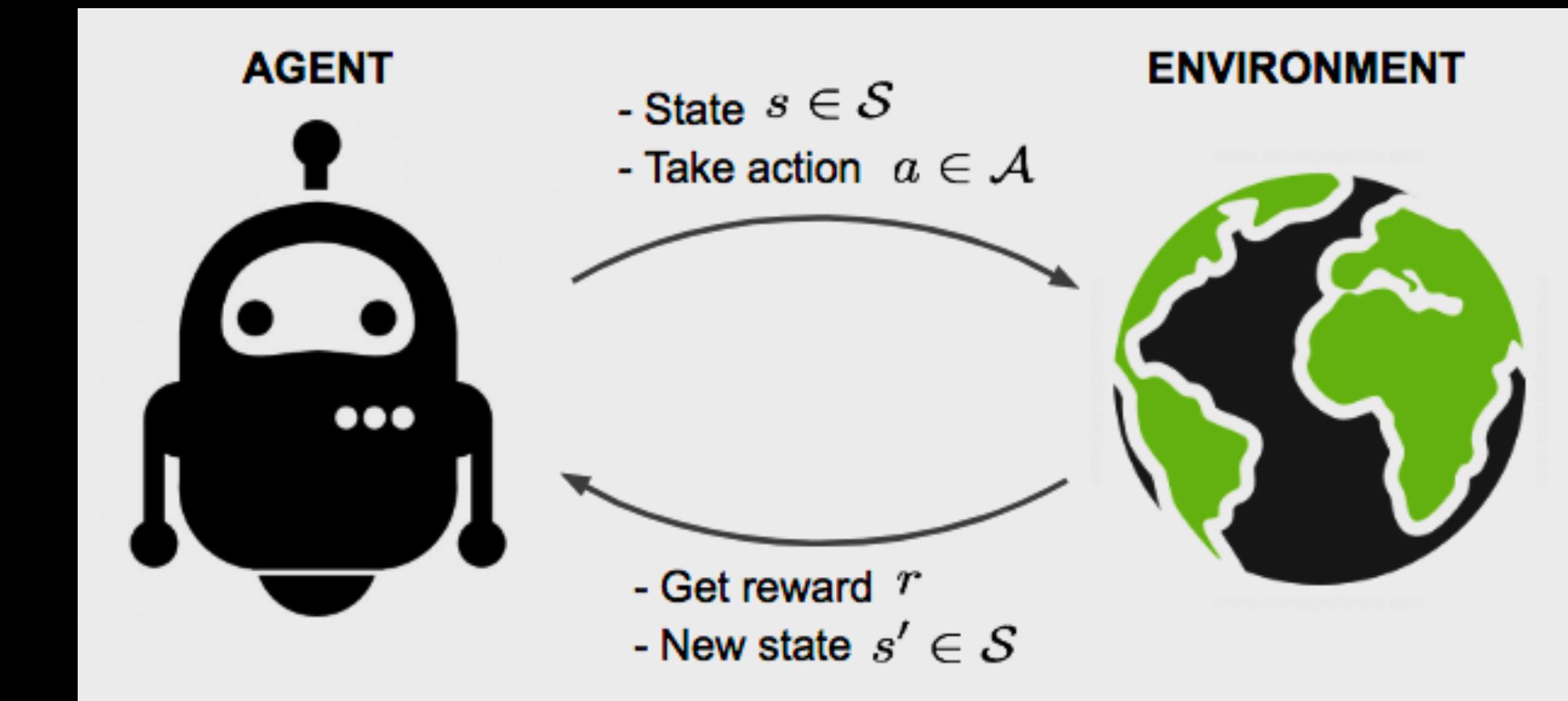
METABot



Follow-The-Leader

Three Main Components

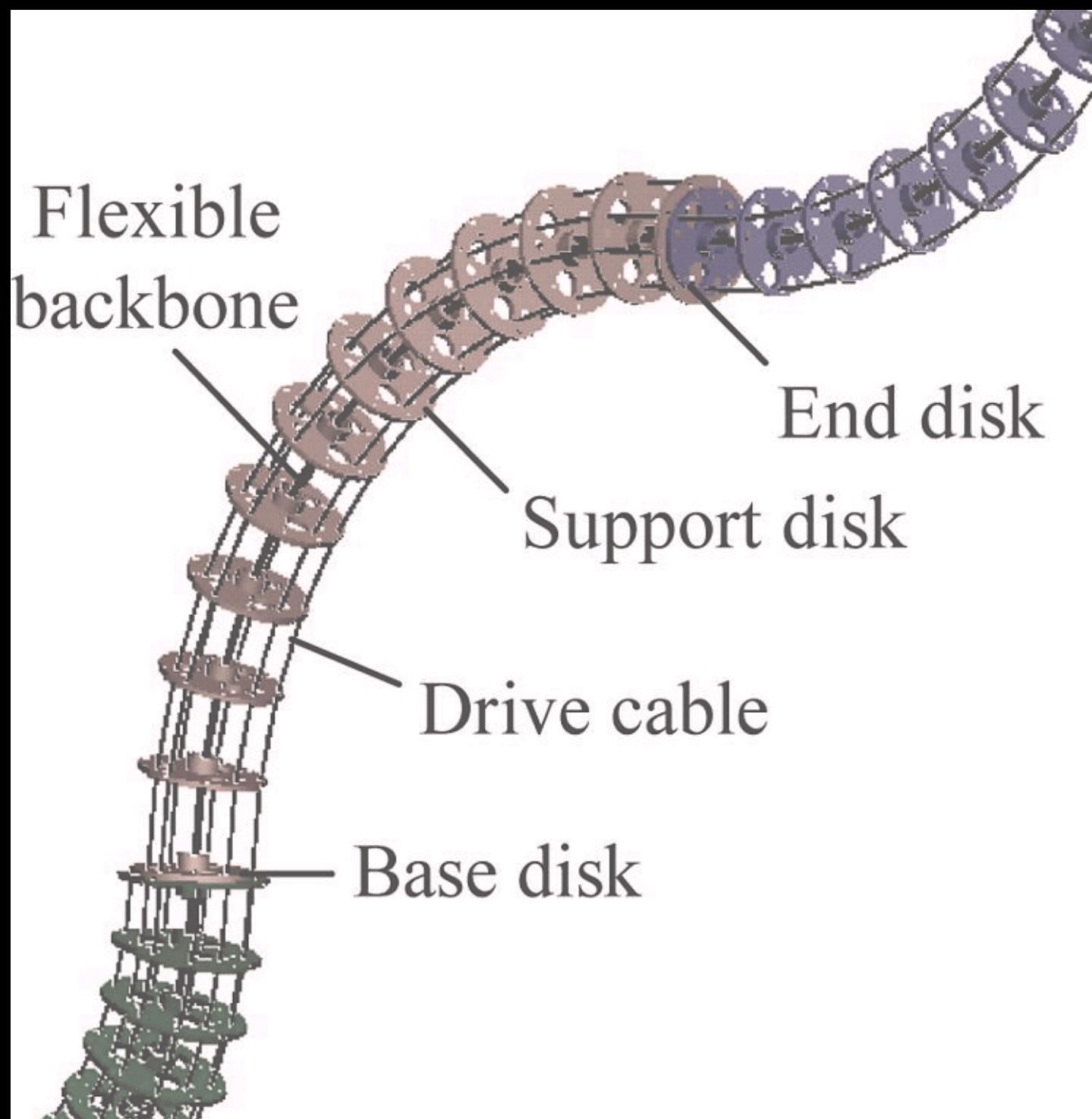
- METABot
- Follow-the-leader
- Reinforcement Learning



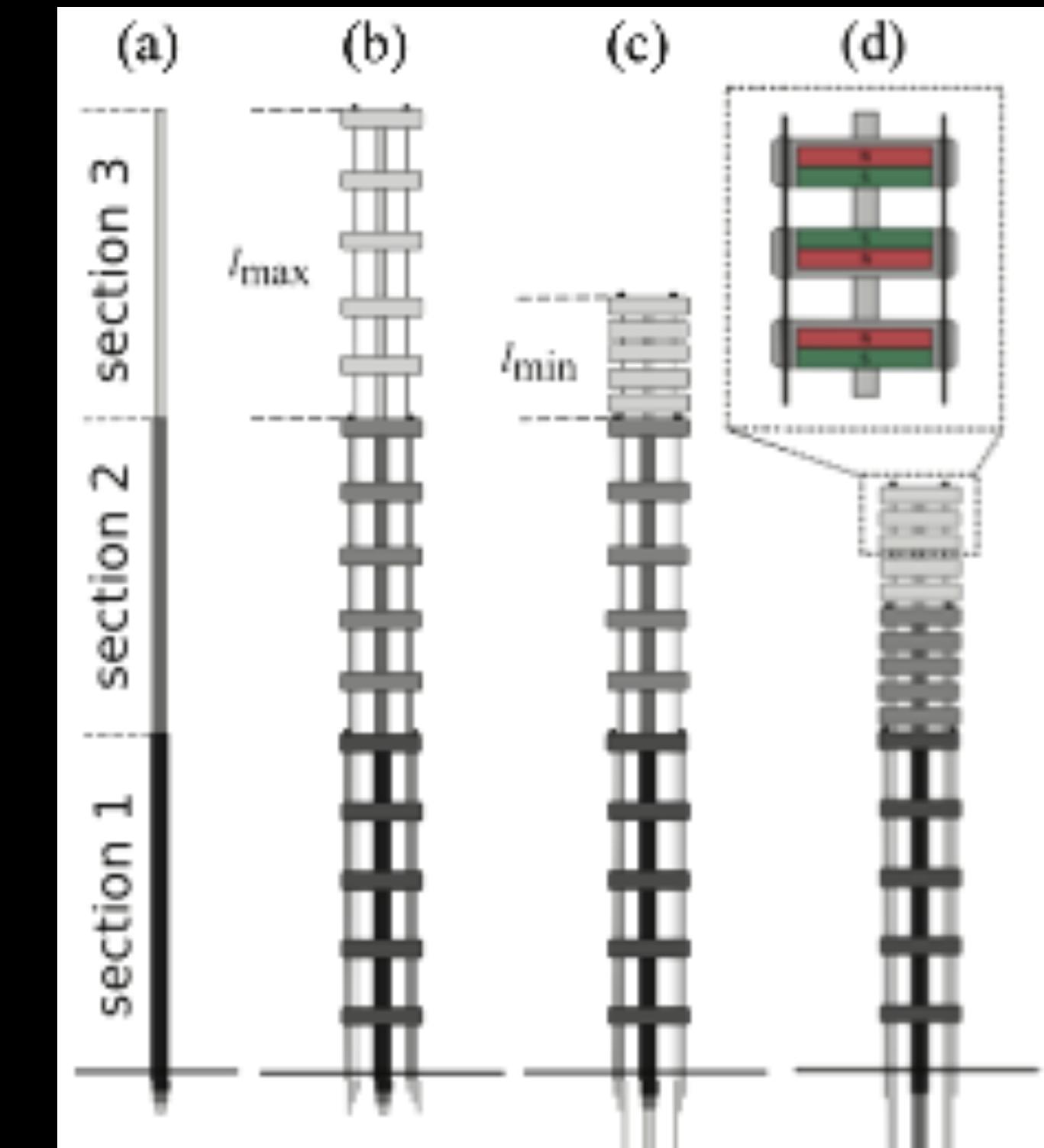
Goal: Autonomous Behaviour Mathematical Behaviour

Reinforcement Leraning

METABot



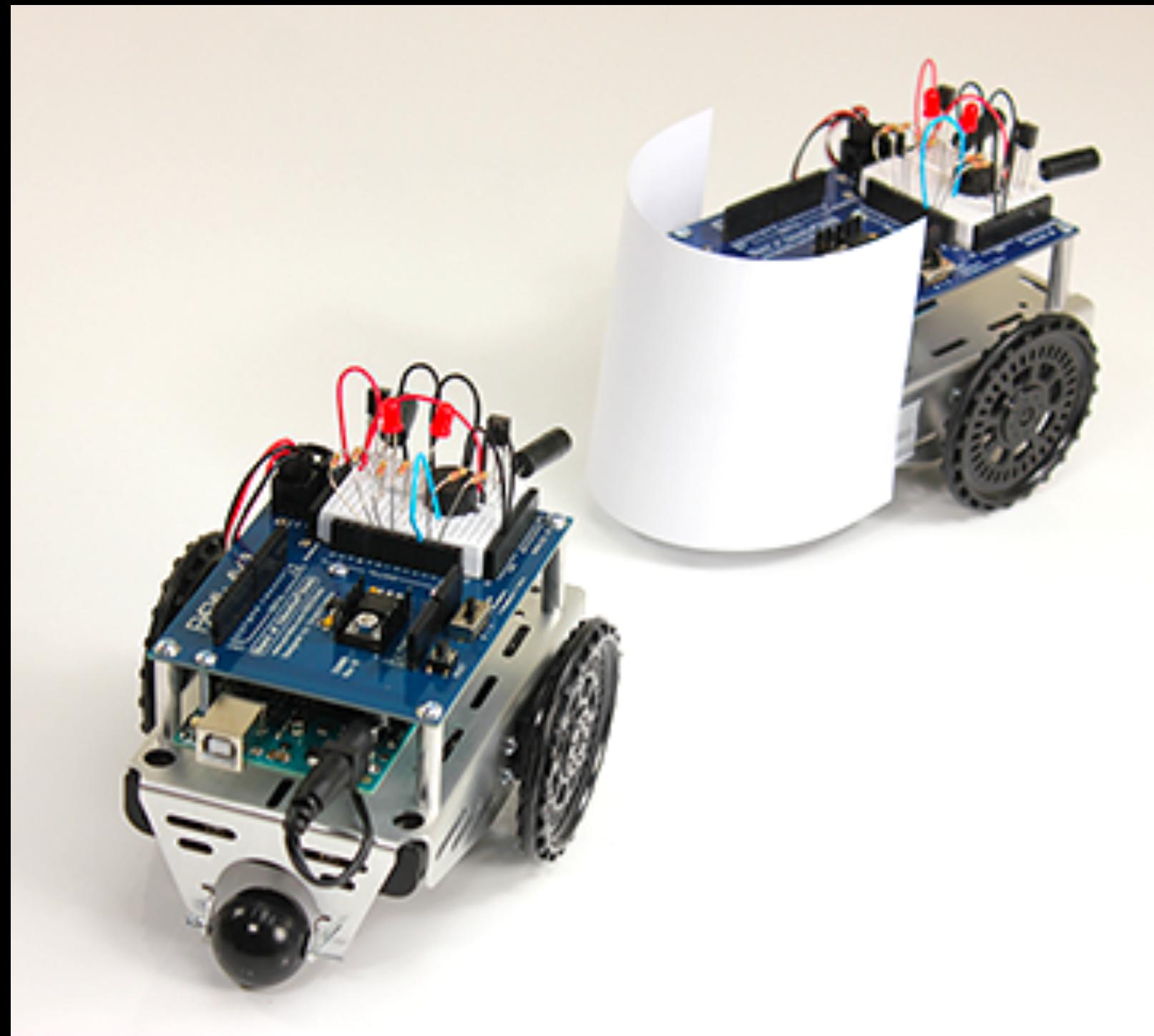
Standart Tendon Driven Robot



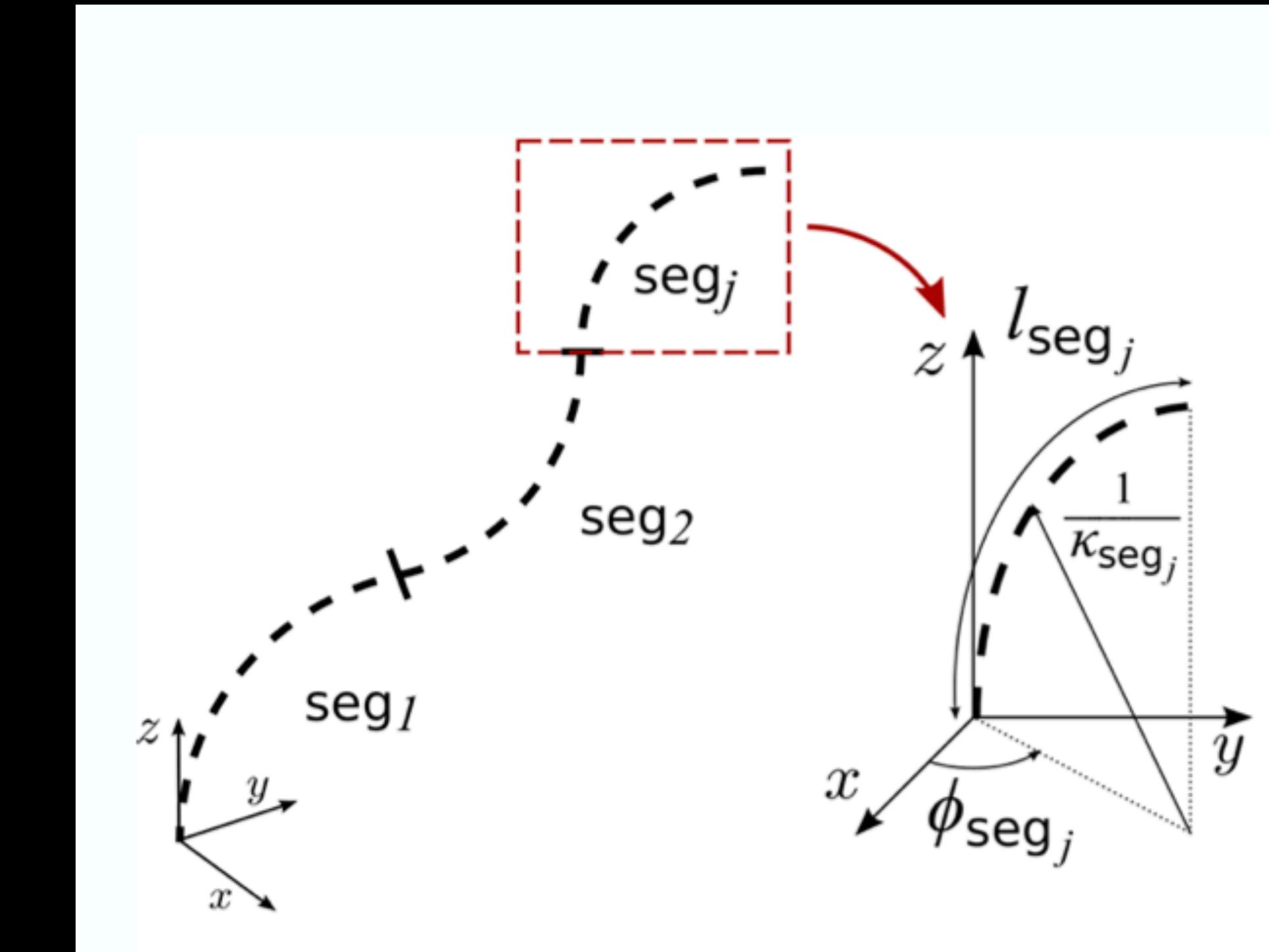
METABot

Follow The Leader

- State of the Art

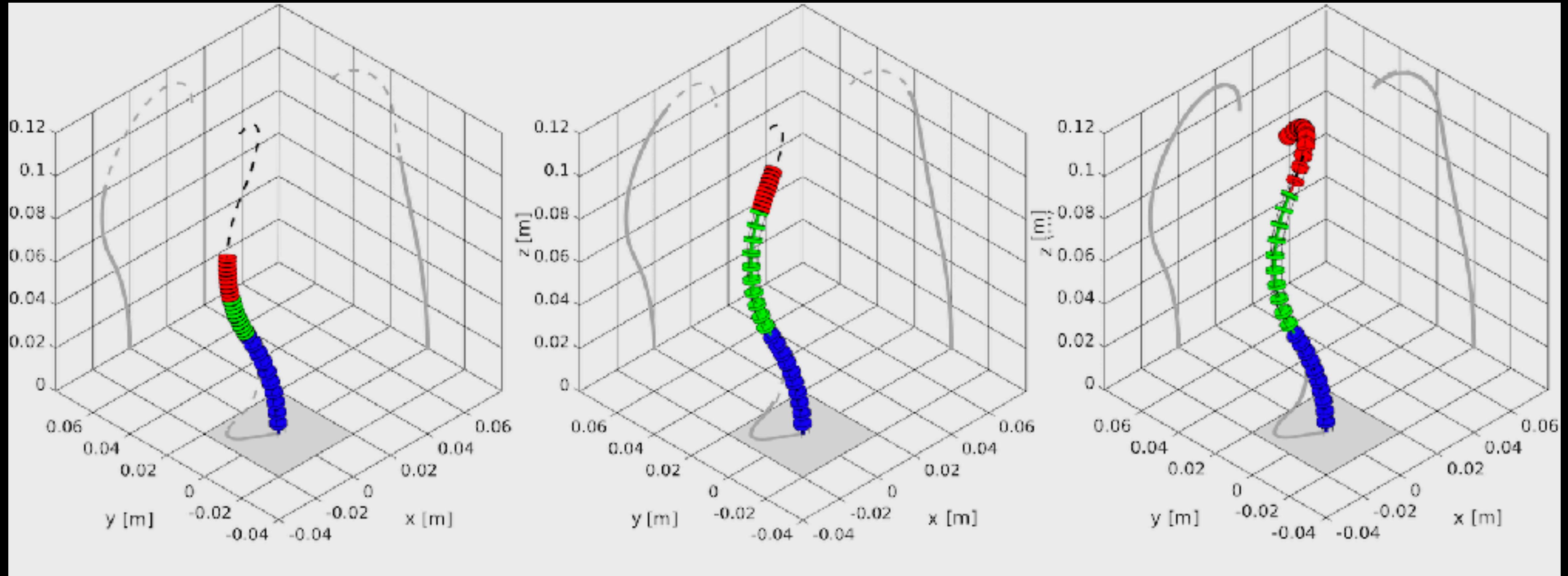


Boe Bots



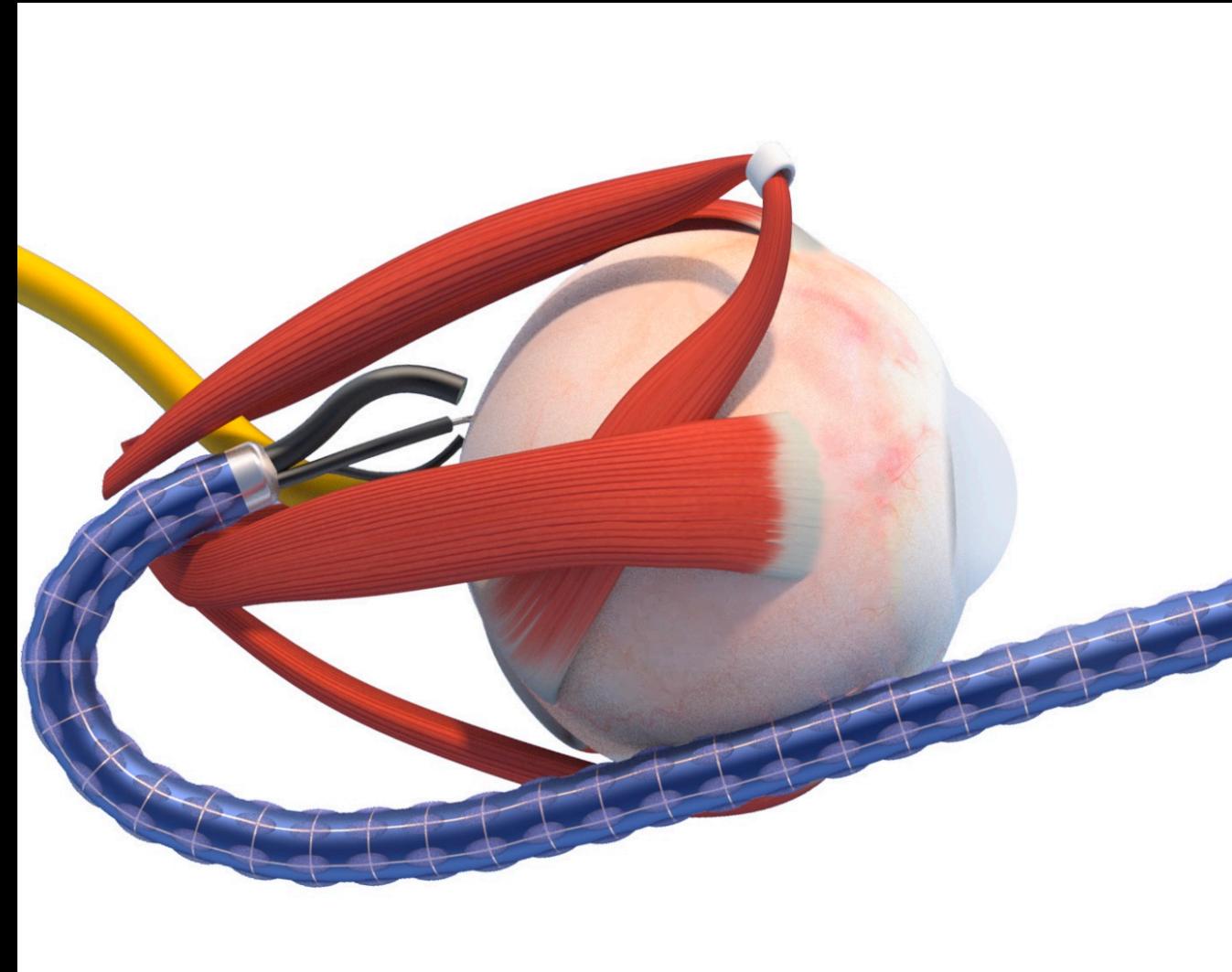
Sequential Section Strategy

Follow the leader



Follow the Leader Motion Basics

Follow the leader



Surgical

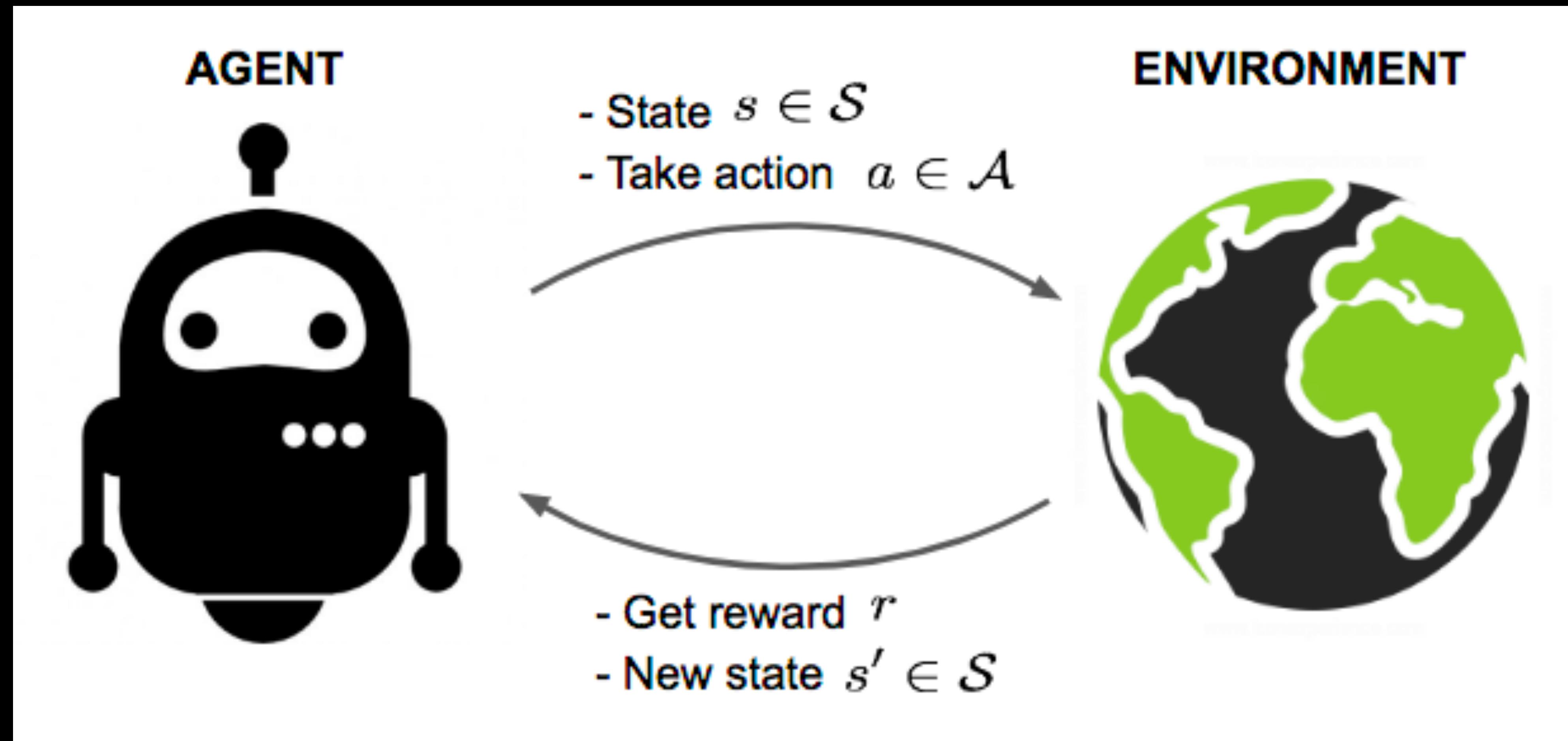


Engine Inspection

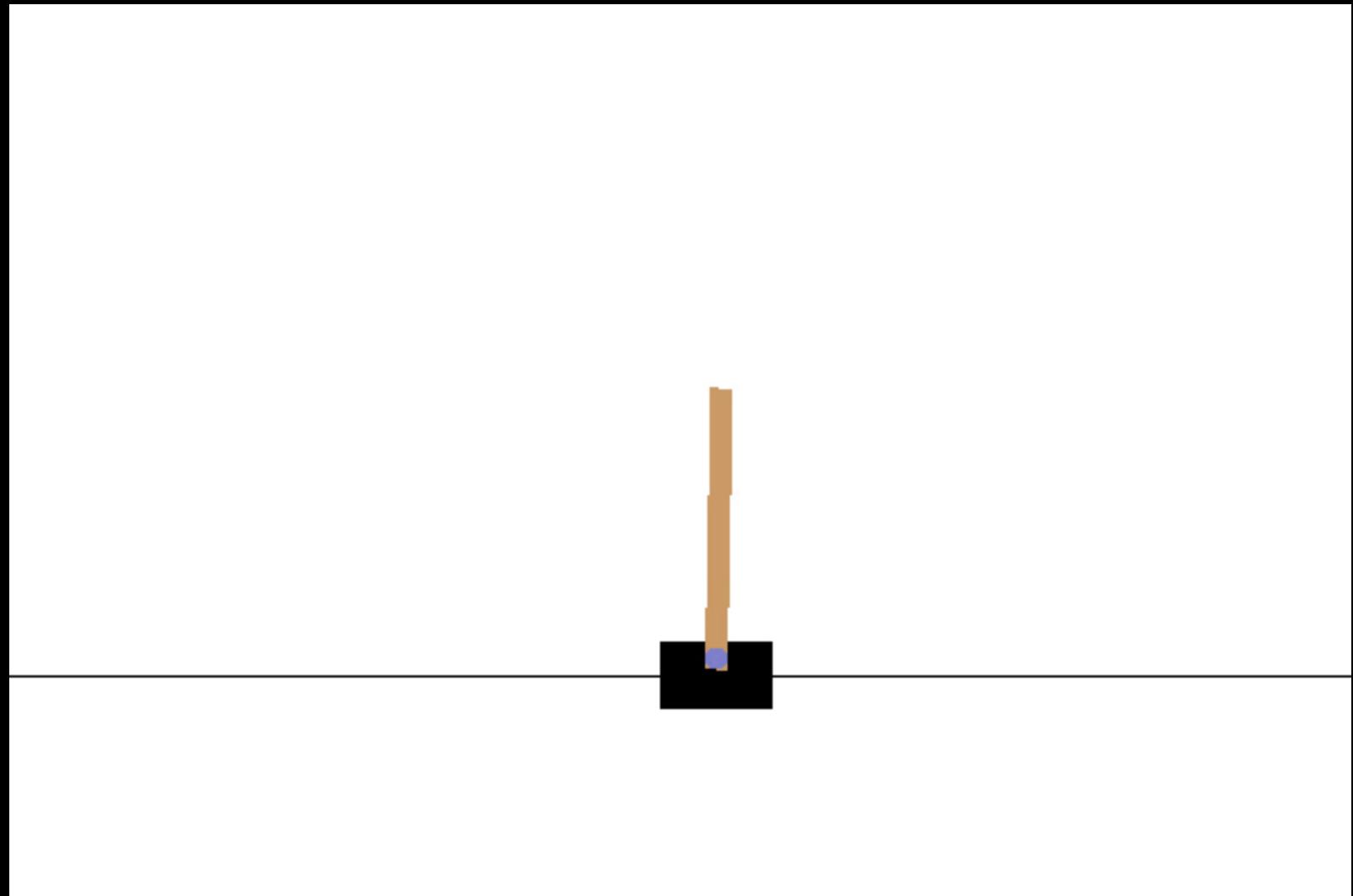


Unknown Environments

Reinforcement Learning



Reinforcement Learning



CartPole

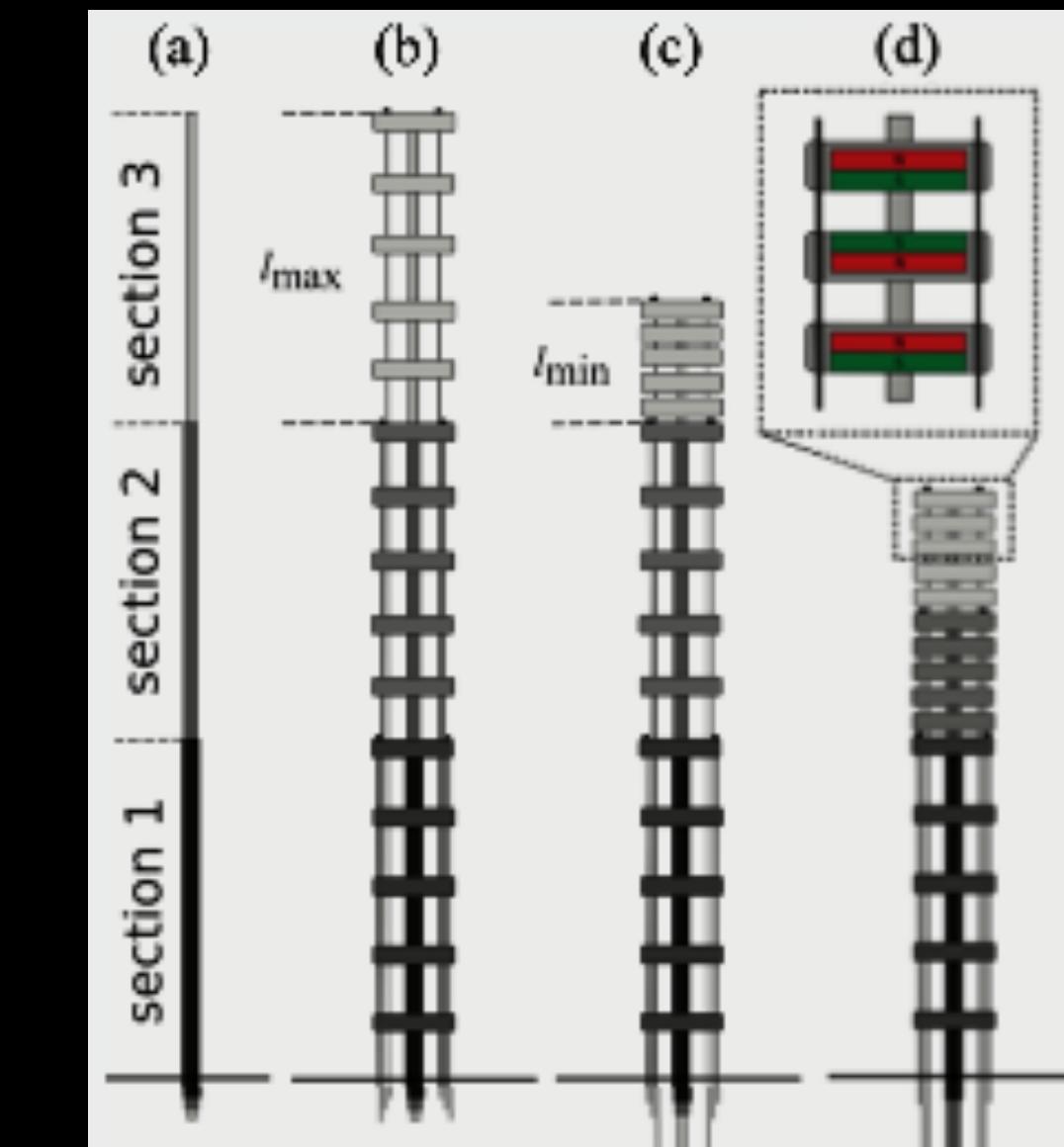


Atari Breakout

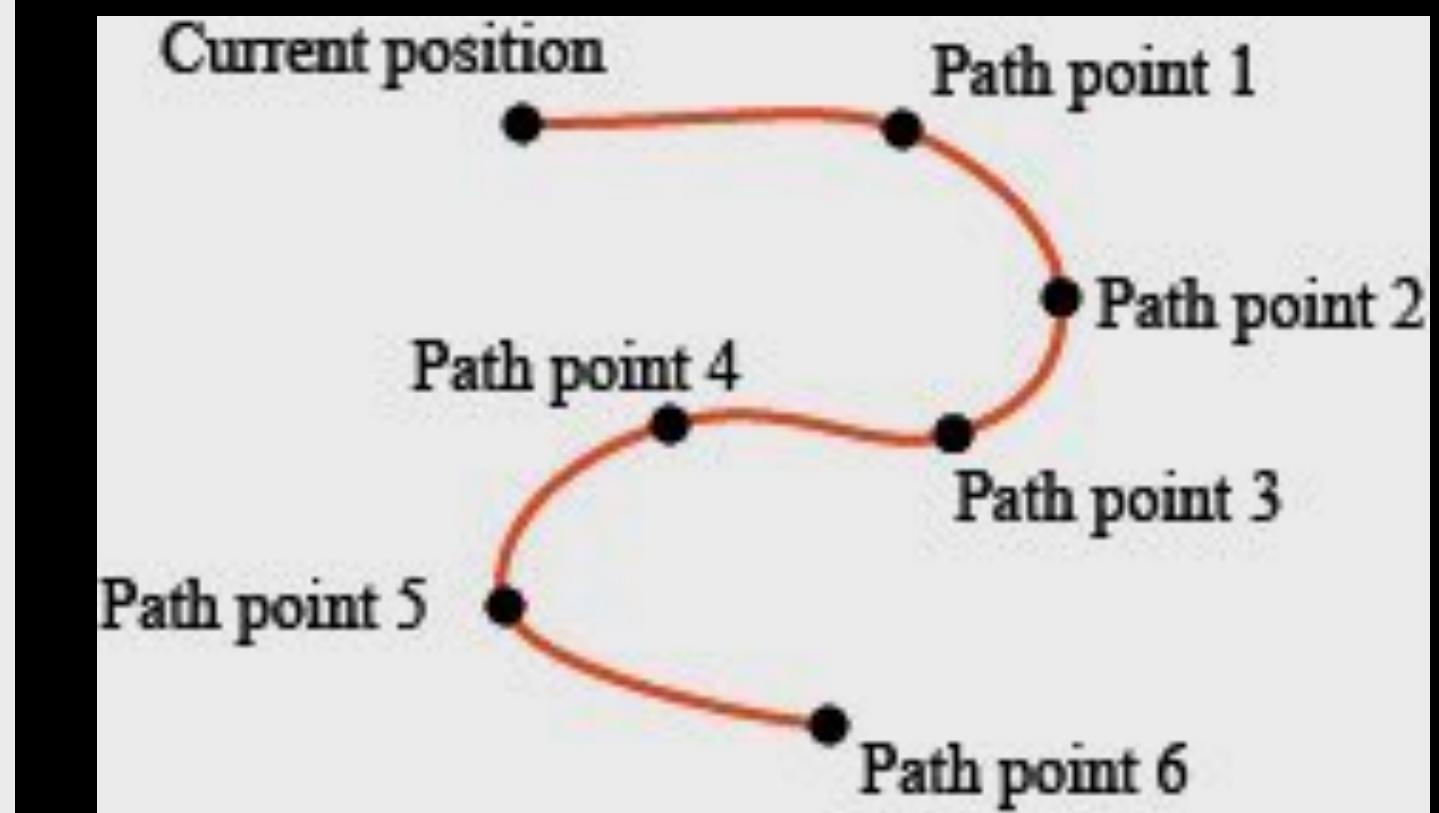


Walking Dummy

METABot Learning to Play



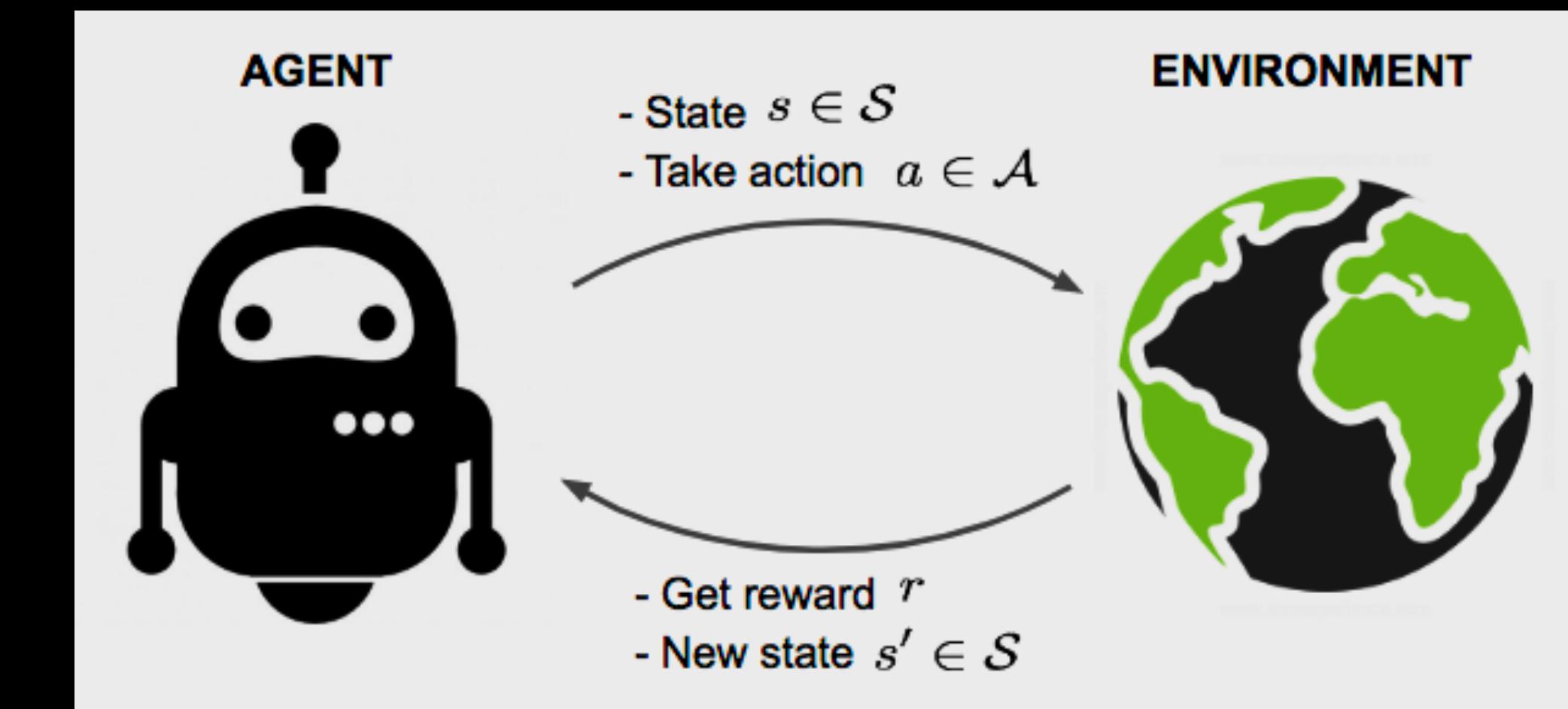
METABot



Follow-The-Leader

Three Main Components

- METABot
- Follow-the-leader
- Reinforcement Learning



Reinforcement Leraning

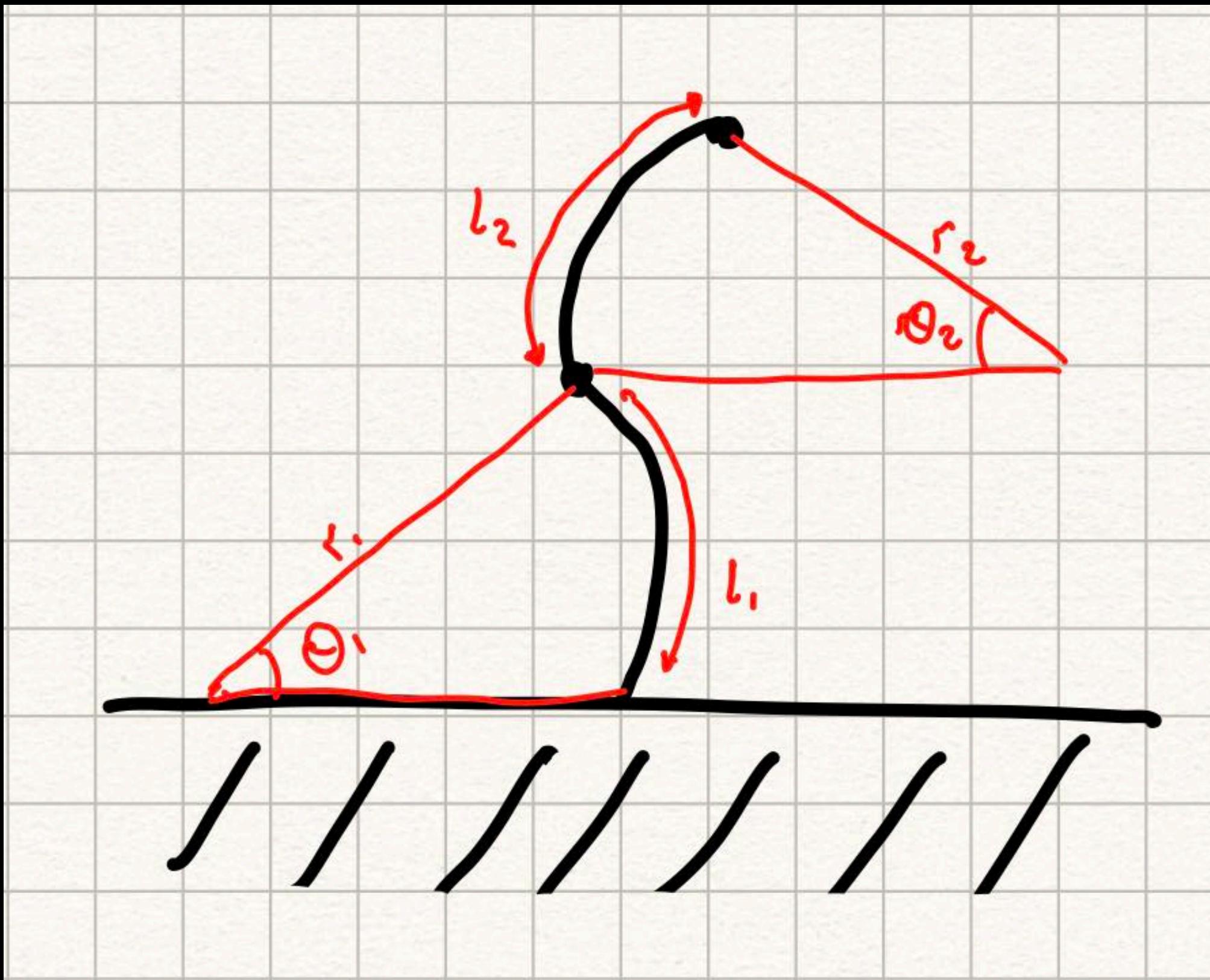
Project Approach

Reinforcement Learning - Environment



Project Approach

Reinforcement Learning - Agent

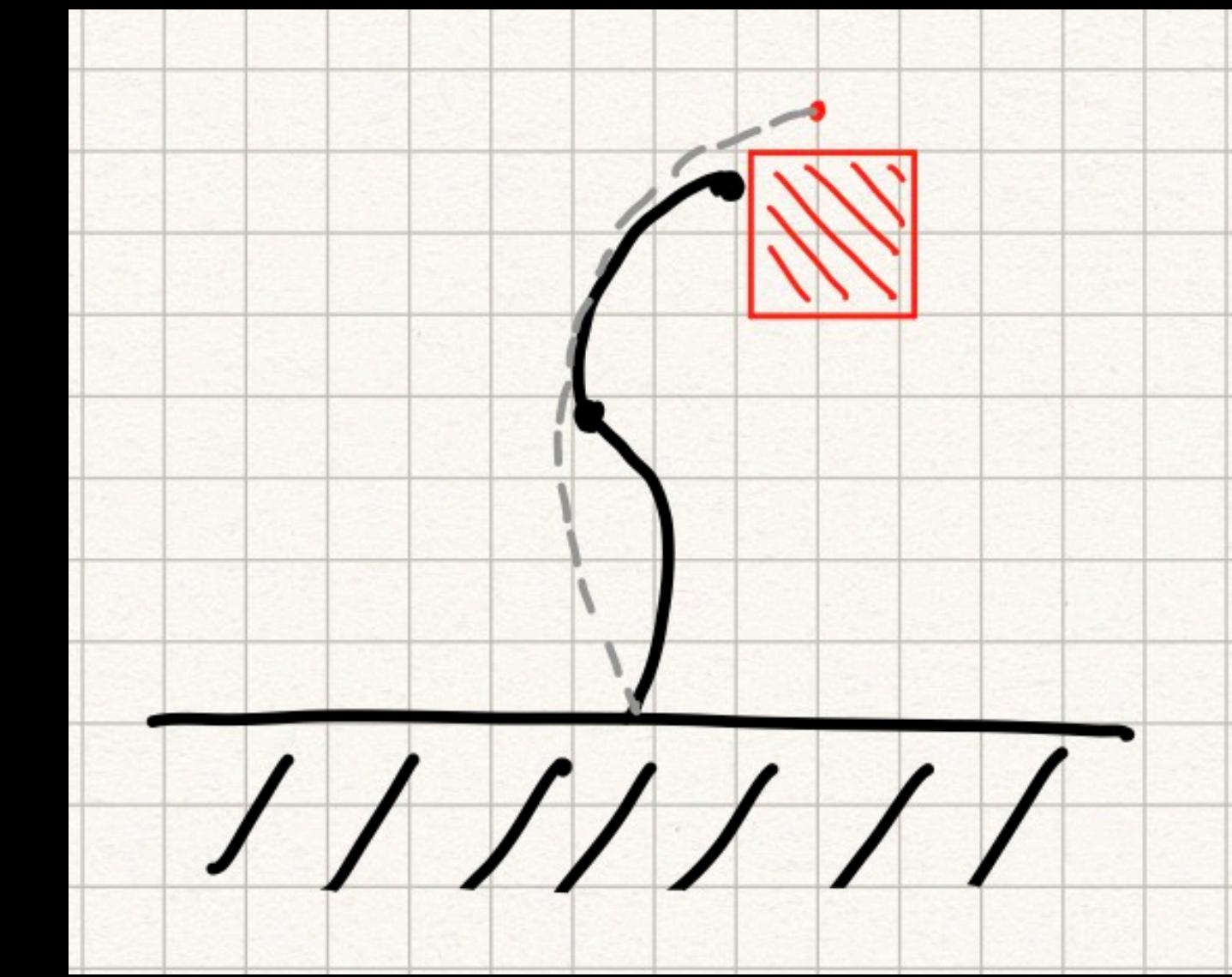
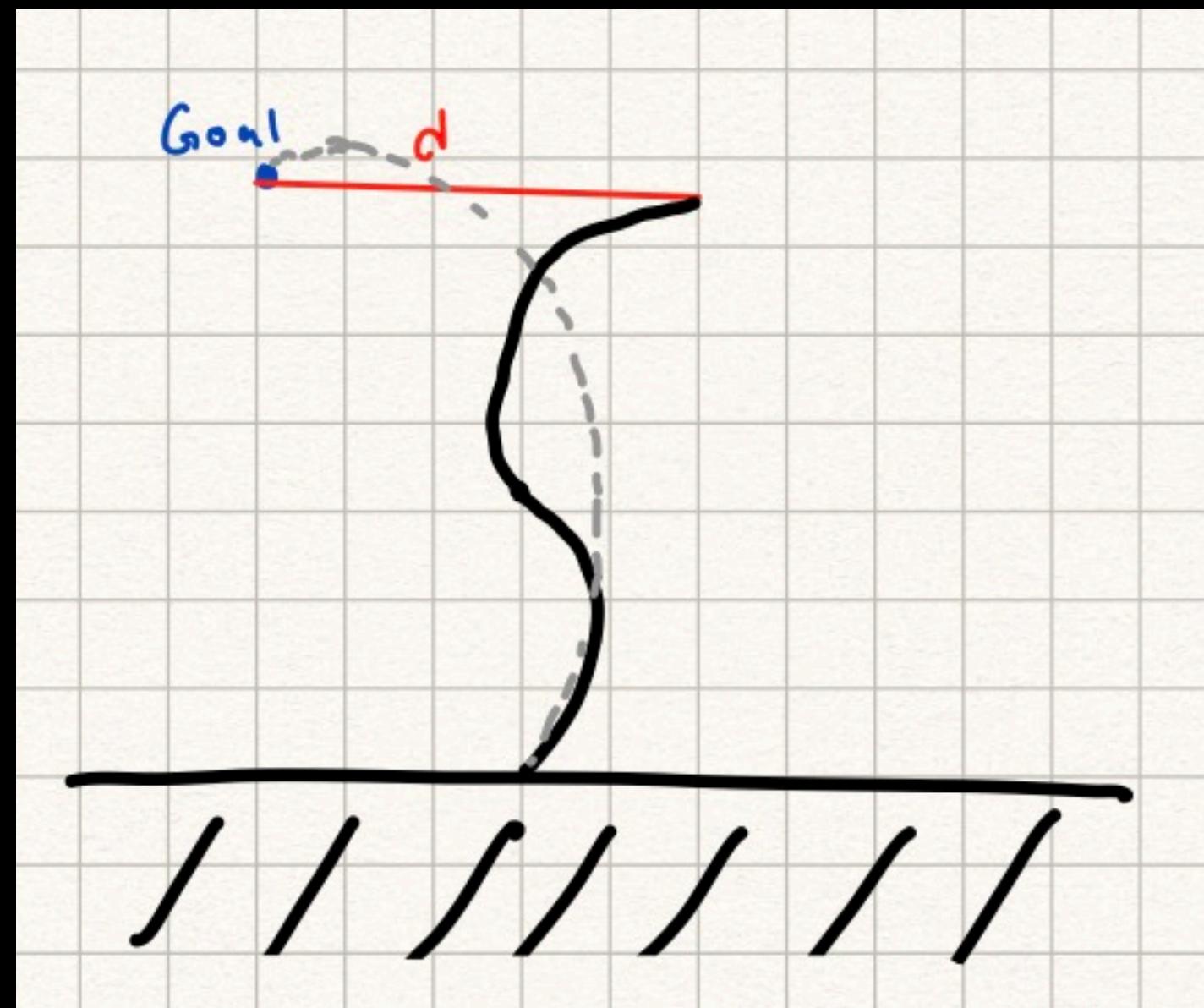
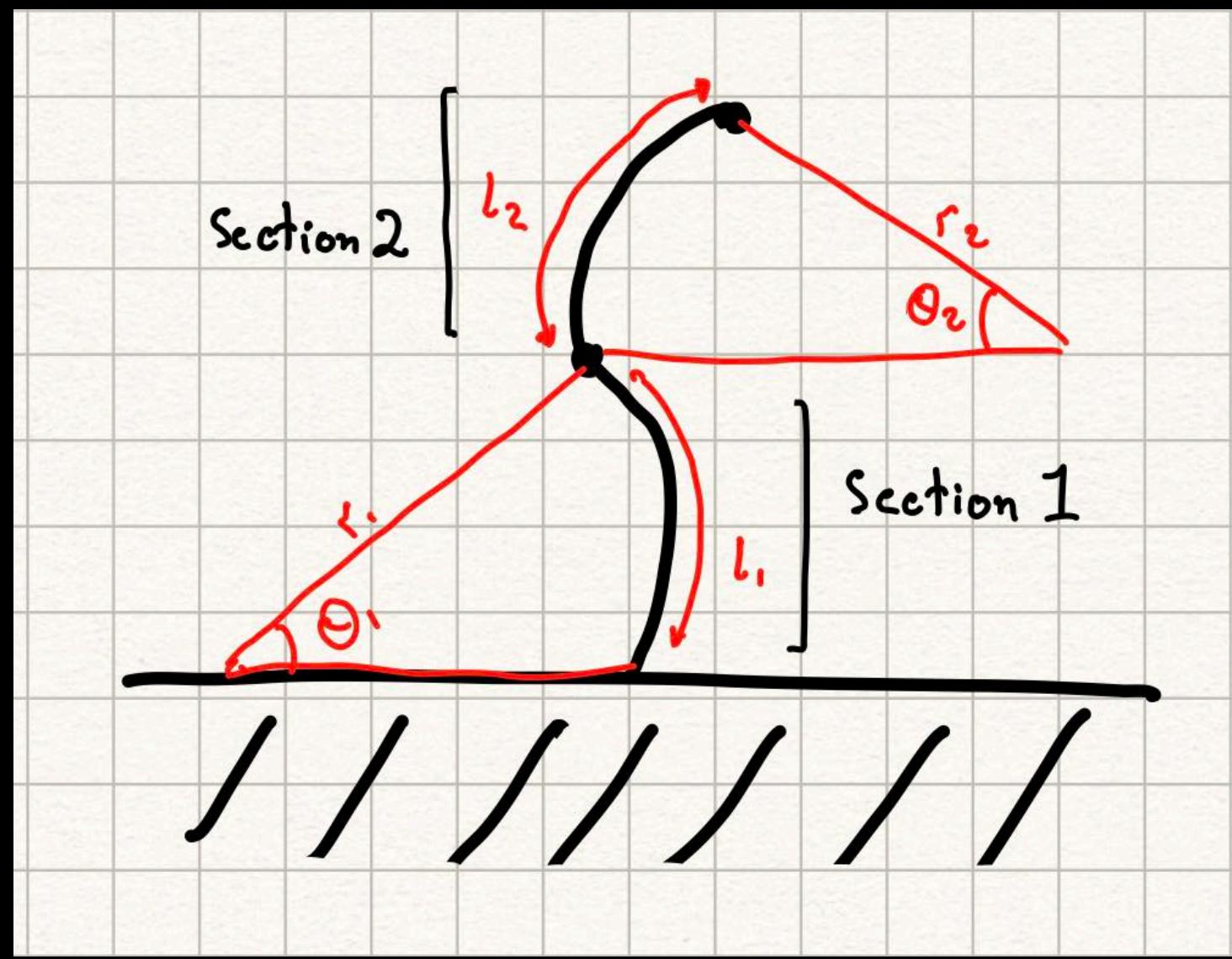


Controller/Actions:
[$(L_1, \Theta_1), (L_2, \Theta_2)$]

Limits:
- Max Segment Length
- Min Segment Length

Project Approach

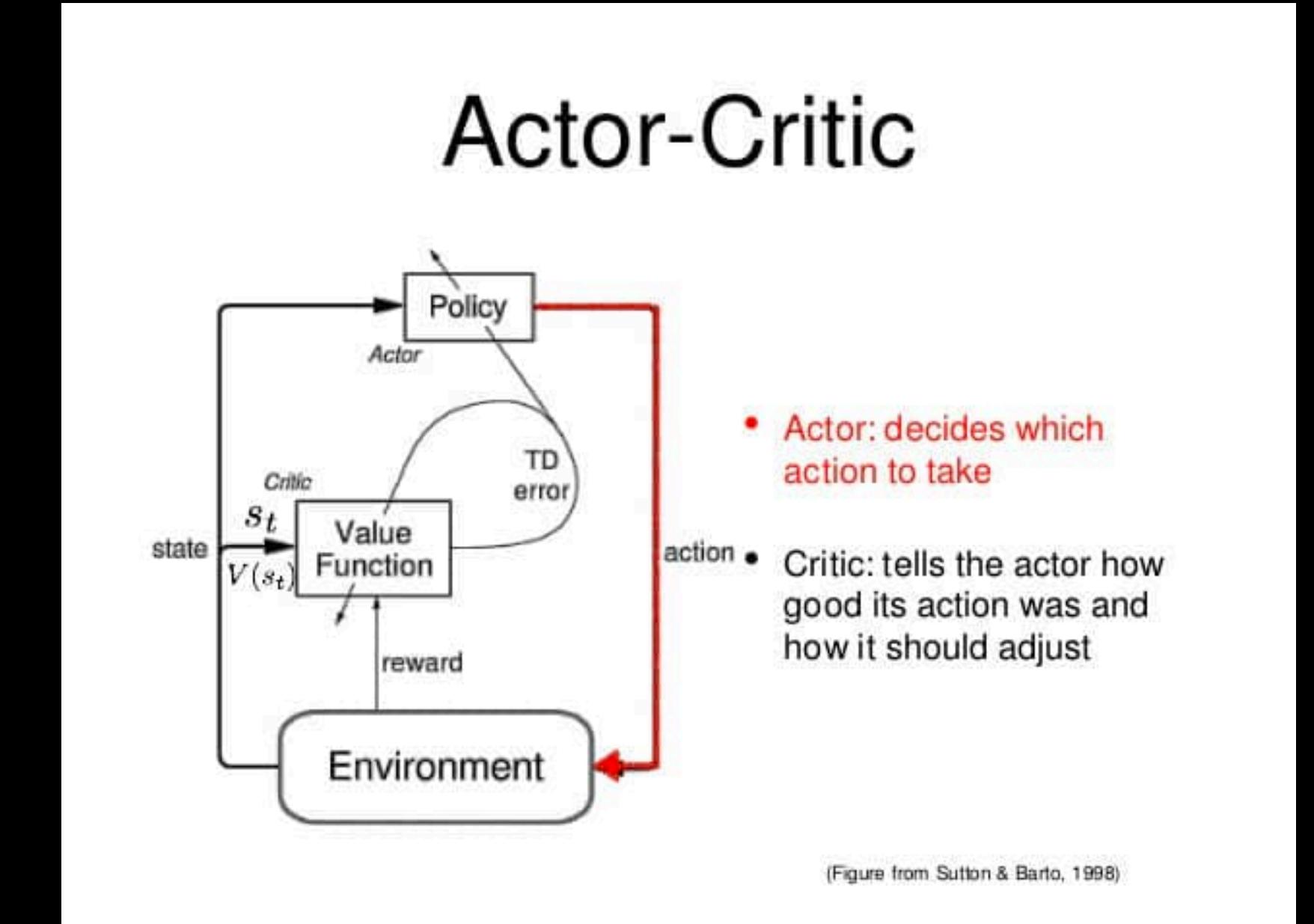
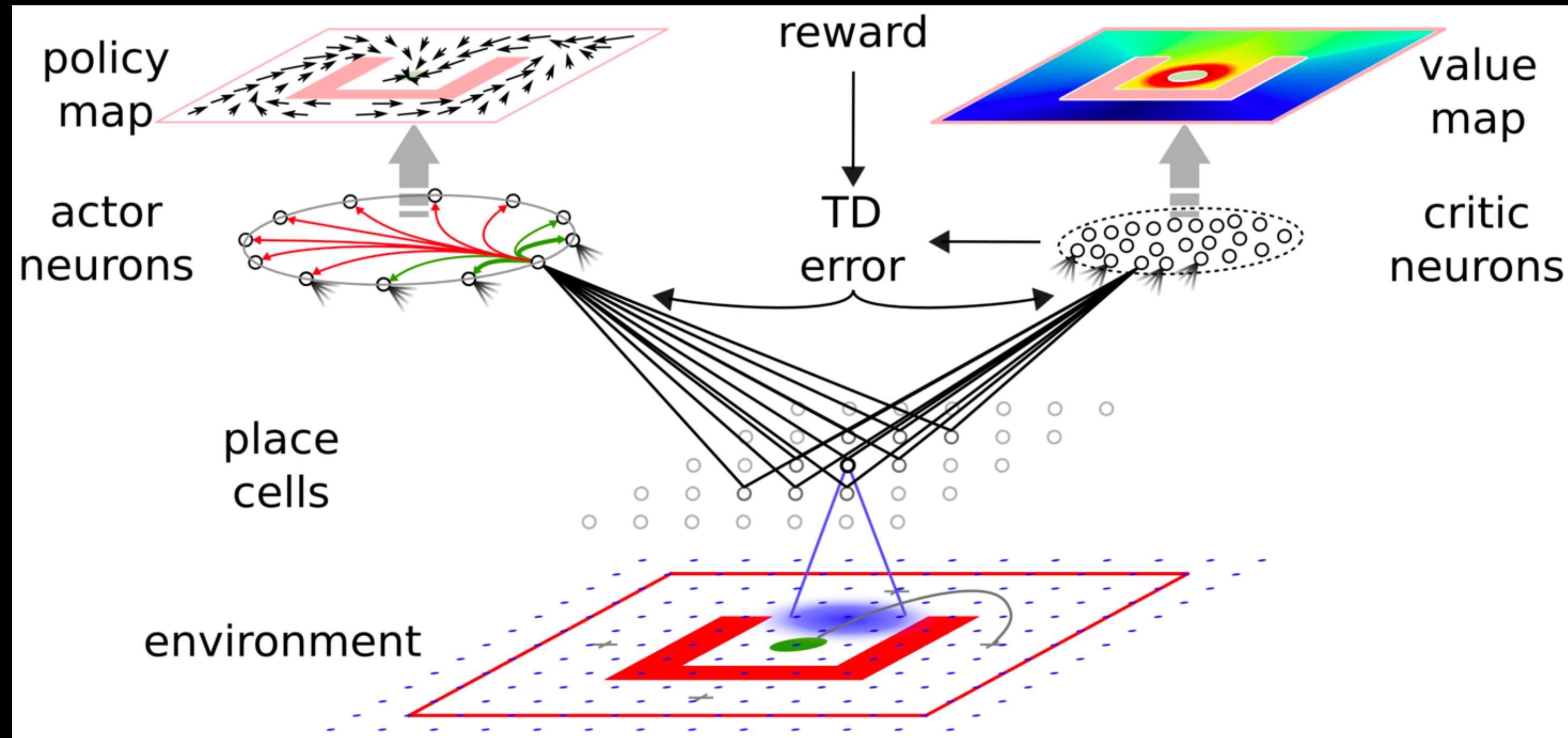
Reinforcement Learning - Reward Function



Reward (Section Configuration , Euclidean Distance, Collision)

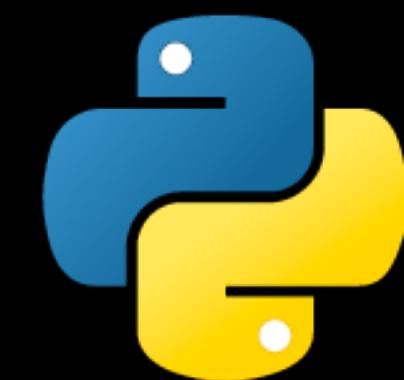
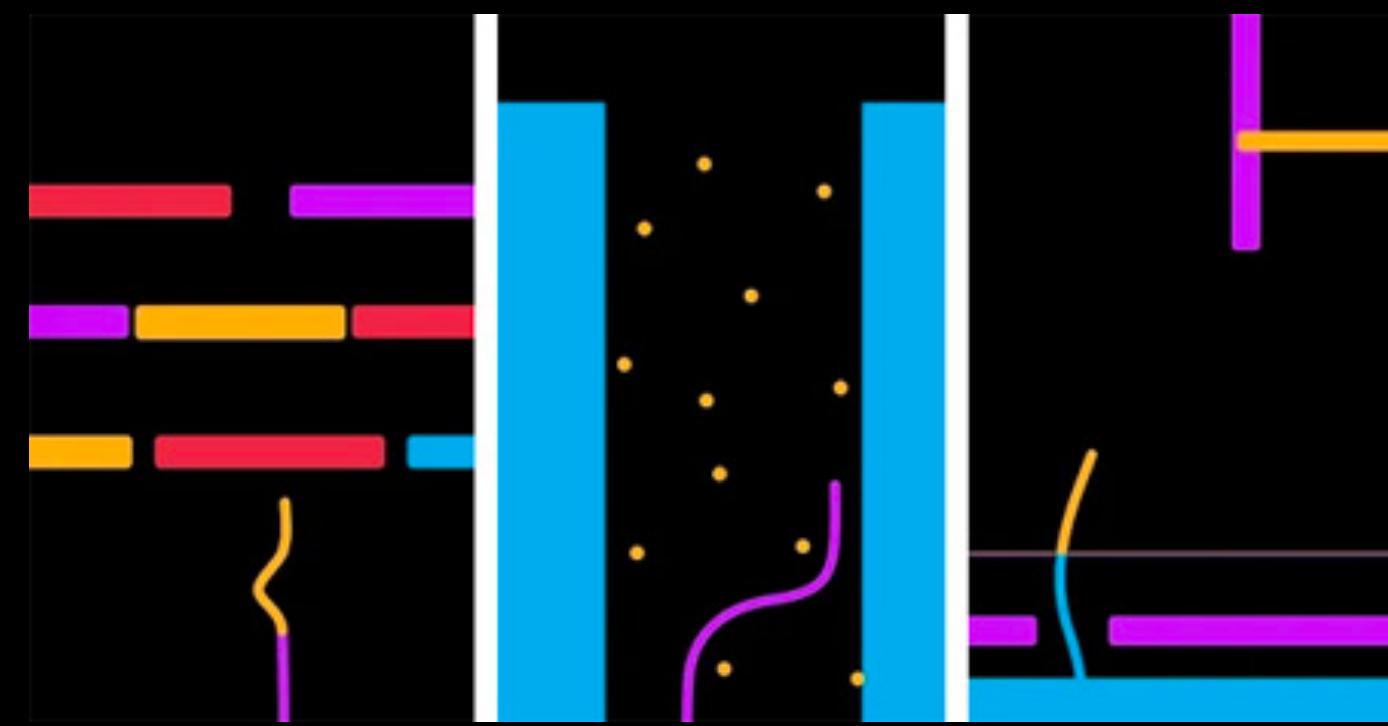
Project Approach

Reinforcement Learning - Neural Network

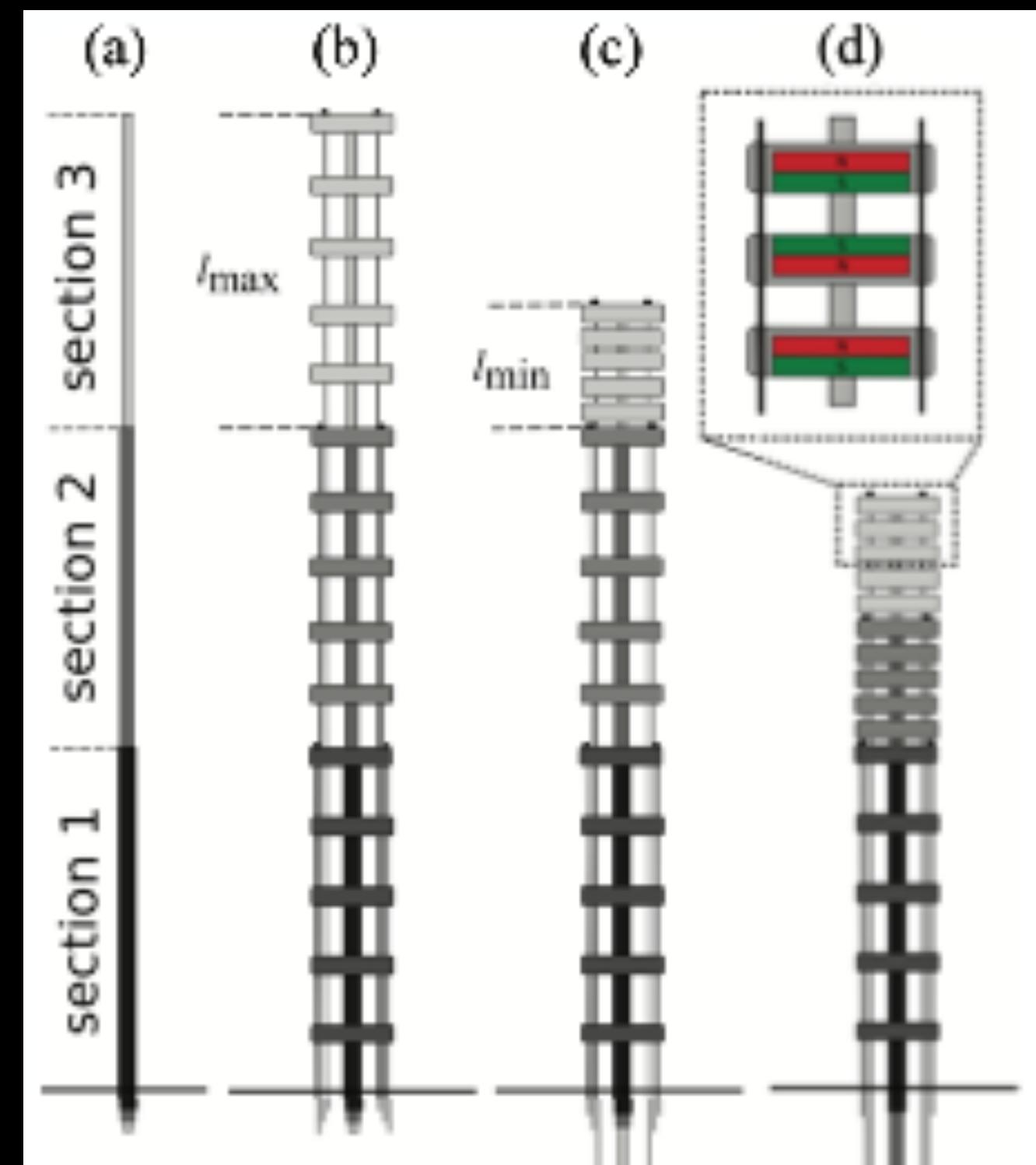


Deep Q Learning with Actor Critic Framework

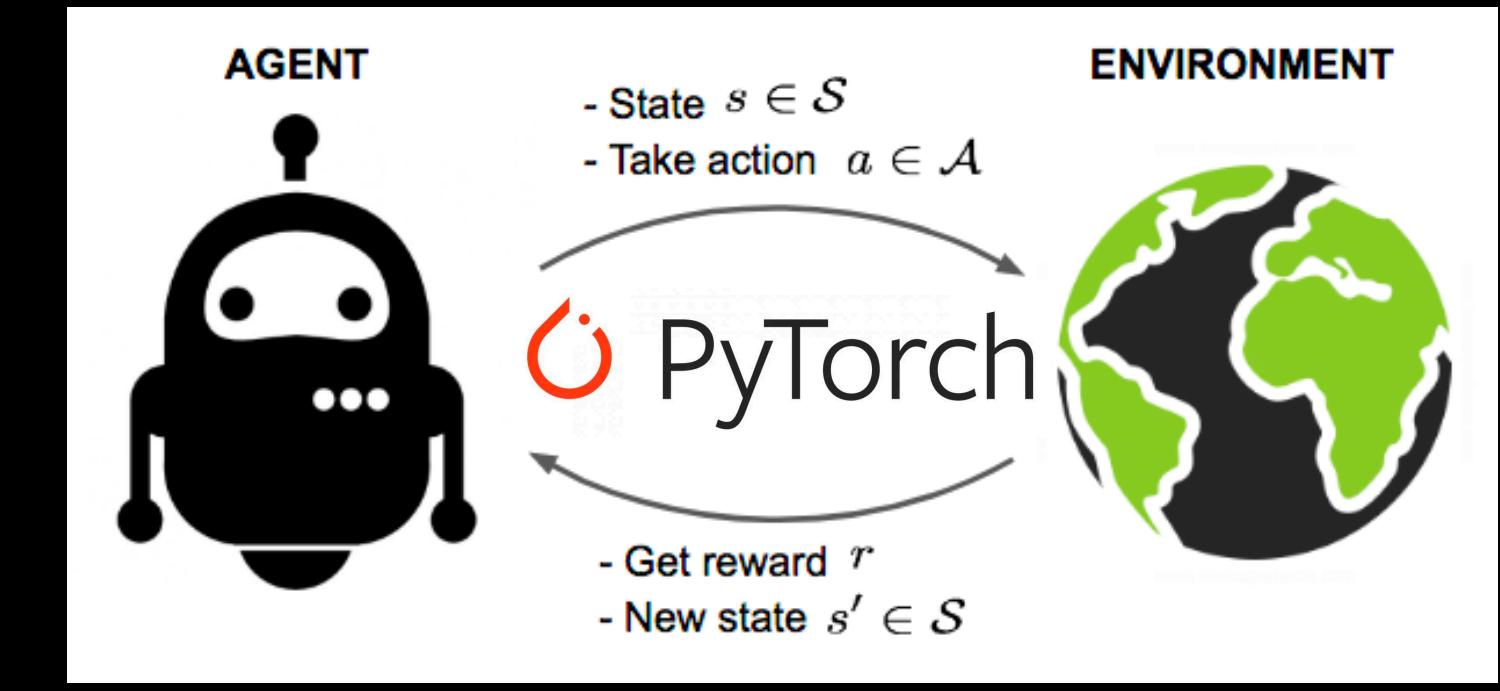
Project Approach



Environment



Prototype



Policy: Follow-The-Leader
Environment: 2-D Plane
Agent Parameters: Arc Parameters

Reinforcement Learning

Time Schedule



Possible Challenges

Reward Function

- Getting Stuck at a local Maximum instead of Global Maximum
- Exploration vs Exploitation Dilemma
- High Variance Episode Variance
- Reward Complexity

Thank You!