Robot Perception and Learning

Introduction and Logistic

Tsung-Wei Ke

Fall 2025

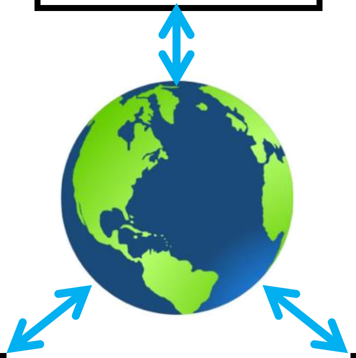
Welcome!

About Me

Research interests:

Perception

What do we see/hear/touch/sense?



Cognition

What do we want?
What did we do?
What and how to solve our problems?

Action

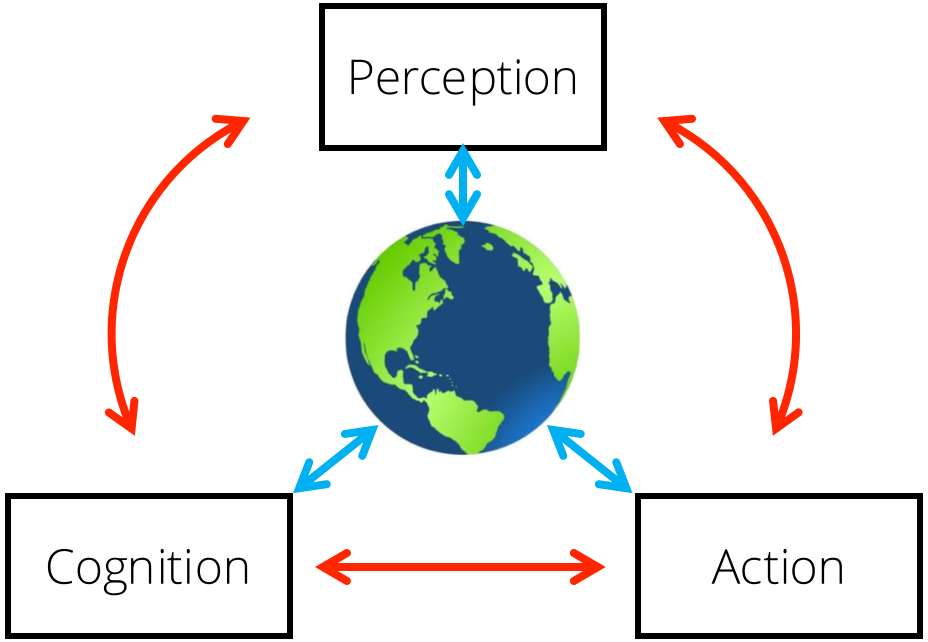
Push, grasp, release, rotate…

Extend my arm for 10 cms, bend my fingers for 180 degree…

Tsung-Wei Ke 柯宗瑋

About Me

Research interests: the trio of perception, cognition and action in the real world



Tsung-Wei Ke 柯宗瑋

About TA: Jia-Wei

Jia-Wei Liao 廖家緯

NTU CSIE PhD Candidate

Research Interests:
- Generative AI in Vision
- Mathematical Modeling
- Data Science

Email: jiawei.ta@gmail.com

Room: CSIE building 505

About TA: Yu-Wen

Yu-Wen Tseng 曾昱文

NTU CSIE PhD

Research Interest:
- Vision Perception for Autonomous Vehicle
- Test-time Adaptation

Email: d12922018@csie.ntu.edu.tw

Room: CSIE building 505

About TA: Chien-Sheng

Chien-Sheng Chiang **江建陞**

NTU CSIE undergraduate

Research Interest:
- Dynamic adaptation
- Reinforcement learning
- Flow model acceleration

Email: b10902063@ntu.edu.tw

Room: CSIE building 544

Administrivia

Enrollment

- I would like to enroll in this class. What should I do?
    - ➢ If you're interested in attending this course, please fill out and submit this form. We will first randomly sample 50-60 people for additional enrollment. We will release more slots if some students withdraw from enrollment

https://forms.gle/FH8NgPywkCCzS4XK6

- I am interested in robotic research but I am not enrolled in this class. What should I do?
    - ➢ We'll release the slides and learning resources online (https://ntu-rpl.github.io/).
    - ➢ Email me/TAs or message us or come to our office hours. We're happy to chat!

- I'm not yet sure if this course is for me…
    - ➢ Hopefully, this lecture can help you decide

Prerequisites

- Theory: Machine Learning, Linear Algebra, Probability Theory
  - ➢ Machine learning: stochastic gradient descent, loss function, optimization, neural network
  - ➢ Linear algebra: matrices, vectors, norms, scalar/vector products, orthogonality, singular value decomposition…
  - ➢ Probability: expectation, independence, Baye's Theorem…
  - ➢ Computer vision: convolutional neural networks, transformer models, 3D vision geometry
- Coding:
  - ➢ Linux system: setting up the required environment, being familiar with bash scripts
  - ➢ Python programming: creating python projects, importing required packages, visualizing your results
  - ➢ Pytorch programming: creating NNs, setting up training & evaluation pipeline
- Hardware: Grading in this course heavily depends on the coding assignments and final project. You should prepare for your own Linux machine with GPUs for the assignments and the final project. Otherwise, you can try cloud platforms for the access of GPUs.

Where Can I Access to GPUs?

- If you are undergraduate students in CSIE department, Meow with 4090s are available
- Otherwise:
  - Google Colab
  - Cloud computing services, such as Google Cloud, Amazon AWS, Microsoft Azure, Runpod, Lambda etc
  - Build your own PCs

Course Objectives

- The goal of this class is to guide you through "what are the essential components for controlling the robot's motion" and "when/how they are used" in robotic research. Going through each component thoroughly is not our goal.

- After this course, I hope you know:
  - ➢ The fundamental theorem of robot kinematics
  - ➢ The classical planning-and-control pipeline in robotics
  - ➢ Why learning is important in robotics, and what are the open research questions in robot learning
  - ➢ What components are under the hood of robot systems

- After this course, I hope you'd be able to:
  - ➢ Set up a simulator to kick-start your robotic research
  - ➢ Train robot policies with imitation learning
  - ➢ Train robot policies with reinforcement learning
  - ➢ Implement a diffusion / flow matching model

Evaluation

- Assignments: 30%
  - 3 assignments, 10% each
  - Individual submissions
  - Throughout the whole semester, students should submit their assignments before the deadline. **The score of the assignment will be multiplied by 0.9 for each additional day of delay. No grace day is granted**. The submission deadline is based on Taiwan's time zone.

- Final Projects: 70%
  - Teams of 3-4 members
  - Project proposal presentation (10%), two milestone report (2 * 10%), a rehearsal presentation (10%) and poster session (30%)
  - **In-person participation of these presentations are required**
  - Poster presentation will be graded by the instructor, TA and students

$$score = \frac{1}{N+2} score_{tw} + \frac{1}{N+2} \sum_{i}^{N} score_{TA,i} + \frac{1}{N+2} \frac{1}{|Students|} \sum_{i \in Students} score_i$$

Policy for Assignments

- You are encouraged to discuss with others, but **do not share your codes with them**! If you wrote the same code as others, you may waive the penalty by refactoring your code in-person within limited time, or otherwise, you'll get 15% total grade penalty (for each assignment).

- Please list your collaborator in the appendix of each assignment

- You are allowed to use AIs **at your own risk**. You are responsible for refactoring the code snippets generated by AIs. You'll get the penalty as long as your submitted codes are the same as others.

Policy for the Final Project

- You can form a team of 3-4 members. If you really want to work alone, come and chat with us.

- Project proposal presentation: You will prepare for a 5-minute oral presentation, describing the topic, experimental setup, todos and expected contribution of each member for the final project.

- Two milestone reports: You will write few pages of report, describing the progress and issues of your ongoing project.

- Rehearsal presentation: You will prepare for a 5-minute oral presentation to rehearse your poster presentation.

- You will prepare a poster to present your final project, describing the overall ideas, methodology and results. Performance and experimental results are not the key to attract attention! Tell a good story to impress your audience.

Tentative Schedule

Week	Course	Announcements
1	Introduction	
2	Robot Kinematics	
3	Motion Planning, Task and Motion Planning	
4	Kinodynamic Motion planning, Trajectory Generation, Feedback Control	
5	Optimal Control	Release HW1
6	Intro. of RL and Imitation Learning	Team matching
7	Markov Decision Process, Policy Iteration and Value Iteration	Release HW2, Course withdrawal ddl, HW1 ddl
8	Monte Carlo, Temporal Difference and Monte Carlo Tree Search	Project proposal presentation
9	Model-free RL: Policy Gradient, Q learning, REINFORCE, Actor Critic	

Week	Course	Announcements
10	On-policy DRL (PPO), Off-policy DRL (SAC), offline RL, RL finetuning	Milestone report
11	Model-based RL	Release HW3, HW2 ddl
12	Sim2real transfer, Real2Sim2Real Learning, Adaptation	Milestone report
13	Intelligent exploration, Visual Imitation Learning	
14	Special topic: Bimanual Dexterous Manipulation	
15	Special topic: Loco-manipulation	Final project rehearsal
17		Poster presentation

Policy for the Final Project

Date	Announcements
10/22	Final project proposal
12/22	Poster session

- If you don't know what to work on, come join the office hours, we're happy to chat with you
- **Think big** and **own** your final project! Don't just consider it as one of the assignments in a class. If you do it well, the final project could end up an awesome paper in a top-tier conference.
- If you want more advice on your final project, come join the office hours or email or message us to arrange for meetings

Policy for course withdrawal

Date	Announcements
10/17	Deadline of Course Withdrawal

- Please be aware of the course withdrawal deadline!
- Any application of course withdrawal after 10/17 **will not** be accepted!
- Please be responsible for your teammates. Piggybackers will be penalized.

- If most of your teammates withdraw, you'd have two options:
  1. Keep working alone. We know it'll be tough, but you'll enjoy at the end of the day.
  2. Team up with others. We'll help you to join other teams.

Communication

- Lecture:
  - Please don't hesitate to raise questions if you have any, I may not know the answers but I'll try my best
- Office hours:
  - Great opportunity to discuss project/research idea, confusion about the lecture, debugging issue, or any difficulty you have
  - Tsung-Wei's OH: 13:00-14:00 pm every Friday; Jia-Wei's OH: 12:00-13:00 pm every Friday; Yu-Wen's OH: 12:00-13:00 pm every Wednesday
- Email:
  - Tsung-Wei's email: twke@csie.ntu.edu.tw
  - Jia-Wei's email: jiawei.ta@gmail.com
  - Yu-Wen's email: d12922018@ntu.edu.tw
  - Chien-Sheng's email: b10902063@ntu.edu.tw
- Discord:
  - Feel free to post if you have any problem/issue. The whole community will try to help you.
  - Please change your username in the following format: [department]-[year]-[name]. For example: [CSIE]-[Ph.D 4 year]-[廖家緯]

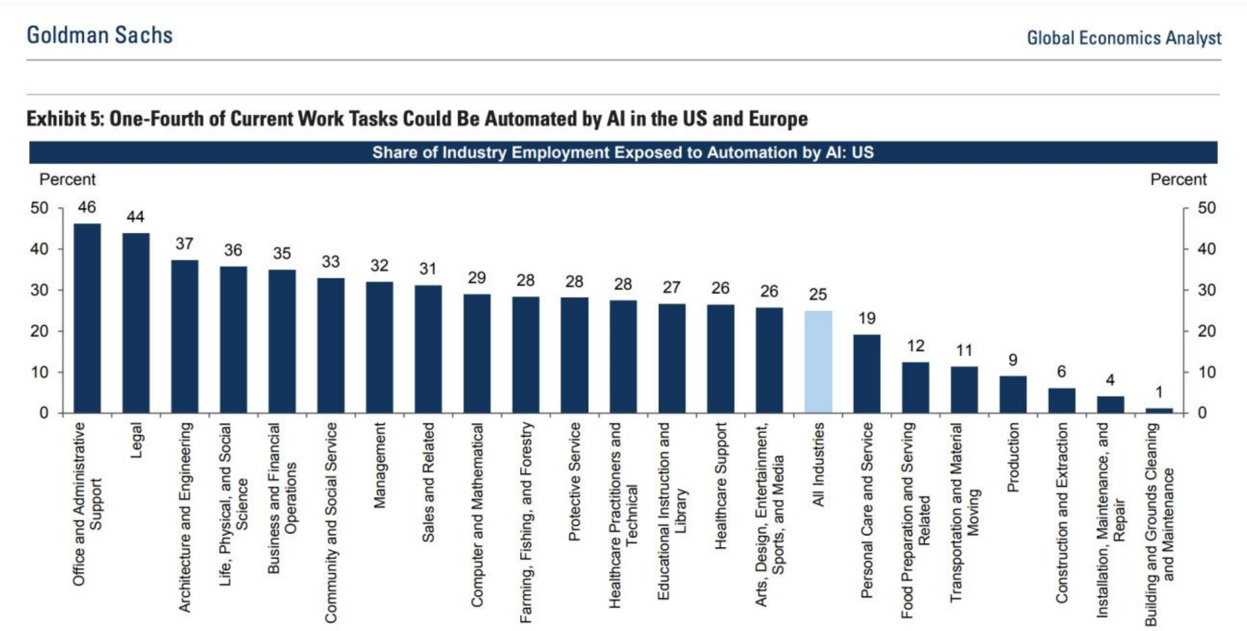
Disclaimer – New(ish) and Evolving Course

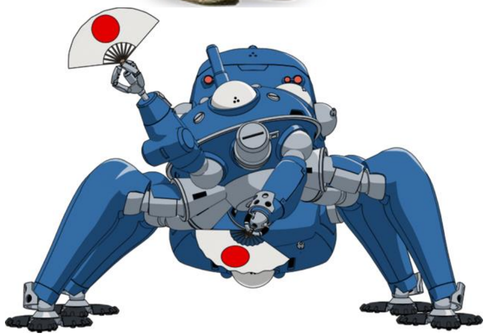
- I'm also learning these contents with you, and there will likely be bugs in material and lecture. Apology in advance! Please don't hesitate to point them out.
- Many contents are borrowed/adapted from other professors' lecture.
- Some math equations / notations might be wrong and confusing. Just point them out if you are confused.

Questions?

Why do we need this class?

How Far We Are to Build Robots with Artificial General Intelligence?





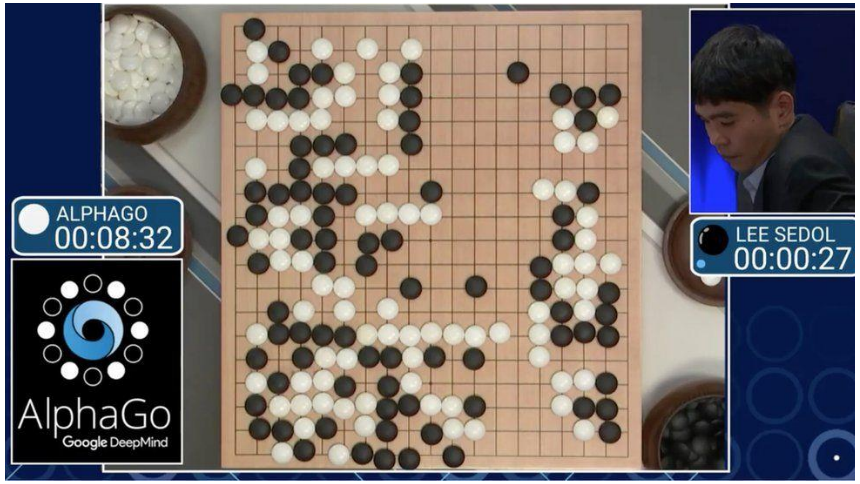
Gemini Claude

Grok Qwen

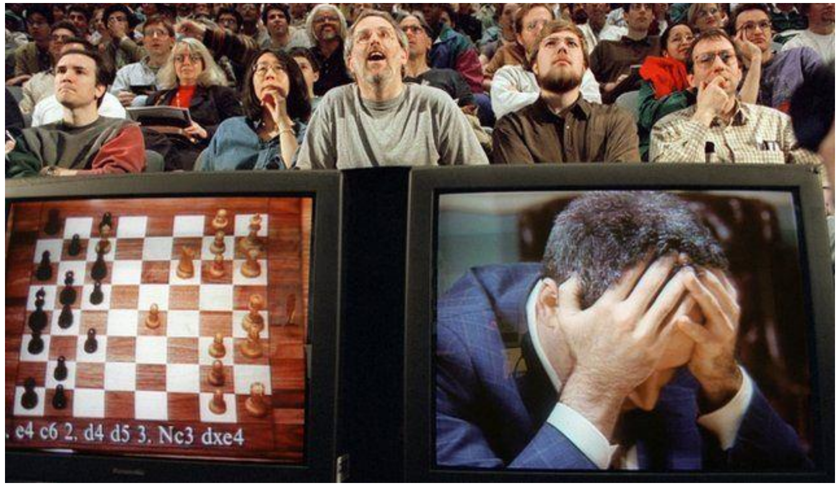
What is Intelligence?

https://youtu.be/JcDBFAm9PPI?si=4yzpsudC6m_zIYd9

Are These All You Need to be Intelligent?

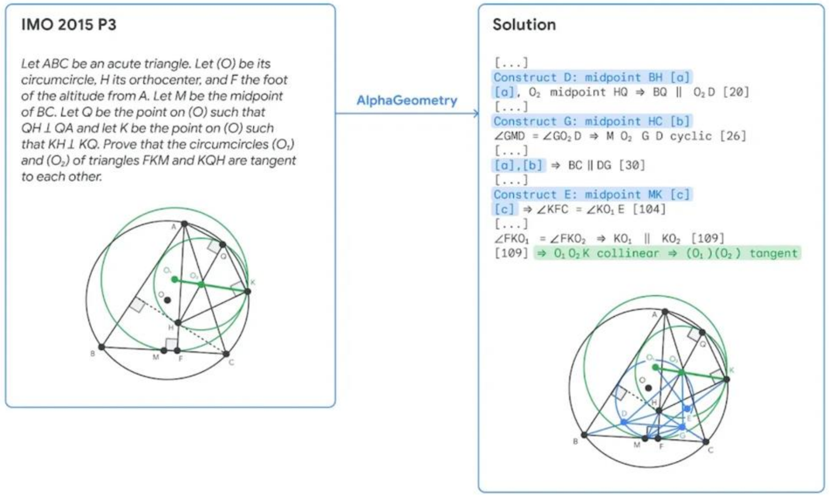


DeepMind's AlphaGO. Image credit https://www.bbc.com/news/technology-35785875



IBM's Deep Blue. Image credit https://www.bbc.com/news/technology-35785875

https://openai.com/index/vpt/



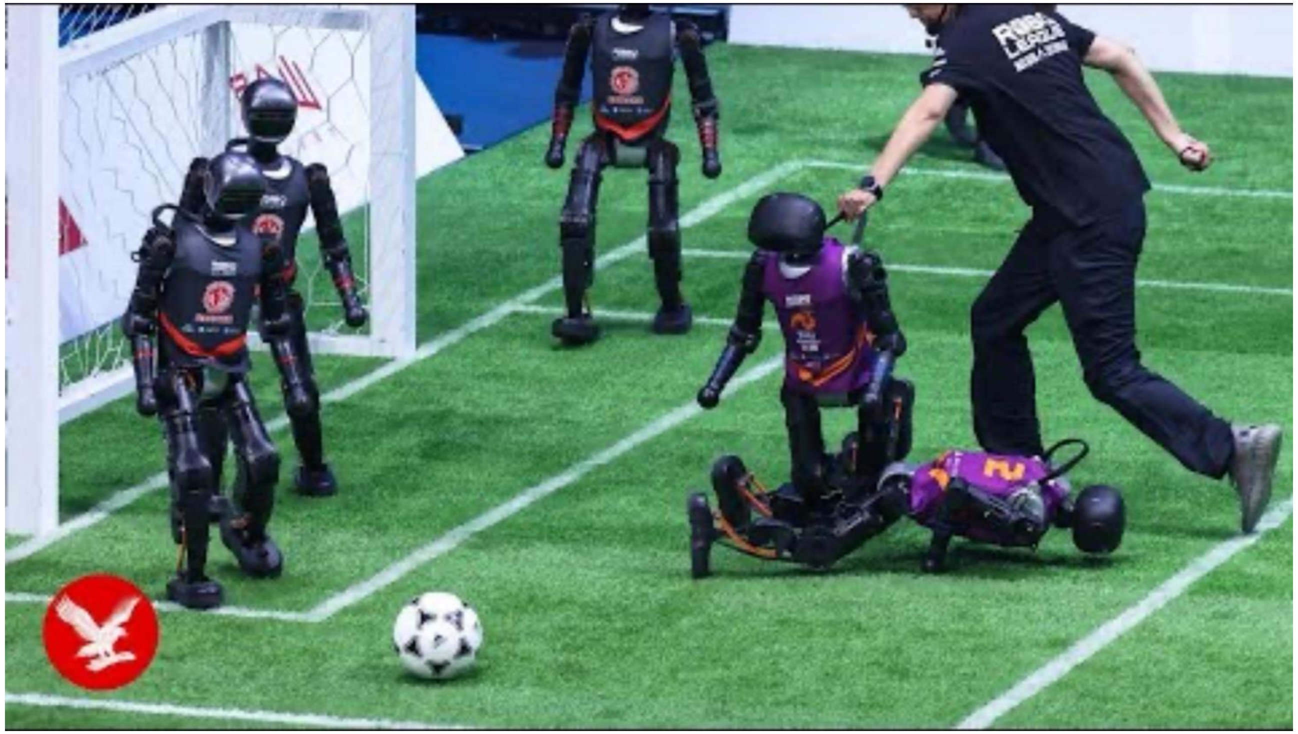
DeepMind's AlphaGeometry

OpenAI's SORA

Meta's SAM2

Wait, Best AIs Can't Do These Tasks Well…

https://www.youtube.com/watch?v=v4vGGHeYtbg



https://youtu.be/bo4d1dH0kZY?si=tNpp7tyPiorjC5lj

https://youtube.com/shorts/8Drm_v3_iG4?si=MdR0ImFrP2EVz4XM

But Human Can Do These...

https://www.youtube.com/watch?v=GL0rbxB9Lqg

https://www.redbull.com/au-en/10-most-epic-free-solo-climbs

https://www.youtube.com/watch?v=5RPNZ0YlGNk

https://youtu.be/VKuU1OoYs50?si=PXIwca8pphq3tdgs

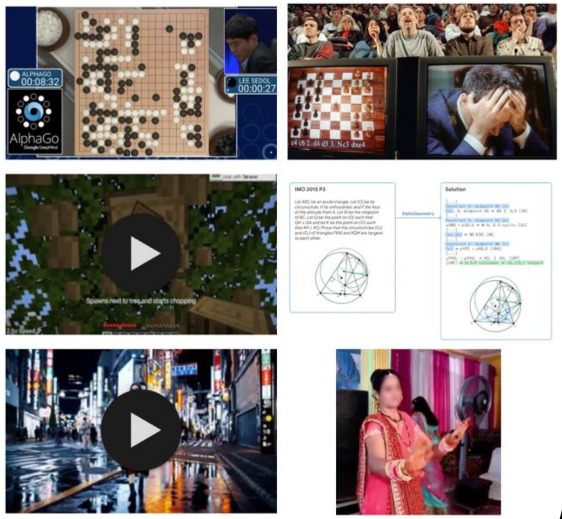
https://youtu.be/3pfhUeLMnsM?si=Spt_WsiiH708ZZRn

https://www.youtube.com/watch?v=dhW4XFGQB4o

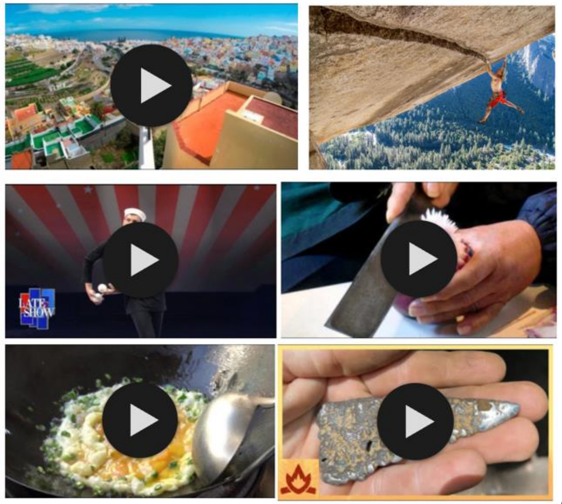
Moravec's Paradox

What makes the difference?
The key is that we are in a physical world

Easy universe



Hard universe



"*We are all prodigious olympians in perceptual and motor areas, so good that we make the difficult look easy. Abstract thought, though, is a new trick, perhaps less than 100 thousand years old. We have not yet mastered it. It is not all that intrinsically difficult; it just seems so when we do it*"

Hans Moravec

"*The main lessons of thirty-five years of AI research is that the hard problems are easy and the easy problems are hard. "The mental abilities of a four-year-old that we take for granted—recognizing a face, lifting a pencil, walking across a room, answering a question—in fact solve some of the hardest engineering problems ever conceived*"

Steven Pinker

Why the Physical Universe is Hard?
Dynamics! Partial Observation! Uncertainty!

https://www.youtube.com/watch?v=RqajKat0v-4

https://www.youtube.com/watch?v=eFsT3QgIvoI

https://youtube.com/shorts/xSqvz-ystw4?si=M_UoOSmRS_Uahtr9

https://youtu.be/L49NBoW5Xw0?si=W8w8qi6guVdo7j_W

https://youtu.be/VXrBowsNFis?si=fY59E0YLPW54UV6W

https://youtu.be/WazuwEkIGzA?si=sEIrYTYu07iOb3kY

Why the Physical Universe is Hard?

1. Dynamics, partial observations and uncertainty
2. There is almost no “reset” of your behavior
3. Rules are unknown, you only learn by “experience”
4. Reward is sparse and ambiguous. “Survival” could be the only shared reward among humankind, which still differs among different culture
5. No perfect simulation

We Want Robots that Work in the Physical Universe

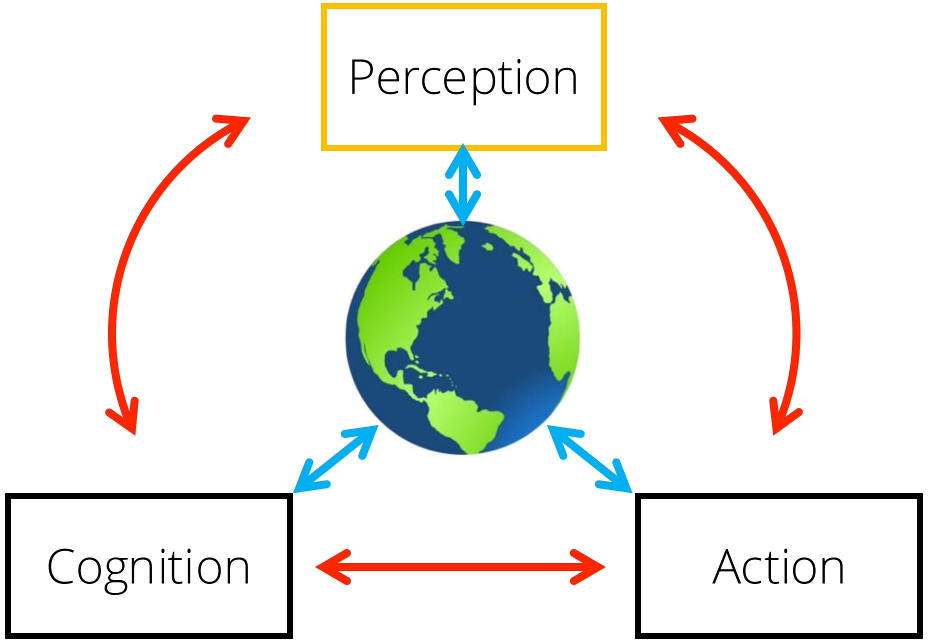
Robots in an easy universe

https://youtu.be/WYnOGAvQEgk?si=W5uwZdIZuwWmKD5w

Robots in the hard universe

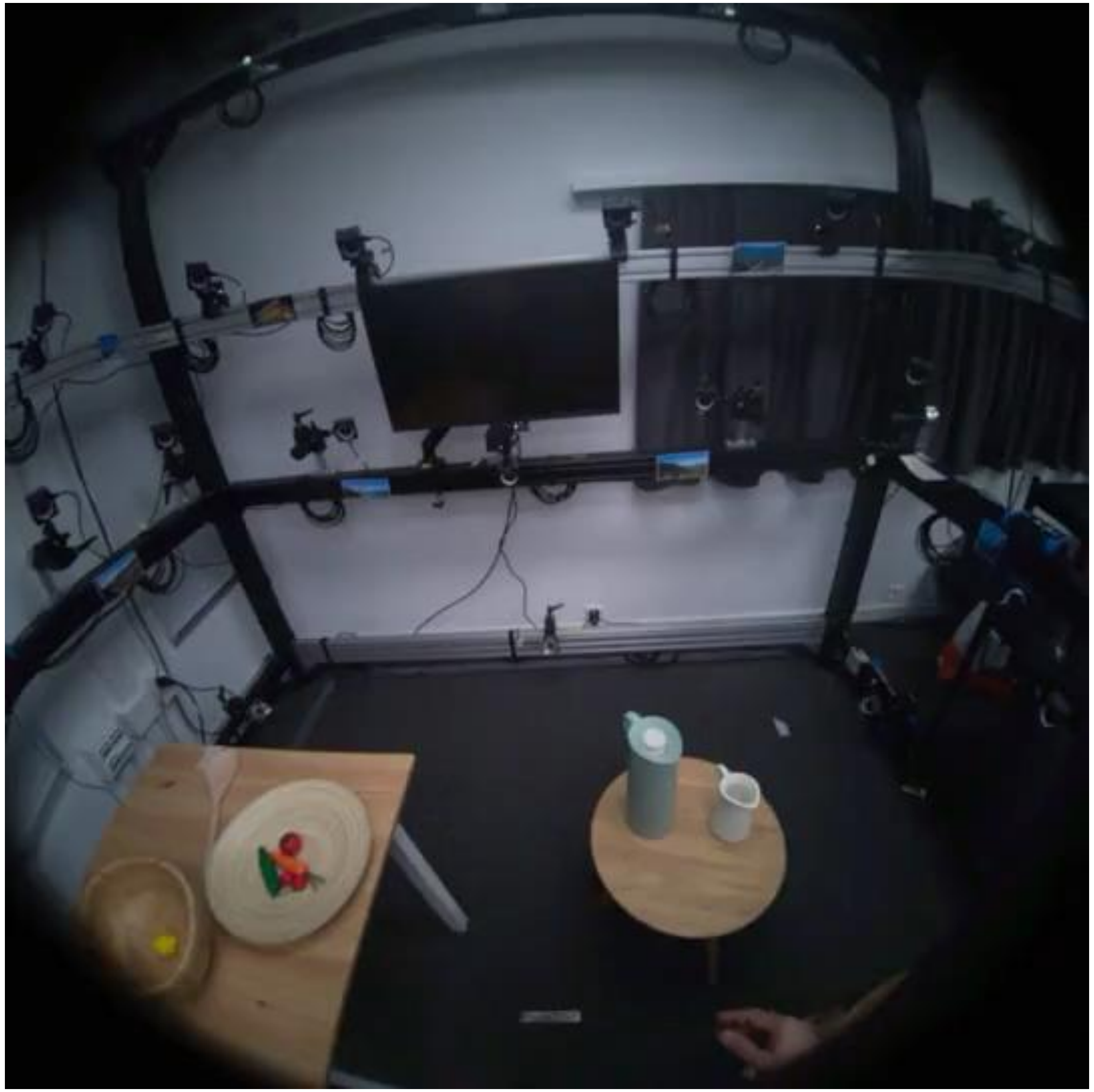
https://youtu.be/QHH3iSeDBLo?si=oJQKP3ig27h03Kpc

To Build a Robot that Works in the Real World, We Need to Study the Trio of Perception, Cognition and Action



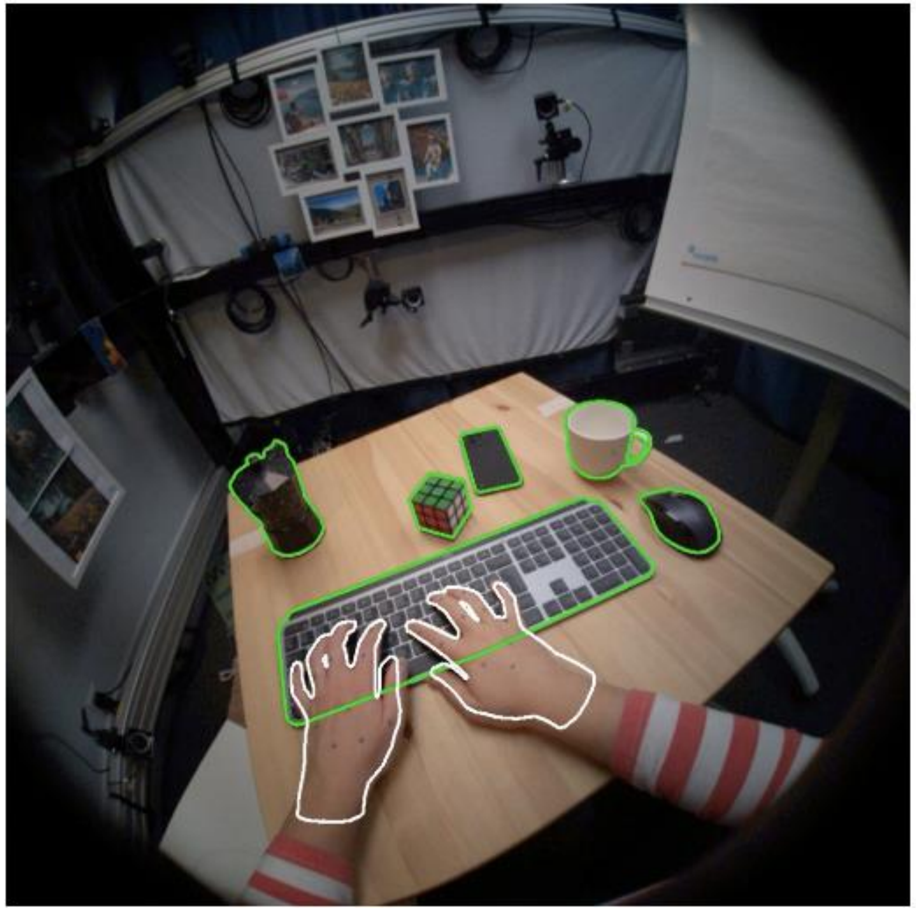
In simplification, we want to answer questions: "What are our goals?" and "What action should we take to achieve our goals in the current environment?".

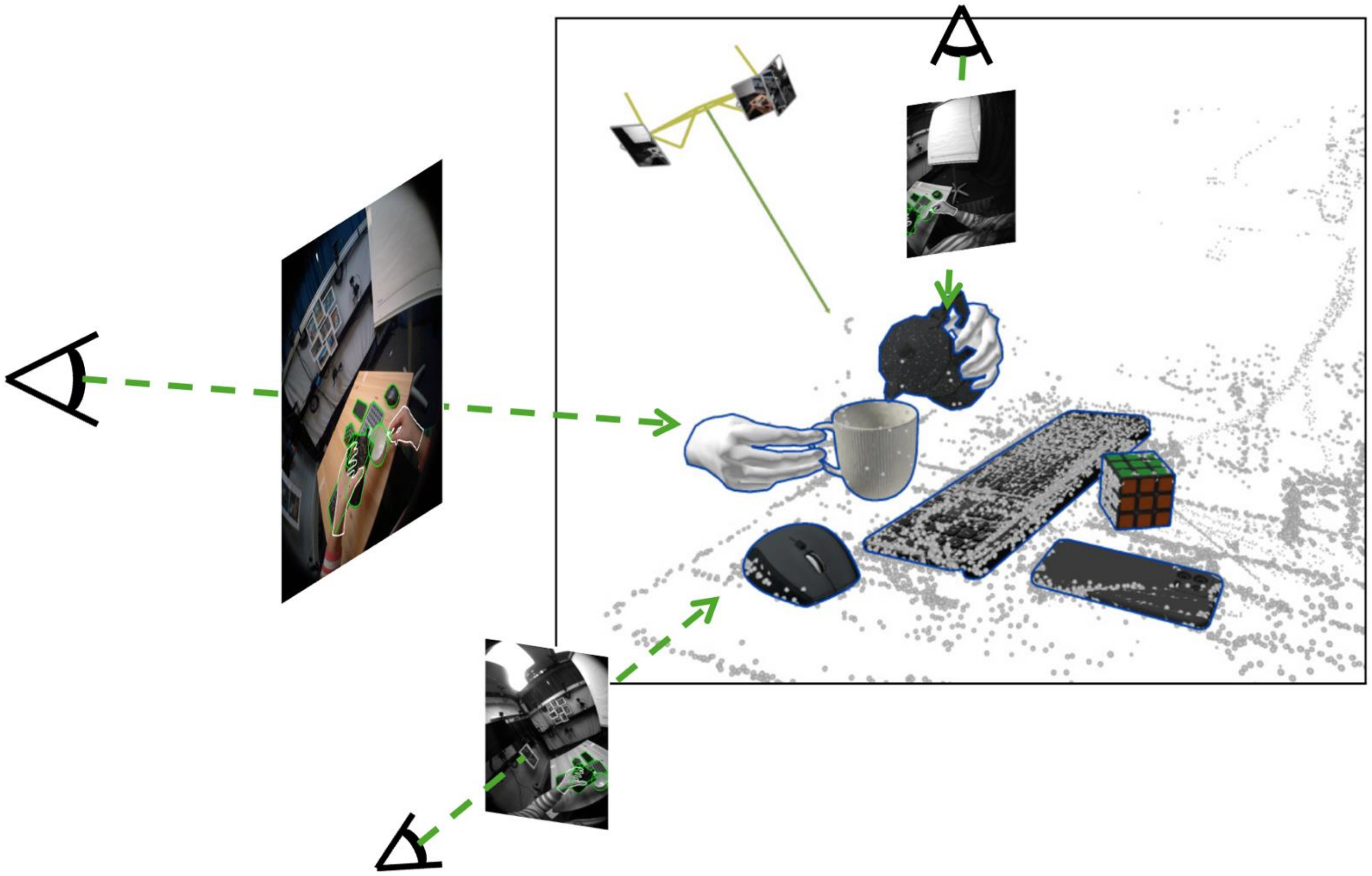
What Are Important Perceptions?



https://explorer.projectaria.com/hot3d-aria/P0001_95eabc0a?st=%220%22

Recognition, Detection and 3D Modeling are Important





HOT3D dataset

Tracking and Pose Estimation are Important

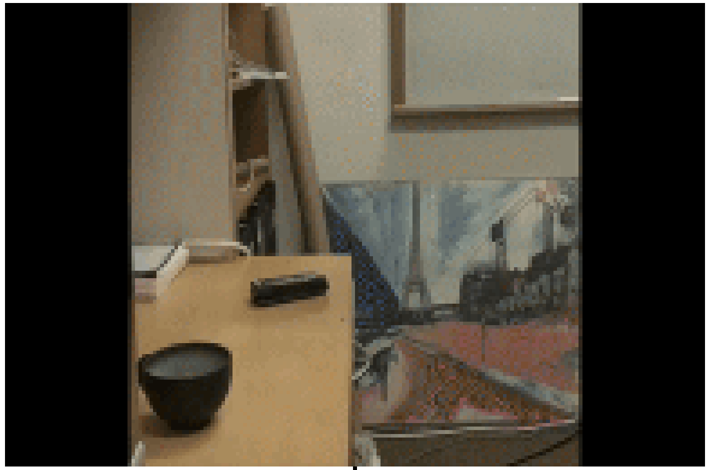
https://explorer.projectaria.com/hot3d-aria/P0001_95eabc0a?st=%220%22

Inferring Physical Properties is Important

Are the cookies well toasted?
Is the plate hot?
Is the plate heavy?
Where to grasp the plate?
Is the glove slippery?

Image credit COCO Dataset

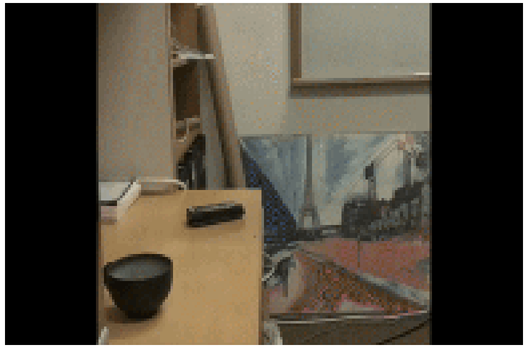
Inferring the Consequence of Actions is Important



Action 1

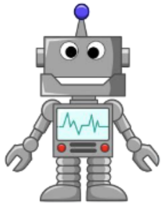
Understanding the Dynamics is Important

Laws of Physics
Environmental parameters
Object shape
Object 3D location (4D motion): $o_1, o_2, \ldots, o_{T-1}$
Object physical properties



World

Action: $a_1, a_2, \ldots, a_{T-1}, a_T$
3D Location (4D motion): $l_1, l_2, \ldots, l_{T-1}$
Shape
Physical properties
…



Perception Is More Than Image/Video

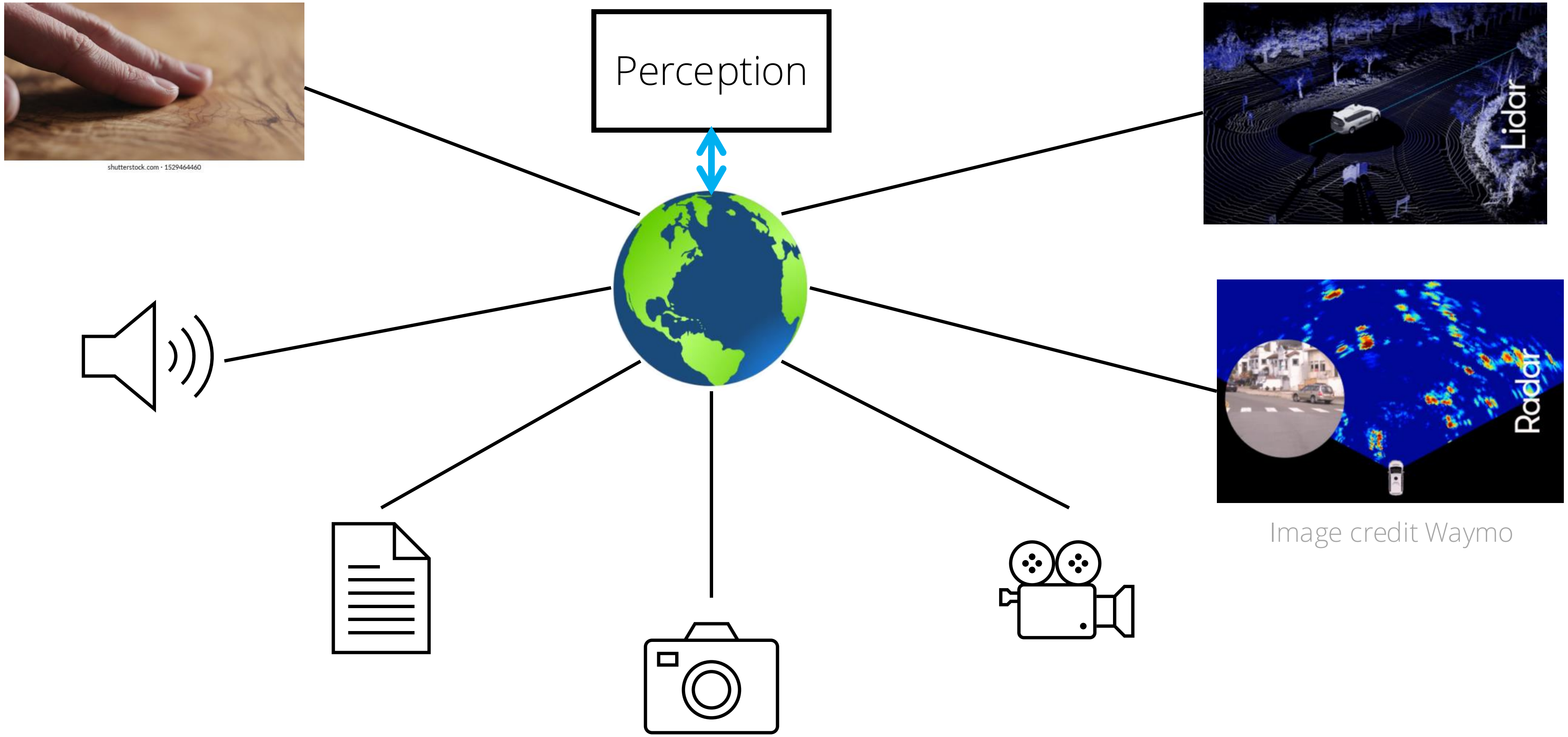
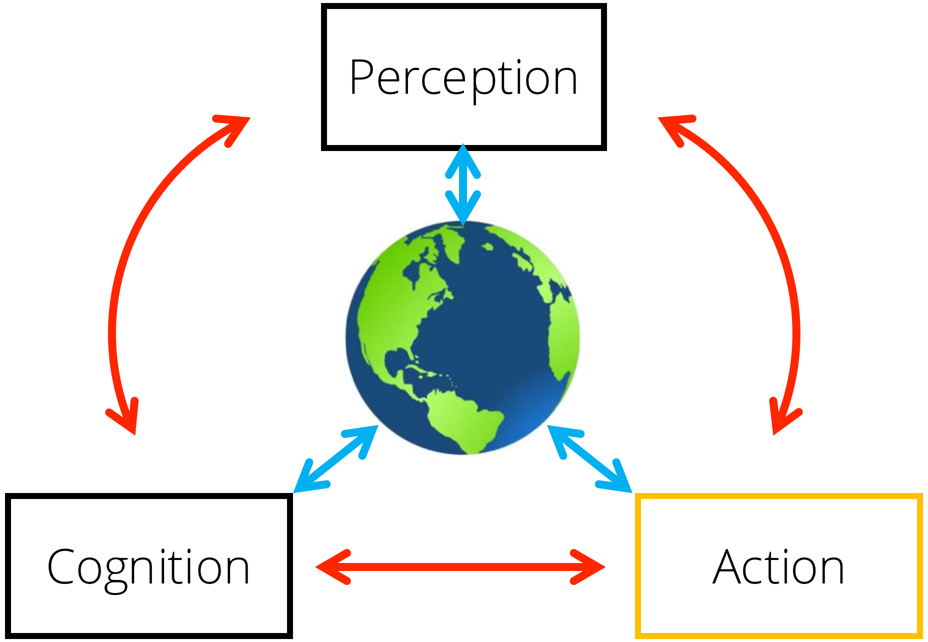


Image credit Waymo

We will Cover the Perception Part in Embodied Vision Course CSIE5421

We will Focus on the Action and Cognition in this Course



In simplification, we want to answer questions: "What are our goals?" and "What action should we take to achieve our goals in the current environment?".

Decision Making: Decide What Action to Take to Achieve Goals (Rewards)

Action: muscle contractions
Rewards: food, attentions

Image credit S. Levine

Action: muscle contractions
Rewards: food

https://www.salon.com/2021/01/16/human-breeding-of-cats-has-made-them-look-like-they-are-always-in-pain/

Action: steering, acceleration
Rewards: win

https://racingnews365.com/championship-standings-after-2024-f1-australian-grand-prix

Action: selecting companies, buying, selling
Rewards: cash

Action: (x, y) location
Rewards: win

https://en.wikipedia.org/wiki/Go_%28game%29

Action: jump, ↑←→↓
Rewards: win

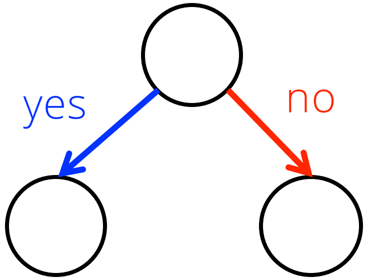
Image credit Nintendo

Action: RGB at each pixel
Rewards: cash, aesthetics

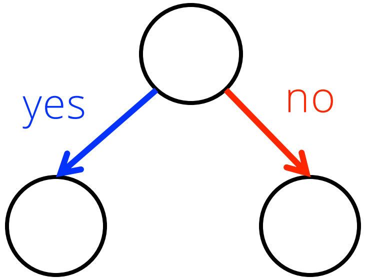
Decision Making: Decide What Action to Take

I.I.D decision making

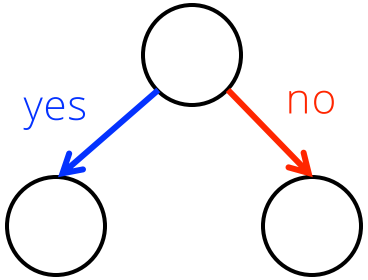
Have a breakfast



Have a lunch

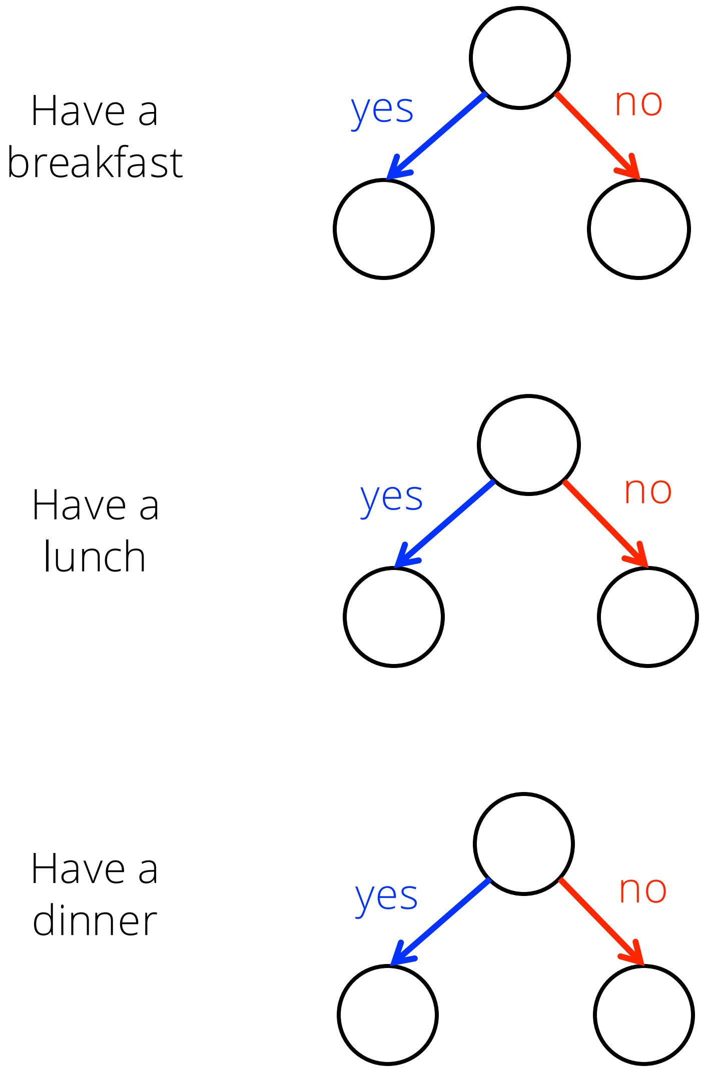


Have a dinner

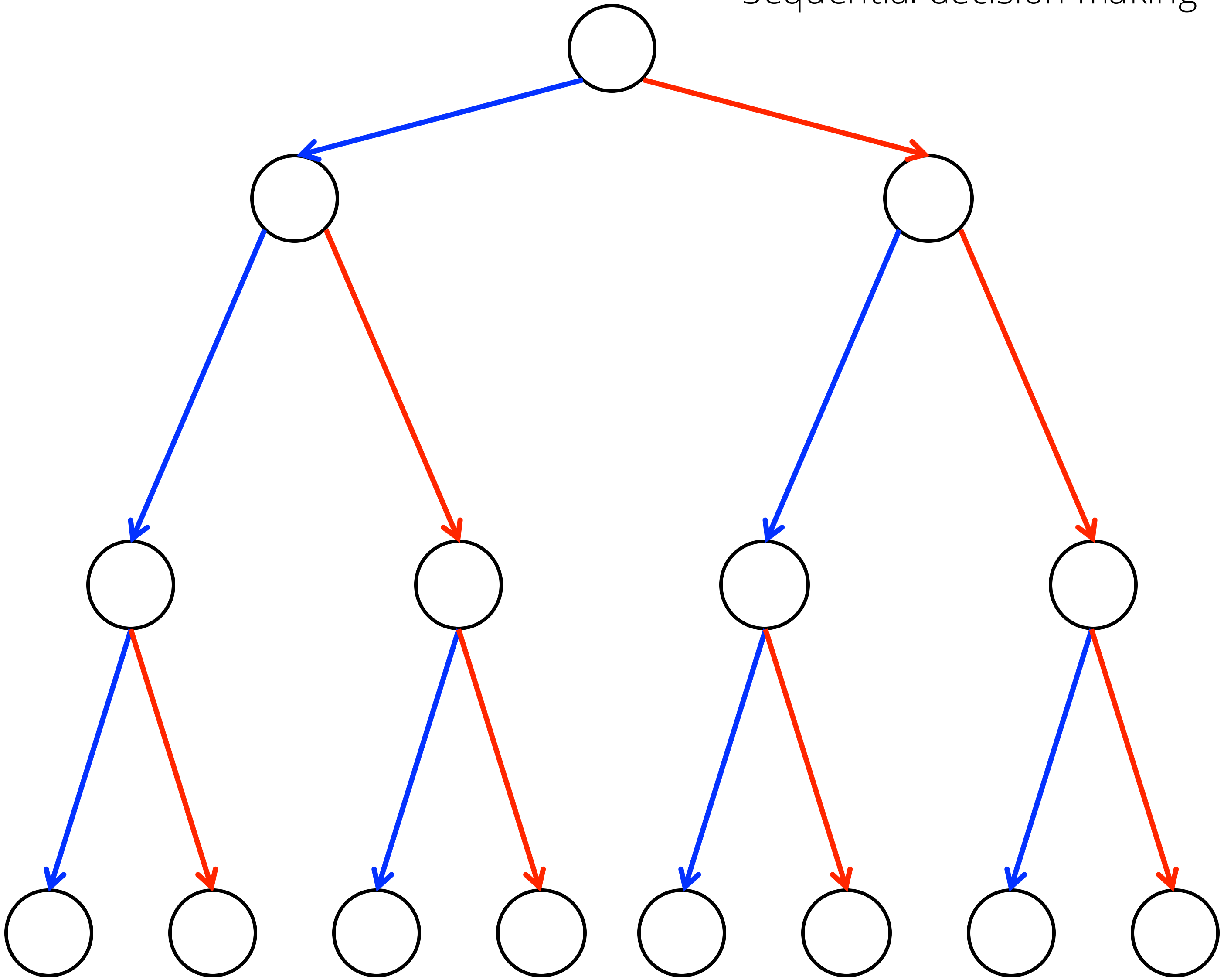


However, Decision Making Problems Are Often Sequential

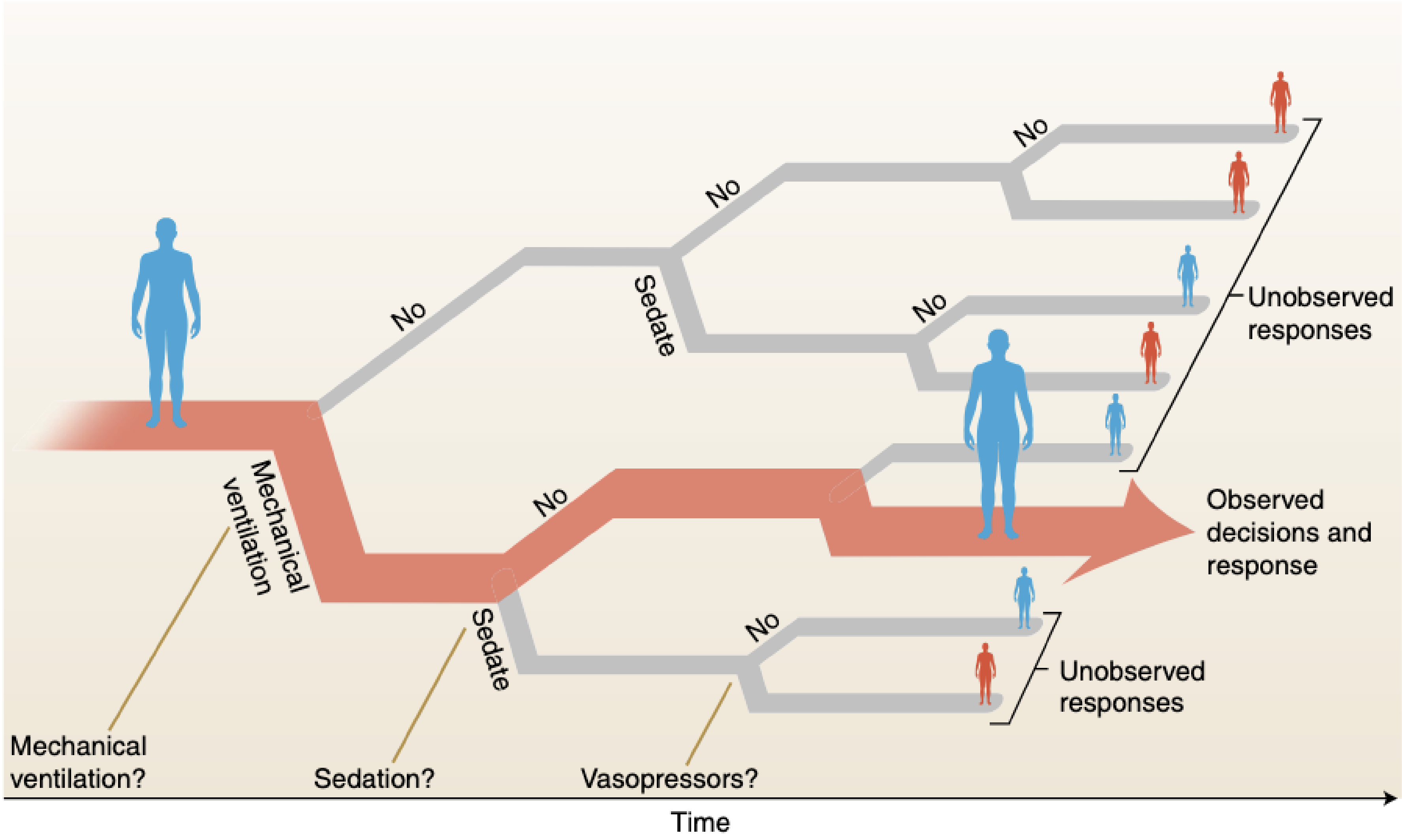
I.I.D decision making



Sequential decision making

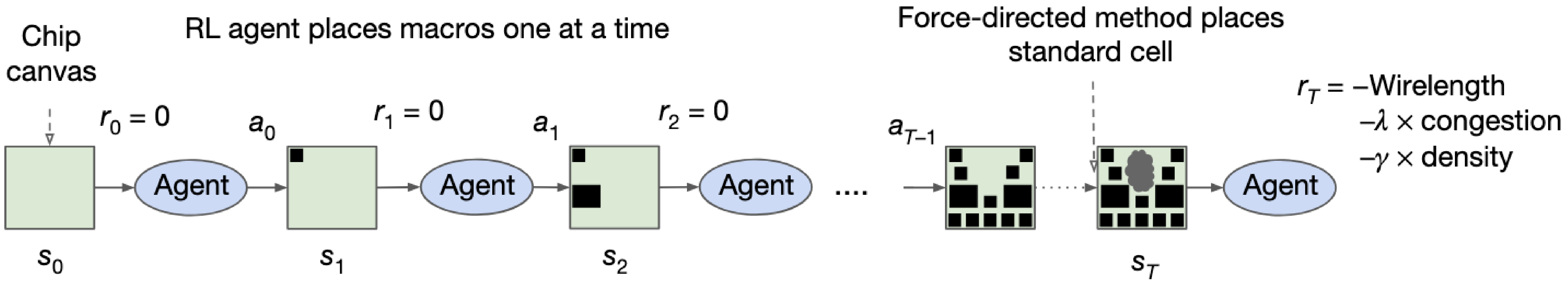


However, Decision Making Problems Are Often Sequential



Guidelines for reinforcement learning in healthcare. Gottesman et al. Nature.

However, Decision Making Problems Are Often Sequential



A graph placement methodology for fast chip design. Mirhoseini et al. Nature.

However, Decision Making Problems Are Often Sequential

https://youtu.be/a_YGPbWJO5g?si=3ZEr3rc5gjQ9JluX

https://youtu.be/6Zbhvaac68Y?si=Ivi7bak6xi-ByGzA

https://youtu.be/I-Y5FHI4JXc?si=mzqvEBCyu-YnYKun

https://www.youtube.com/watch?v=BhMSzC1crr0

https://youtu.be/cM6woUhJsCo?si=dKOi0WOUOLfbgkAk

https://youtube.com/shorts/GYeSEKnHjDI?si=8xSn3p9EBoW4nwBy

Each Decision You Make Influence Future Outcomes

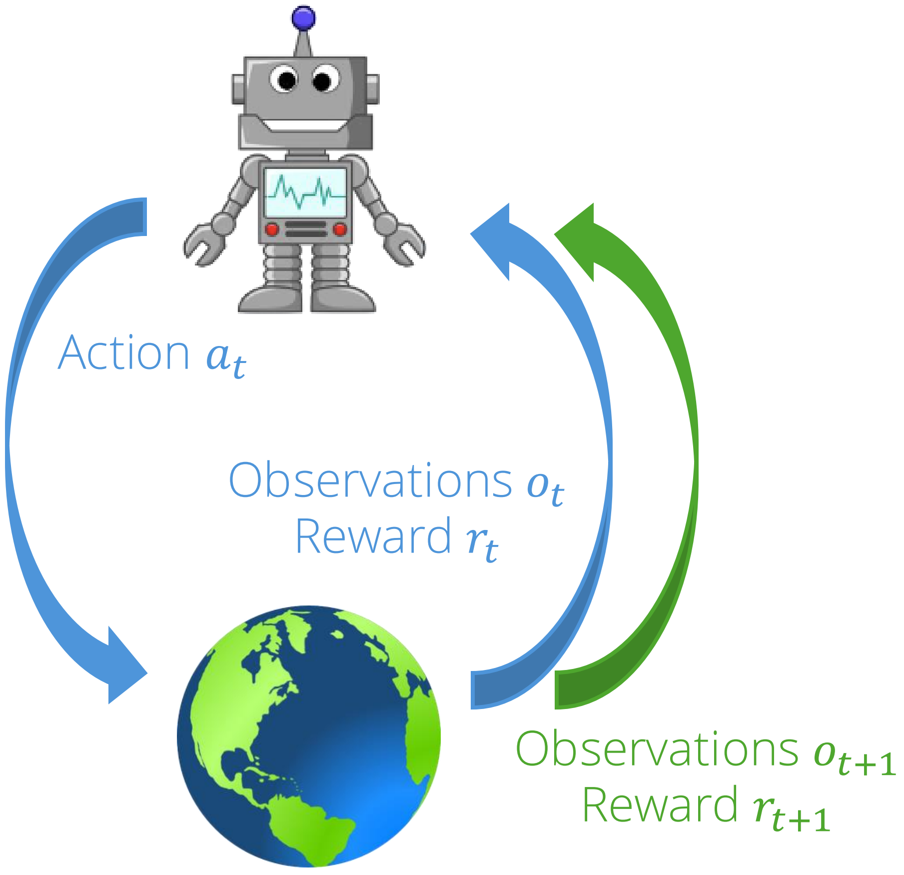
League of Legends

What is my first item?

Age of Empire 2

Where should I drop the town center?
How many villagers on wood?
When should my villagers explore the map?
⋮

Sequential Decision Making Problem



Trajectory: $s_1, a_1, r_1, s_2, a_2, r_2, \dots, s_T, a_T, r_T$

Environment constraints: $s_{t+1} = f(s_t, a_t)$

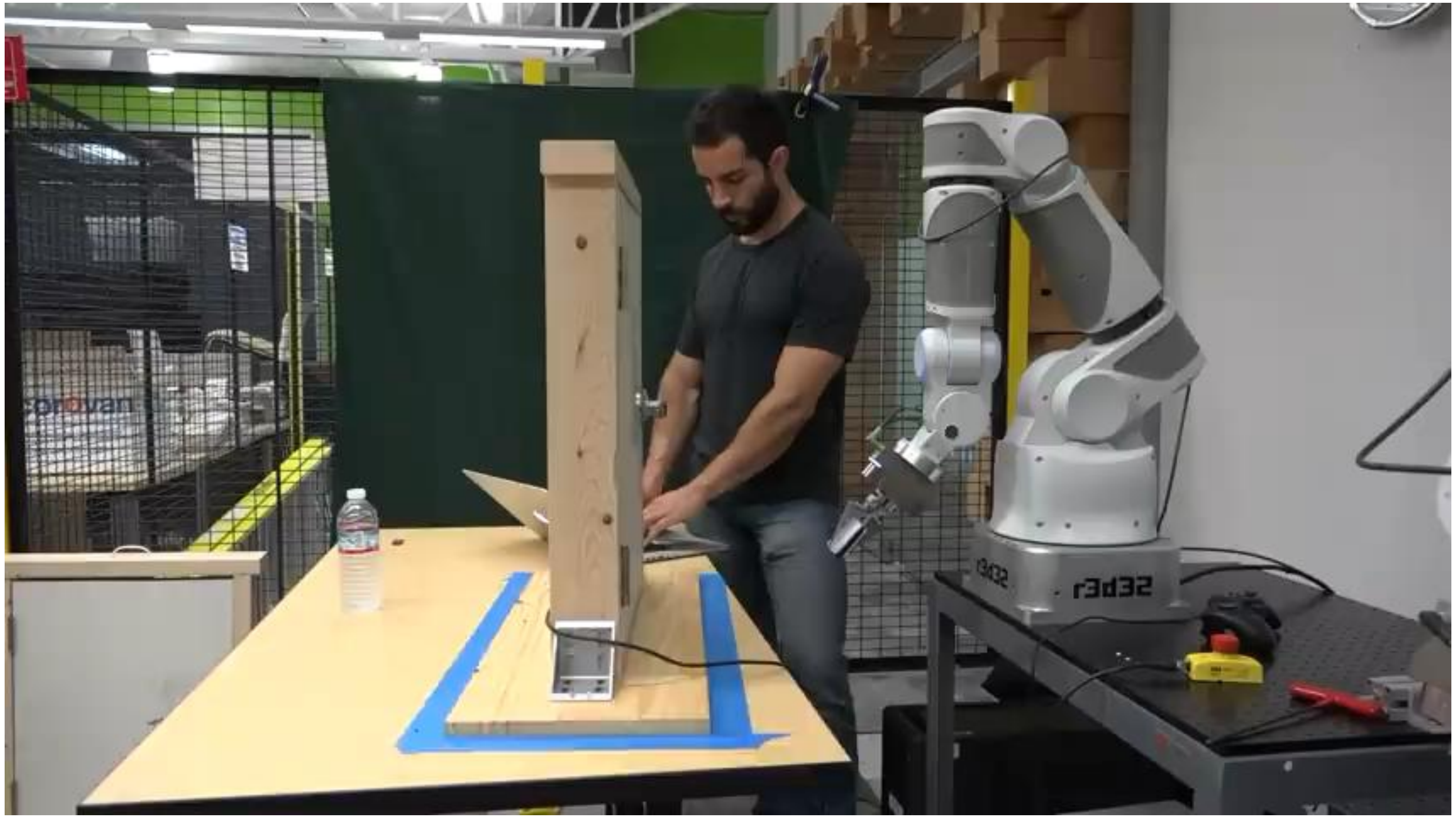
Goal: maximize the total reward

$$\max \sum_{t=1}^{T} r_t$$

Problem: starting with s_1, what is the optimal action trajectory $a_1, \dots a_T$?

Solutions?

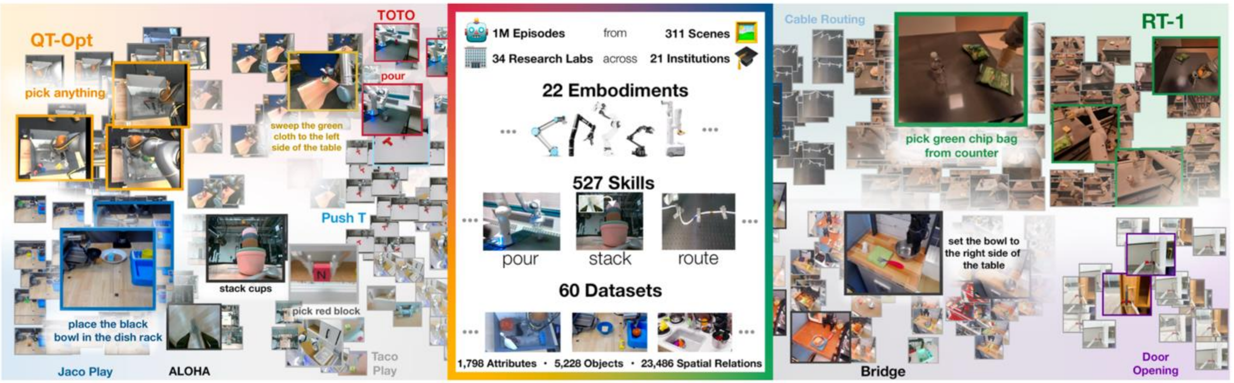
1. Memorize 1 demonstration and replay



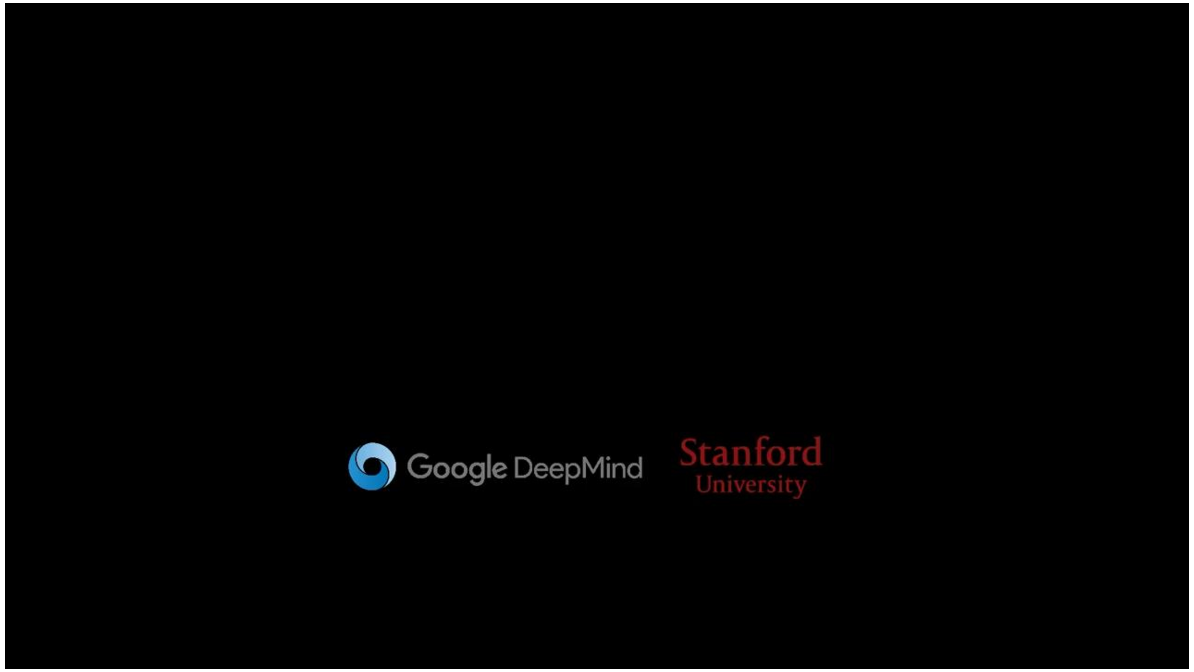
https://youtu.be/9-9udQtO1PY?si=DBV3Zjsv6XZJjgwN

Solutions?

1. Memorize 1 demonstration and replay
2. Imitation learning by fitting many demonstrations



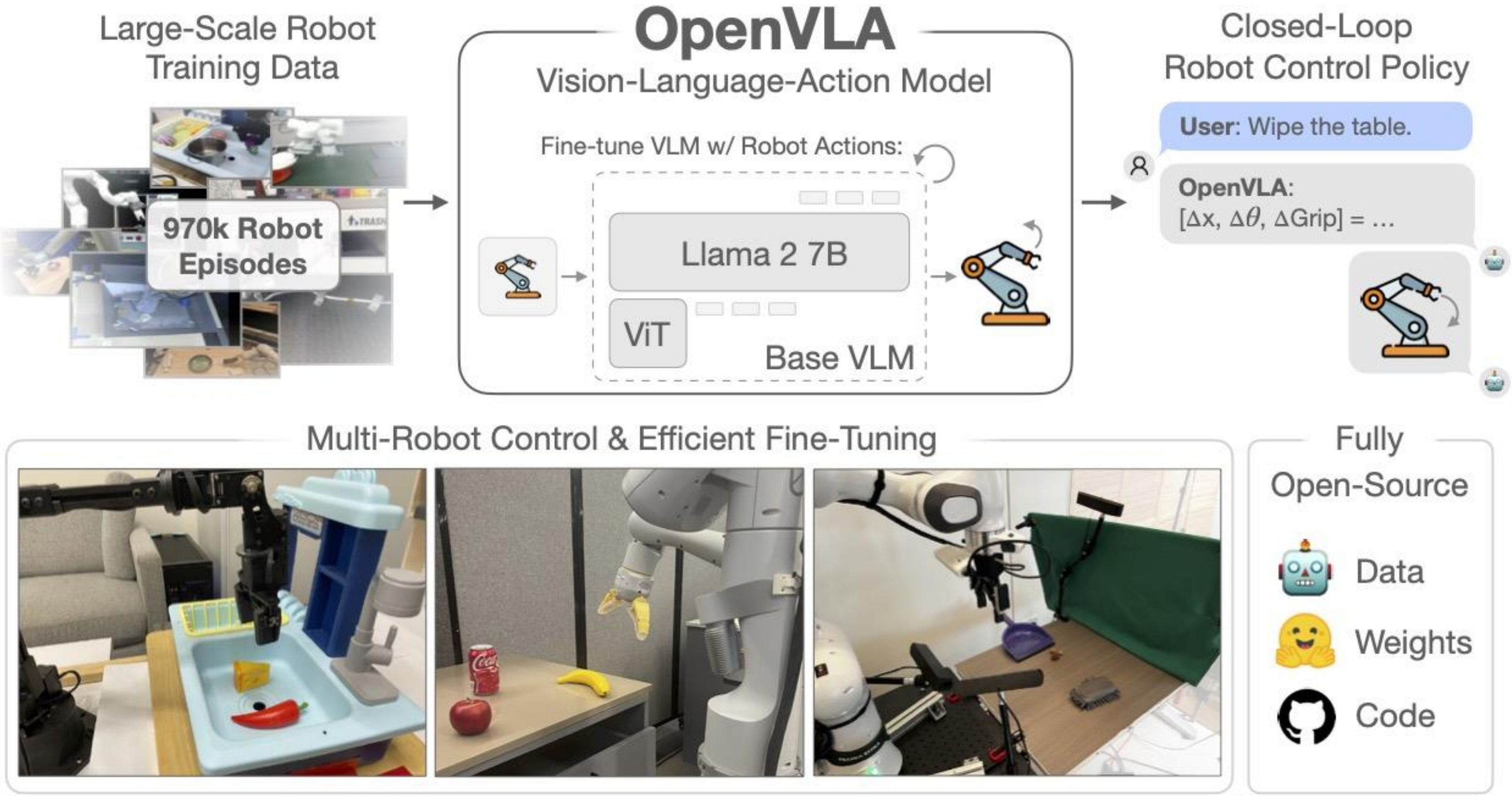
Open X-Embodiment: Robotic Learning Datasets and RT-X Models



ALOHA 2: An Enhanced Low-Cost Hardware for Bimanual Teleoperation

Solutions?

1. Memorize 1 demonstration and replay
2. Imitation learning by fitting many demonstrations with large models



OpenVLA: An Open-Source Vision-Language-Action Model

Solutions?

1. Memorize 1 demonstration and replay
2. Imitation learning by fitting many demonstrations

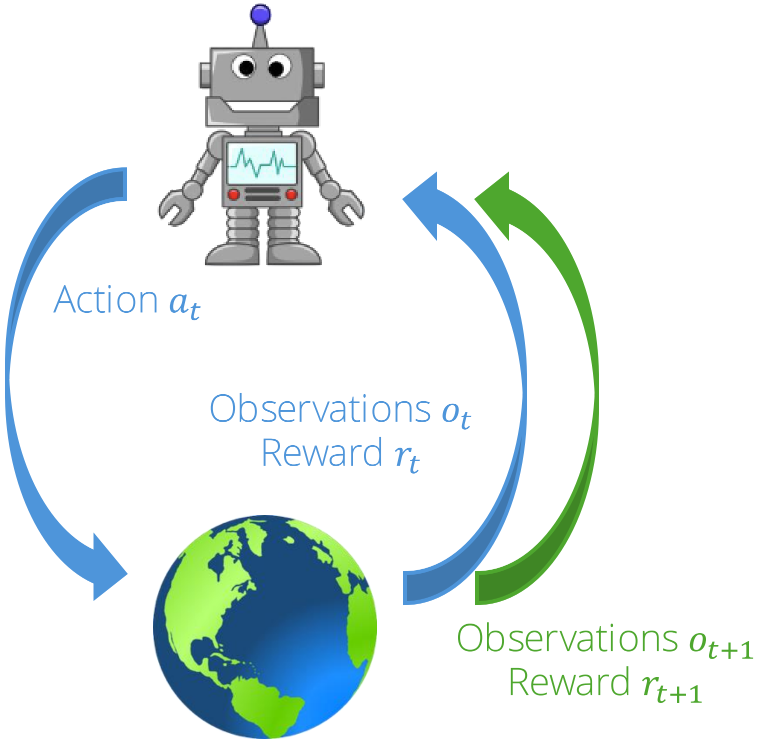
Succeed in the training domain

Fail in the testing domain

RT-1: Robotics Transformer for Real-World Control at Scale

Solutions?

1. Memorize 1 demonstration and replay
2. Imitation learning by fitting many demonstrations
3. Control trajectory optimization if the dynamic rules is known



Known environment constraints:

$$s_{t+1} = f(s_t, a_t)$$

Optimization problem:

$$a_1^*, a_2^*, \dots, a_T^* = \underset{a_1,\dots,a_T}{\operatorname{argmax}} \sum_{t=1}^{T} r_t$$

with

$$s_1 = \bar{s}_1$$
$$s_{t+1} = f(s_t, a_t)$$

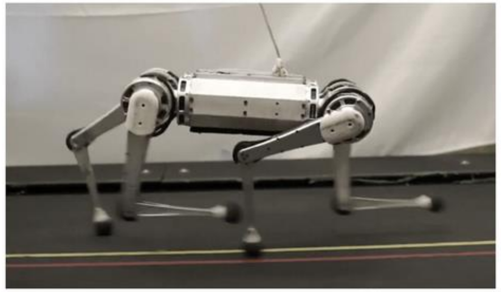
Solutions?

1. Memorize 1 demonstration and replay
2. Imitation learning by fitting many demonstrations
3. Control trajectory optimization if the dynamic rules is known

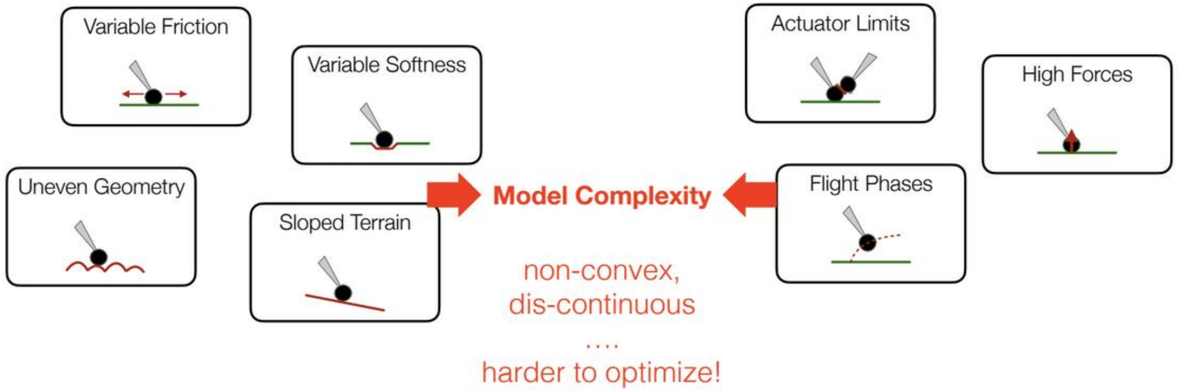
In-the-Wild Locomotion

Lee et al. 2020

Fast Locomotion



Kim et al. 2019

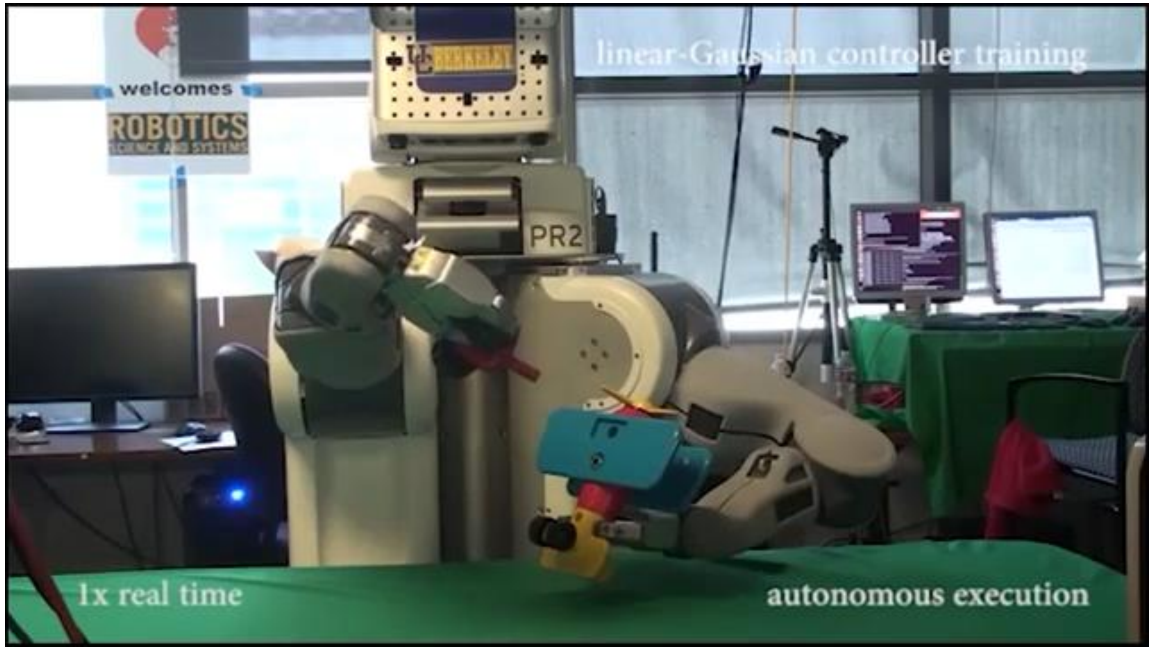


Slide credit P. Agrawal

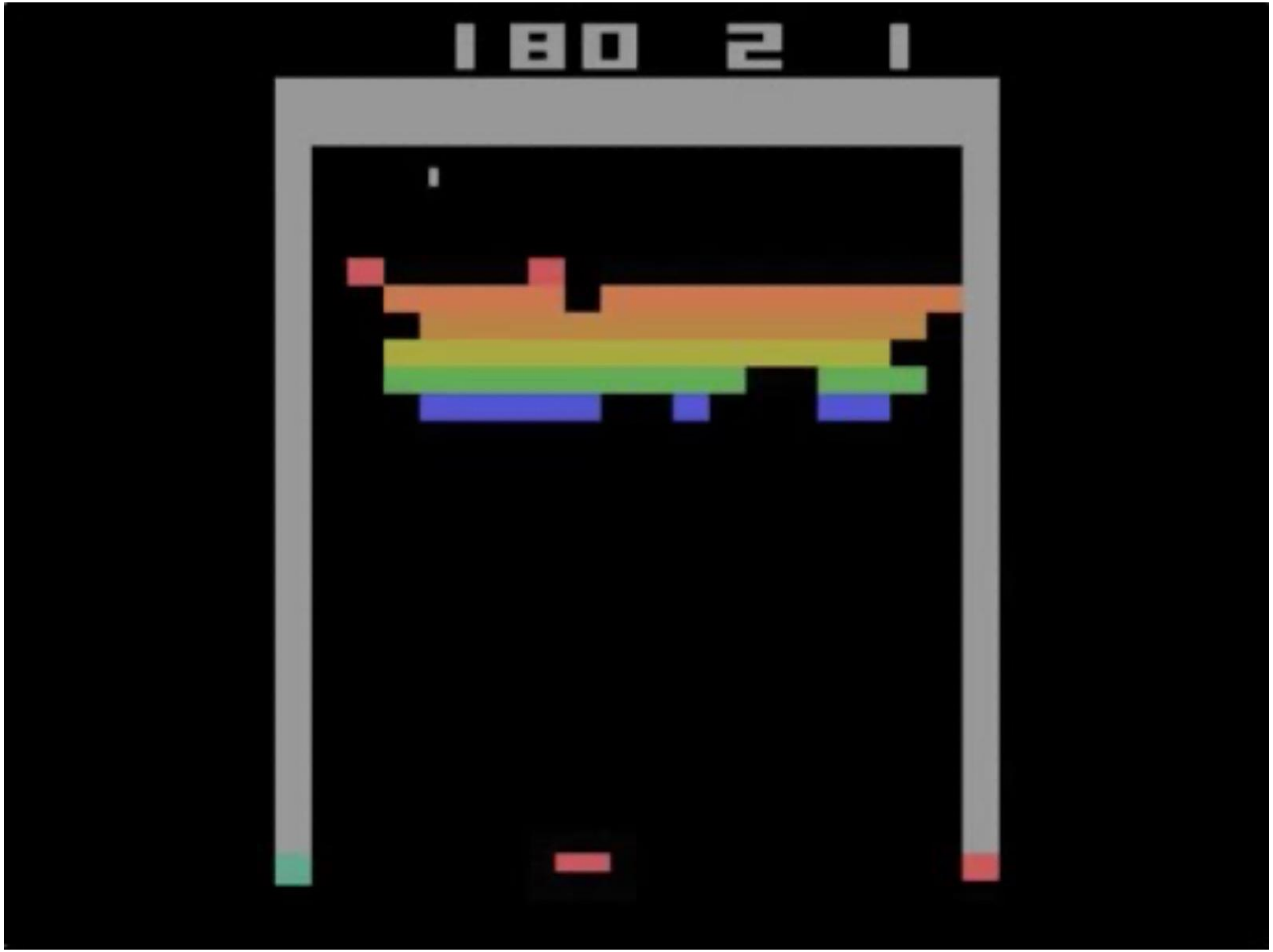
Solutions?

1. Memorize 1 demonstration and replay
2. Imitation learning by fitting many demonstrations
3. Control trajectory optimization if the dynamic rules is known
4. Reinforcement learning by trial and error

https://www.youtube.com/watch?v=kFPcoclV5Vo



https://www.youtube.com/watch?v=JeVppkoloXs&t=11s

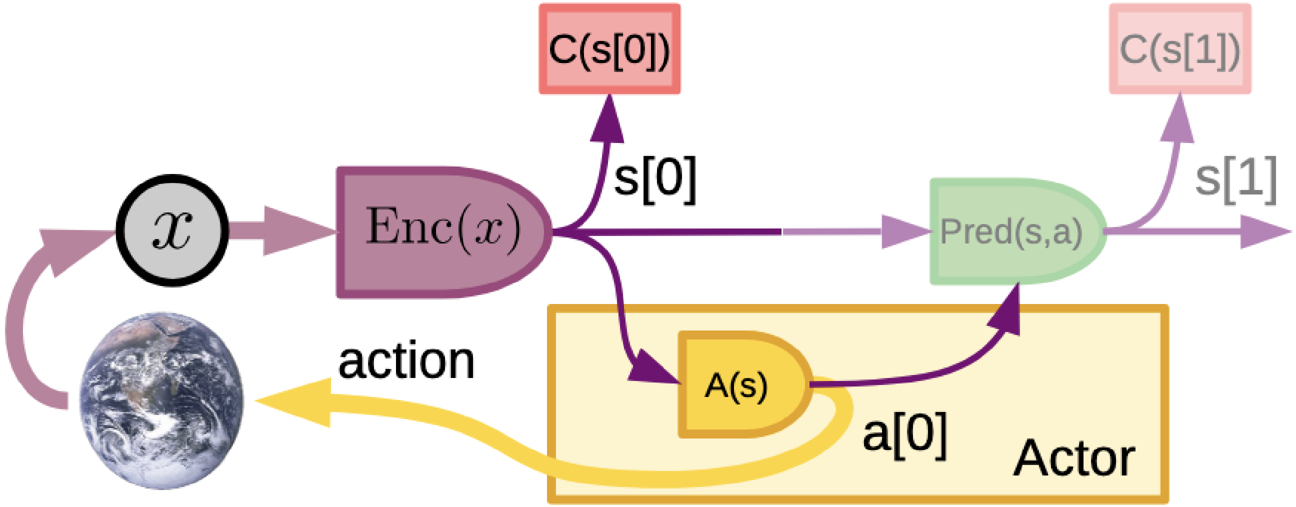

180 2 1

Model-free RL is Sample Inefficient…

https://youtu.be/DcYLT37ImBY?si=dEan3vOdd4fLIbHG

Solutions?

1. Memorize 1 demonstration and replay
2. Imitation learning by fitting many demonstrations
3. Control trajectory optimization if the dynamic rules is known
4. Reinforcement learning by trial and error
5. Reinforcement learning by trial and error + learning a world simulator from experience

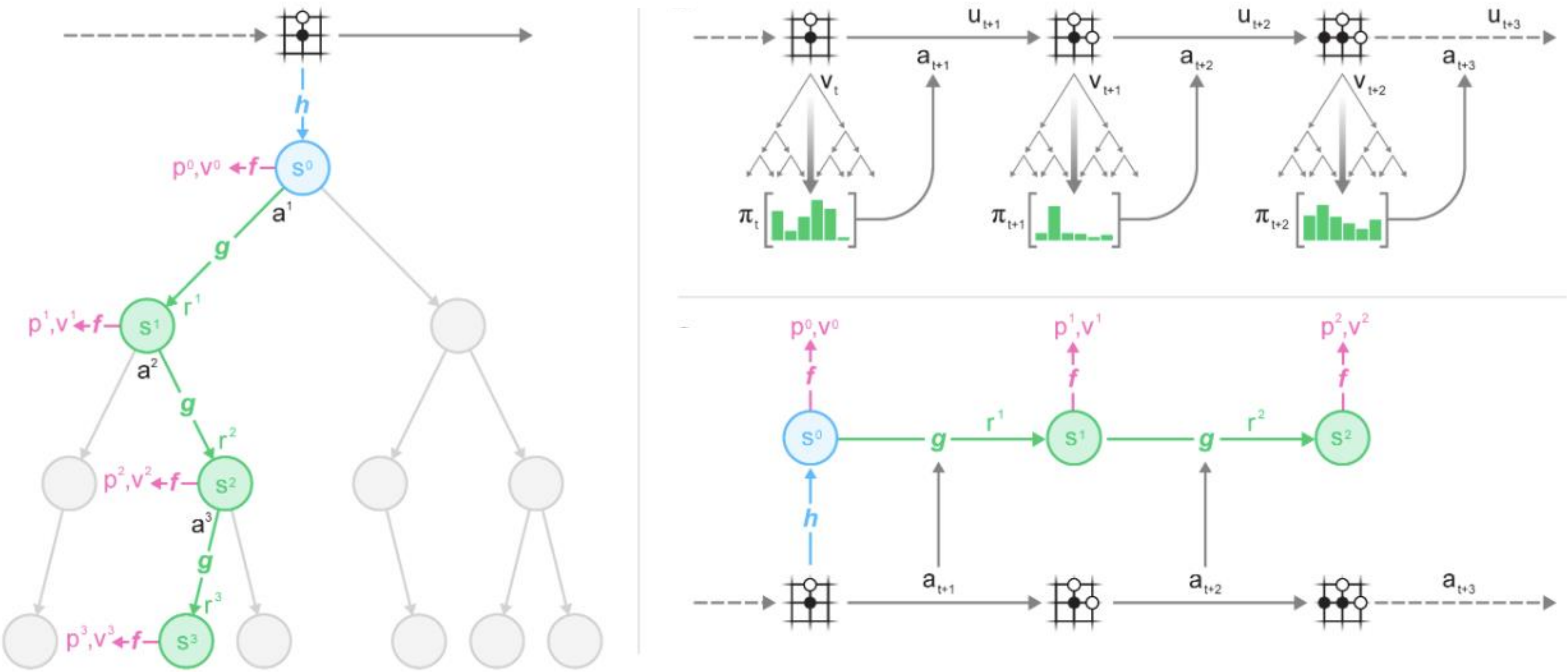


$$\min \| f(s,a) - pred(s,a) \|$$

Model-based RL is More Sample Efficient

A. Plan to obtain optimal actions by simulating outcomes of different actions

B. Interact with the environment with optimal actions



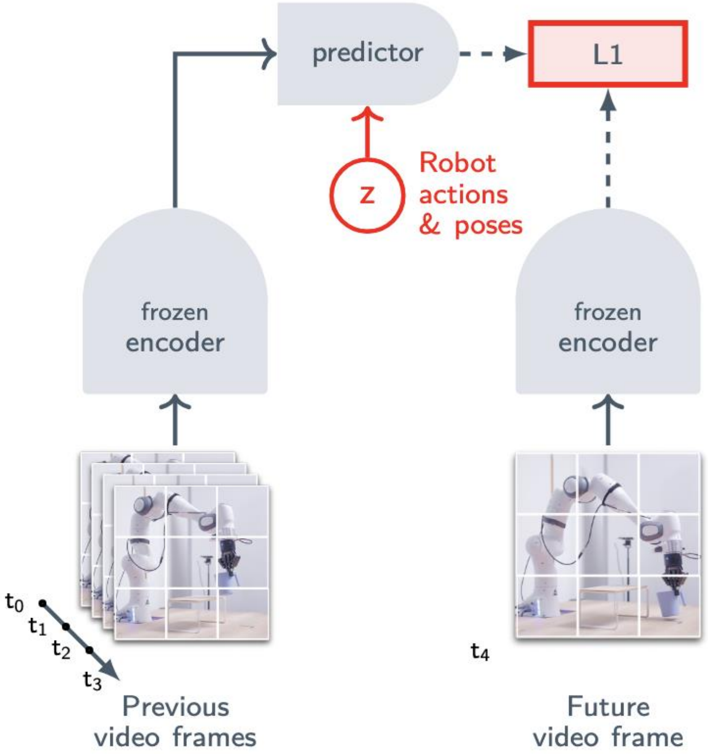
C. Learn the simulator through experience

Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model. DeepMInd.

What Simulators to Learn?

Video generative model

SORA by OpenAI



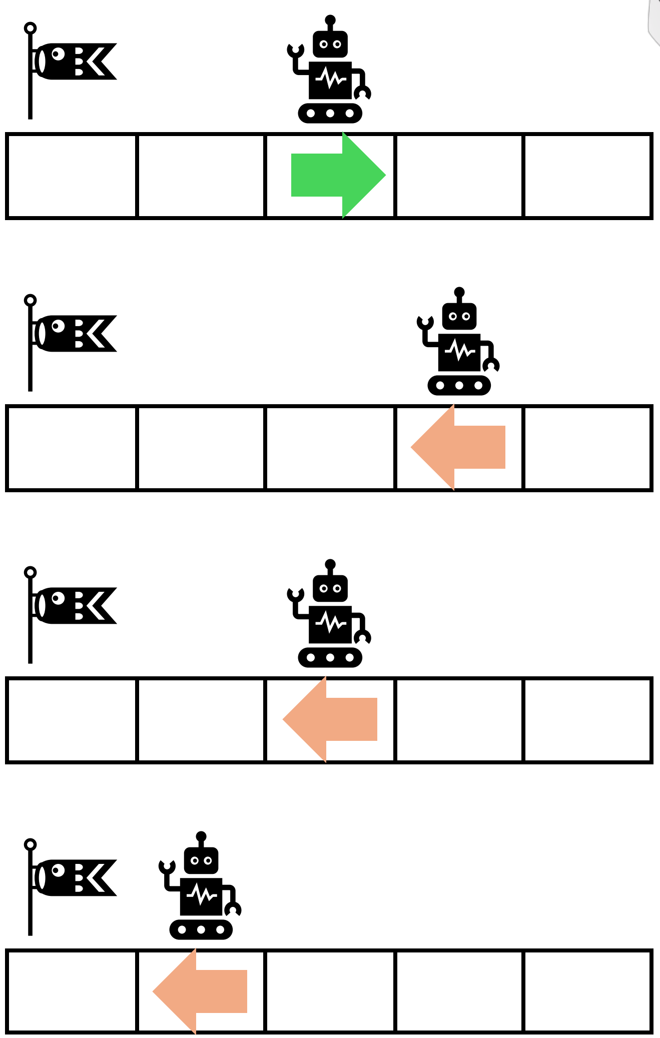
4D generative model

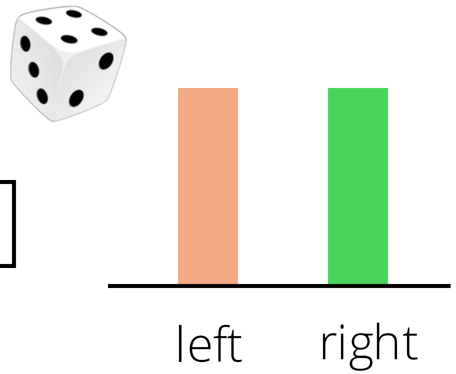
Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis

V-JEPA 2: Self-Supervised Video Models Enable Understanding, Prediction and Planning

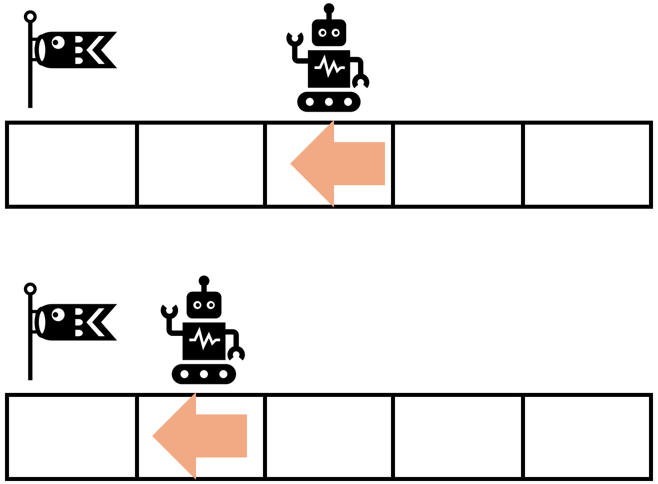
But RL Has More Issues…

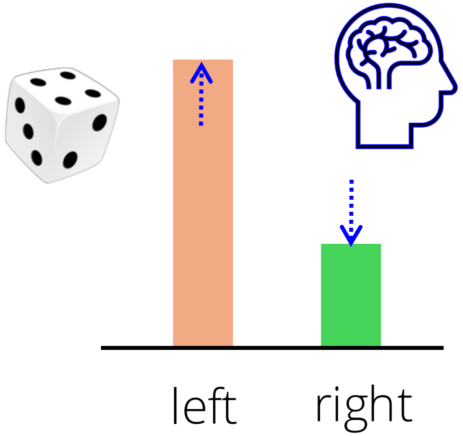
RL Exploration is Ineffective





Wasting steps
on obvious
useless actions



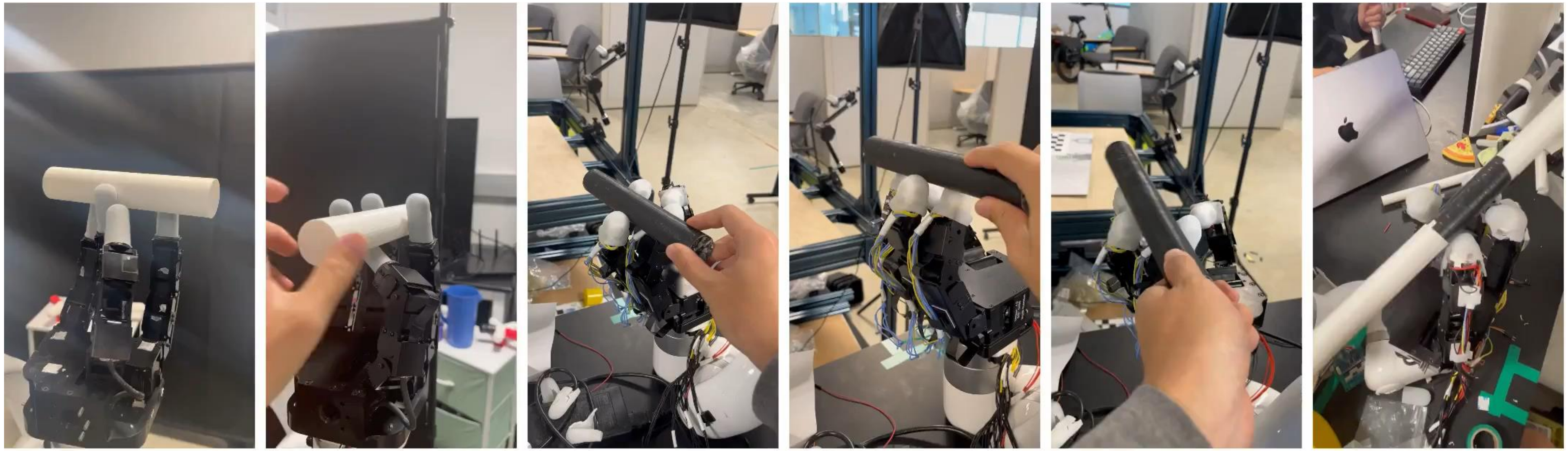


The target is
on the left

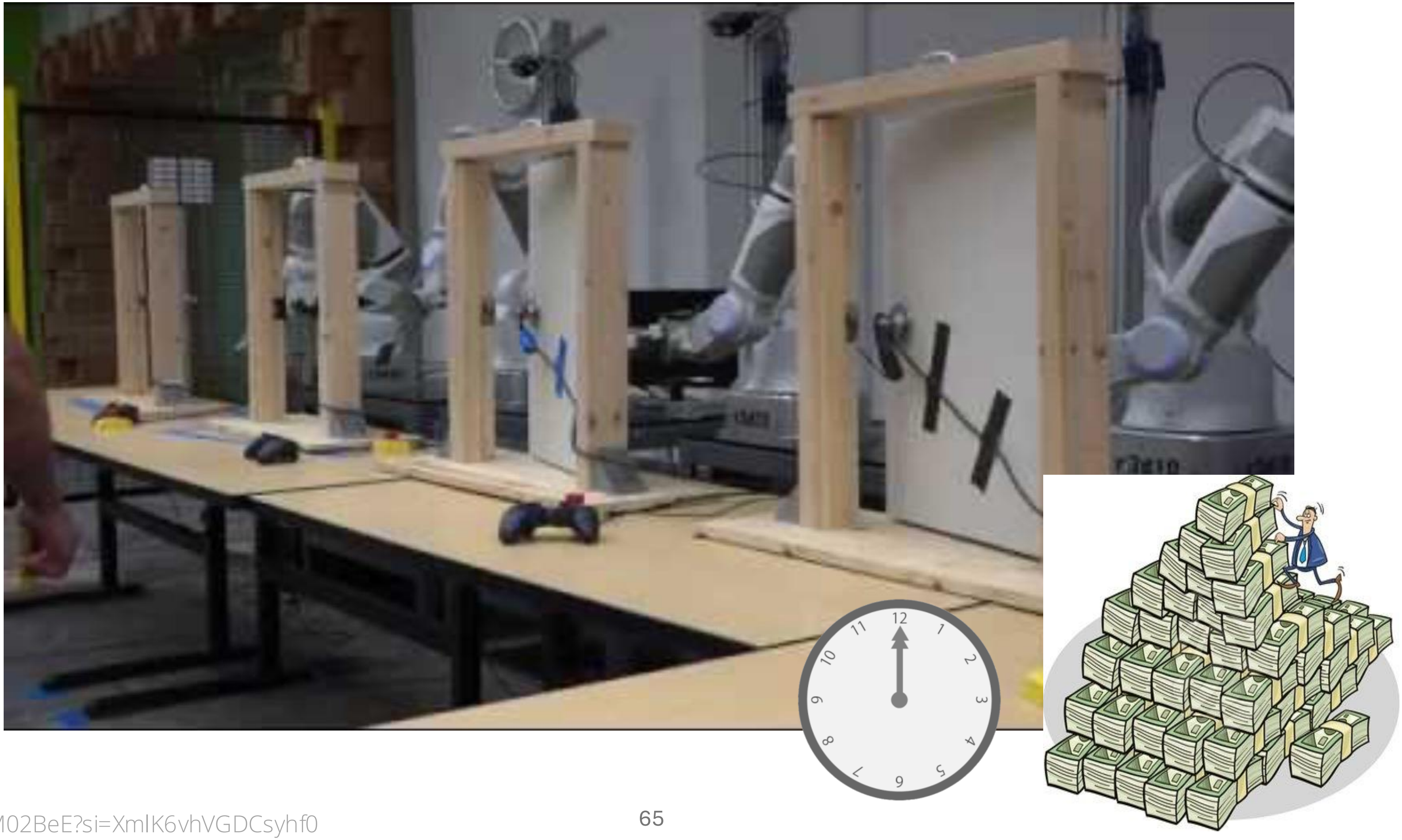
Performing RL in Simulator Has Sim2Real Gap

Learn in Sim

Deploy in Real World



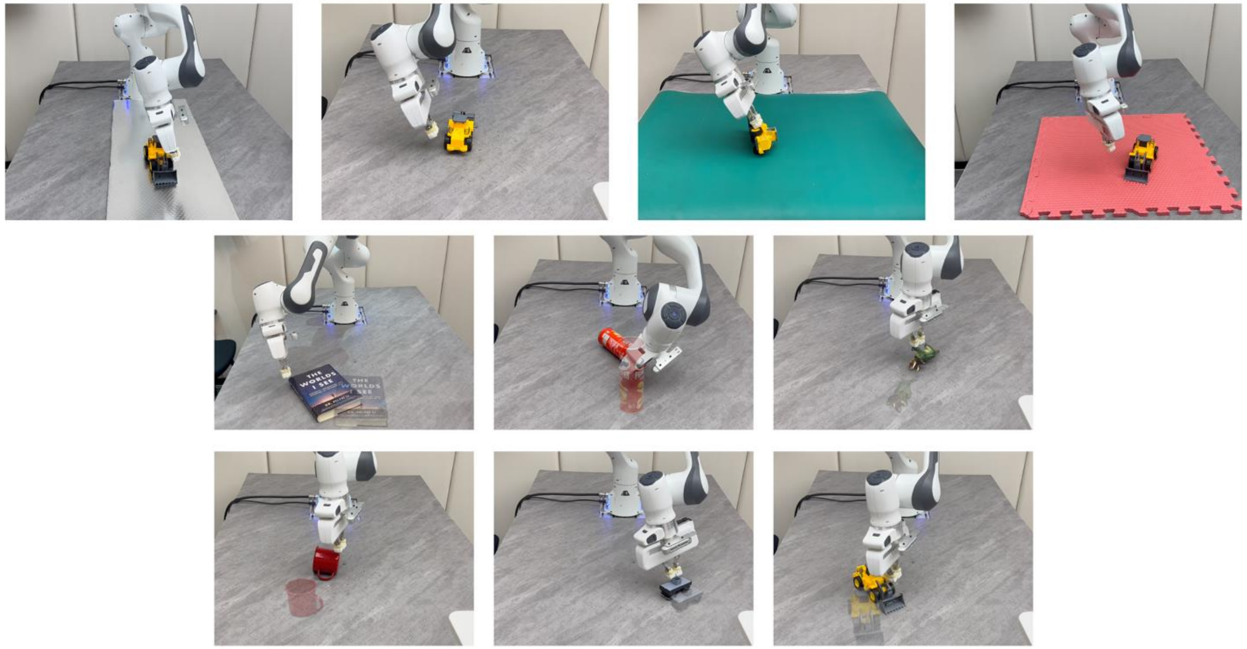
But Real-World RL is Unsolved



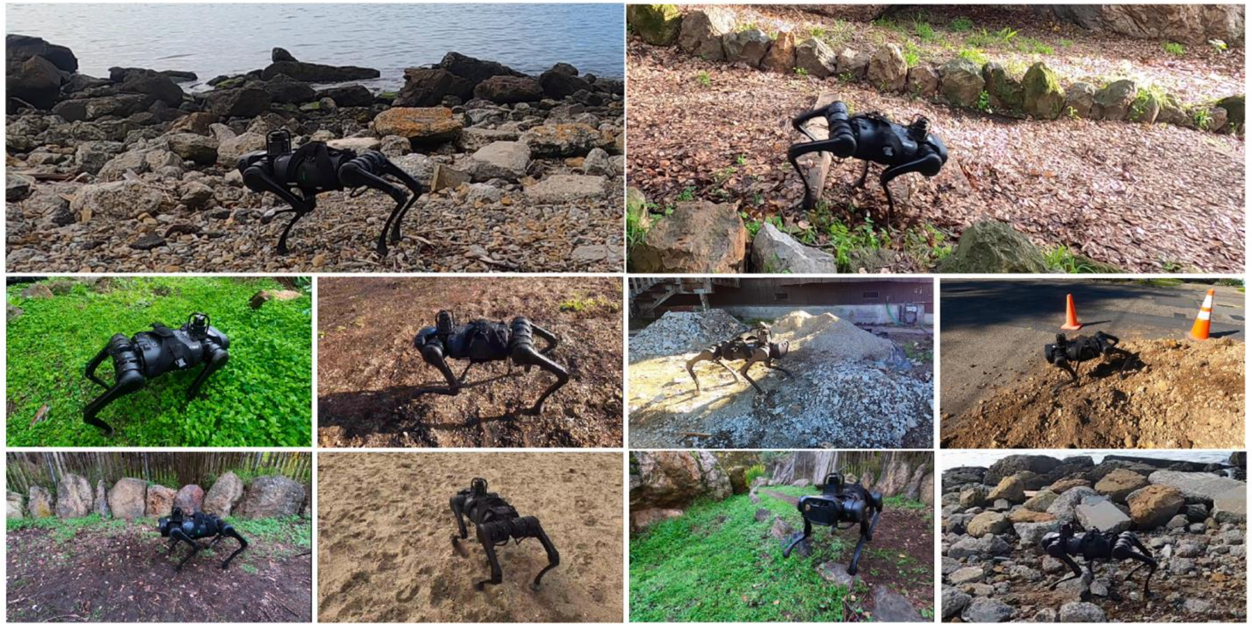
Safety is Essential for Real-world RL

https://youtube.com/shorts/RKUeRVTxzjY?si=BpQ7Fez72DkKfUDf

Even Trained with RL, Policies are not Robust to Varying Dynamics

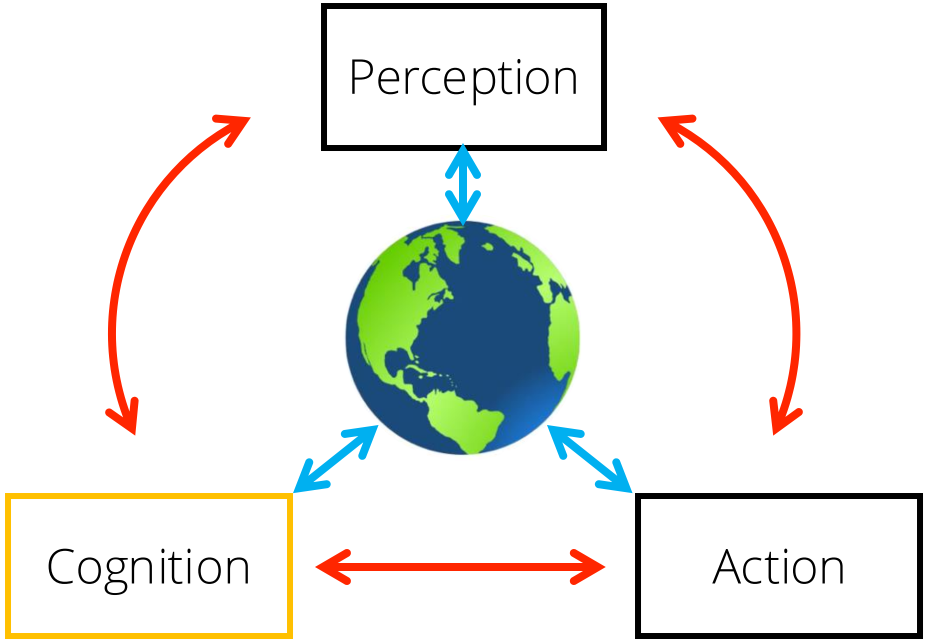


DyWA: Dynamics-adaptive World Action Model
for Generalizable Non-prehensile Manipulation



RMA: Rapid Motor Adaptation for Legged Robots. Kumar et al.

We will Focus on the Action and Cognition in this Course



In simplification, we want to answer questions: “What are our goals?” and “What action should we take to achieve our goals in the current environment?”.

What Are the Goals? Who Decides the Goals?

Action: muscle contractions
Rewards: food, attentions
and ?

Image credit S. Levine

Action: muscle contractions
Rewards: food and ?

https://www.salon.com/2021/01/16/human-breeding-of-cats-has-made-them-look-like-they-are-always-in-pain/

Action: steering, acceleration
Rewards: win

https://racingnews365.com/championship-standings-after-2024-f1-australian-grand-prix

Action: selecting companies,
buying, selling
Rewards: cash and ?

Action: (x, y) location
Rewards: win

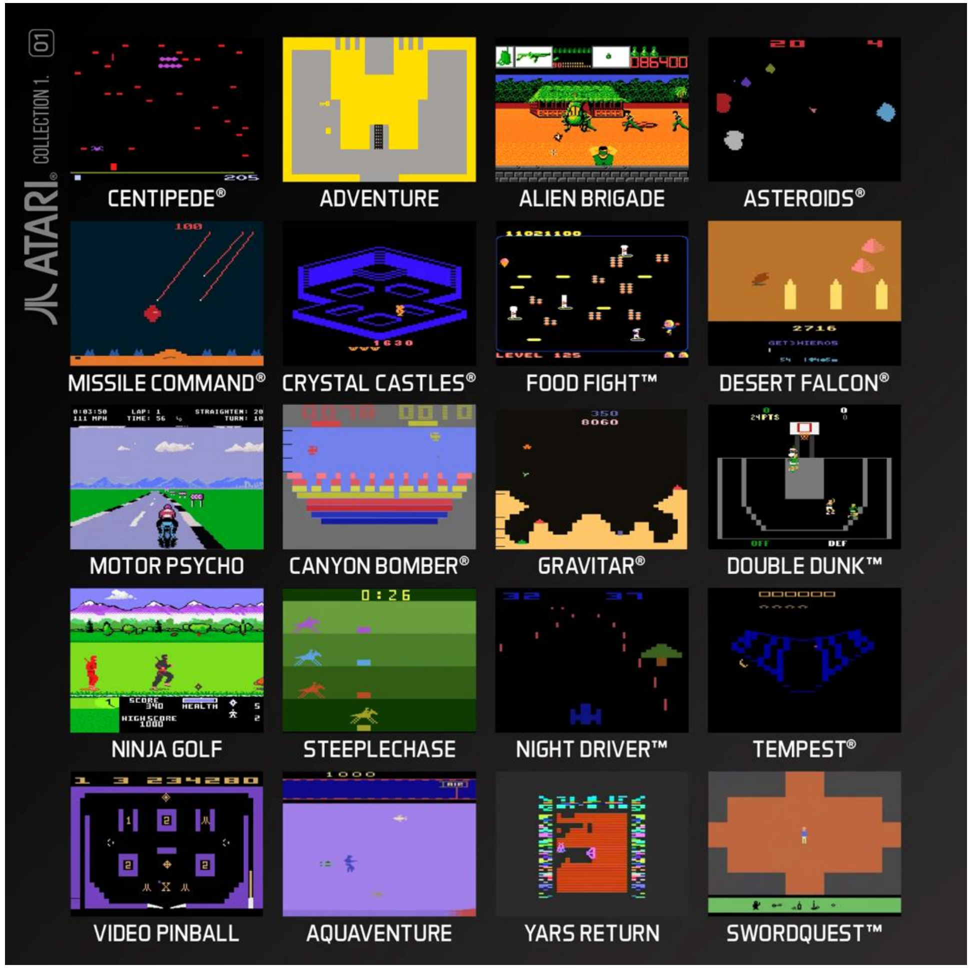
https://en.wikipedia.org/wiki/Go_%28game%29

Action: jump, ↑←→↓
Rewards: win

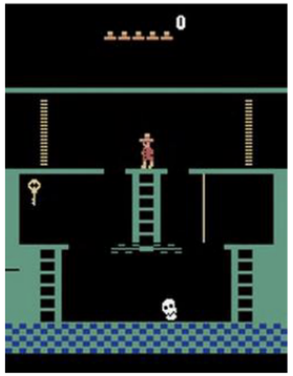
Image credit Nintendo

Action: RGB at each pixel
Rewards: cash, aesthetics
and ?

Tasks with Sparse Rewards May be Difficult…



Montezuma's revenge

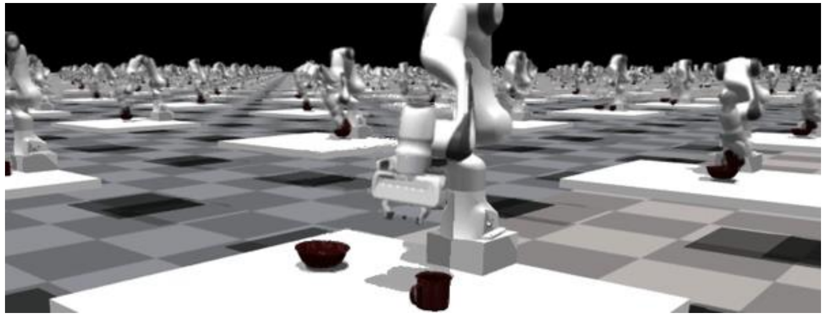


Slide credit S. Levine

- Getting key = reward
- Opening door = reward
- Getting killed by skull = nothing (is it good? bad?)
- Finishing the game only weakly correlates with rewarding events
- We know what to do because we **understand** what these sprites mean!

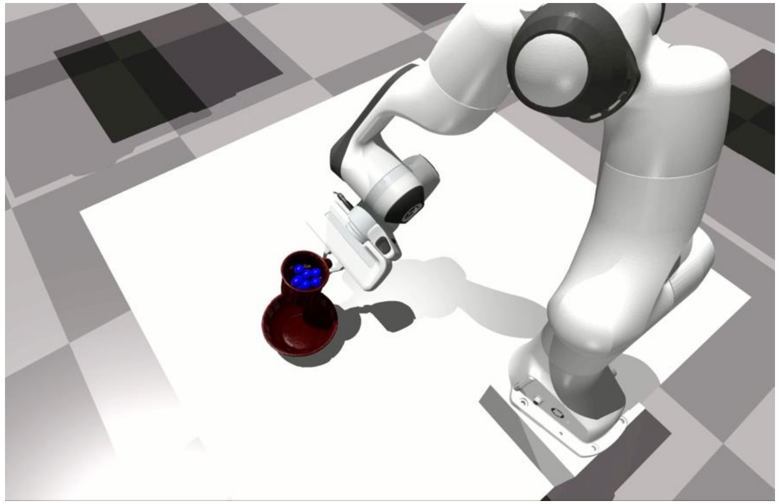
This is also called "credit assignment problem"—which actions attribute to the final success

Tasks with Sparse Rewards May be Difficult…



Policy fails when optimizing with sparse rewards

$$r = \begin{cases} 1, & \textit{if pour water in the bowl} \\ 0, & \textit{otherwise.} \end{cases}$$



Policy fails when optimizing with staged rewards

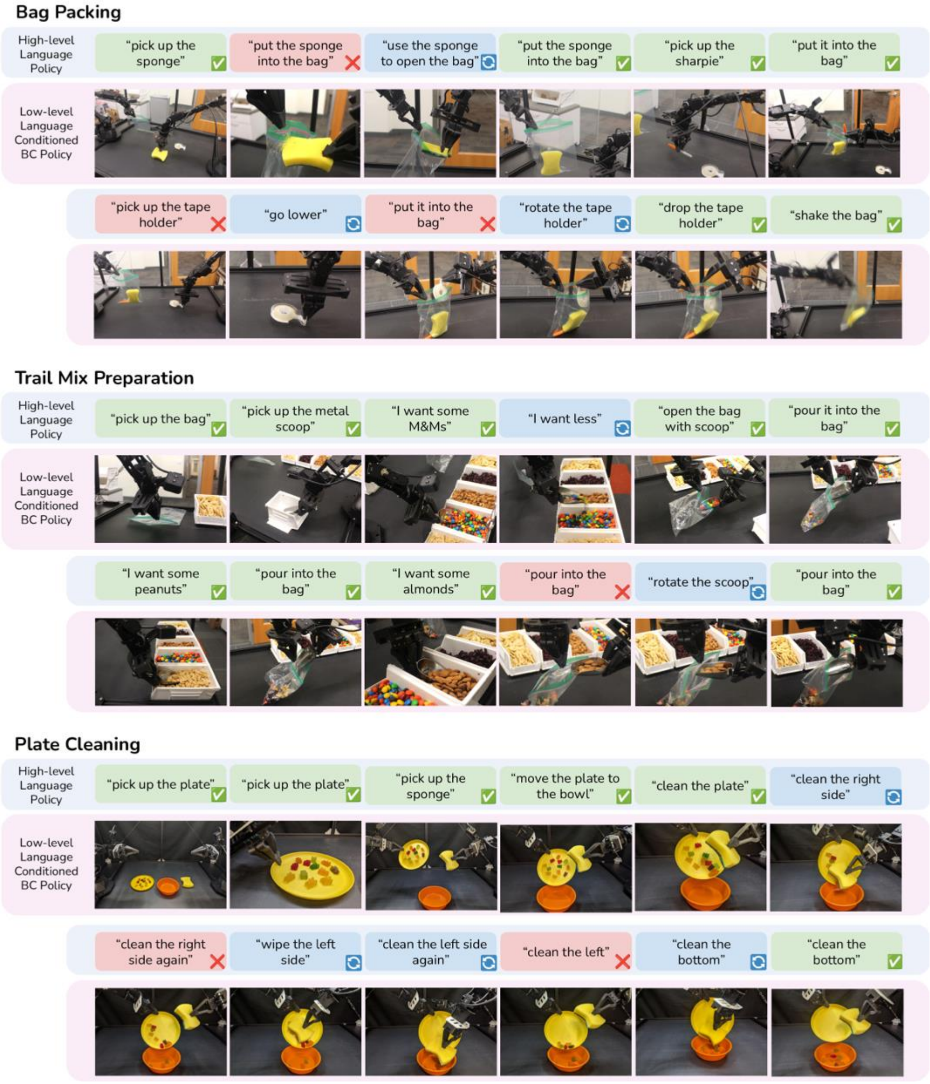
$$r_{reach} = \begin{cases} 1, & \textit{if reach the mug.} \\ 0, & \textit{otherwise.} \end{cases}$$

$$r_{lift} = \begin{cases} 1, & \textit{if lift up the mug.} \\ 0, & \textit{otherwise.} \end{cases}$$

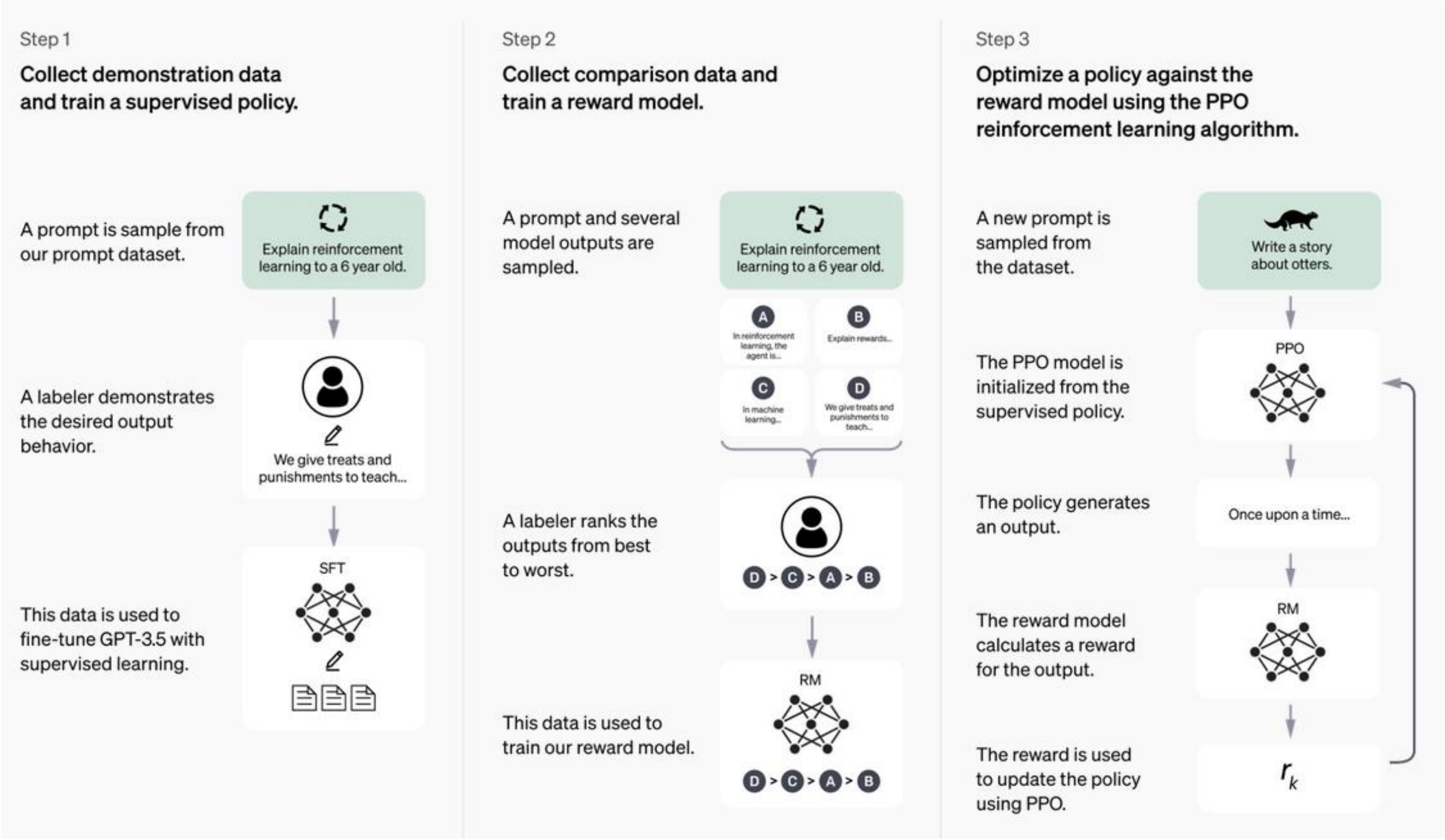
$$r_{pour} = \begin{cases} 1, & \textit{if pour the water in the bowl} \\ 0, & \textit{otherwise.} \end{cases}$$

How is it even possible to cook these dishes with only sparse rewards…

Define Goal with Prior Knowledge: Human Feedback

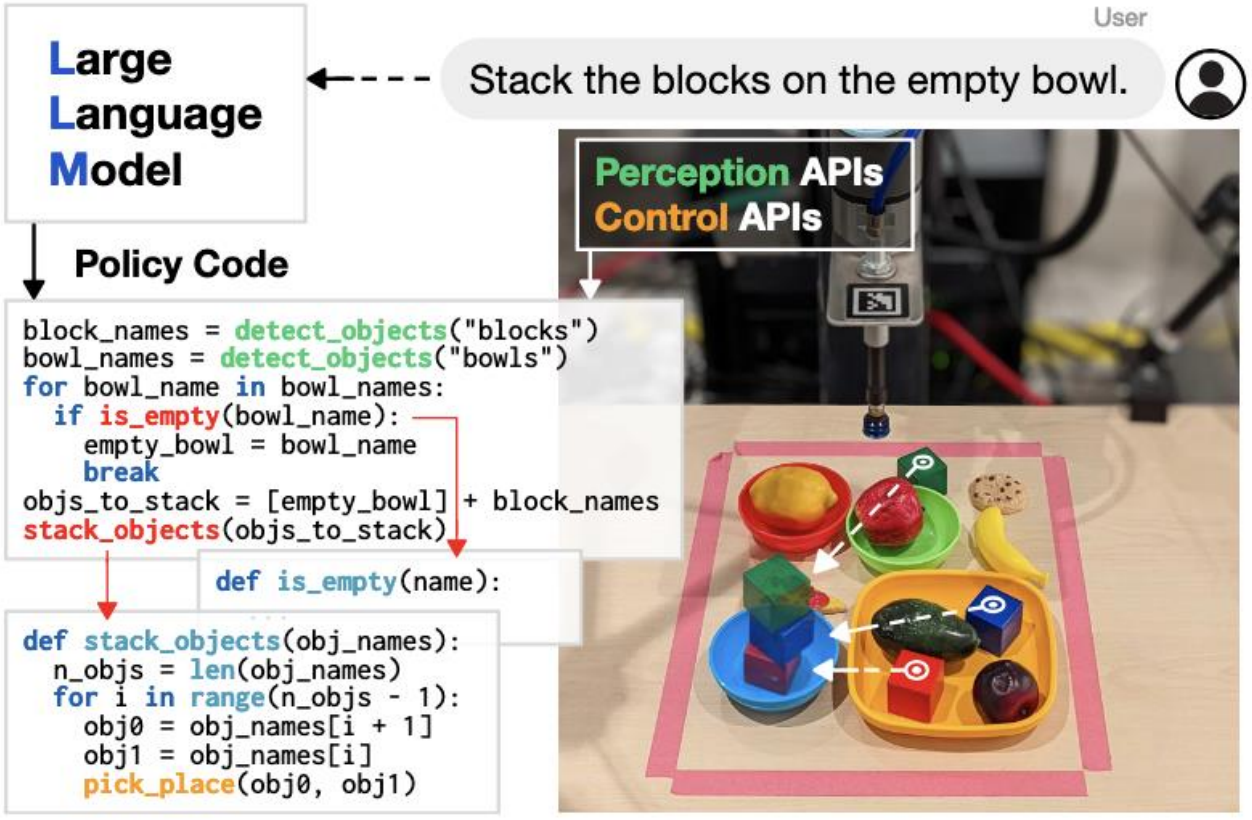


Yell at your robot. Shi et al.

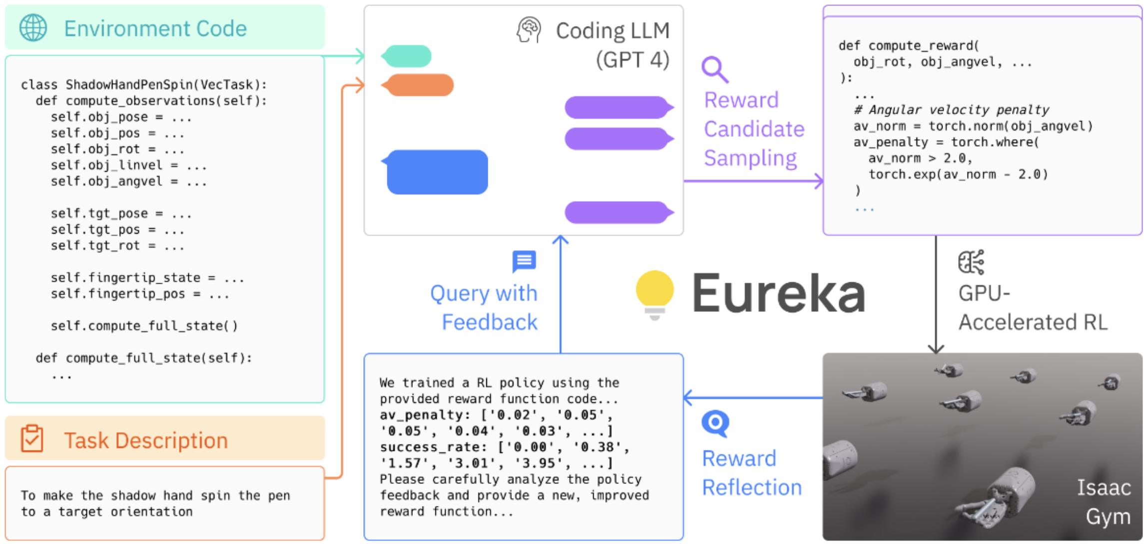


https://openai.com/index/chatgpt/

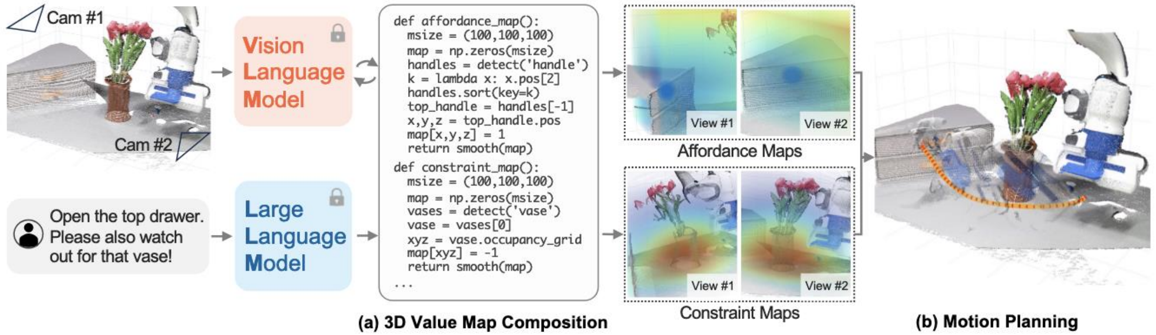
Define Goal with Prior Knowledge: Large Language Models



Code as Policies: Language Model Programs for Embodied Control. Liang et al. ICRA 2023



Eureka: Human-Level Reward Design via Coding Large Language Models. Ma et al



VoxPoser: Composable 3D Value Maps for Robotic Manipulation with Language Models. Huang et al

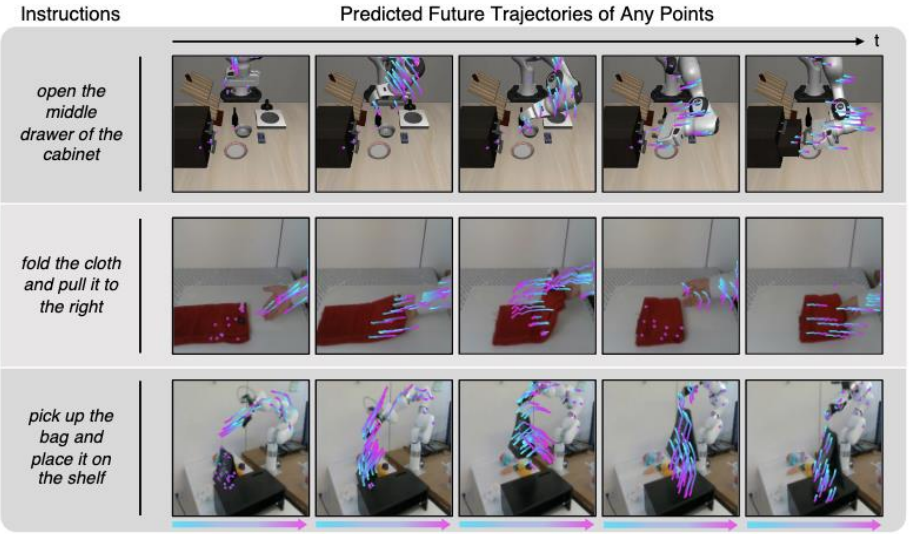
Define Goal with Prior Knowledge: Human Behaviors

Ego-Exo4D dataset

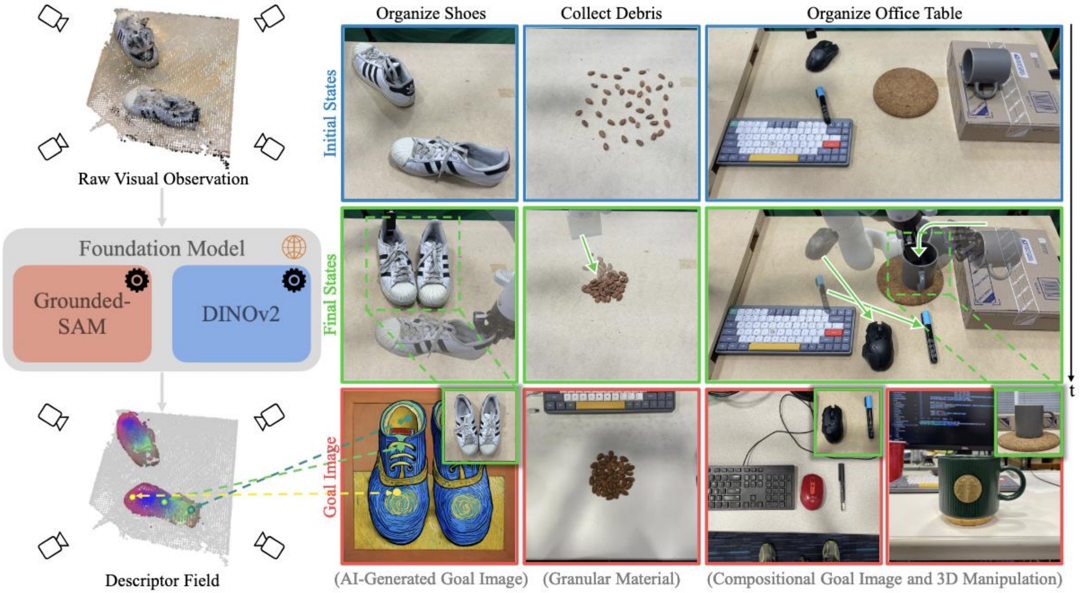
HOT3D dataset

How to Transfer the Knowledge from Human Demonstrations?

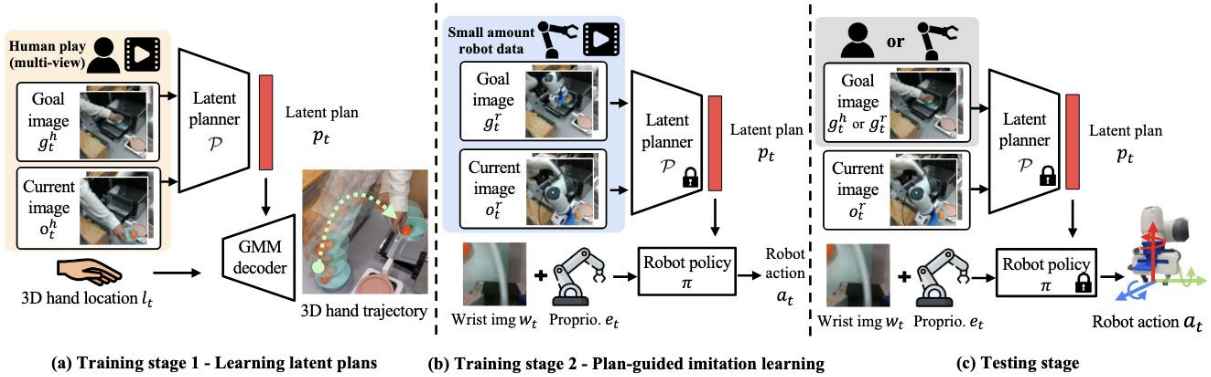
Affordances from Human Videos as a Versatile Representation for Robotics. Bahl et al.



Any-point Trajectory Modeling for Policy Learning. Wen et al.



D3Fields: Dynamic 3D Descriptor Fields for Zero-Shot Generalizable Robotic Manipulation. Wang et al.



MimicPlay: Long-Horizon Imitation Learning by Watching Human Play. Wang et al.

What Are the Goals? Who Decides the Goals?

Action: muscle contractions
Rewards: food, attentions
and ?

Image credit S. Levine

Action: muscle contractions
Rewards: food and ?

https://www.salon.com/2021/01/16/human-breeding-of-cats-has-made-them-look-like-they-are-always-in-pain/

Action: steering, acceleration
Rewards: win

https://racingnews365.com/championship-standings-after-2024-f1-australian-grand-prix

Action: selecting companies,
buying, selling
Rewards: cash and ?

Action: (x, y) location
Rewards: win

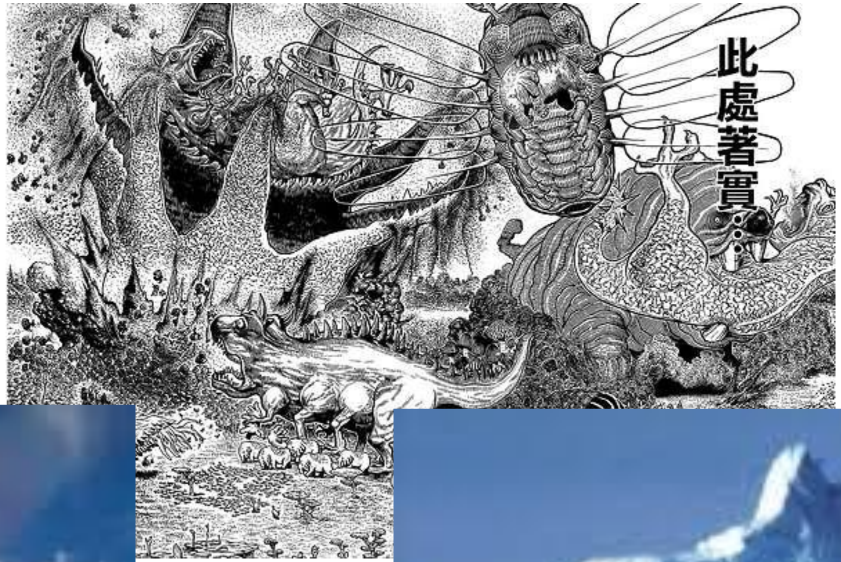
https://en.wikipedia.org/wiki/Go_%28game%29

Action: jump, ↑←→↓
Rewards: win

Image credit Nintendo

Action: RGB at each pixel
Rewards: cash, aesthetics
and ?

Probably Survival is One Common Goal for All Creatures



Curiosity May also be an Important Goal

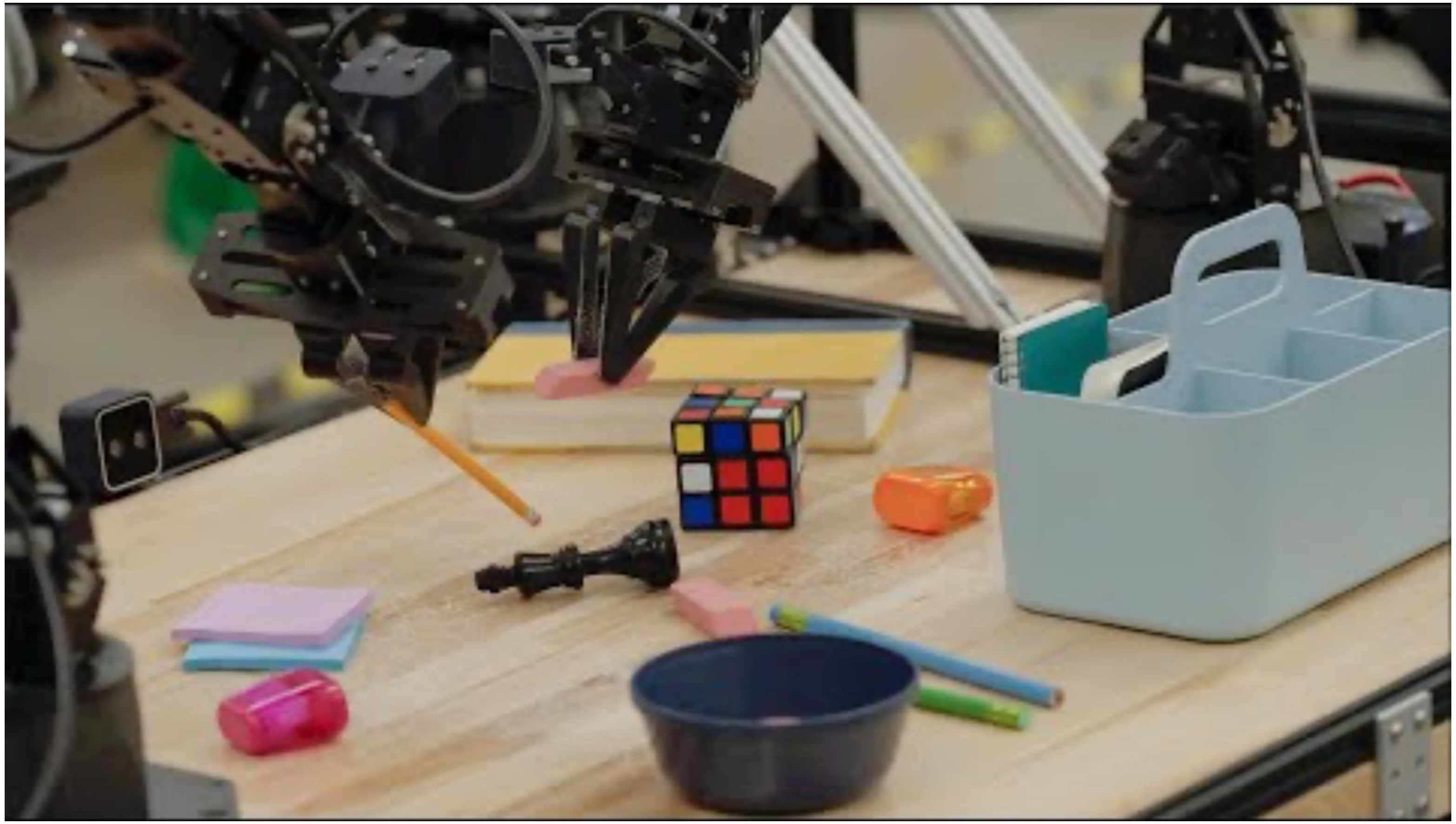
https://youtu.be/8vNxjwt2AqY?si=FKb6wXKsgXvP0j2T

Learning from Babies

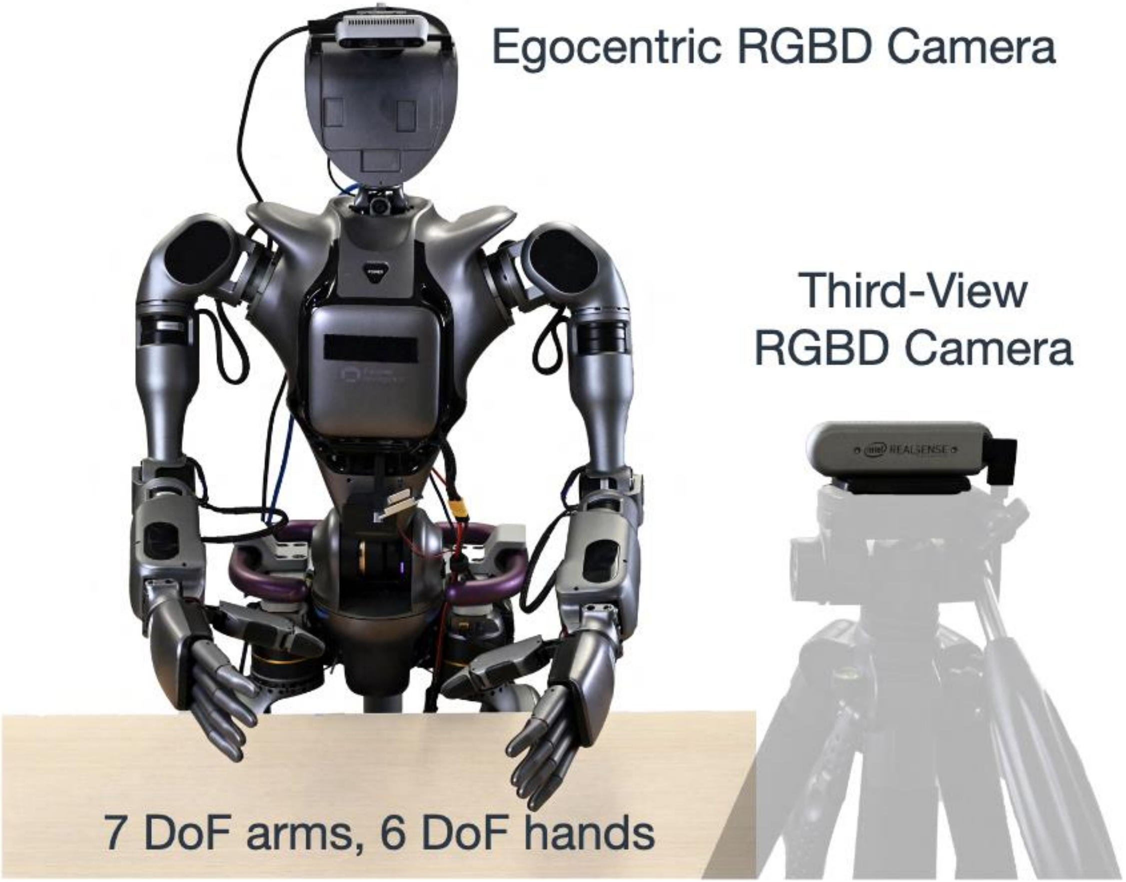
- *Be multi-modal*
- *Be incremental*
- *Be physical*
- *Explore*
- *Be social*
- *Learn a language*

What Kind of Robots Do We Need?

Jaw-Gripper Robots are Simple but Imcompetent



To Perform Daily Tasks, We Need Multi-Arm Multi-Fingered Robots…

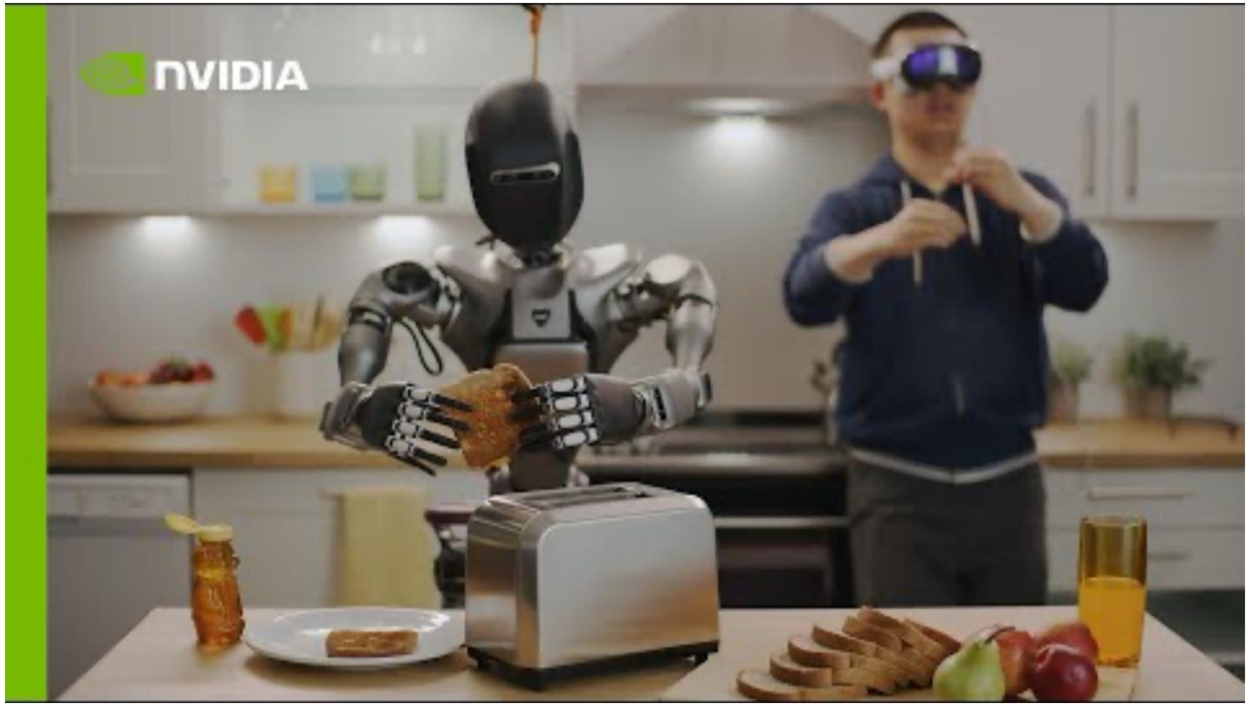


Sim-to-Real Reinforcement Learning for Vision-Based Dexterous Manipulation on Humanoids. Lin et al.

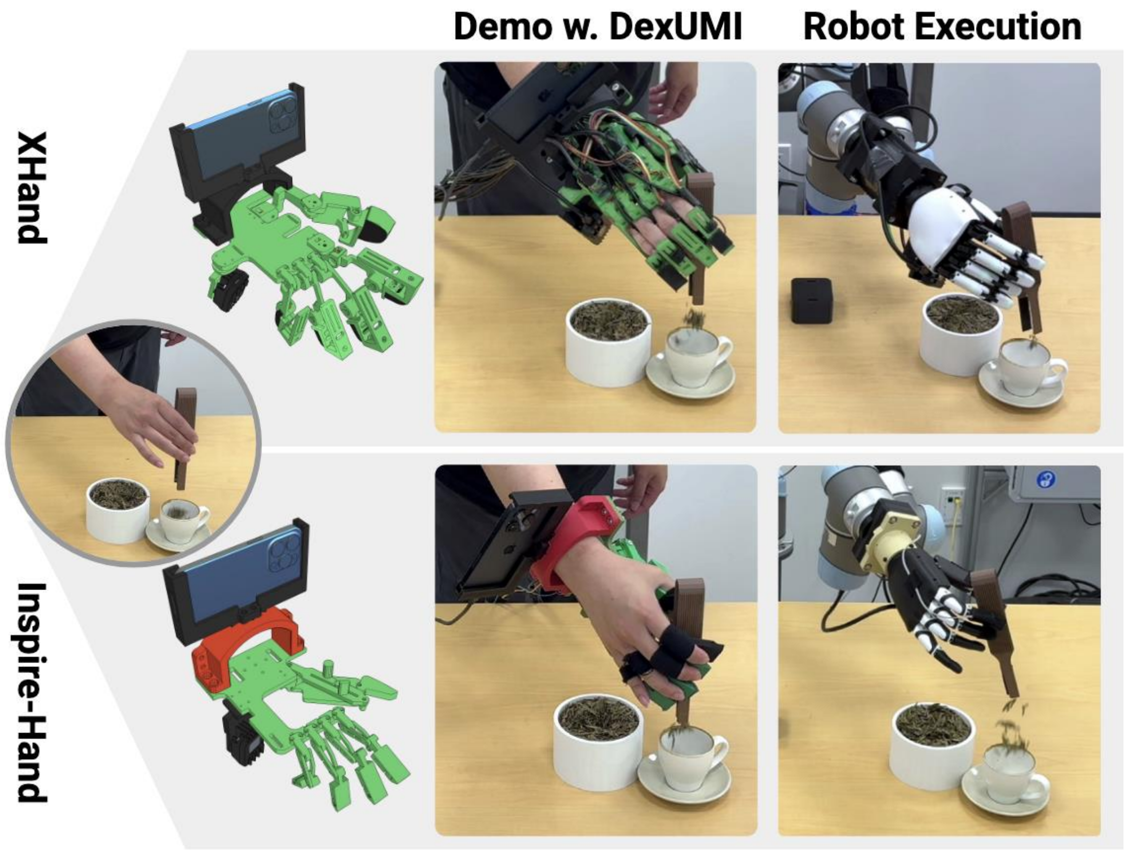
To Perform Daily Tasks, We Need Multi-Arm Multi-Fingered Mobile Robots...

https://youtu.be/HYwekersccY?si=QwZhtTnHWaj3oUcf

Collecting Expert Demonstrations for Multi-arm Multi-Fingered Robots is Expensive...

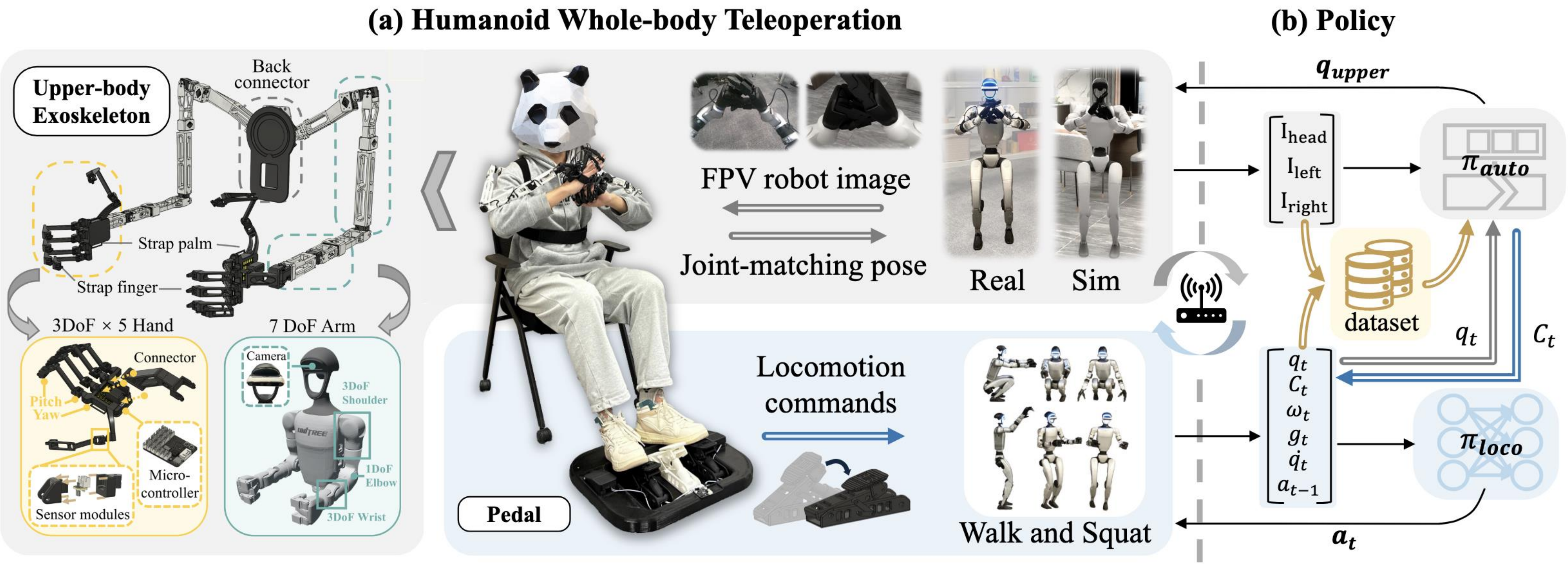


https://youtu.be/Bhg3uOx9ZPw?si=et7L0endzGvGPJz-



DexUMI: Using Human Hand as the Universal Manipulation Interface for Dexterous Manipulation

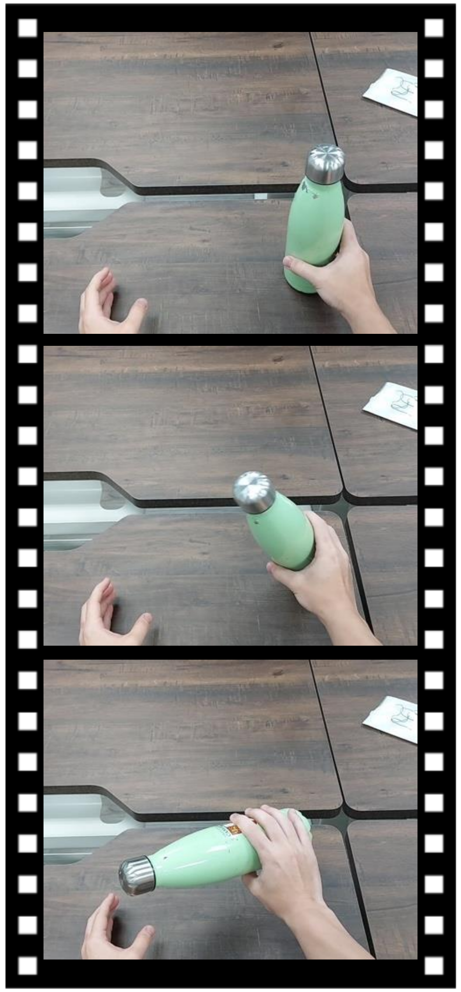
Collecting Expert Demonstrations for Multi-arm Multi-Fingered Mobile Robots is Even More Expensive…



RL Don't Work Out-of-Box with High-Dimensional Actions

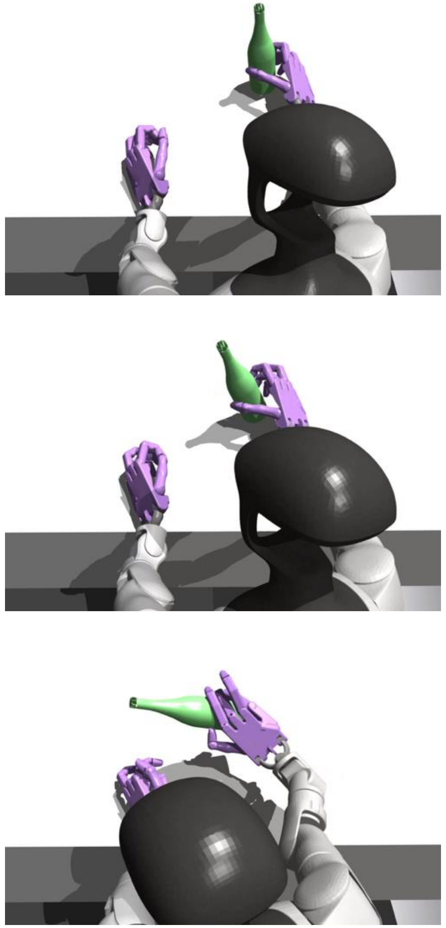
	Fingertip Mapping	Vanilla RL	Ours
Box	13.1±1.8	20.4±1.8	**69.8**±6.6
Micro.	60.6±2.7	56.1±9.8	**100**±0.0
Laptop	9.2±0.5	8.6±1.4	**76.7**±4.1
Coffee.	8.2±0.6	9.2±3.1	**74.8**±3.8
Mixer	8.3±2.1	10.8±3.2	**82.8**±2.1
Notebook	4.5±0.1	4.5±0.3	**64.3**±8.4

Training Multi-arm Multi-Fingered Robots with Human Motion Priors



Input: RGB human videos

Extract motion priors and task rewards



Learn in digital twins

Training Humanoid Robots with Human Motion Priors

What we hope you to know after this class?

From Seeing to Acting, What are Under the Hood?

Optimus by Tesla

- What are the goals? How to obtain the goal?
- How to achieve the goal with low-level actions?
- What are the essential components to control robots?

- What kind of robots do we need?
- What are the limitations of SOTA robots?
- What are the open challenges to build a general robots?

Reminder

- If you're interested in enrolling, please fill out and submit this form. We will first randomly sample 50-60 people for additional enrollment. We will release more slots if some students withdraw their enrollment

https://forms.gle/FH8NgPywkCCzS4XK6

- Any application of course withdrawal after 10/17 **will not** be accepted!