CMPUT 652: Reinforcement Learning with Robots

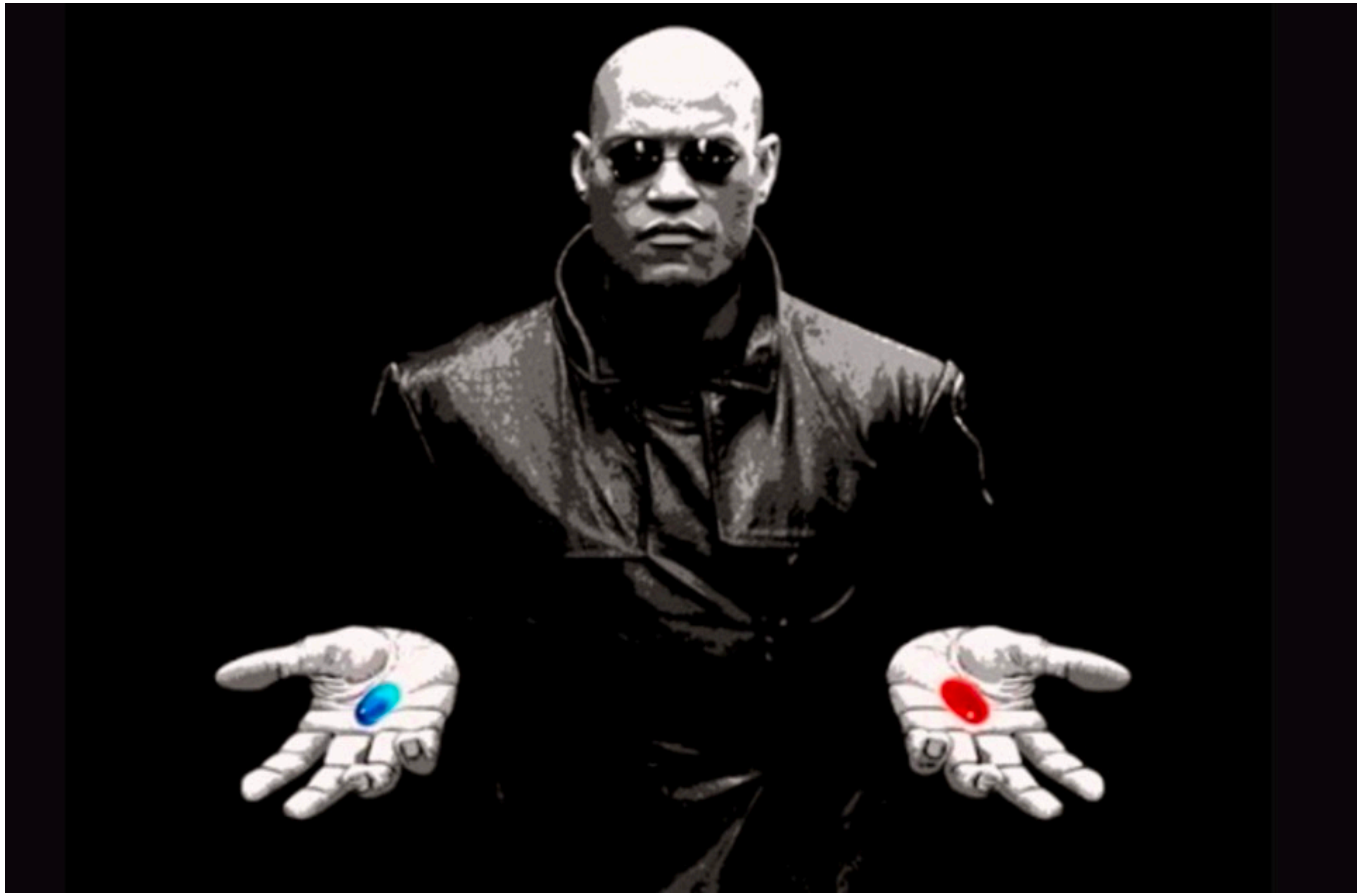
Rupam Mahmood

September 3, 2019

Artificial Intelligence

- ✓ To understand and create goal-oriented behaviors

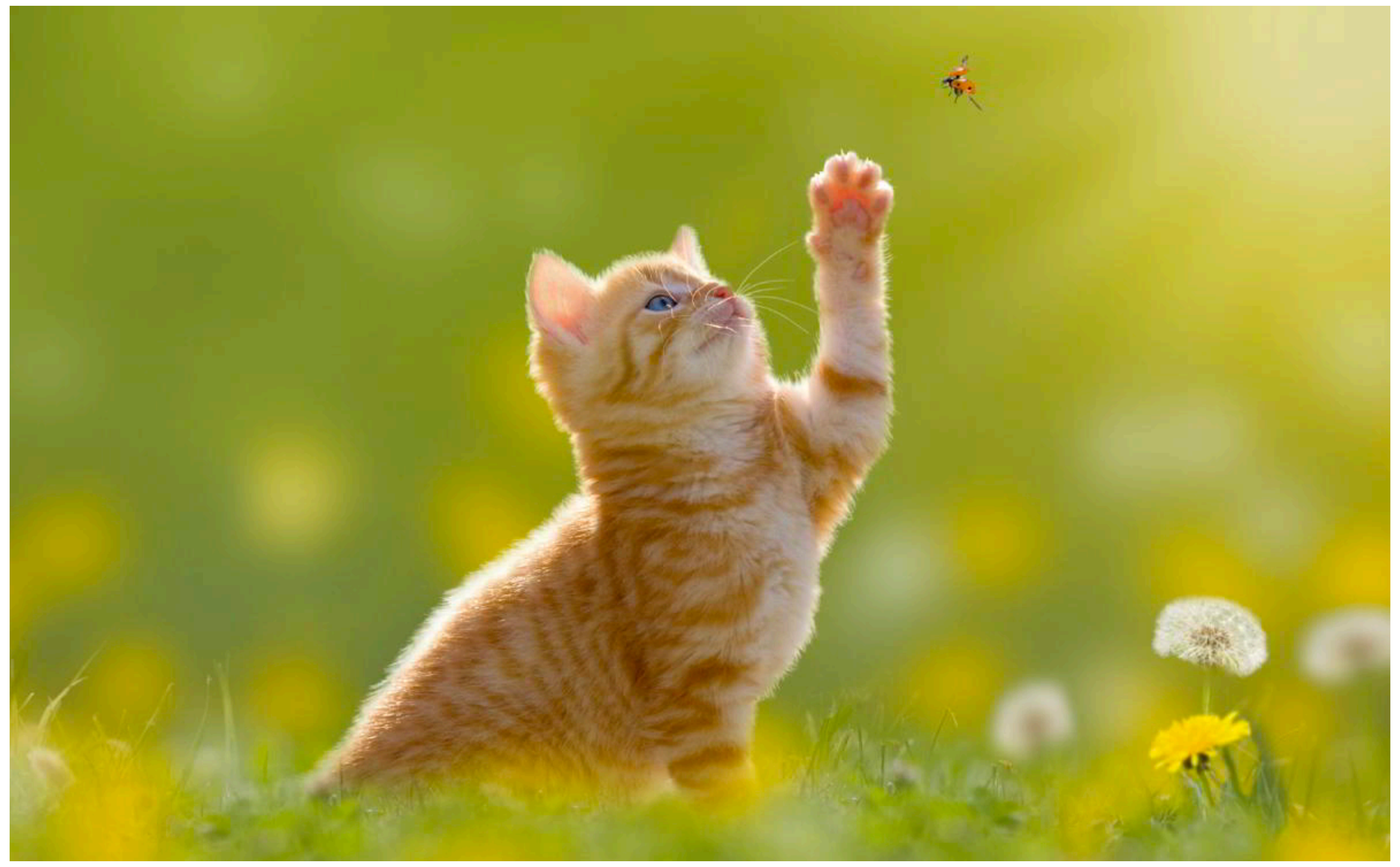
AI is the new electricity



AI is how we understand our deep reality and save our legacy

Goal is an important way to understand intelligence

A gigantic granite boulder



A cat catching a ladybug

Goal is an important way to understand life

Teleology	Teleonomy
Explanation in terms of an essential purpose, for example, eyes are designed to see	The quality of an apparent but not essential purposefulness

Rather than reject this [goal-directedness] idea (as certain biologists have tried to do) it is indispensable to recognise that it is essential to the very definition of living beings. We shall maintain that the latter are distinct from all other structures or systems present in the universe through this characteristic property, which we shall call teleonomy.
- Jacques Monod (1970)

For some phenomena, the length ratio between mechanistic and goal-directed descriptions is too high

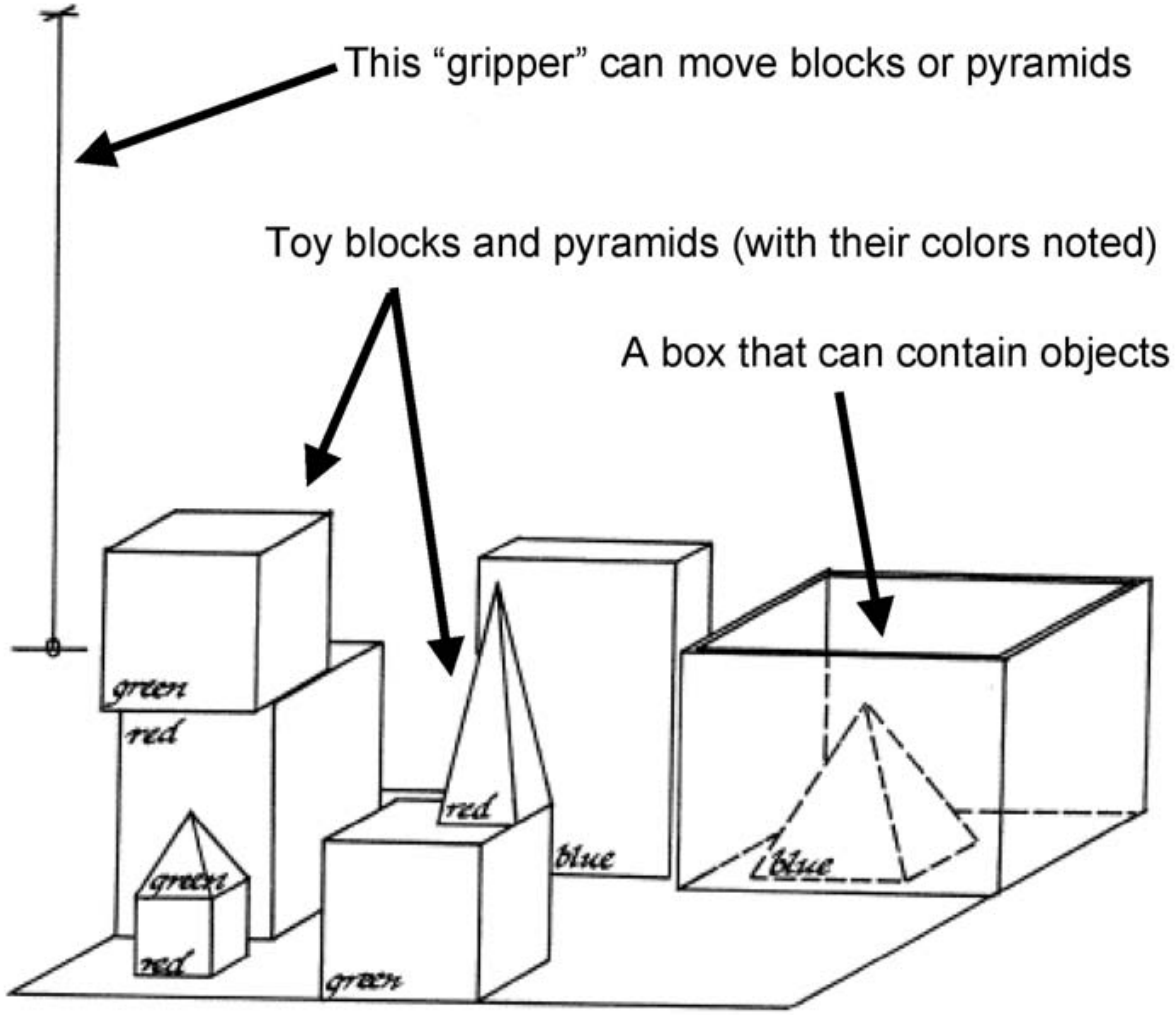
Artificial Intelligence

- ✓ *AI is the science and engineering of making intelligent machines, especially intelligent computer programs.*
- ✓ *Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.* - John Mccarthy
- ✓ AI is a constructive way of understanding animal intelligence.

Intelligence has two big parts

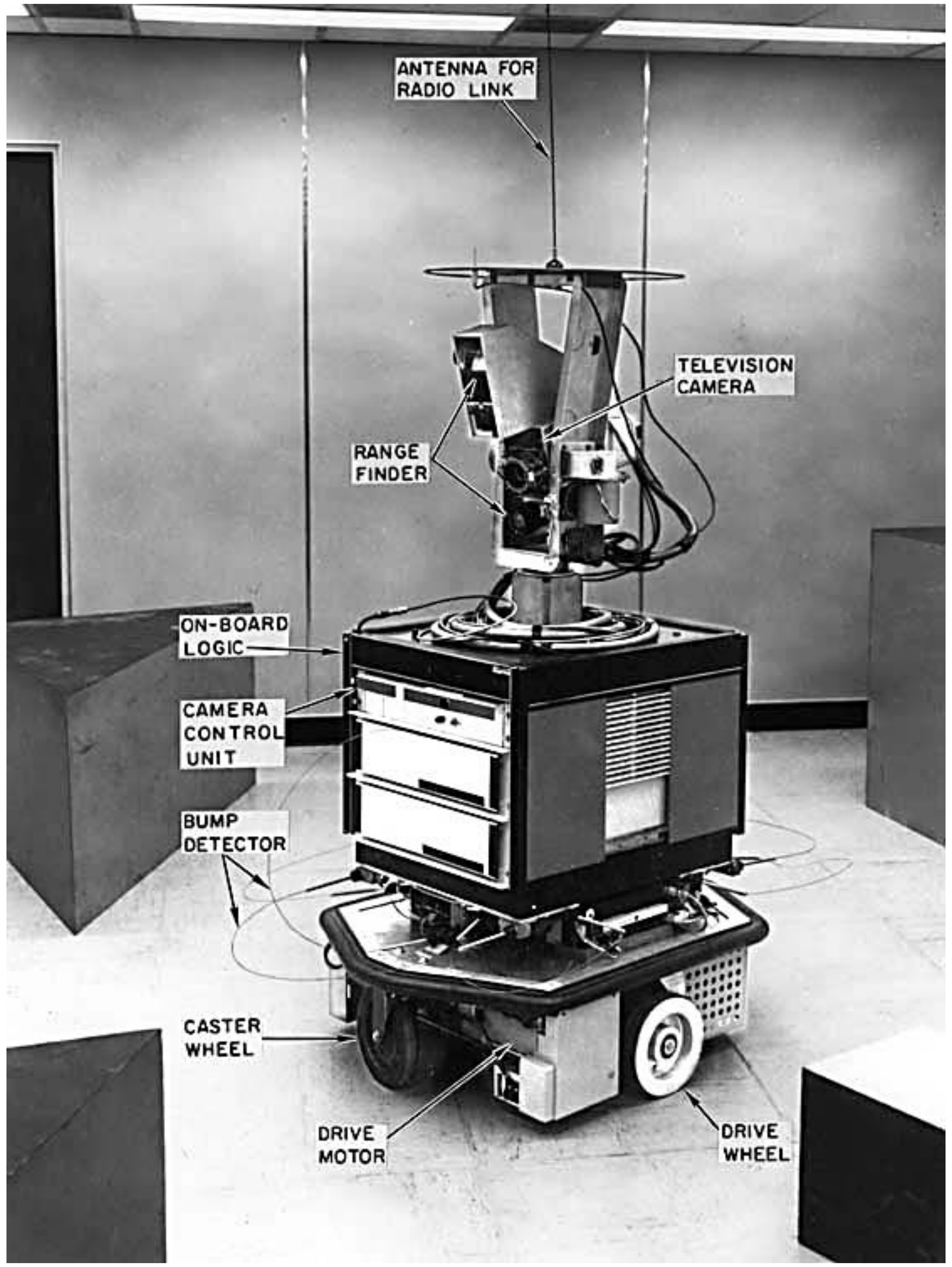
- ✓ Planning: improvement of behavior/knowledge by putting together known information without further experience
- ✓ Learning: improvement through experience

AI Success: SHRDLU (1968–1970)



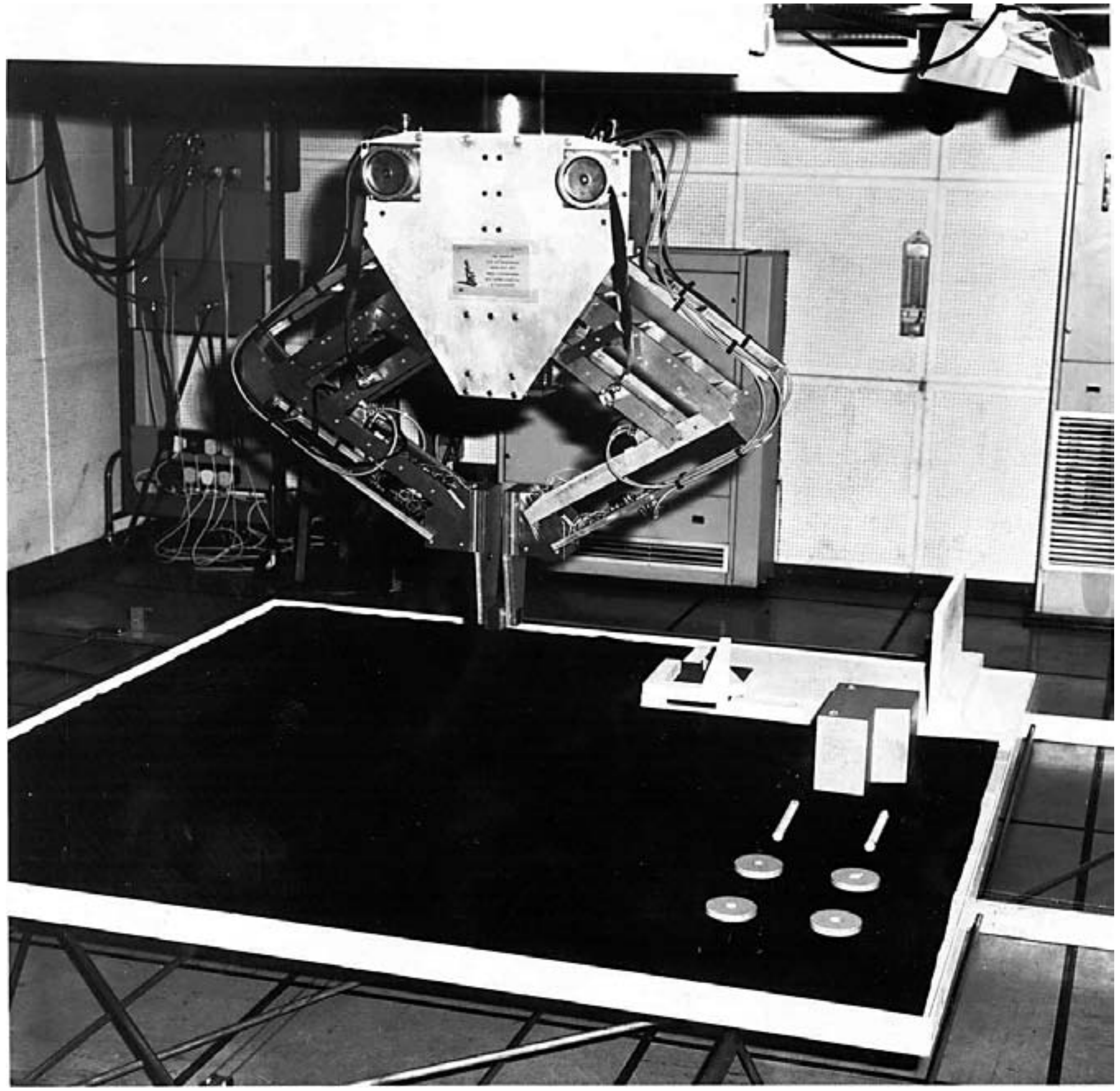
https://www.youtube.com/watch?v=QAJz4YKUwqw

AI Success: Shakey (1966–1972)



https://www.youtube.com/watch?v=7bsEN8mwUB8

AI Success: Freddy II (1973–1976)



https://www.youtube.com/watch?v=3c1Ff99RjkQ

AI Winter: 70s

- ✓ AI solutions did not match expectations
- ✓ “The general purpose robot is a mirage”, does not work in the real world
- ✓ DARPA: "funding people, not projects” -> "mission-oriented direct research, rather than basic undirected research”
- ✓ Moravec’s paradox: *It is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility.*

Fast forward to 21st century

✓ We are reverting to AI's original goal of general-purpose solutions

AlphaGo → **AlphaGo Zero** → **AlphaZero**

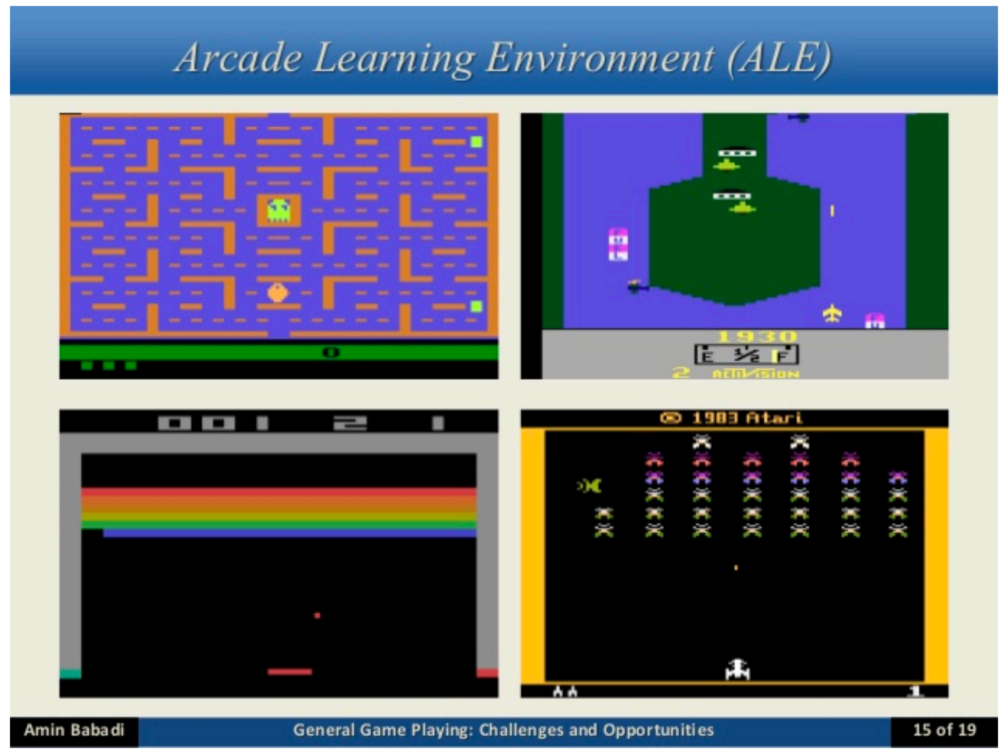
✓ General-purpose solutions also minimize the description length of the scientific theory of mind

Different ways of achieving generality

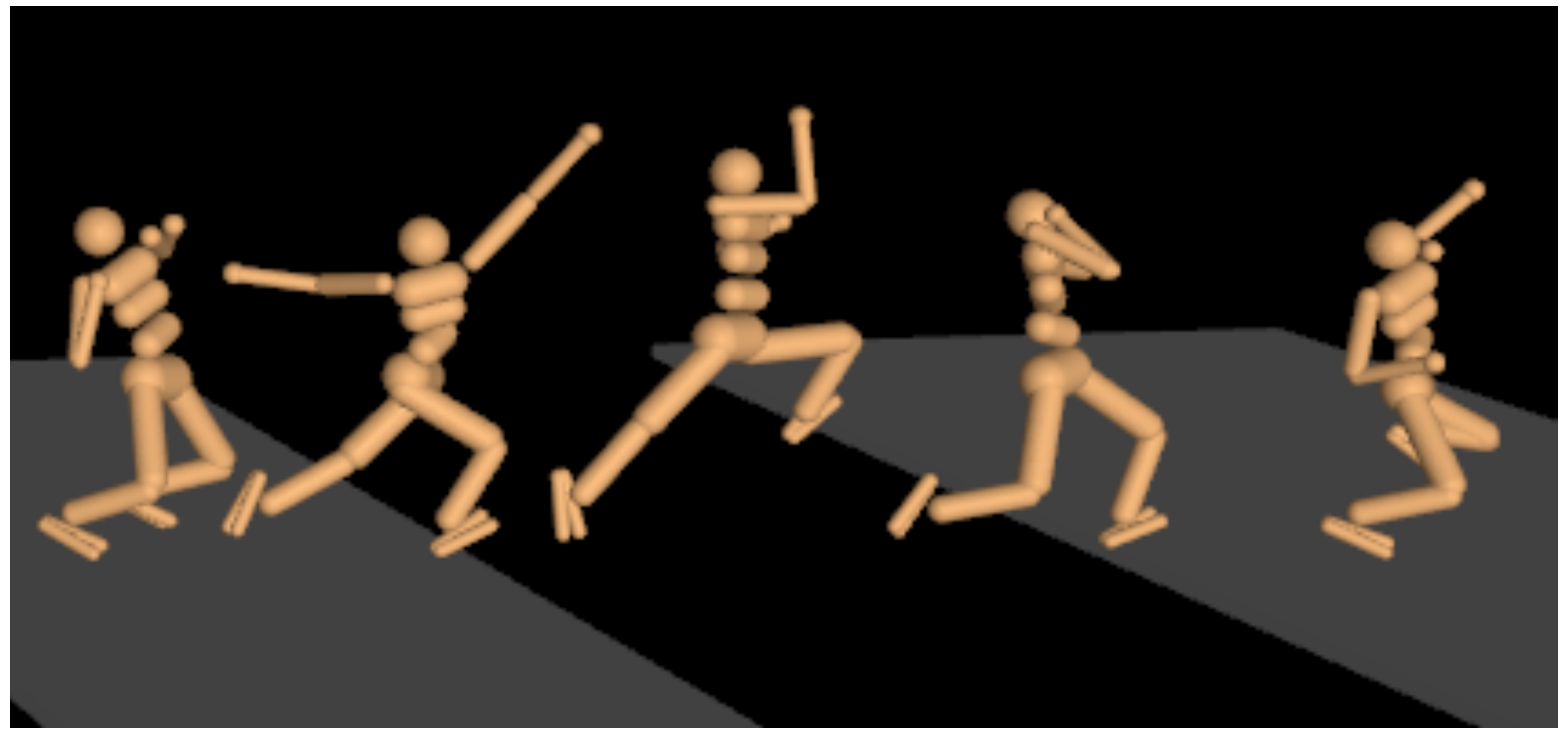
- ✓ Removing reliance on hand-coded solution
- ✓ Removing reliance on humans
- ✓ Adding Learning capabilities for adaptability
- ✓ Aiming for solutions that work for different problems
- ✓ Aiming for solutions that work for different embodiments
- ✓ Aiming for low-level control
- ✓ Aiming for learning from tabula rasa

We are observing a rise in general-purpose solutions

AI beating Go champion



Human-level control of
Atari games

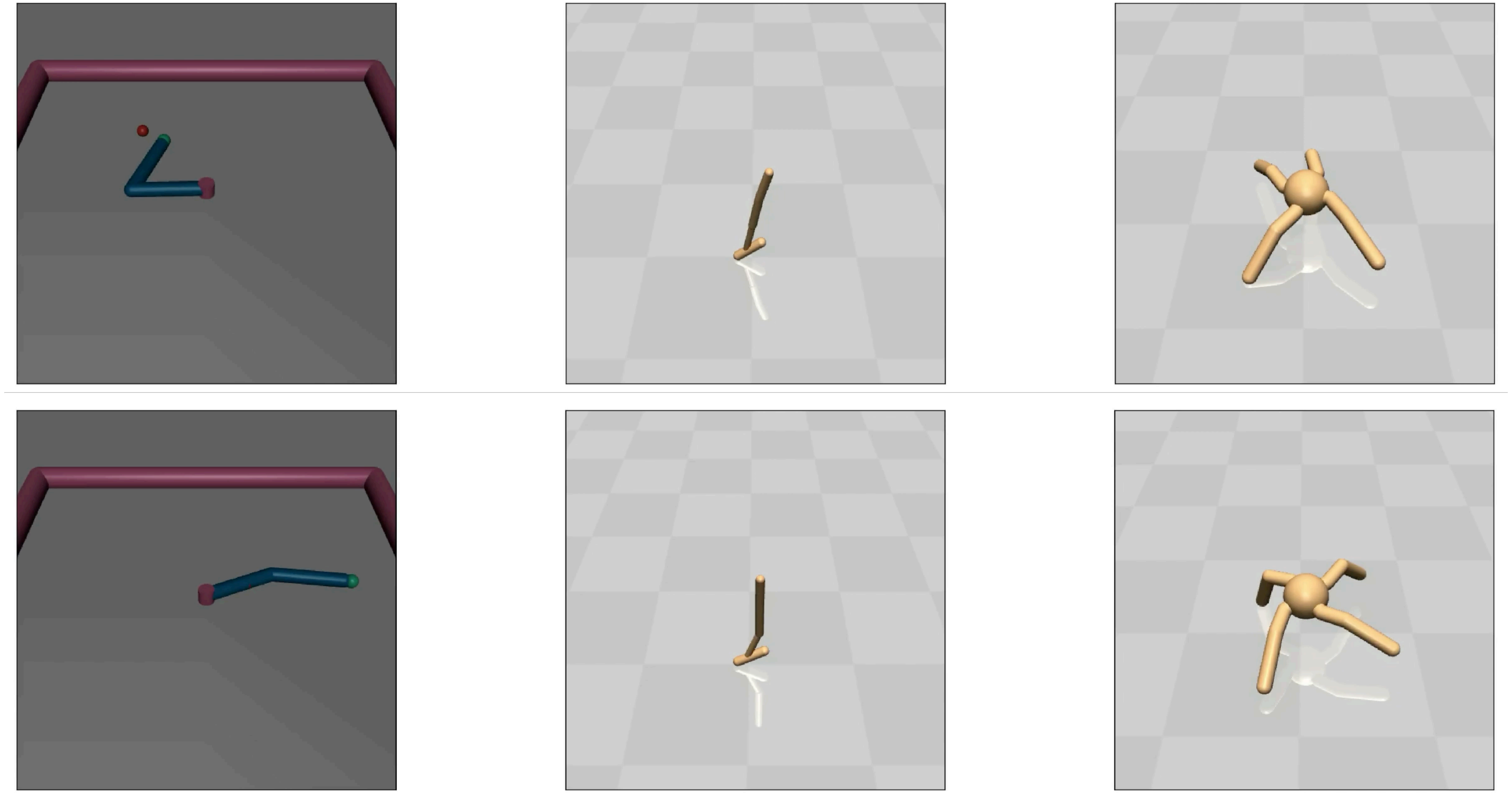


Emergence of locomotion behaviors

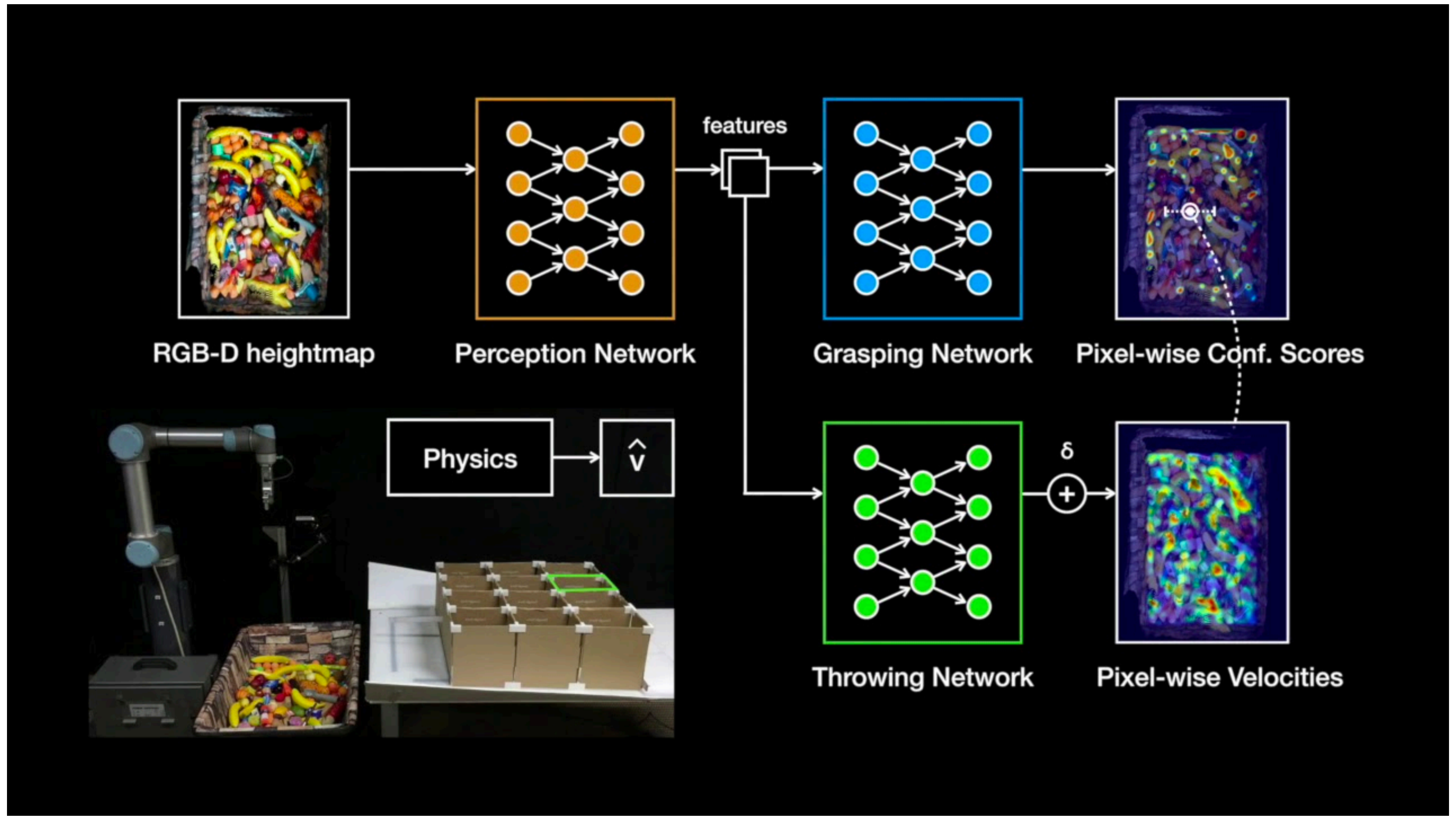
Deep RL agents show excellent general learning capabilities in virtual worlds

Initial behavior

Learning with TRPO



Hybrid approaches are becoming popular in robotics



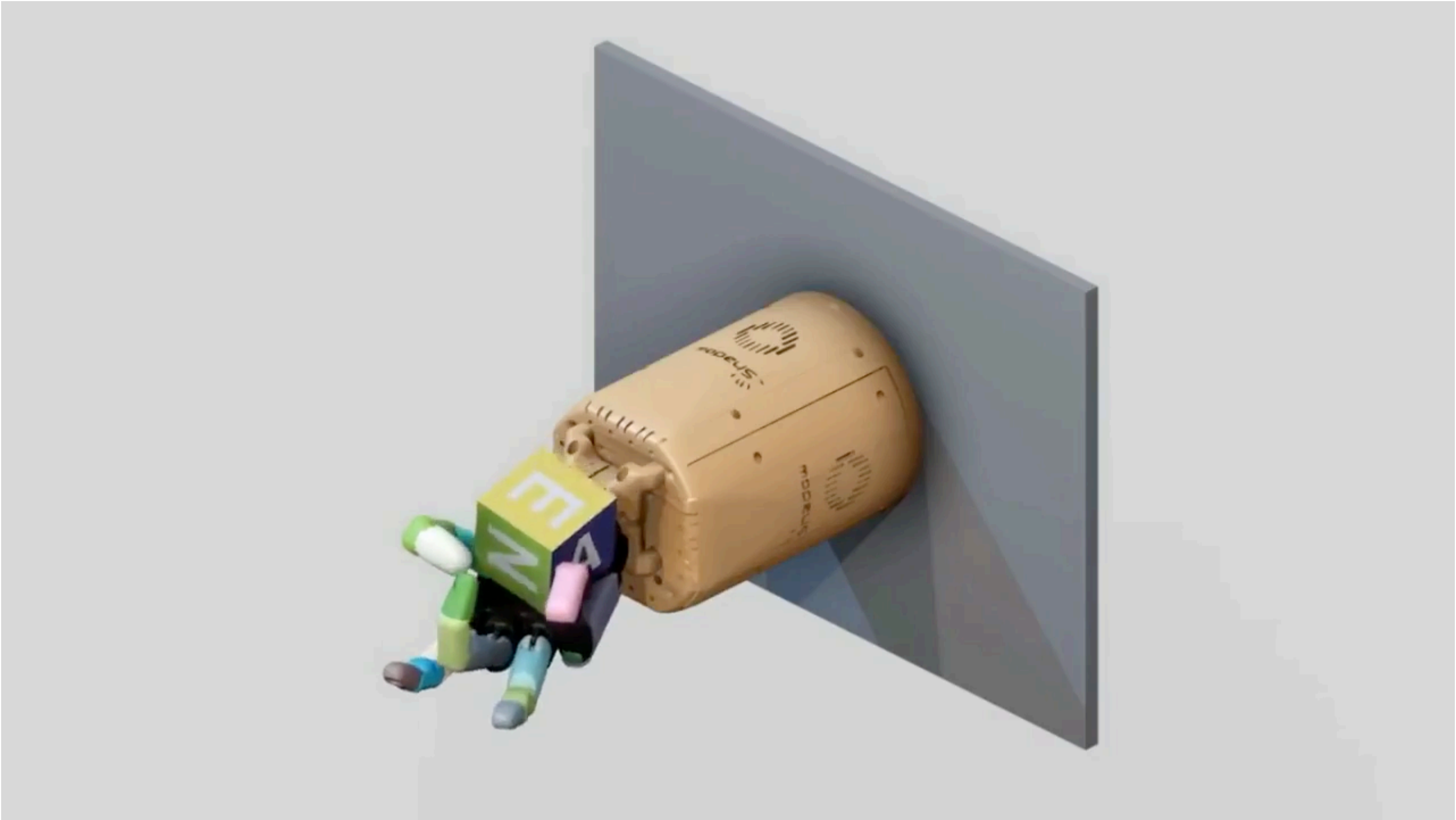
Zeng A, Song S, Lee J, Rodriguez A, Funkhouser T (2019).
TossingBot: Learning to Throw Arbitrary Objects with Residual Physics.

We are also learning more toward learning through imitation learning and sim2real transfer approaches

Imitation learning

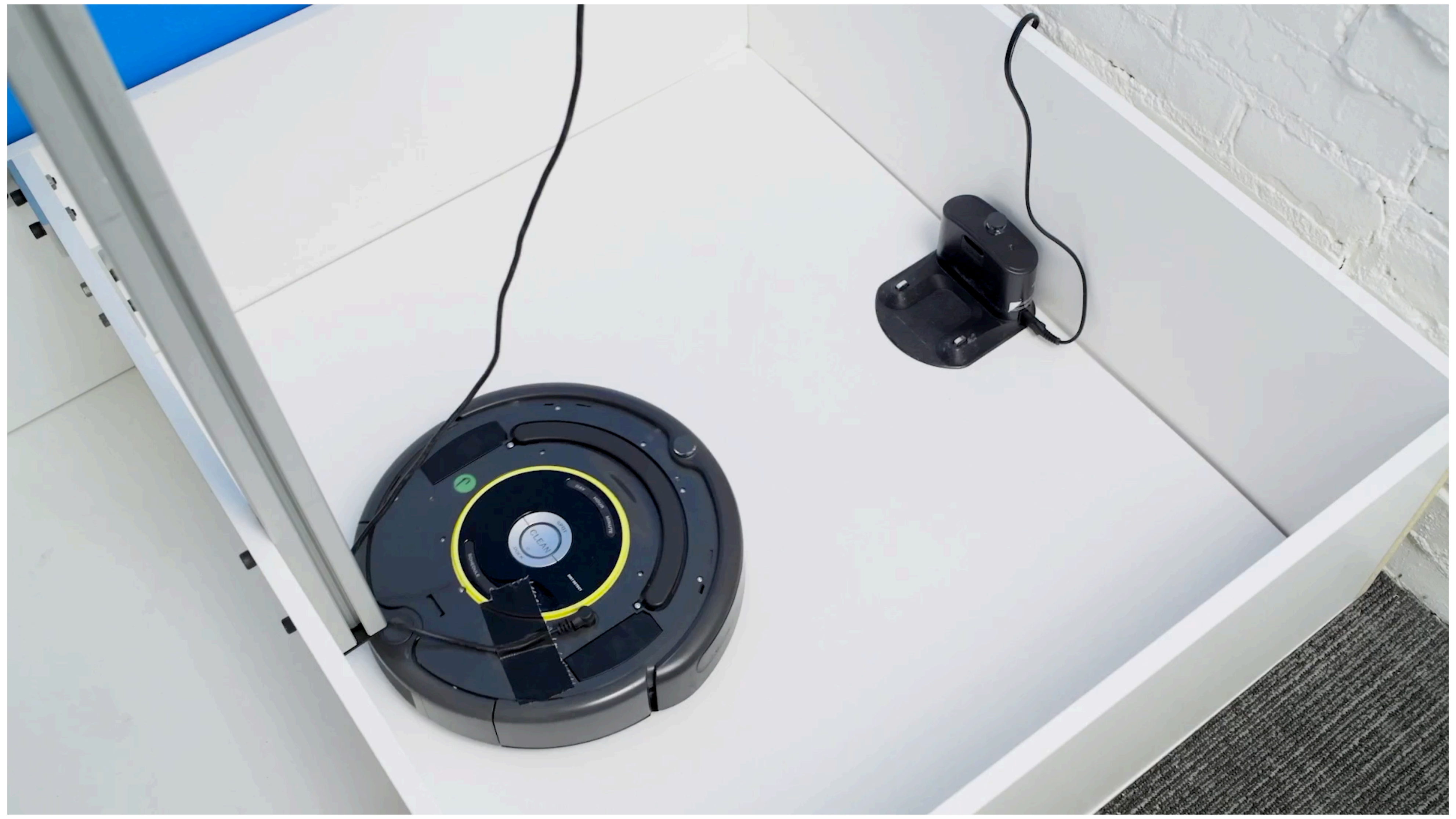
Kormushev P, Calinon S, Caldwell DG (2010).
Robot Motor Skill Coordination
with EM-based Reinforcement Learning

Simulation-to-reality (Sim2Real) transfer

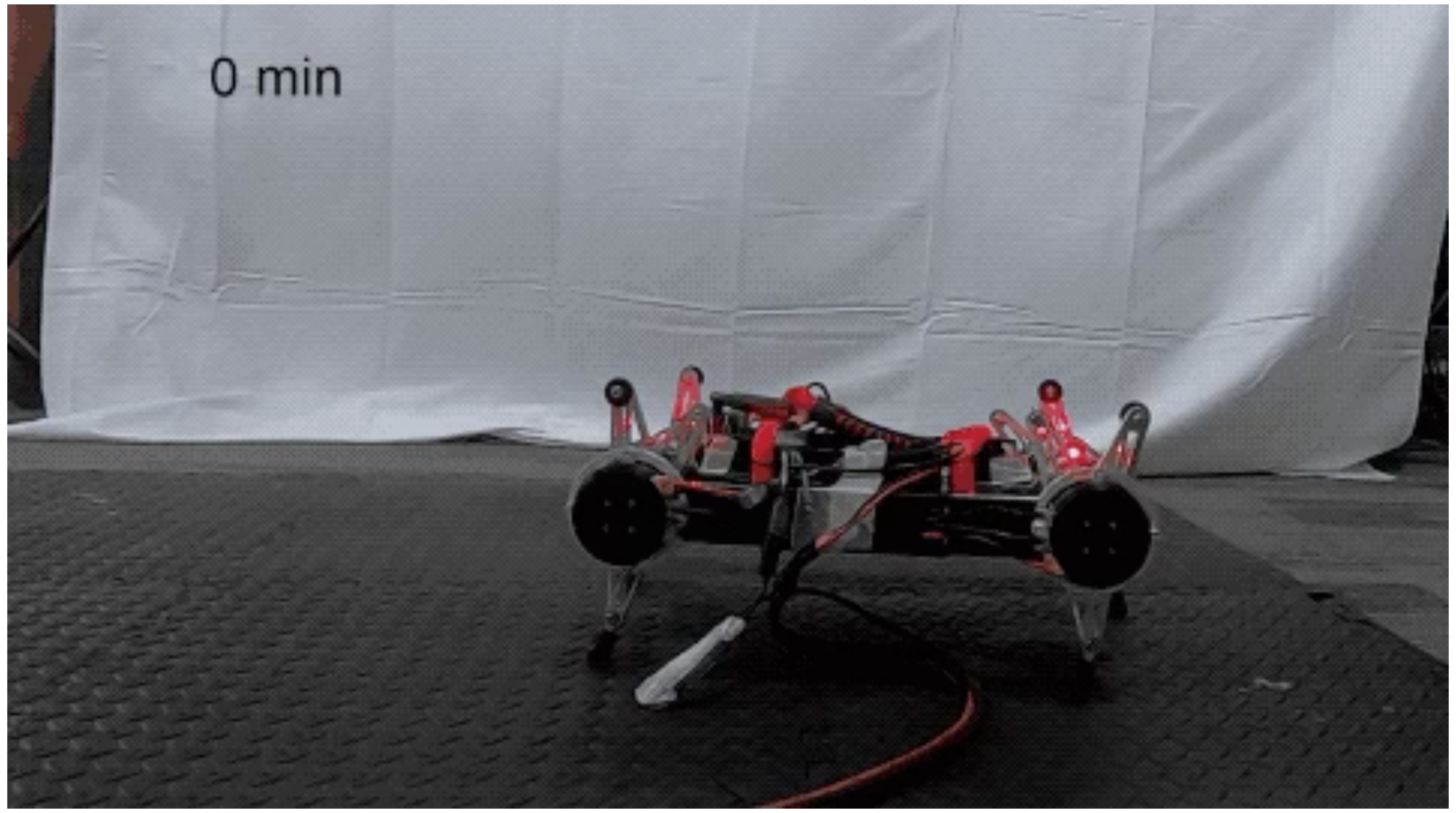


Andrychowicz et al. (2018)
Learning dexterous in-hand manipulation

Learning from scratch in real time is still not common



Mahmood AR, Korenkevych D, Vasan G, Ma W, Bergstra J (2018).
Benchmarking reinforcement learning algorithms on real-world robots.



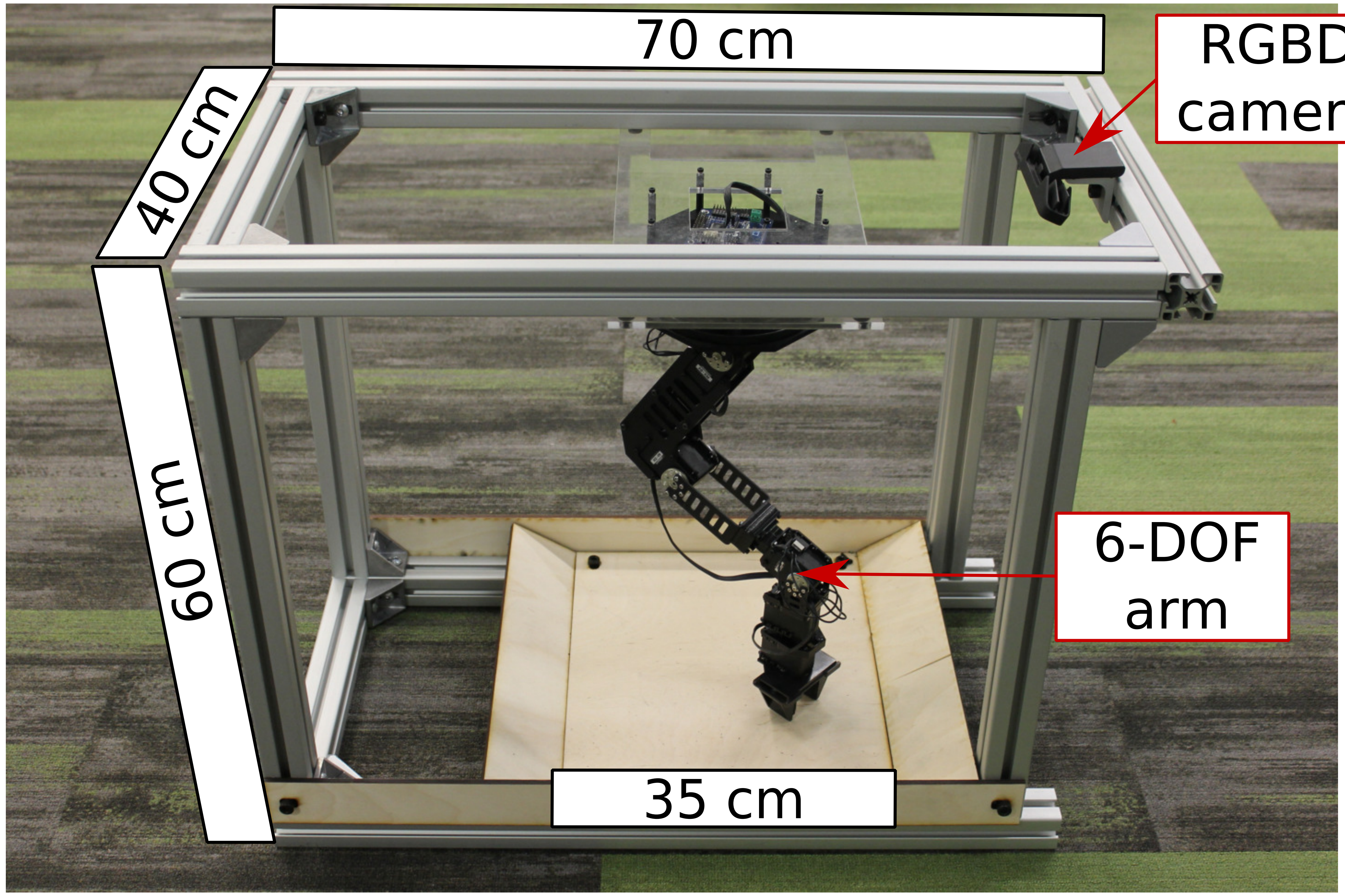
Haarnoja, T., Zhou, A., Ha, S., Tan, J., Tucker, G., & Levine, S. (2018).
Learning to walk via deep reinforcement learning.

There is a growing inequality of ability and knowledge in RL with robots

- ✓ Robot learning is suffering from reproducibility and transfer of knowledge
- ✓ In graduate courses, we moved from Mountain Car to Hopper
- ✓ We need to push the boundary for robot learning

In this course we study the challenges and solutions to real-time learning problems

A reproducibility toolkit for grasping is also available from UC Berkeley

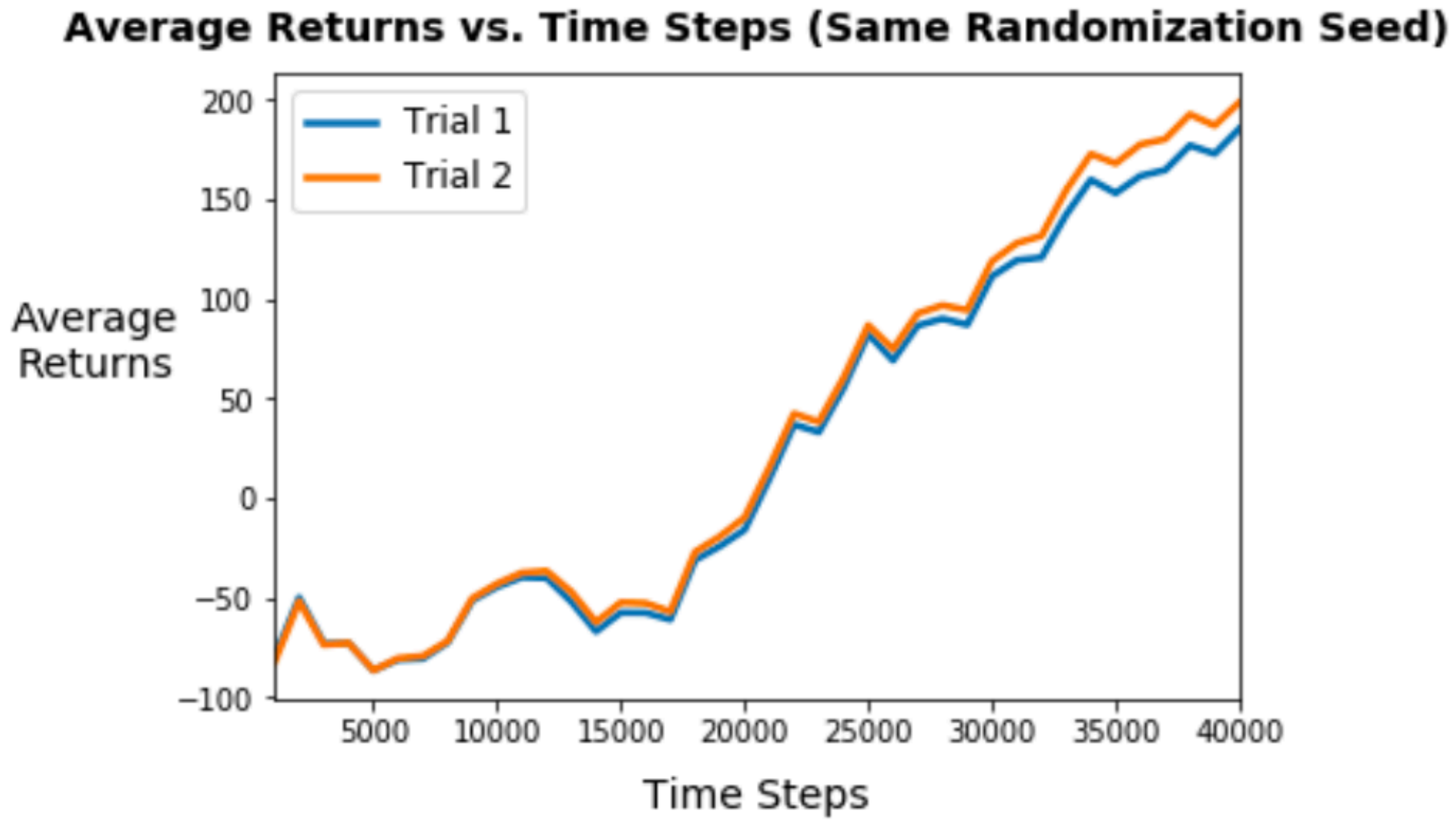


Yang, B., Zhang, J., Pong, V., Levine, S., & Jayaraman, D. (2019).
REPLAB: A Reproducible Low-Cost Arm Benchmark Platform for Robotic Learning.

Kindred introduced *SenseAct*, a benchmark task suite for reproducible RL research with real robots

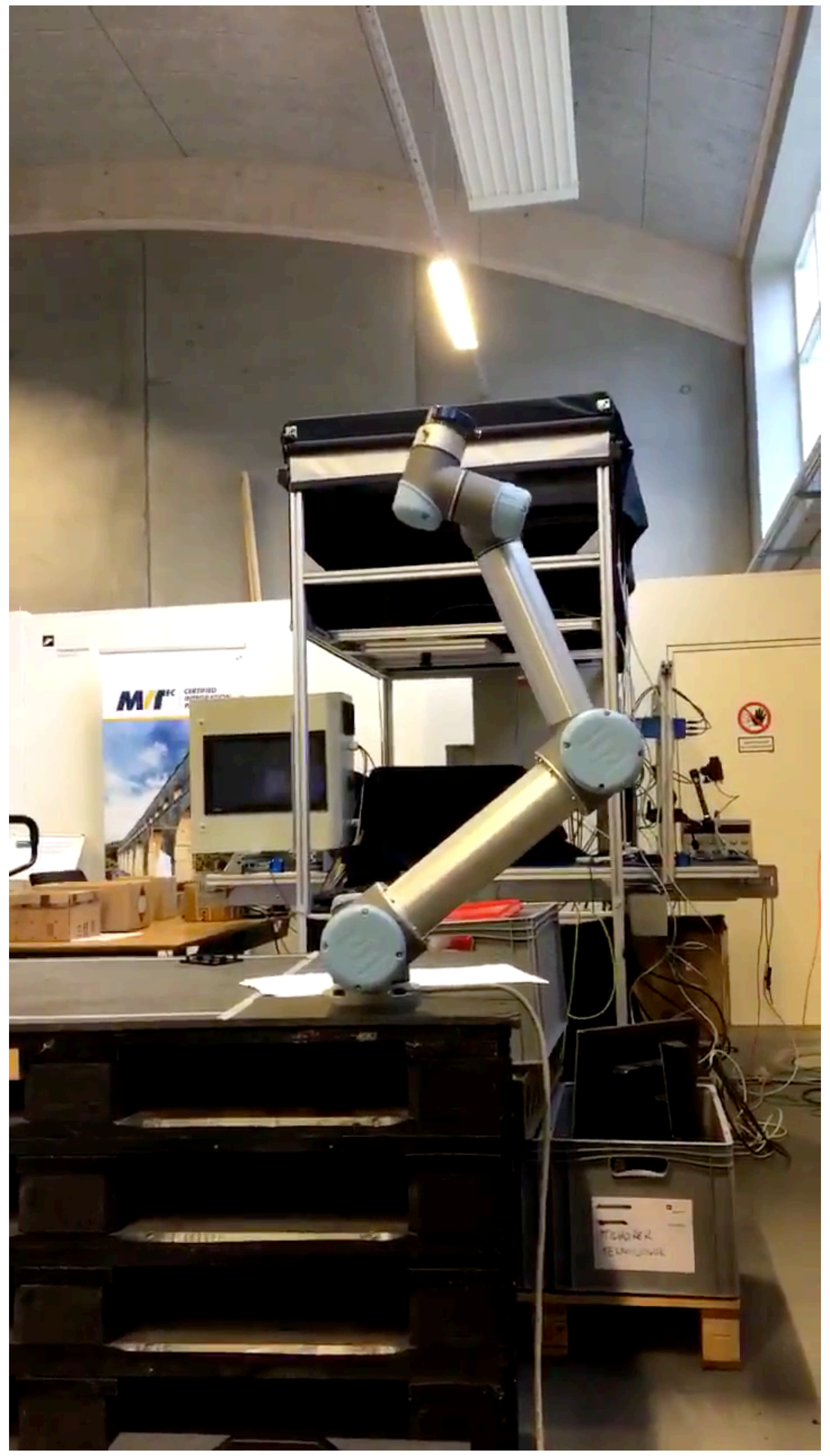
https://github.com/kindredresearch/SenseAct

Different research labs were able to use SenseAct for reproducibility and original research



Two learning curves from different trials using the same randomization seed.

Oliver Limoyo, UTIAS
University of Toronto
Canada



Nicolai Anton Lynnerup
Teknologisk Institut
Denmark

Course page

- ✓ armahmood.github.io/rl-robots-course
- ✓ Assignments at eClass

Course overview

- Stochastic gradient descent and averaging
- Bandits
- Markov decision processes
- Value function approximation
- Policy gradient methods
- Classical control approaches
- Controlling Dynamixels
- Challenges in real-time learning
- Architectures for real-time learning tasks
- Simulation-to-reality transfer
- Learning from demonstration

Teaching Team

- ✓ Instructor: Rupam Mahmood

  Office hours: Monday 4-4:50 @ATH 404

- ✓ Expert: Craig Sherstan

  PhD Candidate in RL w/ Patrick Pilarski

  Intern at Huawei Noah's Ark Lab

Course work and evaluation

- ✓ Class notes 5%
- ✓ Readings 5%
- ✓ Assignment on MDPs and value function 10%
- ✓ Assignment on policy gradient methods 15%
- ✓ Assignment on learning in real-time 15%
- ✓ Course project 50%
  - Proposal presentation 10%
  - Project presentation 10%
  - Project report 30%

Course Materials

- ✓ Textbook: *Reinforcement Learning: An Introduction*, 2nd edition by R Sutton and A Barto, MIT Press.
- ✓ Lectures and other materials will be linked in the schedule
- ✓ Each entry in the schedule will be finalized by the day of the corresponding lecture

Helpful background (e.g.)

- ✓ Probabilities, Linear algebra
- ✓ Python, NumPy
- ✓ PyTorch, TensorFlow
- ✓ Hands-on experience with camera, servomotor I/O
- ✓ Experiments with DQN, A3C
- ✓ Implementation of DQN, A3C

Robot for individual lab

Dynamixel MX-64AT actuator

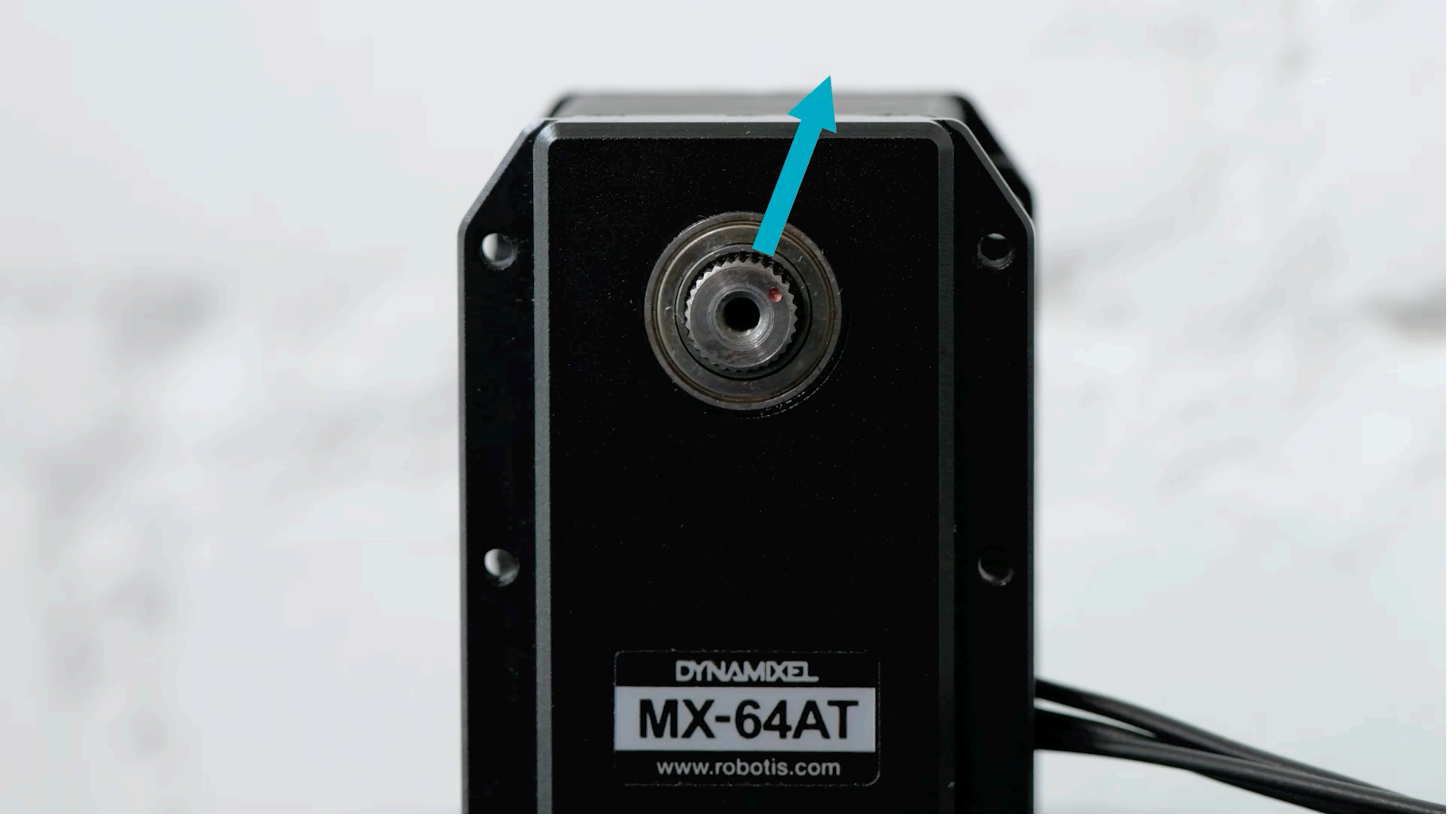
Something similar to . . .

Position, speed & torque control

~$300 USD

Robot for individual lab

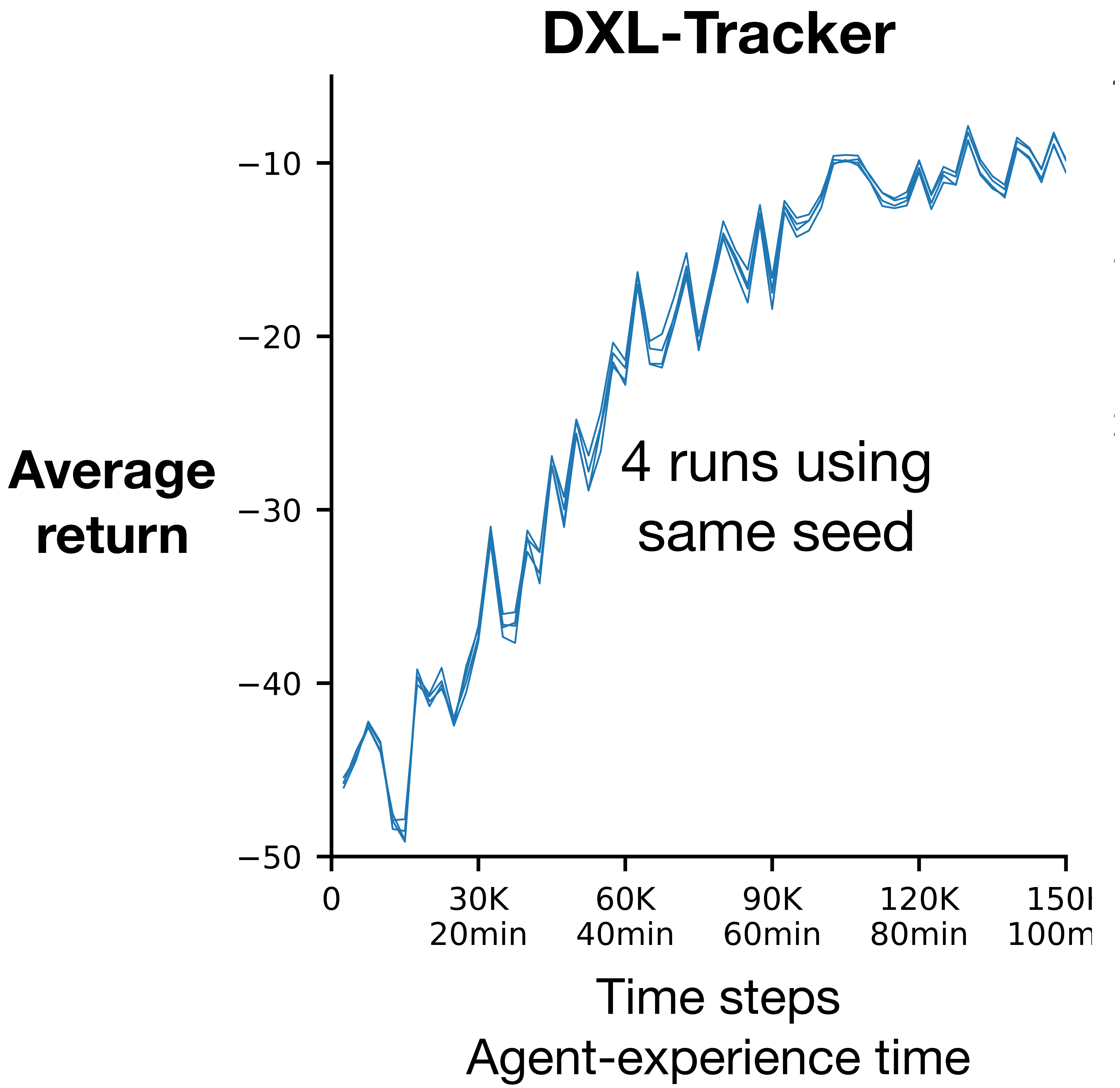
Randomly generated agent



Learned agent

Scripted agent

Robot for individual lab



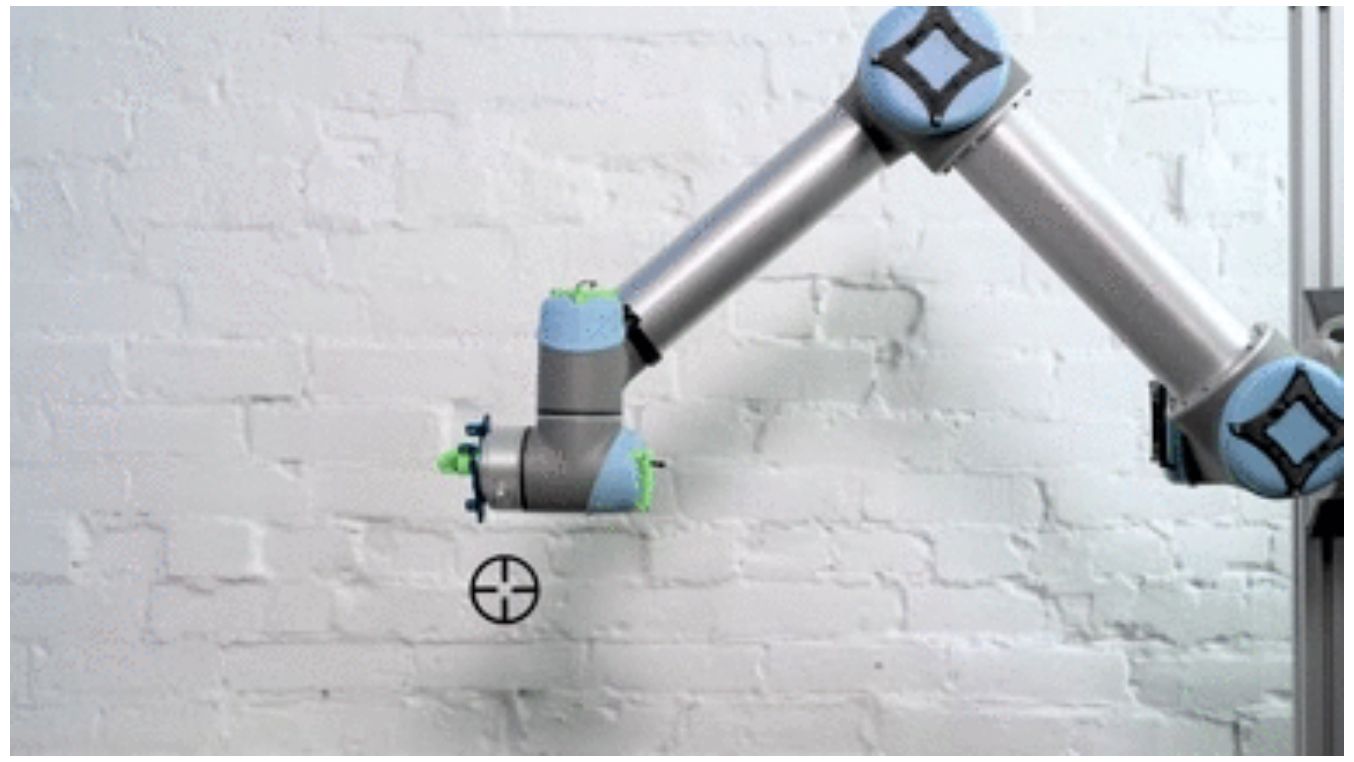
Robots for course project

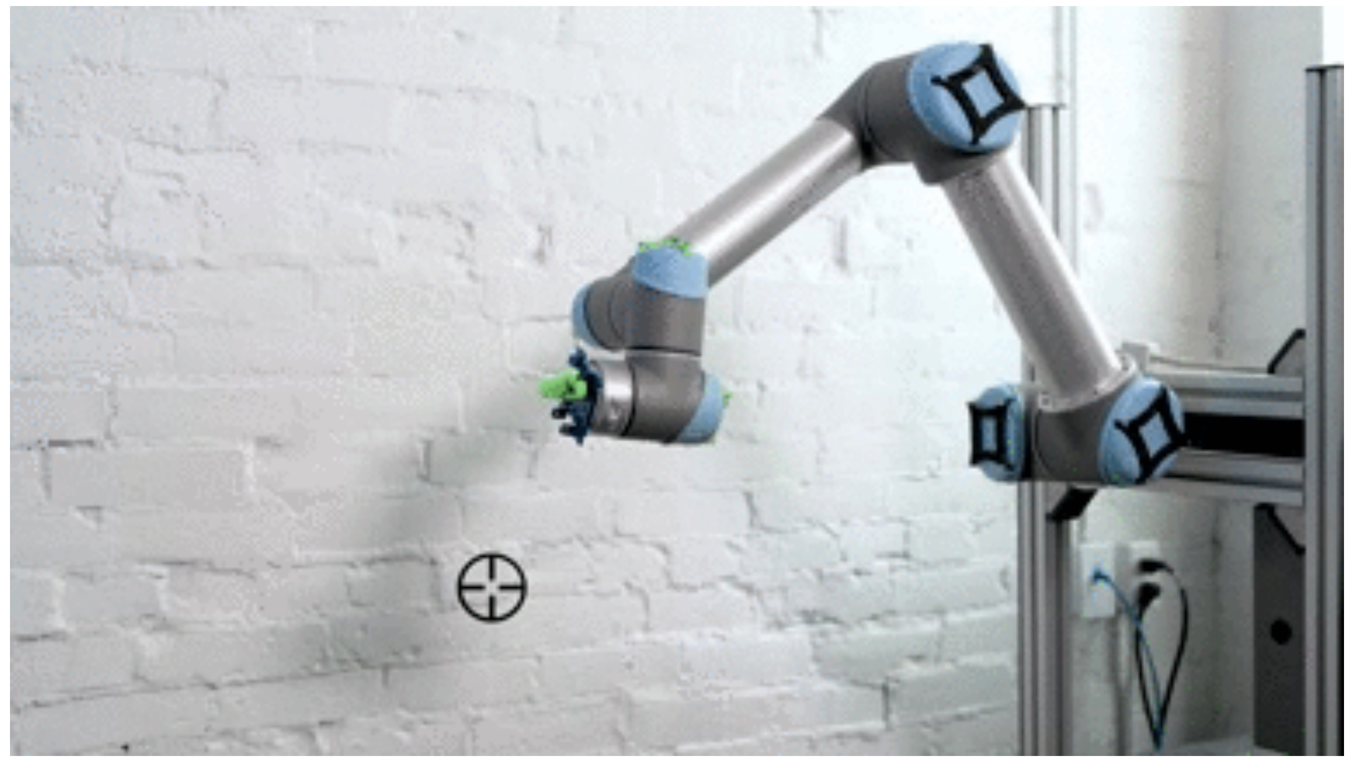
UR5 robotic arm

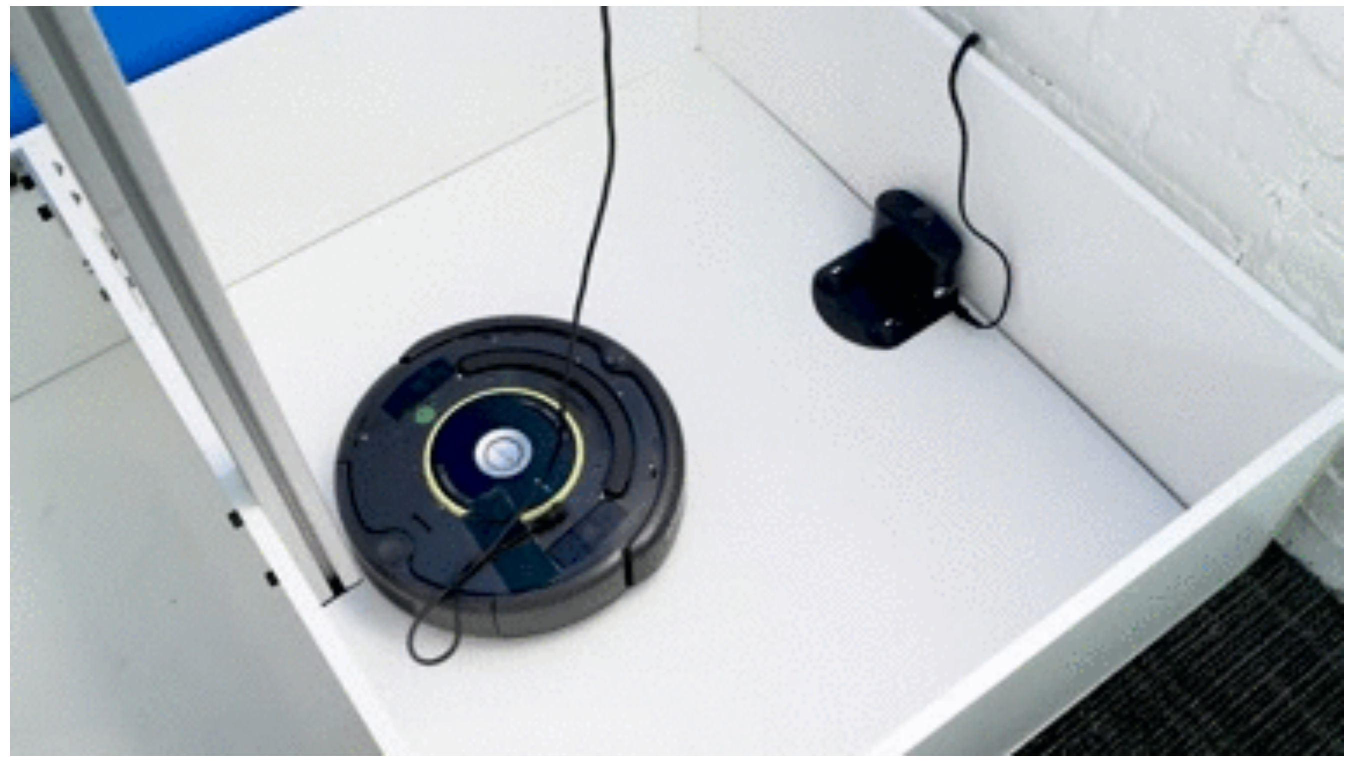
Dynamixel MX-64AT actuator

Create 2 mobile robot base

Robots for course project







Robots for course project