CMPUT 652: RL with Robots

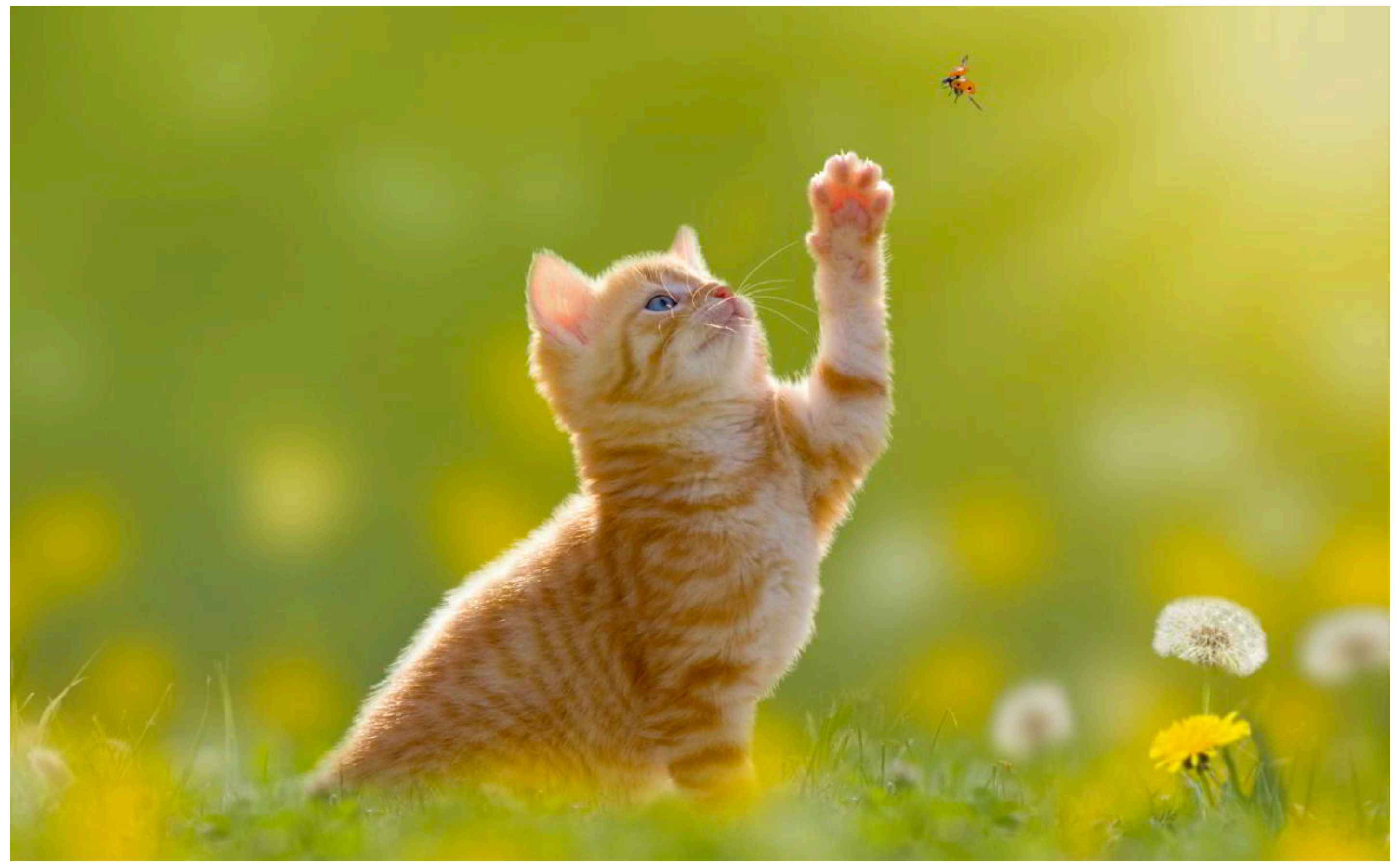
The Road to AI

Rupam Mahmood

September 5, 2019

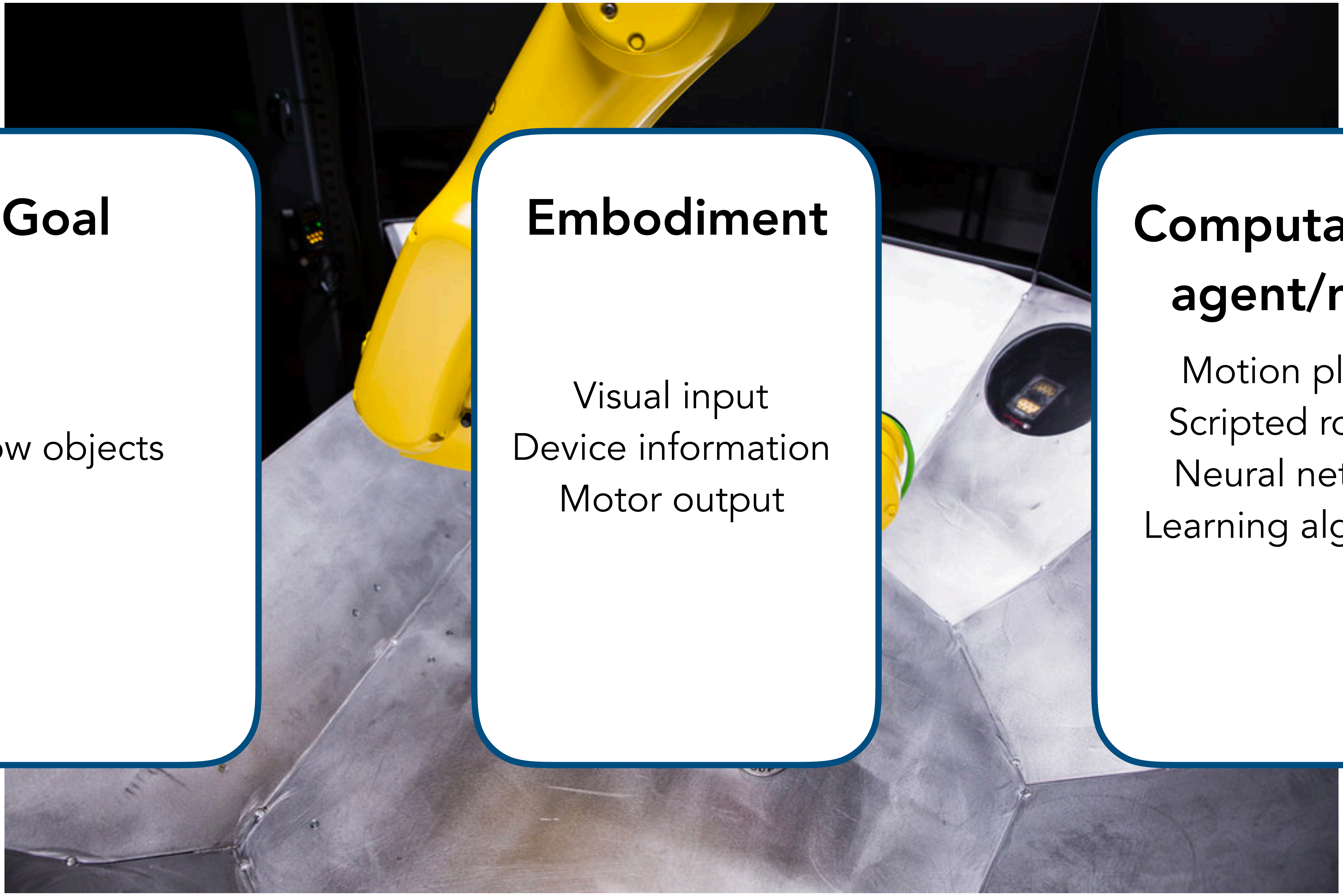
Goal is an important way to understand intelligence

A gigantic granite boulder



A cat catching a ladybug

AI is the study of computational abilities to achieve goals



Goal

Stow objects

Embodiment

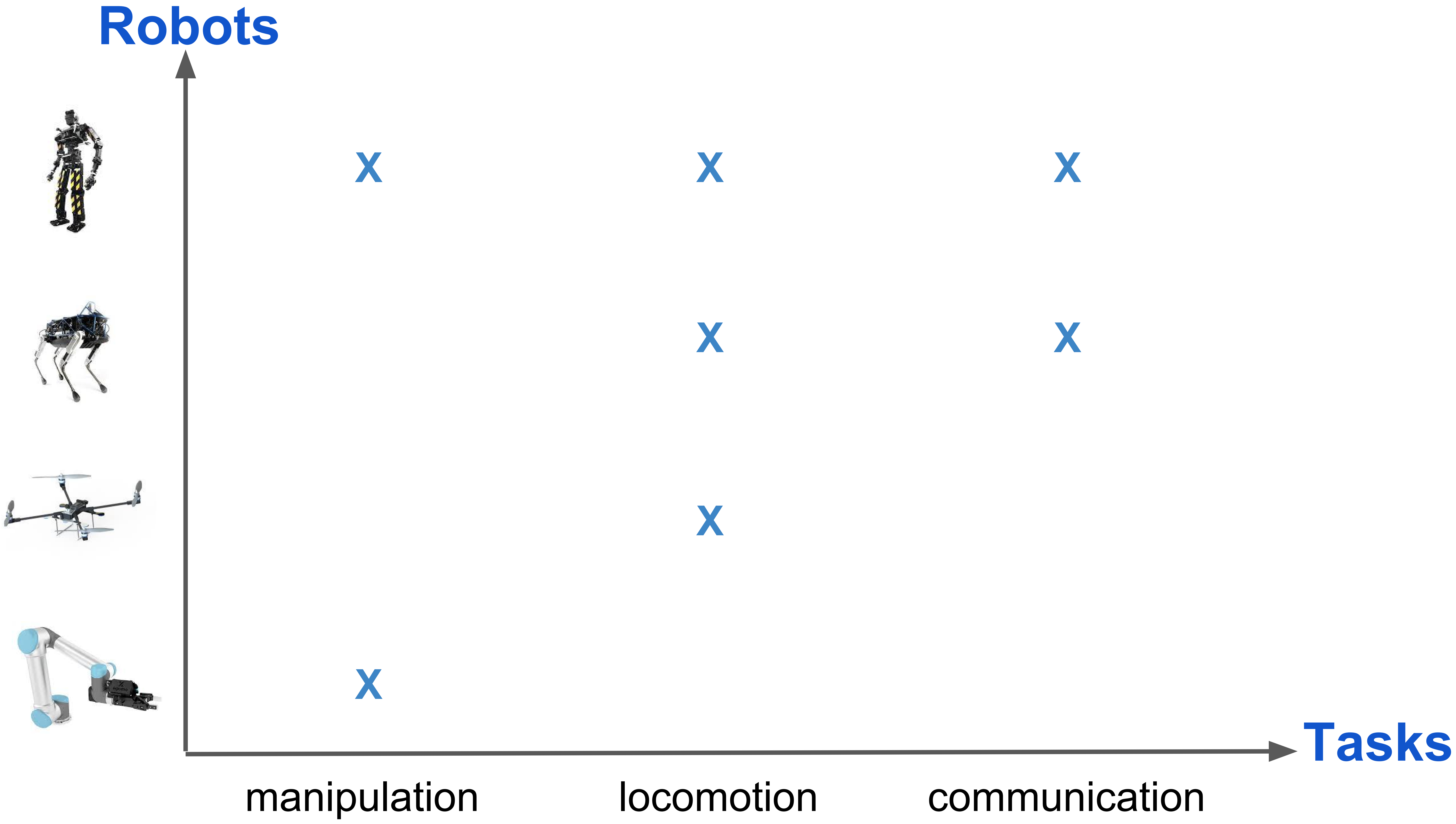
Visual input
Device information
Motor output

Computational agent/mind

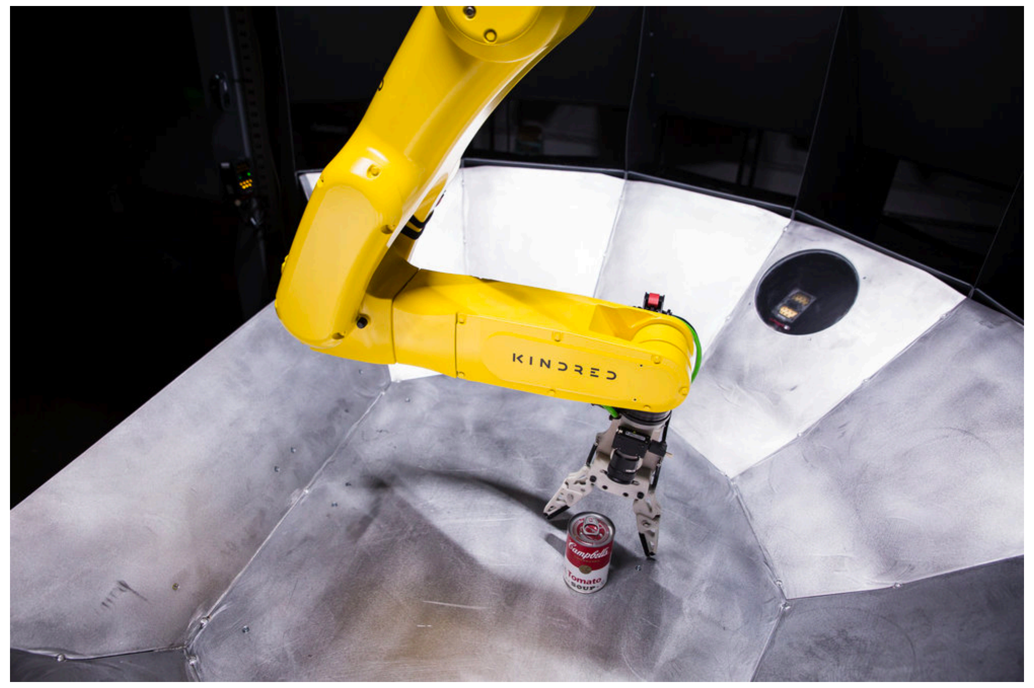
Motion planner
Scripted routines
Neural networks
Learning algorithms

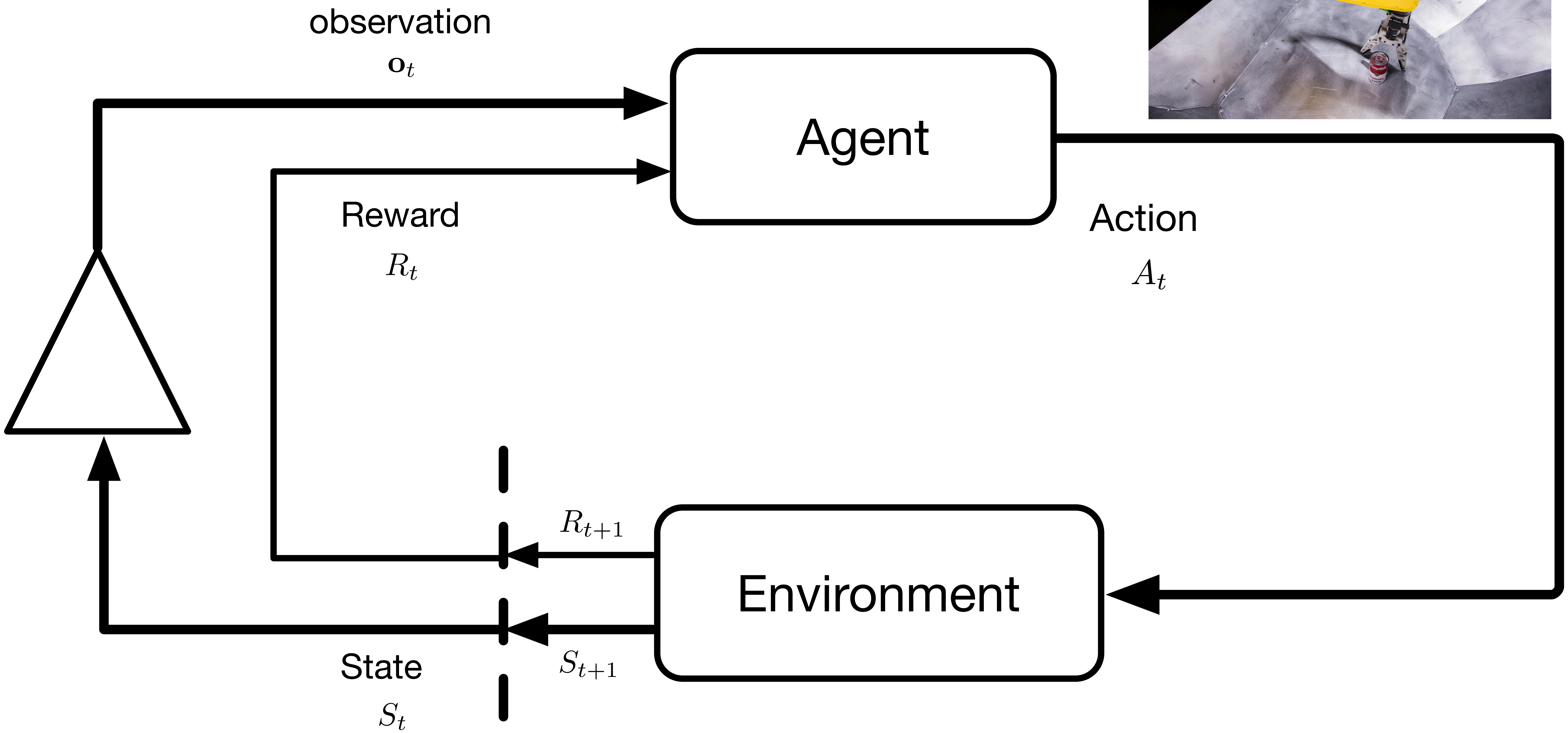
Kindred's robotic product: Sort

Emergence of general learning agents can transform AI applications in robotics

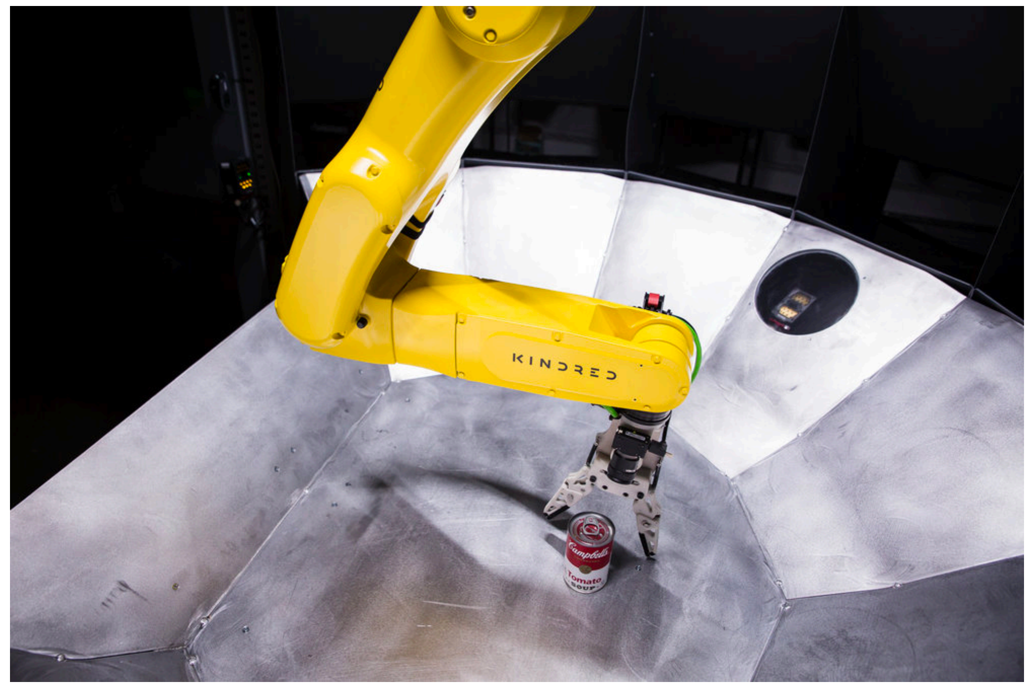


Reinforcement learning provides one of the simplest specifications of an agent





Agents can be studied separately or using a complete physical system



Kindred's robotic product: Sort

Goal

Stow objects

Embodiment

Visual input
Device state
Motor output

Computational agent

Motion planner
Scripted routines
Neural networks
Learning algorithms

using a complete physical AI system

studying separately

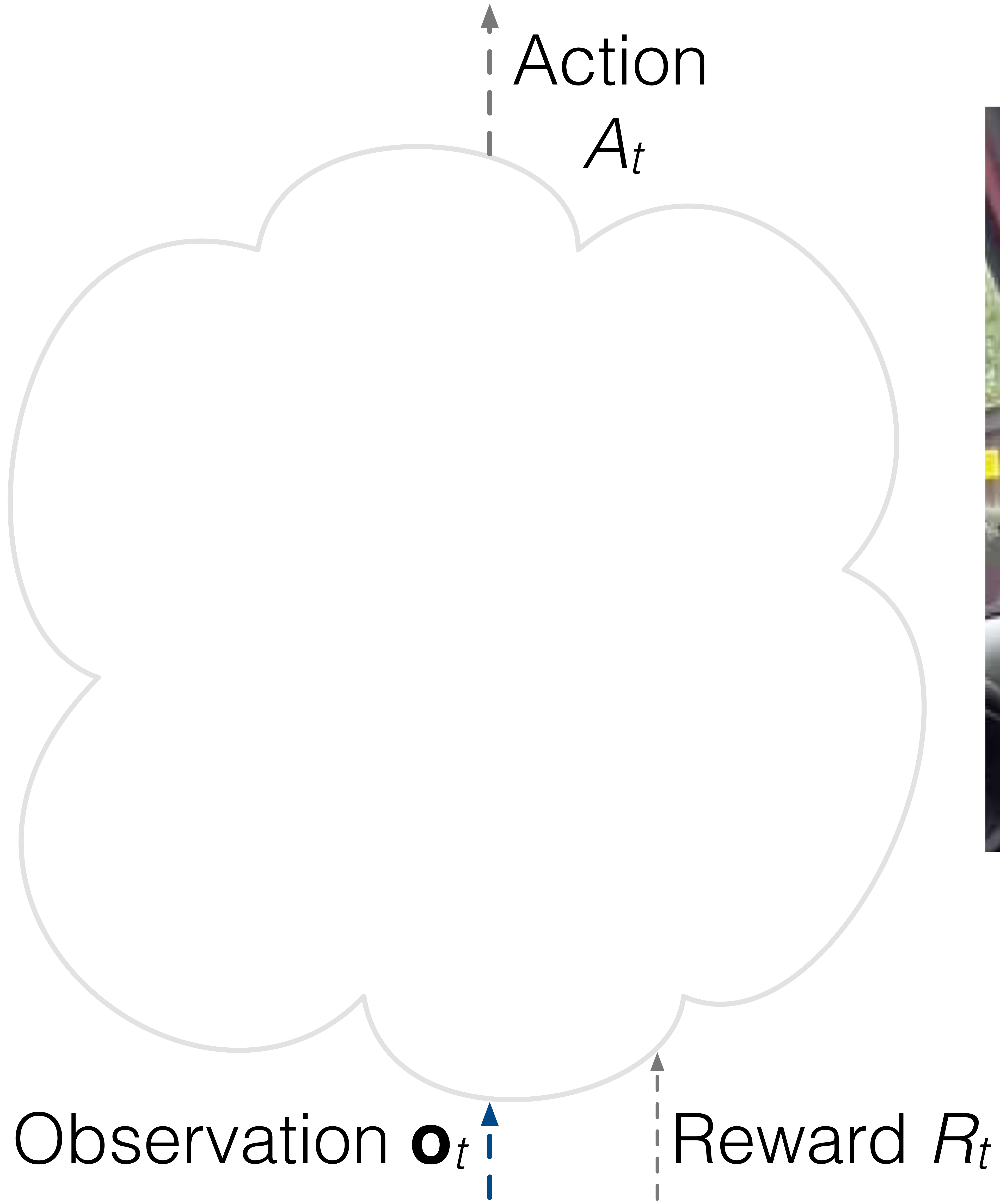
Criteria for a computational agent as a theory of mind

- ✓ General-purpose: Independent of embodiments and goals
- ✓ Efficient: Can achieve goals efficiently
- ✓ Simple: Can be described with simple minimal description
- ✓ Computational: Fully specified in terms of computations
- ✓ Scientific: Developed through a scientific process

Different ways of achieving generality

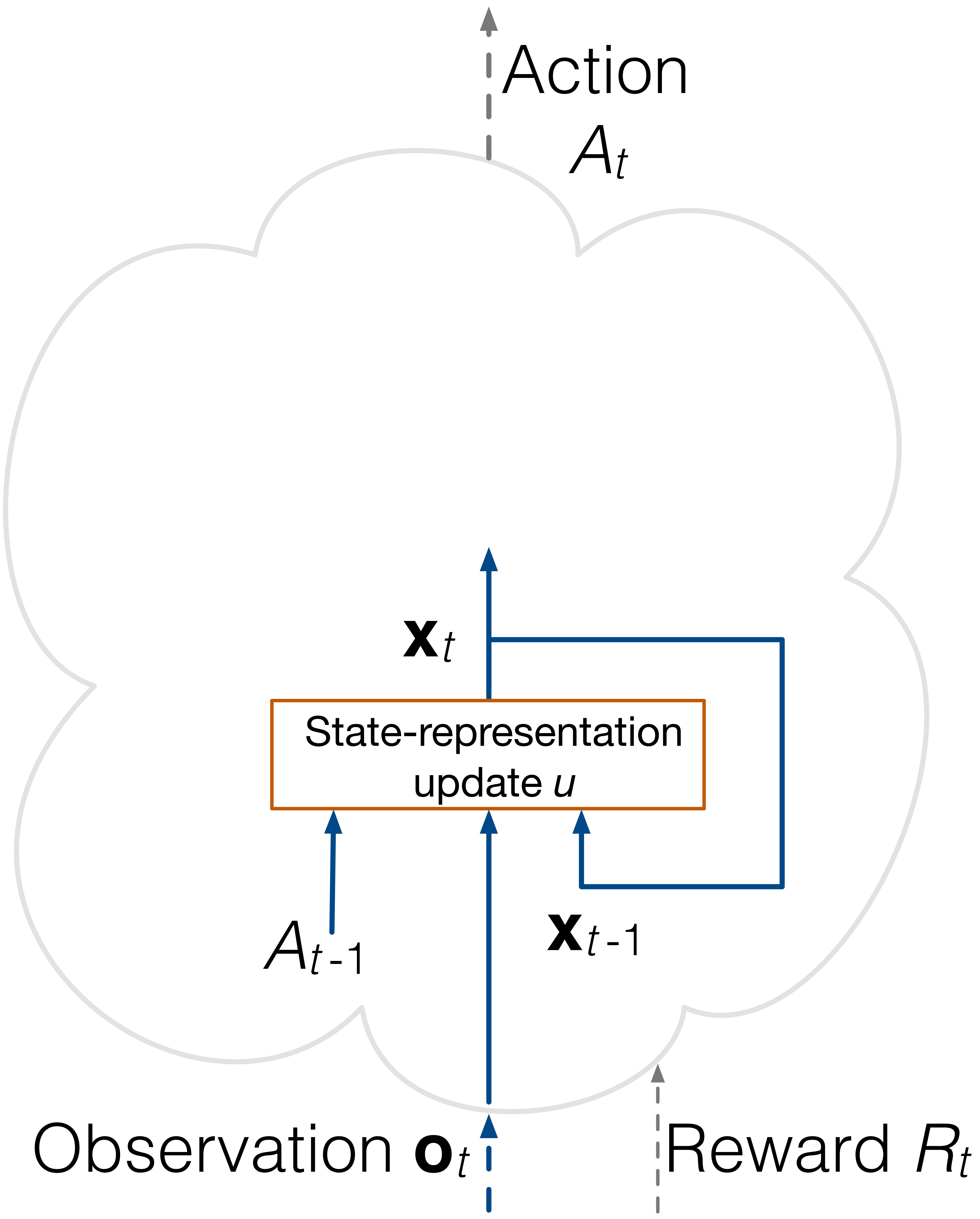
- ✓ Removing reliance on hand-coded solution
- ✓ Removing reliance on humans
- ✓ Adding learning capabilities for adaptability
- ✓ Aiming for solutions that work for different problems
- ✓ Aiming for solutions that work for different embodiments
- ✓ Aiming for low-level control
- ✓ Aiming for learning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

Can mind be specified as a single function?



Mind is a "stateful" process

Minimal processes of mind



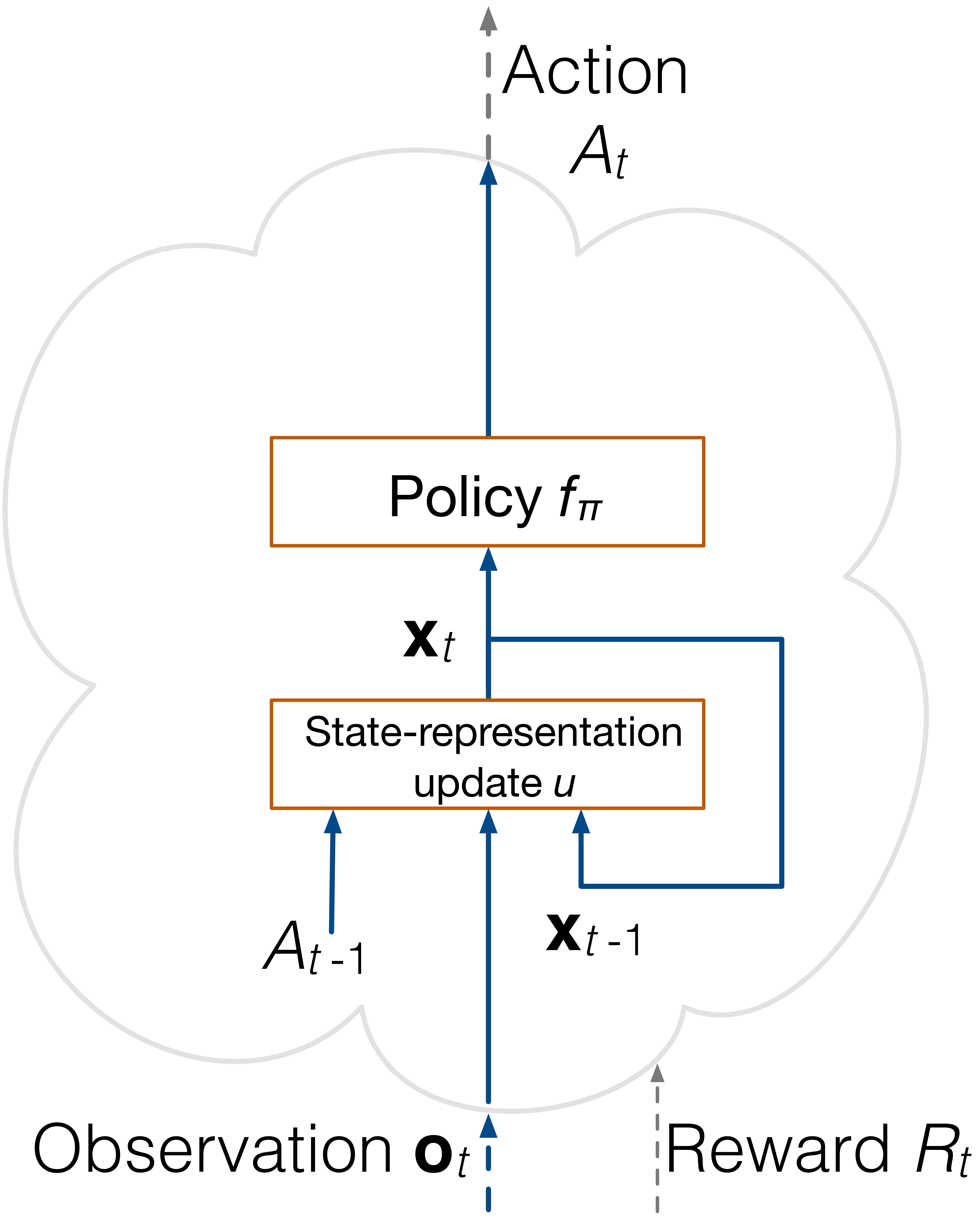
- ✓ State-representation update function:

  $\mathbf{x}_t = f(\mathbf{o}_t, A_{t-1}, \mathbf{o}_{t-1}, A_{t-2}, \ldots)$

- ✓ *Strictly incremental* computation: non-increasing computational and memory requirement per increment

  $\mathbf{x}_t = u(\mathbf{o}_t, A_{t-1}, \mathbf{x}_{t-1})$

Minimal processes of mind



- ✓ State-representation update function:

  $\mathbf{x}_t = u(\mathbf{o}_t, A_{t-1}, \mathbf{x}_{t-1})$

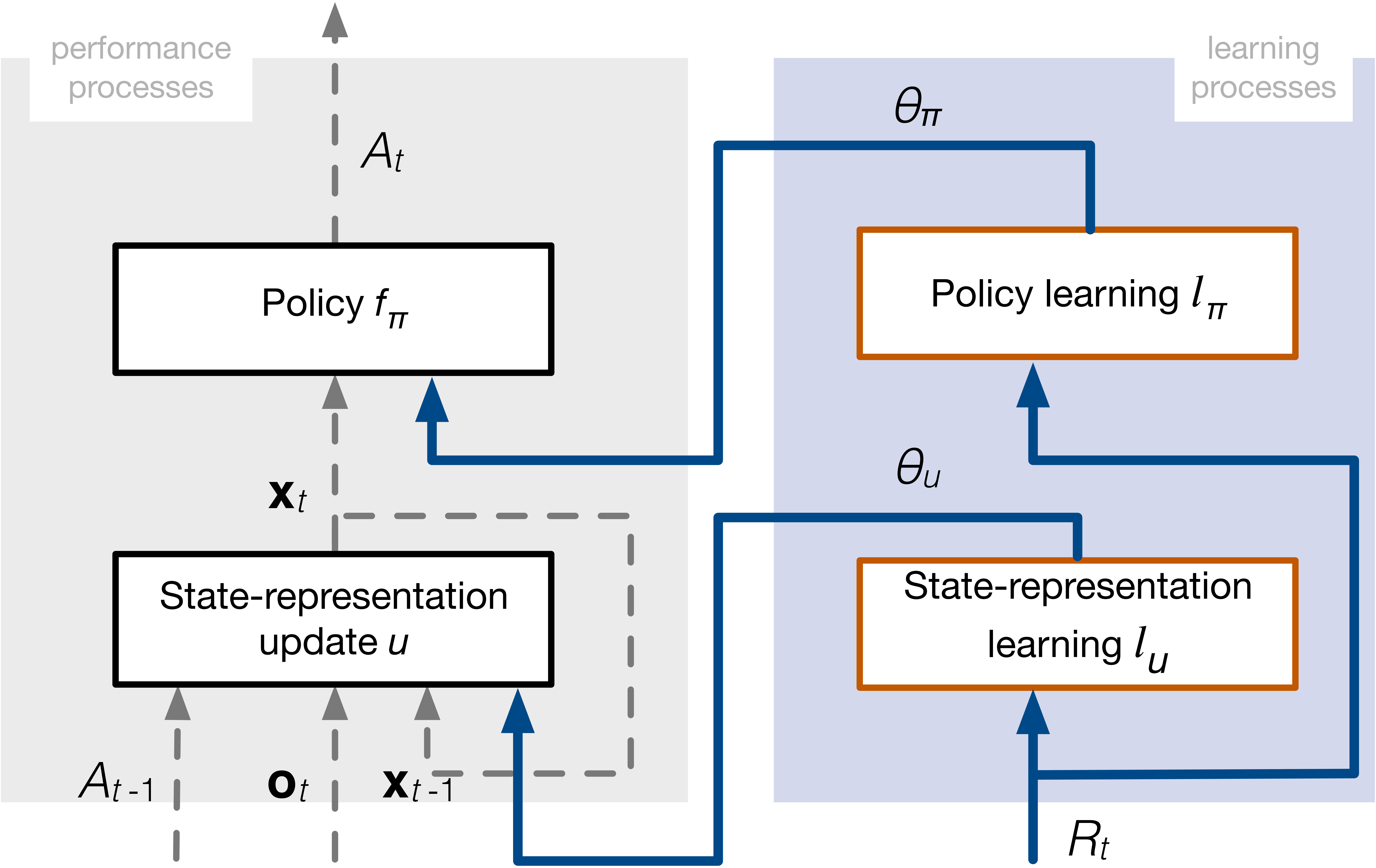
- ✓ Policy function:

  $A_t = f_\pi(\mathbf{x}_t)$

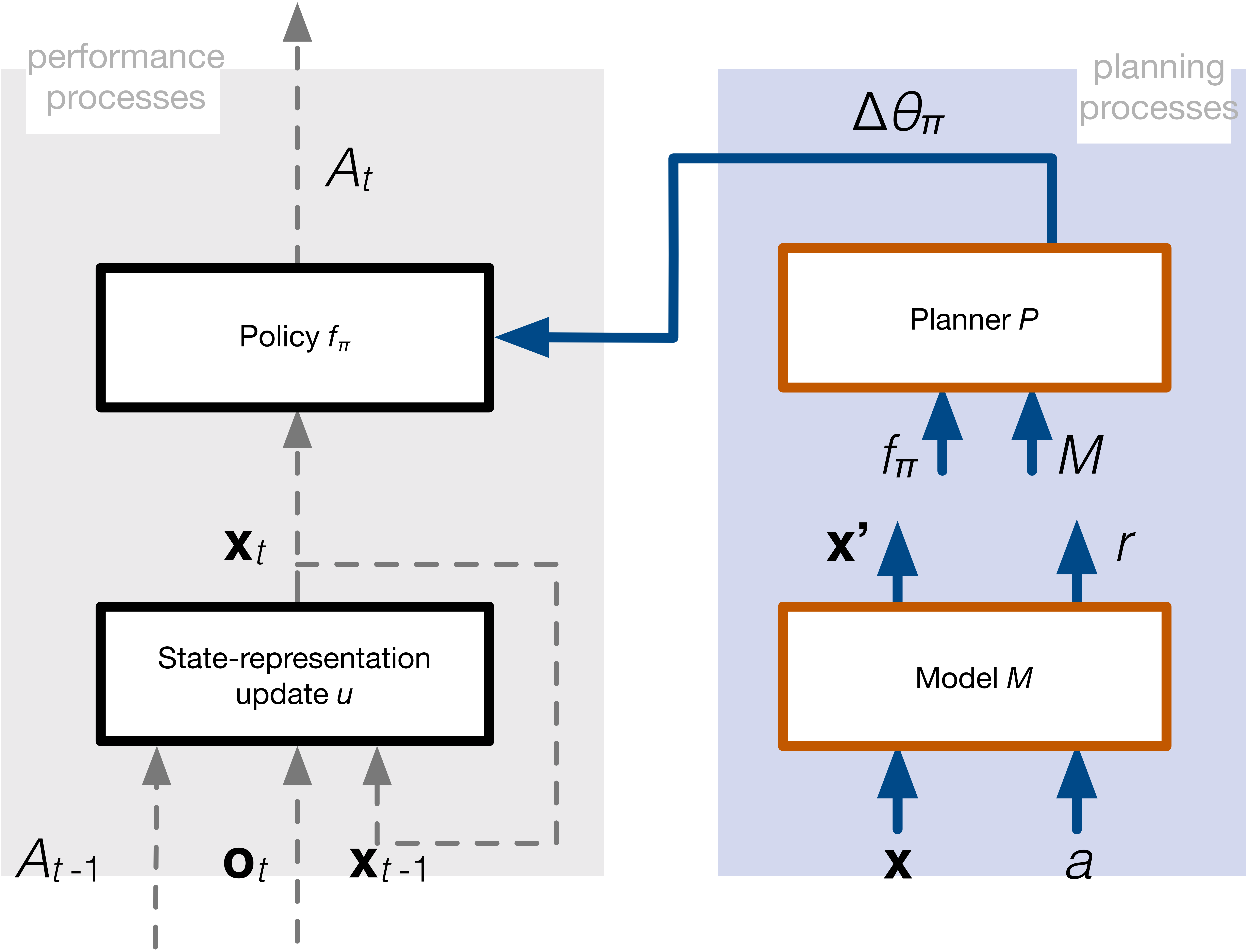
- ✓ These are *performance-processes* of mind

- ✓ These two functions cannot accommodate improvement, e.g., learning

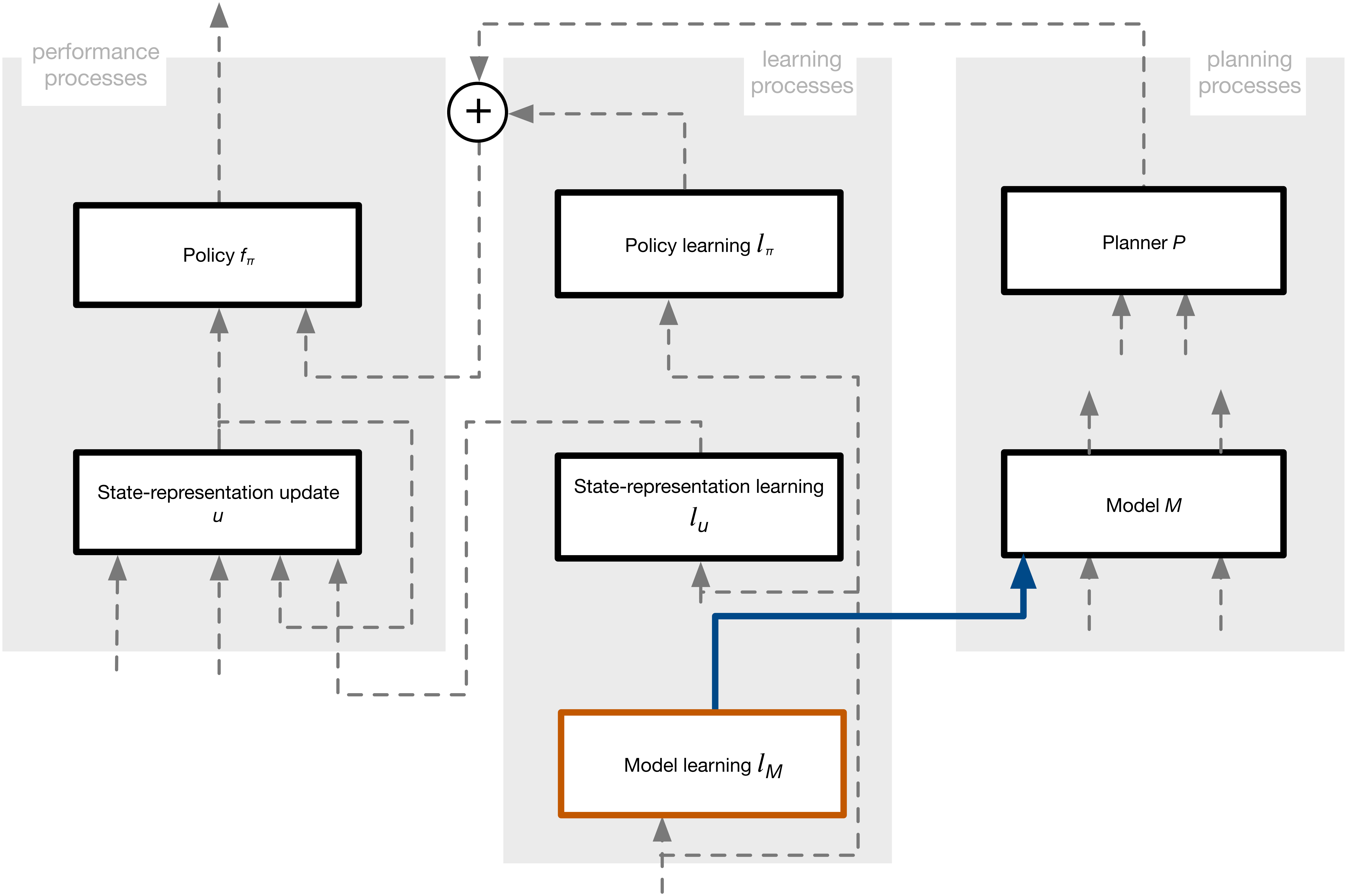
Learning processes in mind



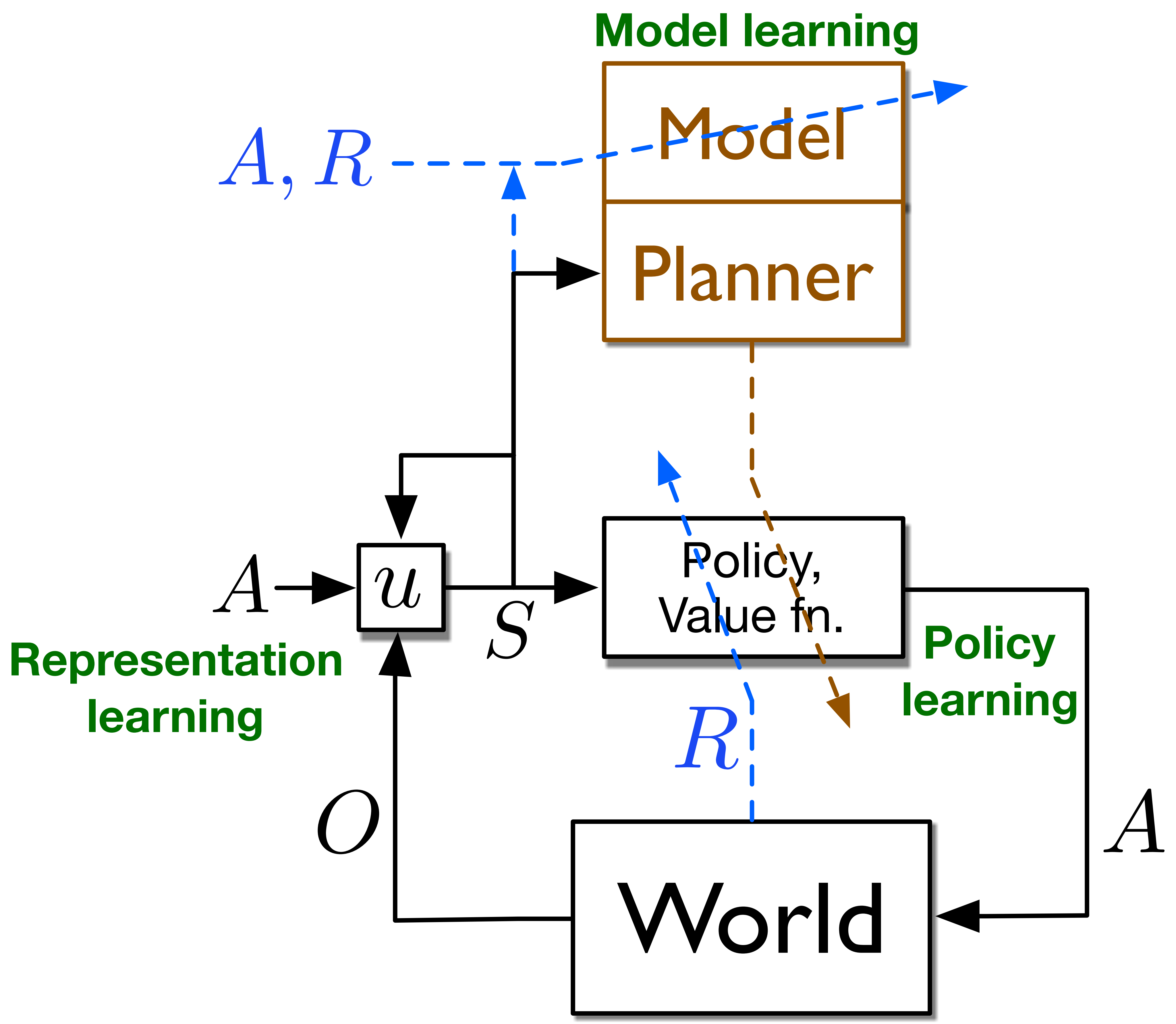
Planning processes in mind



ASCA: A Simple Cognitive Architecture of mind



Learning agents can be studied by decomposing into simpler problems

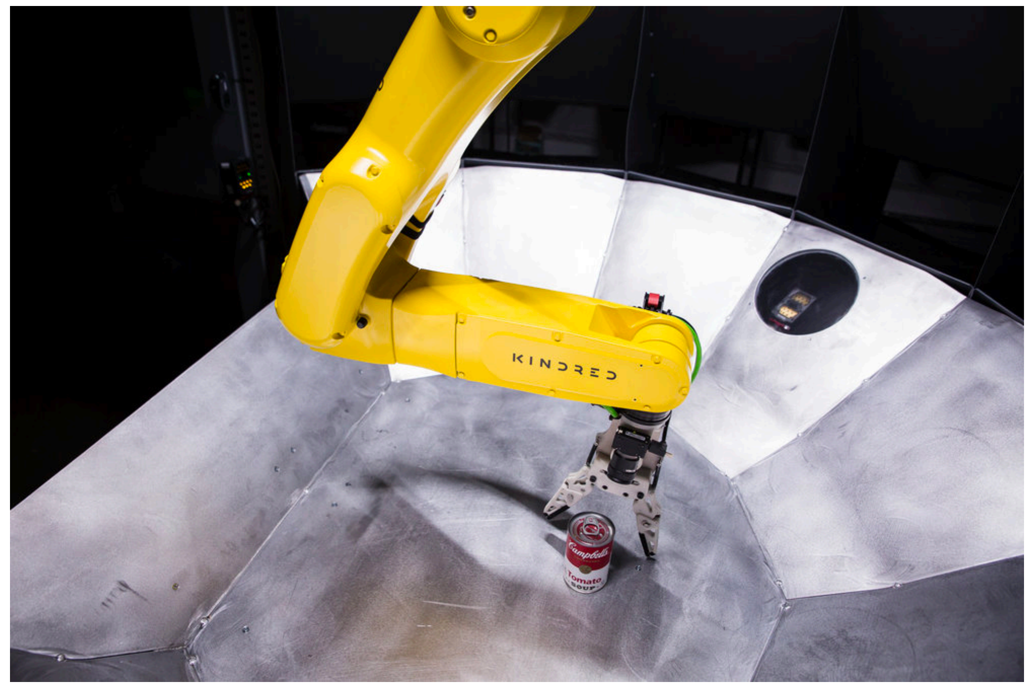


(Sutton & Barto 2018)

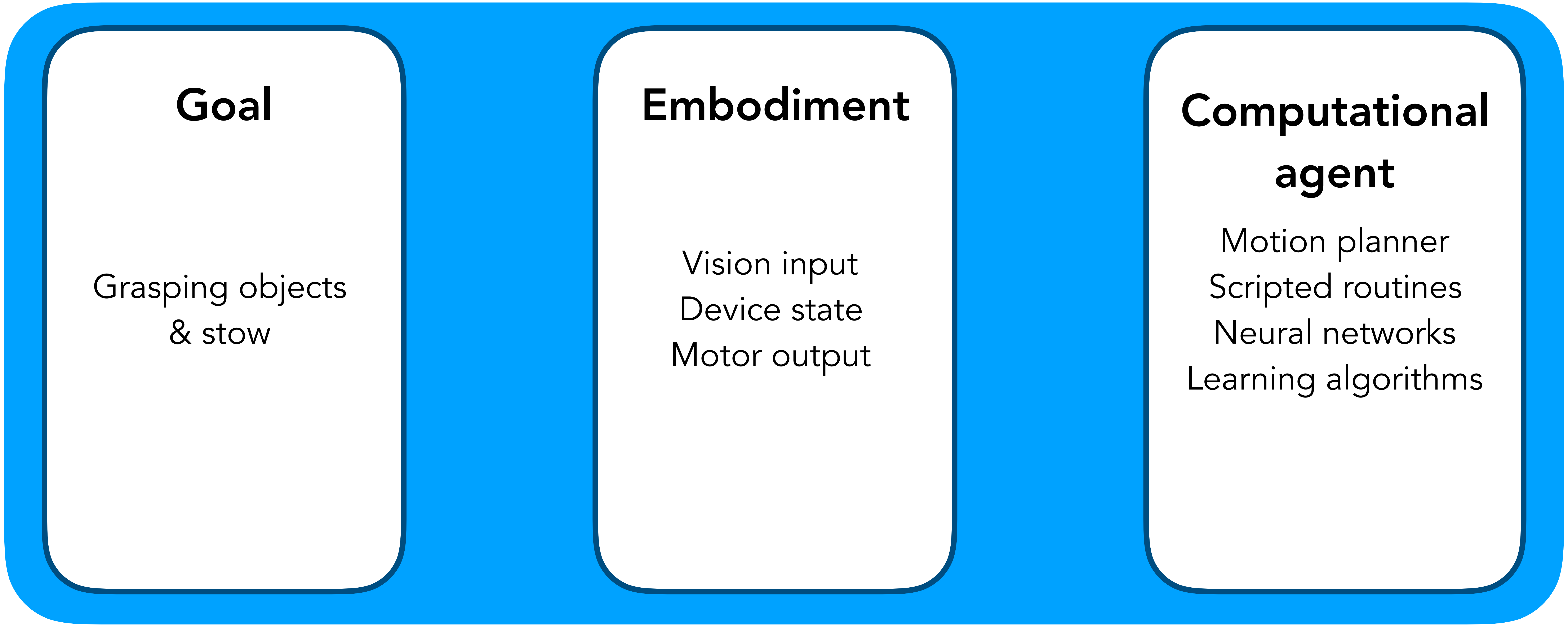
Moravec's paradox

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****.*

Using physical systems in AI research is essential



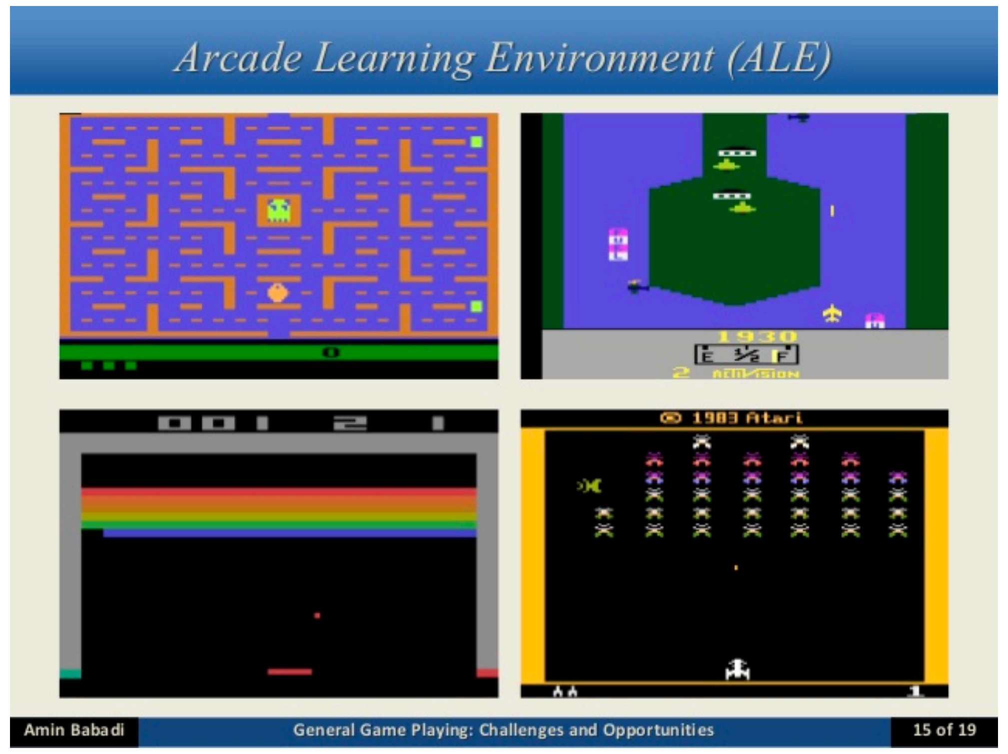
Kindred's robotic product: Sort



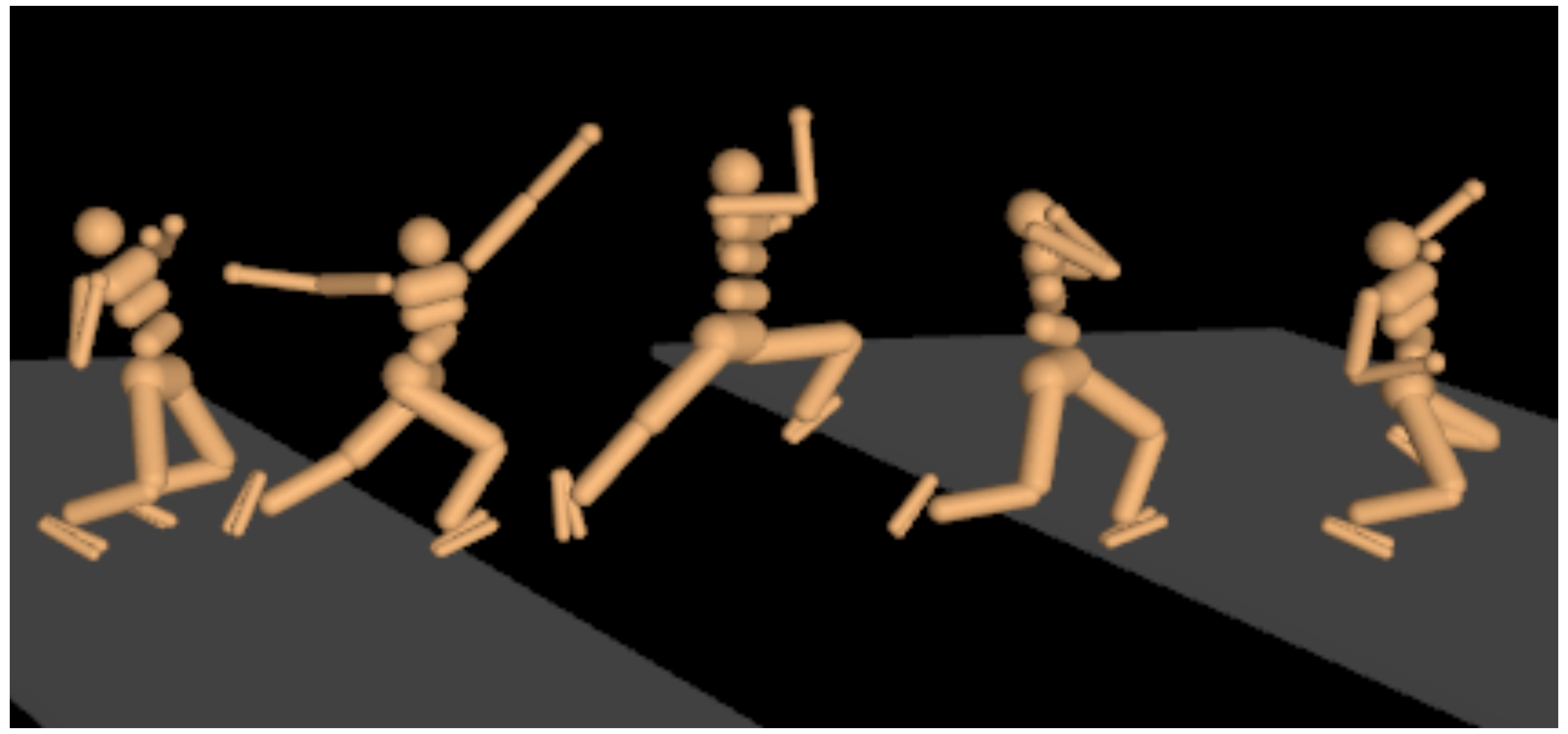
use a complete physical AI system

We are observing a rise in general-purpose solutions

AI beating Go champion



Human-level control of Atari games

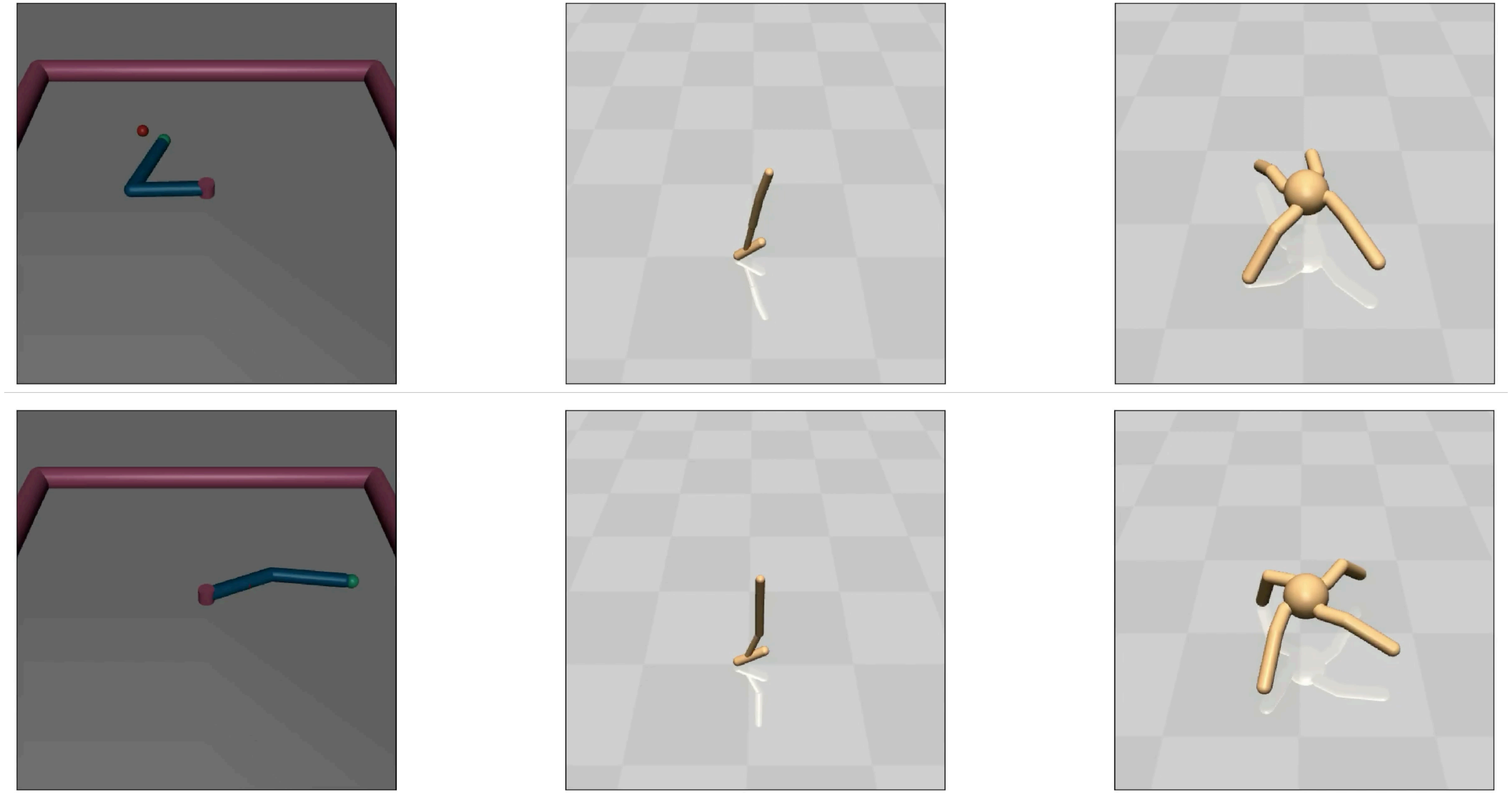


Emergence of locomotion behaviors

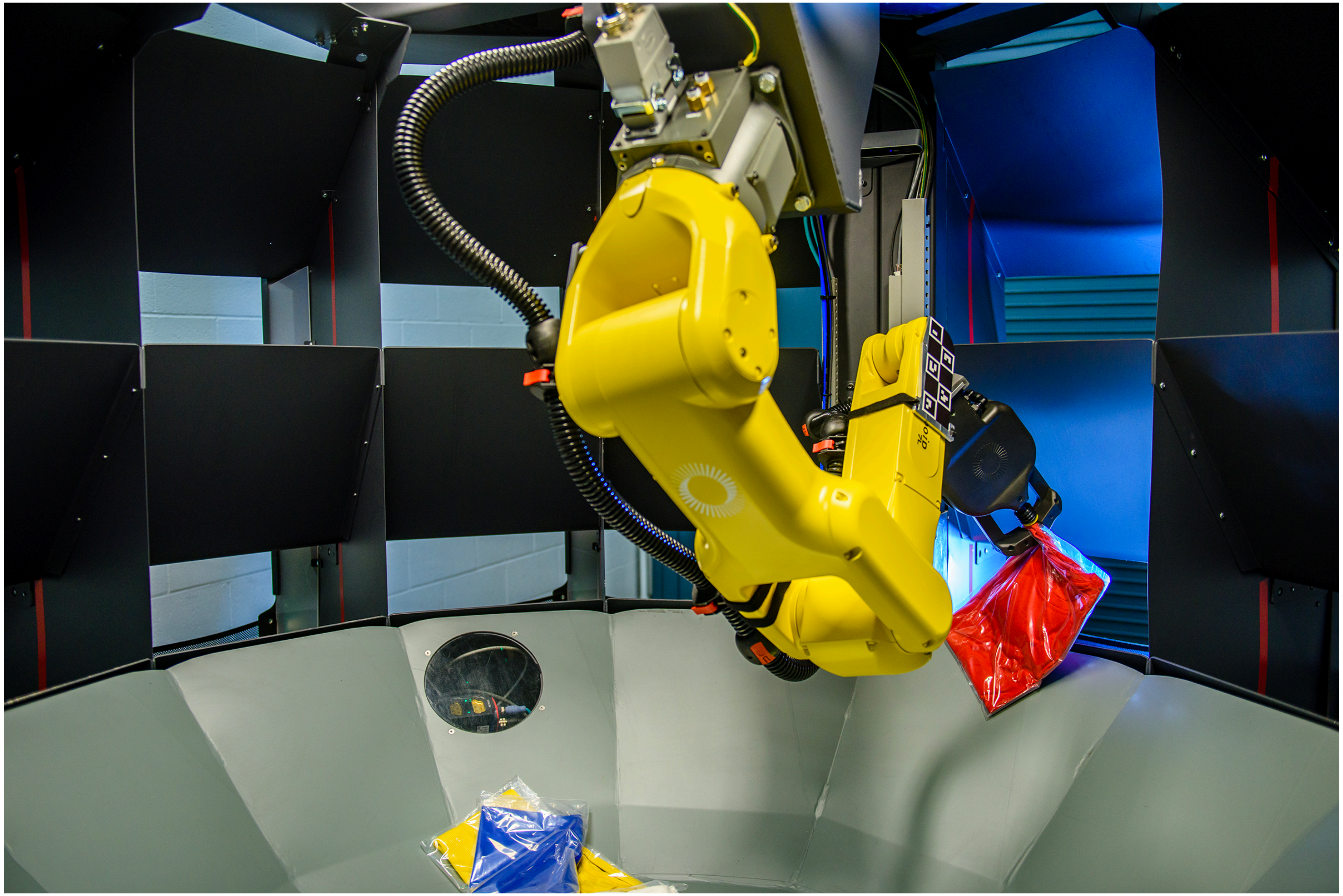
Deep RL agents show excellent general learning capabilities in virtual worlds

Initial behavior

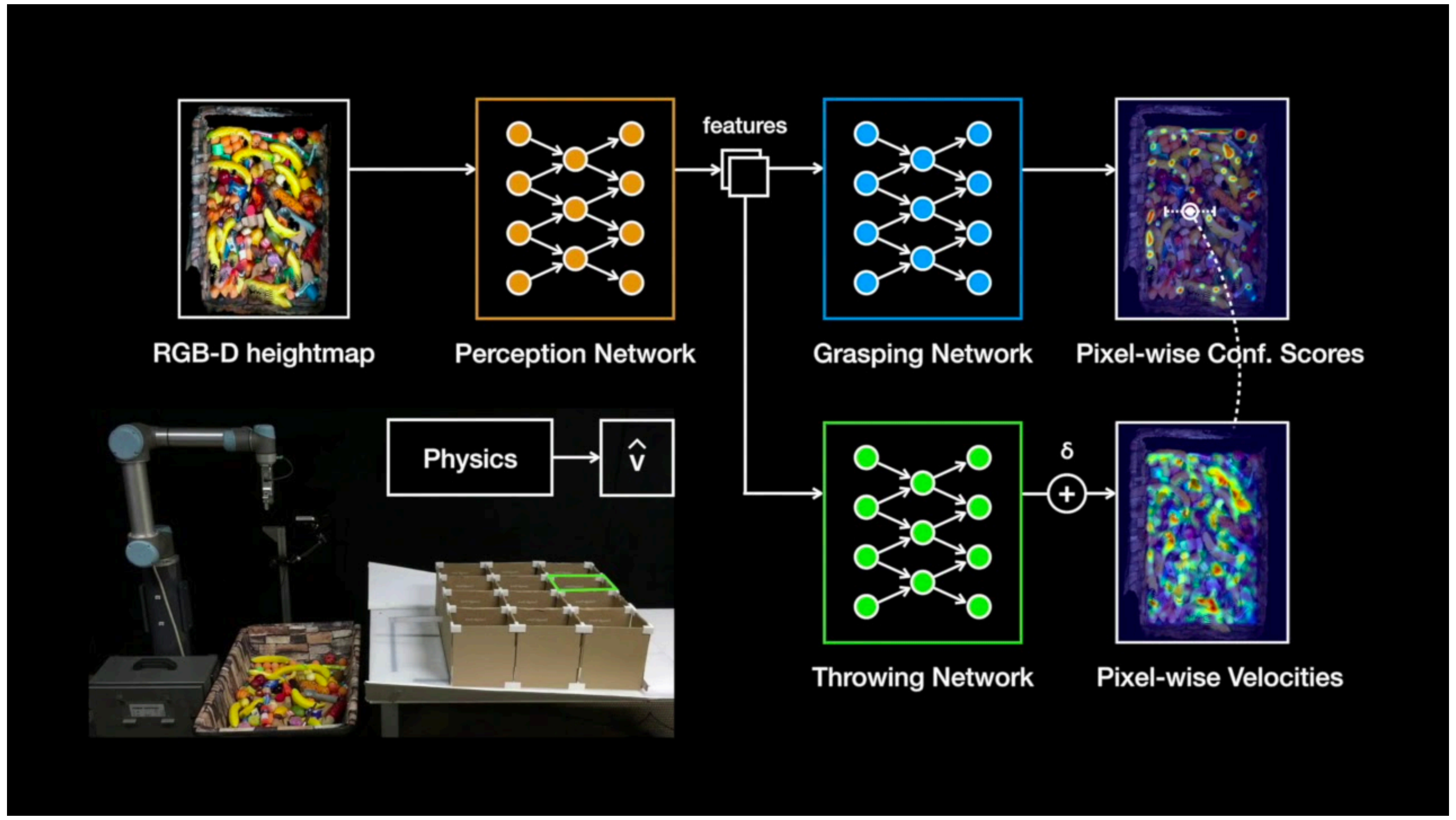
Learning with TRPO



Are deep RL agents ready to control real robots?



Hybrid approaches are becoming popular in robotics



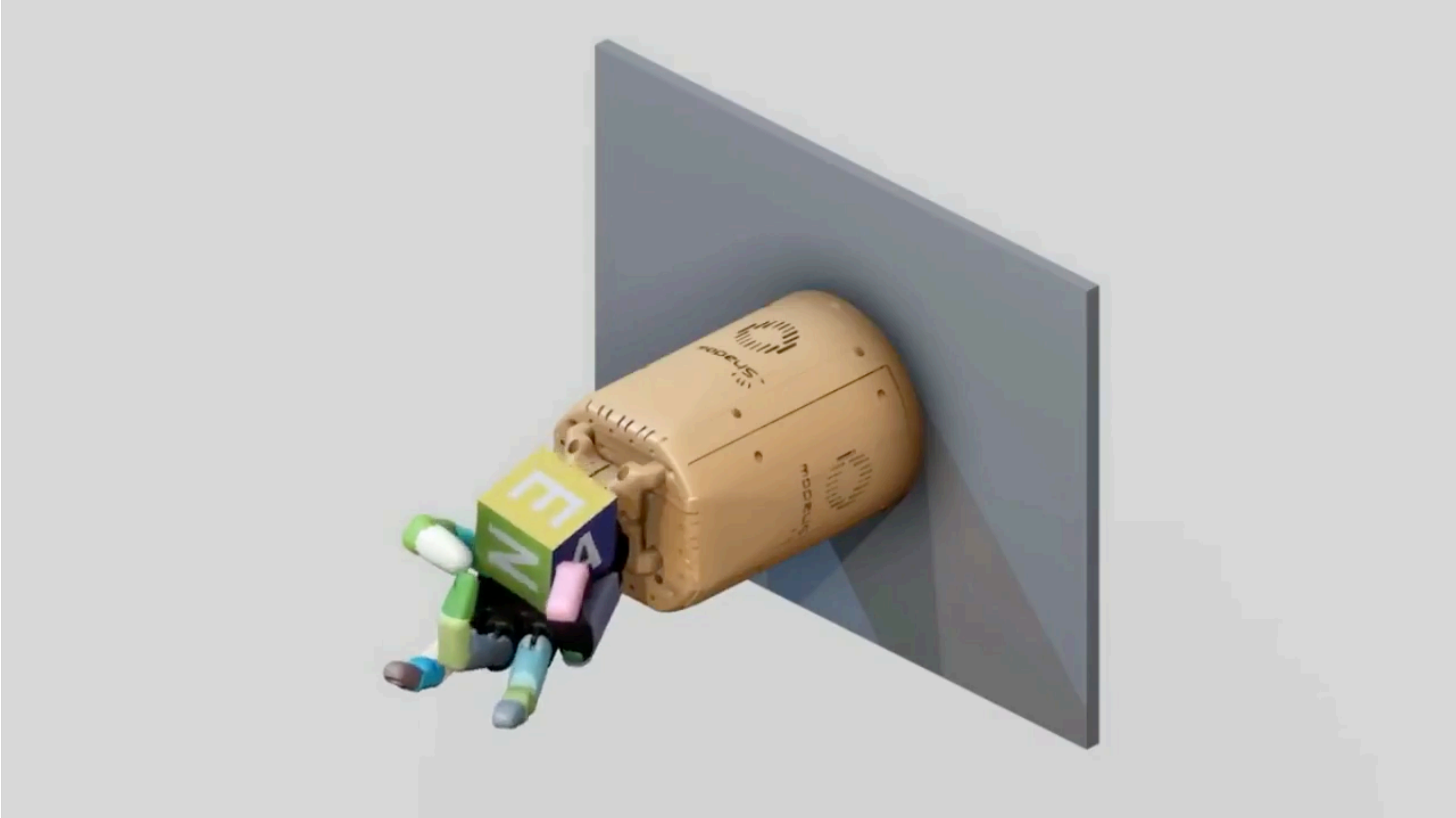
Zeng A, Song S, Lee J, Rodriguez A, Funkhouser T (2019).
TossingBot: Learning to Throw Arbitrary Objects with Residual Physics.

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

Imitation learning

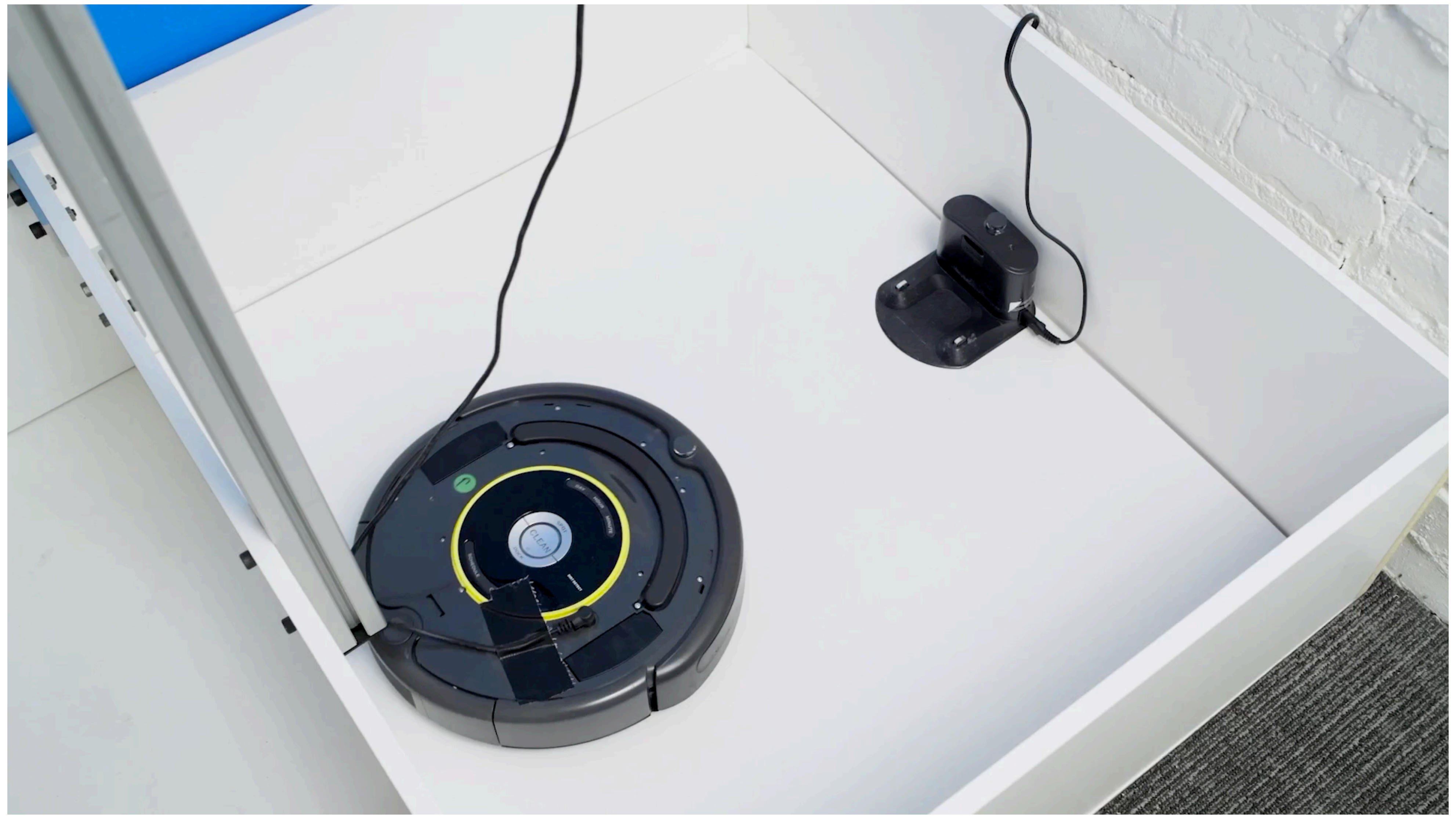
Kormushev P, Calinon S, Caldwell DG (2010).
Robot Motor Skill Coordination
with EM-based Reinforcement Learning

Simulation-to-reality (Sim2Real) transfer

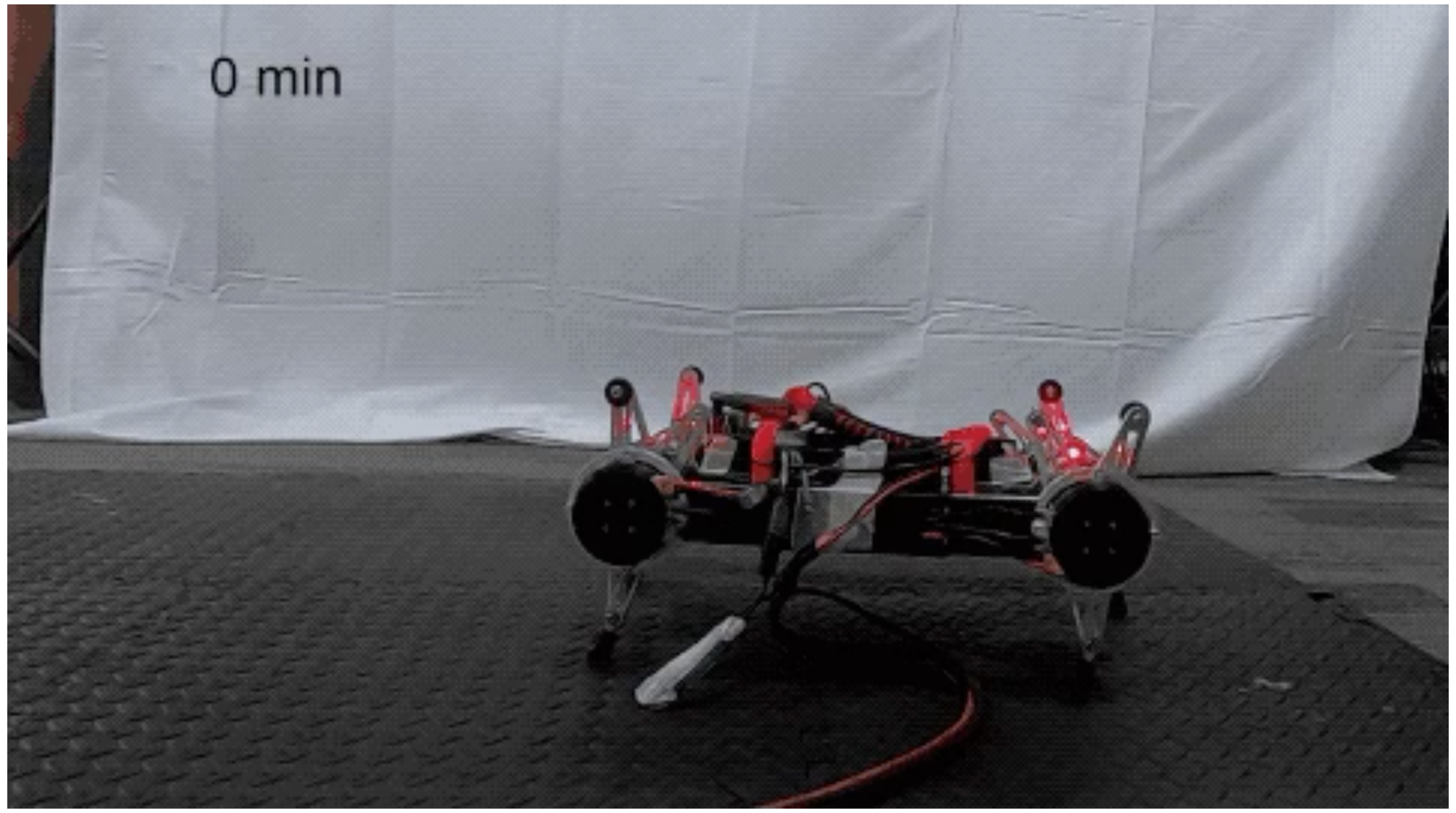


Andrychowicz et al. (2018)
Learning dexterous in-hand manipulation

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.



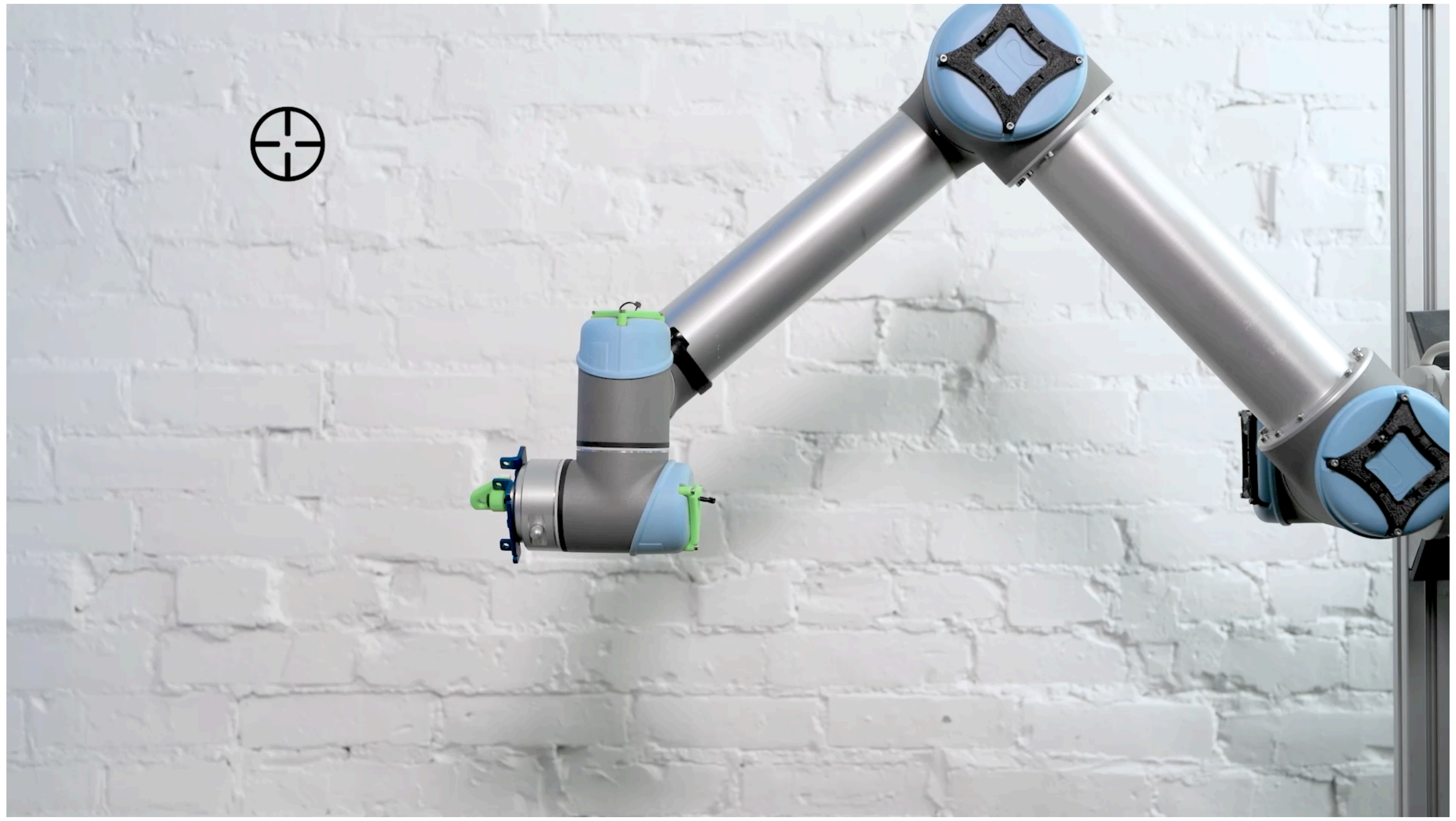
Haarnoja T, Zhou A, Ha S, Tan J, Tucker G, Levine S (2018).
Learning to walk via deep reinforcement learning.

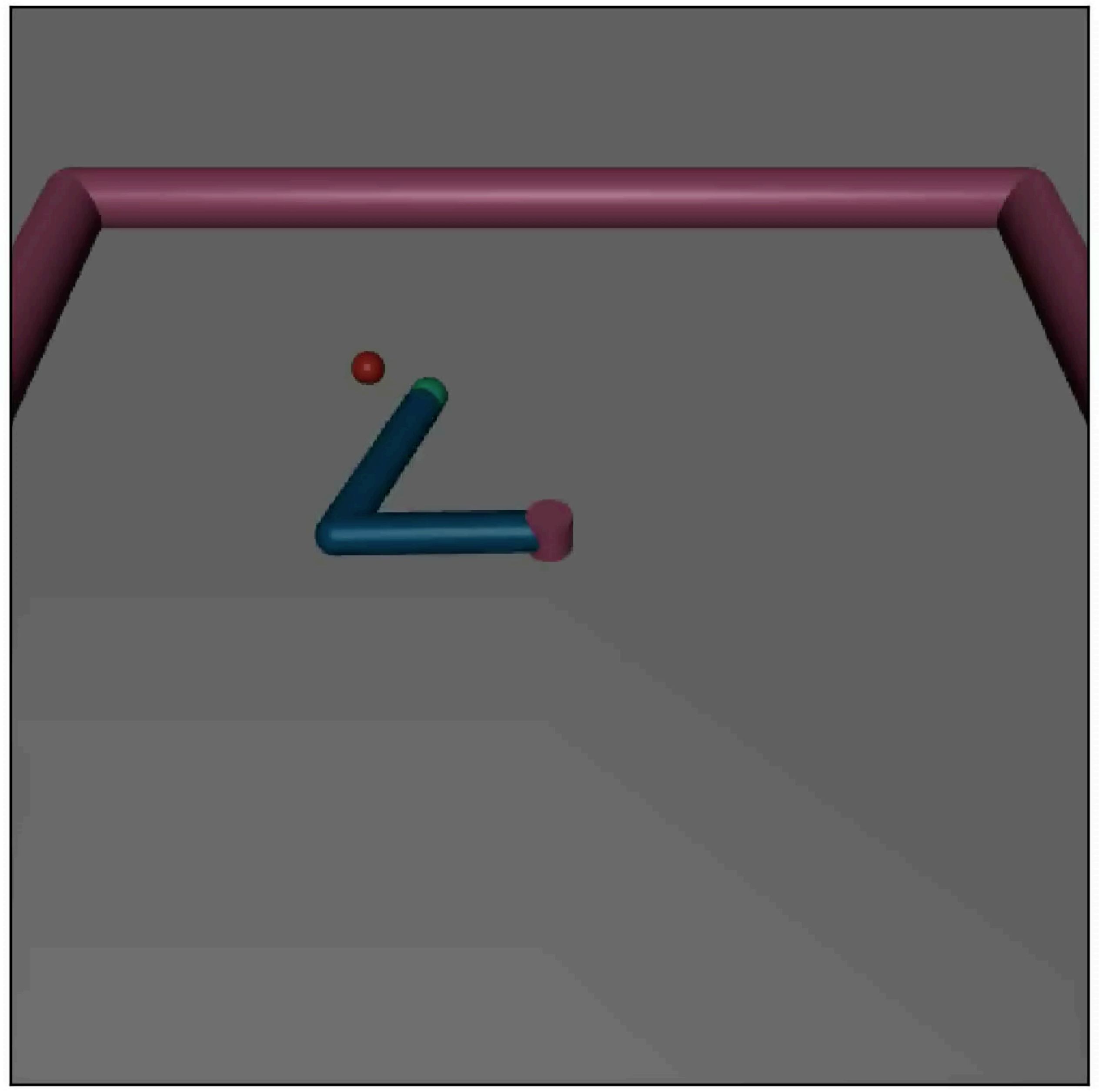
Learning directly with physical robots is substantially more difficult

Commonly-cited factors

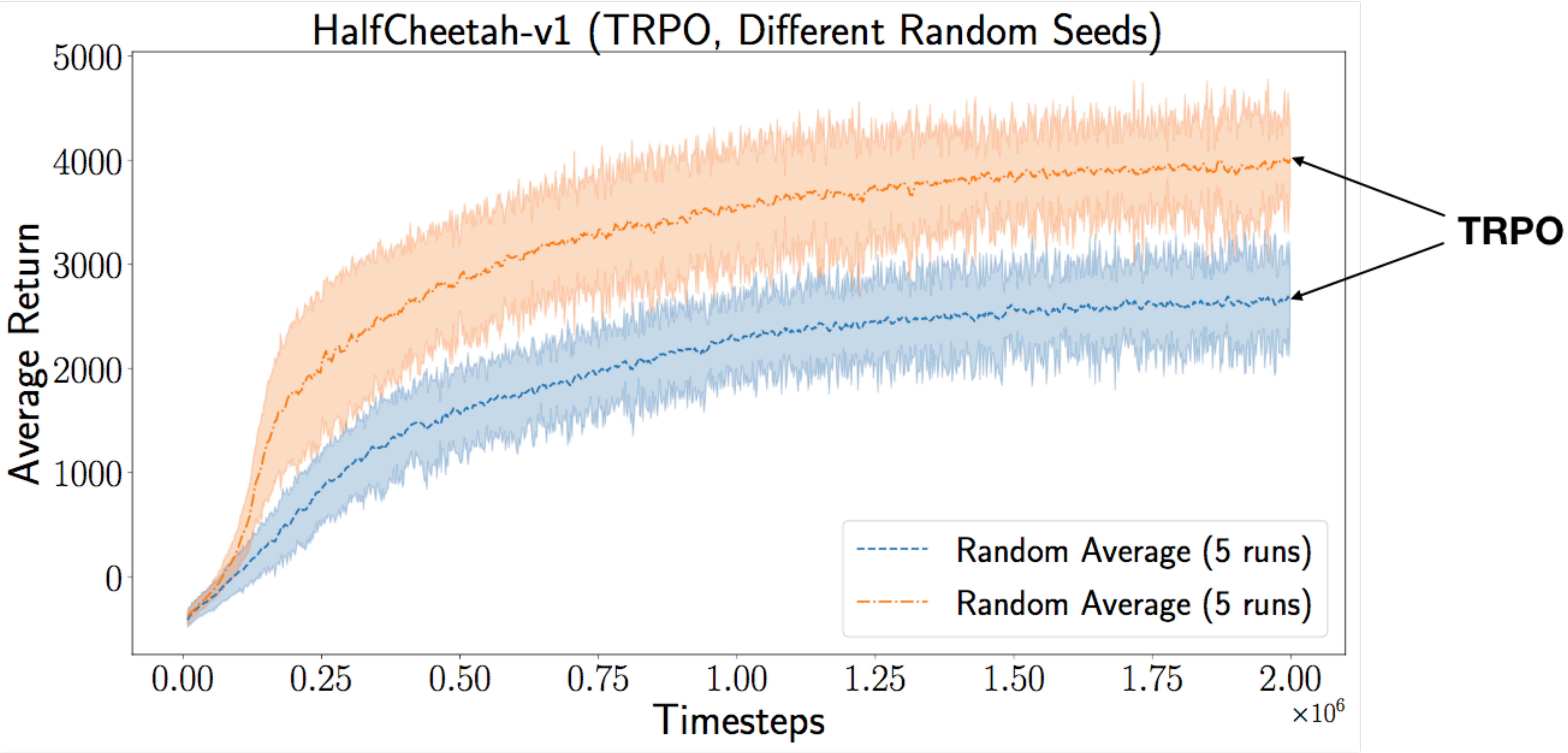
- Slow data-collection rate
- Partial or noisy observations
- Safety and frailty of real devices

Behavior of a random agent



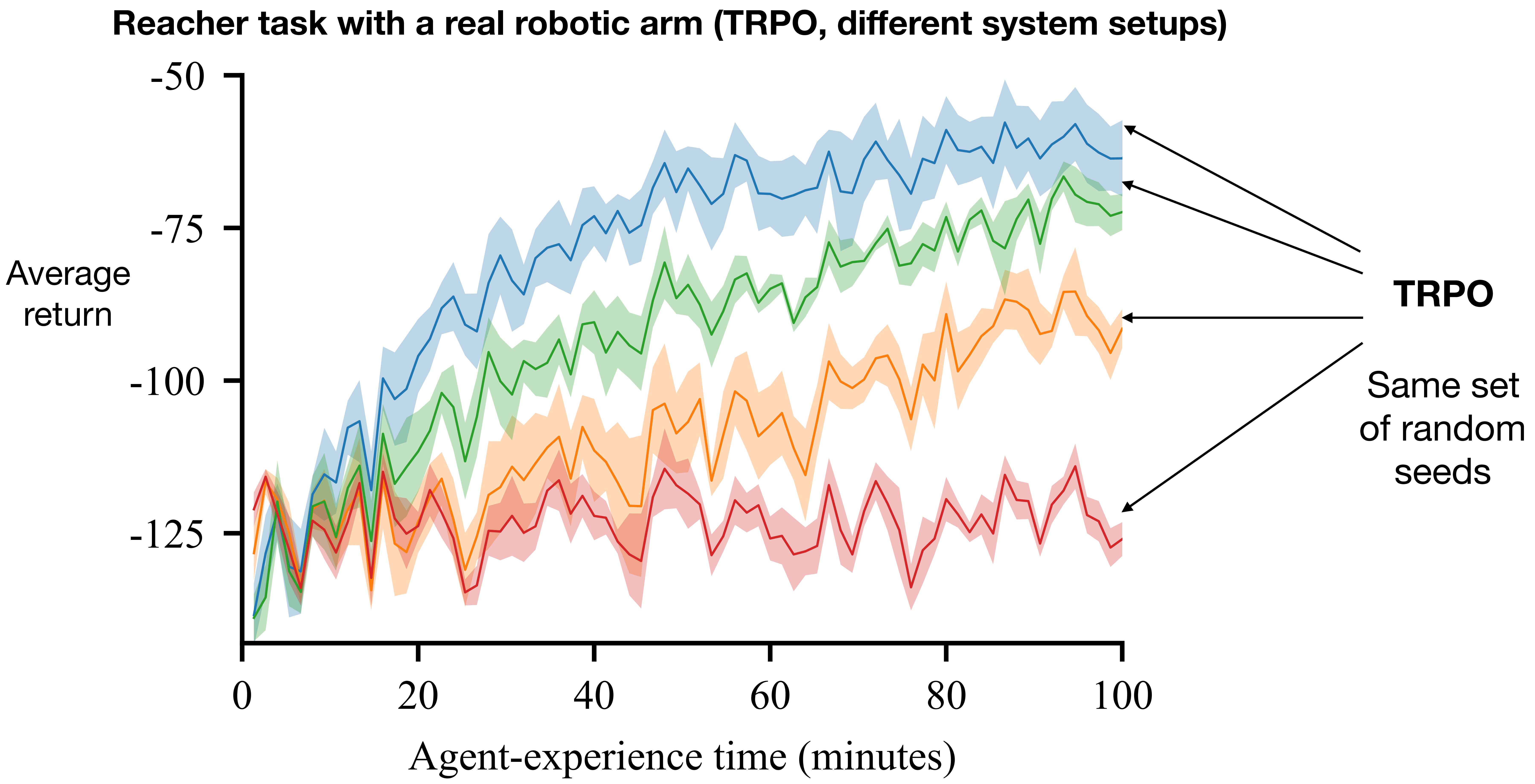


Without proper guidelines, reducing experimental uncertainties is hard even in simulations



Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., Meger, D. (2018). Deep reinforcement learning that matters. In *AAAI*.

We are yet to discover the unknowns involved in experiments with physical robots



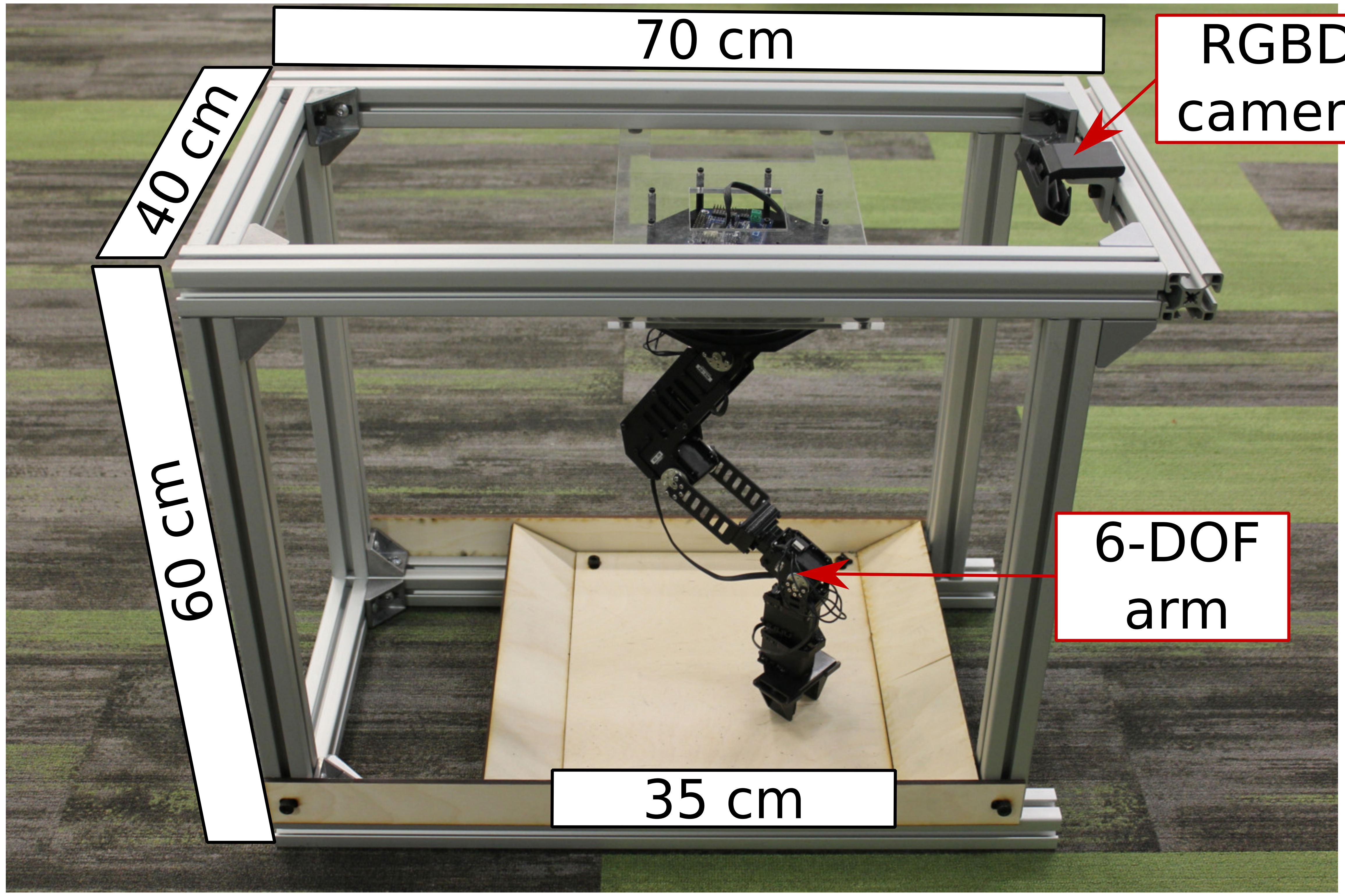
Mahmood, A. R., Korenkevych, D., Komer, B. J., Bergstra, J. (2018). Setting up a reinforcement learning task with a real-world robot. In *IROS*.

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 available from UC Berkeley

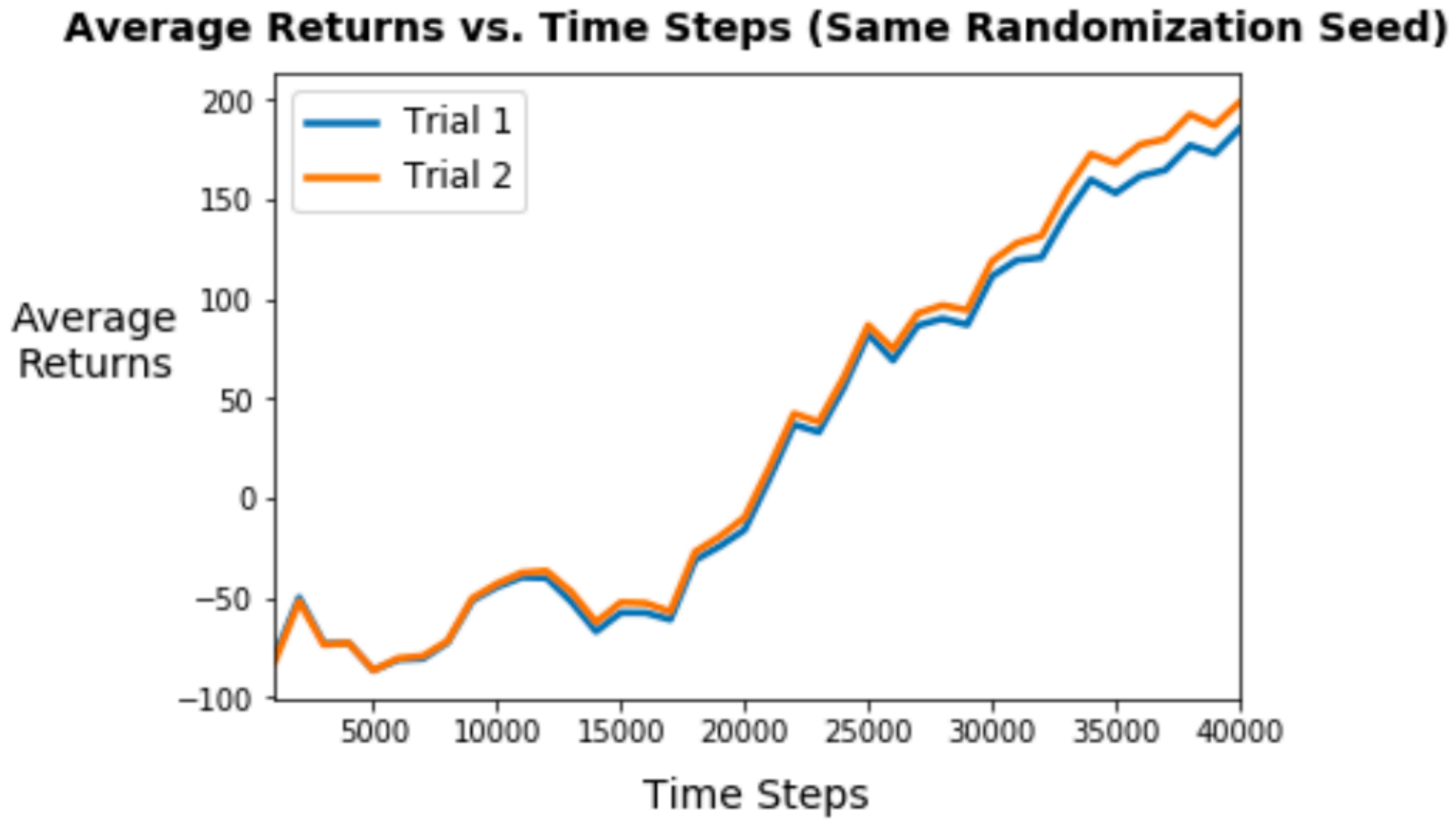


Yang B, Zhang J, Pong V, Levine S, Jayaraman D (2019).
REPLAB: A Reproducible Low-Cost Arm Benchmark Platform for Robotic Learning.

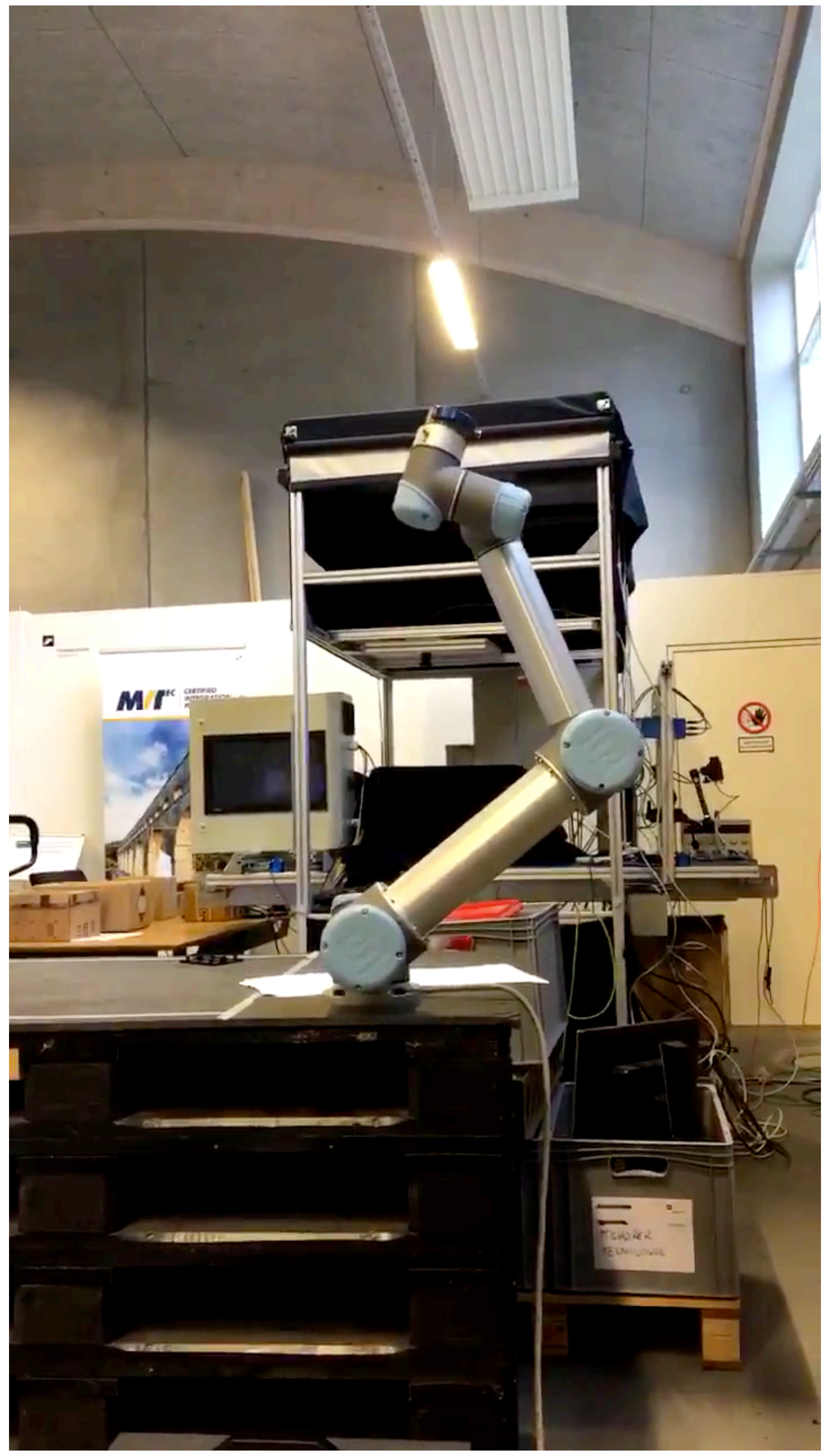
Kindred introduced *SenseAct*, a benchmark task suite for reproducible RL research with real robots

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