Open-ended Agent Learning in the Era of Foundation Models

Jeff Clune

@jeffclune

Canada CIFAR AI Chair
Vector Institute

Associate Professor, Computer Science
University of British Columbia

Senior Research Advisor
DeepMind

Today

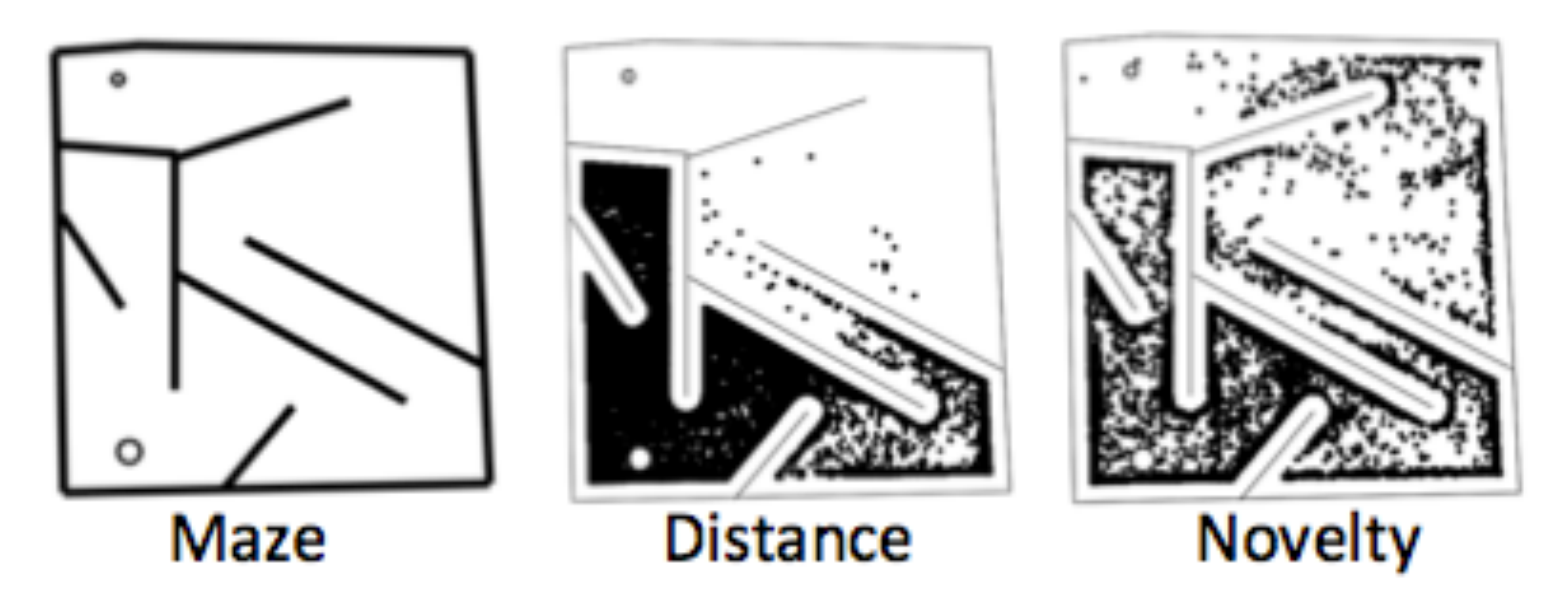
- Motivate
  - Quality-Diversity algorithms
  - Open-ended algorithms
  - AI-generating algorithms
- Discuss new possibilities unlocked by “foundation models”
  - Aka LLMs, LMs, VLMs, Generative AI, Frontier models, GPT, pretraining, etc.

AI Safety

- An important topic
- One I take seriously, and encourage you to as well
  - Bengio, Hinton, et al. 2024, Managing AI Risks in an Era of Rapid Progress. Science
- There are unique safety issues for AI-GAs and OE algorithms
  - Clune 2019, AI-Generating Algorithms
  - Ecoffet, Clune, Lehman 2020, Open Questions in Creating Safe Open-ended AI: Tensions Between Control and Creativity
- We must proceed cautiously and wisely, and prioritize safety
- This talk is not on safety (requires its own focus)
  - But please ask questions

A Paradox

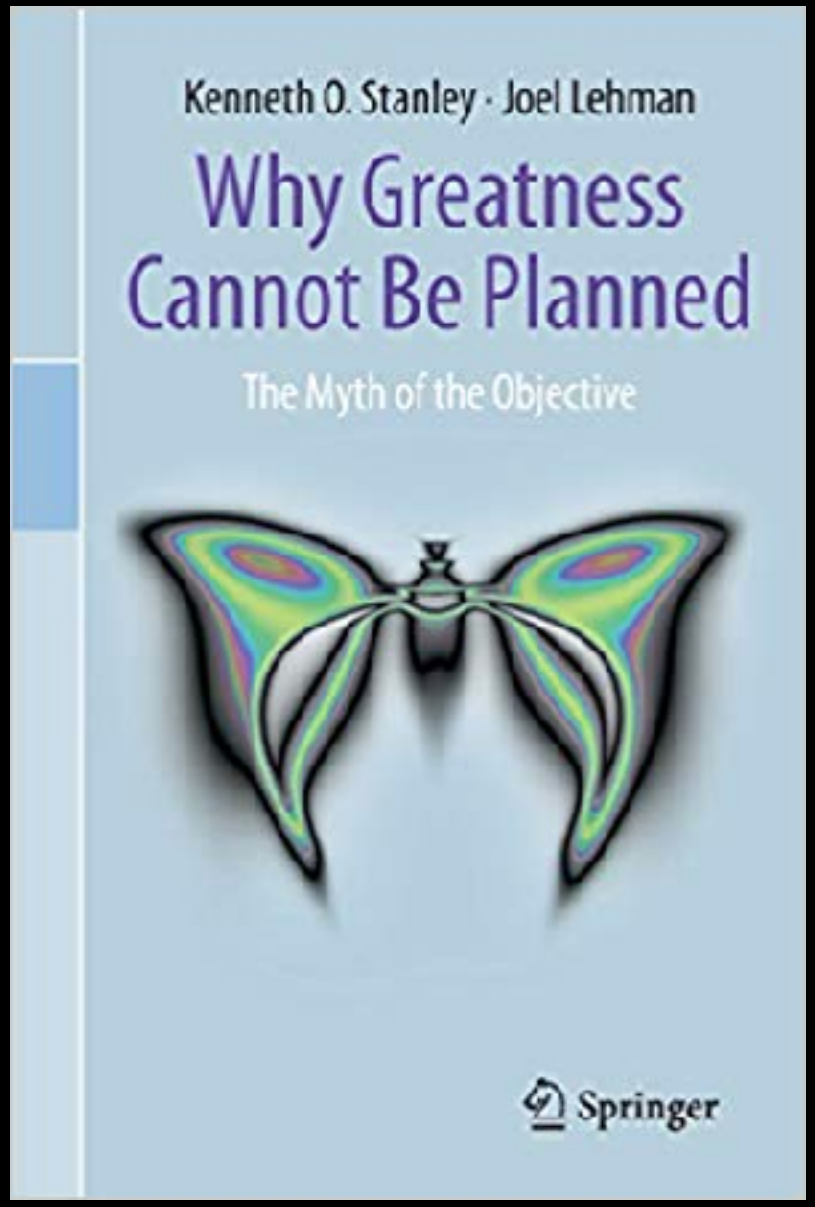
- If you try too hard to solve a hard problem, you'll fail
- If you ignore the objective, you're more likely to succeed

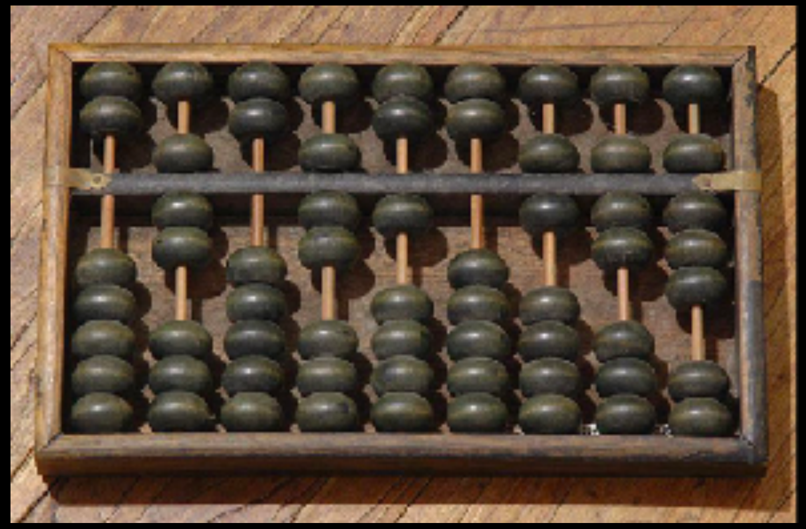


Maximize Reward: Fails
Maximize Novelty: Succeeds

Novelty Search: Lehman & Stanley

Key for Science & Technological Innovation: Generating Problems, Goal Switching



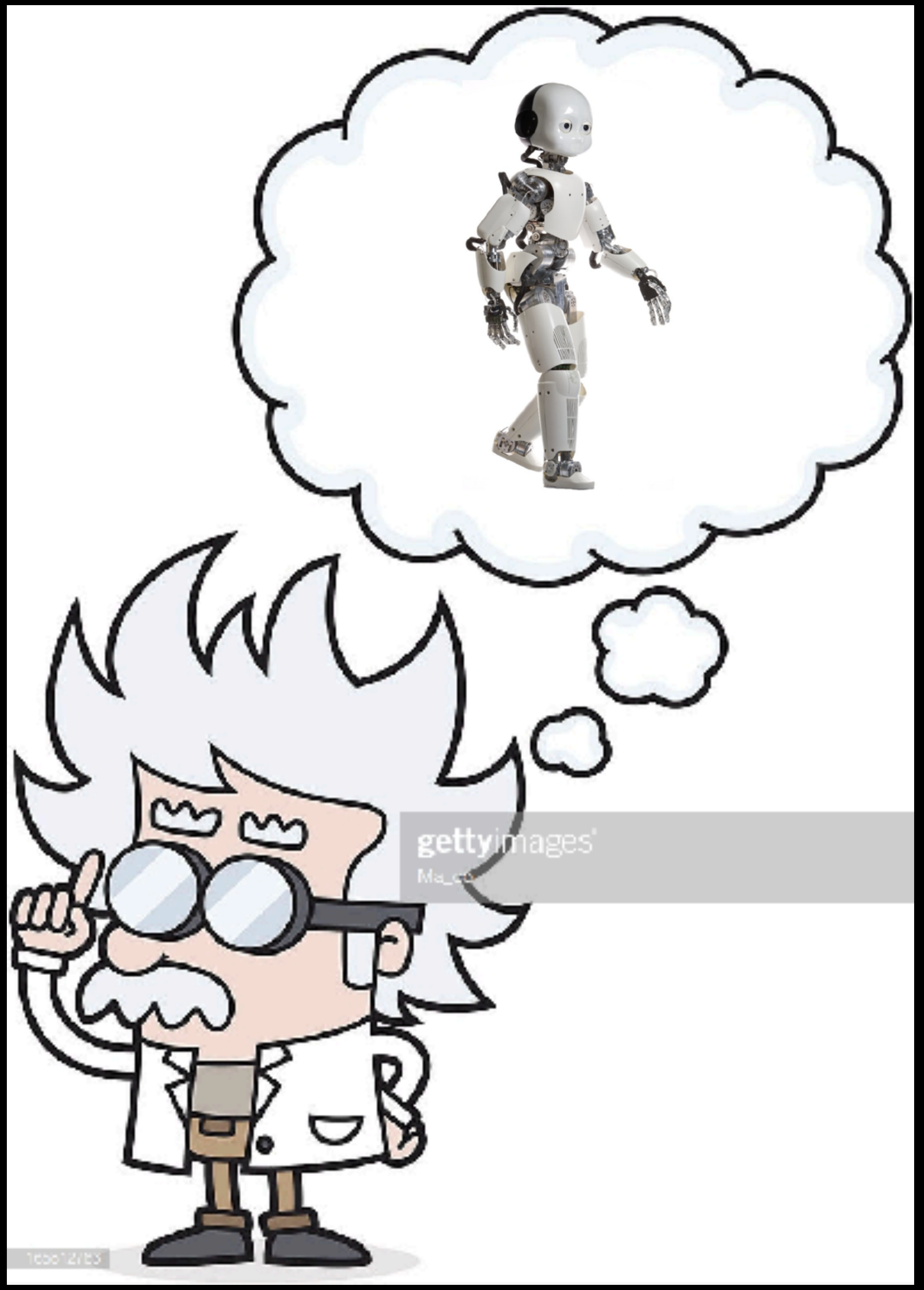


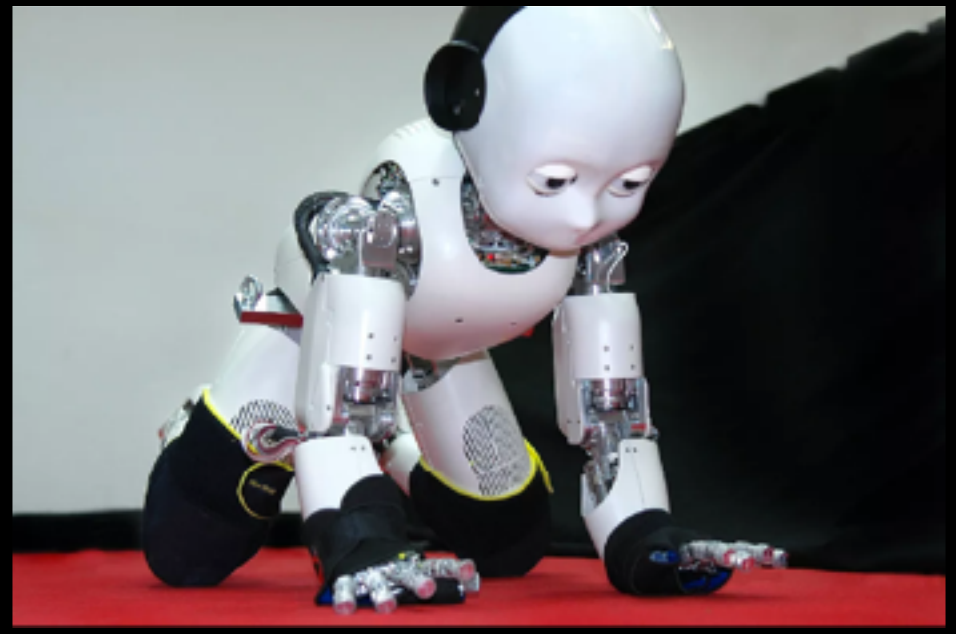
Conjecture: The only way to solve hard problems may be by creating problems while you solve them and goal switching between them

"Goal Switching"

Nguyen, Yosinski & Clune 2016

- When trying to solve task A, if you make progress on task B
  - also start optimizing for B

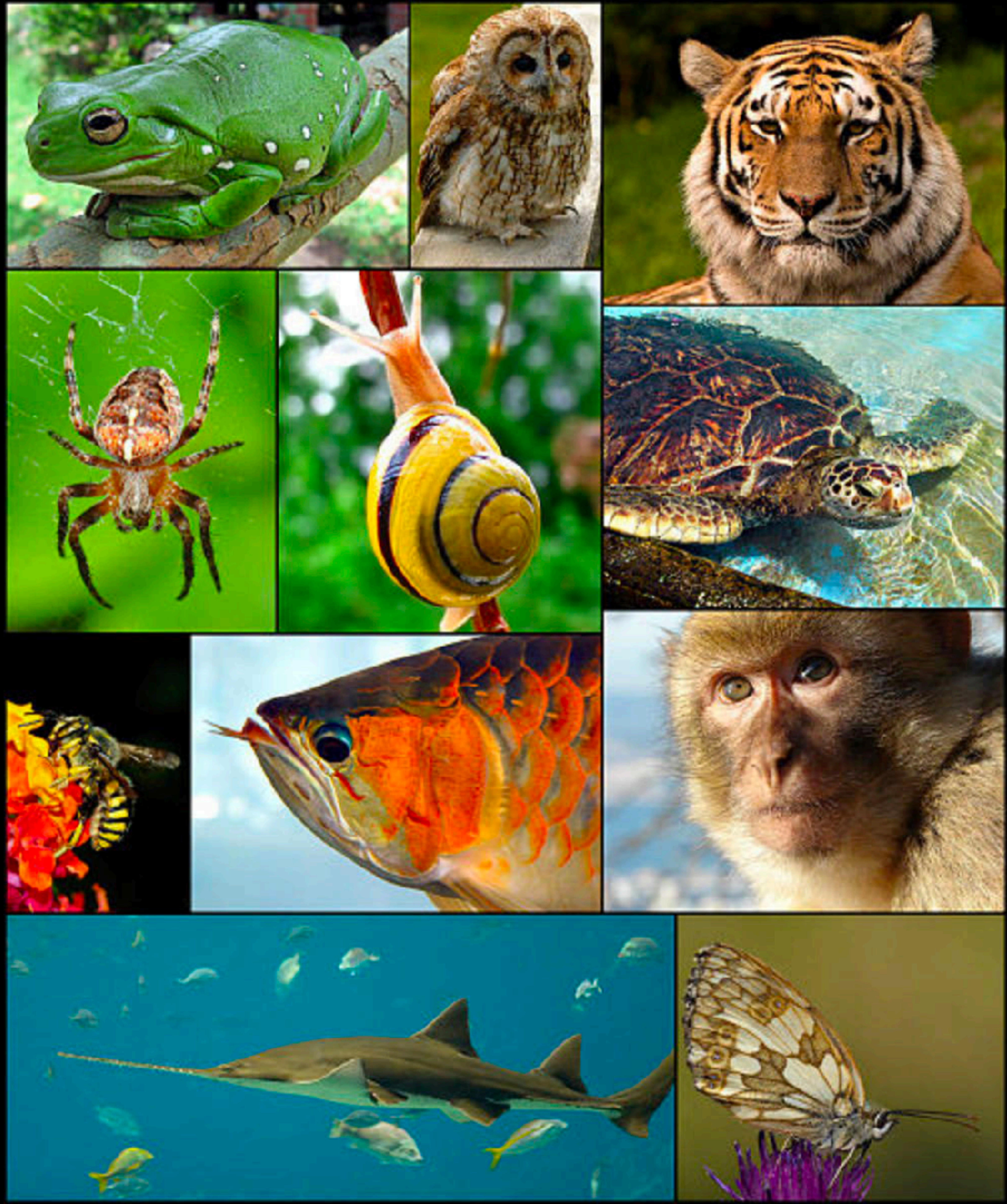




We want our algorithms to capture serendipitous discoveries

Quality Diversity Algorithms

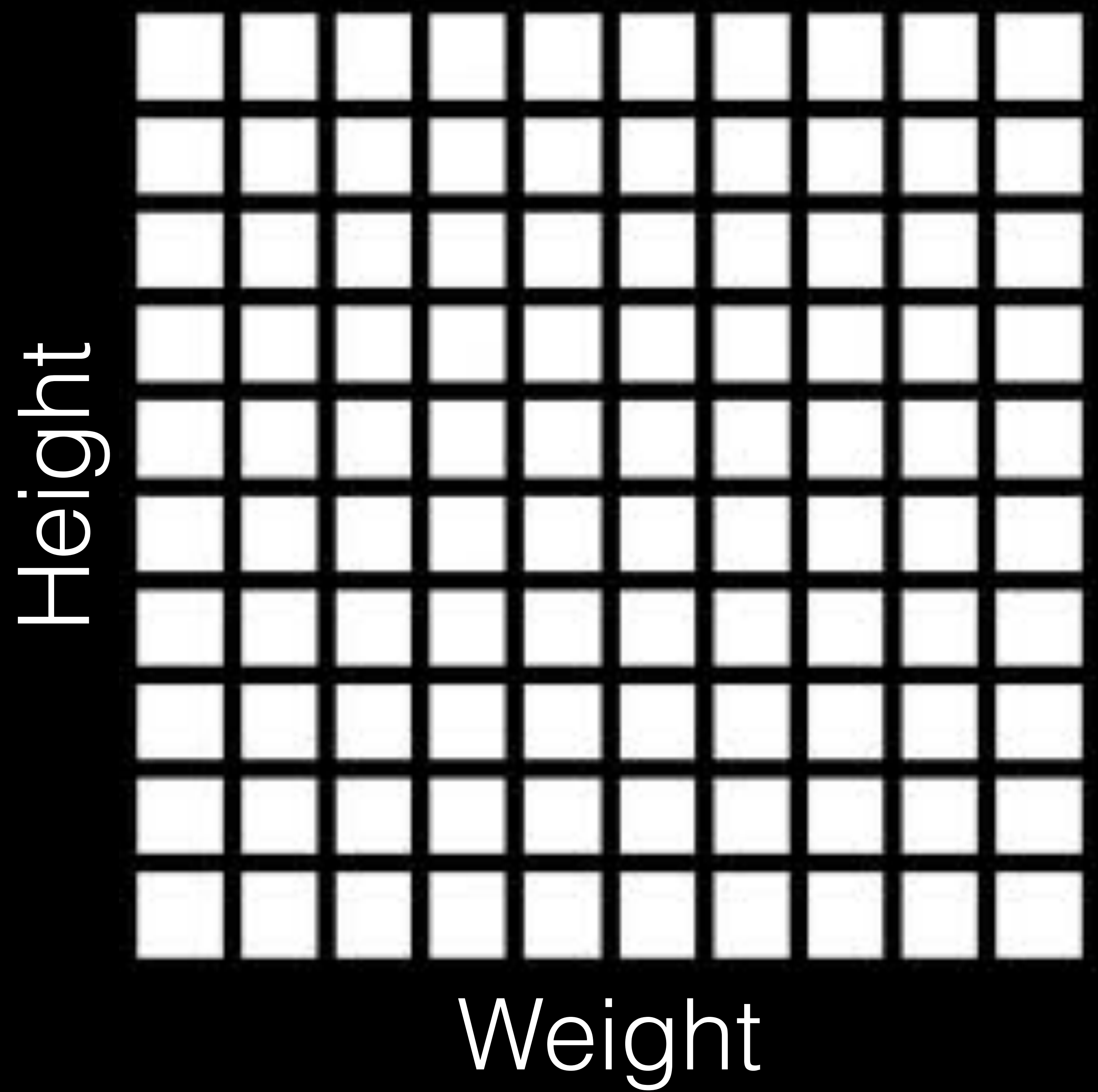
- a diverse set of high-performing agents (policies)



MAP-Elites

Mouret & Clune 2015, arXiv

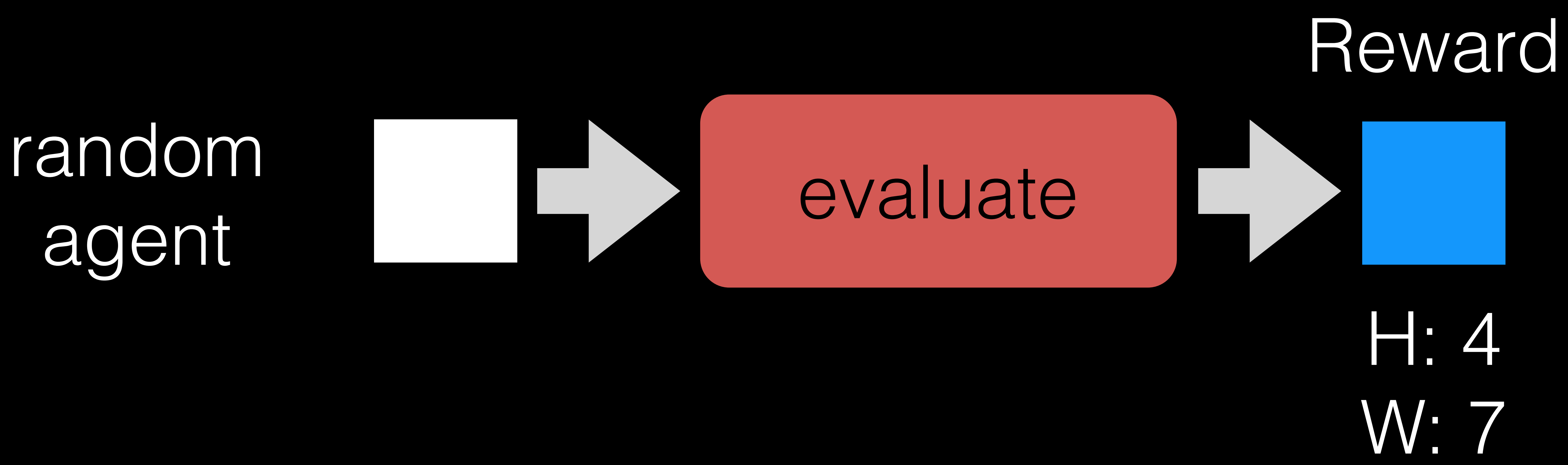
- Multi-dimensional archive of phenotypic elites
  - Choose dimensions of interest in behavior space
  - Discretize
  - Perturb, locate, replace if better, repeat

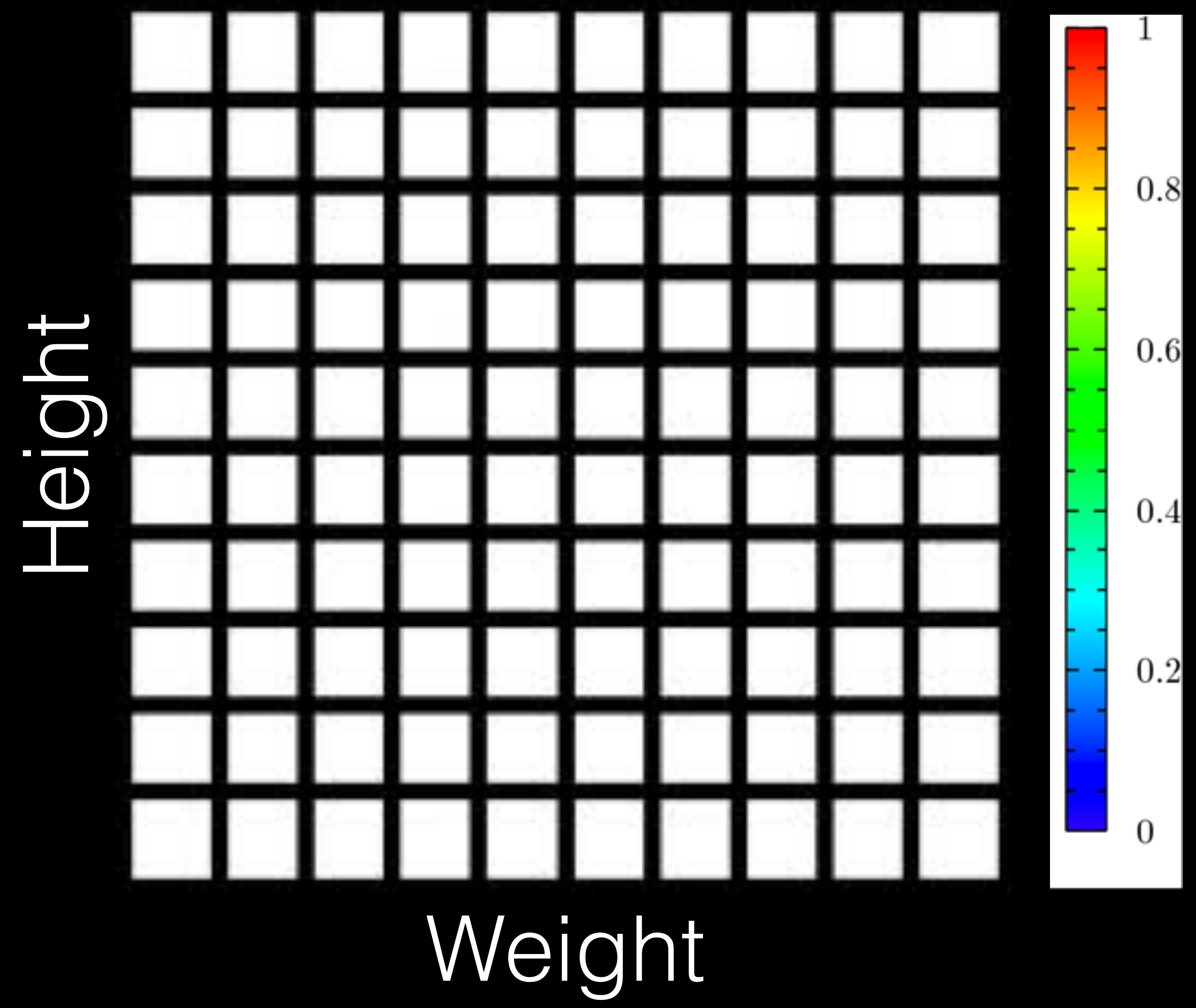


MAP-Elites

Mouret & Clune 2015, arXiv

- Multi-dimensional archive of phenotypic elites
  - Choose dimensions of interest in behavior space
  - Discretize
  - Perturb, locate, replace if better, repeat





MAP-Elites

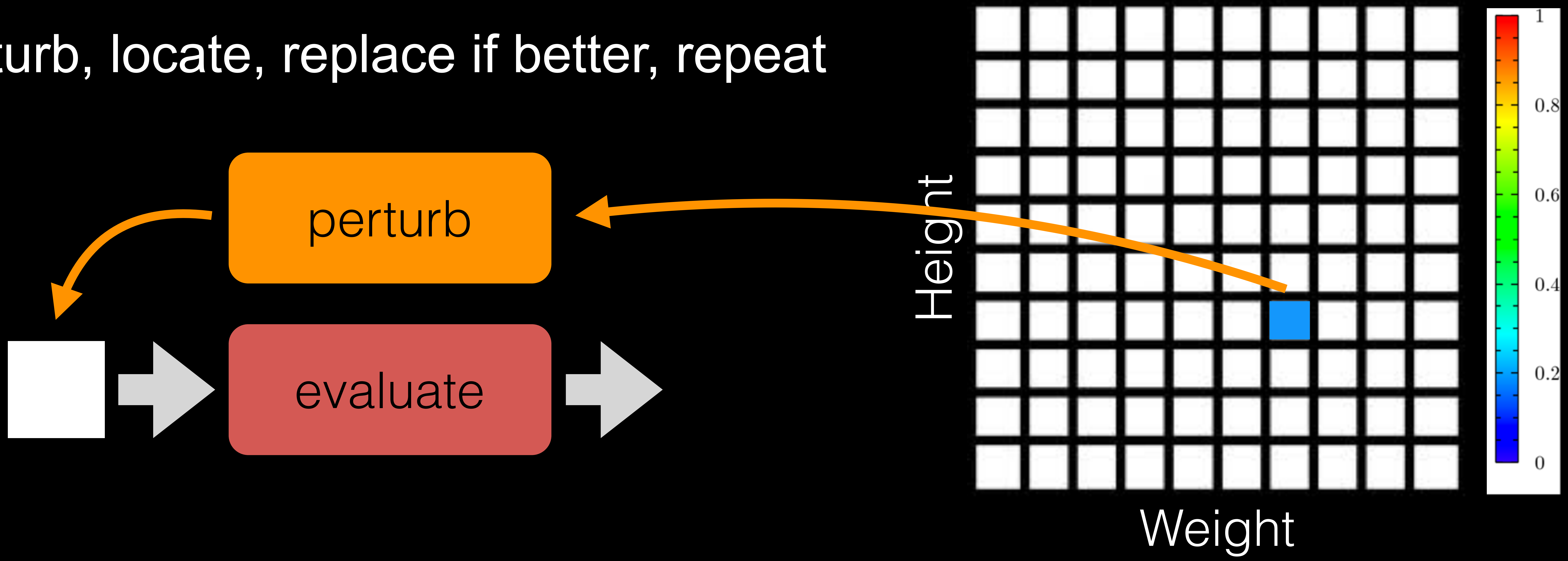
Mouret & Clune 2015, arXiv

- Multi-dimensional archive of phenotypic elites
  - Choose dimensions of interest in behavior space
  - Discretize
  - Perturb, locate, replace if better, repeat

MAP-Elites

Mouret & Clune 2015, arXiv

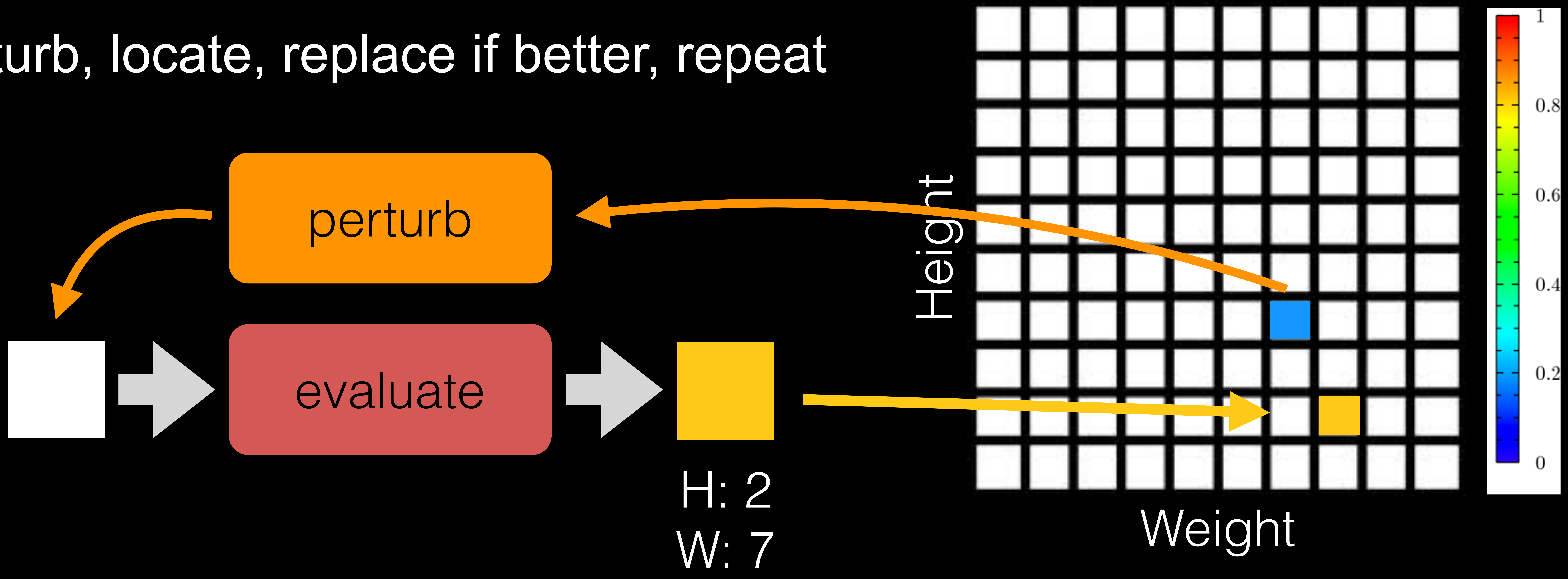
- Multi-dimensional archive of phenotypic elites
  - Choose dimensions of interest in behavior space
  - Discretize
  - Perturb, locate, replace if better, repeat



MAP-Elites

Mouret & Clune 2015, arXiv

- Multi-dimensional archive of phenotypic elites
  - Choose dimensions of interest in behavior space
  - Discretize
  - Perturb, locate, replace if better, repeat

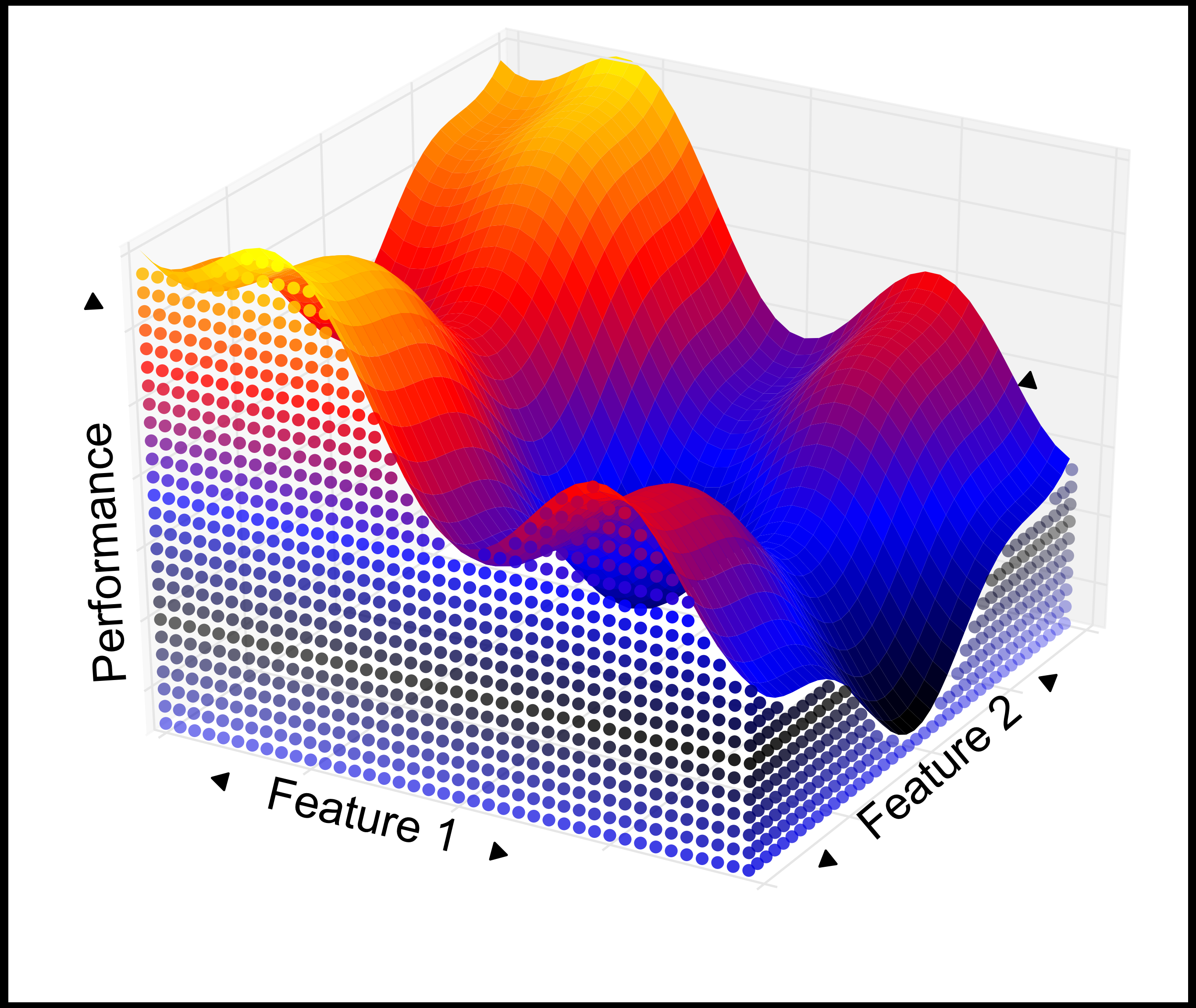


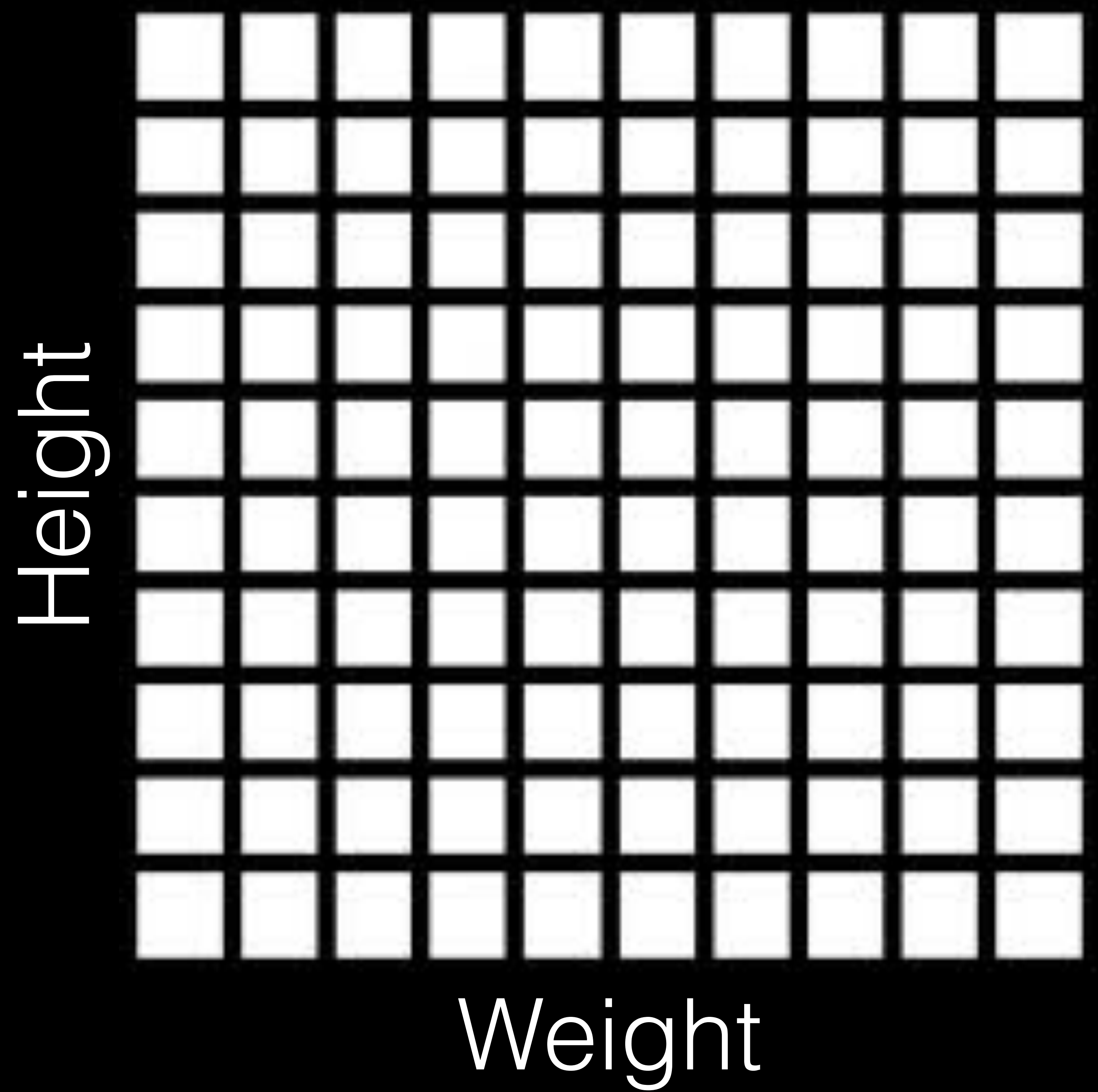
MAP-Elites

Mouret & Clune 2015, arXiv

- Multi-dimensional archive of phenotypic elites
  - Choose dimensions of interest in behavior space
  - Discretize
  - Perturb, locate, replace if better, repeat

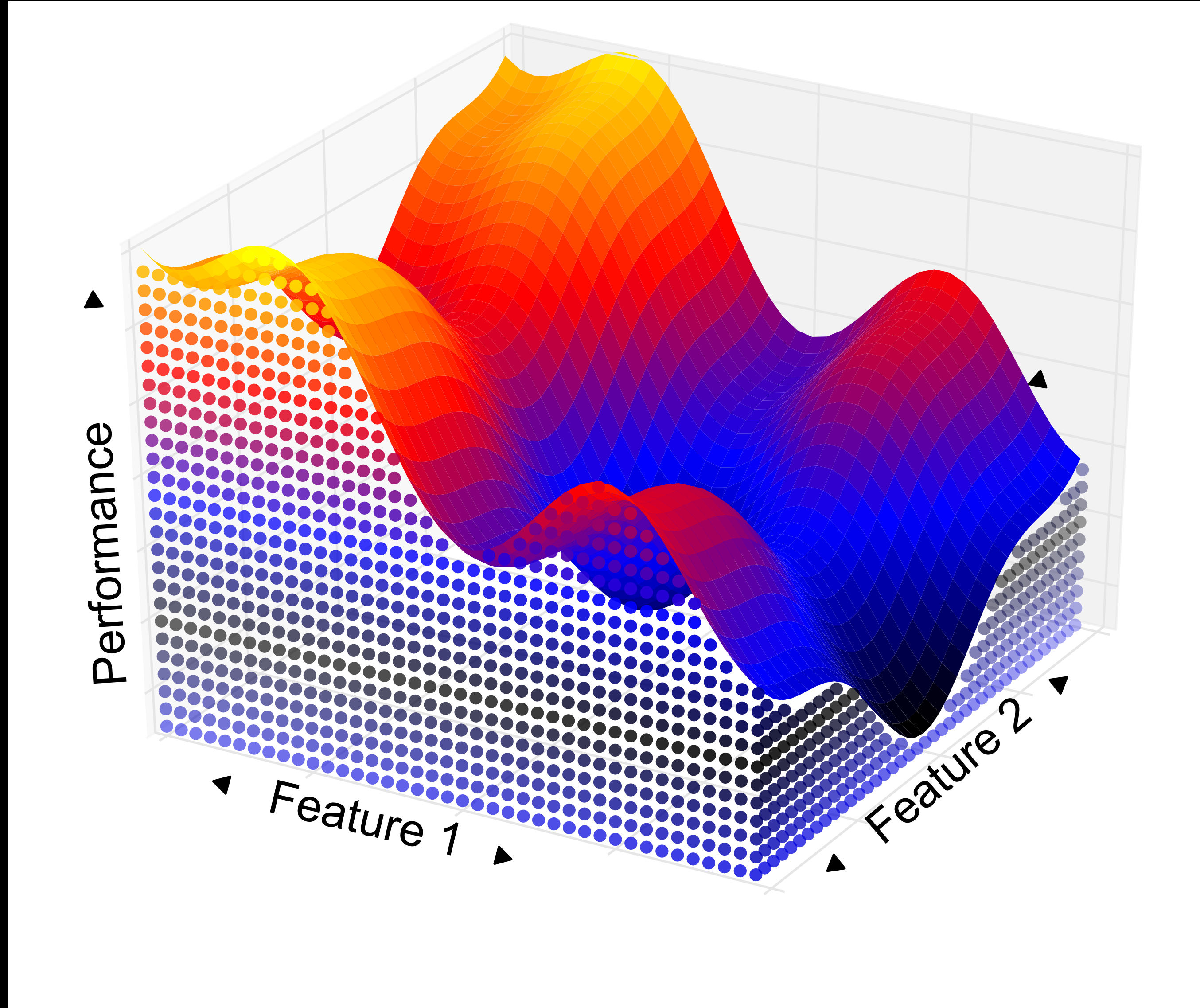
Set of diverse, high-quality solutions





Soft Robots Problem

Mouret & Clune 2015, arXiv



- Dimensions
  - number of voxels
  - % bone (dark blue)

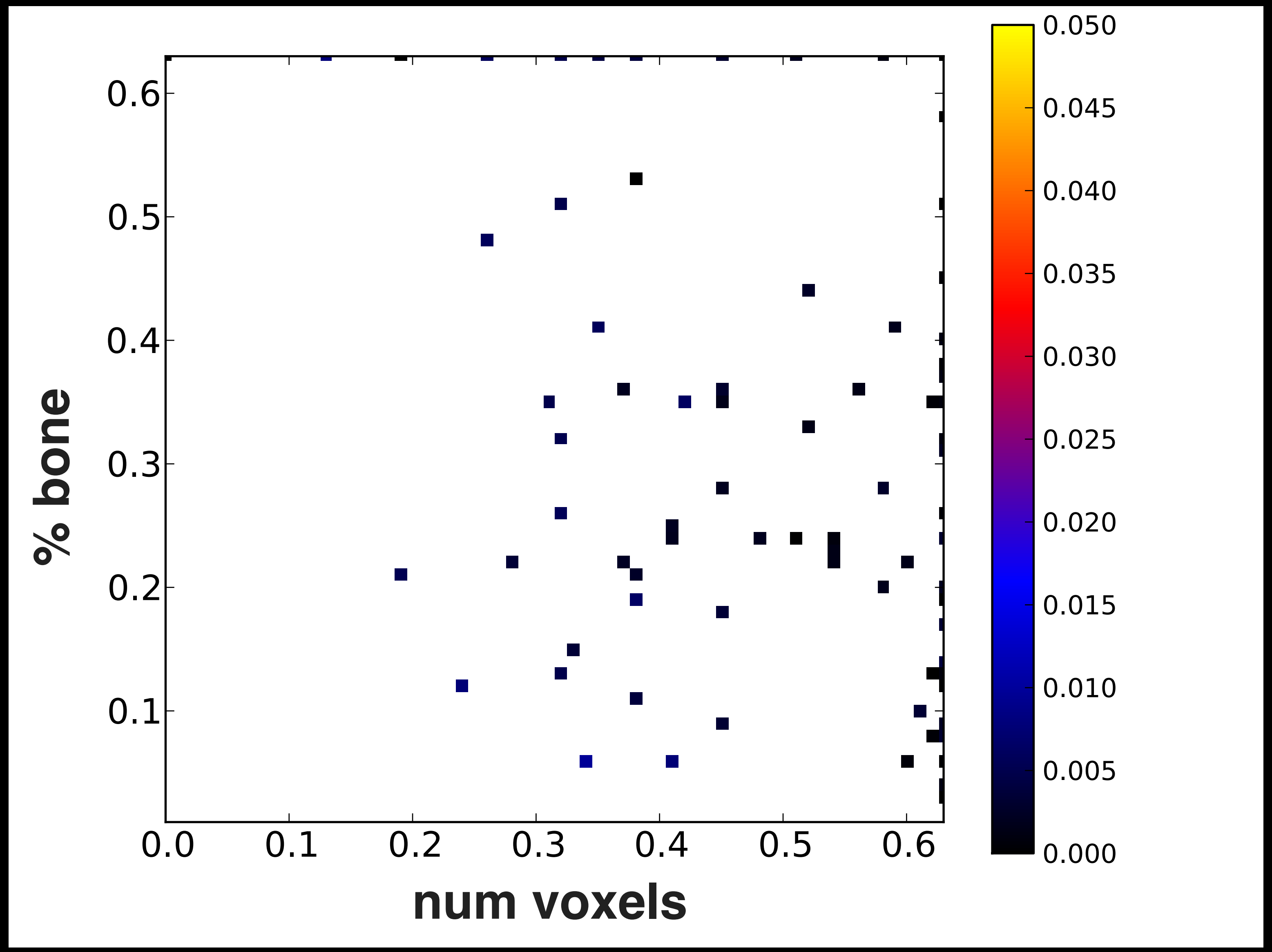
Cheney, MacCurdy, Clune, Lipson 2014

Qualitatively Different

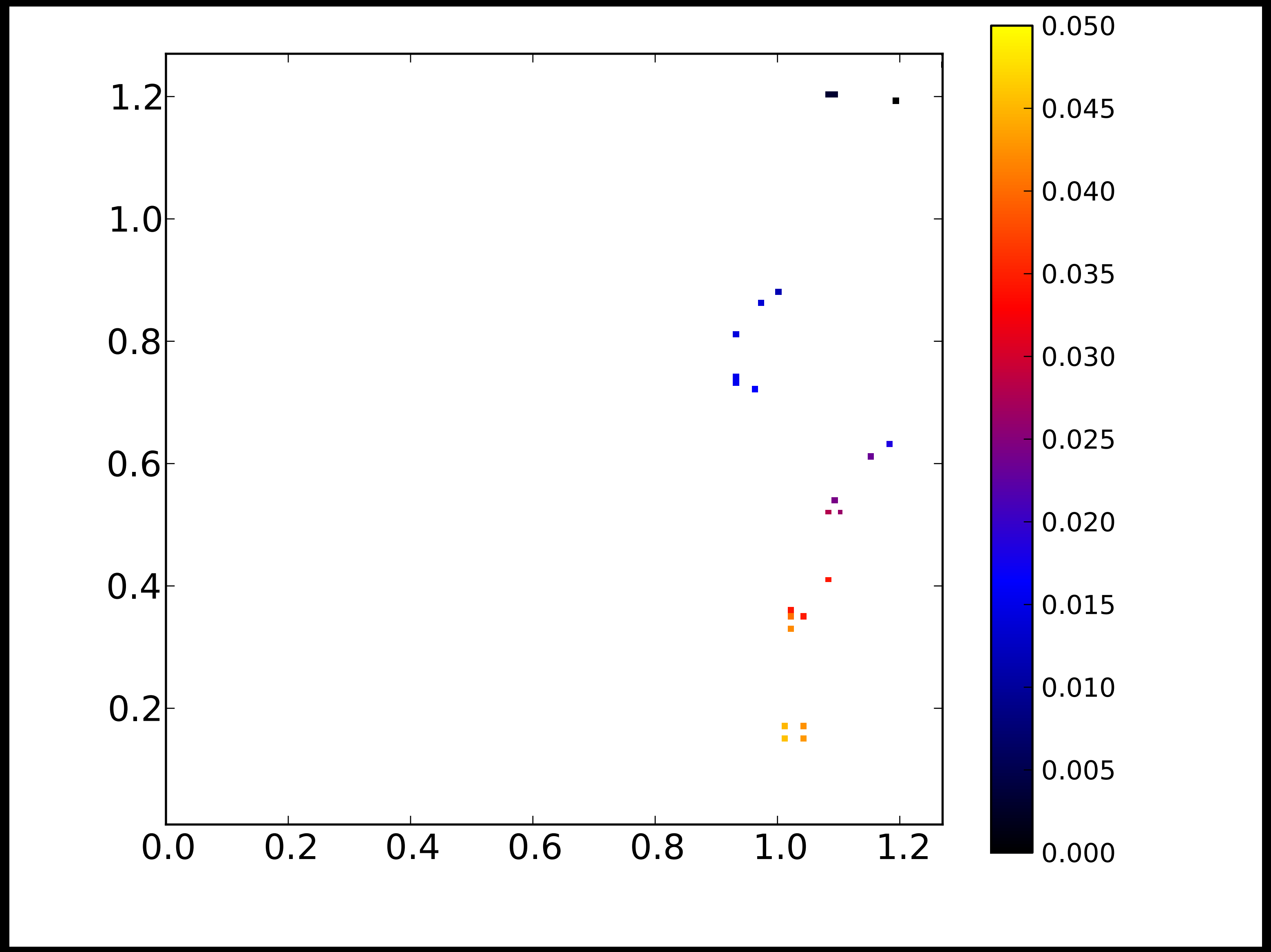
Mouret & Clune 2015, arXiv

soft robots problem

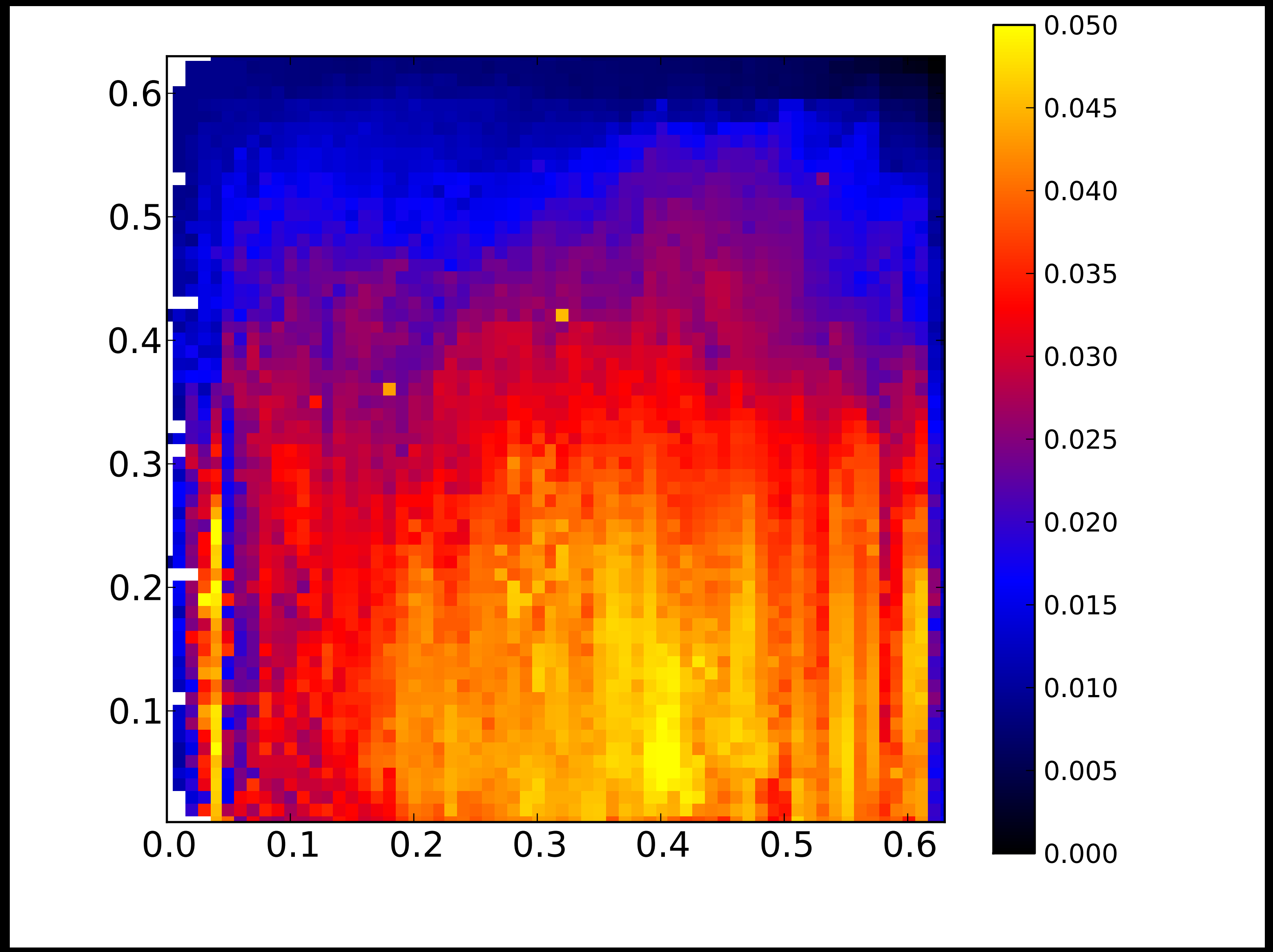
Classic Optimization



Classic + Diversity



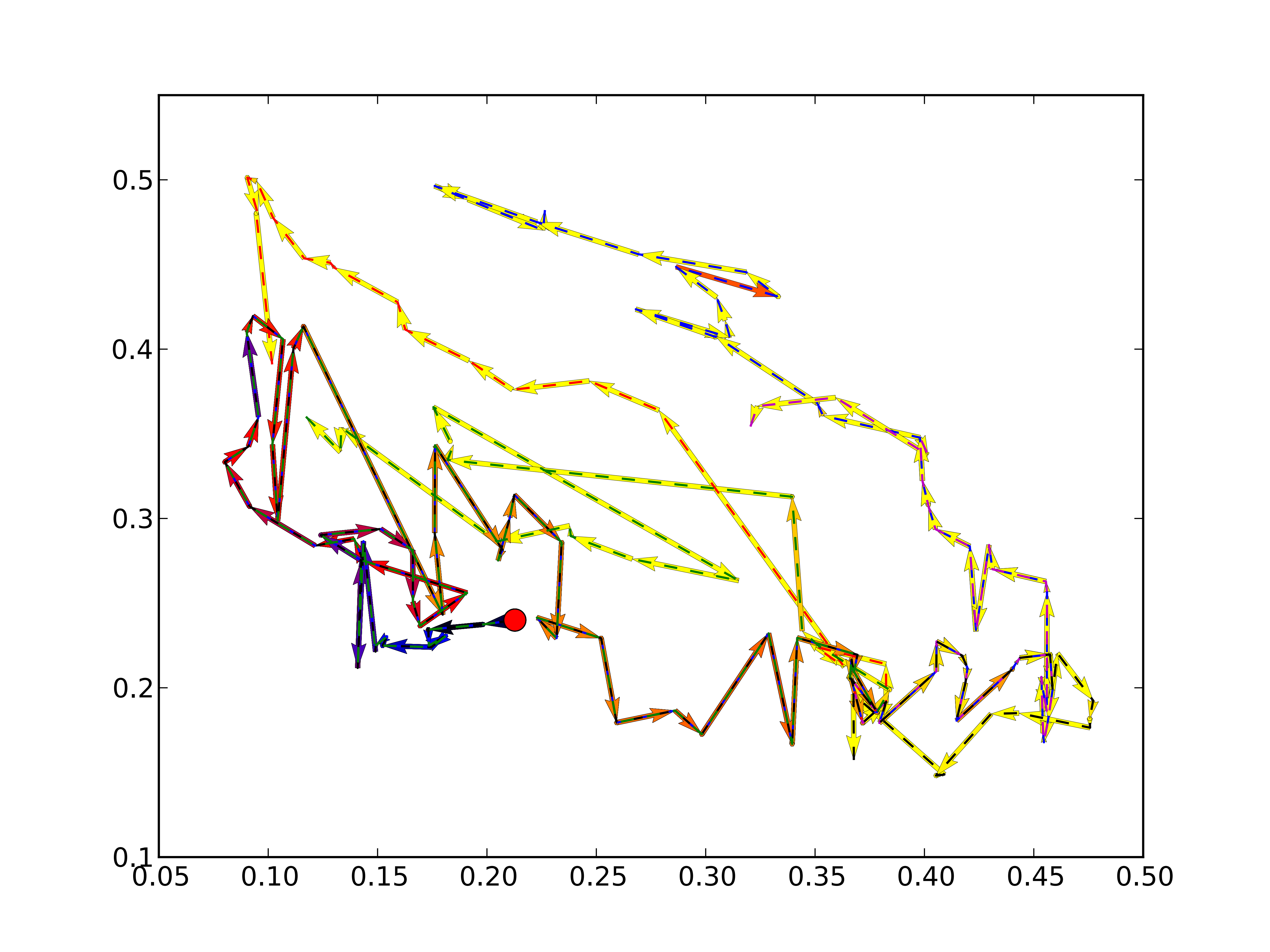
MAP-Elites



same # evals!

Often finds a better max than max-focused algorithms

Goal Switching is Critical



Lineage of a Final Solution

circle = iteration 0, color = reward

Robots that adapt like animals
2015

Produces rapid adaptation

Antoine Cully
UPMC Université
France

Jeff Clune
University of Wyoming
USA

Danesh Tarapore
UPMC Université
France

Jean-Baptiste Mouret
UPMC Université
France

Go-Explore

A new approach for hard-exploration problems

Adrien Ecoffet

Joost Huizinga

Joel Lehman

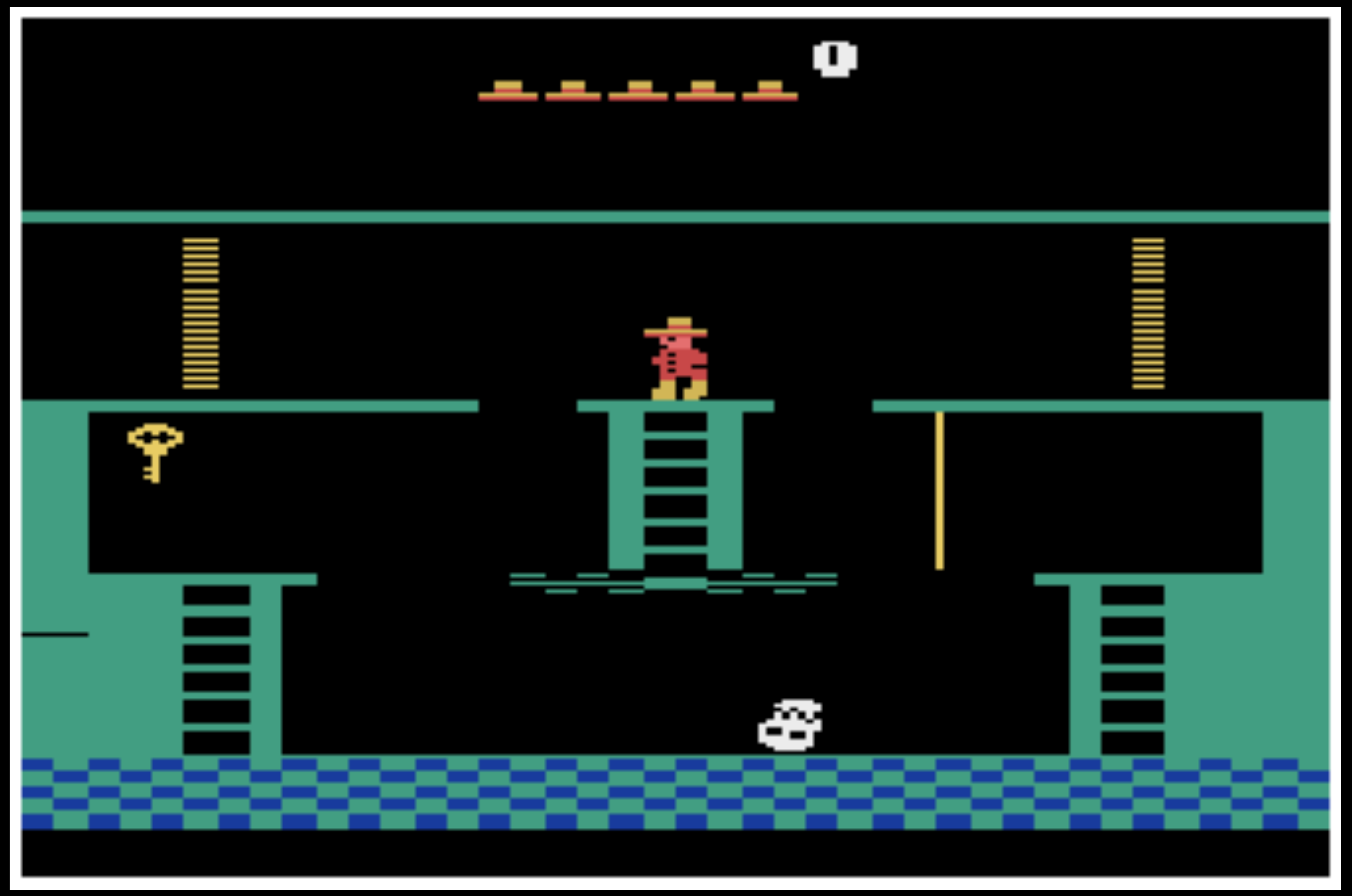
Ken Stanley

Jeff Clune

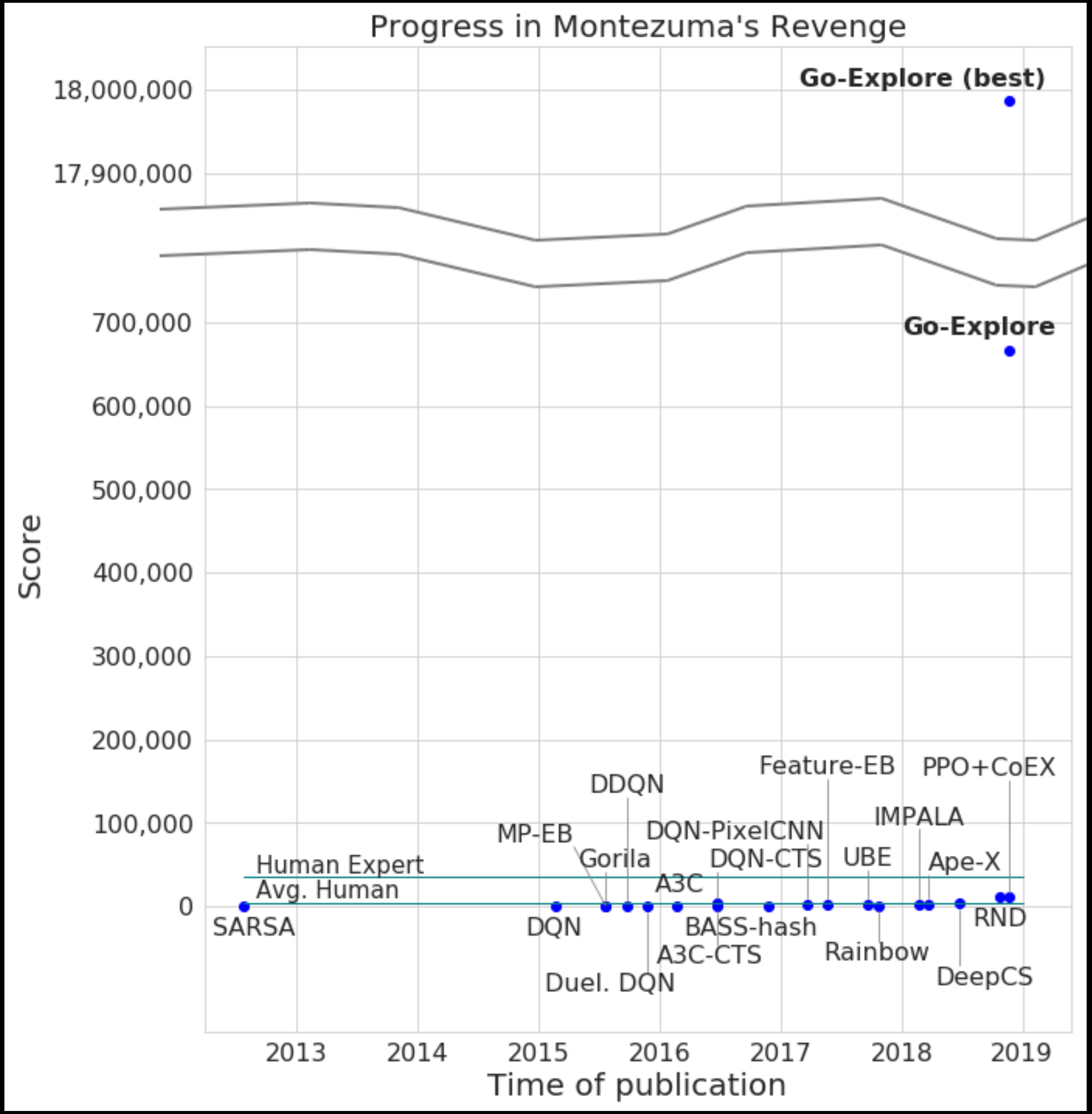
Nature 2021

Grand Challenge in RL: Effective Exploration

- Hard-exploration problems
  - Sparse-rewards
    - rare feedback
    - Montezuma's Revenge
      - Was a (mini) grand challenge



Results on Montezuma's Revenge



- Avg: 660,000+ points
- Best:
  - scores ~18 million
  - solved 1,141 levels
- Beats human world record
  - ~1.2 million

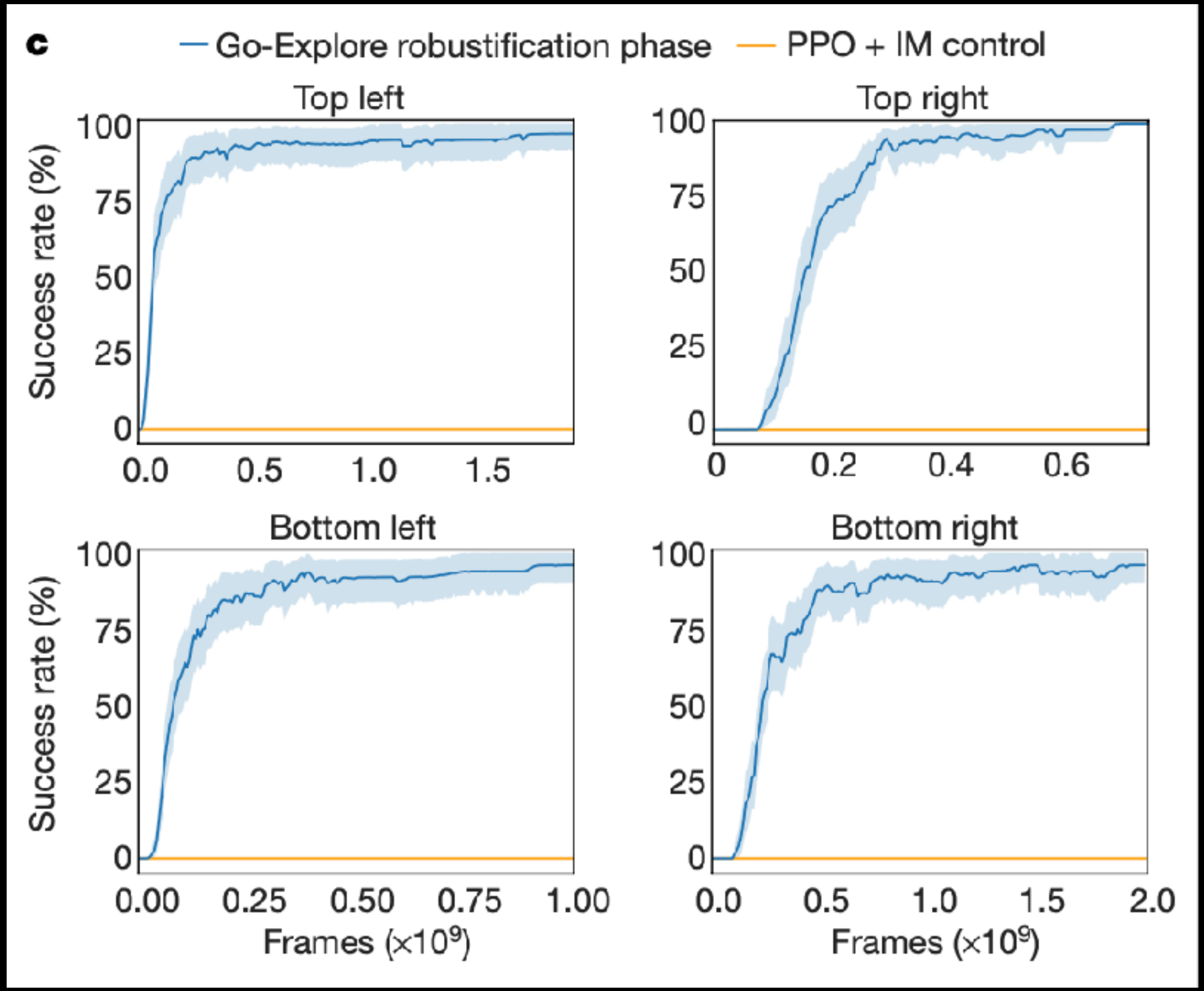
Solved all Atari games

solved = super-average-human (definition from DQN paper)

Simulated Hard-Exploration Robotics Task

Target: Top Left

Target: Top Right



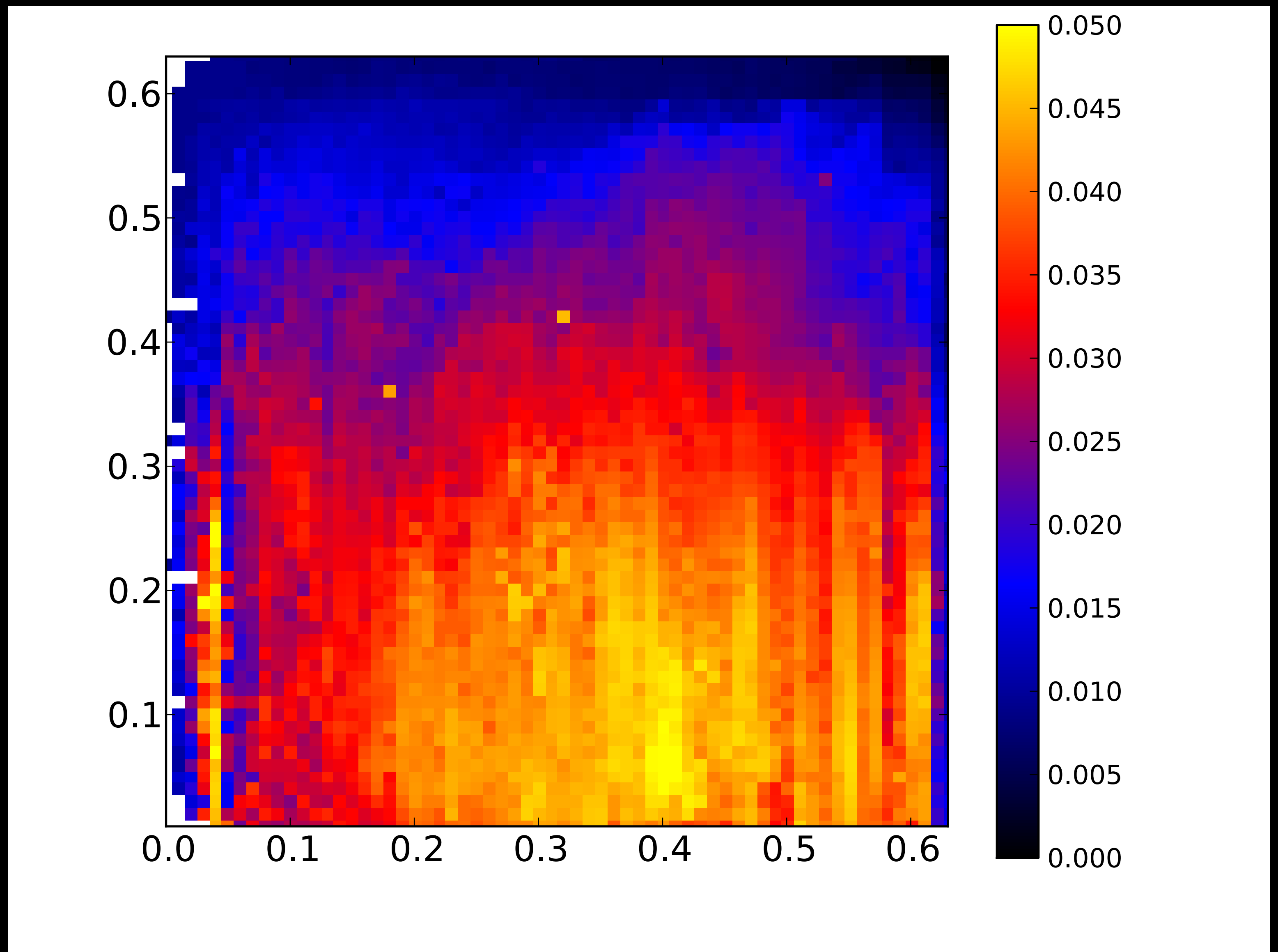
PPO+Intrinsic Motivation fails due to derailment

What's missing?

- QD algorithms
  - a diverse set of high-performing agents (policies)
  - exploration and goal-switching

- But they're stuck in a single environment

MAP-Elites



Traditional ML

- We pick challenges and solve them

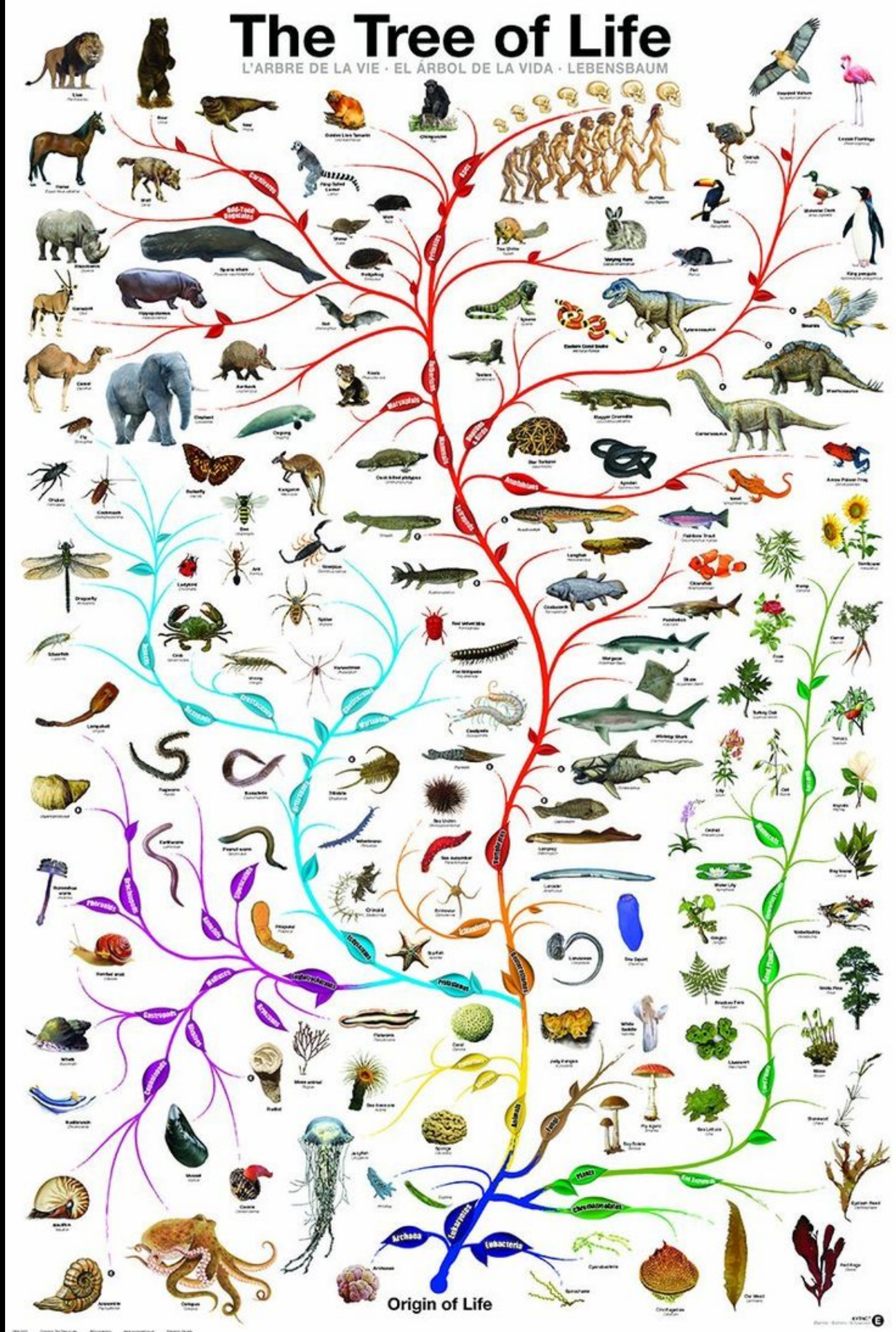



Starcraft

Dota 2

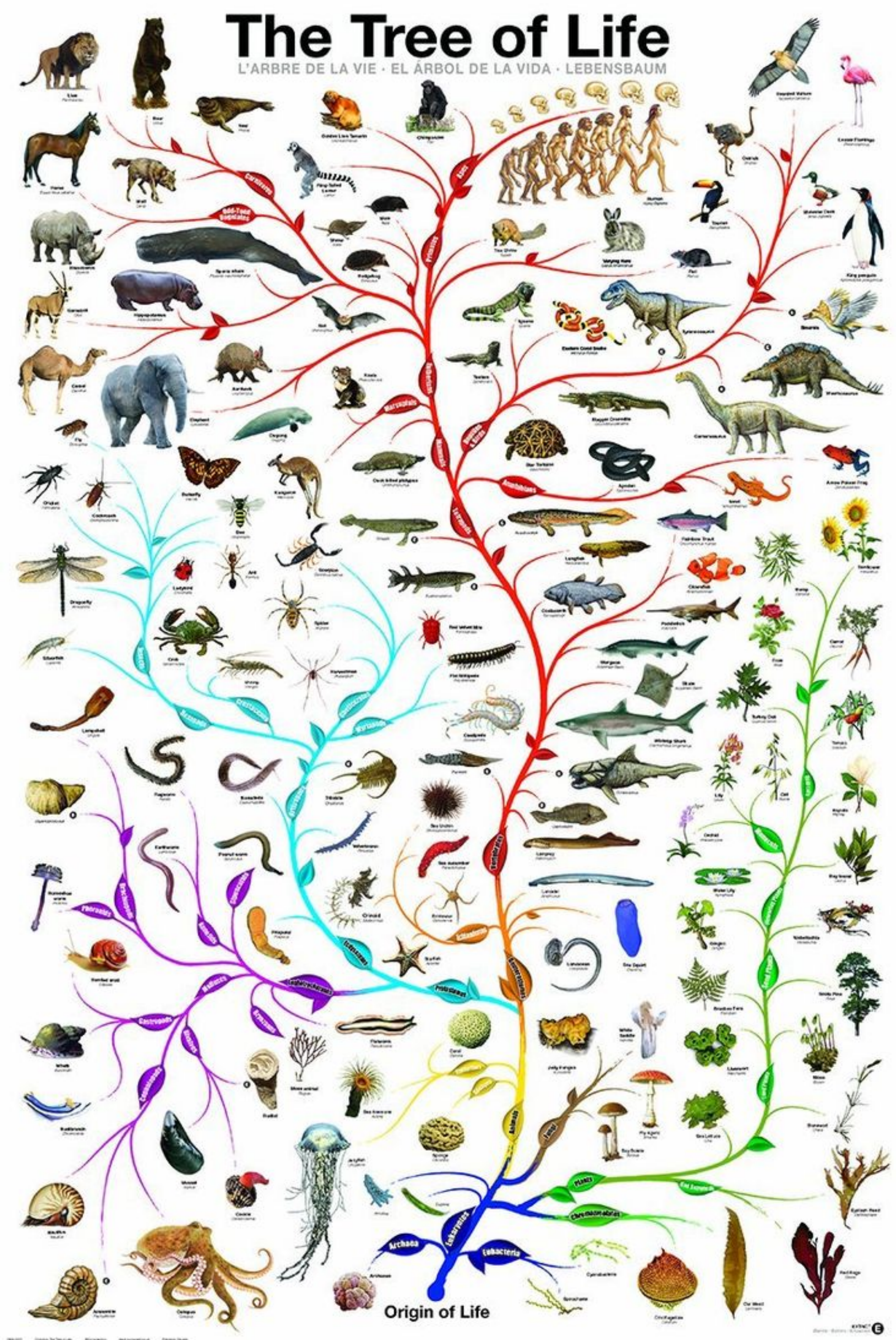
Open-Ended Algorithms

- Endlessly innovate
- Examples
  - Natural evolution
  - Human culture (science, technology, art)
- Can we make algorithms that do this?
  - Worth running for a billion years?



Intriguing Possibility

- Could the algorithm generate its own challenges and solve them?
  - niche/challenge/opportunity/problem:
    - tree leaves
  - solution:
    - giraffes
    - caterpillars



Paired Open-Ended Trailblazer (POET)

Rui Wang

Joel Lehman

Jeff Clune*

Ken Stanley*

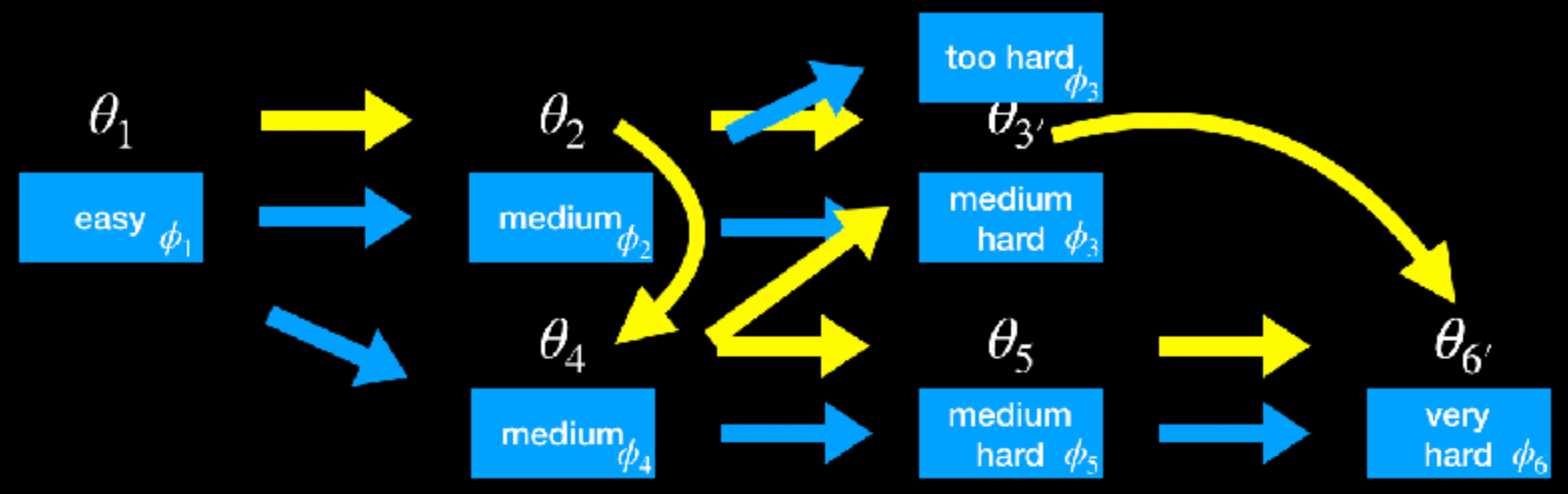
*Co-senior authors

GECCO 2019

Endlessly Generating Increasingly Complex and Diverse Learning Environments and their Solutions

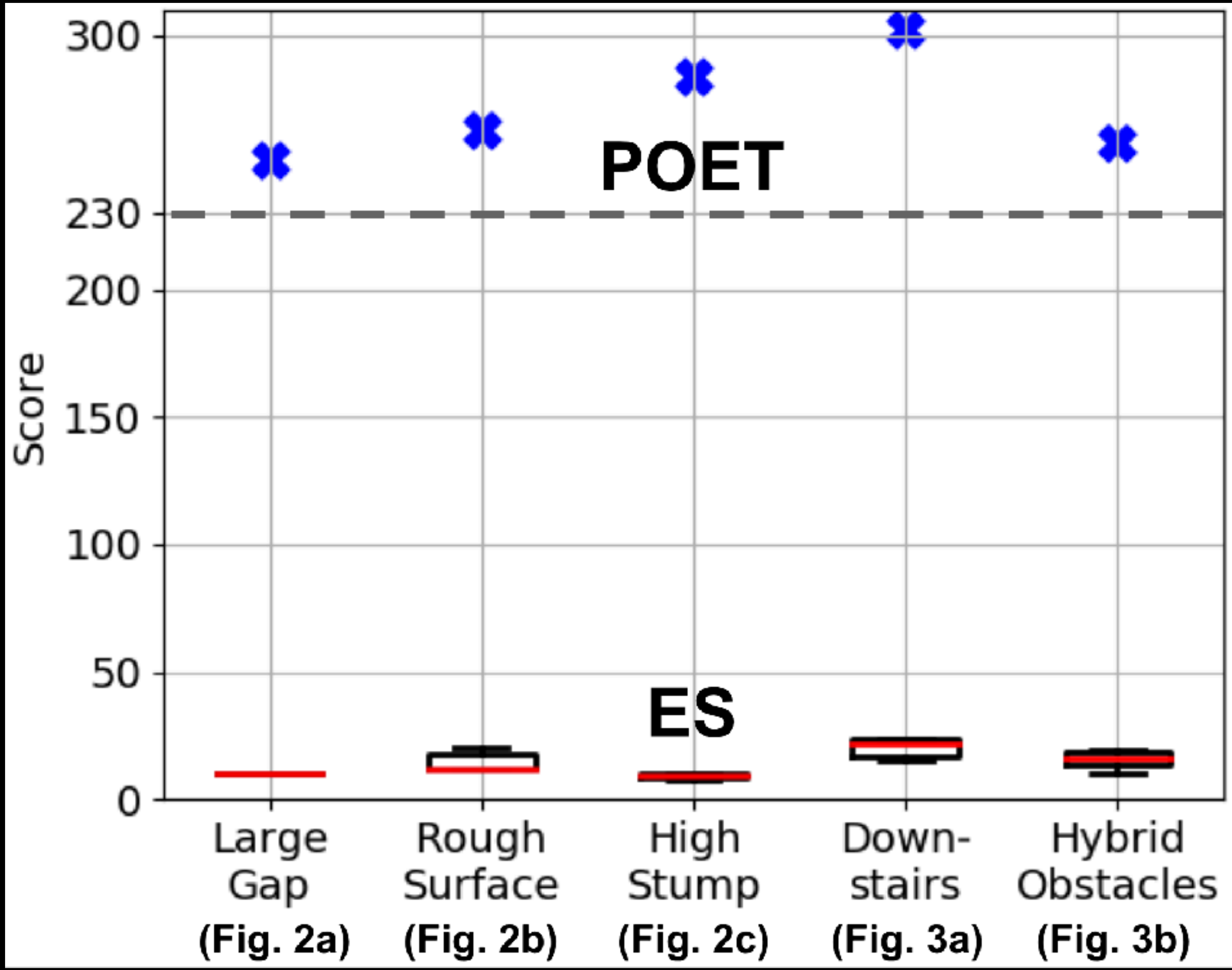
POET

- Encode environments
- Periodically
  - Generate new learning environments
    - add to population IF
      - not too easy, not too hard
      - novel
- Optimize agents to solve each one
  - allow goal-switching




It starts off easy...

Direct Optimization Fails



Hand-Designed Curricula Fail Too (see paper)

Producing all this diversity (and more) in a single run

Enhanced POET. R Wang, J Lehman, A Rawal, J Zhi, Y Li, J Clune, K Stanley (2020) ICML

Enhanced POET. R Wang, J Lehman, A Rawal, J Zhi, Y Li, J Clune, K Stanley (2020) ICML

POET-Inspired Work Outdoors

Joonho et al. Science Robotics. 2020

POET: Conclusions

- Shows benefits of
  - QD
  - generating challenges while solving them

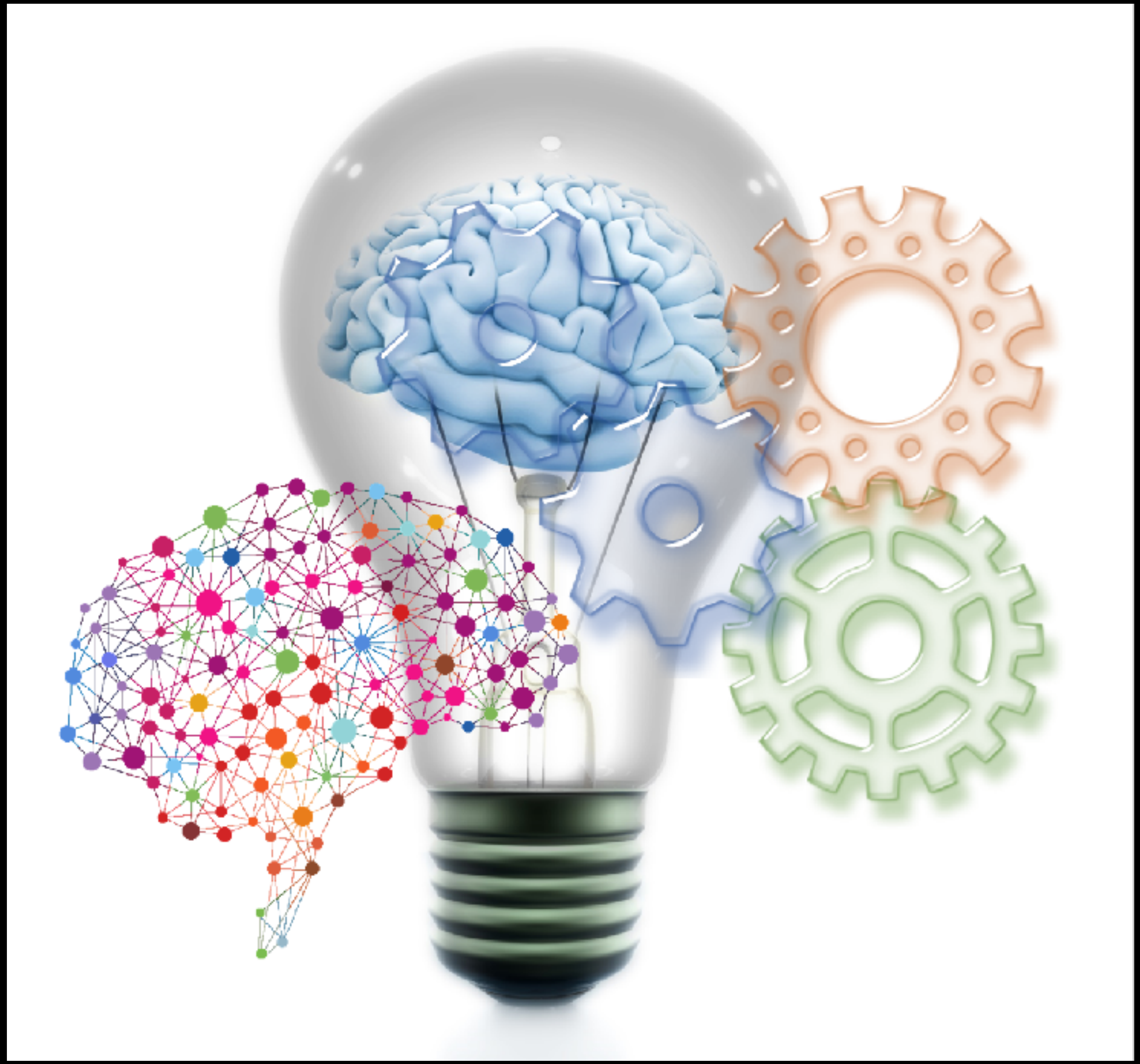
POET

- Pro: auto-generates tasks
- Con: small, hand-chosen distribution of environments/tasks

Artificial General Intelligence (AGI)

or Human-level AI, if you prefer

- If we should make it
  - We have a long way to go
  - How will we get there?



Clear Machine Learning Trend:
Hand-designed pipelines are ultimately outperformed by learned solutions

hand designed $\longrightarrow$ learned

- Features
- Architectures
- Hyperparameters & data augmentation
- RL algorithms

suggests alternate path

AI-Generating Algorithms

Clune 2019

- Lots of pre-training
  - produces a sample-efficient, intelligent agent
- Existence proof
  - Earth

AI-Generating Algorithms

Clune 2019

Three Pillars

1. Meta-learn architectures
2. Meta-learn learning algorithms
3. Generate effective learning environments

Handcrafting each is slow, limited by our intelligence/time

Better to learn them. Let ML+compute do the heavy lifting

AI-Generating Algorithms

Clune 2019

Three Pillars

1. Meta-learn architectures
2. Meta-learn learning algorithms
3. Generate effective learning environments
    - Quality Diversity algorithms
    - Open-Ended algorithms
    - e.g. POET

Today

- Motivate QD, open-ended and AI-generating algorithms
- Discuss new possibilities unlocked by “foundation models”

Today

- Motivate QD, open-ended and AI-generating algorithms
- Discuss new possibilities unlocked by “foundation models”

OMNI

Open-endedness via Models of human Notions of Interestingness



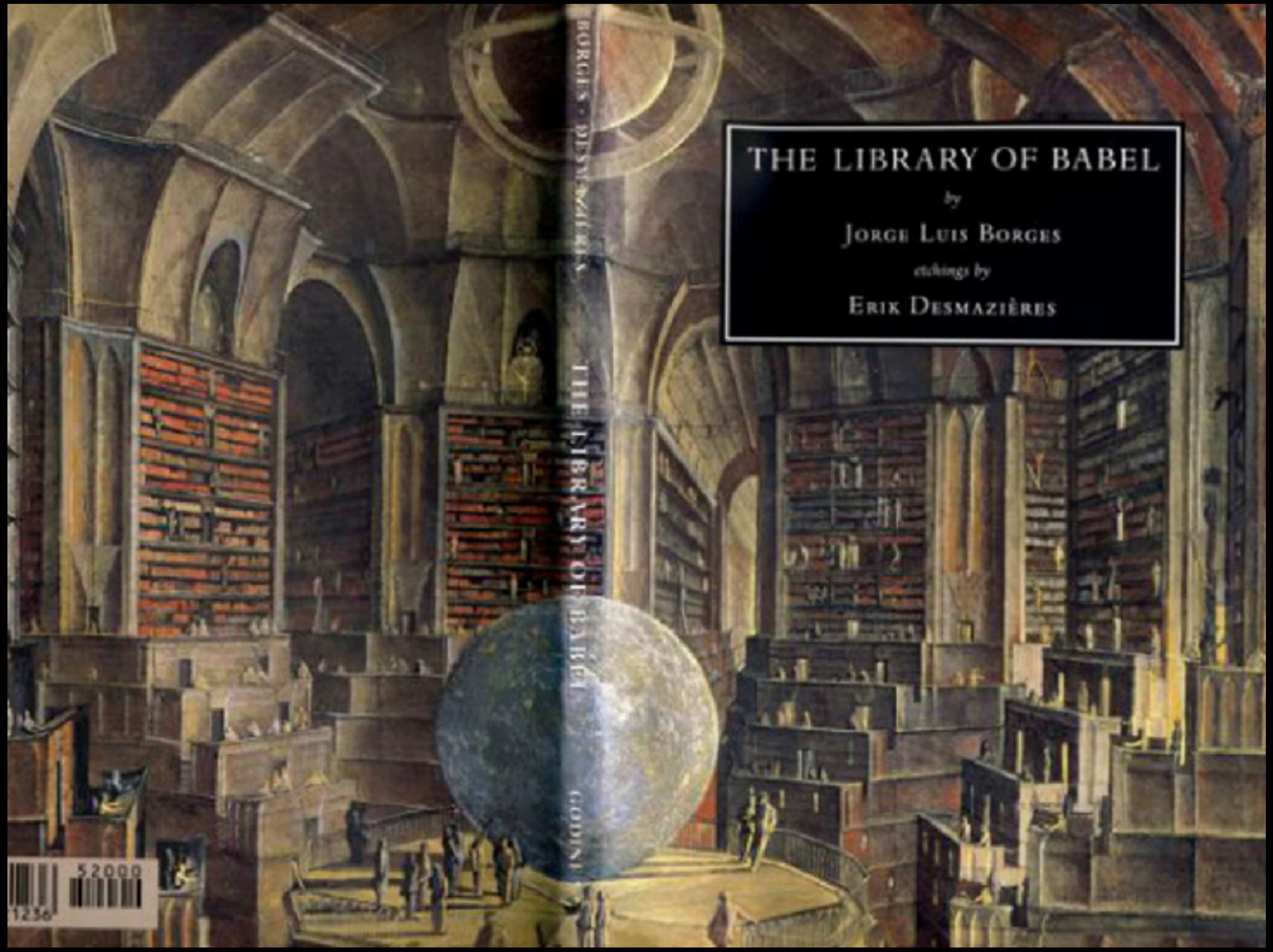

Jenny Zhang* Joel Lehman Ken Stanley Jeff Clune*

UBC, Vector Institute

To Produce Open-Ended Algorithms / AI-GAs

- We need a vast task/environment space

All Tasks Described by Words



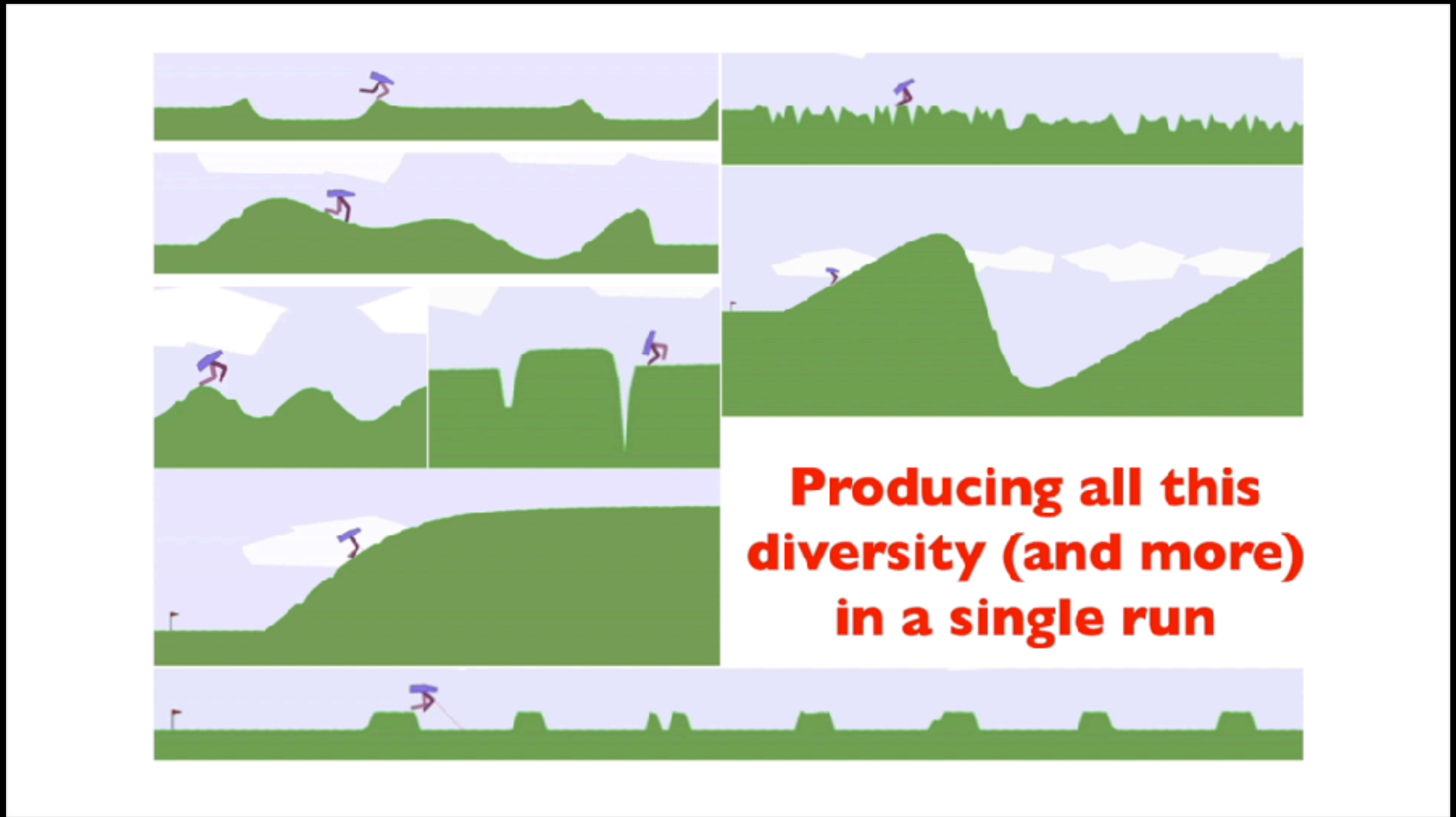
Many tasks: impossible, nonsense, too hard, or too easy

Tasks Must be New & Interesting

How do we find them?

First Challenge: Tasks Must Have Learning Progress

- Filters out impossible, too hard, and too easy tasks
  - e.g. POET



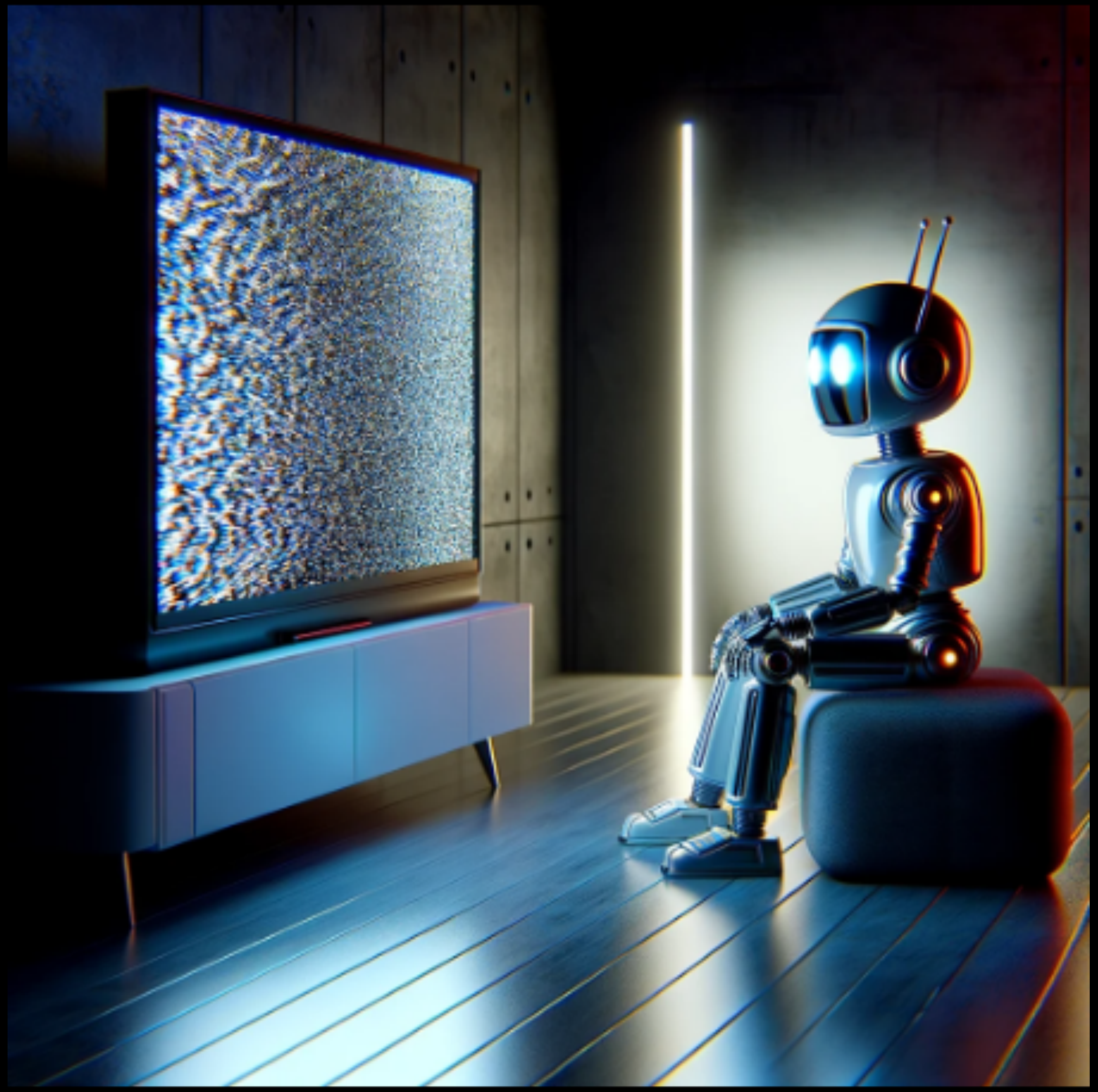
POET Wang et al. 2019; Enhanced POET Wang et al. 2020

Problem

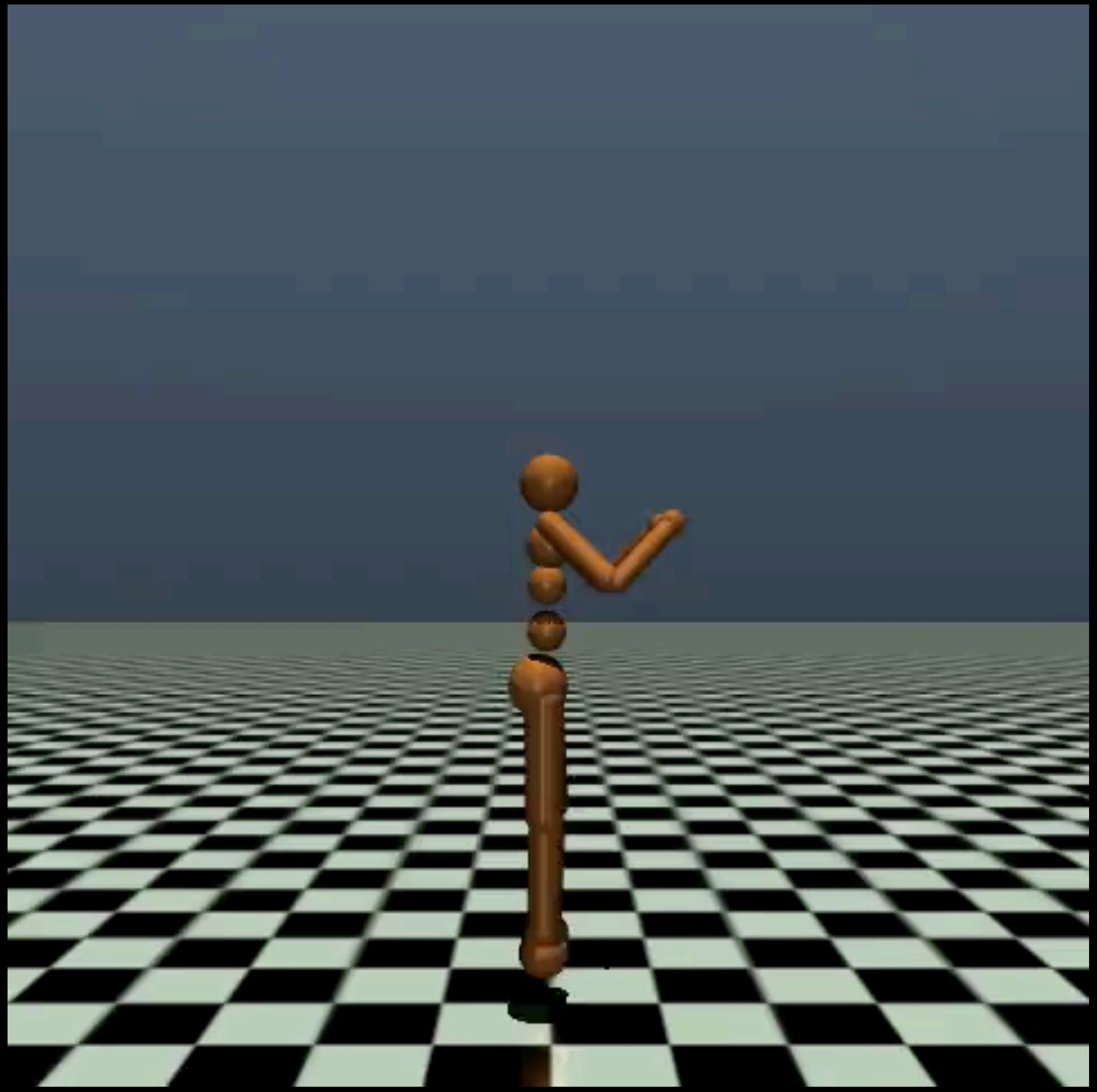
- Easy to have learning progress without being interestingly new
  - 3.14
  - 3.14159265358979323
  - 3.141592653589793238462643383279502884197
- Needed: way to automatically determine what is learnable and interestingly new

Problem: Quantifying Interestingness

- Hand-coded new/interesting/novelty metrics produce pathologies
- “When a measure becomes a target, it ceases to be a good measure” – Goodhart’s law



Find new states = noisy TV problem



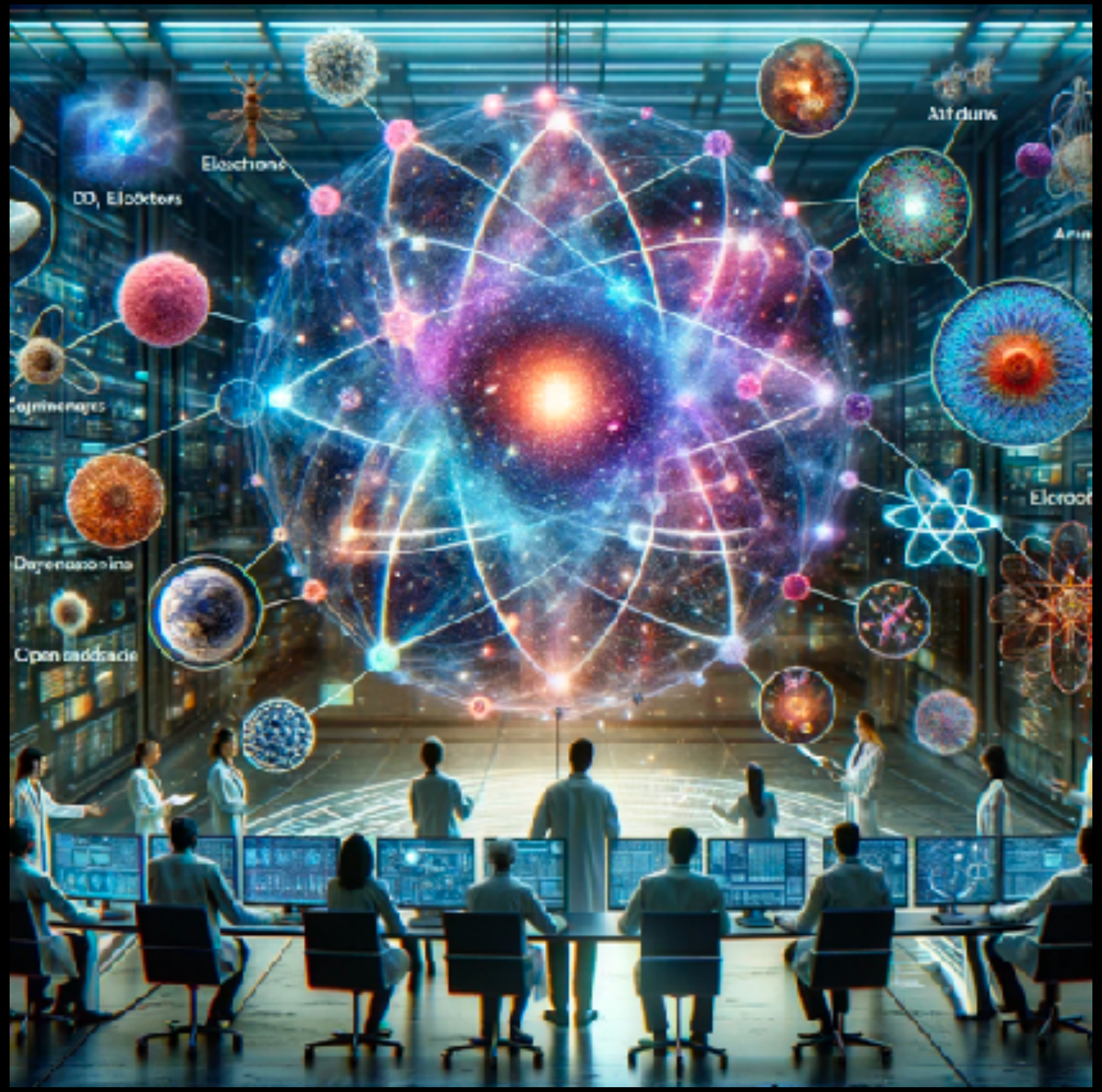
Do new movements = endless twitching

Problem: Quantifying Interestingness

- Past approaches (hand coding) do not work

We cannot define it, but we know it when we see it
- Justice Stewart, paraphrased

- Automatically identifying interestingness is the most important problem in open-endedness



An answer!
Foundation Models

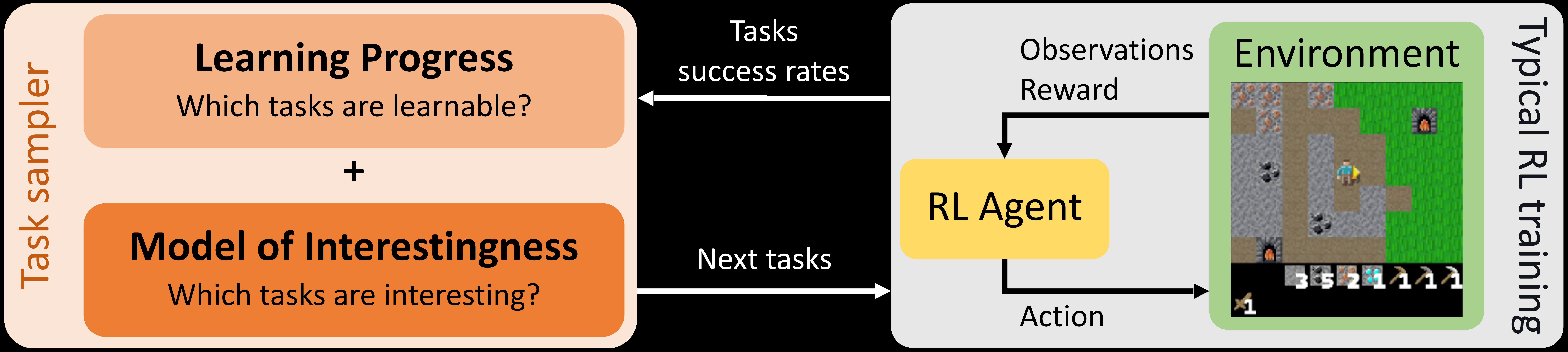
- FMs already know what is interesting
  - They have read the entire Internet
  - Including LOTs of humans talking about what is interesting vs. not
  - Have common sense
- Insight: we can just ask them!

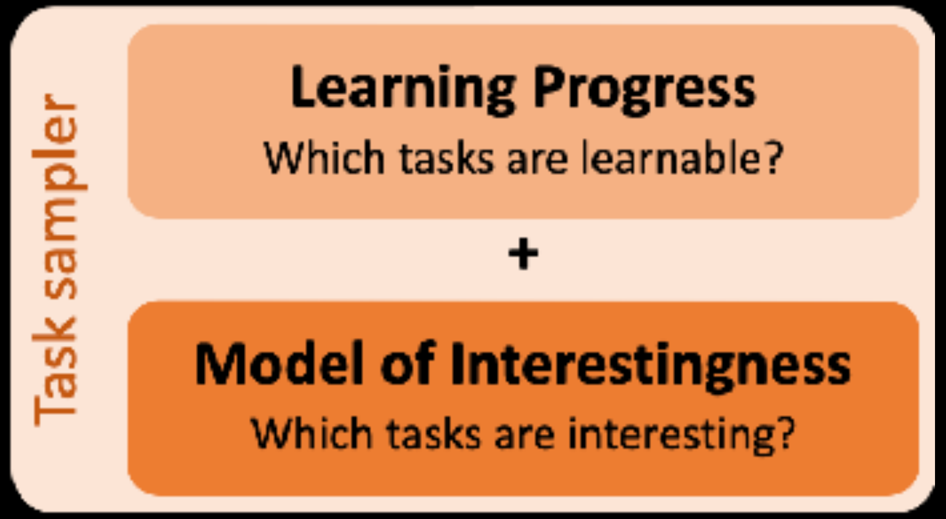
OMNI

Open-endedness via Models of human Notions of Interestingness

Methods

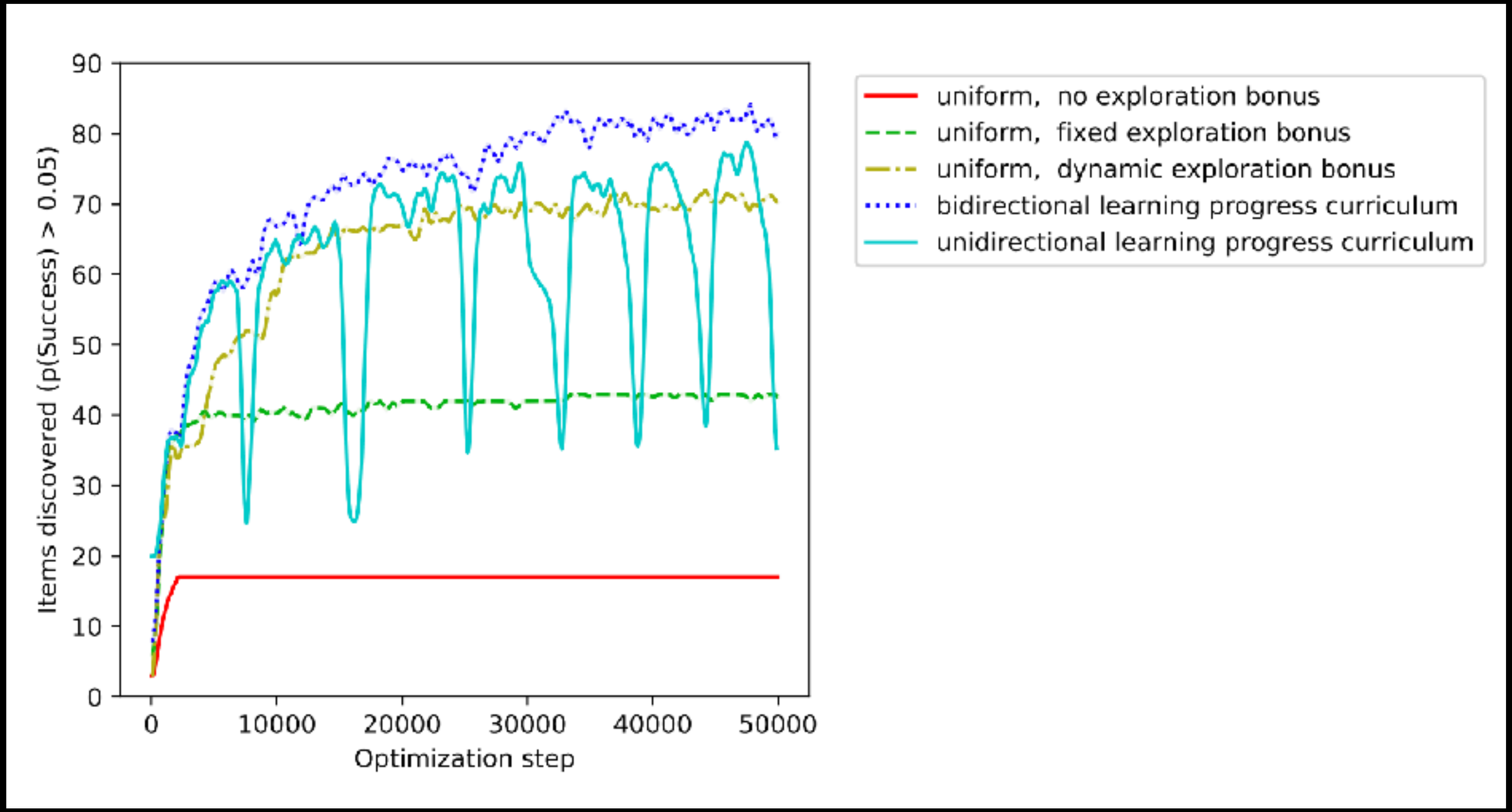
- Reinforcement Learning, Task-conditioned
- Task sampler
  - High learning progress
  - Are interesting



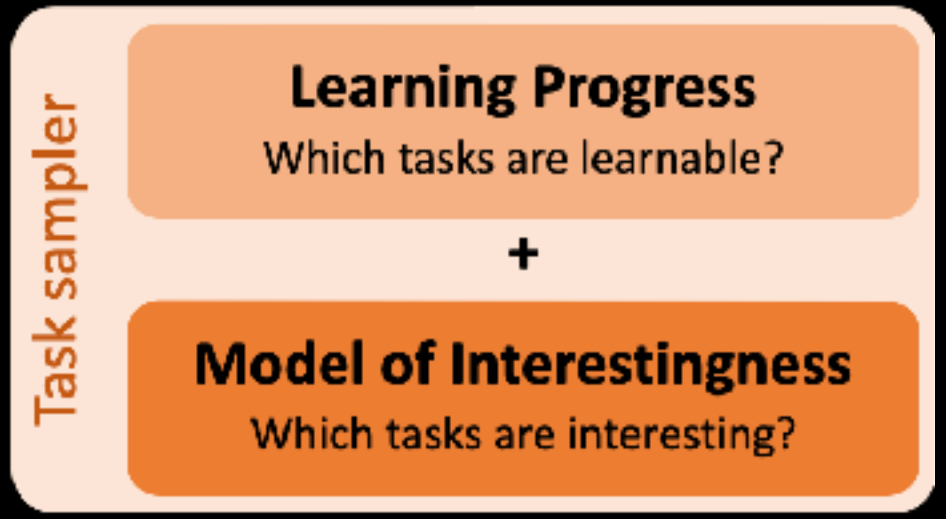


Learning Progress

- Sample tasks with higher learning progress
  - Using Kanitscheider et al. 2021



[1] Kanitscheider, Ingmar, et al. "Multi-task curriculum learning in a complex, visual, hard-exploration domain: Minecraft." *arXiv preprint arXiv:2106.14876* (2021).



Model of Interestingness (MoI)

- GPT

Goals

The ultimate goal that [the agent] would like your help with is to learn as many interestingly different skills as possible ...

Environment description

The grid world has objects in six distinct colors ... Each task is a sequence of instructions ...

Tasks the agent currently does well:

“open some door”, “go to some object”

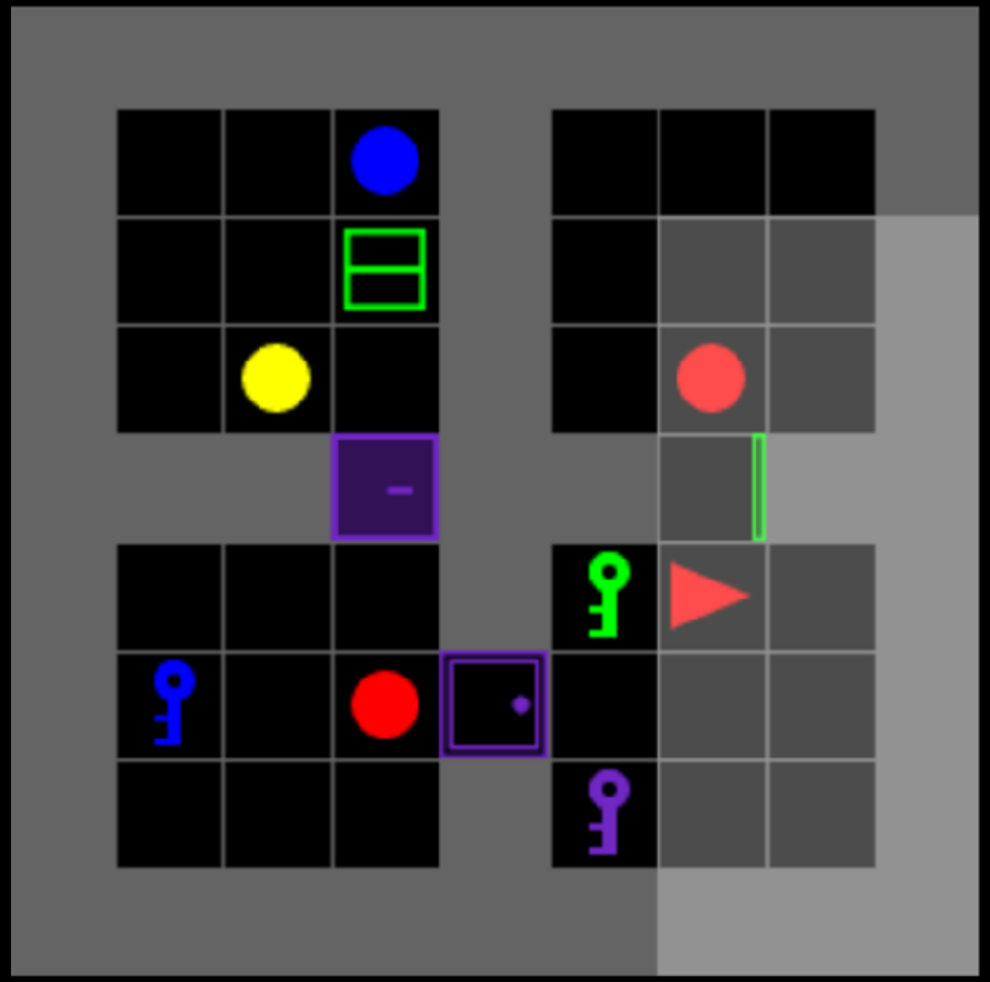
Predict which of these tasks are interesting:

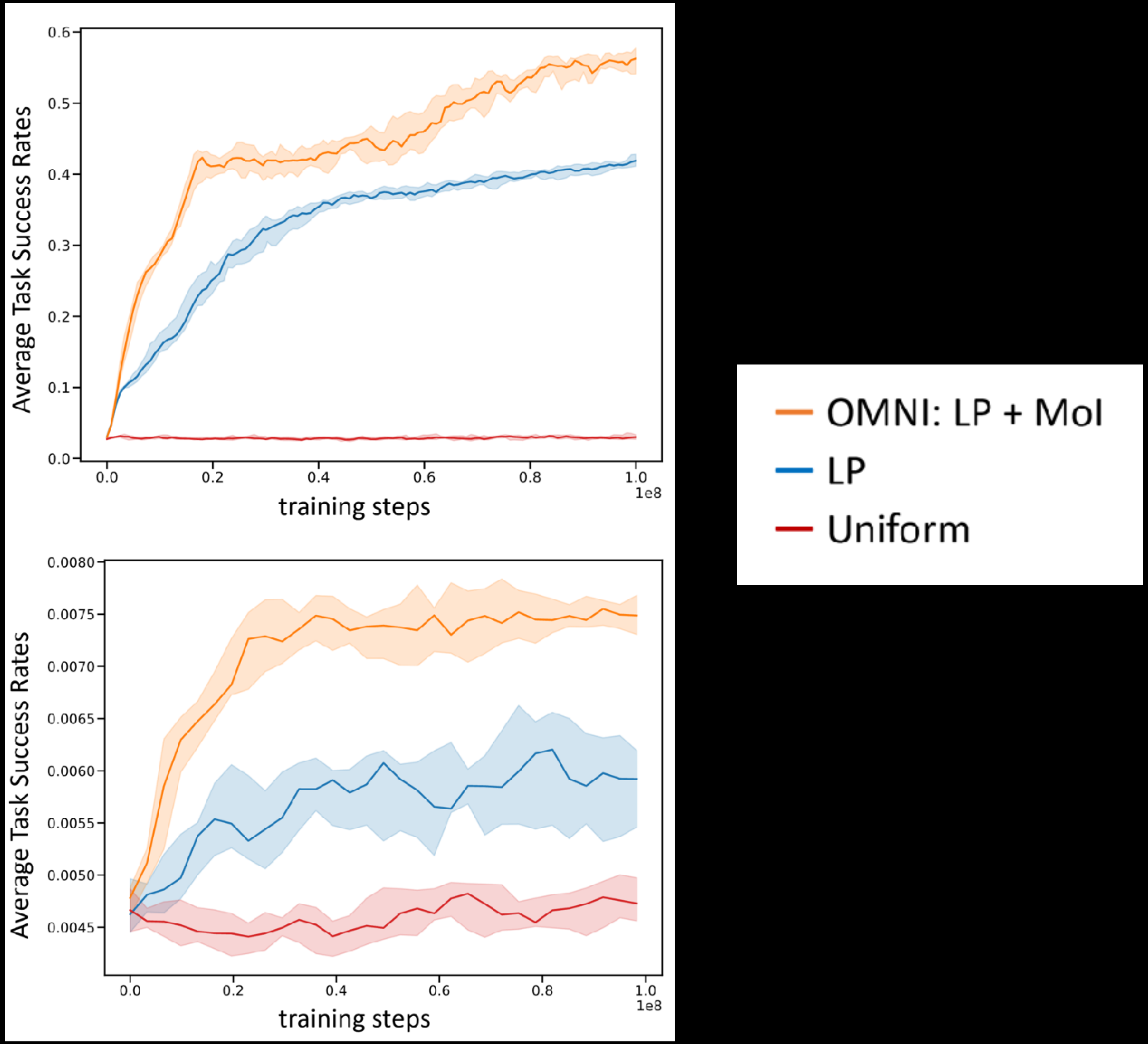
“pick up some object”, “open some door, then go to some object”

Results

Crafter

BabyAI





Results on a Human-Predefined Set of Interesting Tasks

Infinite Task Space

- More complex domain
  - Kitchen env (Thor)
- Any task specifiable in natural language
- Challenge: requires a "universal reward function"
  - says if any task has been completed

Solution: ask GPT for reward function code

- Prompt
  - Generate the next new, interesting, learnable tasks
  - and code to check if each is done

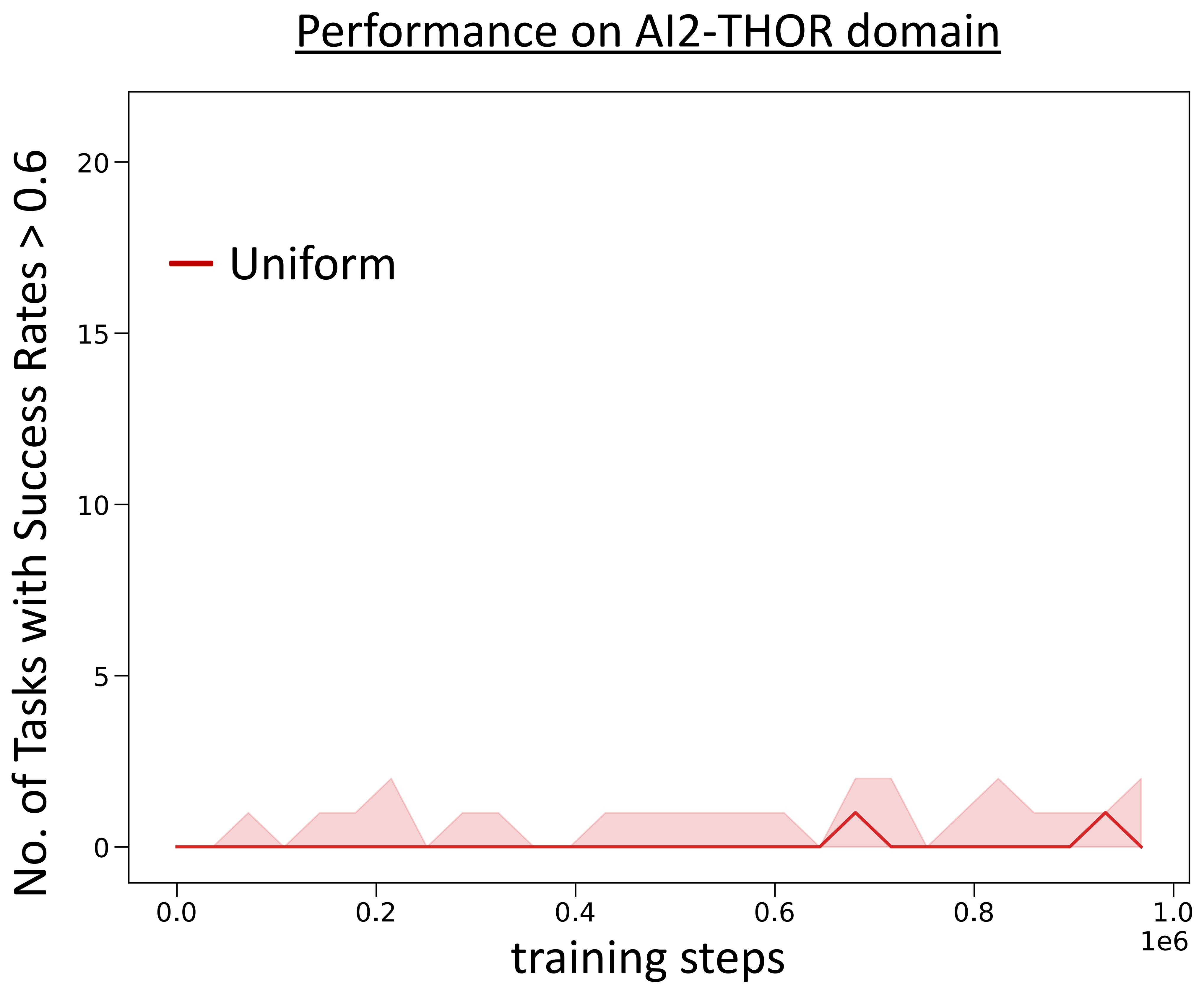
Next tasks

```
1. Pick up the knife.
2. …
```

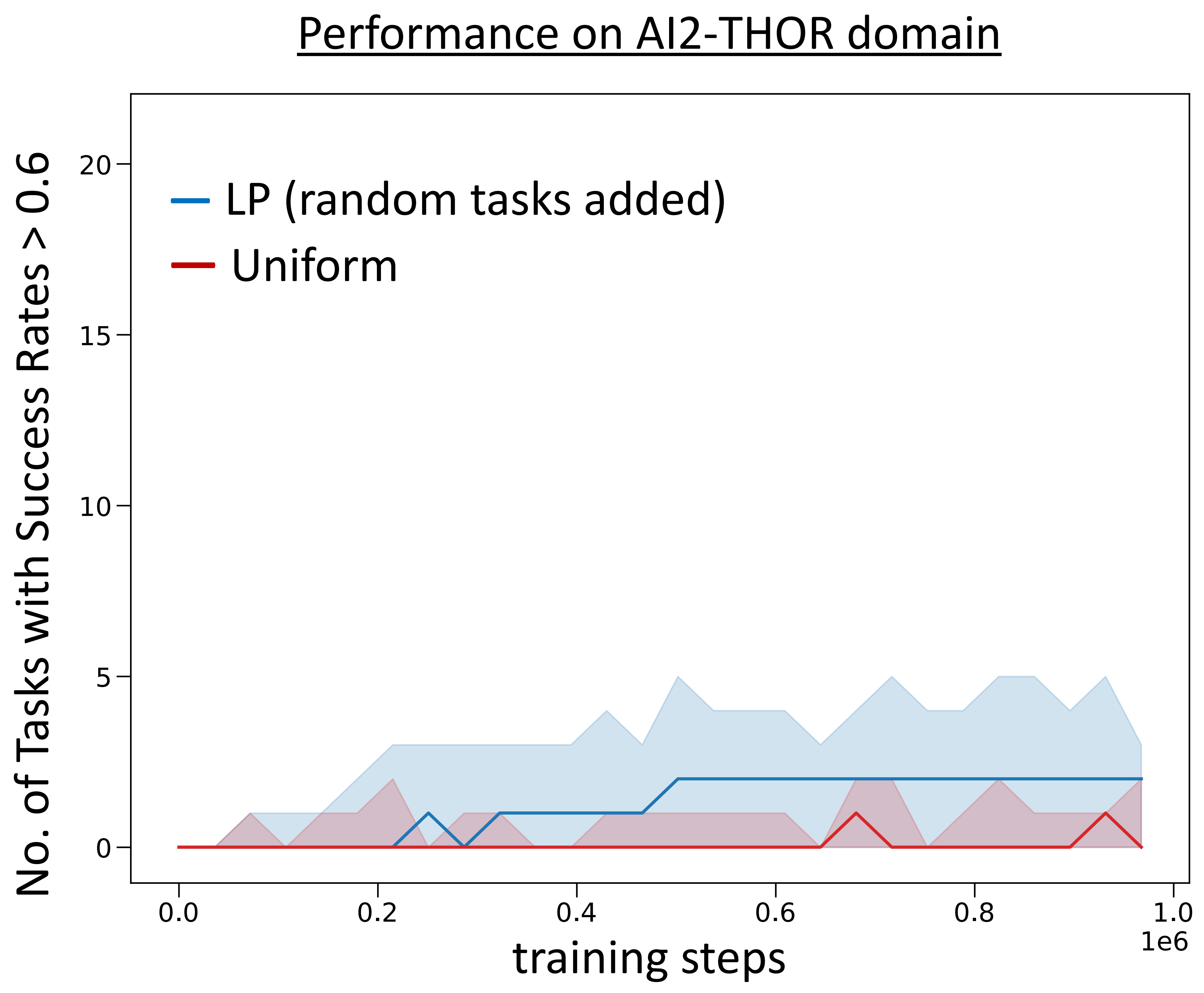
Code to check if the task is completed

```
1. [[obj_attributes("Knife", "visible": True)], [obj_attributes("Knife", "isPickedUp": True)]]
2. …
```

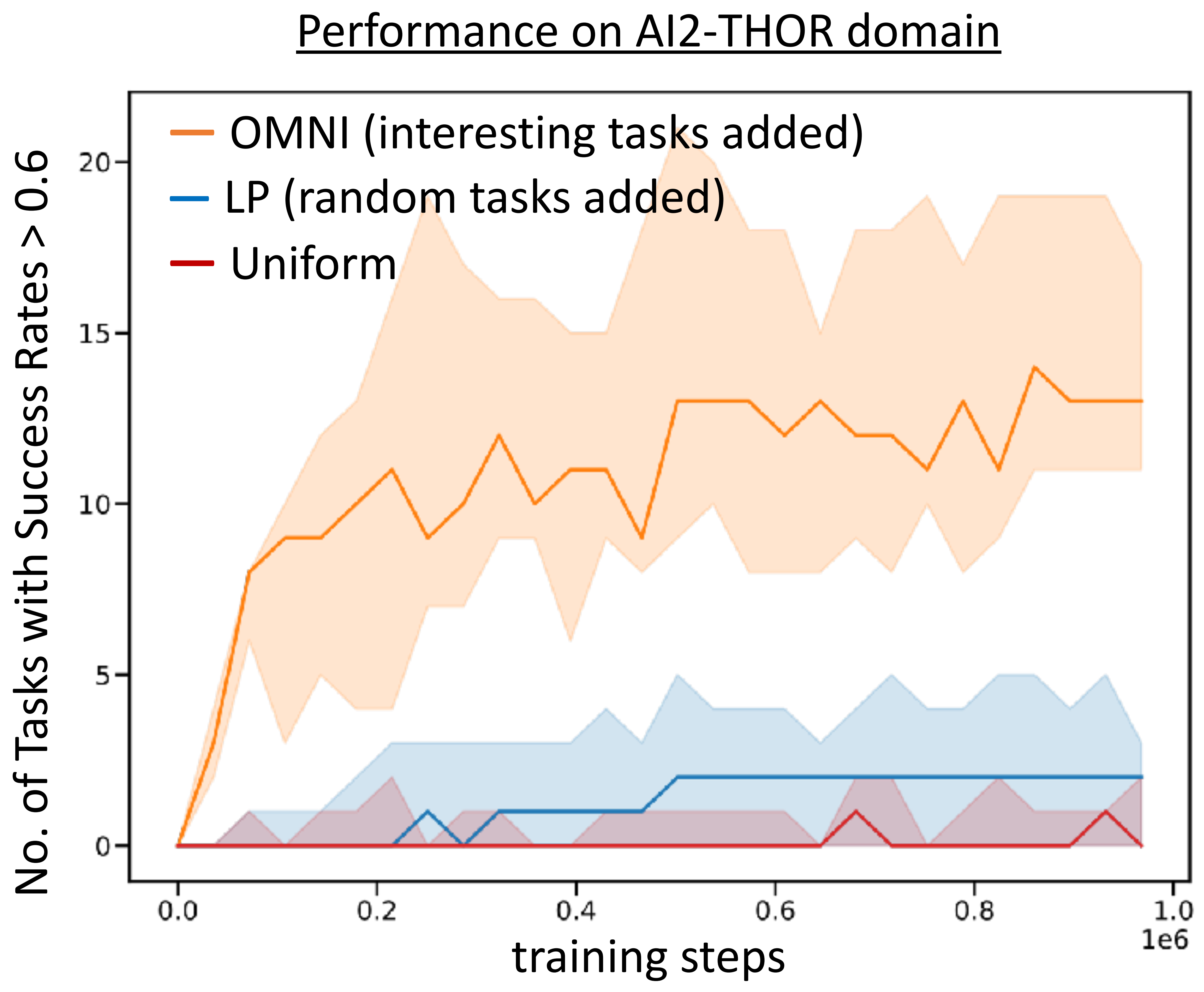
Infinite Task Space



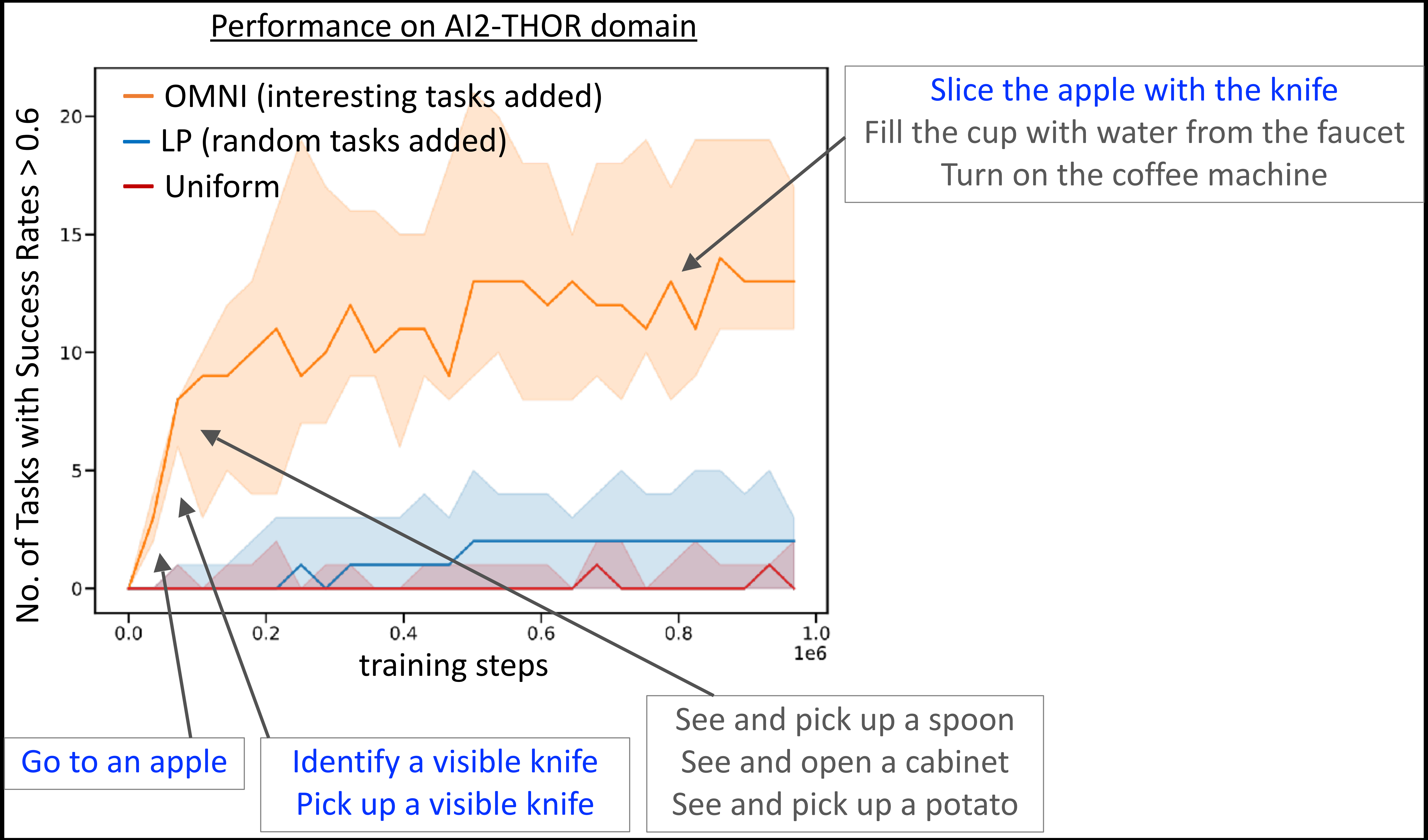
Infinite Task Space



Infinite Task Space



Infinite Task Space



OMNI

- Pro: infinite task space
- Con: still one type of environment

- Needed: Darwin Complete environmental search space (Clune 2019)
  - Can express any environment

Darwin Complete

Clune 2019, AI-Generating Algorithms

- A search space that can express any environment

- So far, I've only thought of two
  1. Code!

OMNI-EPIC

Open-endedness via Models of human Notions of Interestingness via Environments Programmed In Code

Jenny Zhang

Maxence Faldor

Antoine Cully

Jeff Clune

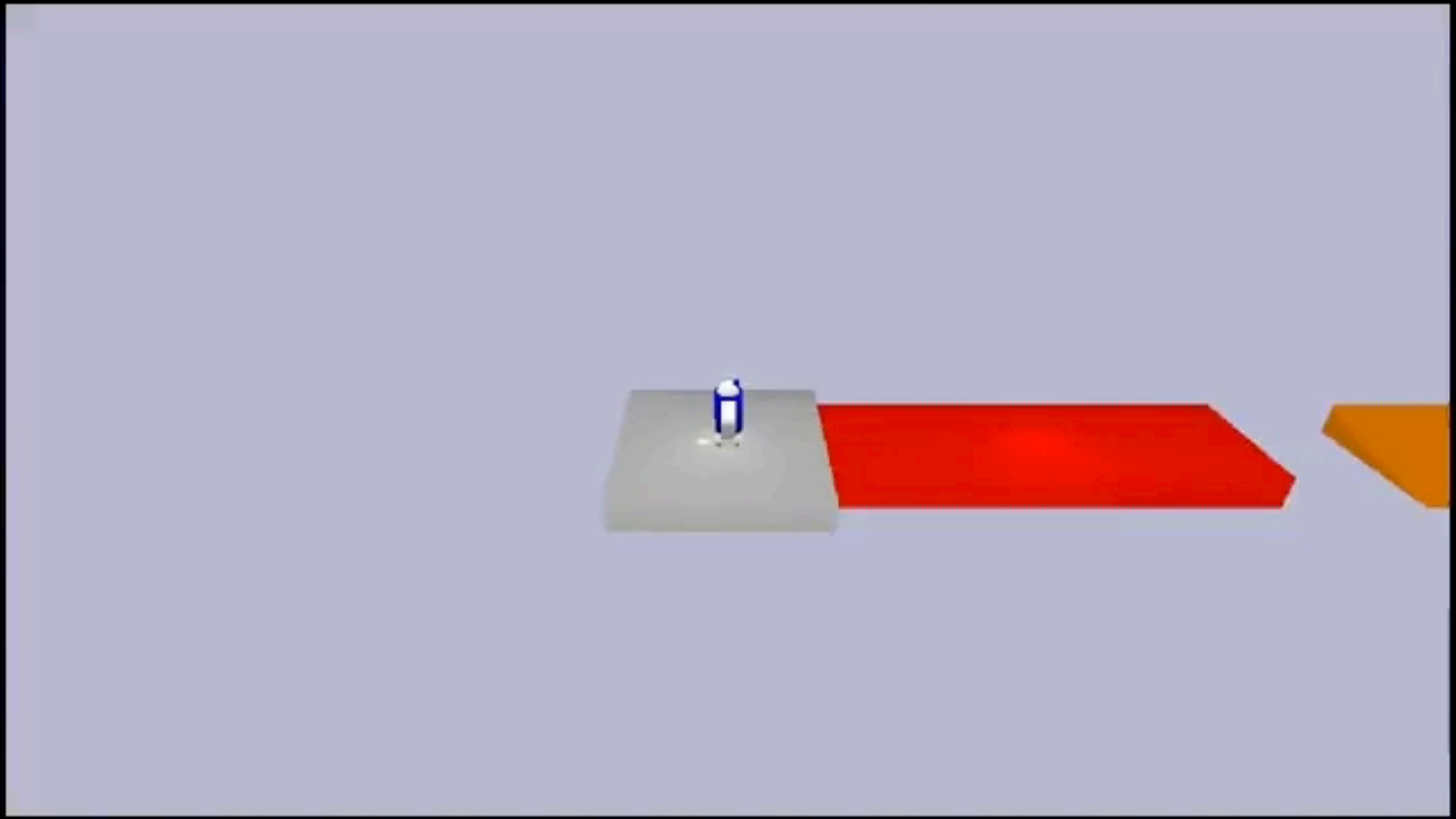
Idea

- Have OMNI produce code for new environments and reward functions
- Is thus Darwin Complete!

Seed 1

Cross a pride-colored bridge with gaps to reach a platform

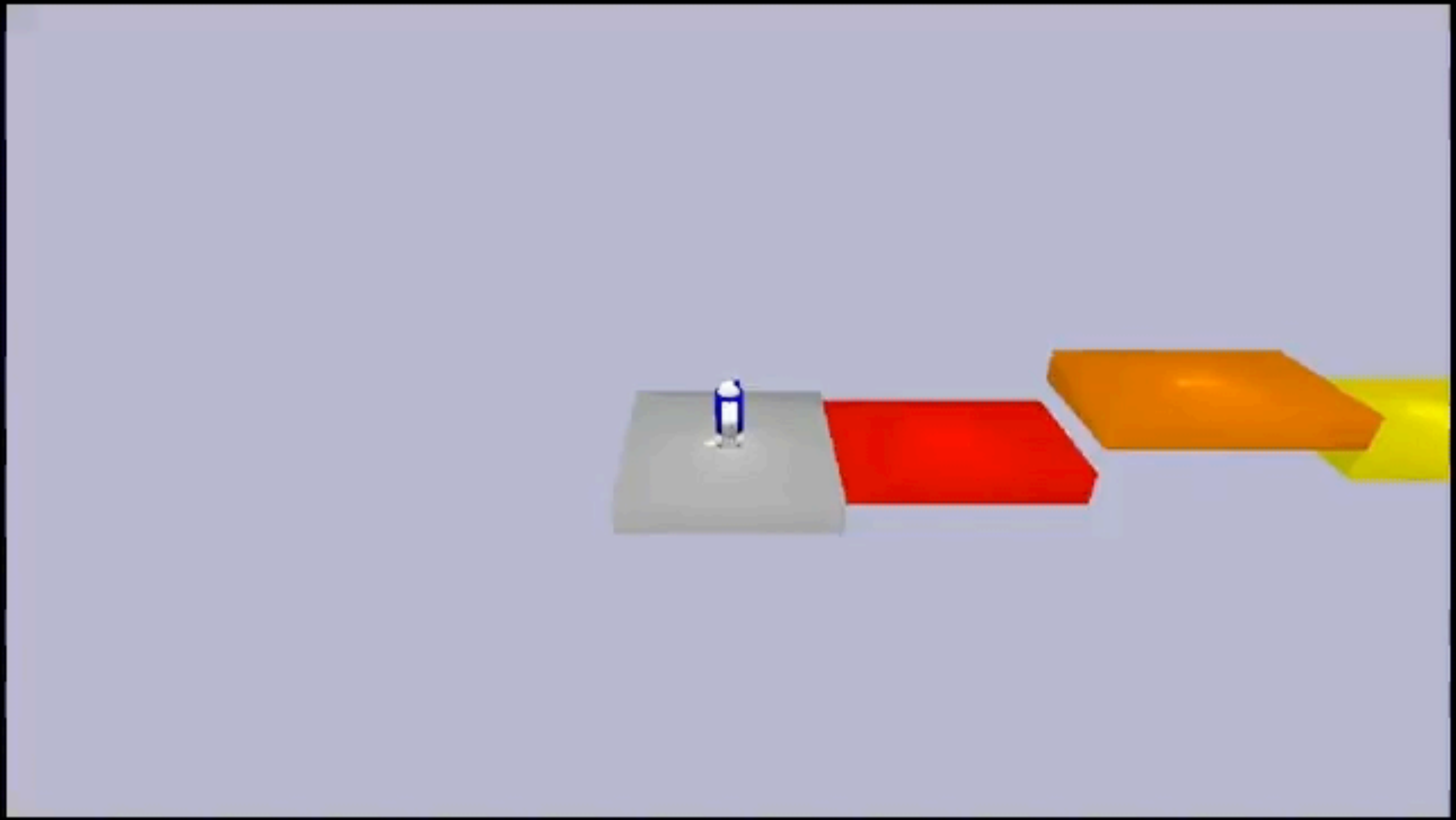
success detector: true
human review: success



Task 1

Cross a bridge with moving segments to reach a platform

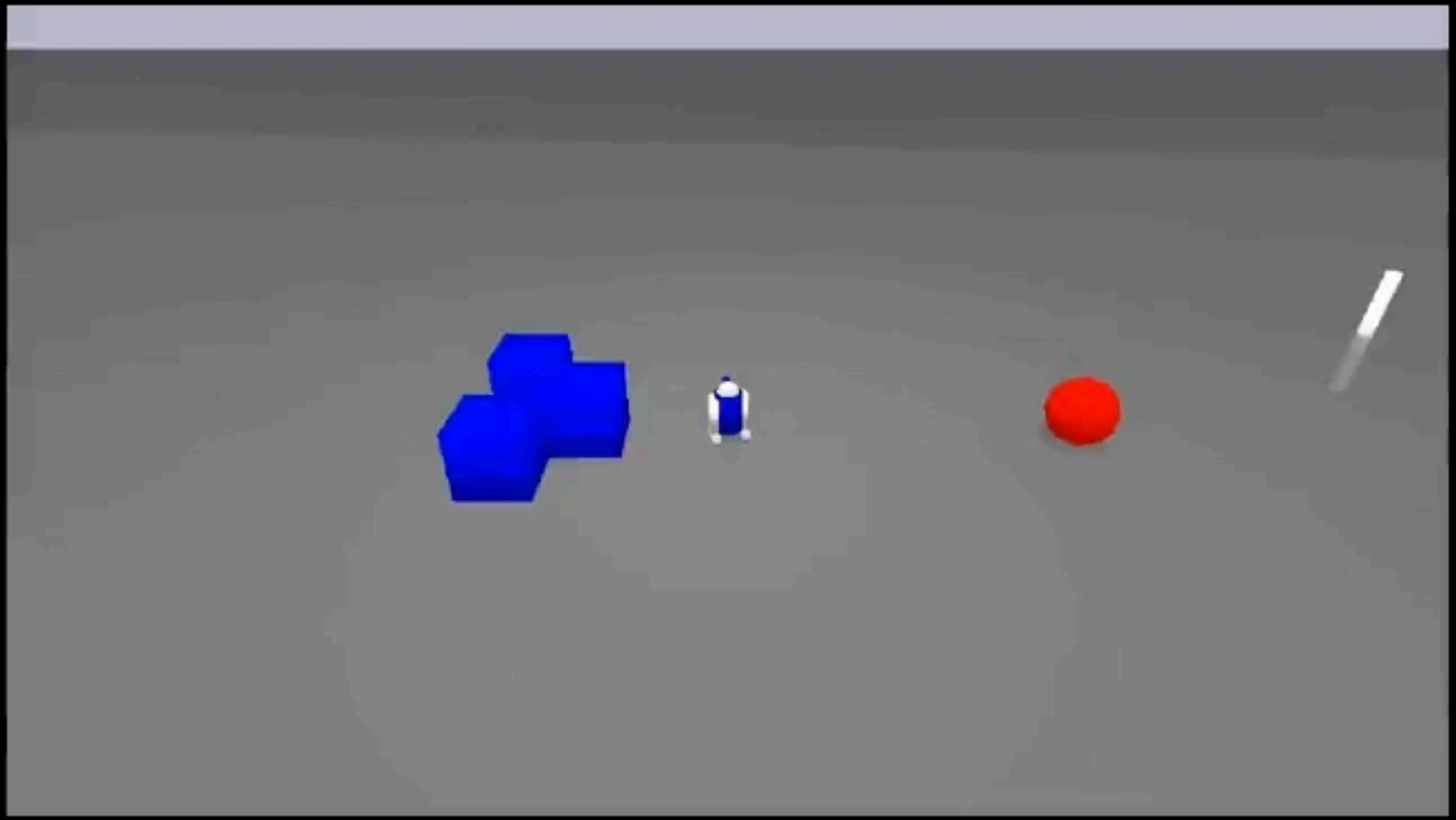
success detector: true
human review: success



Task 3

Kick a ball into a goal with obstacles

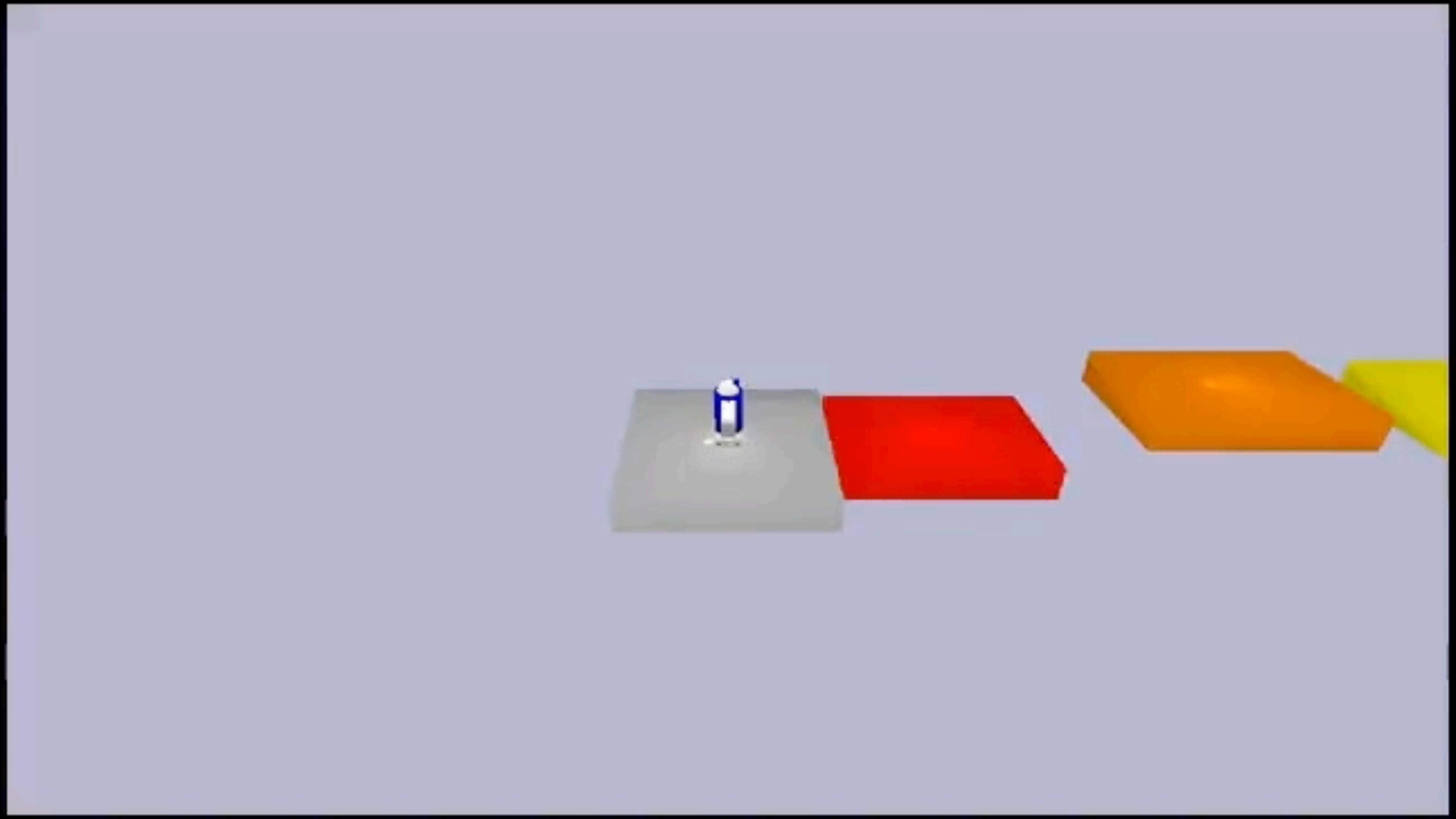
success detector: true
human review: success



Task 4

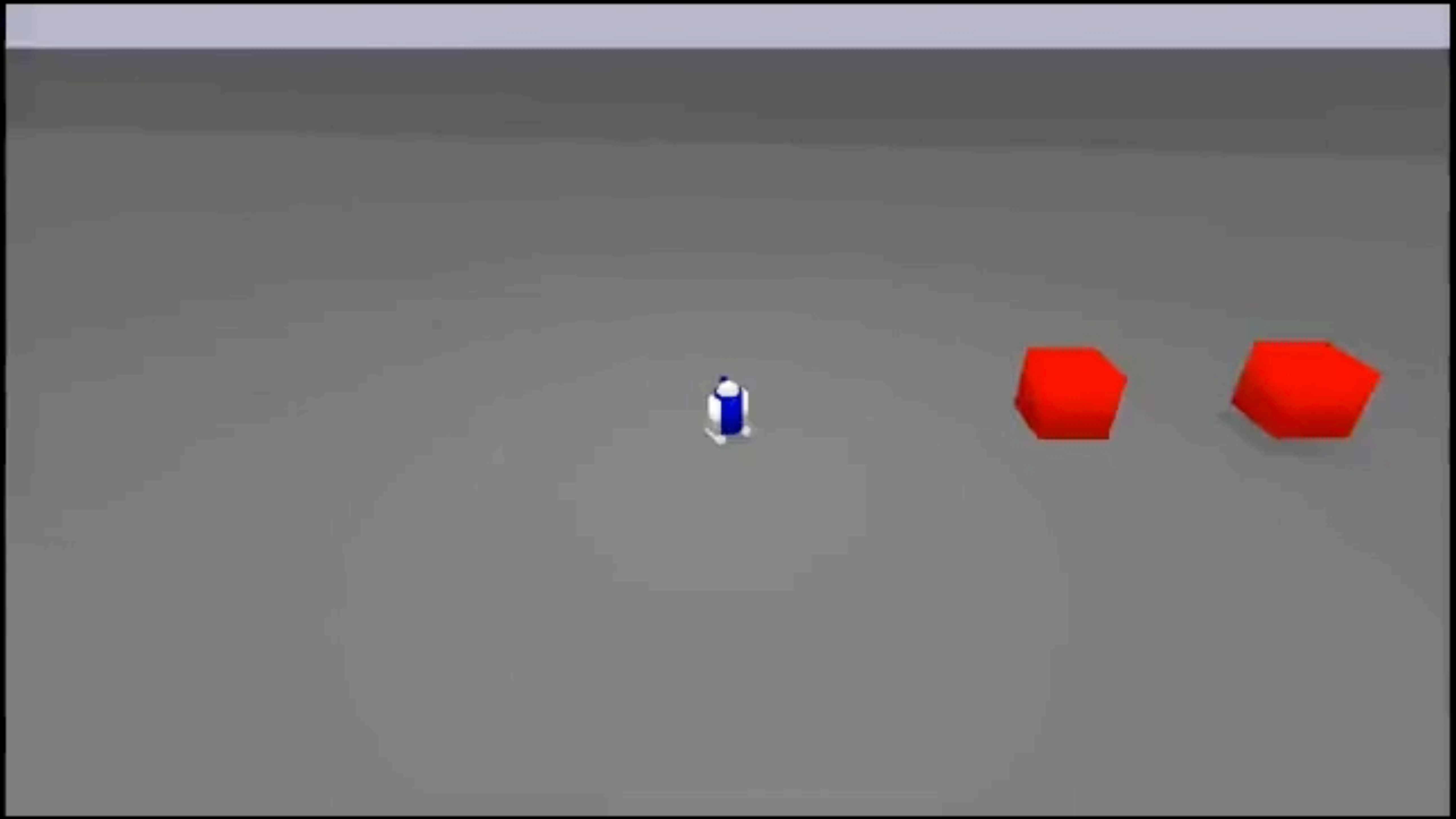
Cross a rainbow bridge with moving segments and gaps to reach a platform

success detector: true
human review: success



Task 8

Navigate a dynamic obstacle course to reach a target

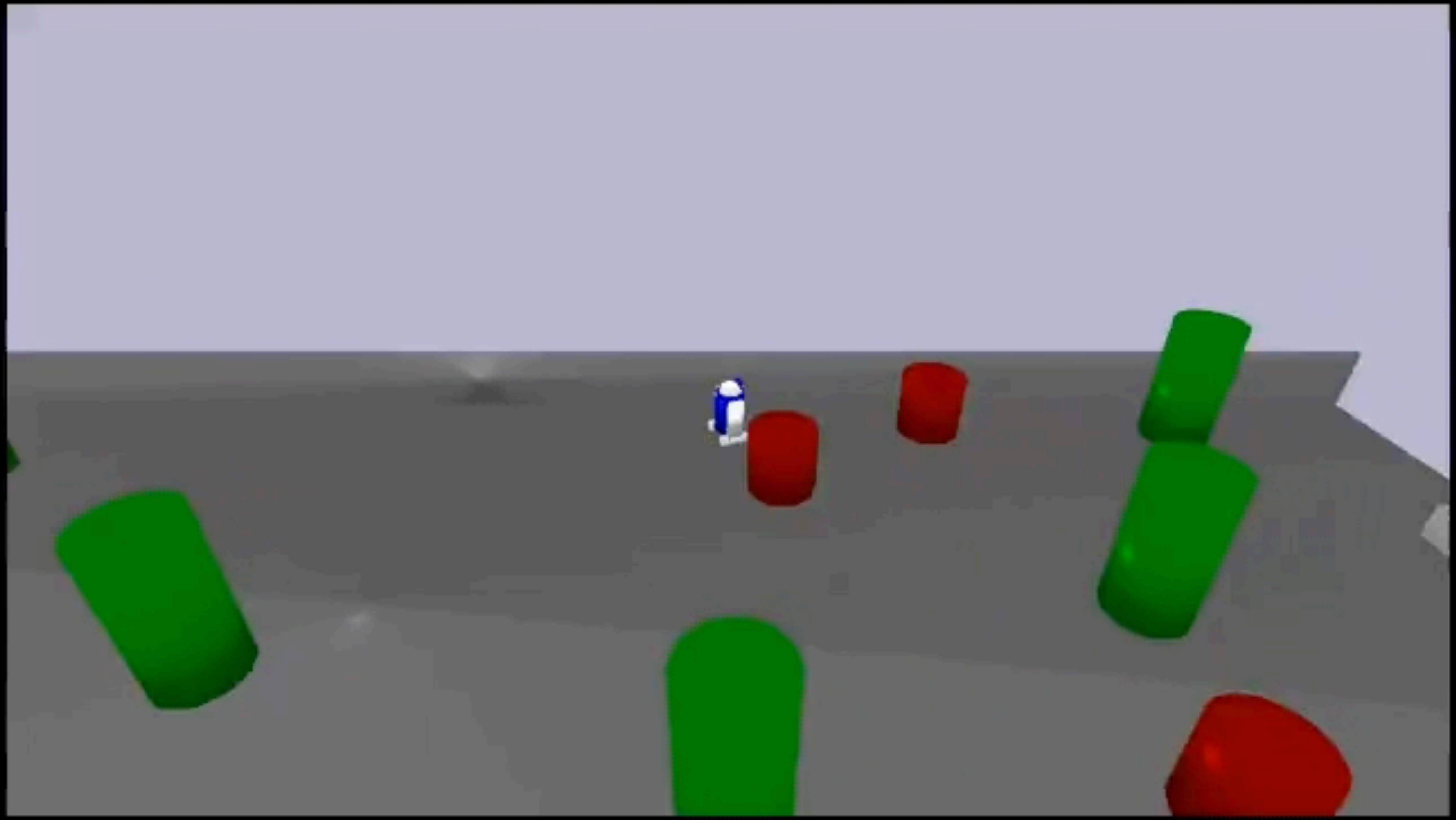


success detector: true
human review: success

Task 12

Navigate a terrain with obstacles to reach a target zone

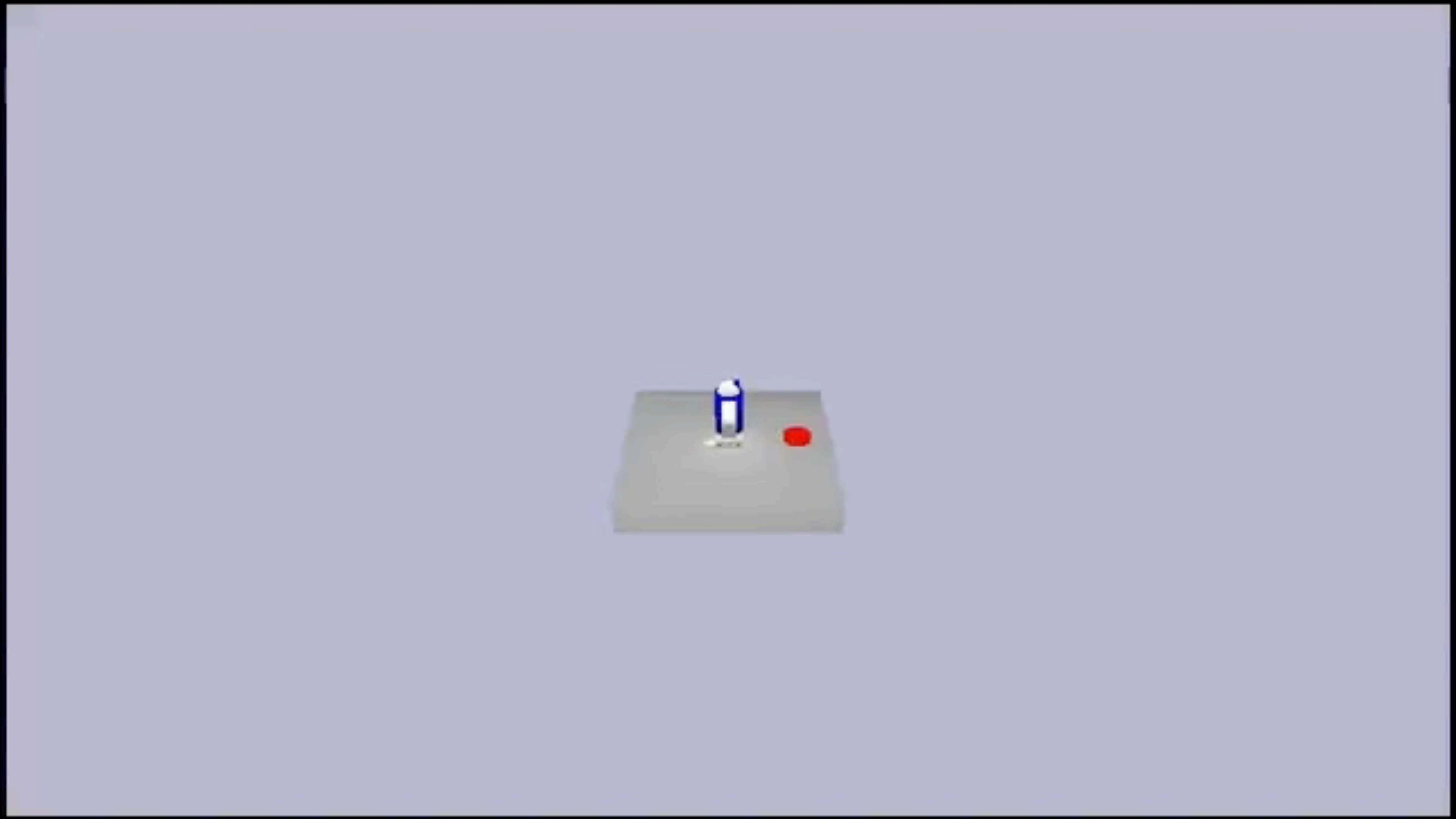
success detector: false
human review: fail



Task 17

Touch a lever or a button to activate a drawbridge

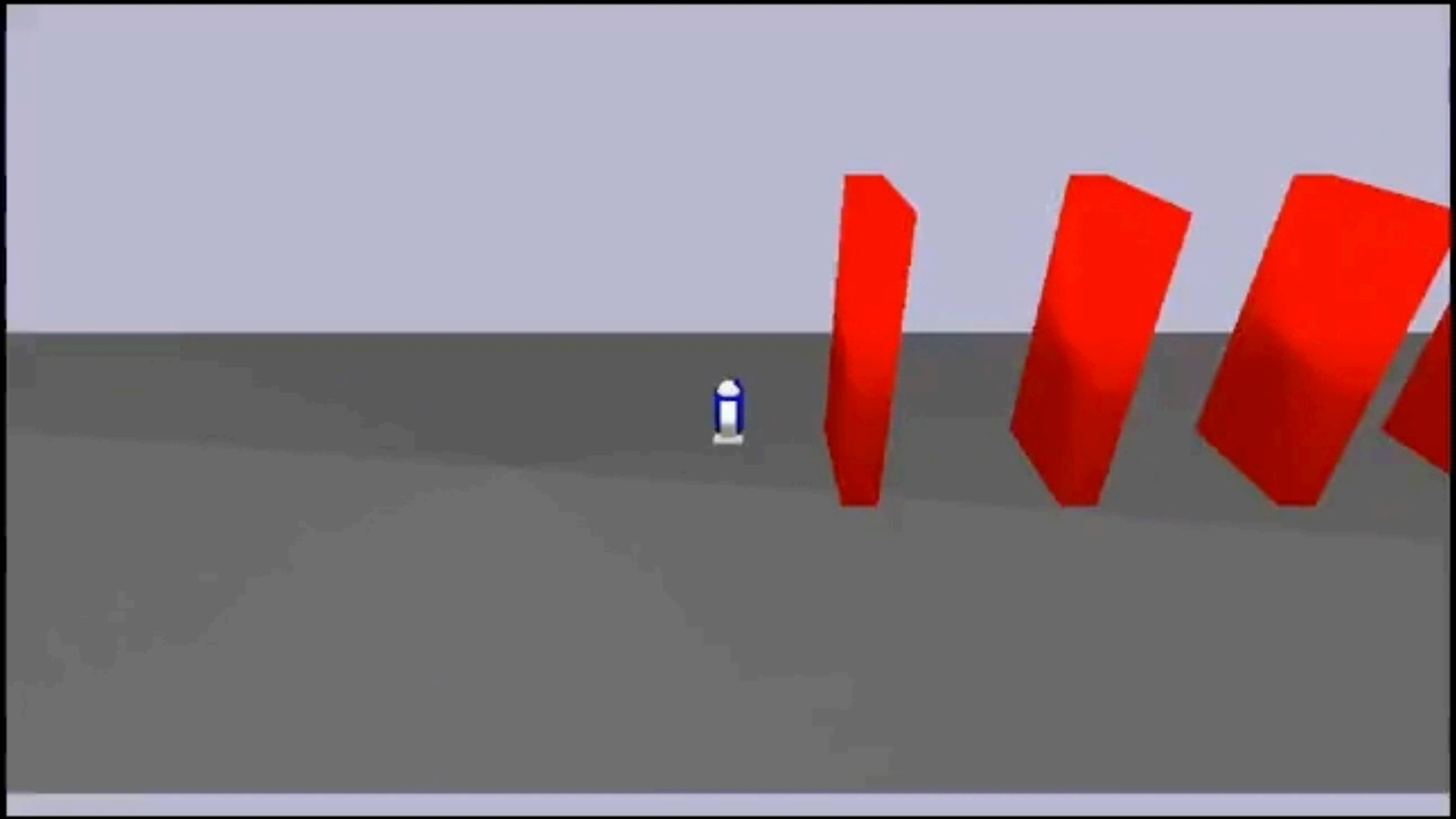
success detector: true
human review: success



Task 20

Push a domino to start a chain reaction

success detector: true
human review: success



Task 21

Task: Destroy a tower of blocks

success detector: true
human review: success

Task 84
Busy Restaurant Table Service.
1 20 40 60 80 100 120 140 160 180 200
Generation Number
Successfully learned
ZZZ Not interesting
Task 90
Outdoor Retrieval Challenge.
Task 23
Search and Rescue in a Multi-Room Environment.
Task 76
Construction Site Cleanup.
Seed 1
Cross a pride-colored bridge with gaps to reach a platform.
Task 103
Conveyor Belt Item Delivery.
Task 155
Dynamic Target Pushing.
Task 88
Bounce Shot Scoring.
Seed 2
Ascend a series of stairs to reach a platform.
Task 178
Coordinate with another moving ball to hit a target.
Task 197
The robot must navigate through the platforms, jumping over the gaps.
Task 196
Dodgeball Arena.
Task 176
Push two balls into separate goals while avoiding static obstacles.
Task 122
Kick ball through moving gates to zones marked by moving targets.
Seed 3
Kick a ball into a goal.
Task 74
Receive pass and kick ball past moving obstacles and goalie into goal.
Task 60
Simplified 1-vs-1 robot soccer.

OMNI: Conclusion and Future Work

- OMNI: Open-endedness via Models of human Notions of Interestingness
- Finally (!!) enables us to focus on interesting tasks
  - FMs have learned that ineffable concept from humans!
- OMNI-EPIC: generate any possible environment
- Excited to scale OMNI and see its creativity: billions of years?

Darwin Complete

Clune 2019, AI-Generating Algorithms

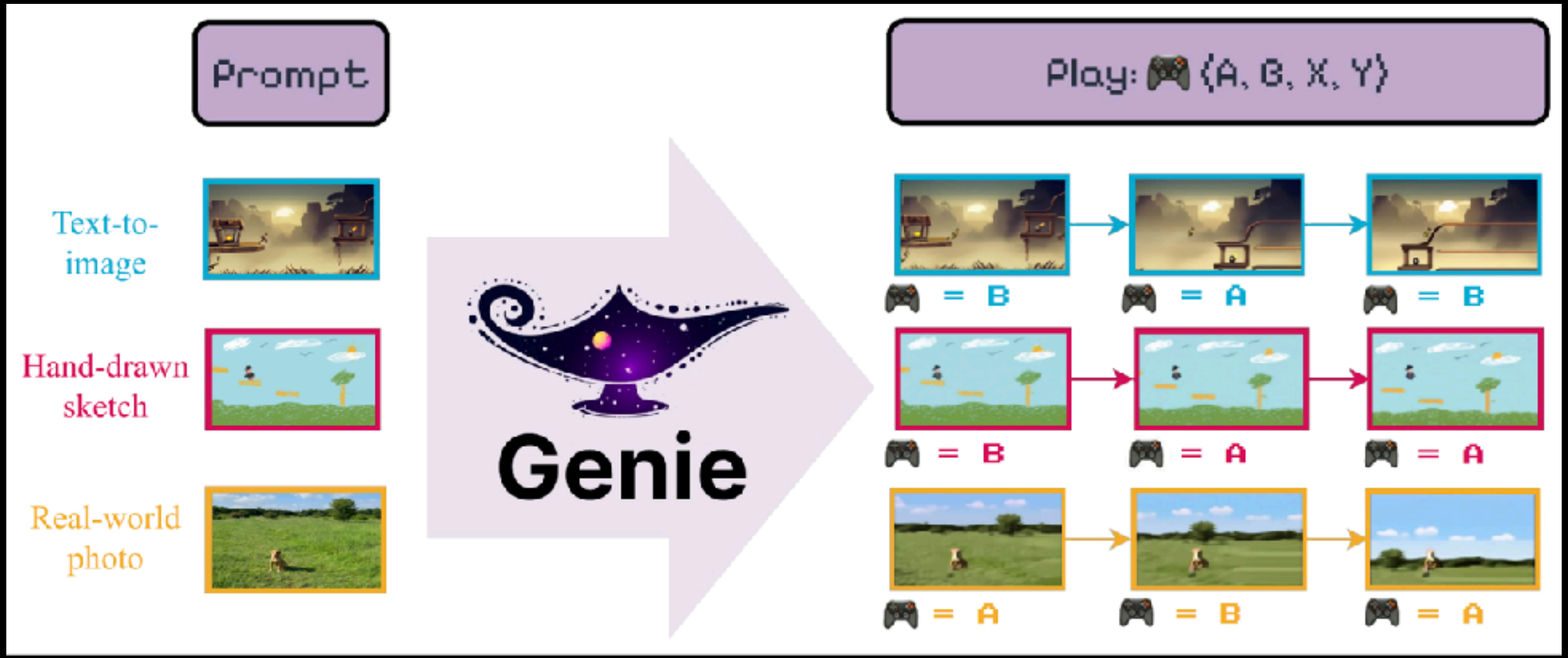
- Can express any environment
- So far, I've only thought of two
  - Original Paper: Deep Neural Network
  - Recently: Code! (OMNI-EPIC)

Genie: Generative Interactive Environments

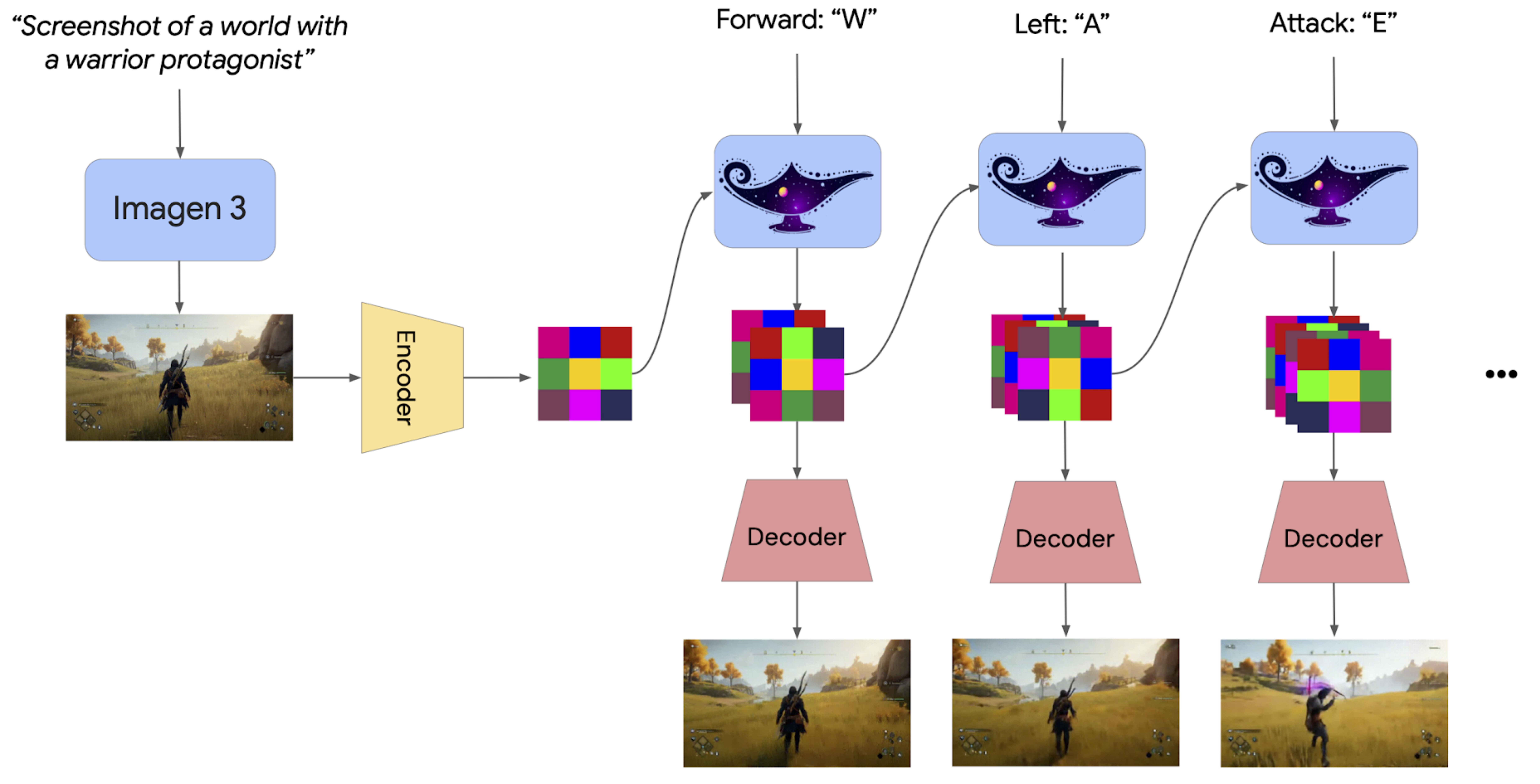
Google DeepMind

Jake Bruce*, Michael Dennis*, Ashley Edwards*, Jack Parker-Holder*, Yuge (Jimmy) Shi*, Edward Hughes, Matthew Lai, Aditi Mavalankar, Richie Steigerwald, Chris Apps, Yusuf Aytar, Sarah Bechtle, Feryal Behbahani, Stephanie Chan, Nicolas Heess, Lucy Gonzalez, Simon Osindero, Sherjil Ozair, Scott Reed, Jingwei Zhang, Konrad Zolna, Jeff Clune, Nando de Freitas, Satinder Singh, Tim Rocktäschel*

- A world model that can generate any environment (a video game)
  - conditioned on any of the following: text, image, sketch, etc.
- Trained from online videos



Best Paper, ICML 2024



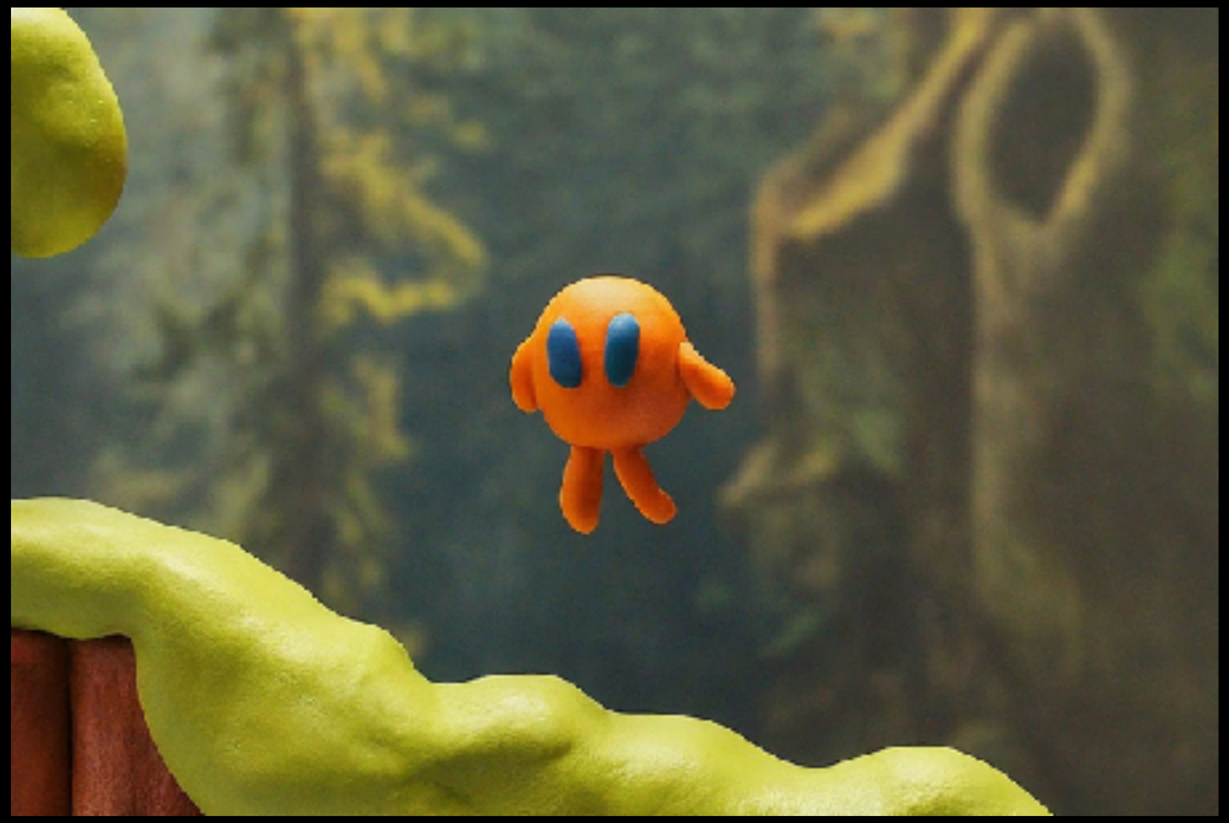
"Screenshot of a world with a warrior protagonist"
Imagen 3
Encoder
Forward: "W"
Left: "A"
Attack: "E"
Decoder
Decoder
Decoder

Genie: Generative Interactive Environments

Generate any environment via text, image, sketch, etc.

From Online Videos with *No Labels*

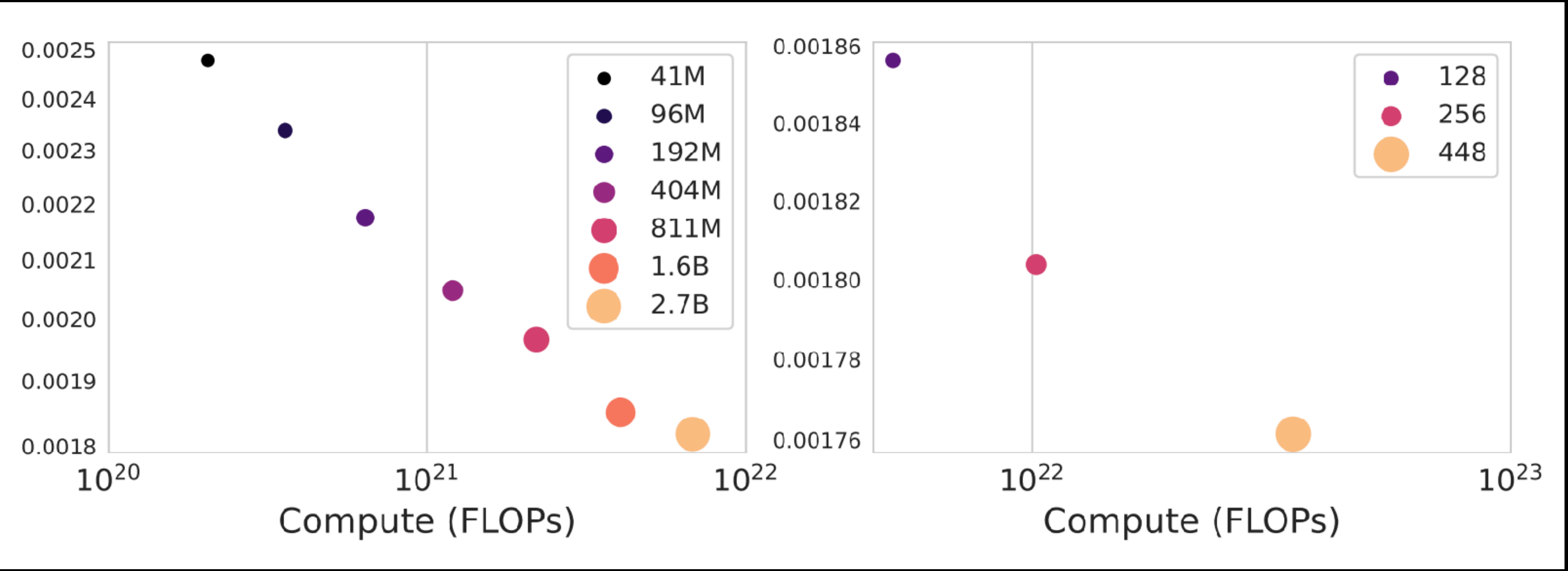
Generate a playable world
set in a futuristic city



Jake Bruce*, Michael Dennis*, Ashley Edwards*, Jack Parker-Holder*, Yuge (Jimmy) Shi*, Edward Hughes, Matthew Lai, Aditi Mavalankar, Richie Steigerwald, Chris Apps, Yusuf Aytar, Sarah Bechtle, Feryal Behbahani, Stephanie Chan, Nicolas Heess, Lucy Gonzalez, Simon Osindero, Sherjil Ozair, Scott Reed, Jingwei Zhang, Konrad Zolna, Jeff Clune, Nando de Freitas, Satinder Singh, Tim Rocktäschel*

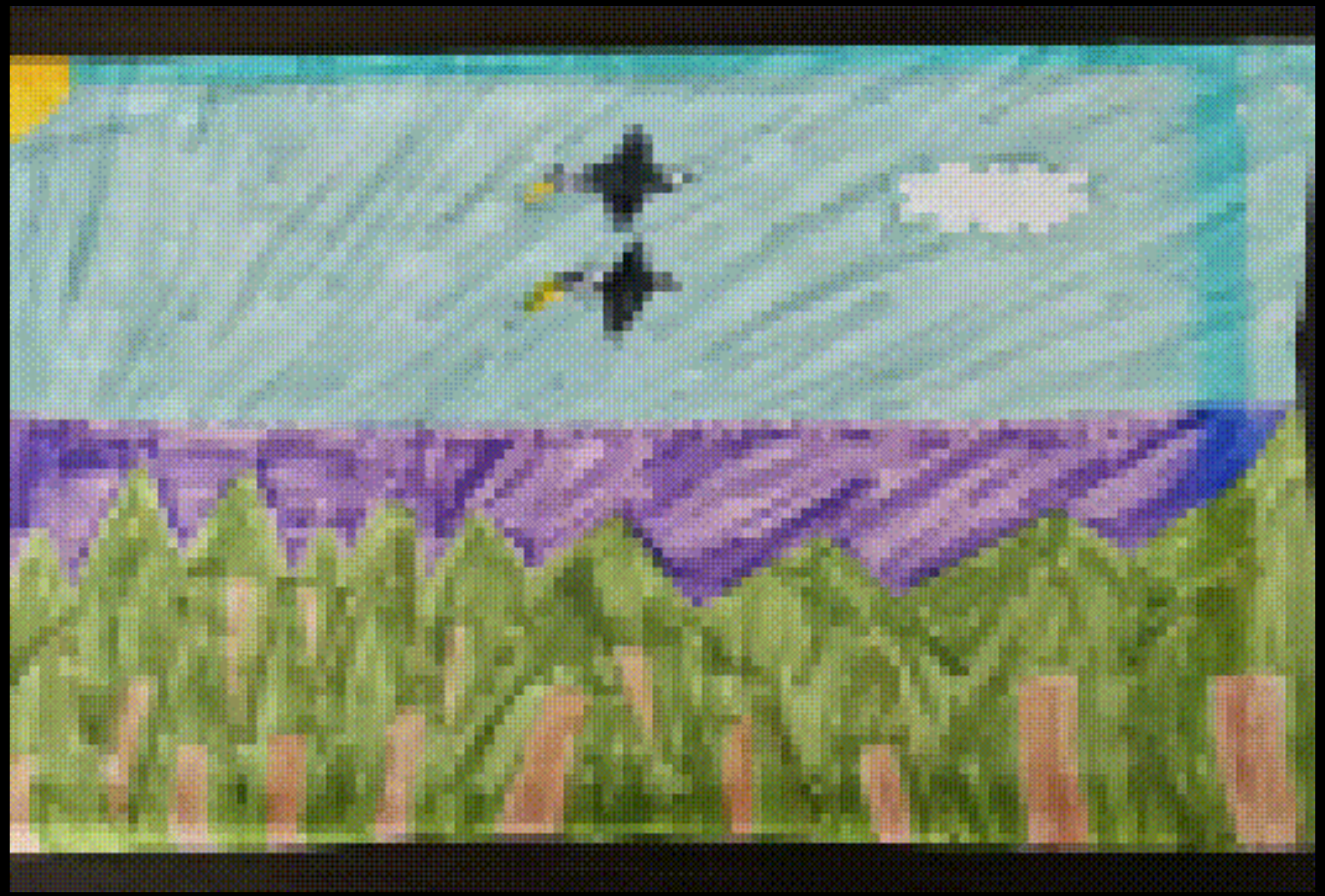
https://sites.google.com/view/genie-2024

Scaling Laws: More compute = better



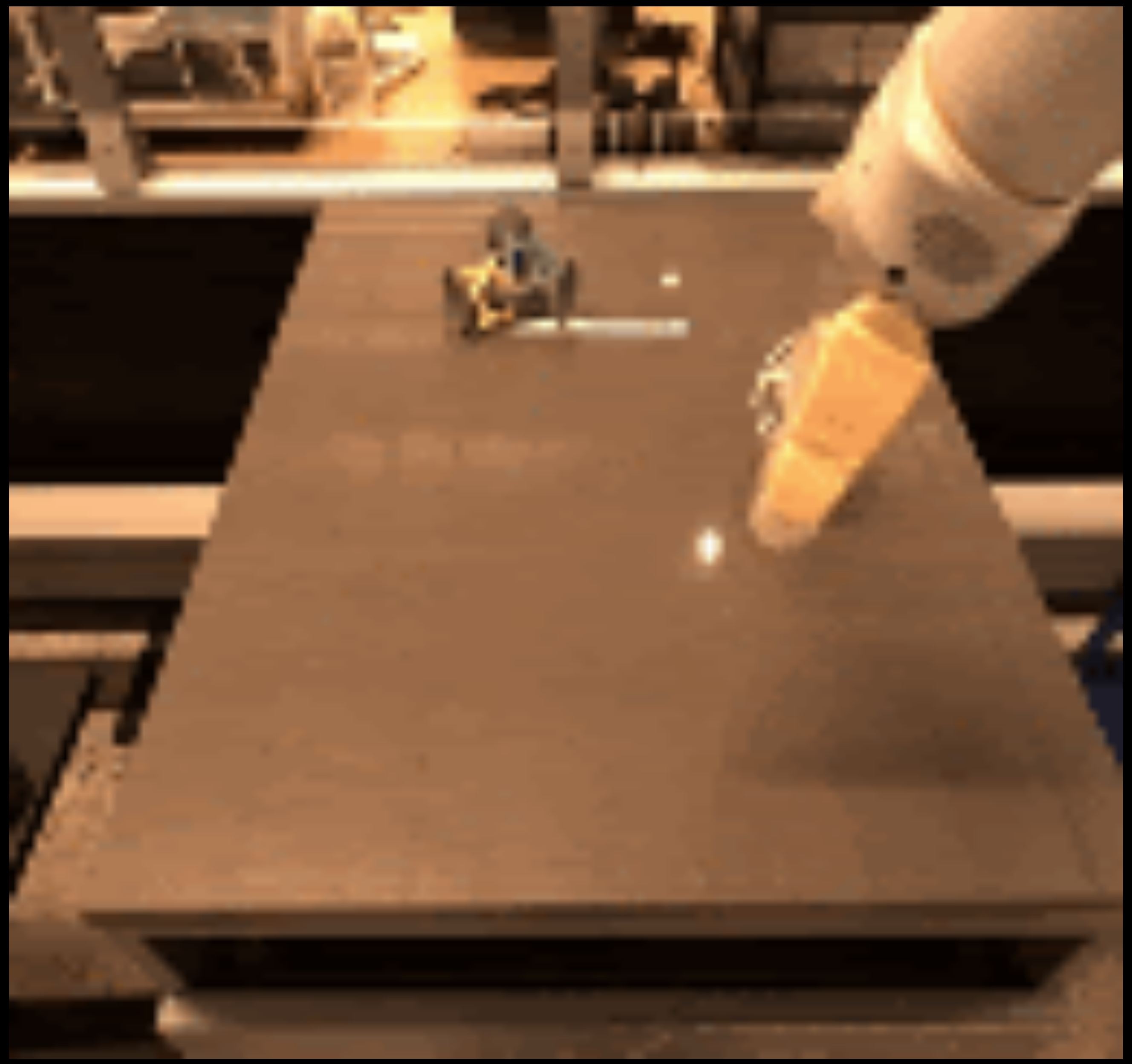
Results from Genie 1

Genie: Generative Interactive Environments



Video Games, by Seneca and Caspian Clune
(Ages 7 and 9)

Genie: Generative Interactive Environments



Can generate robotics training environments too

https://sites.google.com/view/genie-2024

Foundation World Model

- Genie can be considered a “foundation world model”
- Lots of analogies to Foundation Models
  - General: can generate any environment
  - Foundational: Lots can be build upon it
  - Similar methods apply
    - It can be fine-tuned
    - RLHFed, etc.
- Scaling laws apply
- Etc.

Genie 2: A large-scale foundation world model

Generate a playable world
on a spaceship

https://deepmind.google/discover/blog/genie-2-a-large-scale-foundation-world-model

Go through the left door

Go through the center door

Go through the right door


Q W E
A S D

Imagine

- Open-ended algorithms + Genie
  - Including Genie 3, 4, 5…10
- Incredibly exciting and powerful
- “Genie 2 could enable future agents to be trained and evaluated in a limitless curriculum of novel worlds” - Parker-Holder et al. 2024

AI-Generating Algorithms

Clune 2019

Three Pillars

1. Meta-learn architectures
2. Meta-learn learning algorithms
3. Generate effective learning environments
    - Quality Diversity algorithms
    - Open-Ended algorithms
        - e.g. POET, OMNI, OMNI-EPIC
    - Foundation World Models
        - e.g. Genie

AI-Generating Algorithms

Clune 2019

Three Pillars

1. Meta-learn architectures
2. Meta-learn learning algorithms
3. Generate effective learning environments
    - Quality Diversity algorithms
    - Open-Ended algorithms
        - Eg. POET, OMNI, OMNI-EPIC
    - Foundation World Models

Video Pre-Training (VPT)

Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos

Bowen Baker[*†] bowen@openai.com

Ilge Akkaya[*†] ilge@openai.com

Peter Zhokhov[*†] peterz@openai.com

Joost Huizinga[*†] joost@openai.com

Jie Tang[*†] jietang@openai.com

Adrien Ecoffet[*†] adrien@openai.com

Brandon Houghton[*†] brandon@openai.com

Raul Sampedro[*†] raulsamg@gmail.com

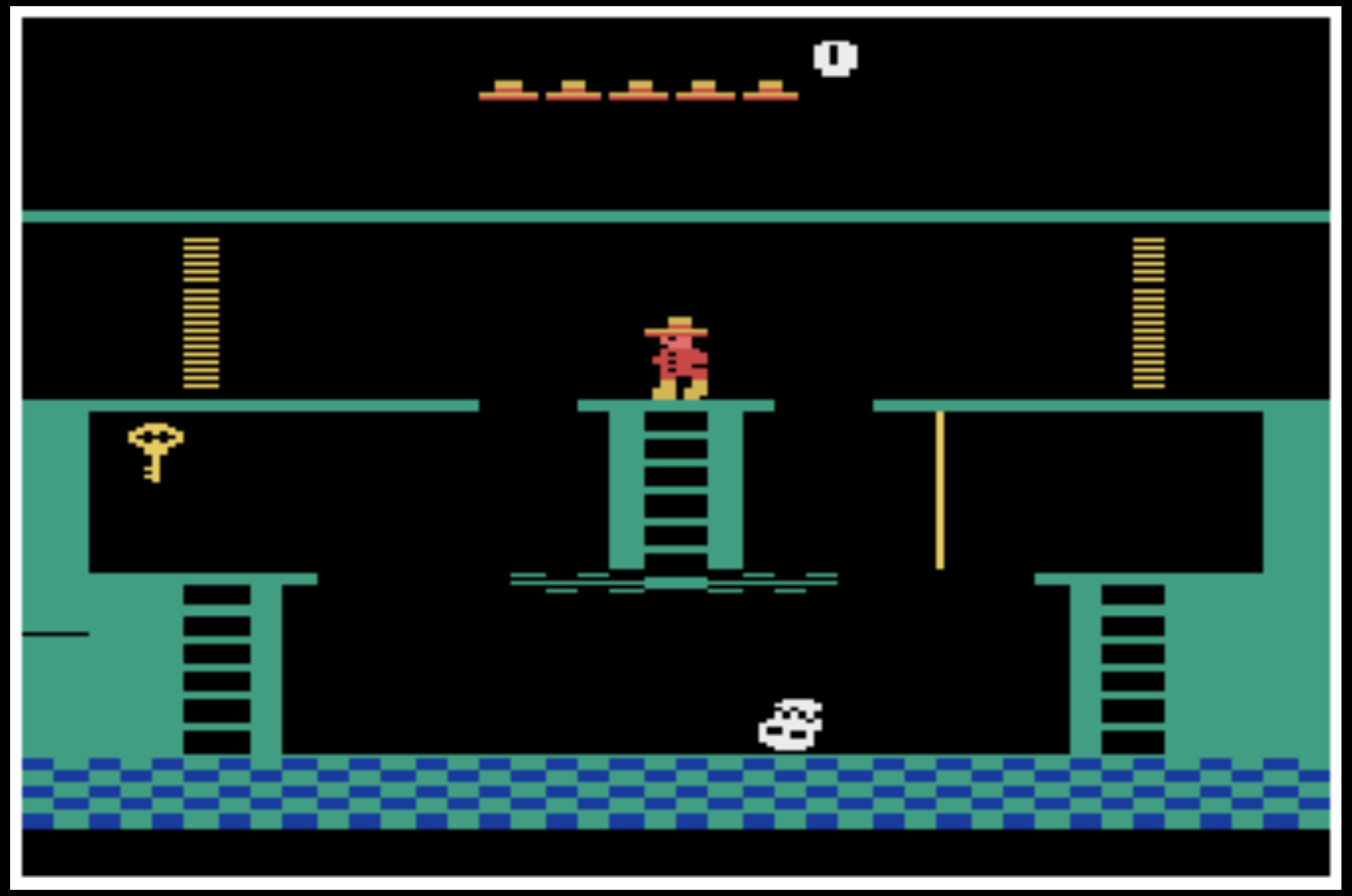
Jeff Clune[*†‡] jclune@gmail.com

NeurIPS 2022

OpenAI

Hard Exploration Prevents Open-Ended Learning

- RL often fails to learn because exploration is too hard
- Prevents open-ended algorithms
- In theory, meta-learning from scratch can address this, but it might take forever



Can we make AI-GAs without a planet-sized computer?

Yes

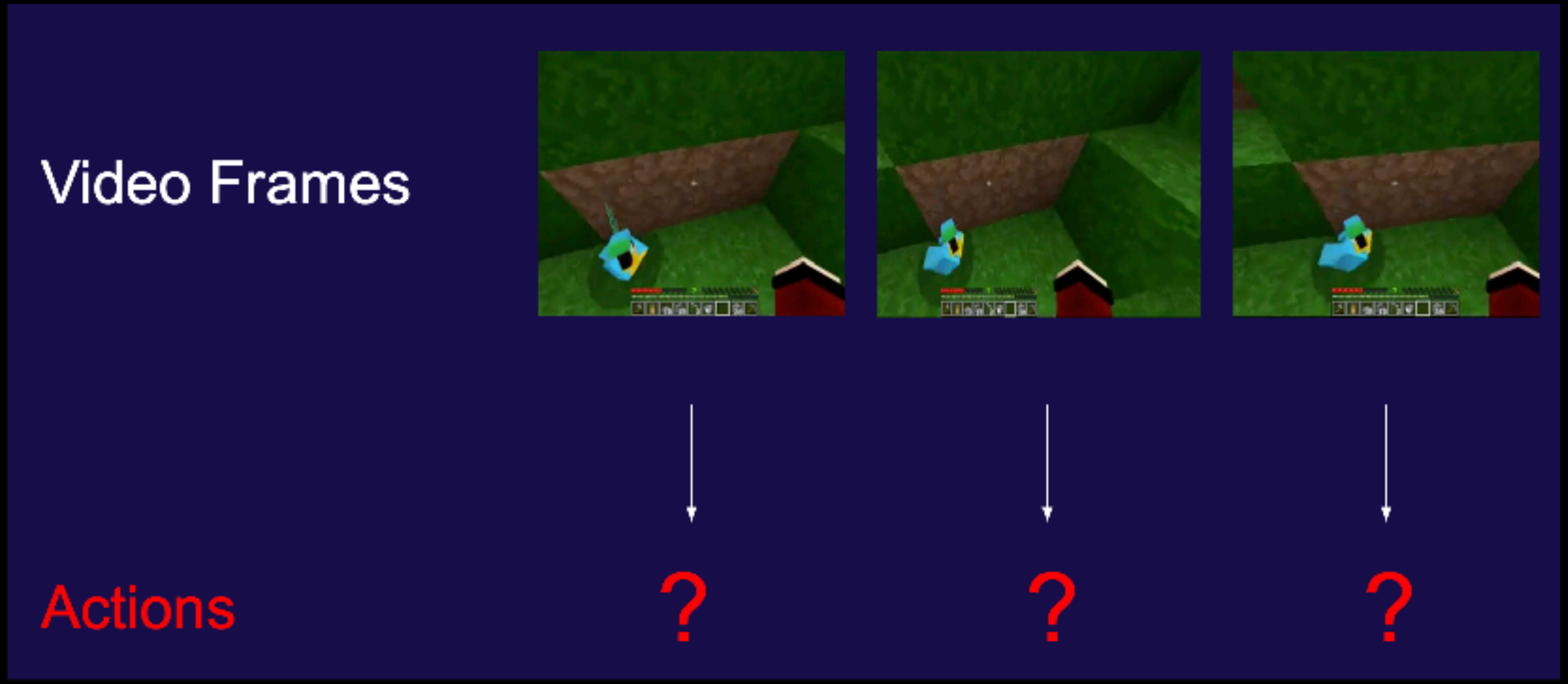
(to borrow from Newton)

AI sees further by standing on the shoulders of giant~~s~~ human datasets

Solution? Internet videos/tutorials

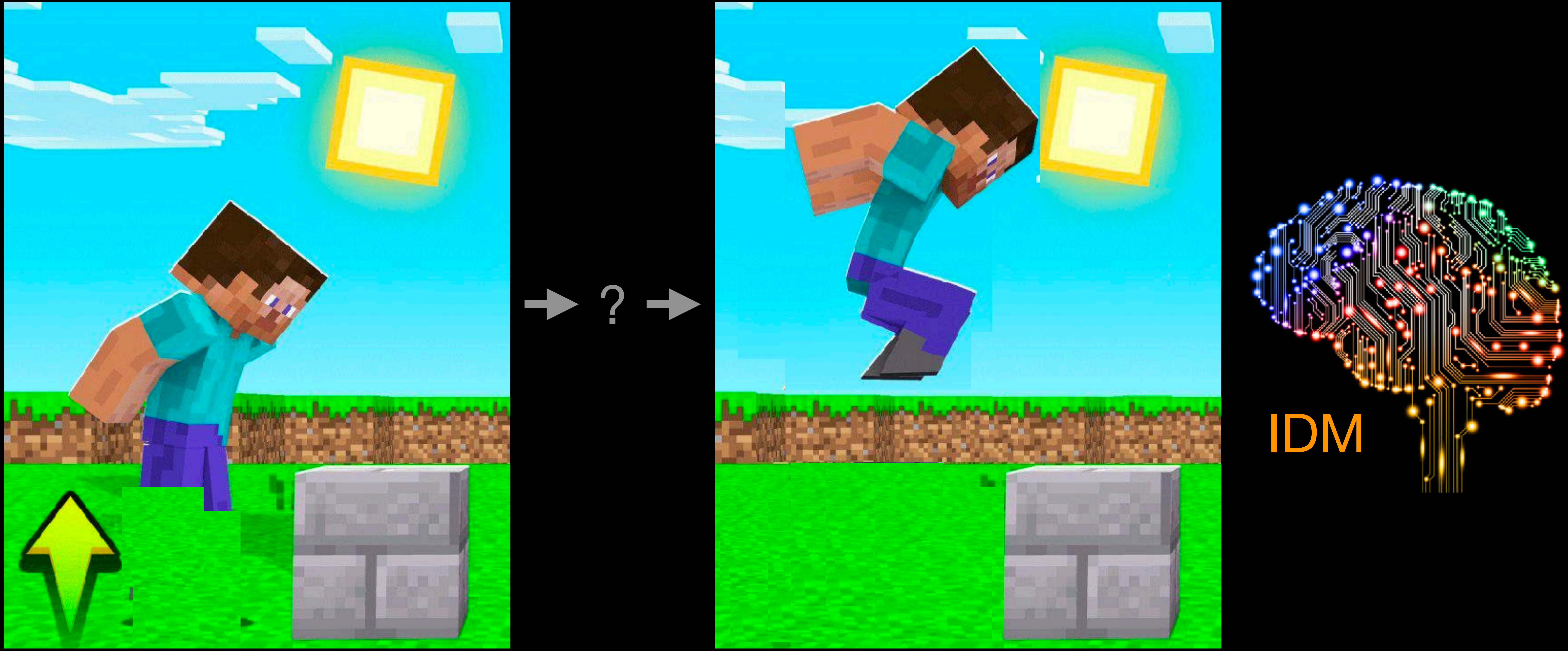
Challenge

- Videos are unlabeled
  - unlike text, music, and images, you do not get labels “for free” in online data
    - text: next word
    - music: next note
    - images: next pixel
  - we need actions taken at each timestep



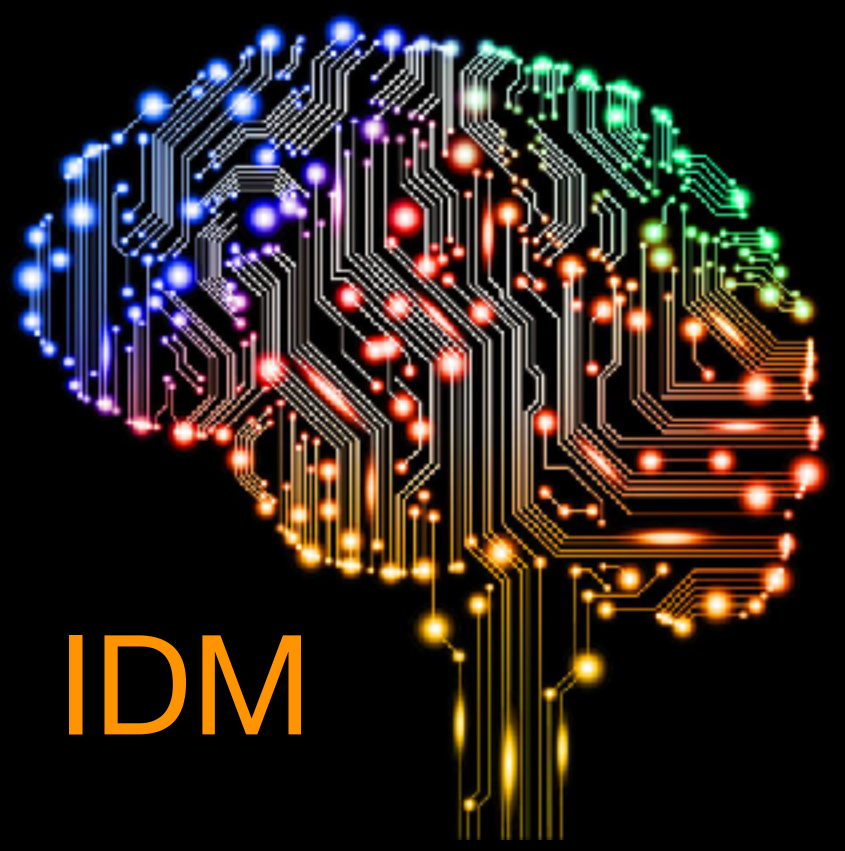
Train “Inverse Dynamics Model”

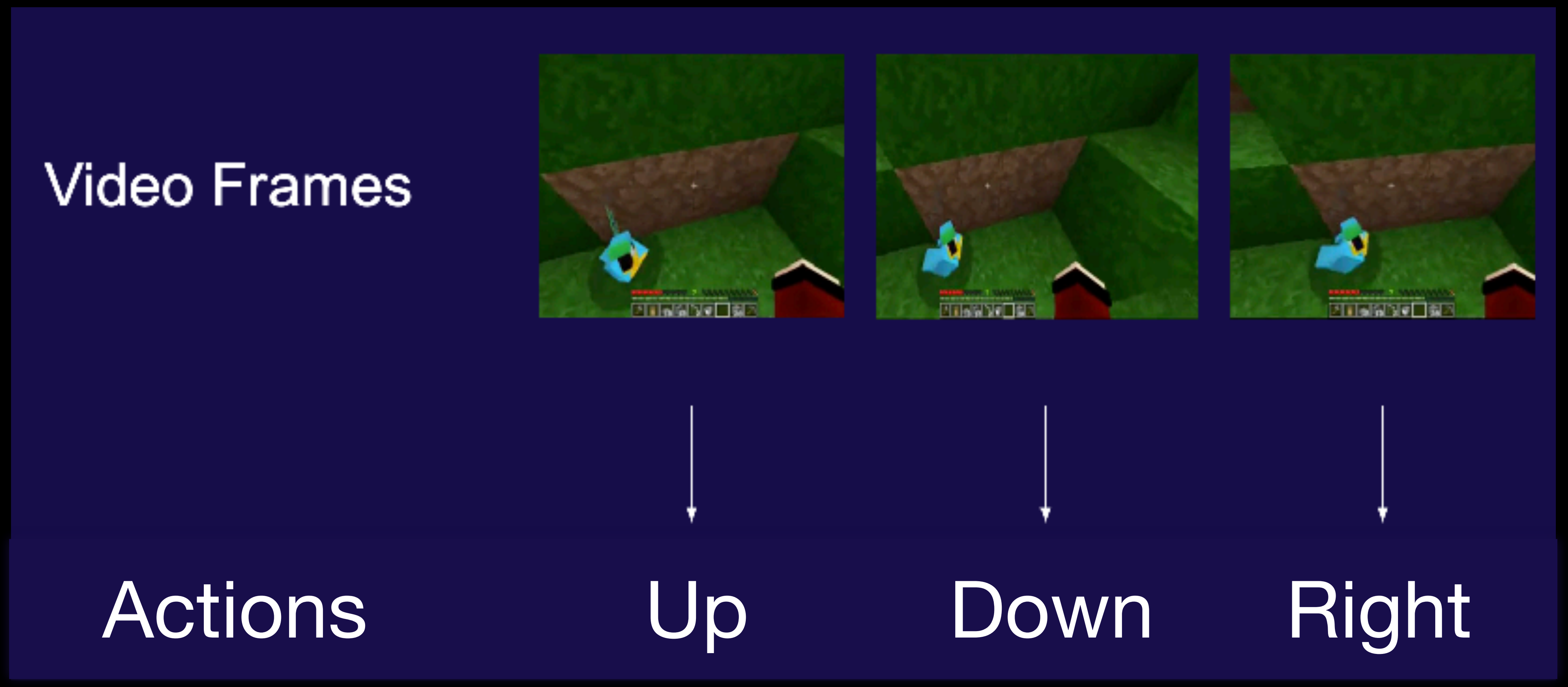
- Learns to predict action taken based on past and future



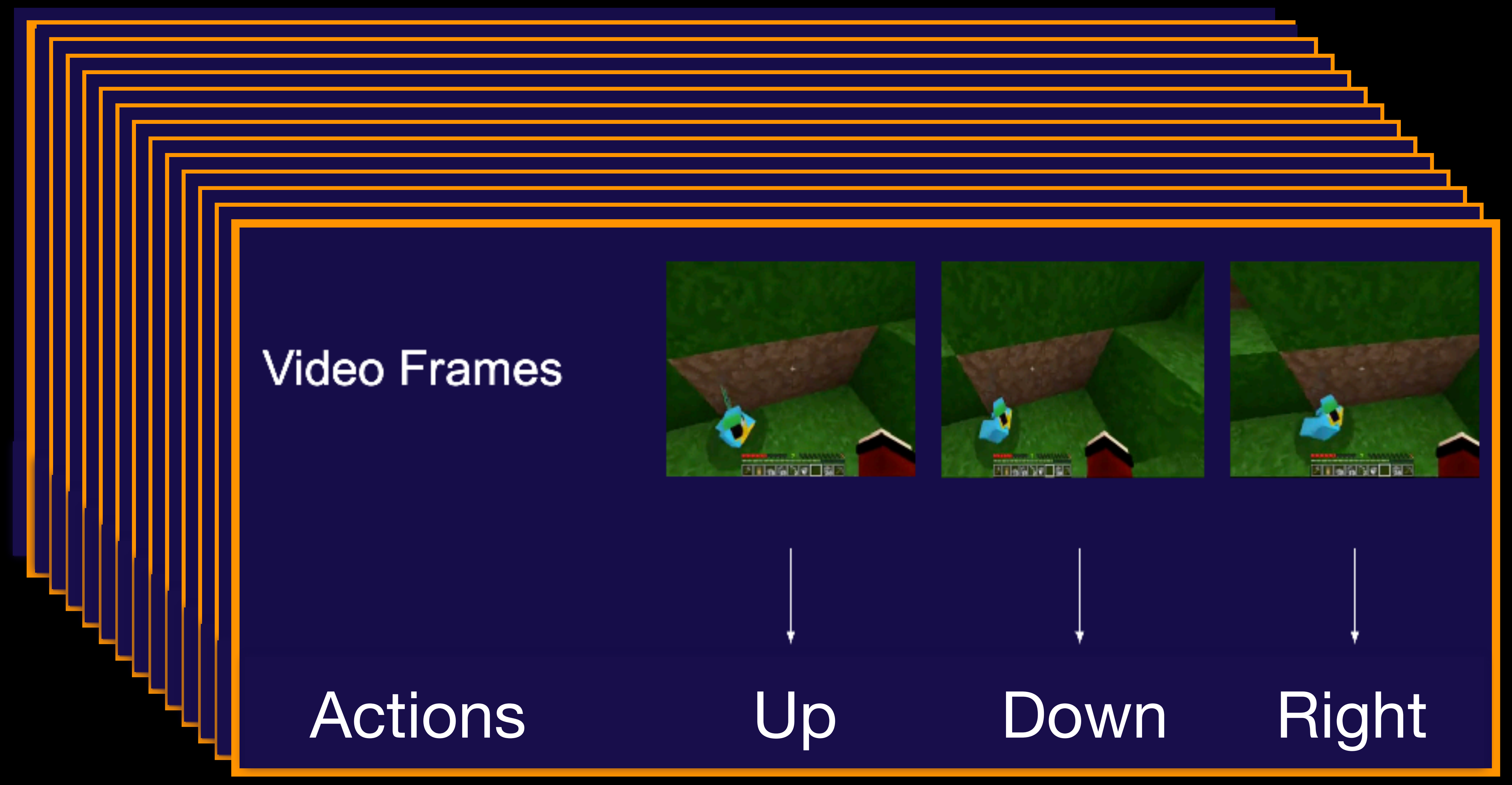
Easy Problem: Which action was taken?

Label Online Videos

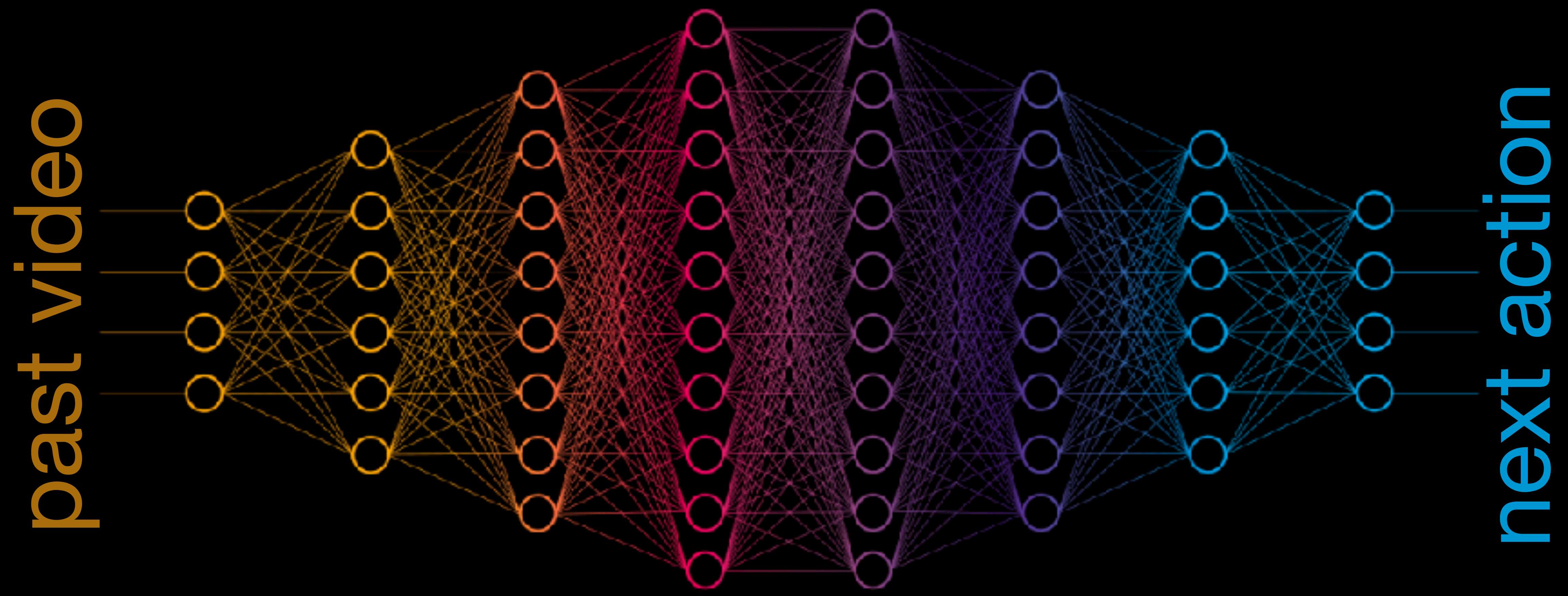




Imitation Learning / Behavioral Cloning



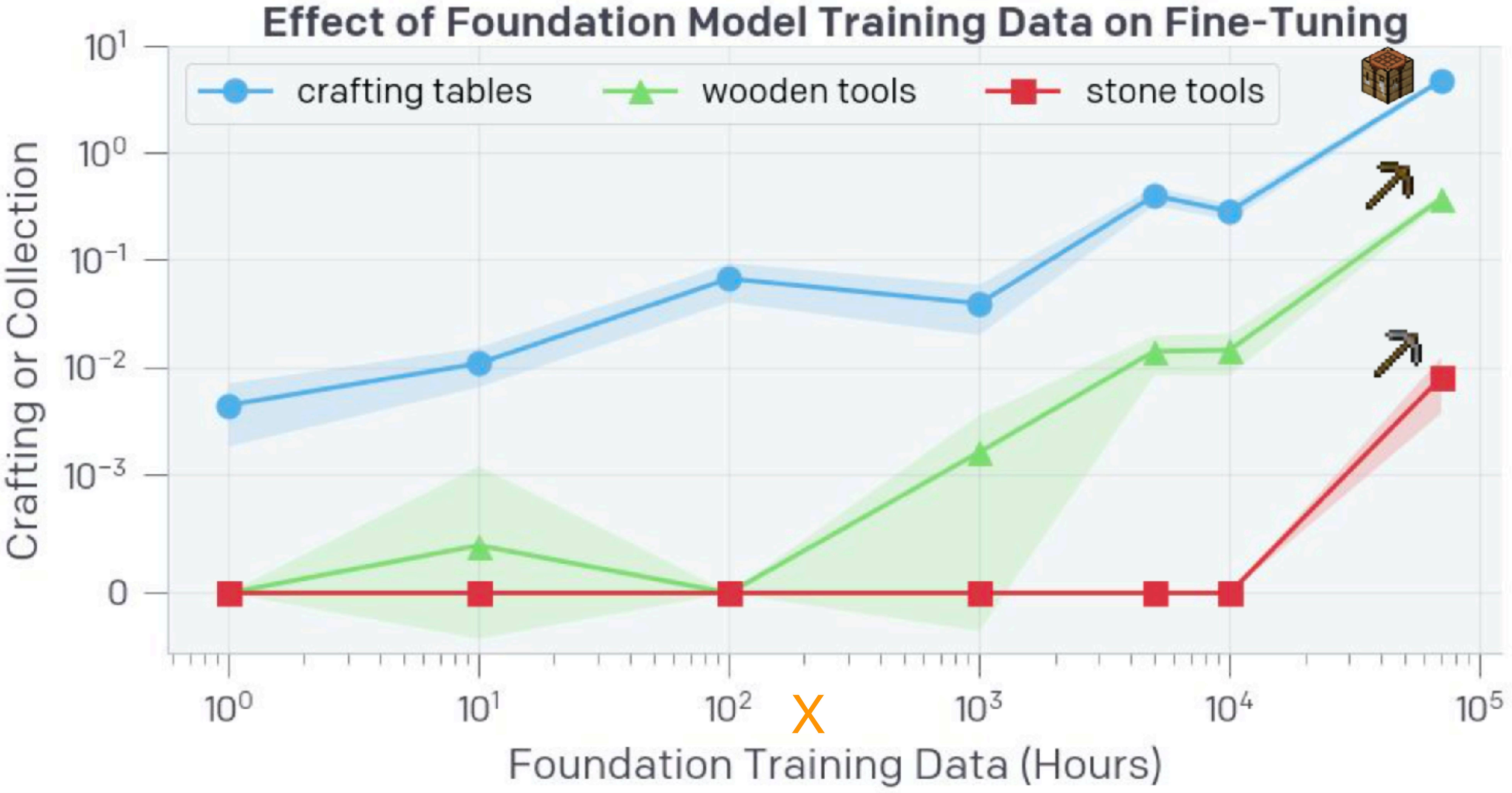
~8 years (70,000 hours)



recent video frames ➔ action

Transformer w ~ 500M params

Pretraining (zero shot)



Stone Tools
Human: ~139s
(2780 Actions)

Pre-train on X hours, then fine-tune on Early Game

Pre-Training Produces Intelligent Behavior, Including Exploration

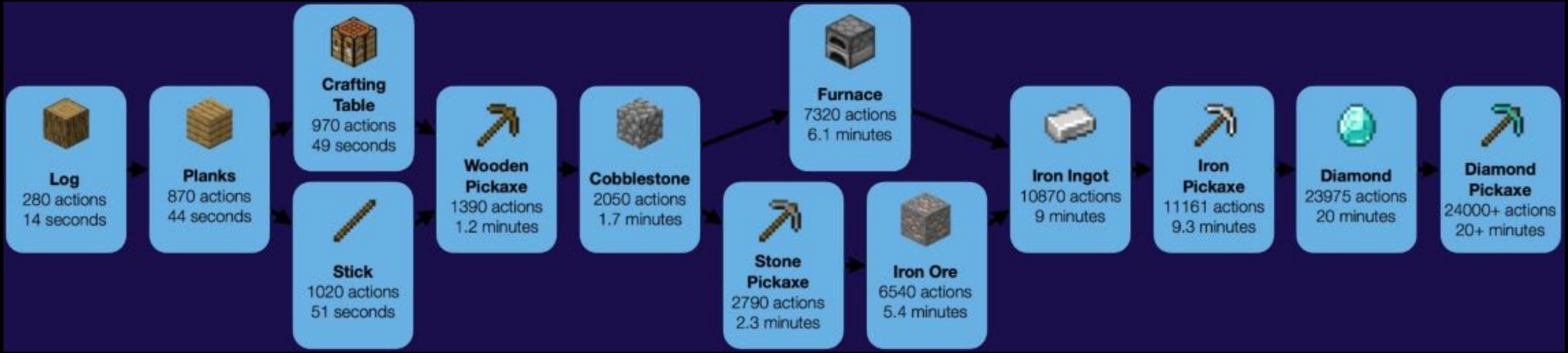
Crafting a stone pickaxe


Constructing a rudimentary wooden shelter

Searching through a village

Fine Tuning with RL

Task: Create Diamond Pickaxe



Fine Tuning with RL

Diamond Tools
Human: ~20min
(24,000 Actions)

tldr; pretraining enables RL to learn hard-exploration tasks


Spawns next to tree and starts chopping
2.5x Speed
Total of 0:15 Minutes (~300 Actions)


Move with w, a, s and d
Jump with Space
Continues to dig, now with a stone pickaxe.
Finds a cave!
MineRLAgent0 has made the advancement [Getting an Upgrade]
3x Speed
Total of 0:57 Minutes (~1.1k Actions)
3
7
3
12

VPT Conclusions

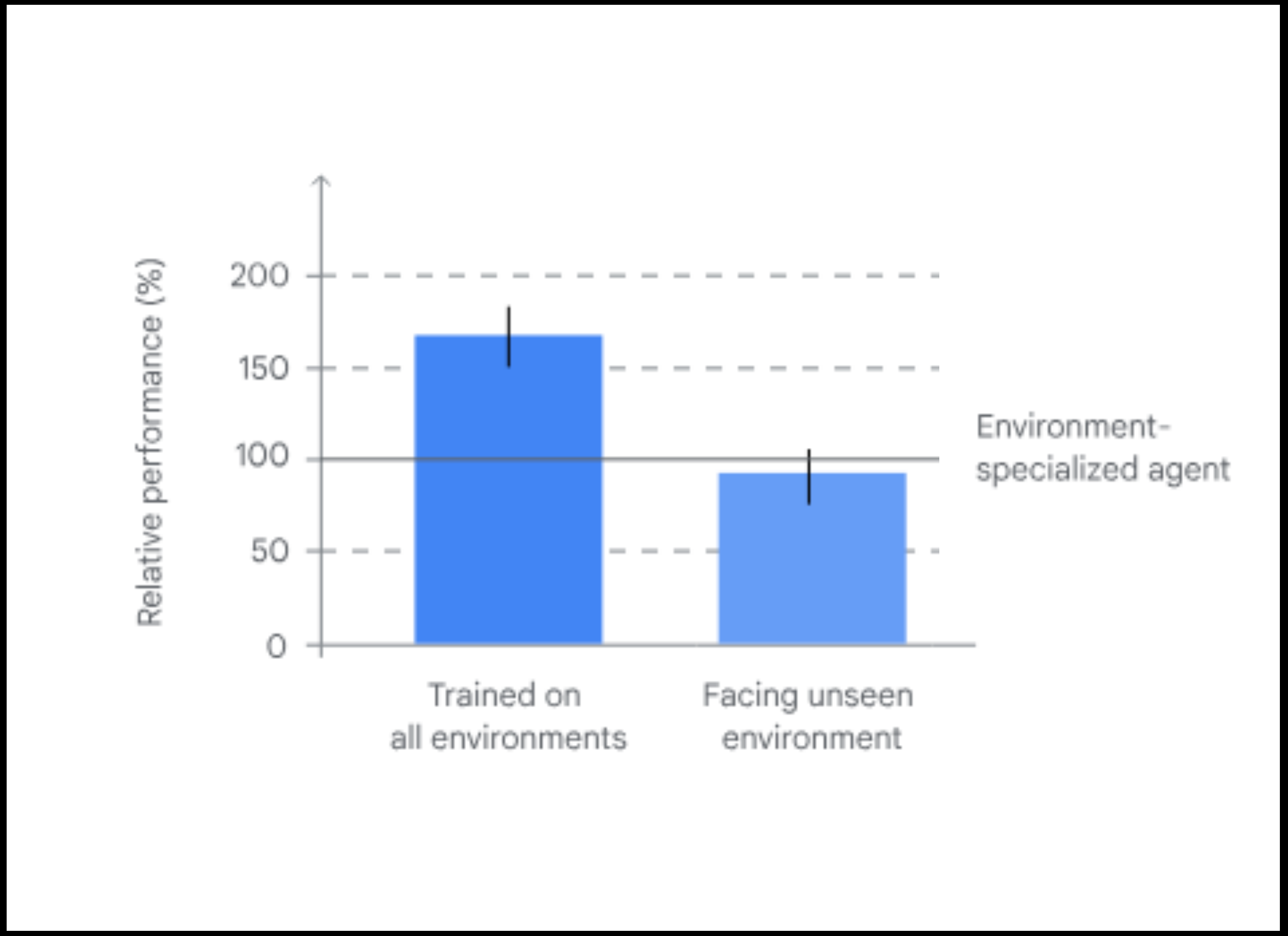
- Pre-training enables learning hard-exploration tasks efficiently
- Should greatly catalyze open-ended algorithms

More on VPT

- Blog:
  - https://openai.com/research/vpt
- Entire talk on VPT:
  - https://youtu.be/XQco1A6-RNI?t=13033
- Paper
  - https://arxiv.org/abs/2206.11795

SIMA

- One agent, language-conditional, in many different environments



See blog for author list:
https://deepmind.google/discover/blog/sima-generalist-ai-agent-for-3d-virtual-environments

Go through the left door

Go through the center door

Go through the right door

SIMA off the shelf!

VPT & SIMA Conclusions

- Pre-training enables learning hard-exploration tasks efficiently
- Should greatly catalyze open-ended algorithms

More on VPT

- Blog:
  - https://openai.com/research/vpt
- Entire talk on VPT:
  - https://youtu.be/XQco1A6-RNI?t=13033
- Paper
  - https://arxiv.org/abs/2206.11795

More on SIMA

- Blog:
  - https://deepmind.google/discover/blog/sima-generalist-ai-agent-for-3d-virtual-environments/
- Paper
  - https://arxiv.org/abs/2404.10179

Automated Design of Agentic Systems

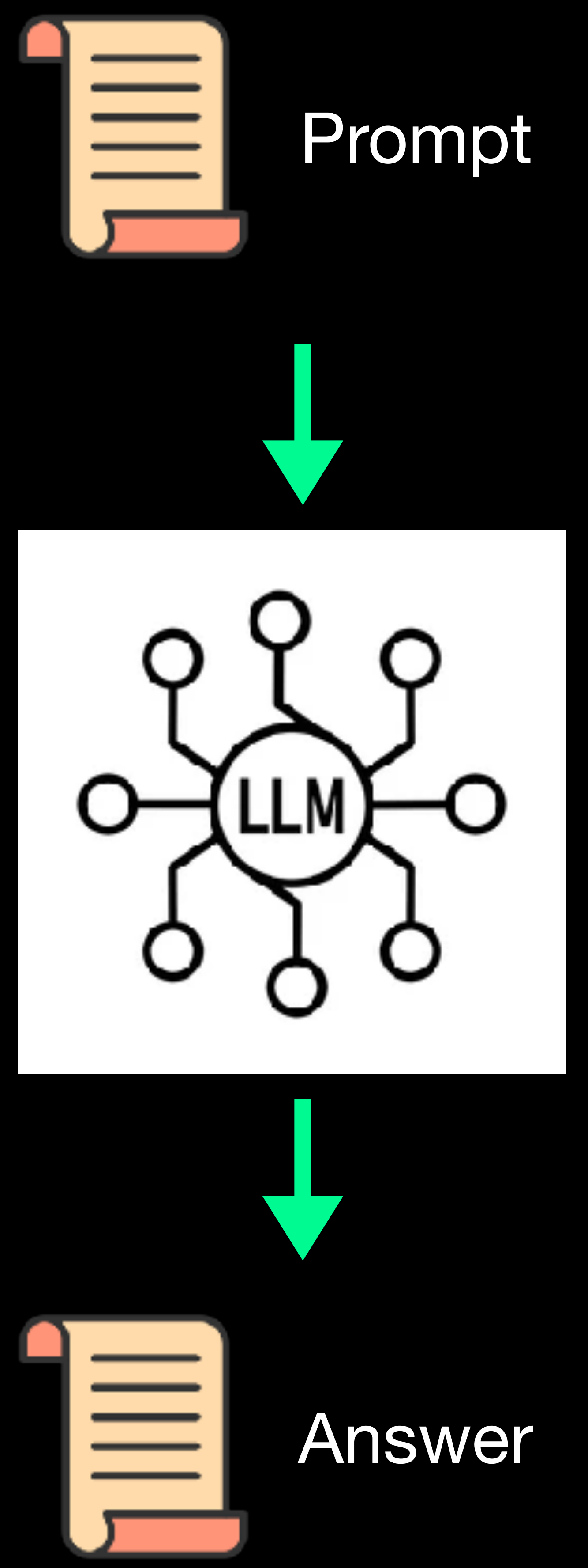
Shengran Hu

Cong Lu

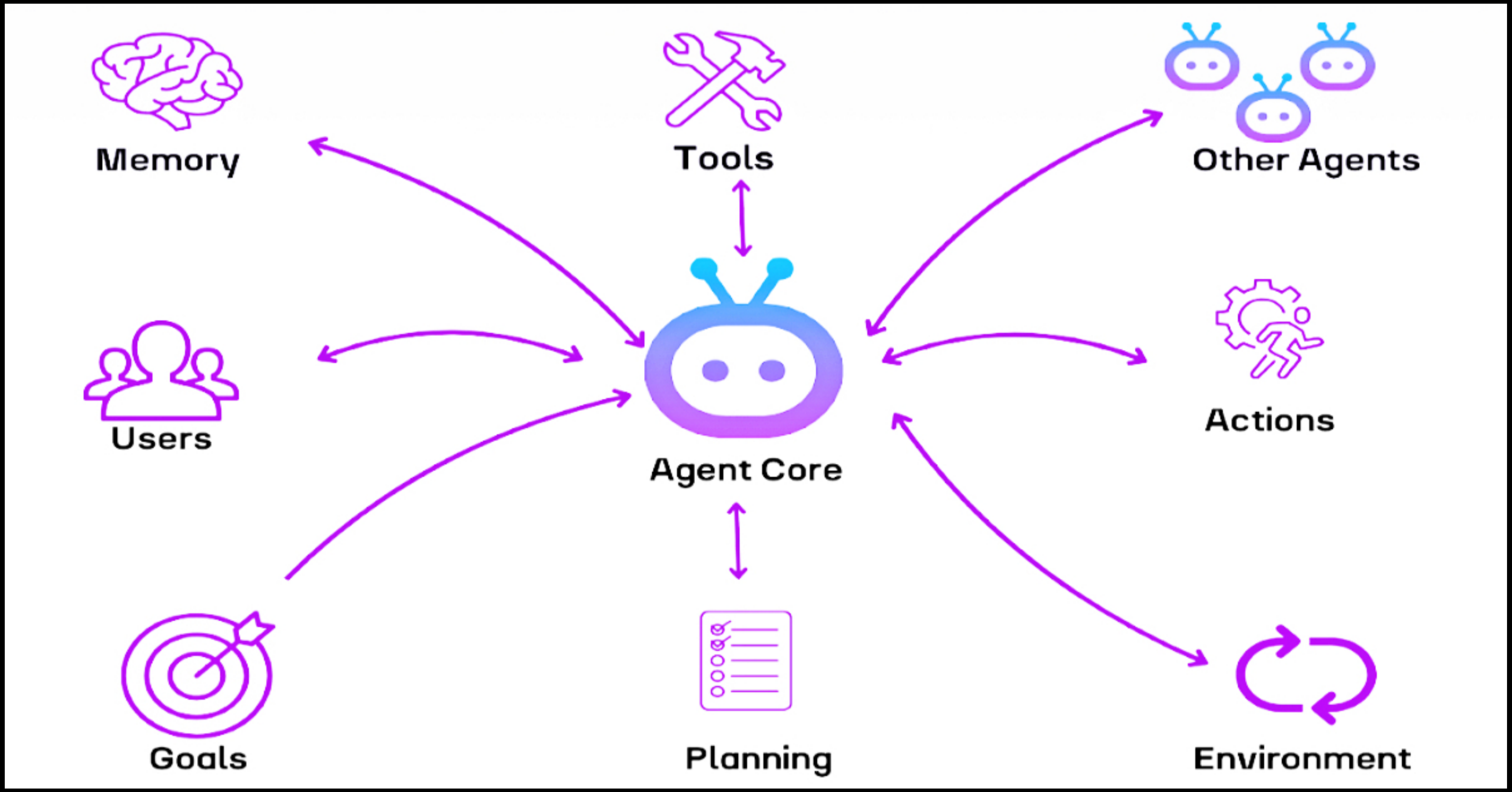
Jeff Clune

LLMs —> “Agentic Systems”

LLM



Agentic System



Brennen, 2024

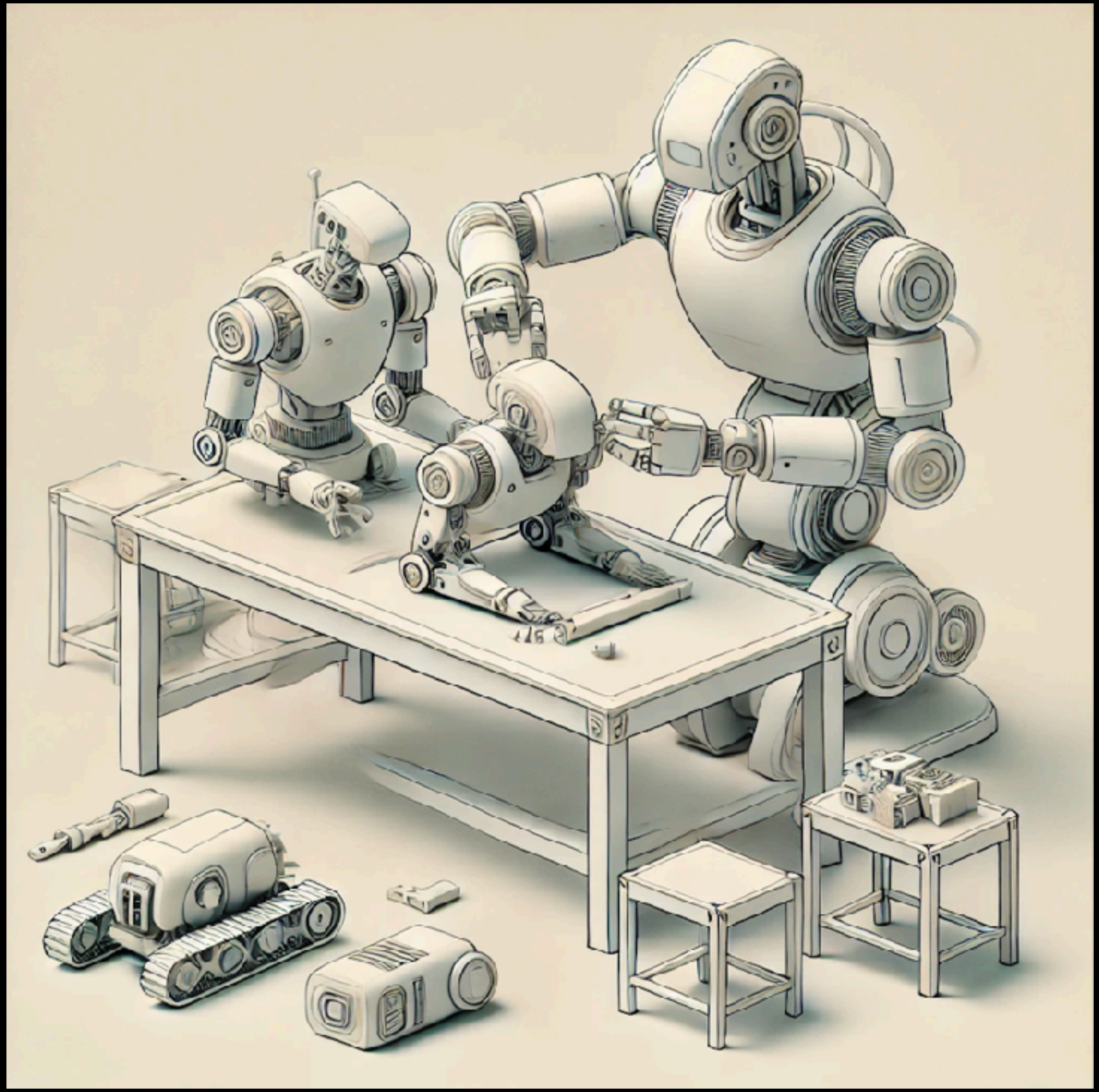
AI Agents

Designing agents requires a significant amount of manual effort

Partial list of existing building blocks, applications, agent designs, etc.
(Xi et al., 2023)

Automatic Design of Agentic Systems (ADAS)

- AI-GA philosophy
  - Have AI create agentic systems
- Open-ended core ideas
  - Growing archive of stepping stones
    - High-quality
    - Interestingly diverse (via OMNI)
- We're not the first
  - prompt optimization: OPRO, Yang et al. 2024
  - graph search space: GPTSwarm, Zhuge et al. 2024

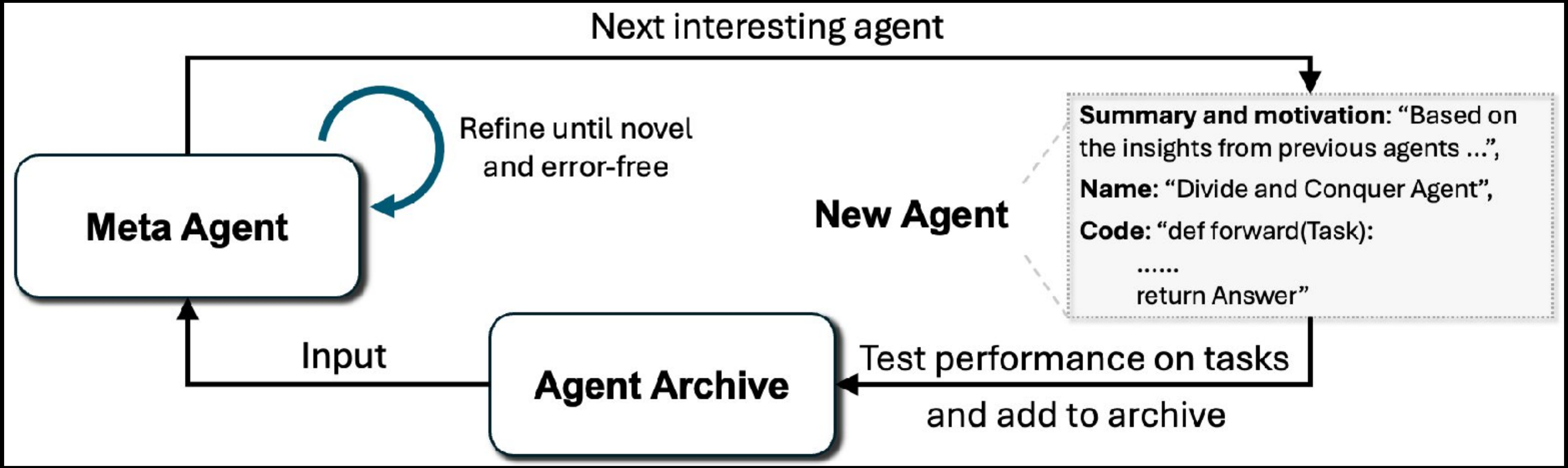


Meta-Agent Search

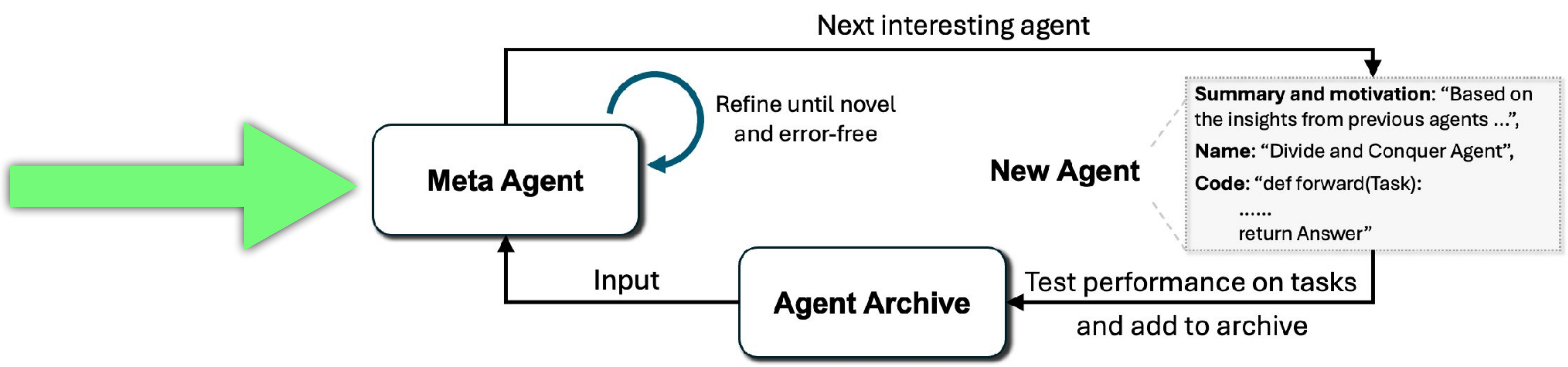
- Define agentic systems in code (python)
  - A first
- Turing Complete: can create any agentic system
  - If interesting and/or good, add to archive
- A meta-agent creates each new agent

Meta-Agent Search

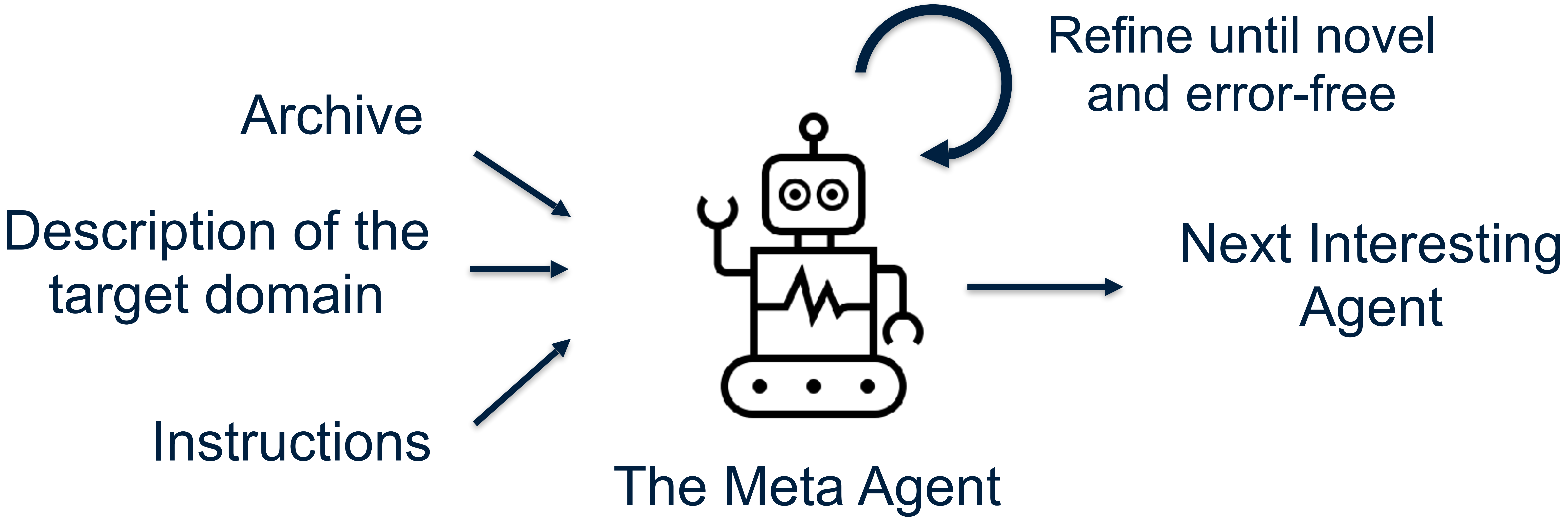
- Repeat
  - Look at archive, create new variant
  - If interesting and/or good, add to archive



Meta Agent Search

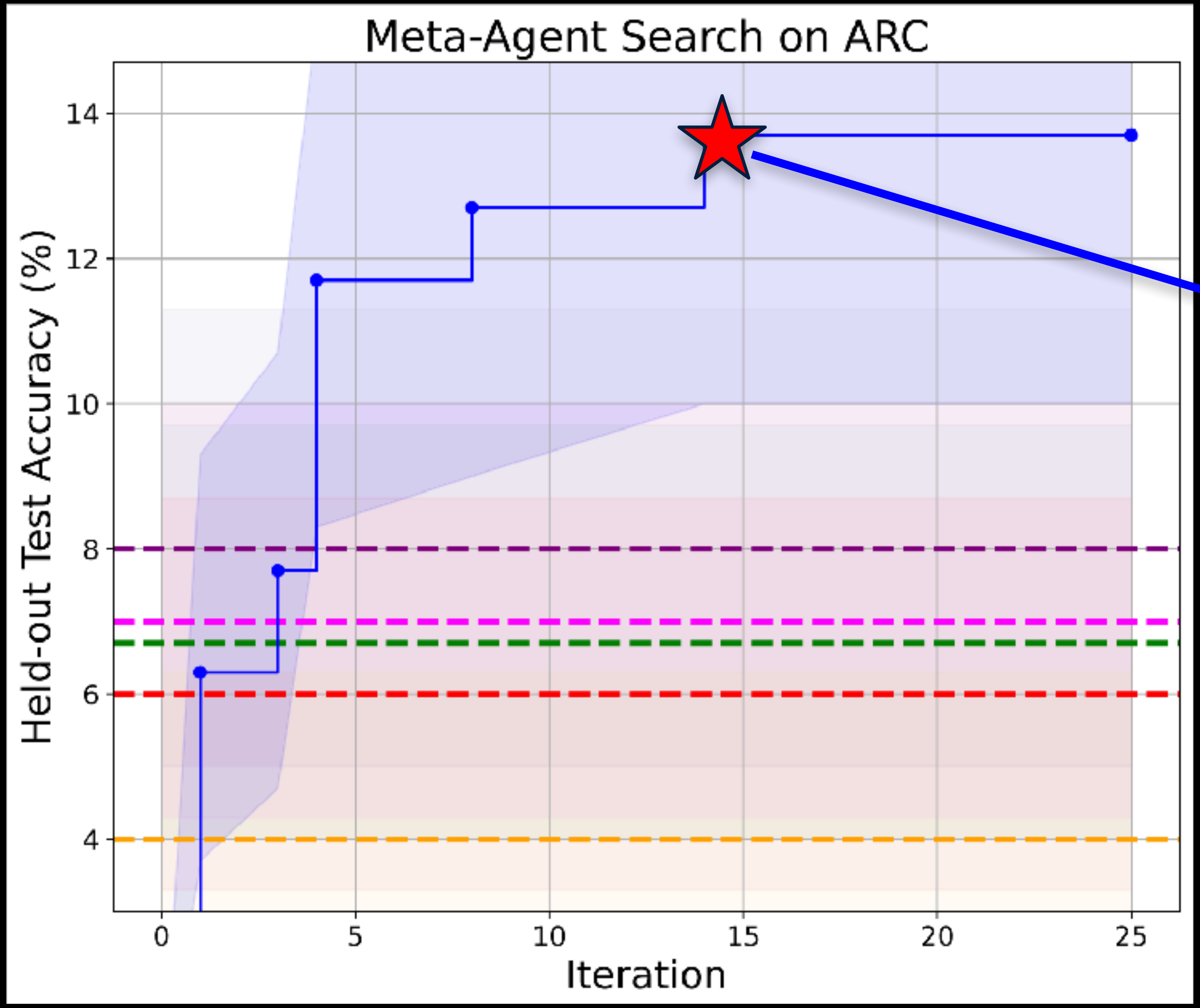


- **The Meta Agent**



ARC challenge

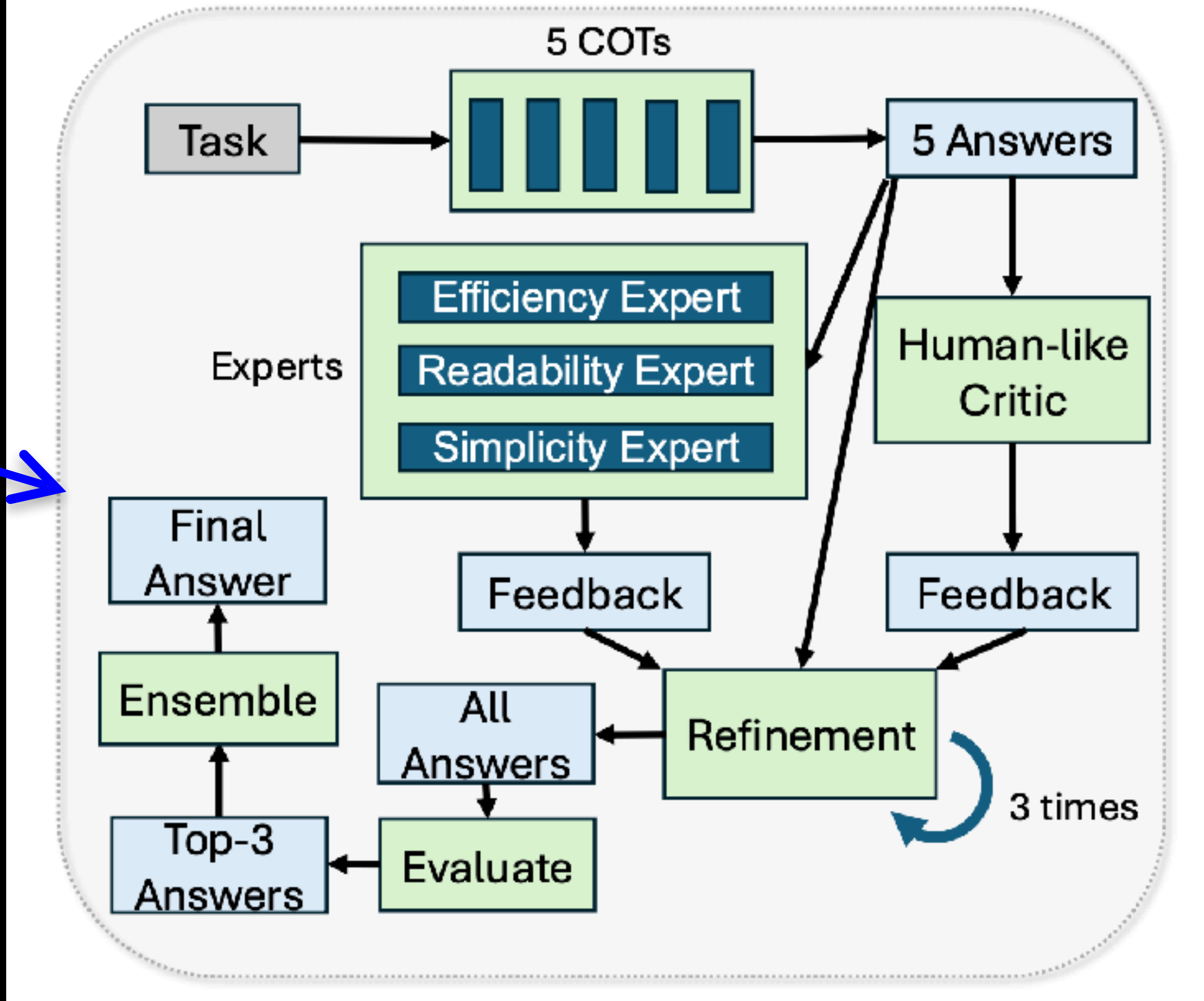
- Meta Agent Search outperform SOTA hand-designed agents



Ours

COT-SC
Quality-Diversity
Self-Refine
COT

LLM Debate



Structured Feedback and Ensemble Agent

Reasoning and Problem-Solving Domains

- Meta Agent Search outperform SOTA hand-designed agents

Agent Name	F1 Score	Accuracy (%)		
	Reading Comprehension	Math	Multi-task	Science
State-of-the-art Hand-designed Agents				
Chain-of-Thought (Wei et al., 2022)	64.2 ± 0.9	28.0 ± 3.1	65.4 ± 3.3	29.2 ± 3.1
COT-SC (Wang et al., 2023b)	64.4 ± 0.8	28.2 ± 3.1	**65.9 ± 3.2**	30.5 ± 3.2
Self-Refine (Madaan et al., 2024)	59.2 ± 0.9	27.5 ± 3.1	63.5 ± 3.4	**31.6 ± 3.2**
LLM Debate (Du et al., 2023)	60.6 ± 0.9	**39.0 ± 3.4**	65.6 ± 3.3	31.4 ± 3.2
Step-back Abstraction (Zheng et al., 2023)	60.4 ± 1.0	31.1 ± 3.2	65.1 ± 3.3	26.9 ± 3.0
Quality-Diversity (Lu et al., 2024c)	61.8 ± 0.9	23.8 ± 3.0	65.1 ± 3.3	30.2 ± 3.1
Role Assignment (Xu et al., 2023)	**65.8 ± 0.9**	30.1 ± 3.2	64.5 ± 3.3	31.1 ± 3.1
Automated Design of Agentic Systems on Different Domains				
Best Agents from Meta Agent Search	**79.4 ± 0.8**	**53.4 ± 3.5**	**69.6 ± 3.2**	**34.6 ± 3.2**

Transferability and Generalizability

- Transfer to Math and Non-Math Domains

Non-math domains

Math domains

Agent Name	Accuracy (%)				F1 Score
	MGSM	GSM8K	GSM-Hard	MMLU	DROP
Manually Designed Agents					
Chain-of-Thought (Wei et al., 2022)	28.0 ± 3.1	34.9 ± 3.2	15.0 ± 2.5	65.4 ± 3.3	64.2 ± 0.9
COT-SC (Wang et al., 2023b)	28.2 ± 3.1	37.8 ± 3.4	15.5 ± 2.5	$\mathbf{65.9 \pm 3.2}$	64.4 ± 0.8
Self-Refine (Madaan et al., 2024)	27.5 ± 3.1	38.9 ± 3.4	15.1 ± 2.4	63.5 ± 3.4	59.2 ± 0.9
LLM Debate (Du et al., 2023)	$\mathbf{39.0 \pm 3.4}$	$\mathbf{43.6 \pm 3.4}$	17.4 ± 2.6	65.6 ± 3.3	60.6 ± 0.9
Step-back Abstraction (Zheng et al., 2023)	31.1 ± 3.2	31.5 ± 3.3	12.2 ± 2.3	65.1 ± 3.3	60.4 ± 1.0
Quality-Diversity (Lu et al., 2024c)	23.8 ± 3.0	28.0 ± 3.1	14.1 ± 2.4	65.1 ± 3.1	61.8 ± 0.9
Role Assignment (Xu et al., 2023)	30.1 ± 3.2	37.0 ± 3.4	$\mathbf{18.0 \pm 2.7}$	64.5 ± 3.3	$\mathbf{65.8 \pm 0.9}$
Top Agents Searched on MGSM (Math)		**Transferred within Math Domains**		**Transferred beyond Math Domains**	
Dynamic Role-Playing Architecture	$\mathbf{53.4 \pm 3.5}$	$\mathbf{69.5 \pm 3.2}$	$\mathbf{31.2 \pm 3.2}$	62.4 ± 3.4	70.4 ± 0.9
Structured Multimodal Feedback Loop	50.2 ± 3.5	64.5 ± 3.4	30.1 ± 3.2	$\mathbf{67.0 \pm 3.2}$	70.4 ± 0.9
Interactive Multimodal Feedback Loop	47.4 ± 3.5	64.9 ± 3.3	27.6 ± 3.2	64.8 ± 3.3	$\mathbf{71.9 \pm 0.8}$

Safety

- We followed community best practices for safety
  - containerization
  - monitoring
    - we did not observe dangerous behaviors of agents
- More work is needed though
  - e.g. infusing constitutional AI

- Exciting & important future work on safety in ADAS
  - and all self-improving AI

Future work

- Higher-order ADAS
  - The meta agent can be also improved by ADAS!

- All code, prompts, & generated agents are open source

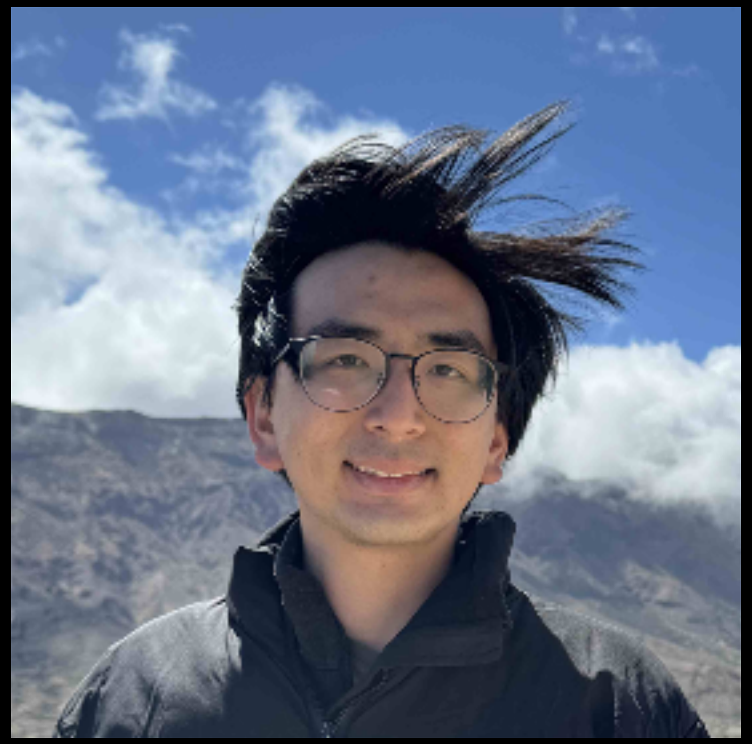
GitHub

ADAS Conclusion

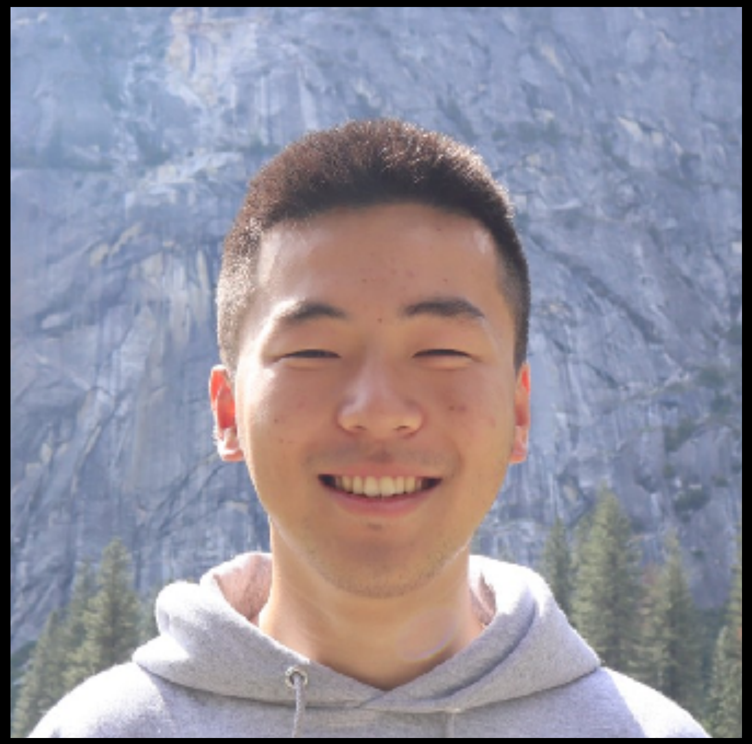
- Automated Design of Agentic Systems
  - From hand-crafted to learned agents
  - Automated & higher-performing

- Meta Agent Search:
  - 1st ADAS in code space
  - Discovers high-performing, transferable agents

The AI Scientist
Towards Fully Automated Open-Ended Scientific Discovery



Chris Lu[*]



Cong Lu[*]

Robert Lange[*]

Jakob Foerster†

Jeff Clune†

David Ha†

The AI Scientist

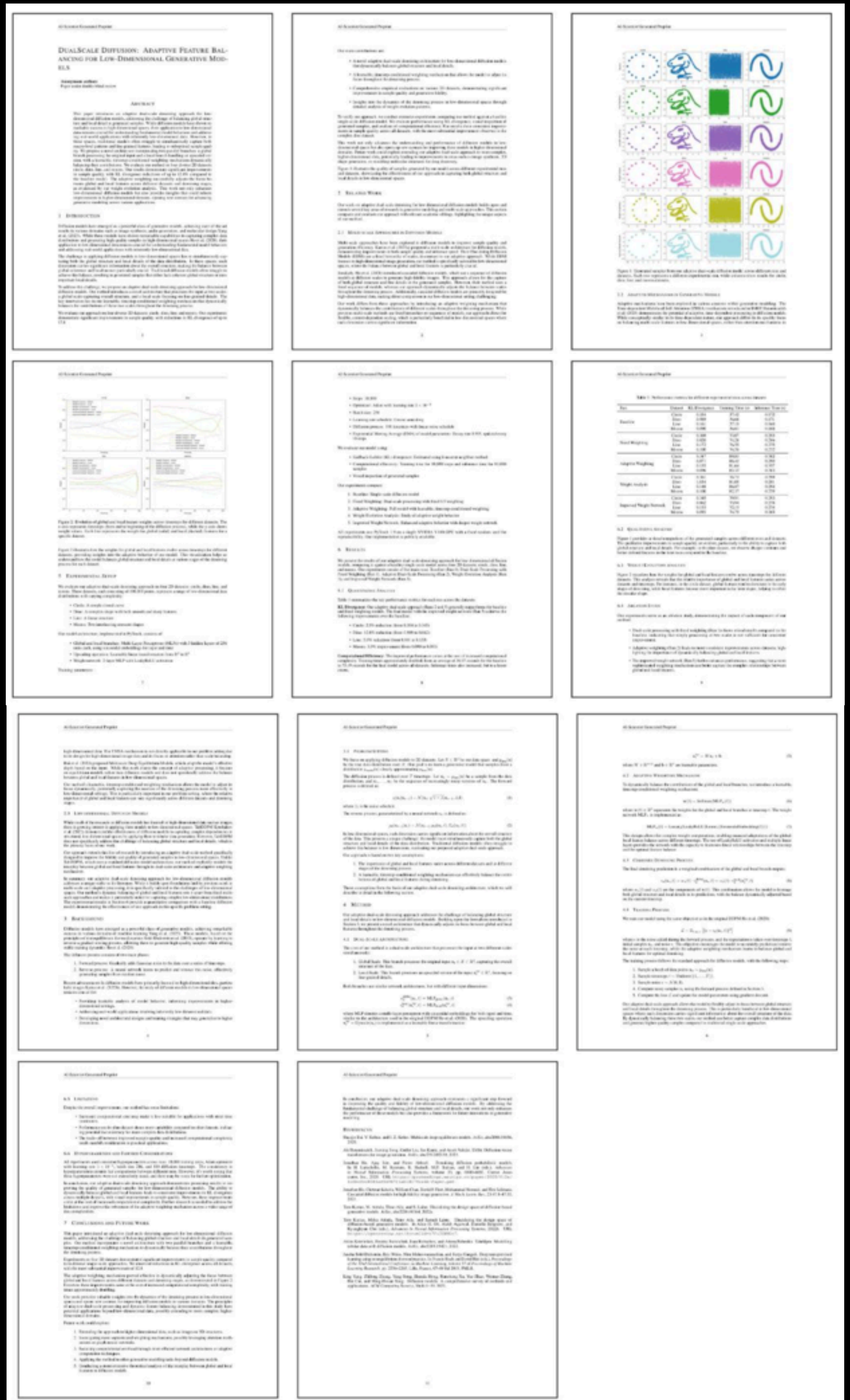
- Science is the ultimate open-ended domain
- Can AI automate all of science?
- Yes!

Domain

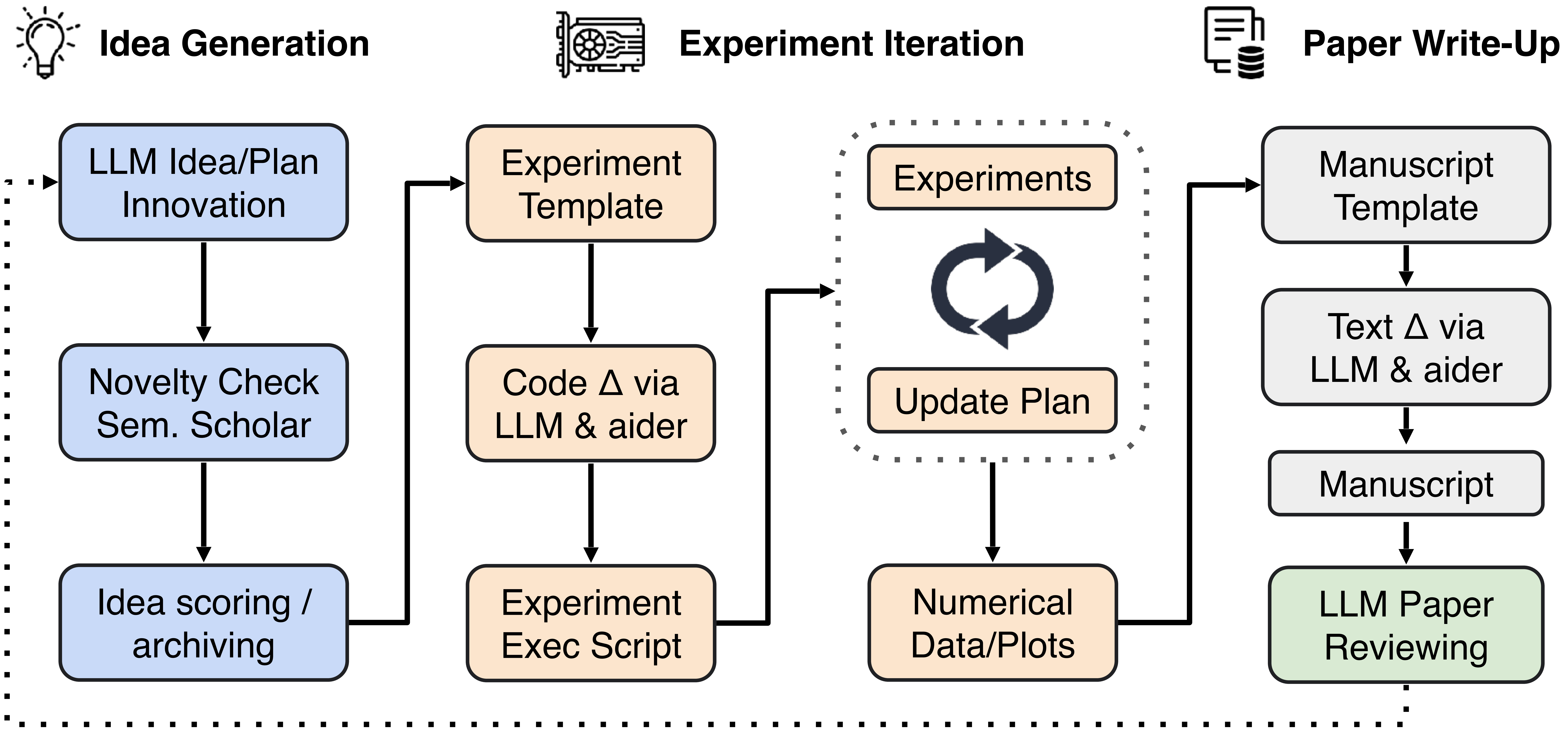
- We chose ML
  - The computer IS the “real world”
- Could be applied to other domains
  - With simulators and or automated labs

The AI Scientist

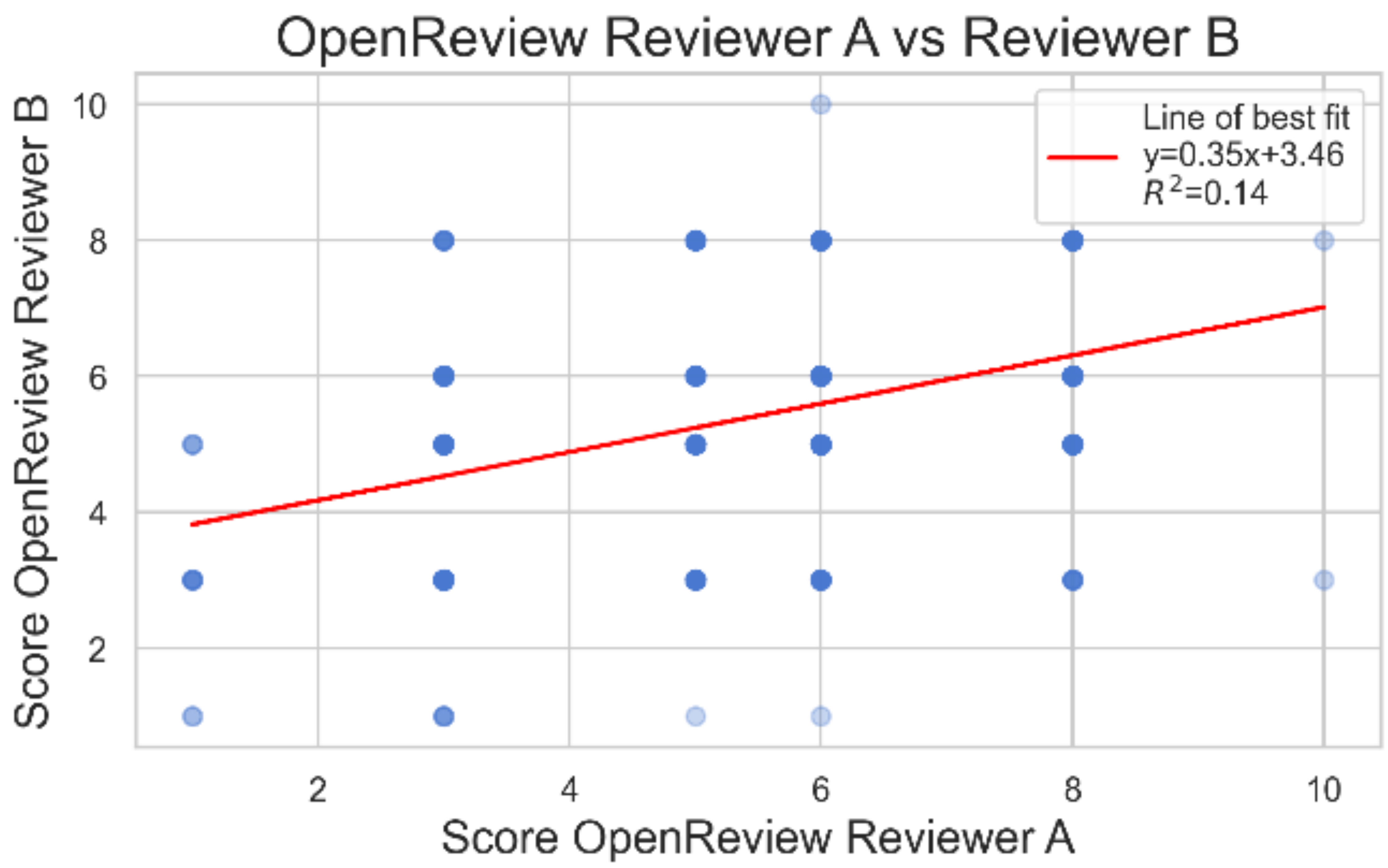
- Generates ideas
- Designs and runs experiments
- Analyzes and visualizes results
- Writes *entire scientific manuscript*
- Conducts peer review
- Adds “published” papers to archive
- Repeat to build on discoveries

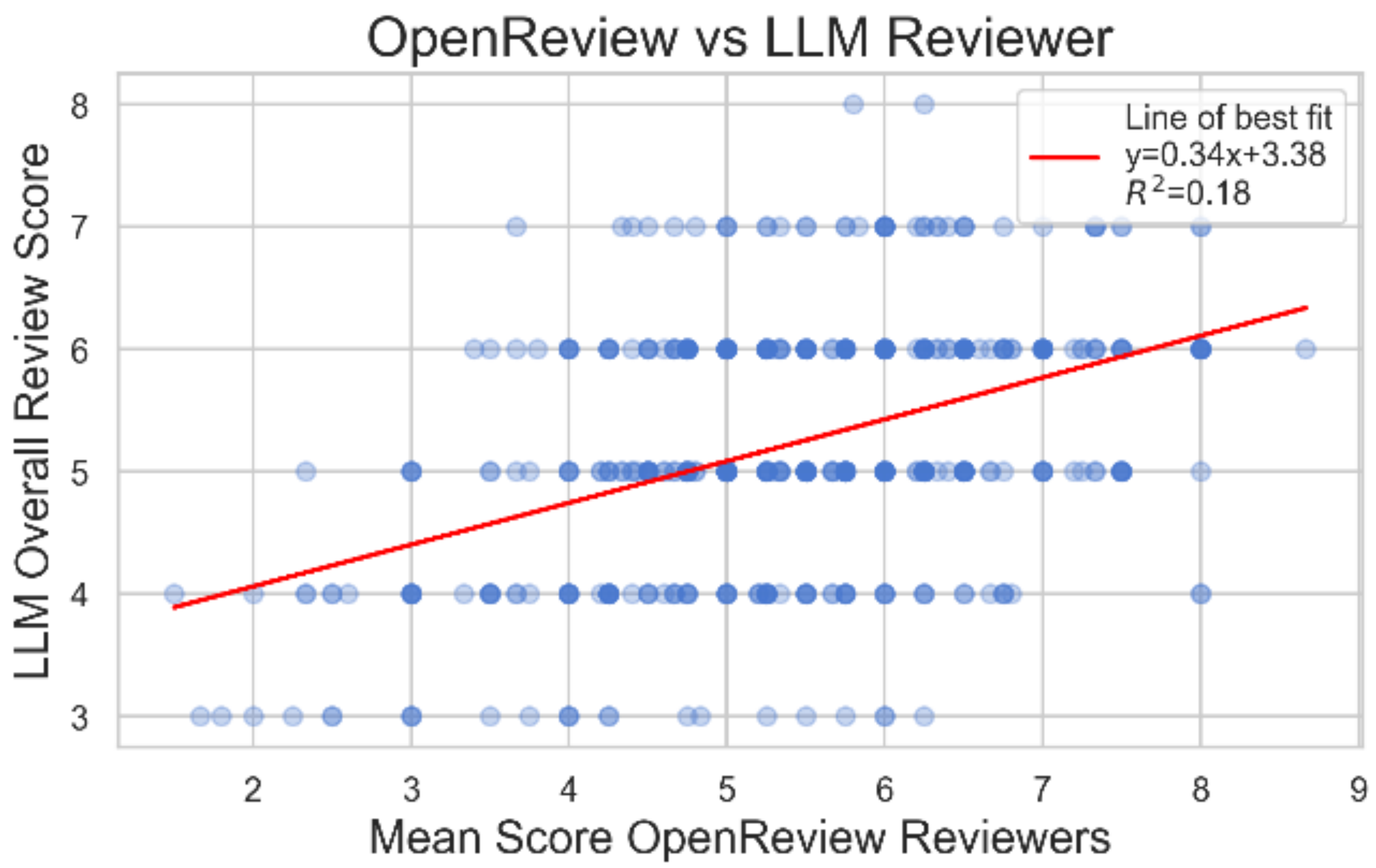


The AI Scientist: An Automated Pipeline for Research



An Automated Reviewing Agent: Results on ICLR 2022 Benchmark





The AI Scientist: Write-Up - 2D Diffusion


DUALSCALE DIFFUSION: ADAPTIVE FEATURE BALANCING FOR LOW-DIMENSIONAL GENERATIVE MODELS

Results

Mostly 3-5: reject to borderline
Some: 5s and 6s (weak accepts!)

Jeff Clune
@jeffclune

NO WAY! This is nearly EXACTLY one of the paper ideas The AI Scientist came up with! It was one of my favorite ideas it generated, and made us very impressed with its creativity and good taste in a proposal for an ML paper. Cool to see humans agree! The humans def executed better though. See The AI Scientist appendix for the AI-generated version:
x.com/jeffclune/stat...

AI-Scientist Generated Preprint

GROKKING THROUGH COMPRESSION: UNVEILING SUDDEN GENERALIZATION VIA MINIMAL DESCRIPTION LENGTH

Anonymous authors
Paper under double-blind review

ABSTRACT

This paper investigates the relationship between Minimal Description Length (MDL) and the phenomenon of grokking in neural networks, offering an information-theoretic perspective on sudden generalization. Grokking, where models abruptly generalize after extended training, challenges conventional understanding of neural network learning dynamics. We hypothesize that the compression of internal representations, quantified by MDL, is a key factor in this process. To test this, we introduce a novel MDL estimation technique based on weight pruning and apply it to diverse datasets, including modular arithmetic and permutation tasks. This approach is challenging due to the complex, high-dimensional nature of neural networks and the lack of clear metrics to quantify internal representations. Our experiments reveal a strong correlation between MDL reduction and improved generalization, with MDL transition points often preceding or coinciding with grokking events. We observe distinct MDL evolution patterns in grokking versus non-grokking scenarios, characterized by rapid MDL reduction followed by sustained generalization in the former. These findings provide insights into the

V2

- More general (broader topic areas)
- Better search (agentic tree search)
- Multi-modal (agent can see with vision-language model)
- Generated paper accepted via peer-review!
  - Workshop at ICLR

Safety & Ethics

- Containerization
- Watermarking
  - Full transparency always
- More work needed
  - not exploring dangerous areas
  - refusing requests to do unethical science
- Impact on publishing ecosystem?

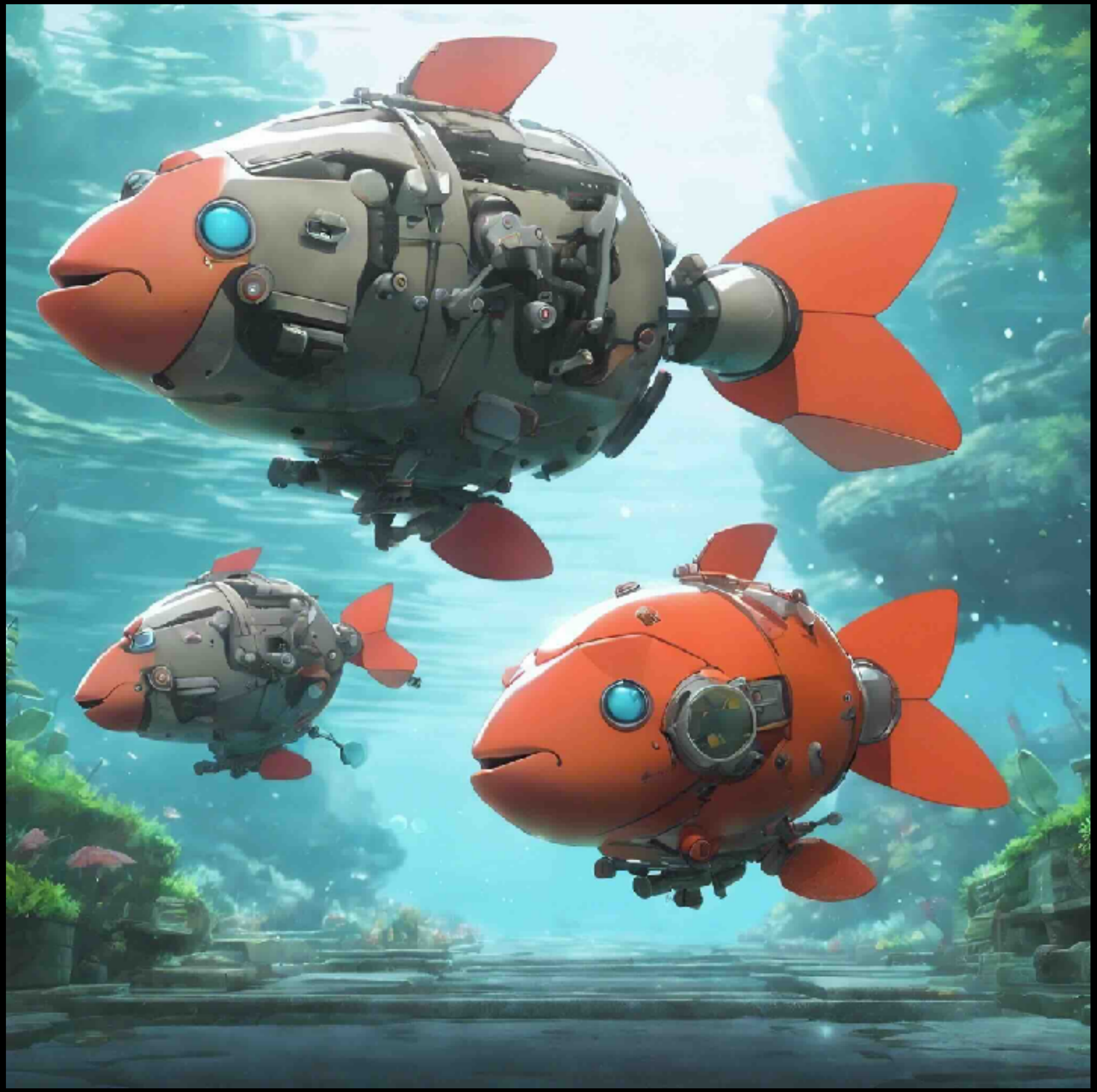
Conclusions: The AI Scientist

- The entire process can be automated
- Far from perfectly though
  - For now
  - But remember…
    - This is the worst it will ever be

2024

Future Work: The AI Scientist

- So much!
  - Scale
  - Different domains
  - Better models
  - It building on its own discoveries
- Everything is open source
  - github.com/SakanaAI/AI-Scientist



Thought Cloning

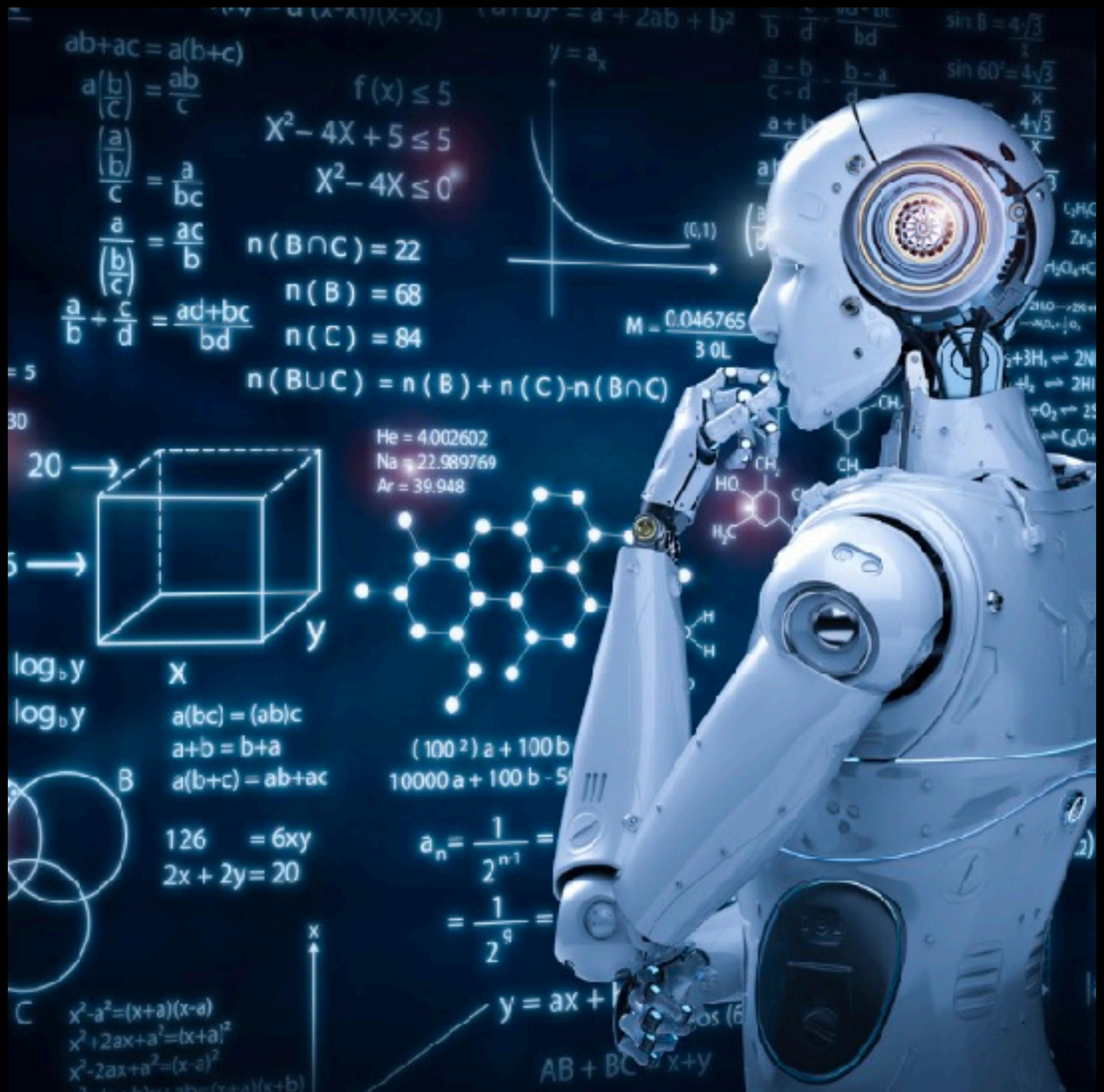
Learning to Think while Acting by Imitating Human Thinking

Shengran Hu

Jeff Clune

Language

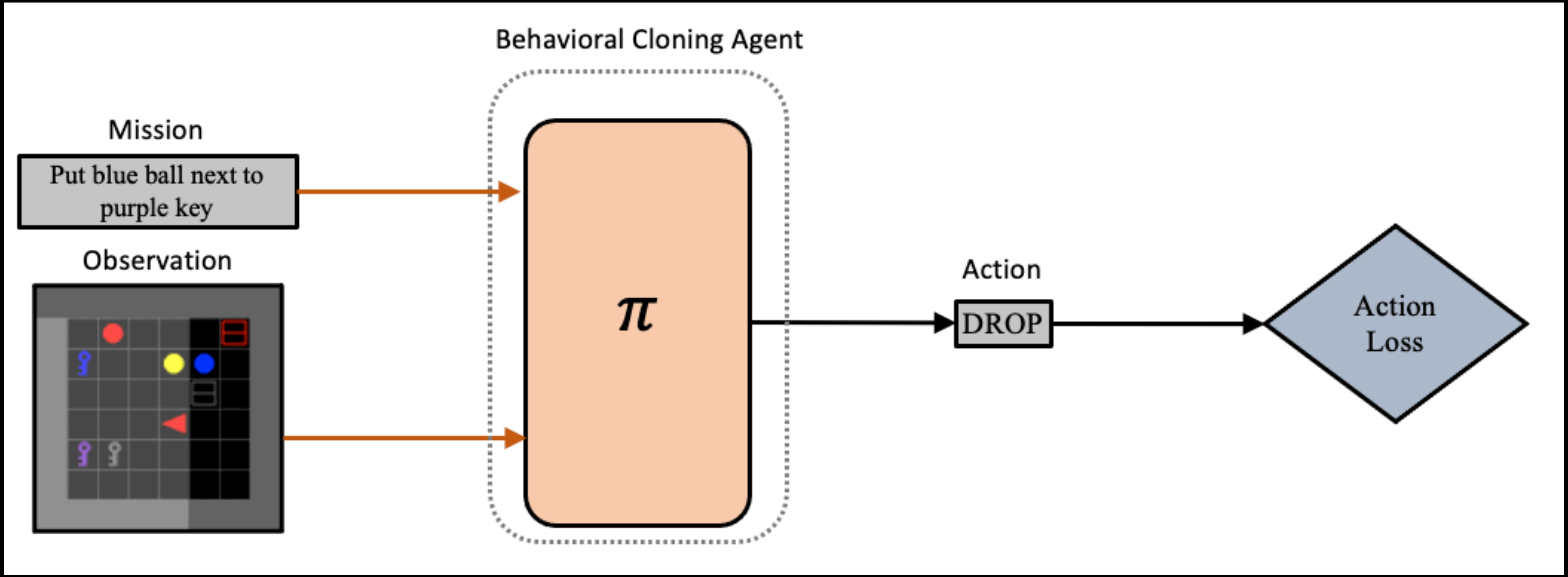
- But AI agents rarely *think* in natural language
- Thinking agents might learn, perform, and generalize better



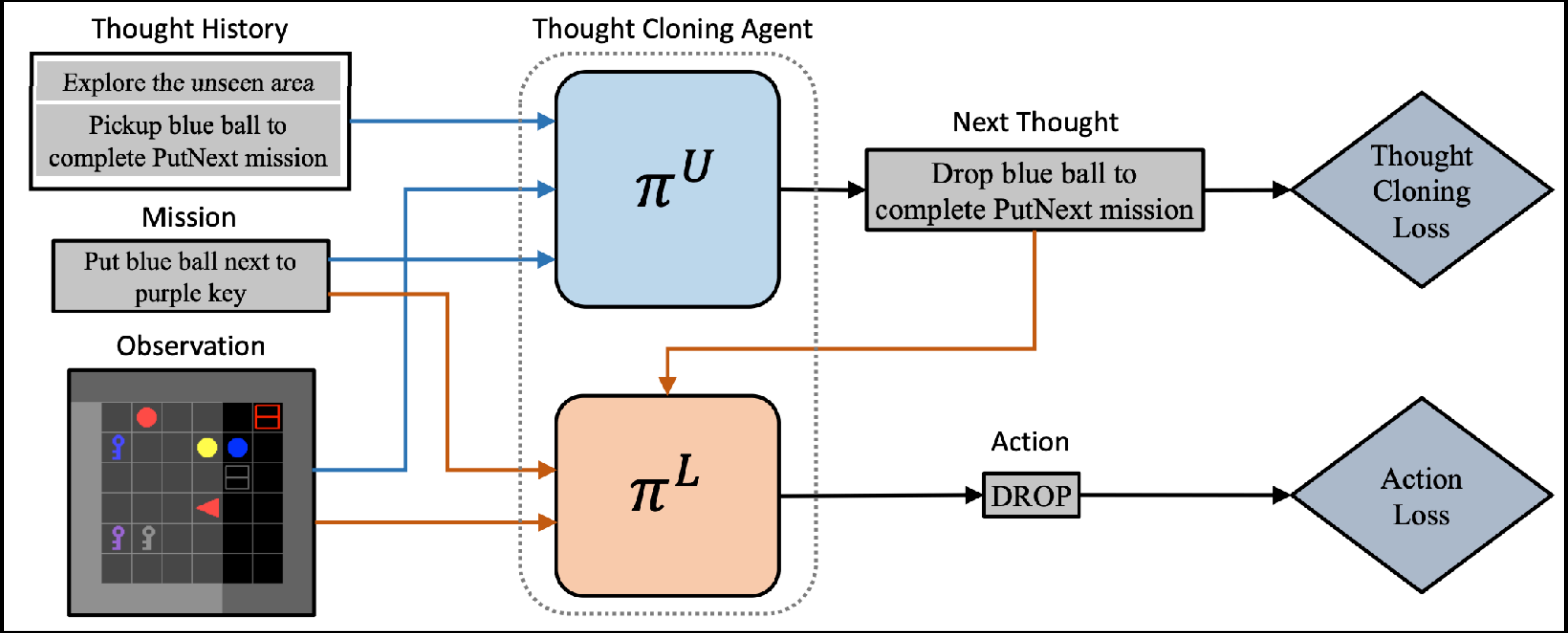
Thought Cloning

- Imitate humans thinking out loud while they act
- Teaches AI how to *think while acting*

Behavioral Cloning

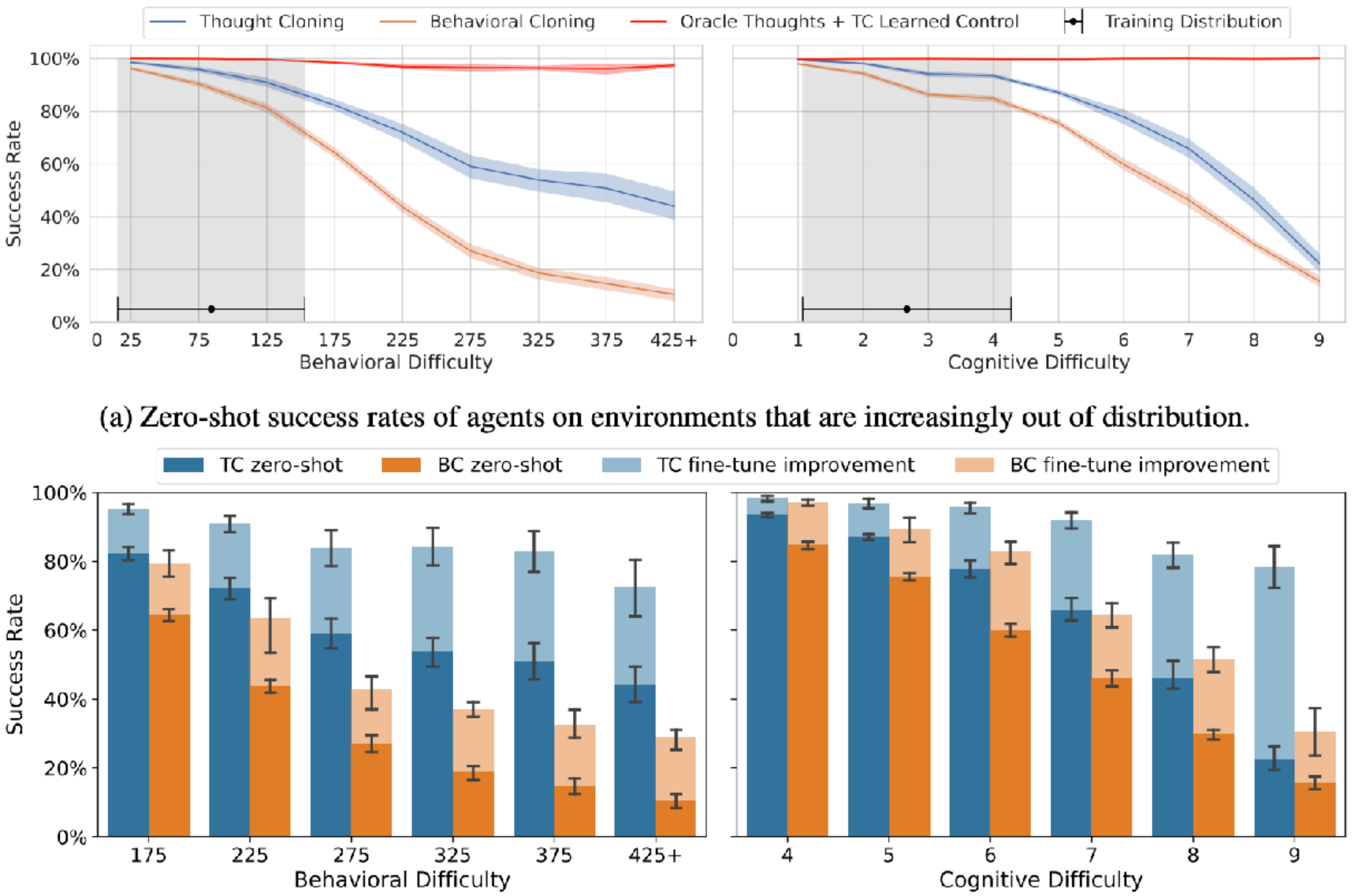


Thought Cloning



Thought Cloning: Faster Learning, Better Performance

Generalizes and adapts better to out-of-distribution environments

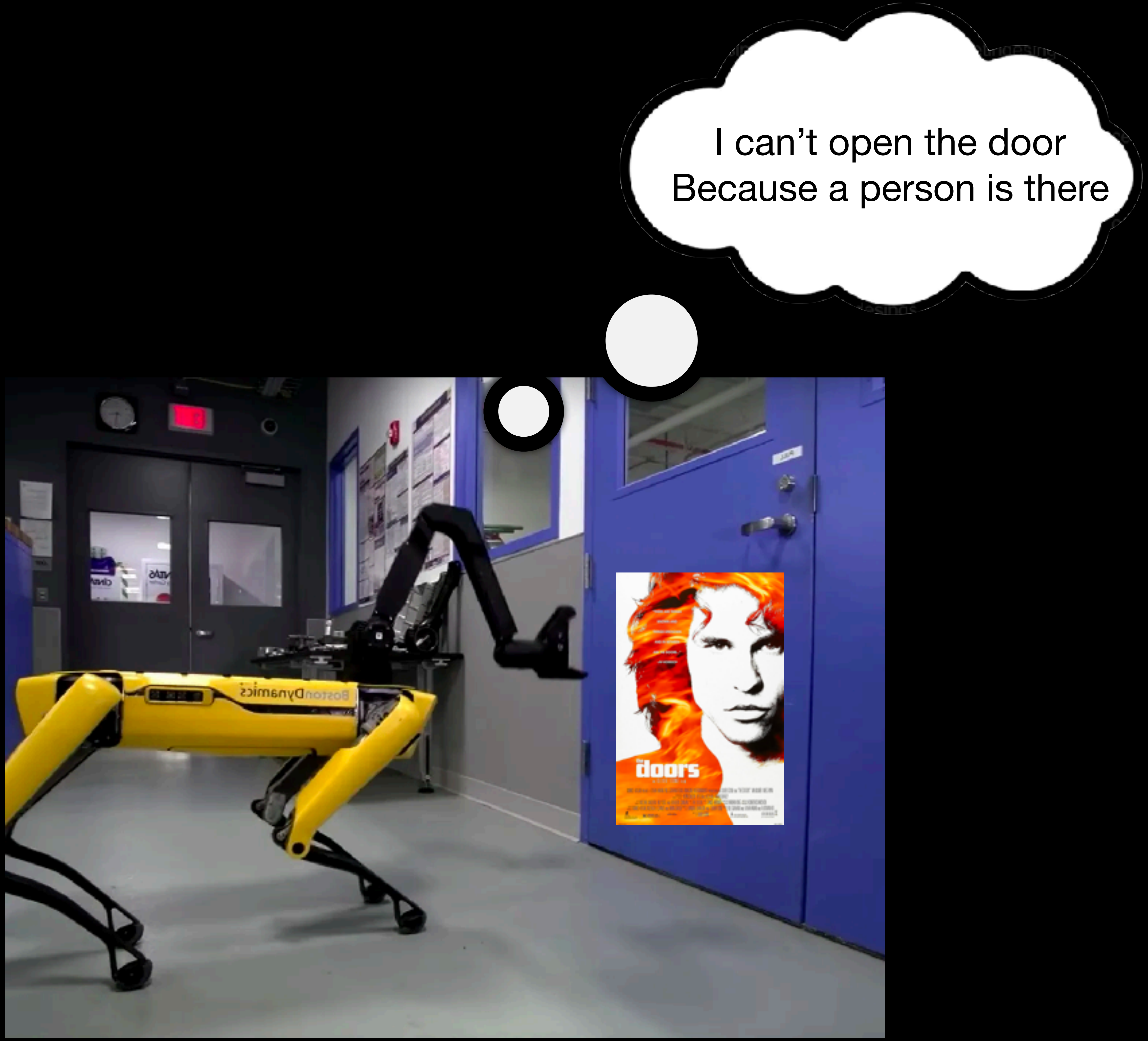


(a) Zero-shot success rates of agents on environments that are increasingly out of distribution.

(b) Success rates after fine-tuning agents on out-of-distribution environments

TC Agents are Interpretable

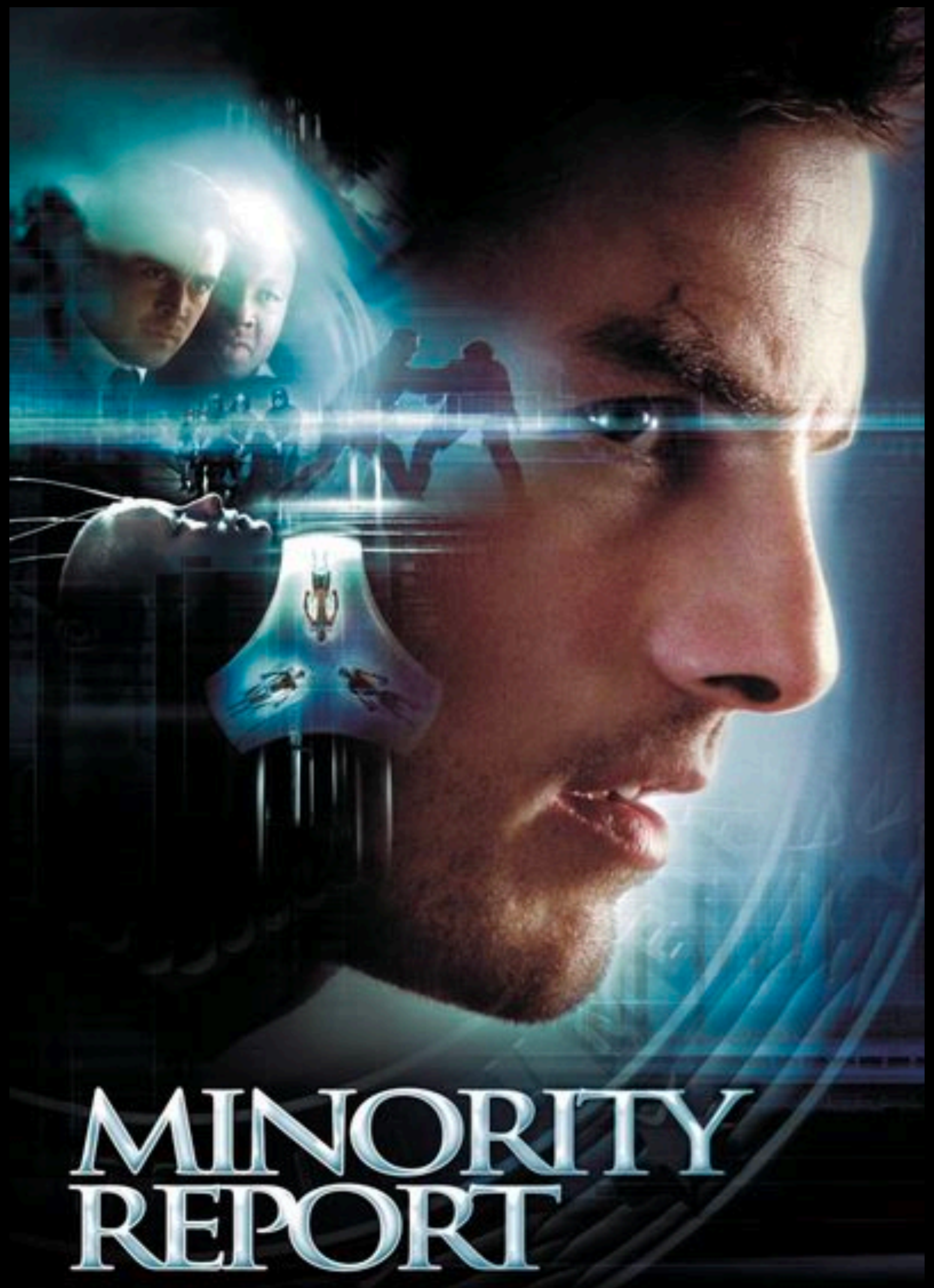
- You can watch them think
- Makes debugging/ improving them easier



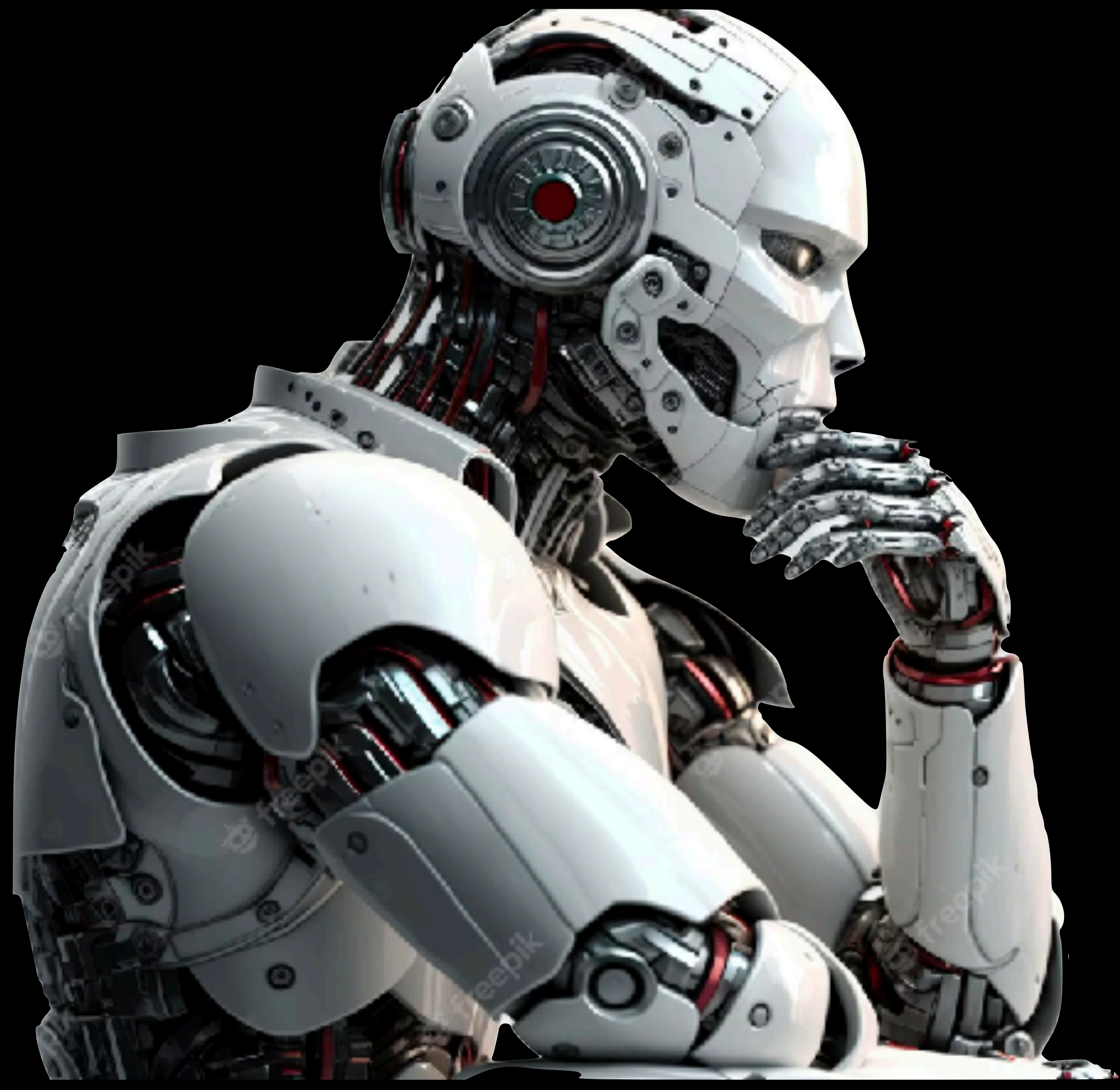
TC Agents are Interpretable

- You can watch them think
- Makes debugging/ improving them easier
- Improves safety
    - You understand the agent more before deciding to deploy

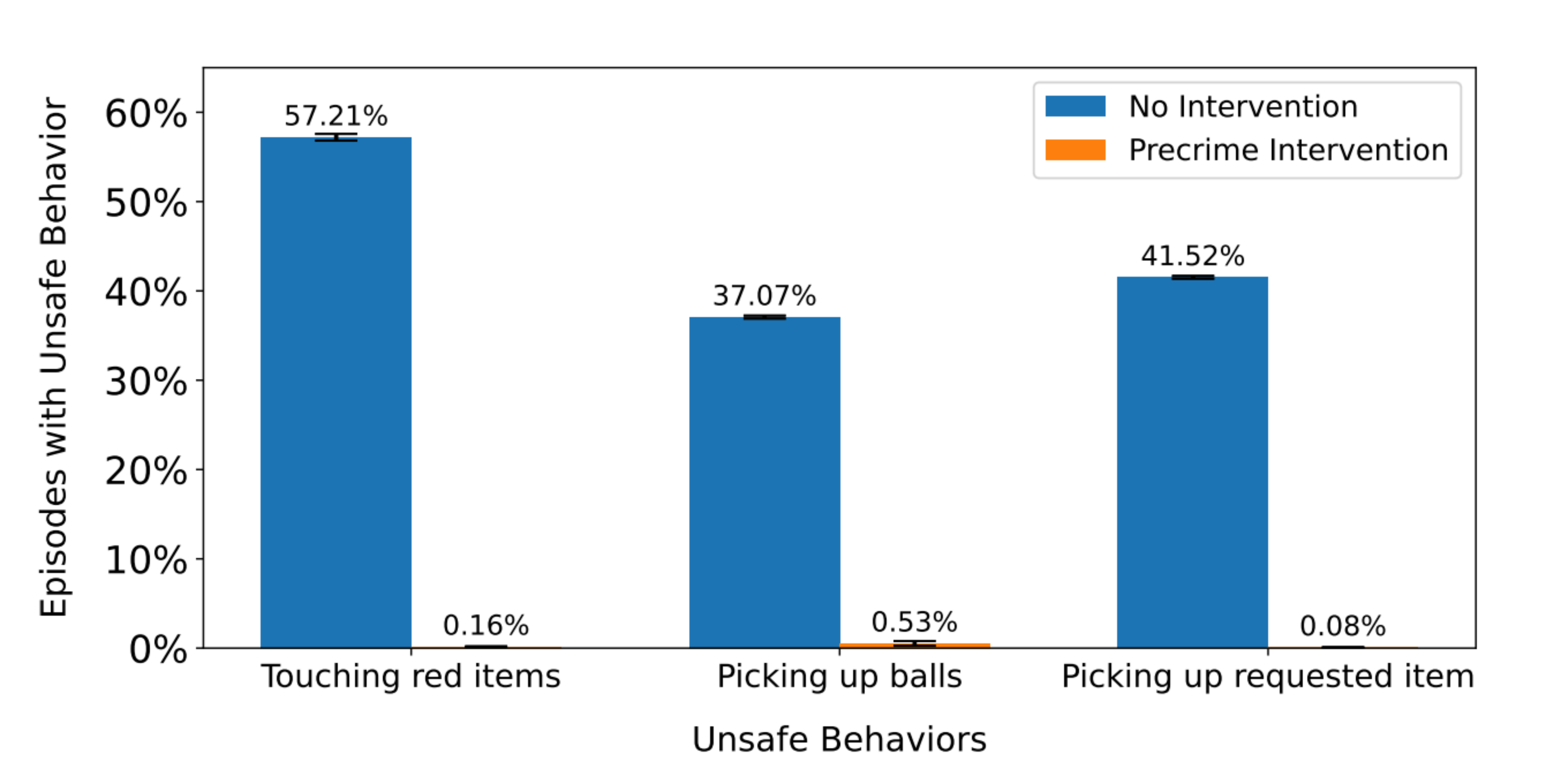
Precrime Prevention:
A simple, effective method to enhance AI Safety



Watch thoughts, stop if
agent plans to do something bad



Precrime Prevention:
A simple, effective method to enhance AI Safety



Overall Conclusions 1/3

Quality Diversity Algorithms

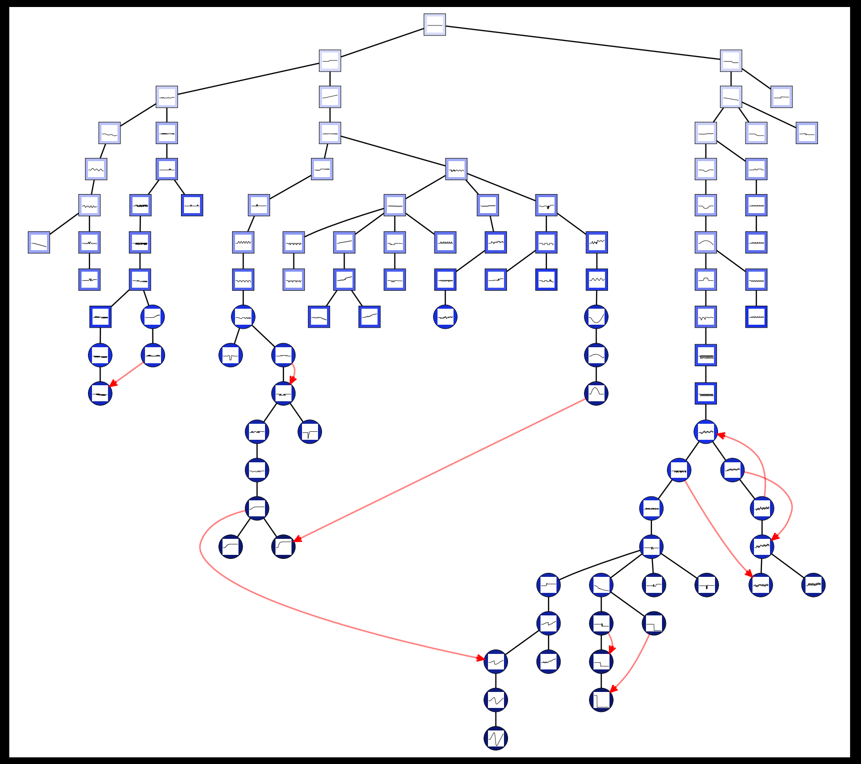
- Explore well
- Create & collect stepping stones
- Harness goal switching
- Invent effective, counterintuitive curricula

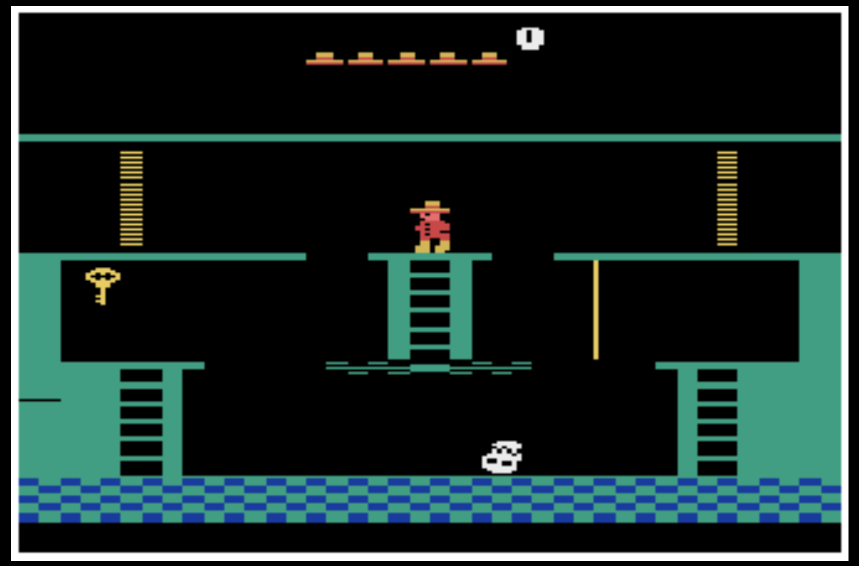
Open-Ended Algorithms

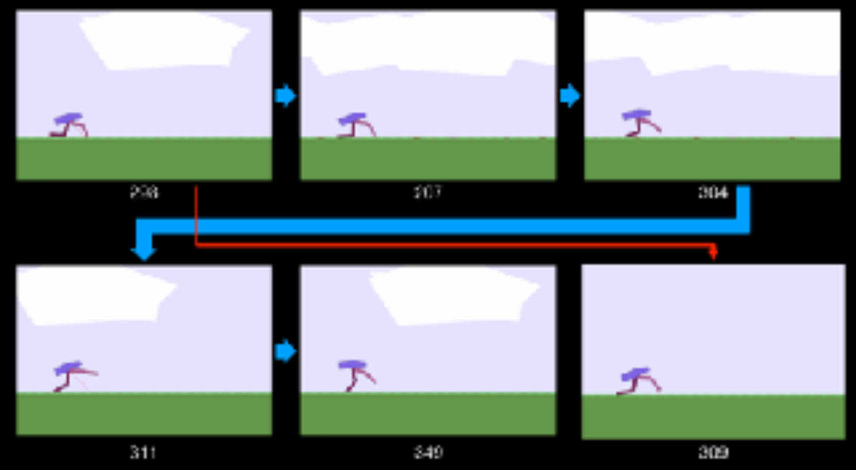
- Goal: create algorithms that *endlessly* learn & innovate
  - Like natural evolution and human culture

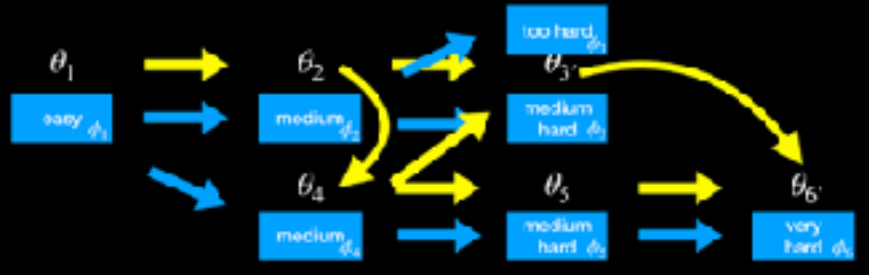
AI-Generating Algorithms

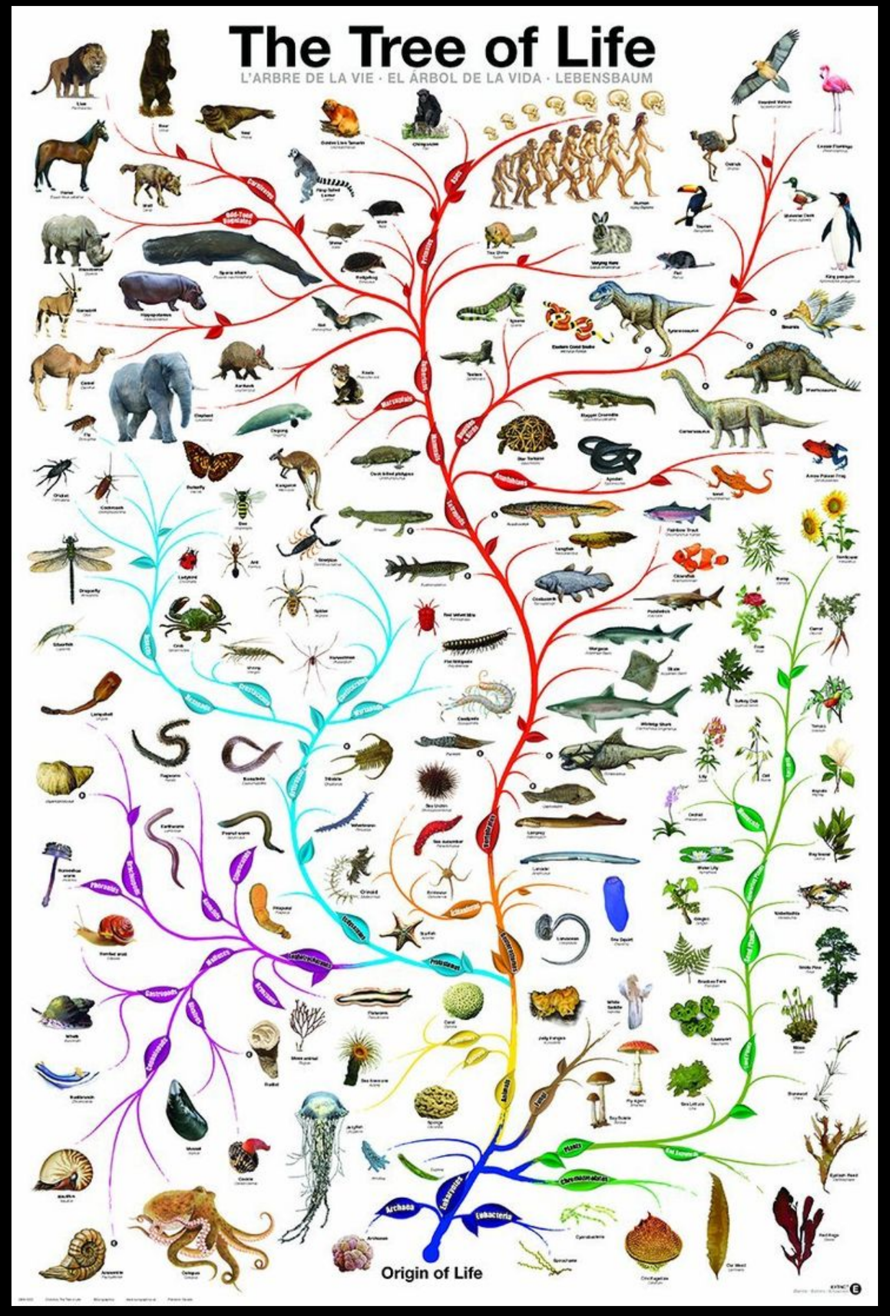
- Open-endedness to produce AGI
- Likely to be fastest path to AI











Overall Conclusions 2/3

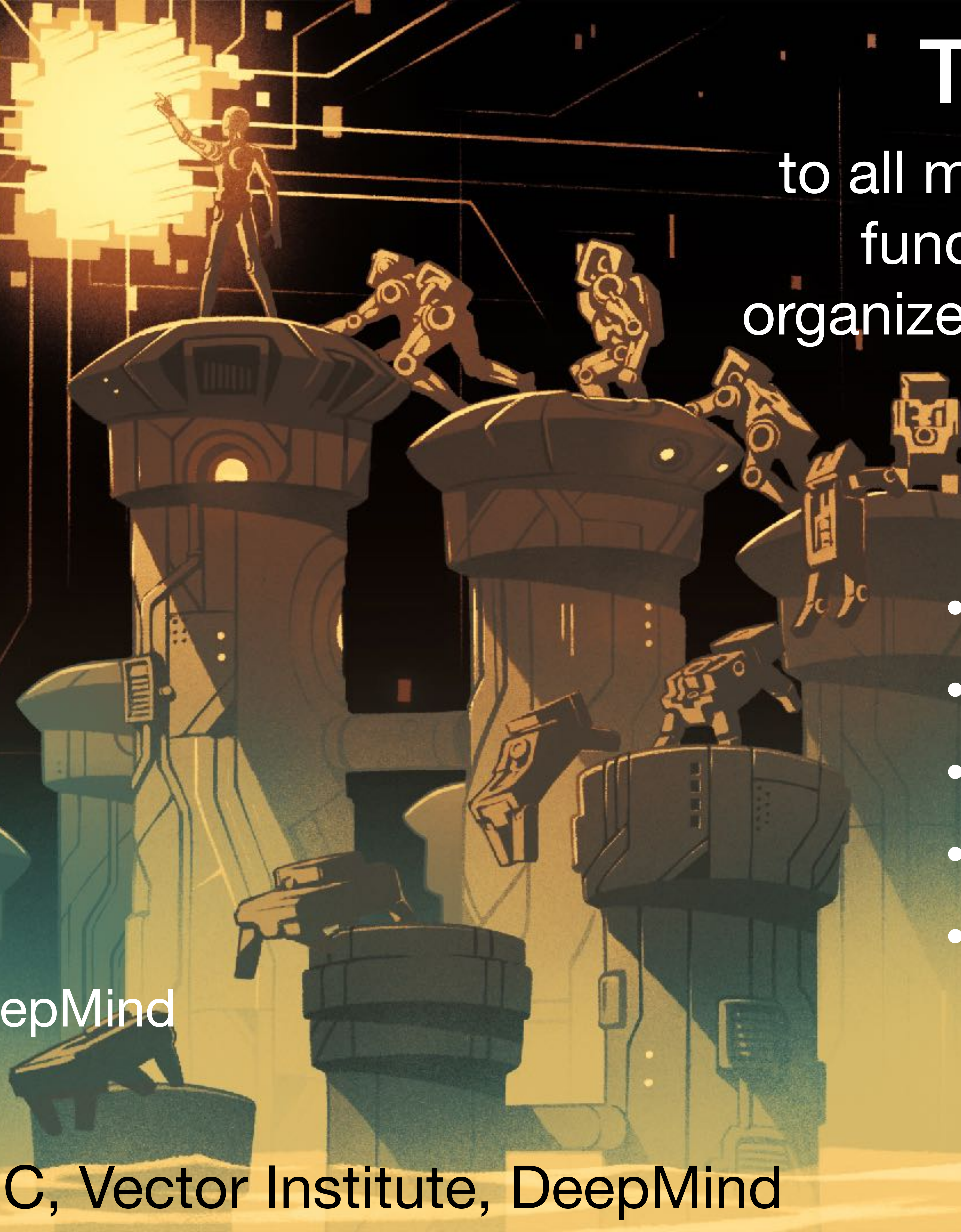
- Open-endedness and AI-GAs are the next wave in AI
- Via Foundation Models, OE algorithms can stand on the shoulders of giant human datasets
  - Great catalysts
  - Both
    - Environments (Genie)
    - Agents (VPT, Sima)
      - Evolution took 3.5 billion years to produce humans
      - VPT/Sima allow us to leapfrog to the interesting part

Overall Conclusions 3/3

- OMNI-EPIC: we can finally pursue what is interesting!
  - In Darwin Complete search spaces
  - Avoids the pitfall of trying to define “interesting”
  - Profound opportunities for QD,OE, AI-GAs, etc.
- The AI Scientist
  - Same set of exciting ideas applied to agents and science itself!
- + Thought Cloning
  - Adds performance, generalization, adaptability
  - Also interpretability, debugability, and safety

Thanks!

to all my collaborators,
funding sources,
organizers, and all of you!



Main collaborators

- Jenny Zhang
- Joel Lehman
- Ken Stanley
- Maxence Faldor
- Antoine Cully
- Shengran Hu
- Cong Lu
- VPT Team @ OpenAI
- Genie & SIMA teams @ DeepMind
- AI Scientist team

Funding Sources

- NSERC
- Schmidt Futures
- Open Philanthropy
- Vector Institute
- CIFAR

UBC, Vector Institute, DeepMind