Endlessly Generating Increasingly Complex and Diverse Learning Environments and their Solutions through the Paired Open-Ended Trailblazer (POET)

Jeff Clune

Senior Research Manager

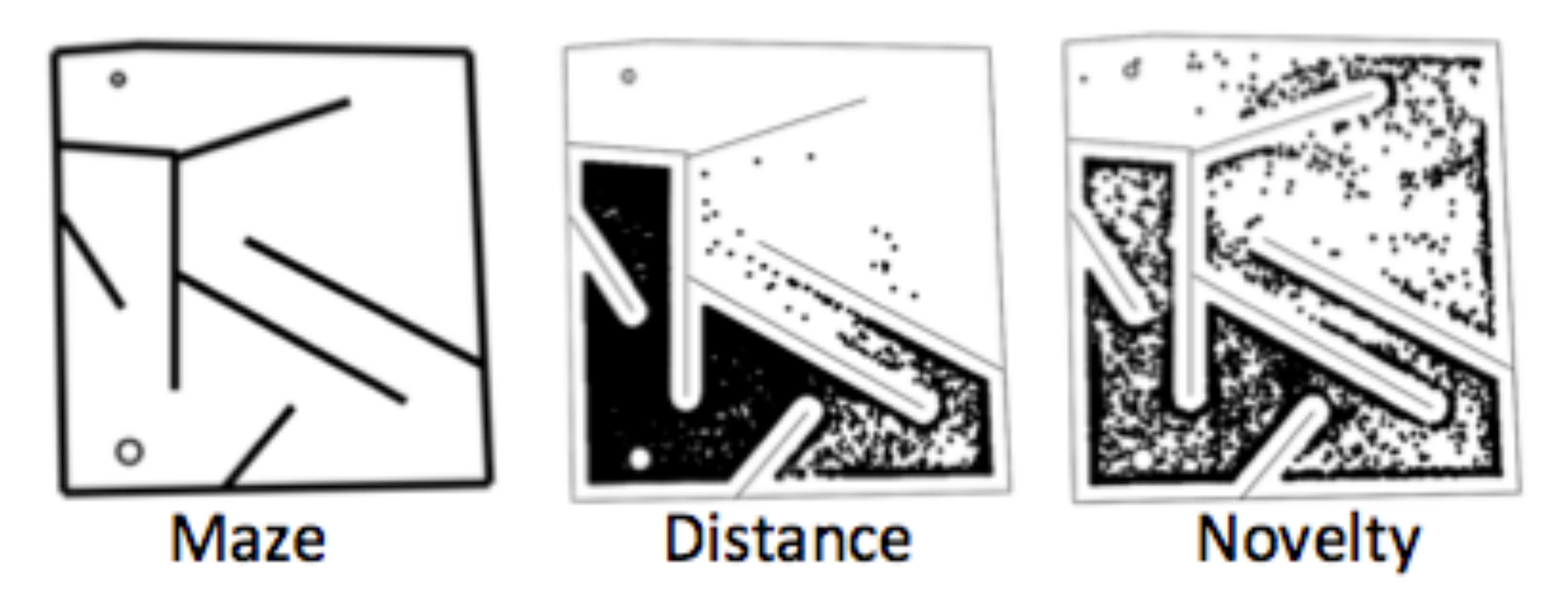
UBER AI Labs

Harris Associate Professor, Computer Science

University of Wyoming

A Paradox

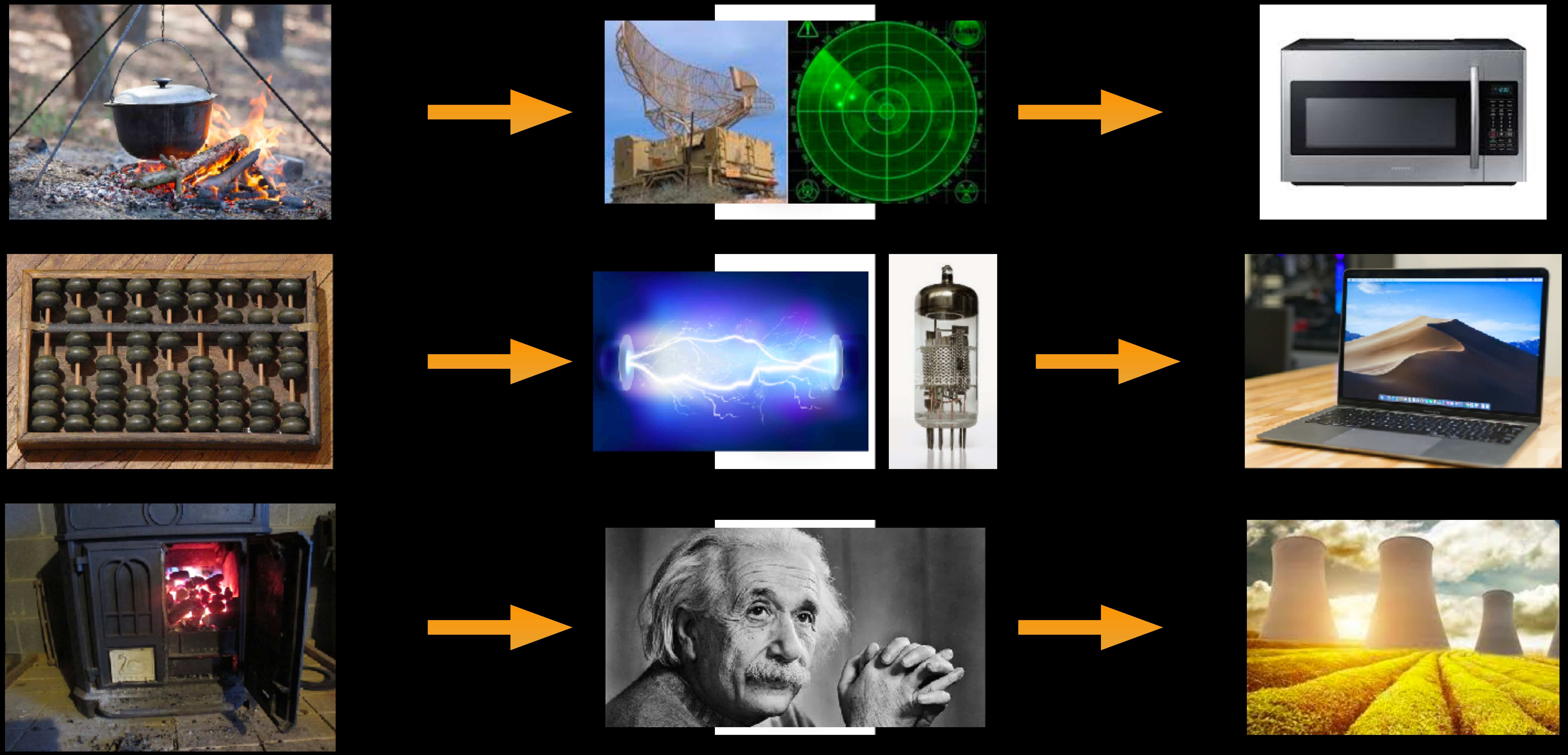
- If you try too hard to solve a hard problem, you'll fail
- If you ignore the objective, you're more likely to succeed

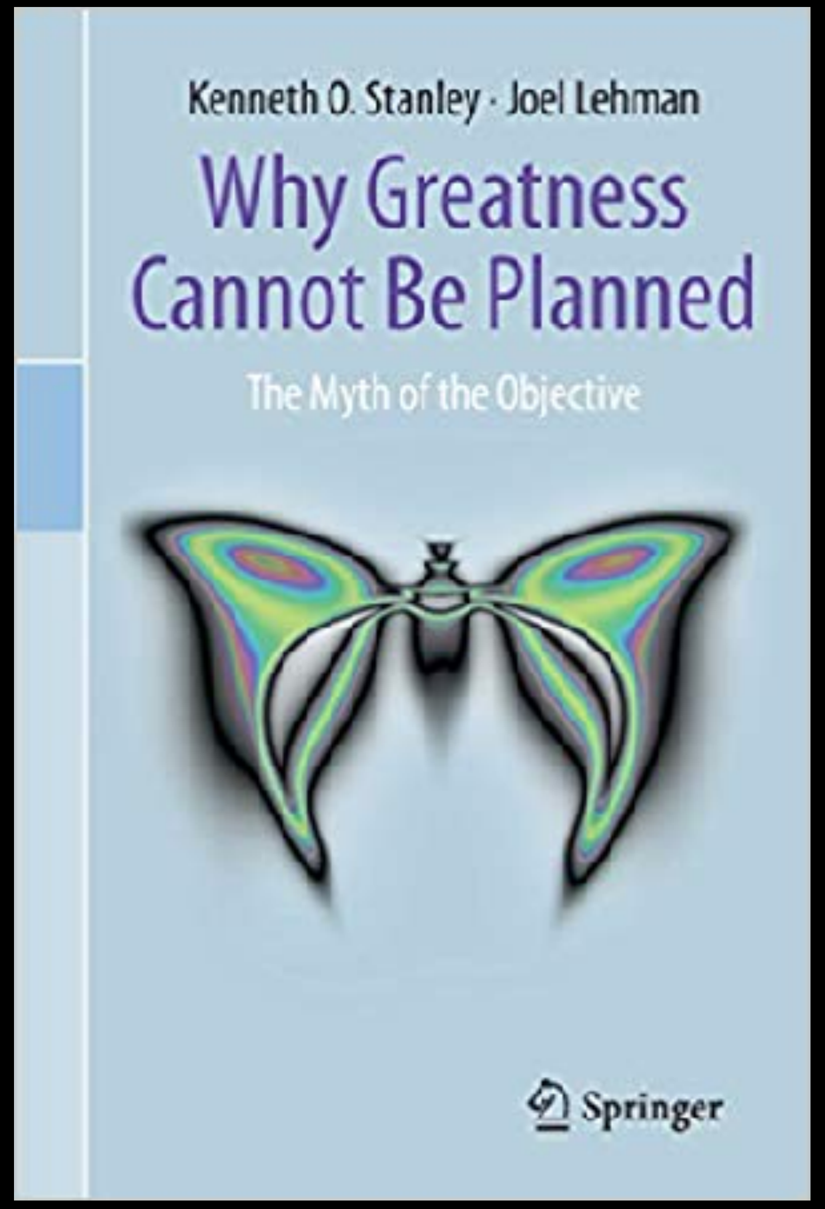


Maximize Reward: Fails
Maximize Novelty: Succeeds

Novelty Search: Lehman & Stanley

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



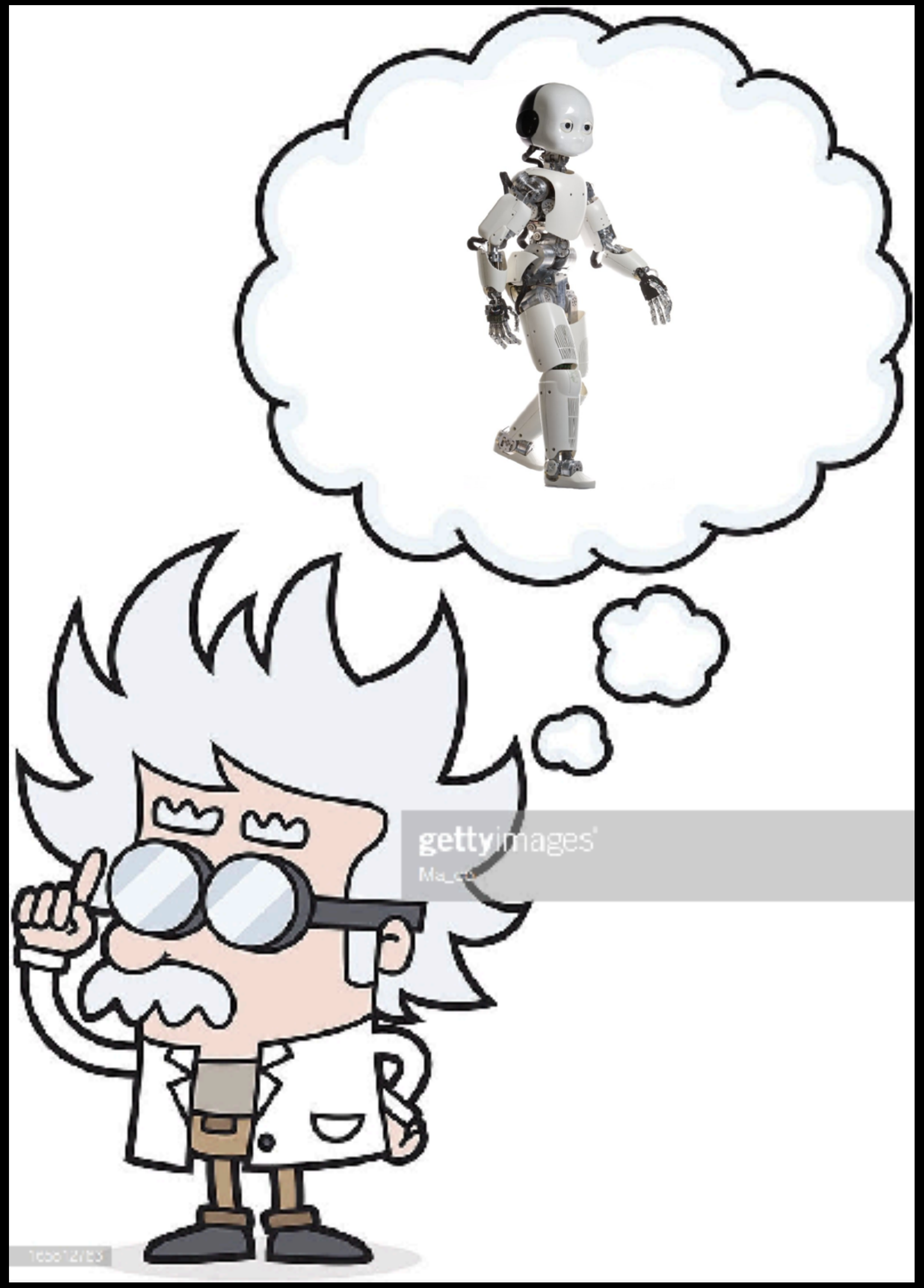


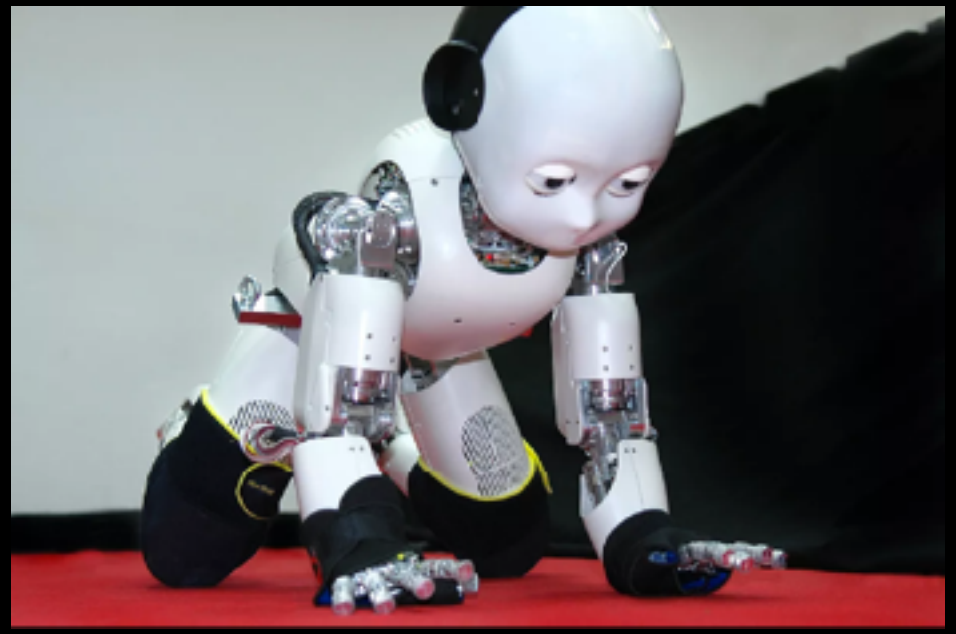
Conjecture: The only way to solve hard problems may be by creating problems while you solve them and goal switching between them

“Goal Switching”

Nguyen, Yosinski & Clune 2016

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





Quality Diversity Algorithms

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

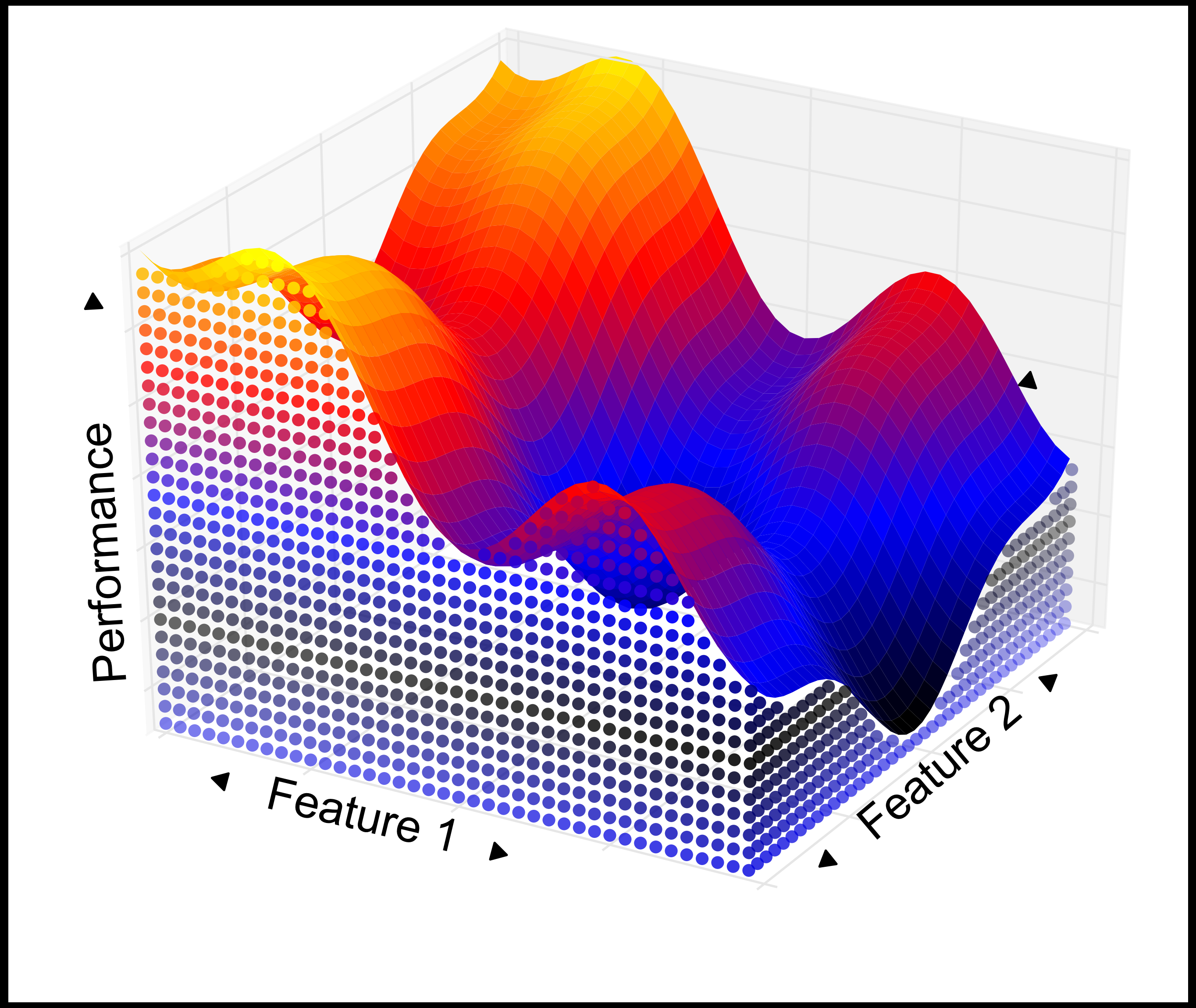
Lehman & Stanley 2011
Mouret & Clune 2015
Pugh et al. 2016

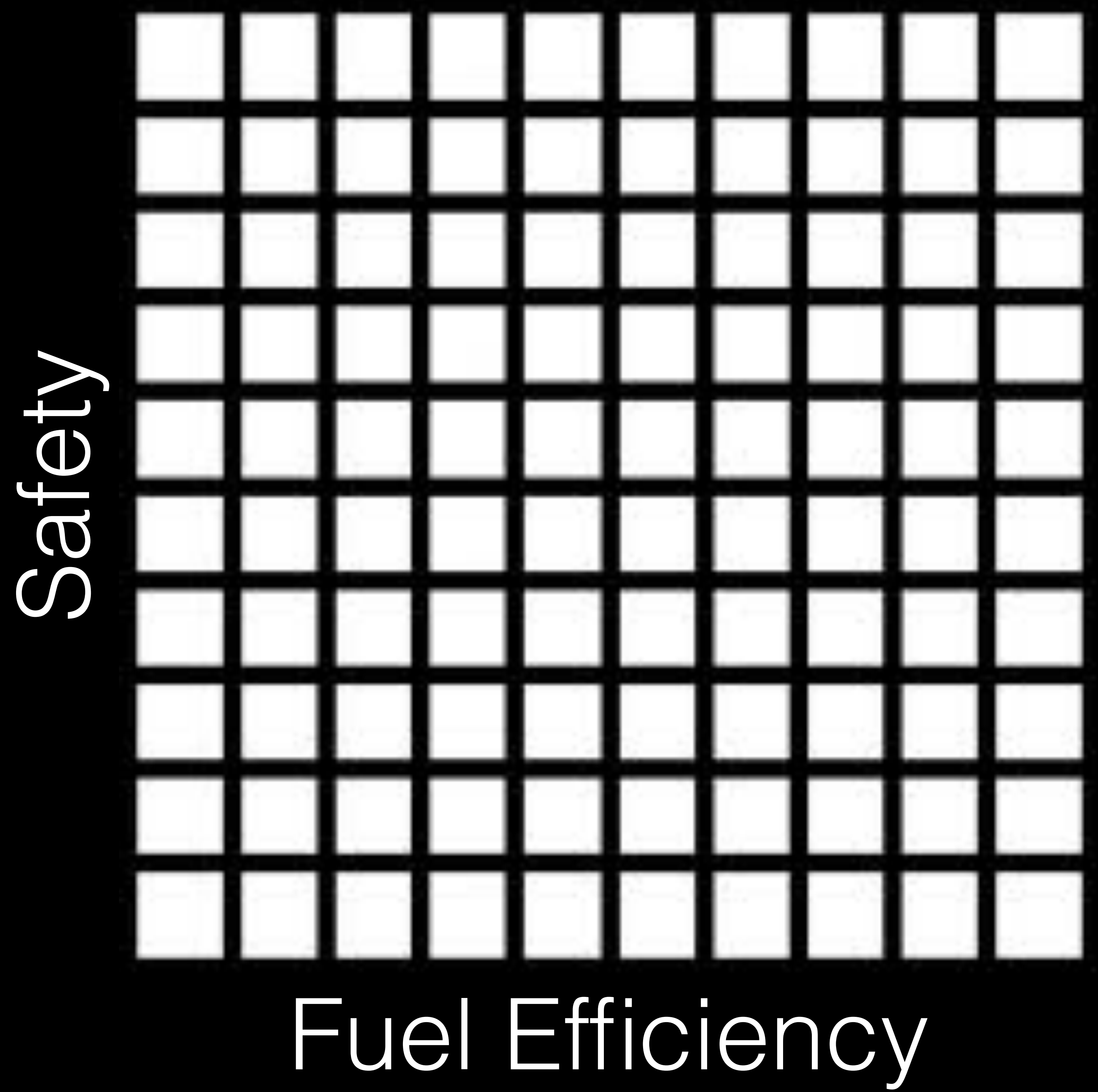
MAP-Elites

Mouret & Clune 2015, arXiv

- Choose dimensions of interest, discretize them
- Search for highest-performing policy in each cell

Set of diverse, high-quality solutions



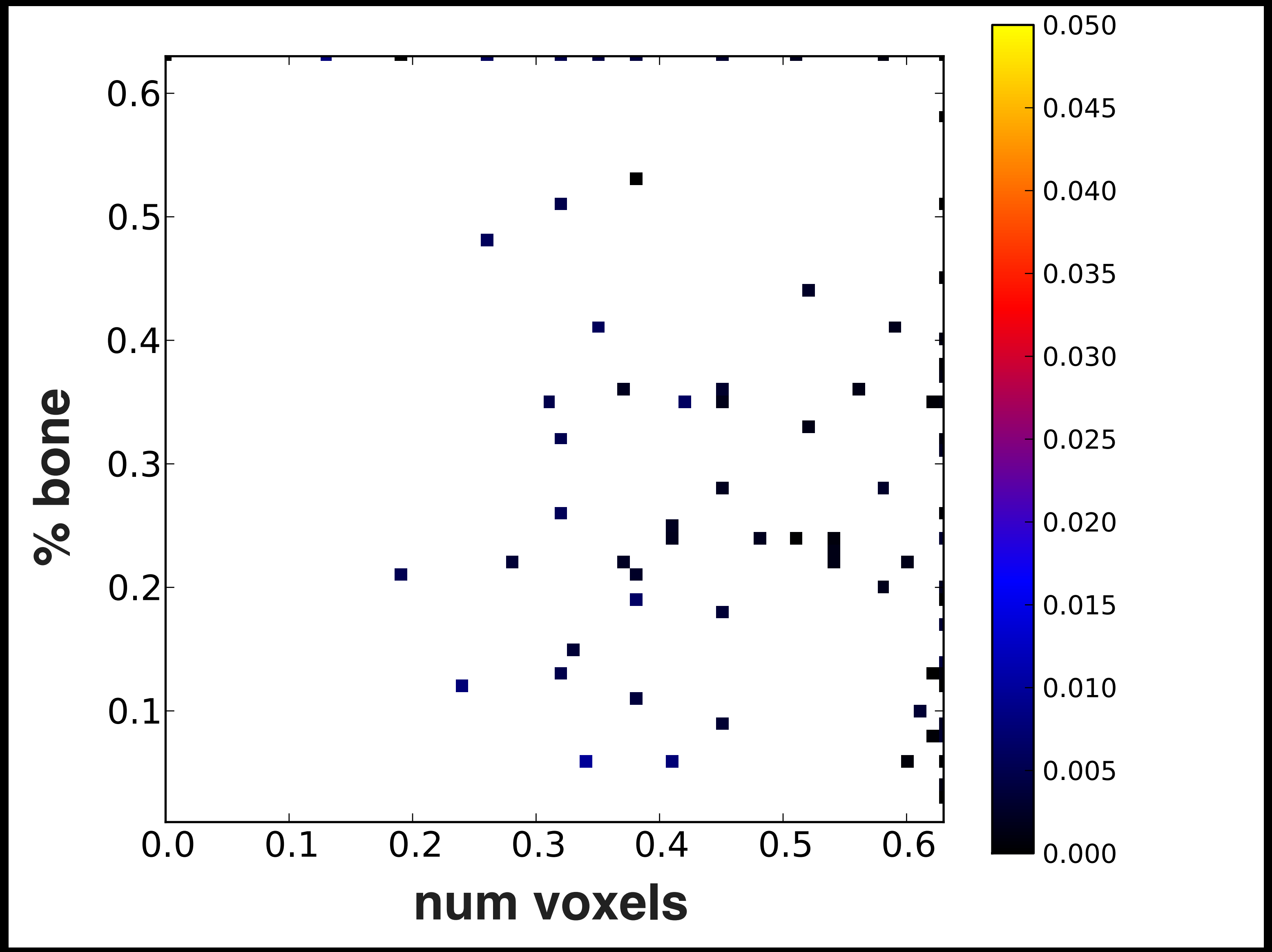


Qualitatively Different

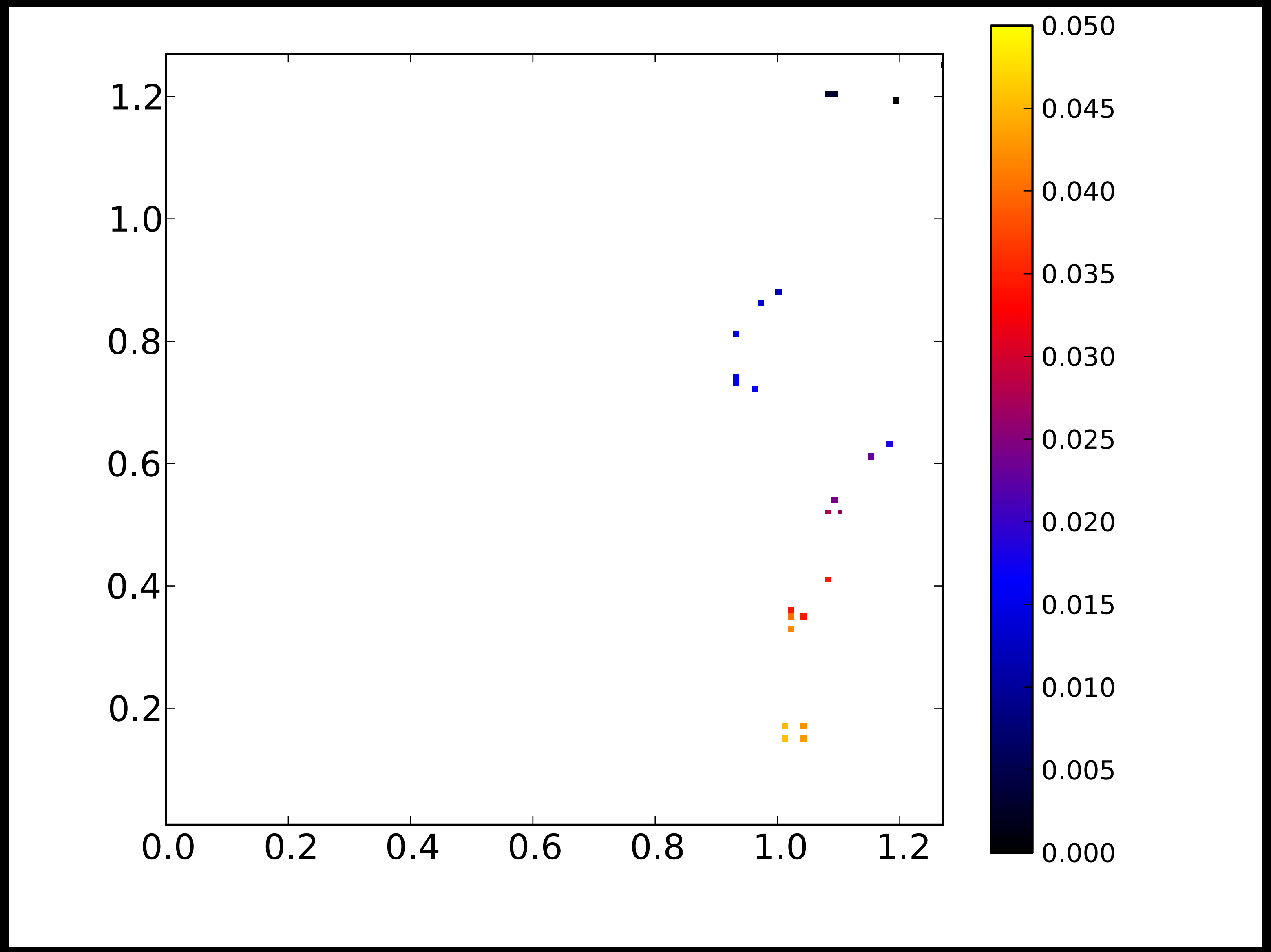
Mouret & Clune 2015, arXiv

soft robots problem

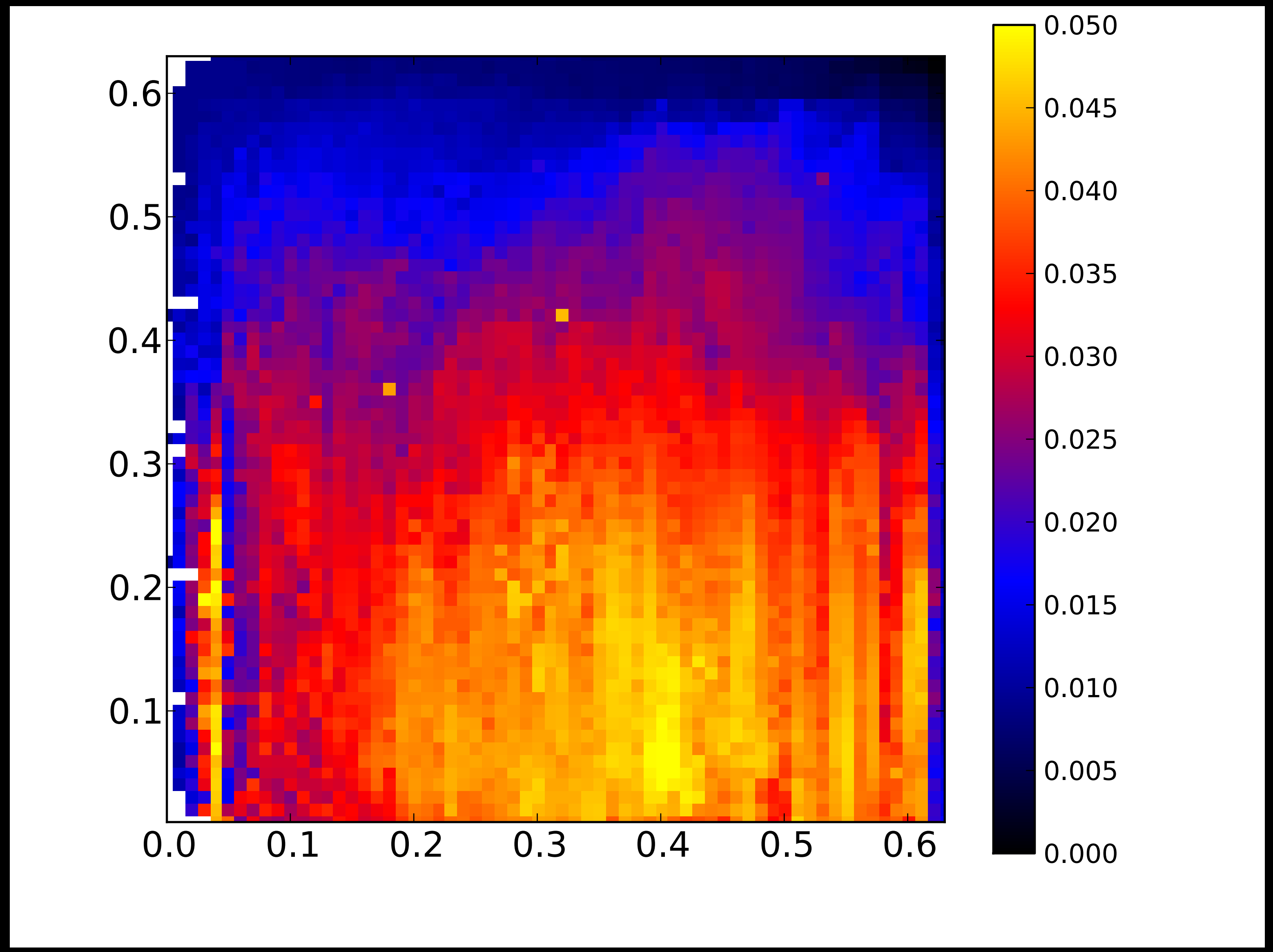
Classic Optimization



Classic + Diversity



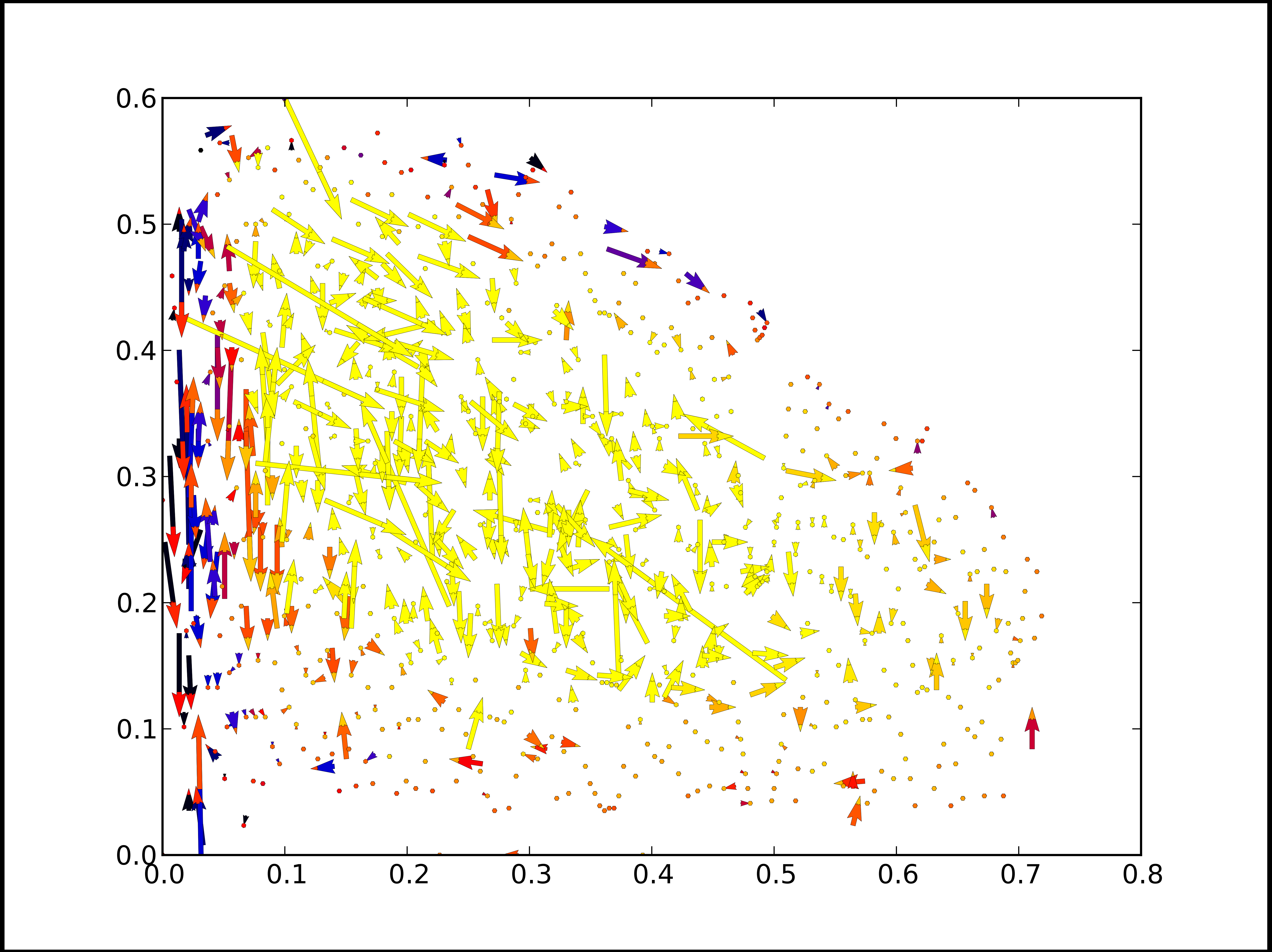
MAP-Elites



same # evals!

Often finds a better max than max-focused algorithms!

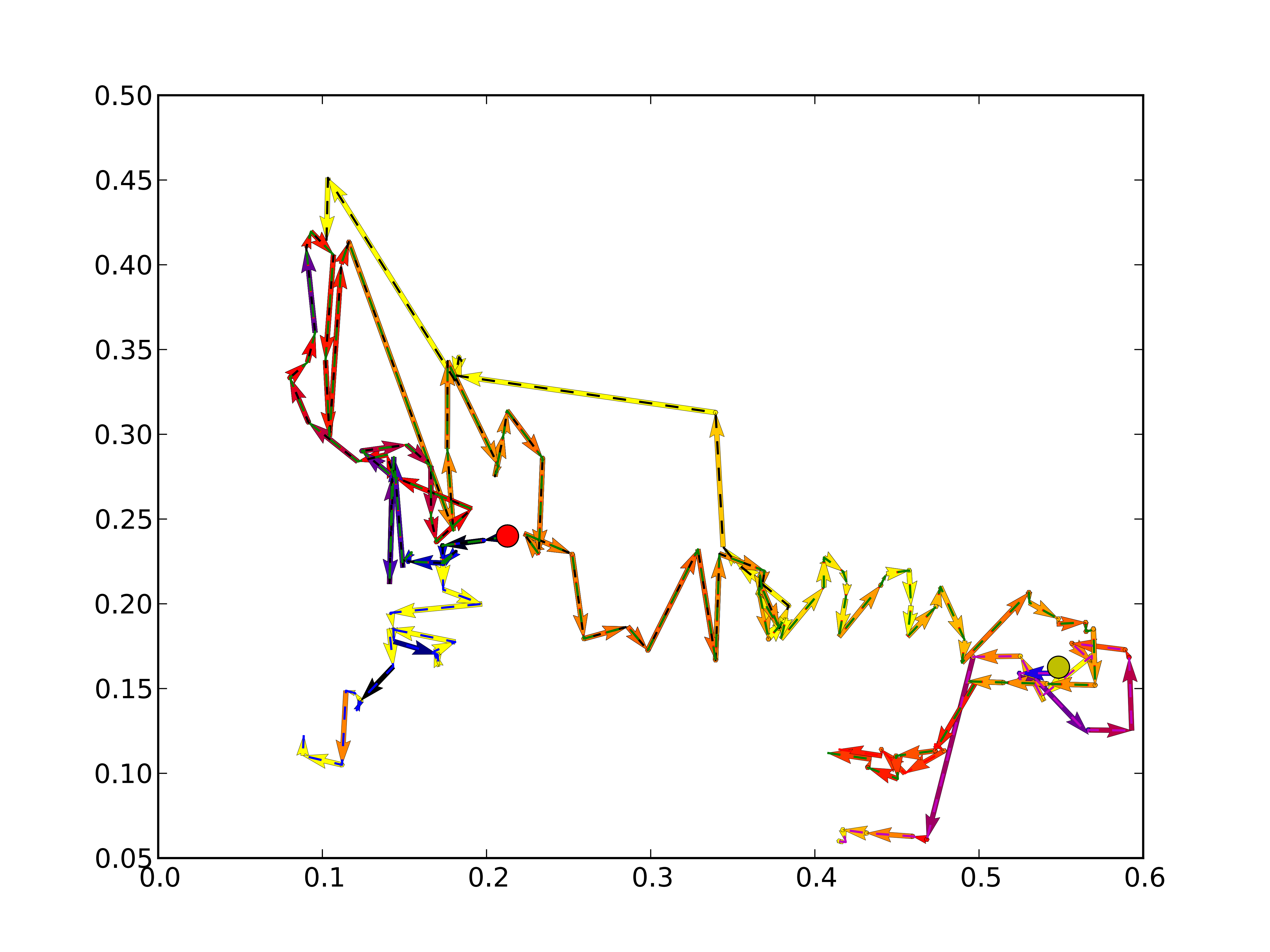
Goal Switching is Critical

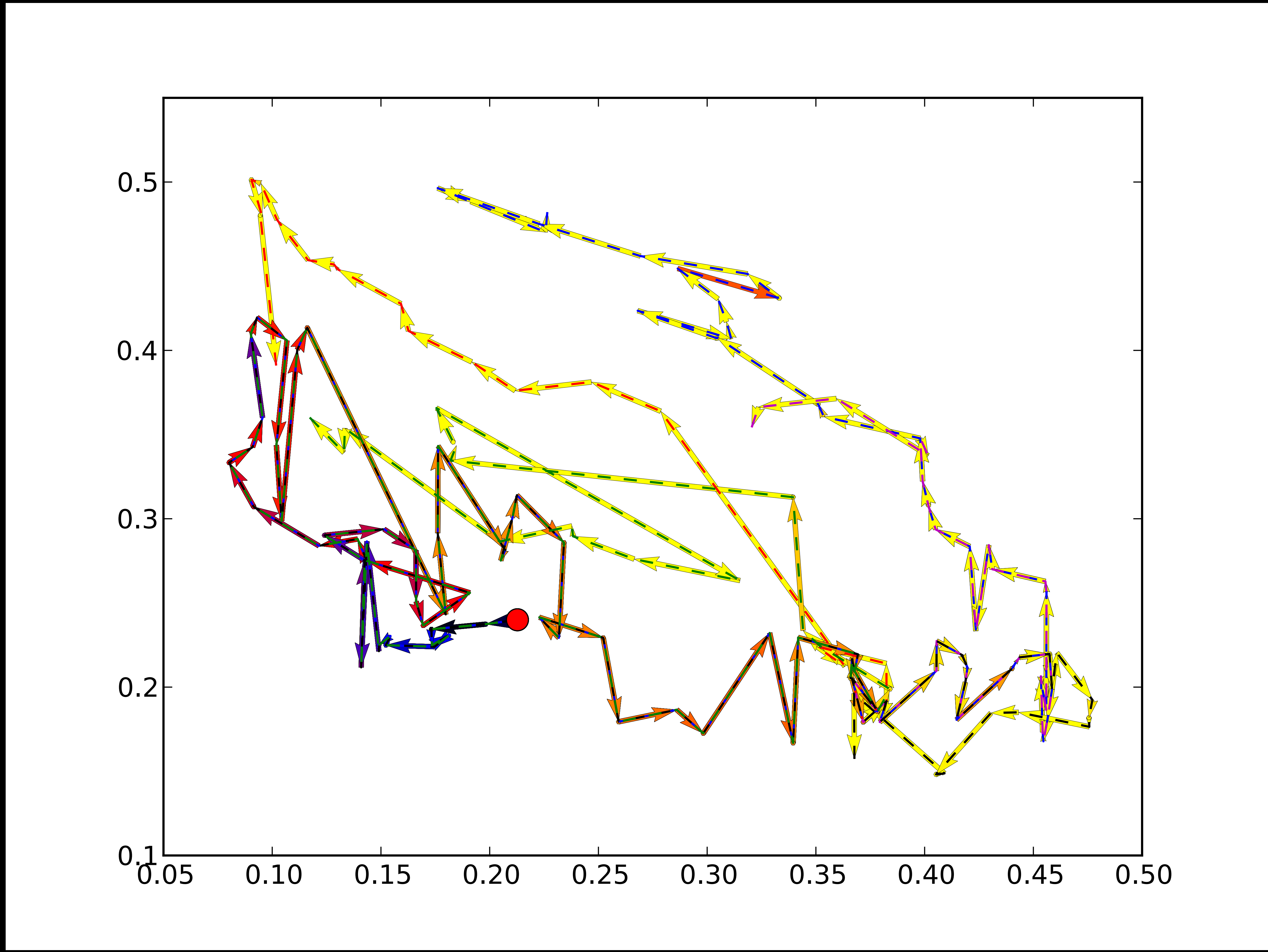


retina problem

color = reward

MAP-Elites Lineages of a Few Final Solutions





Circles are iteration 0, color = reward

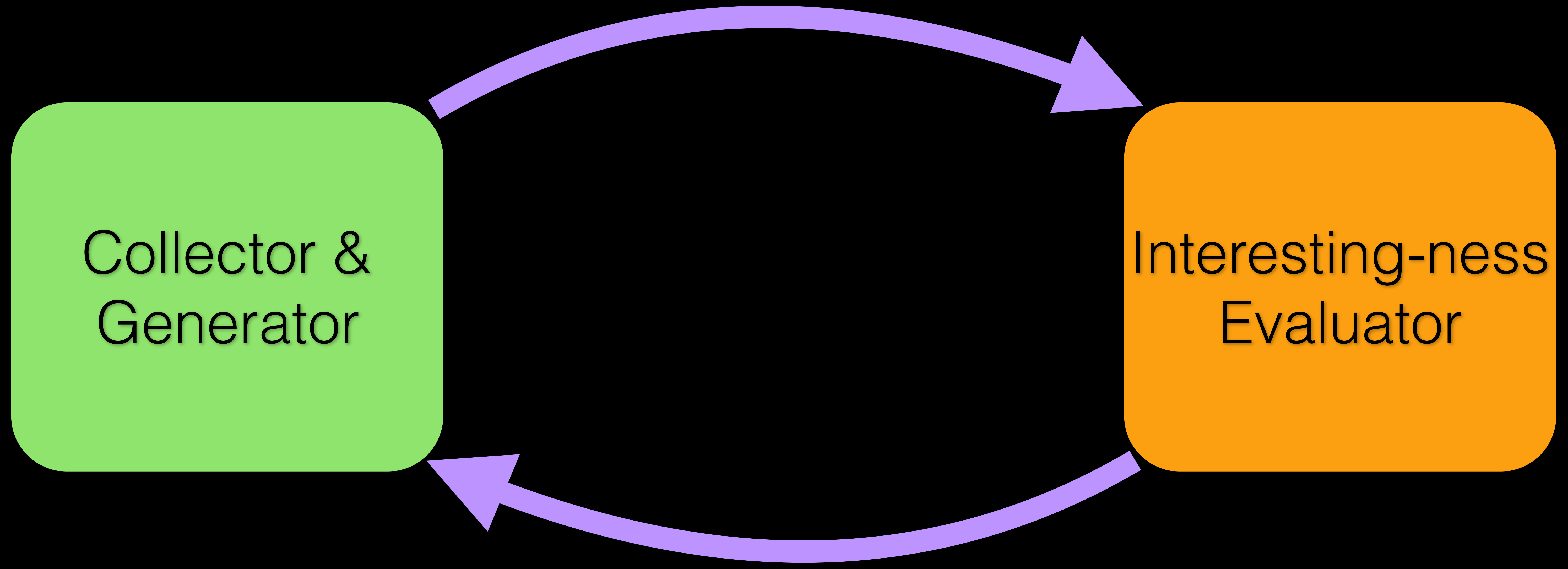
Innovation Engines

Nguyen, Yosinski & Clune 2015

- Nature, Culture, & QD algorithms are *Innovation Engines*
  - generate permutations of previous interesting things
  - if interesting, keep them
  - repeat

Innovation Engines

Nguyen, Yosinski & Clune 2015

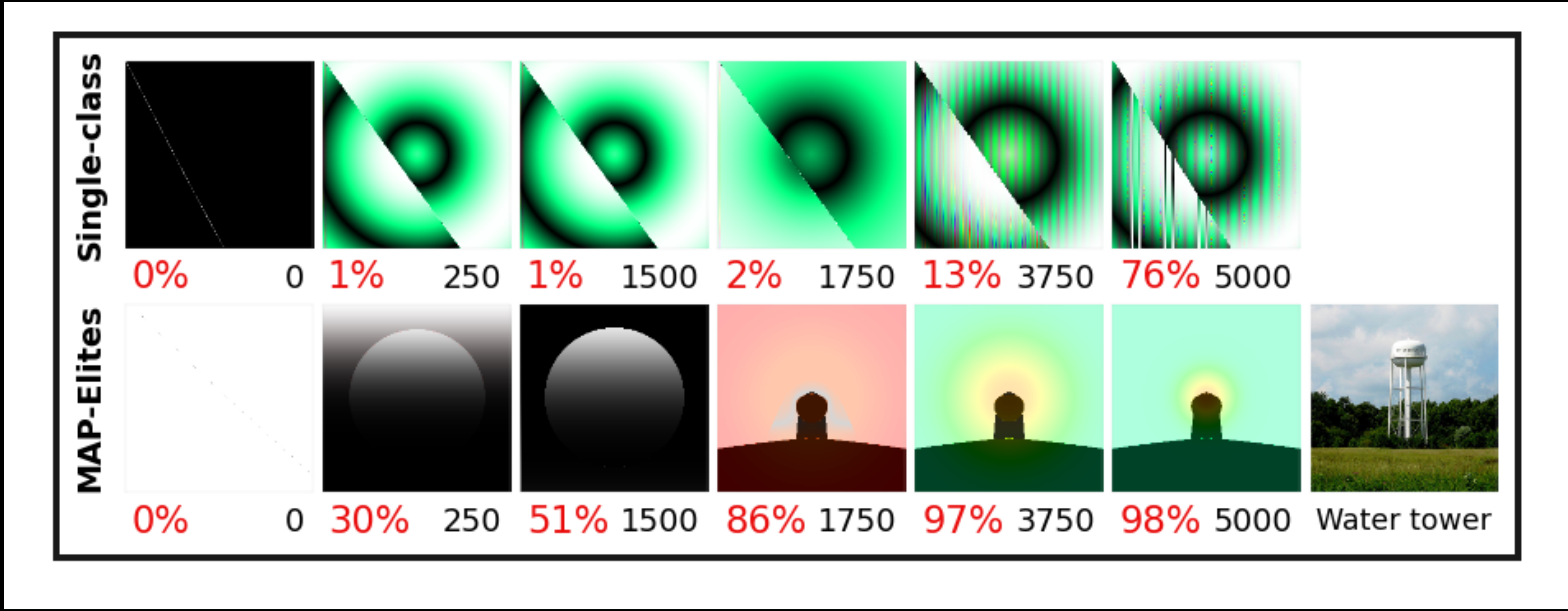


MAP-Elites
one bin per ImageNet class

AlexNet

Encodings: Small CPPN networks

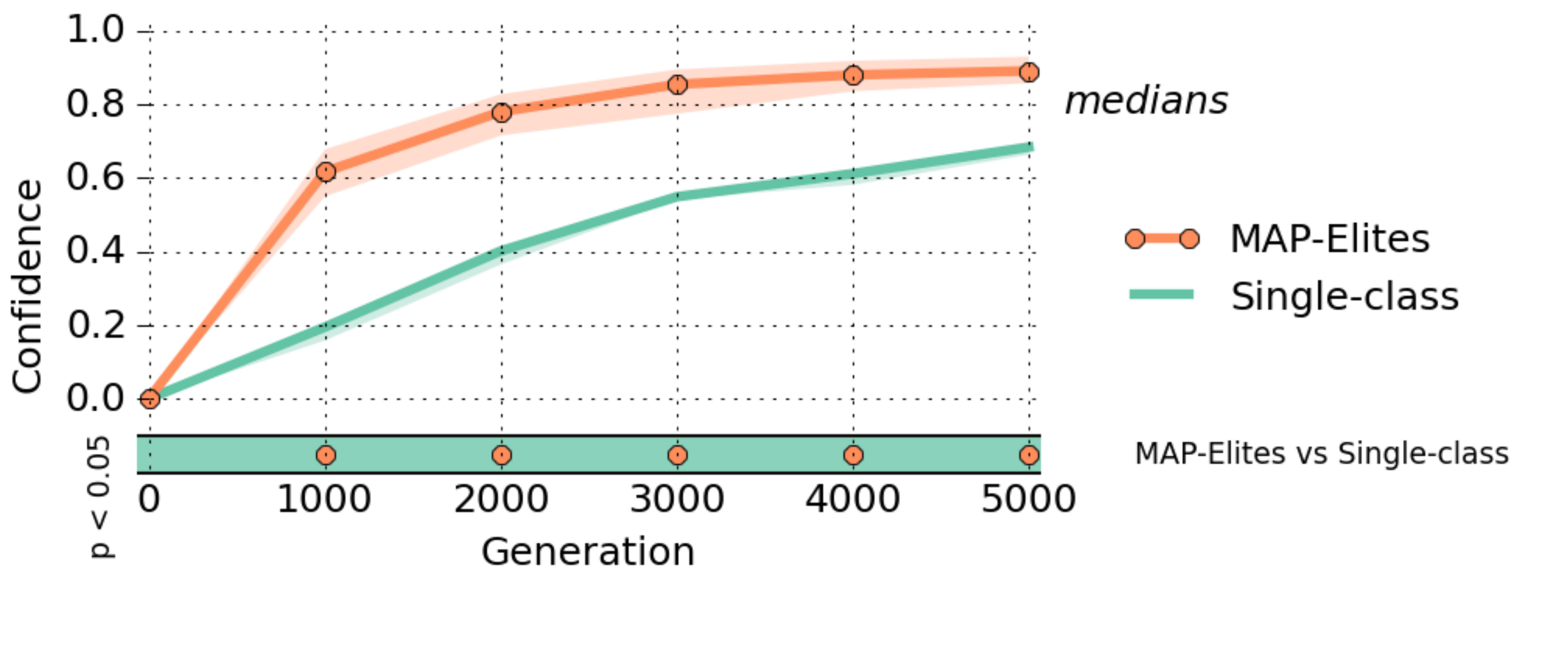
Goal Switching



Nguyen, Yosinski & Clune 2015

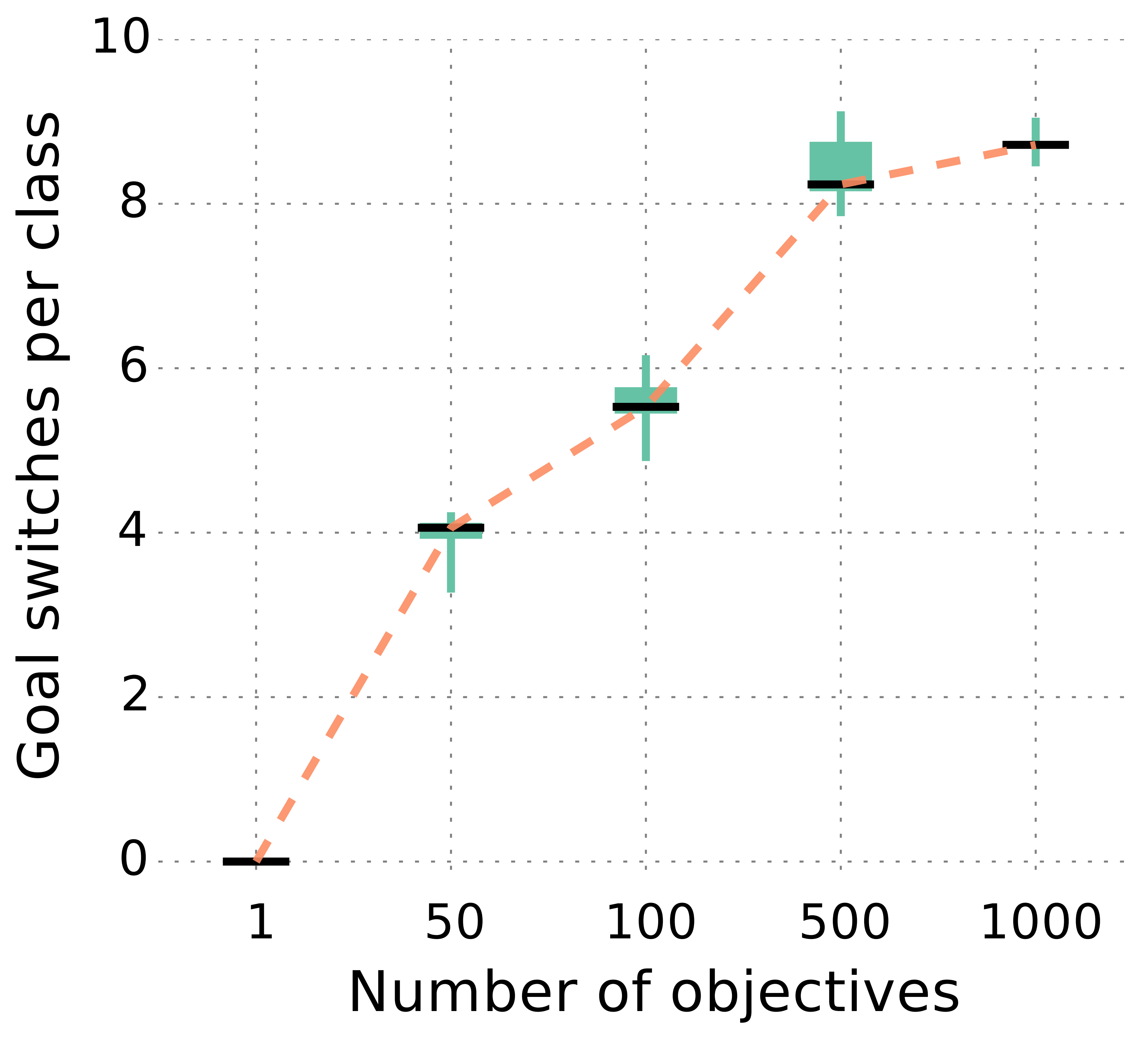
Goal Switching

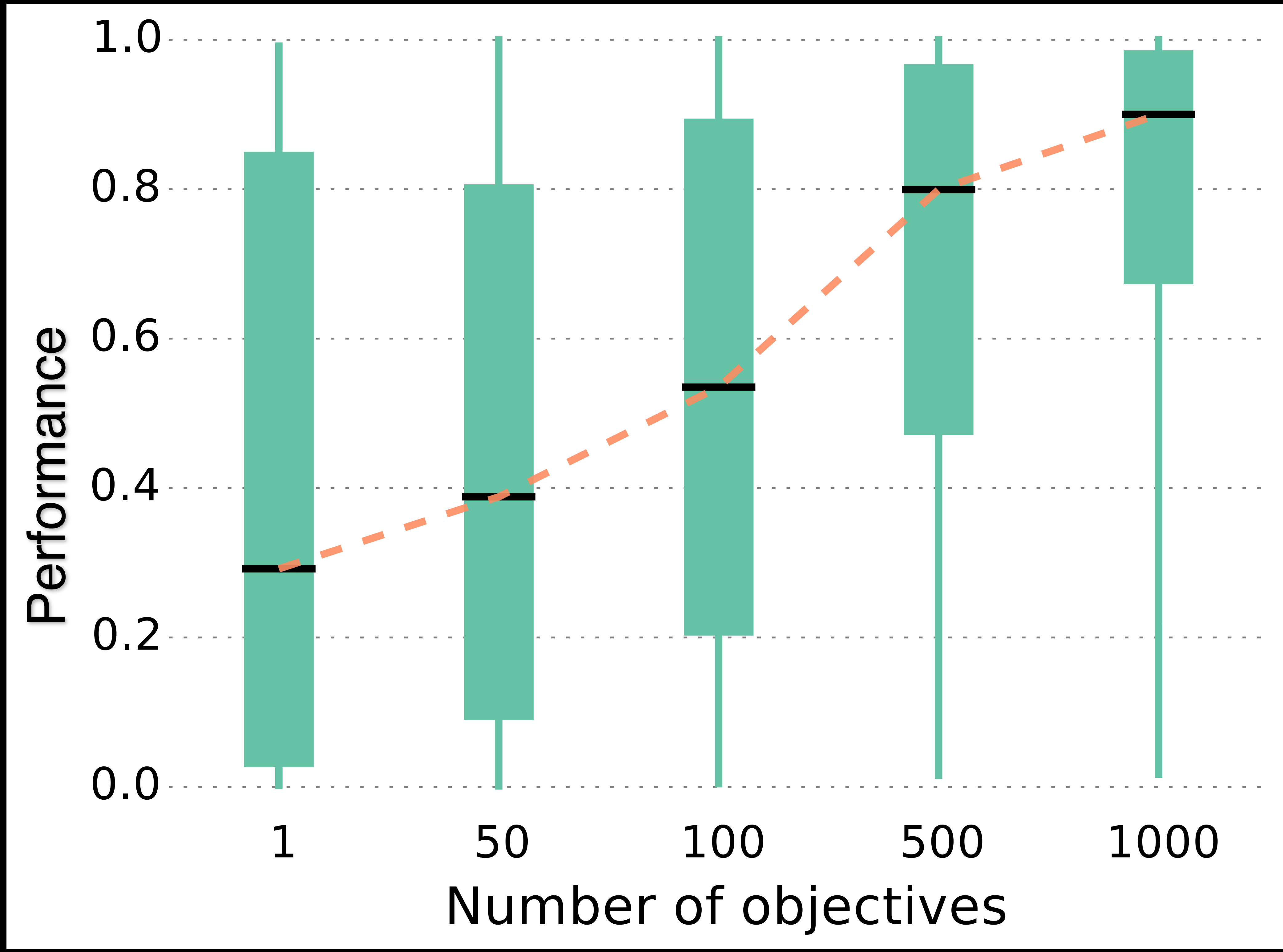
- Many-class MAP-Elites vs. One-class MAP-Elites



Nguyen, Yosinski & Clune 2015

Goal Switching

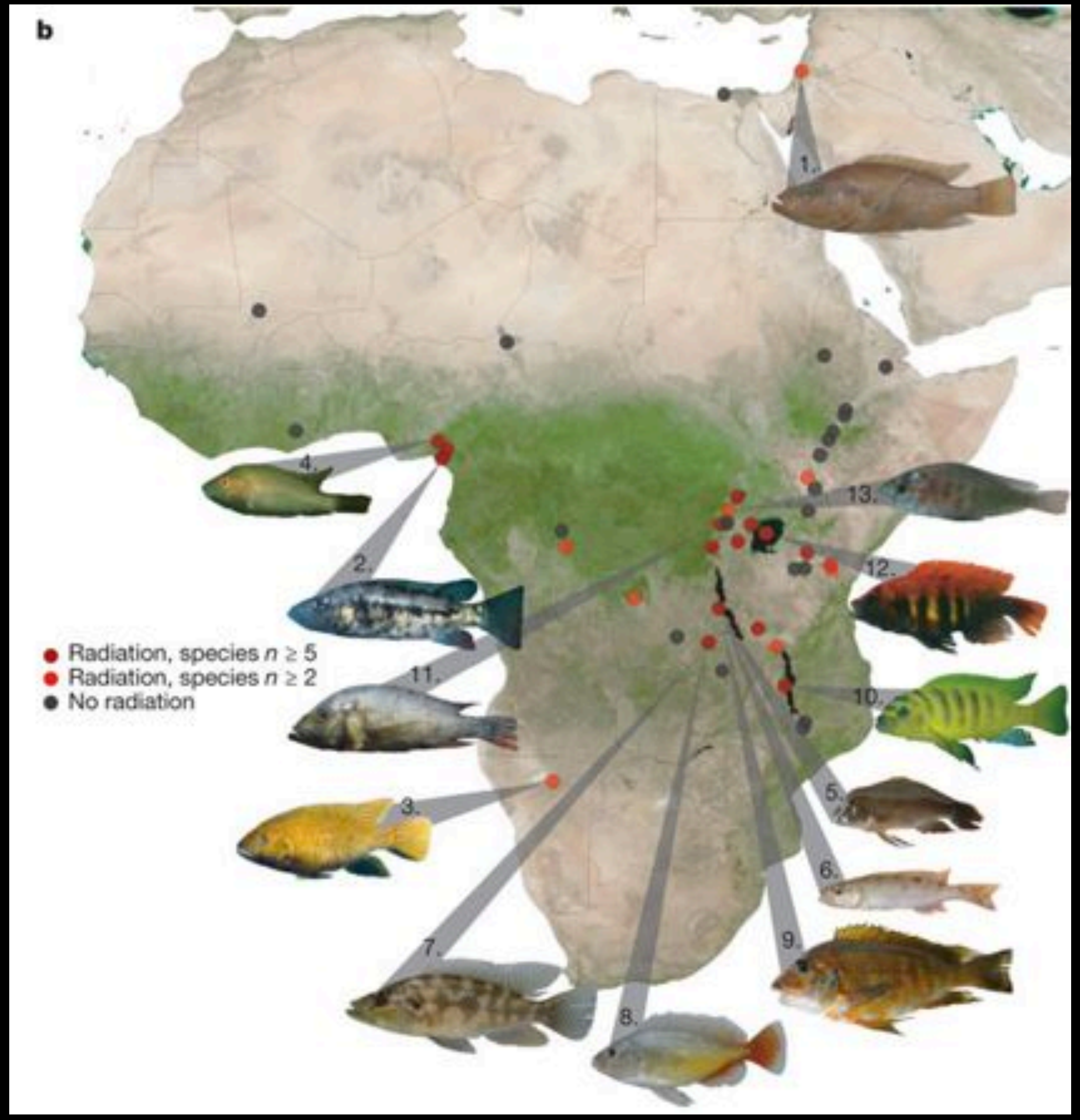


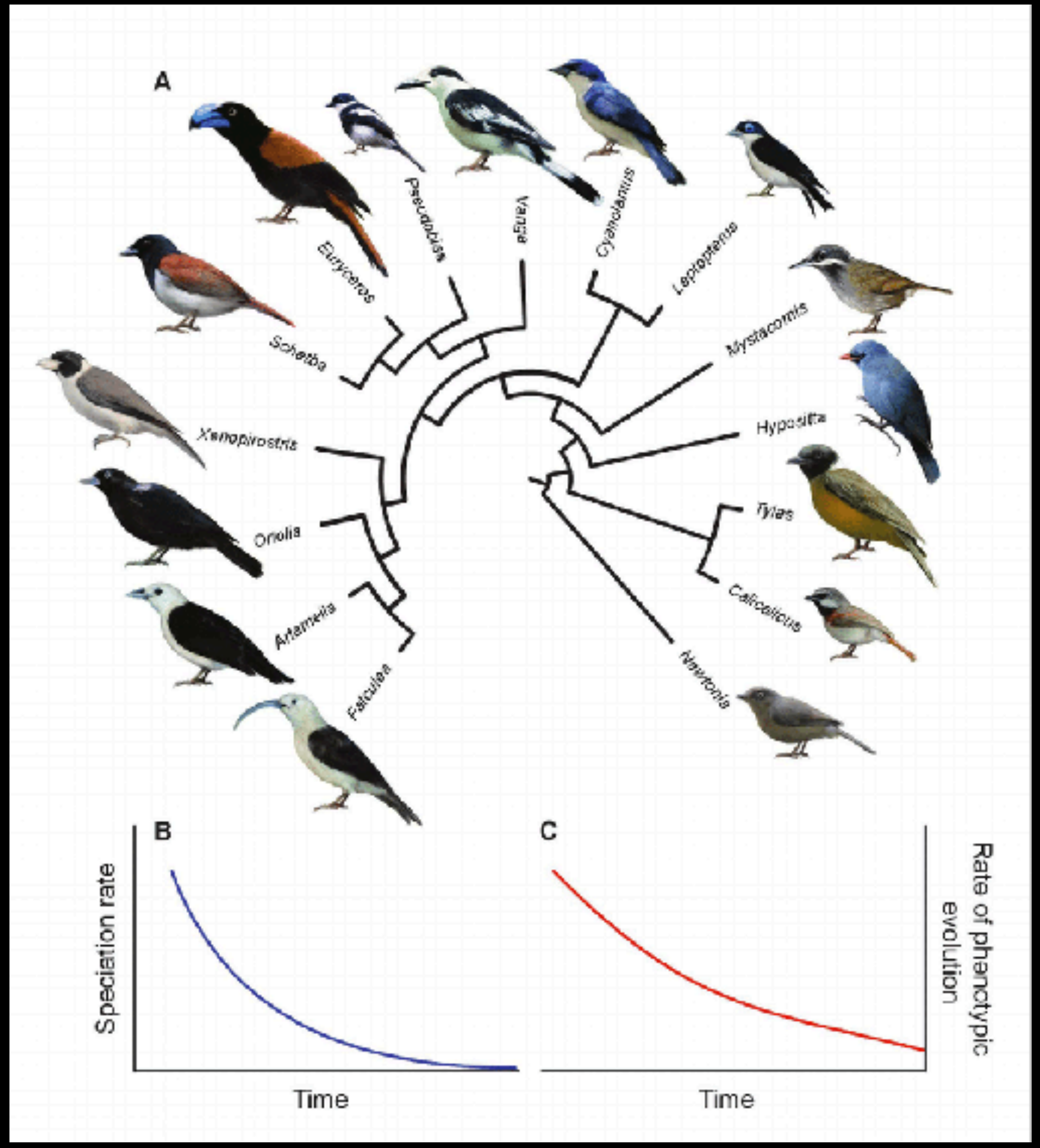


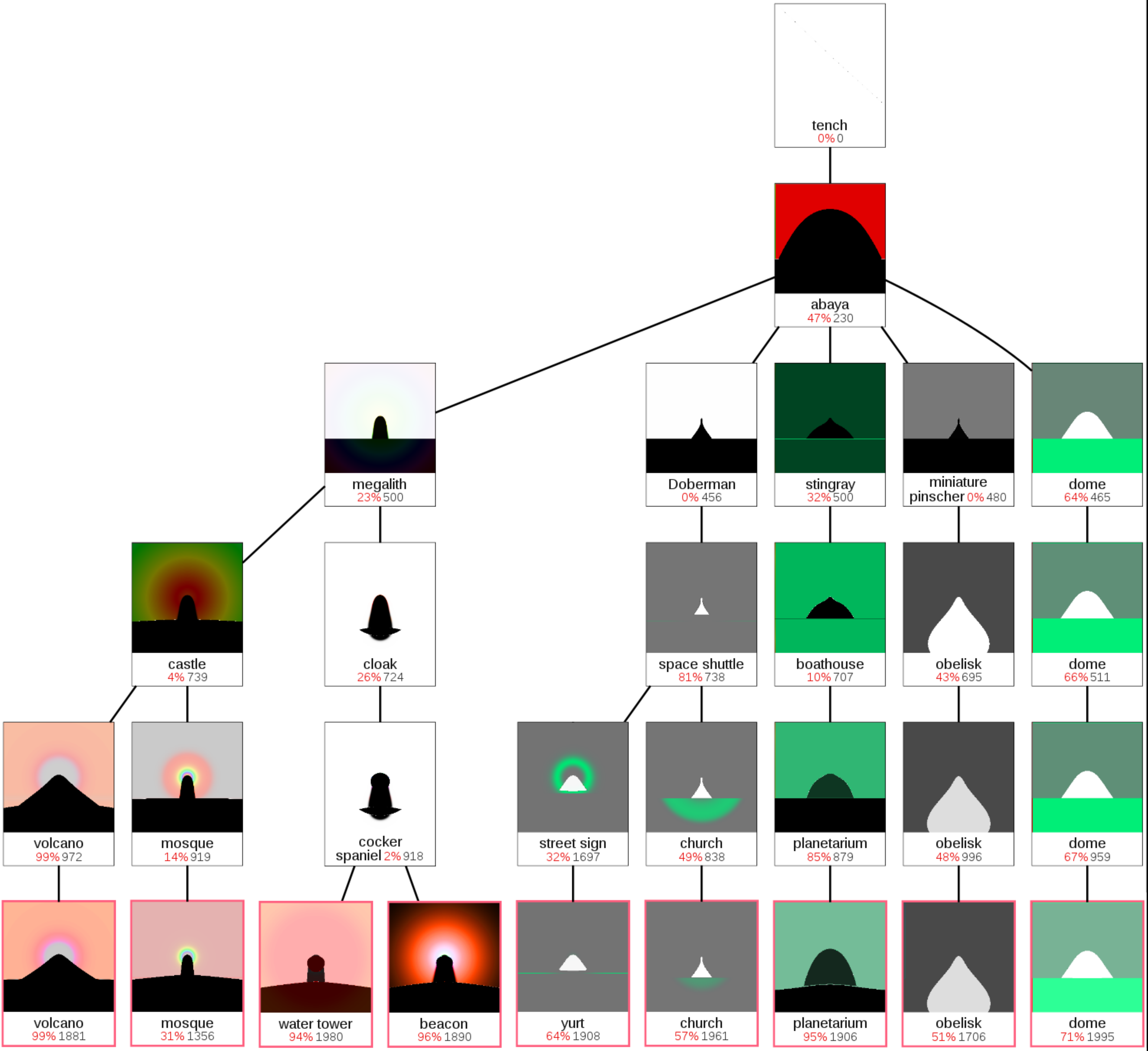
Nguyen, Yosinski & Clune 2015

Goal Switching Enables Good Ideas to Spread

- Fundamental advances spread to other problems/niches
- Then are built upon to solve that specific problem
- "Adaptive Radiations"







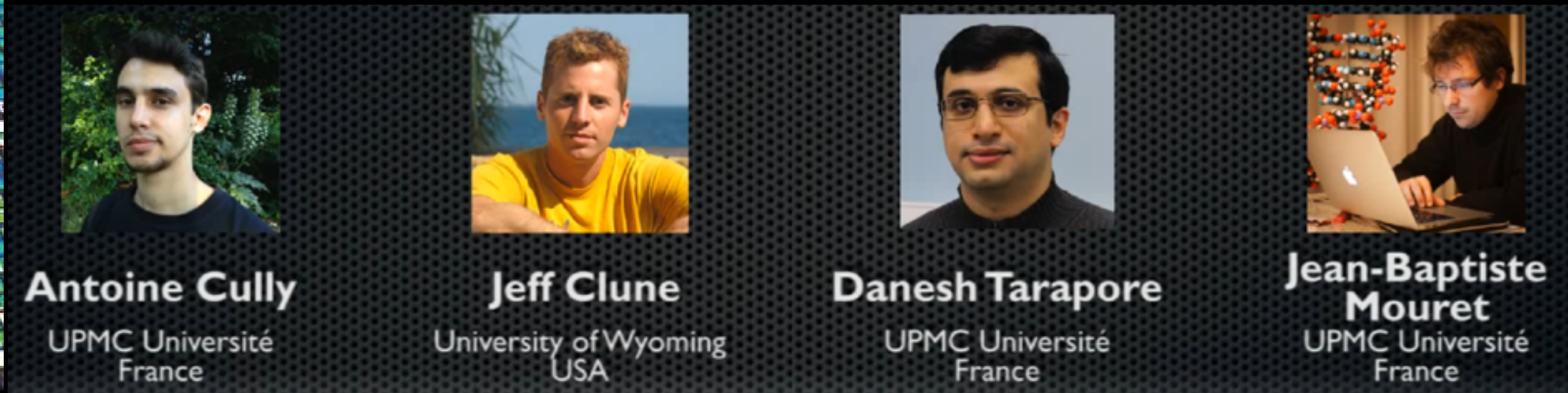
Adaptive Radiations in QD!

Innovation Engines

Nguyen, Yosinski & Clune 2015

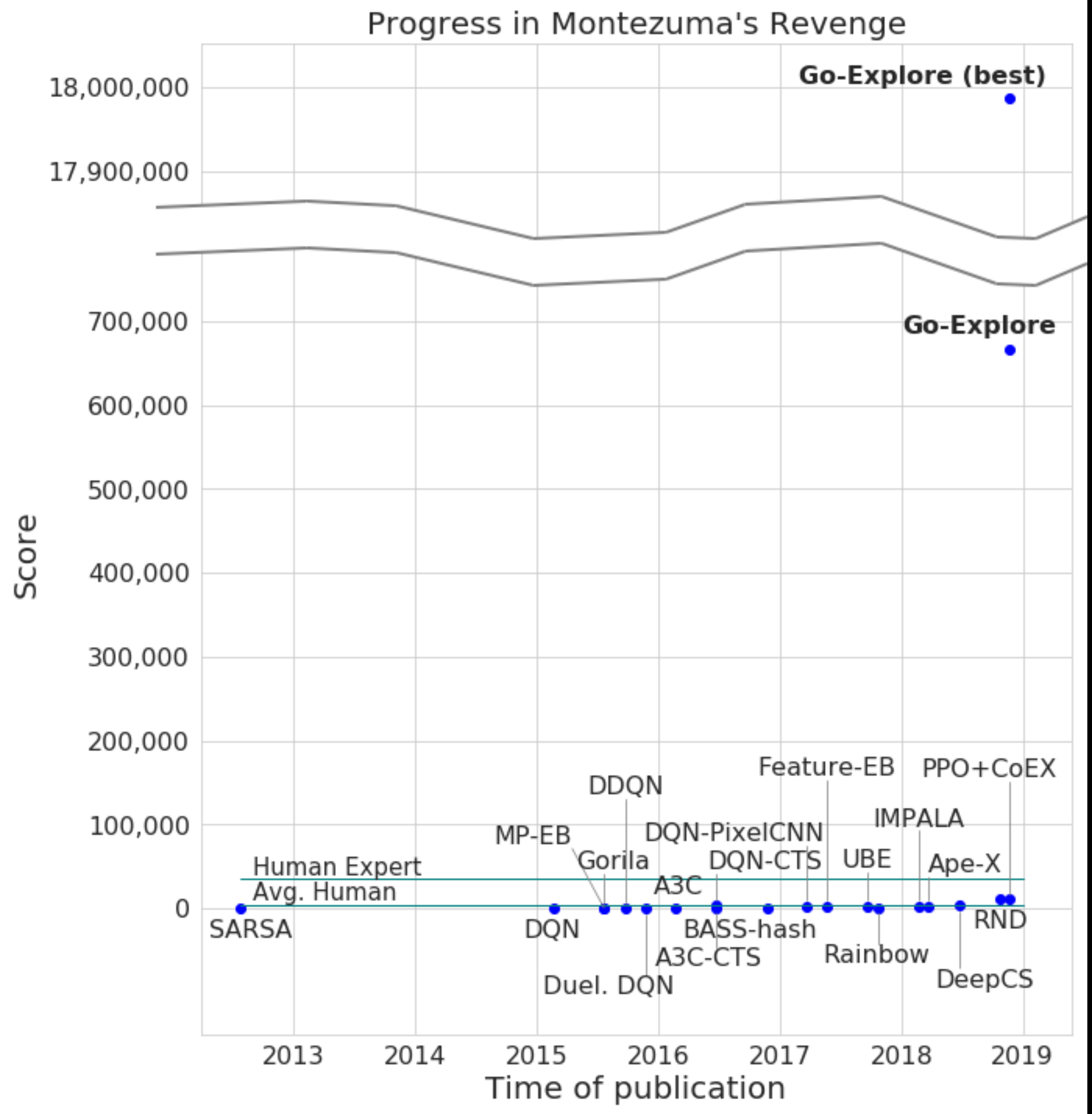
Robots that adapt like animals

2015

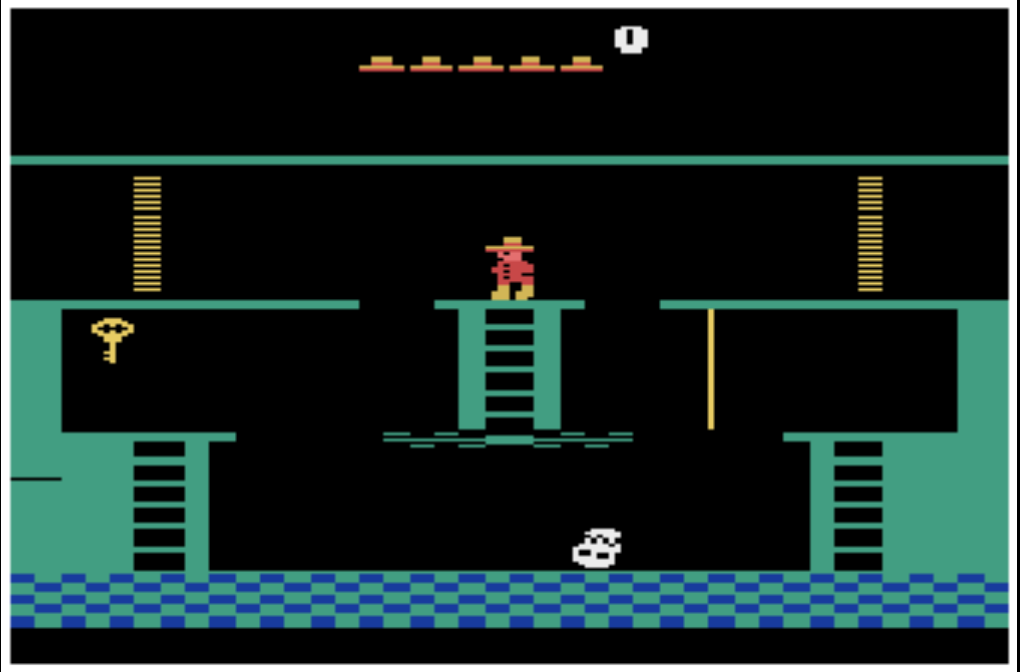


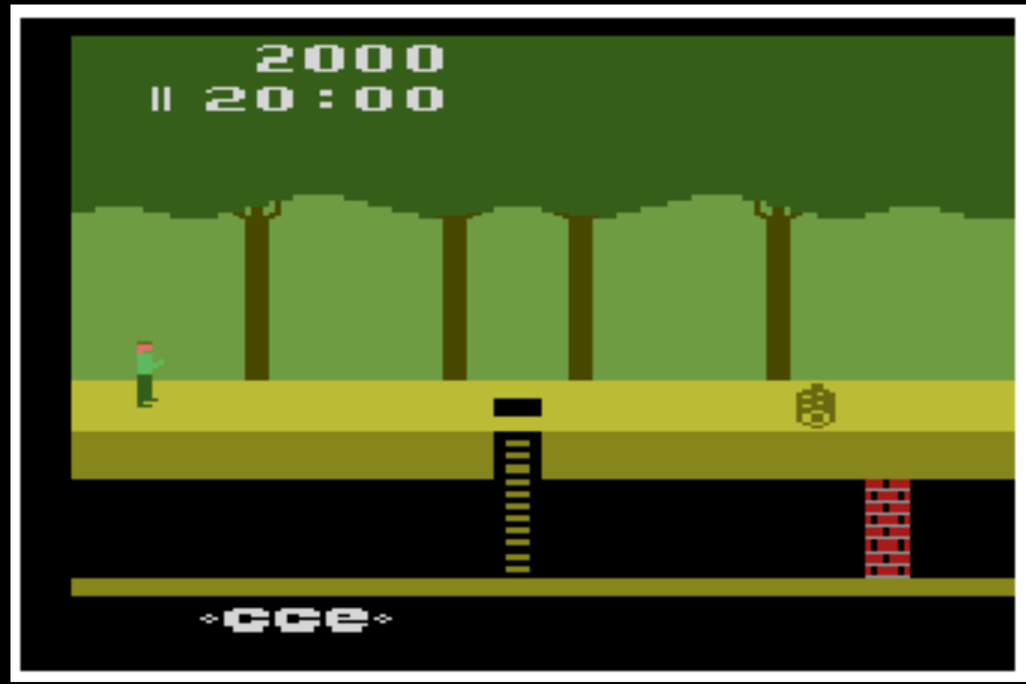
Go-Explore: Solves Montezuma's Revenge

Ecoffet, Huizinga, Lehman, Stanley, & Clune 2019, arXiv



- Average score: 660,000
- Best Go-Explore policy
  - scores ~18 million
  - solved 1,141 levels
- Beats human world record
  - 1,219,200



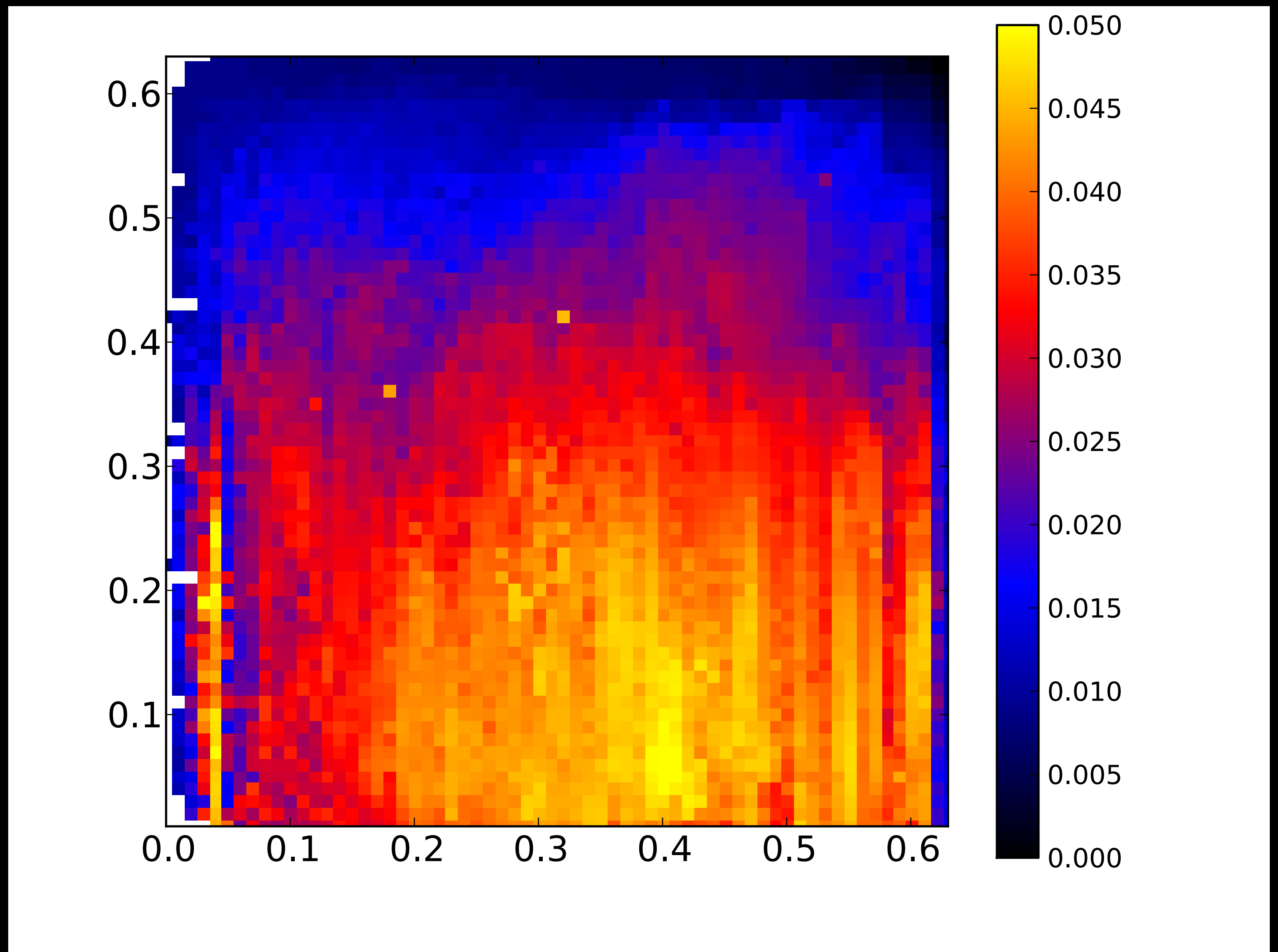


Caveat: Exploits deterministic training

What’s missing?

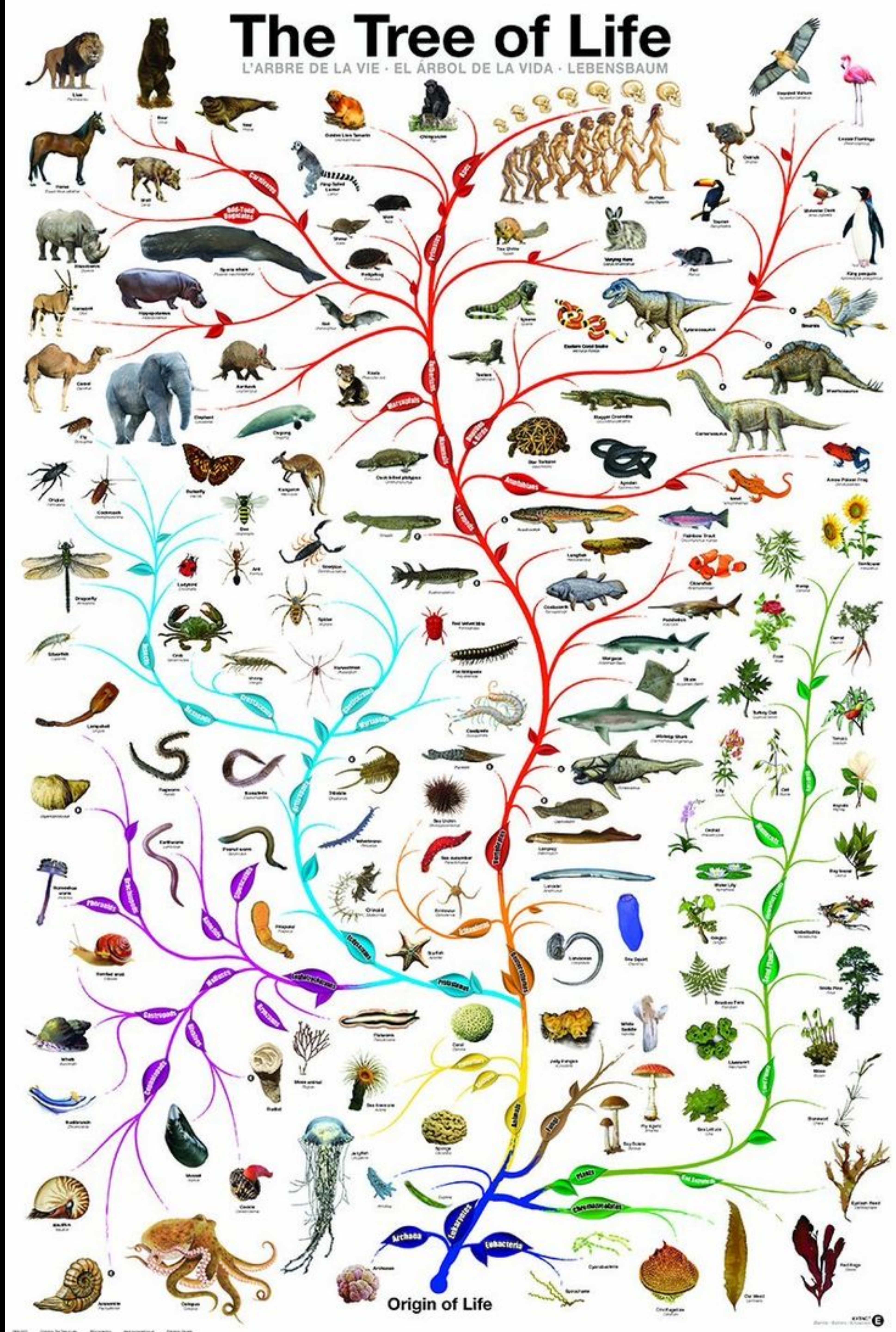
- QD algorithms
  - a diverse set of high-performing agents (policies)
  - goal-switching
  - conduct divergent search: find out what is possible within an environment
- But their ability to innovate is constrained
  - stuck in a single environment

MAP-Elites



Open-Ended Algorithms

- Endlessly innovate
- Examples
  - Natural evolution
  - Human culture (science, technology, art)
- Can we make algorithms that do this?



Traditional ML

- We pick challenges and solve them



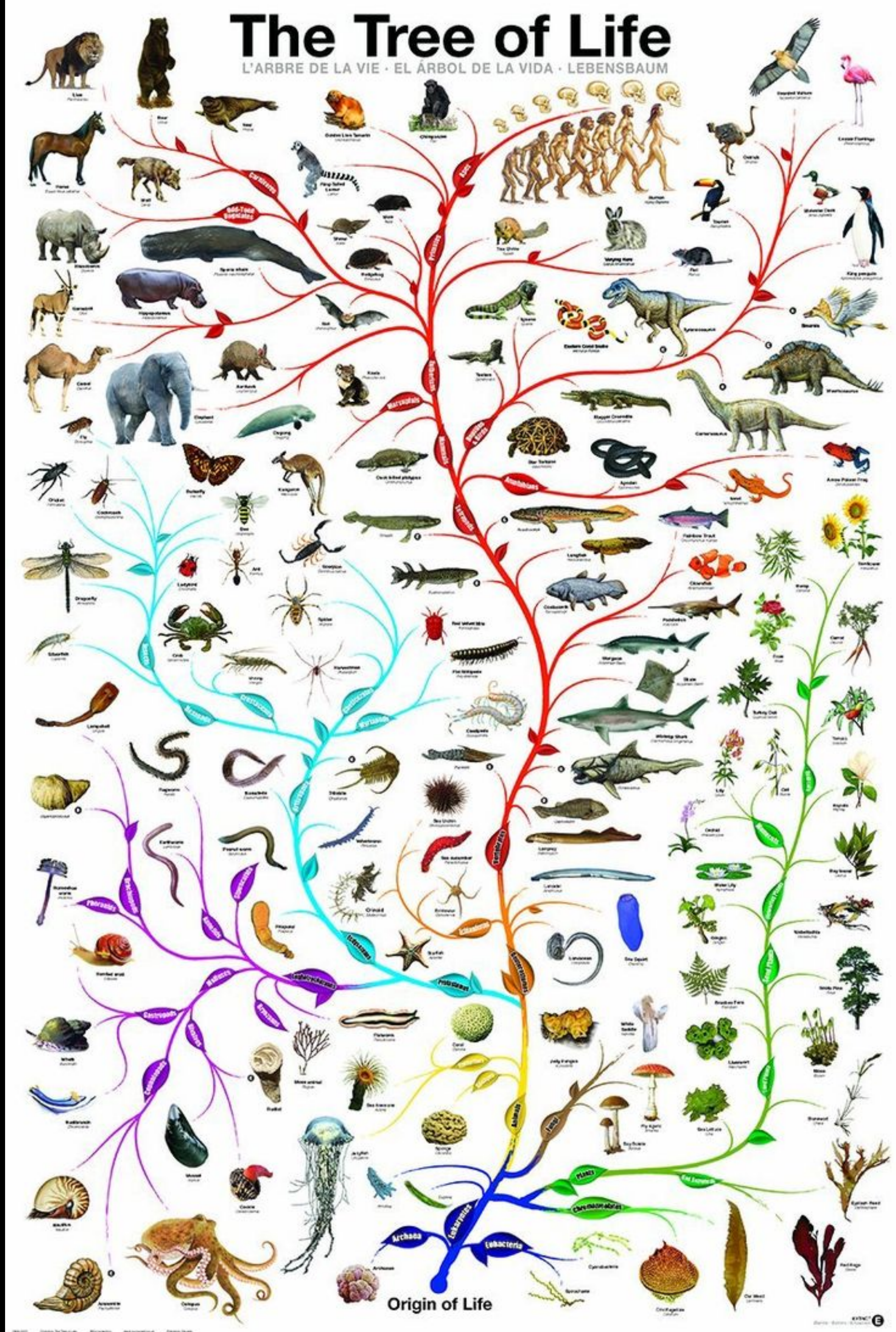

Starcraft

Dota 2

Intriguing Possibility

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

Interesting after a billion years?



Paired Open-Ended Trailblazer (POET)

Rui Wang

Joel Lehman

Jeff Clune*

Ken Stanley*

*Co-senior authors

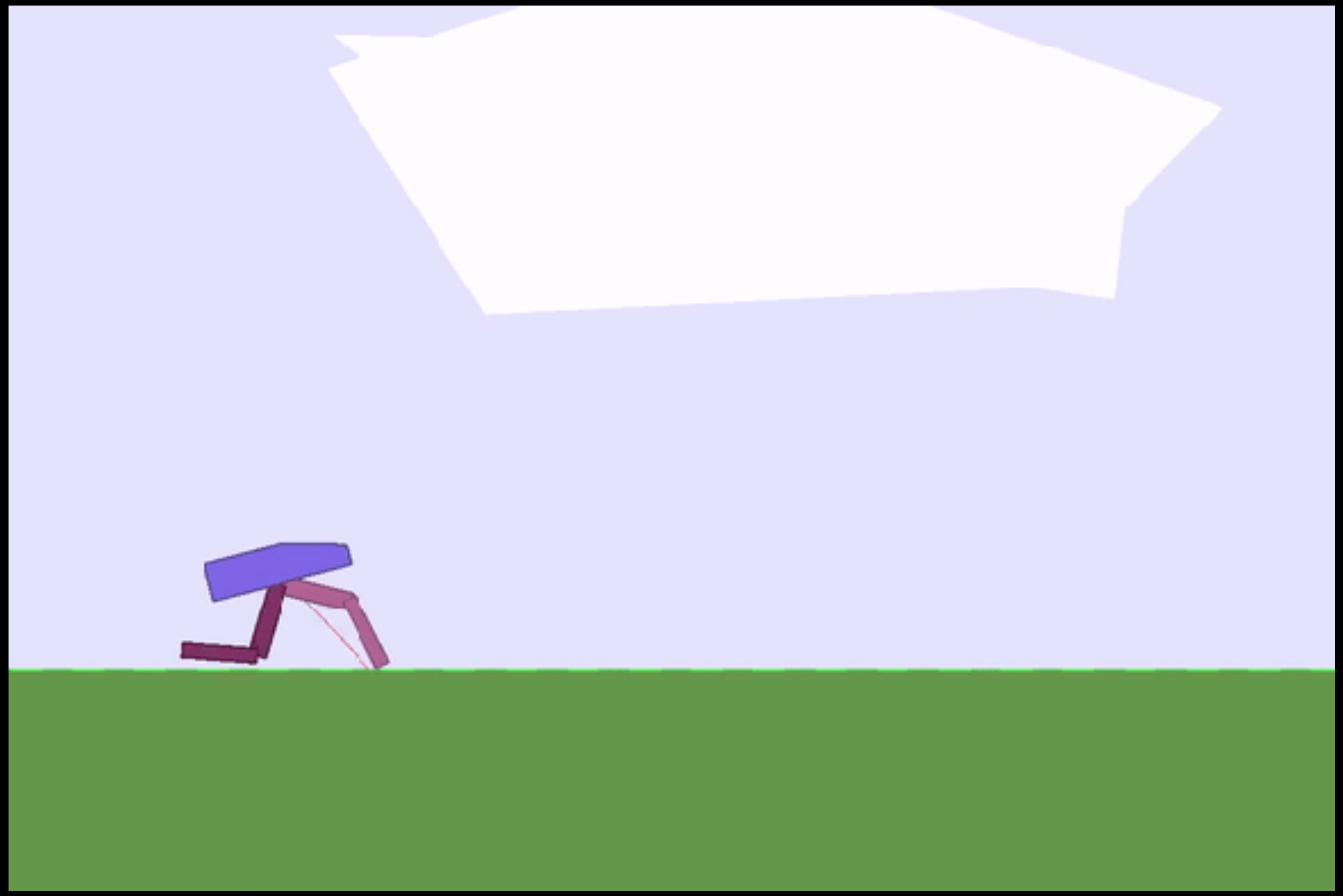
GECCO 2019

Endlessly Generating Increasingly Complex and Diverse Learning Environments and their Solutions

POET

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

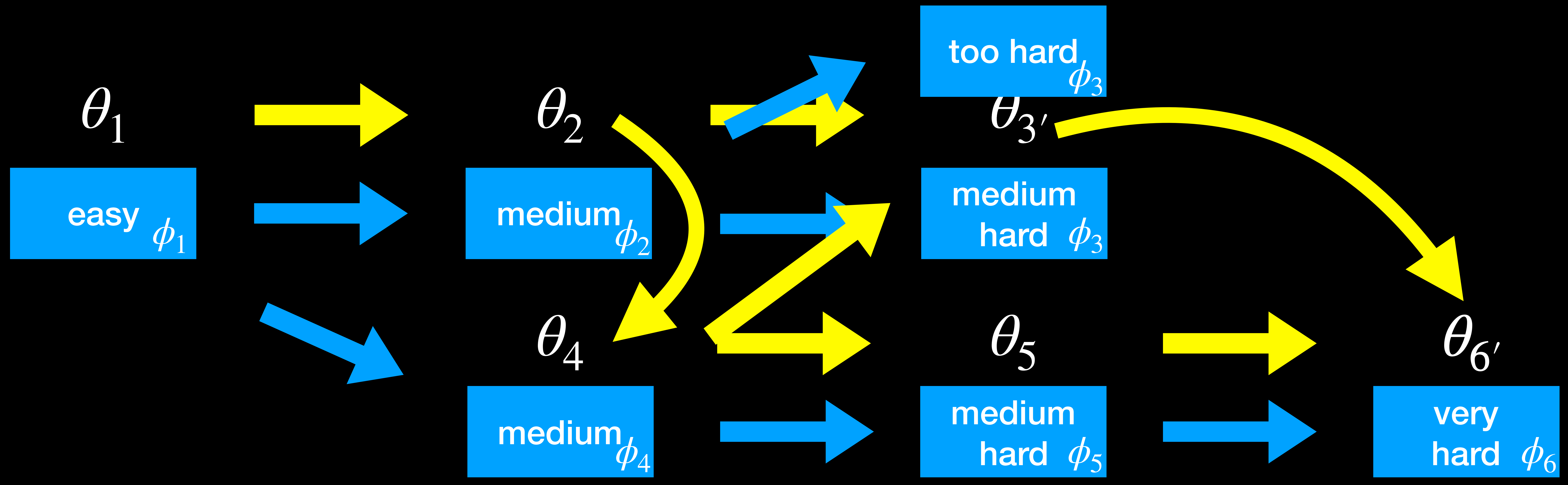
Task: Obstacle Courses



- Run fast
- Don't fall
- Reduce torque
- Balance hull

- 4 motors
- 10 LIDAR sensors
- 14 other proprioceptive state variables

POET



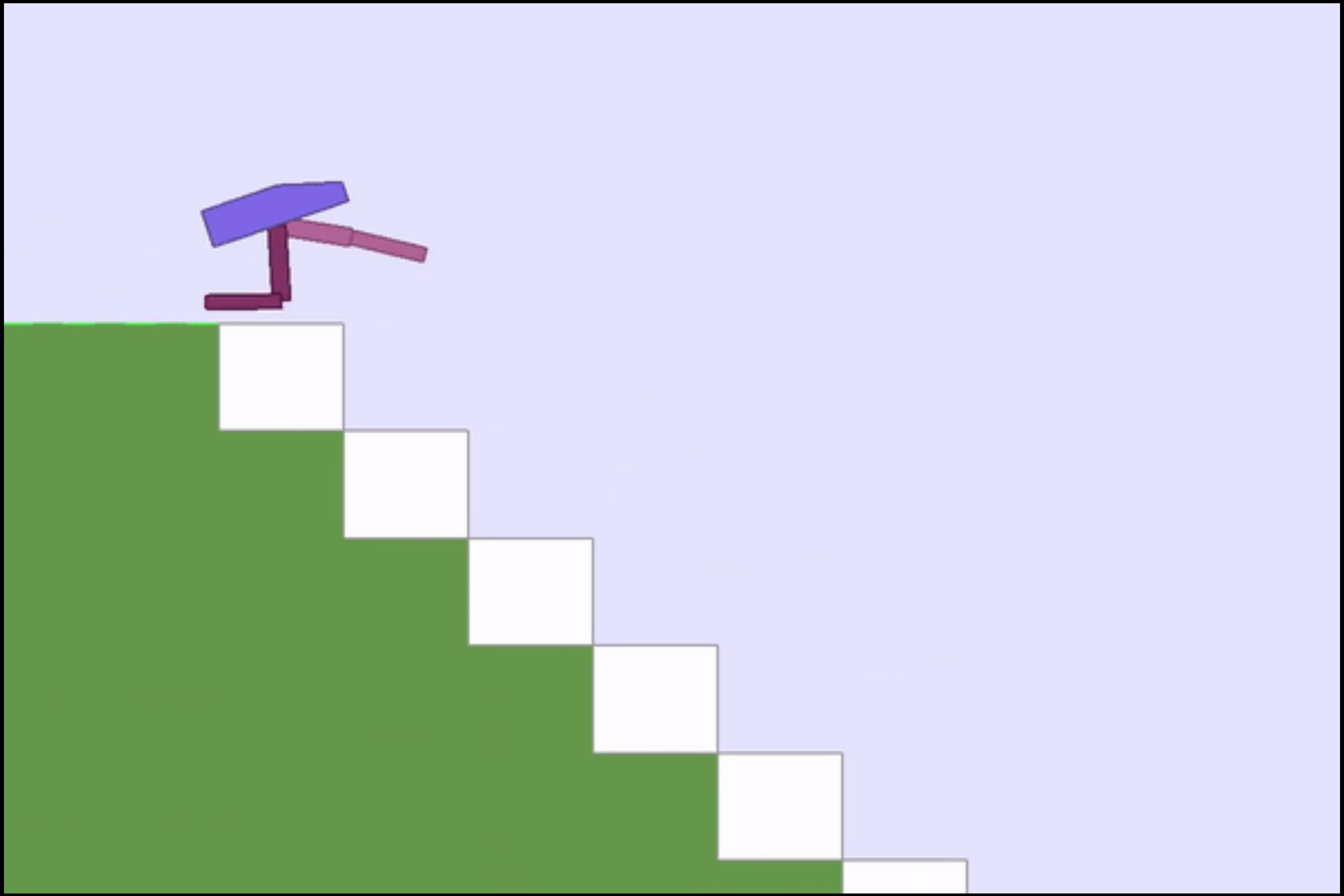
θ_1
easy ϕ_1
θ_2
medium ϕ_2
too hard ϕ_3
$\theta_{3'}$
medium hard ϕ_3
θ_4
medium ϕ_4
θ_5
medium hard ϕ_5
$\theta_{6'}$
very hard ϕ_6

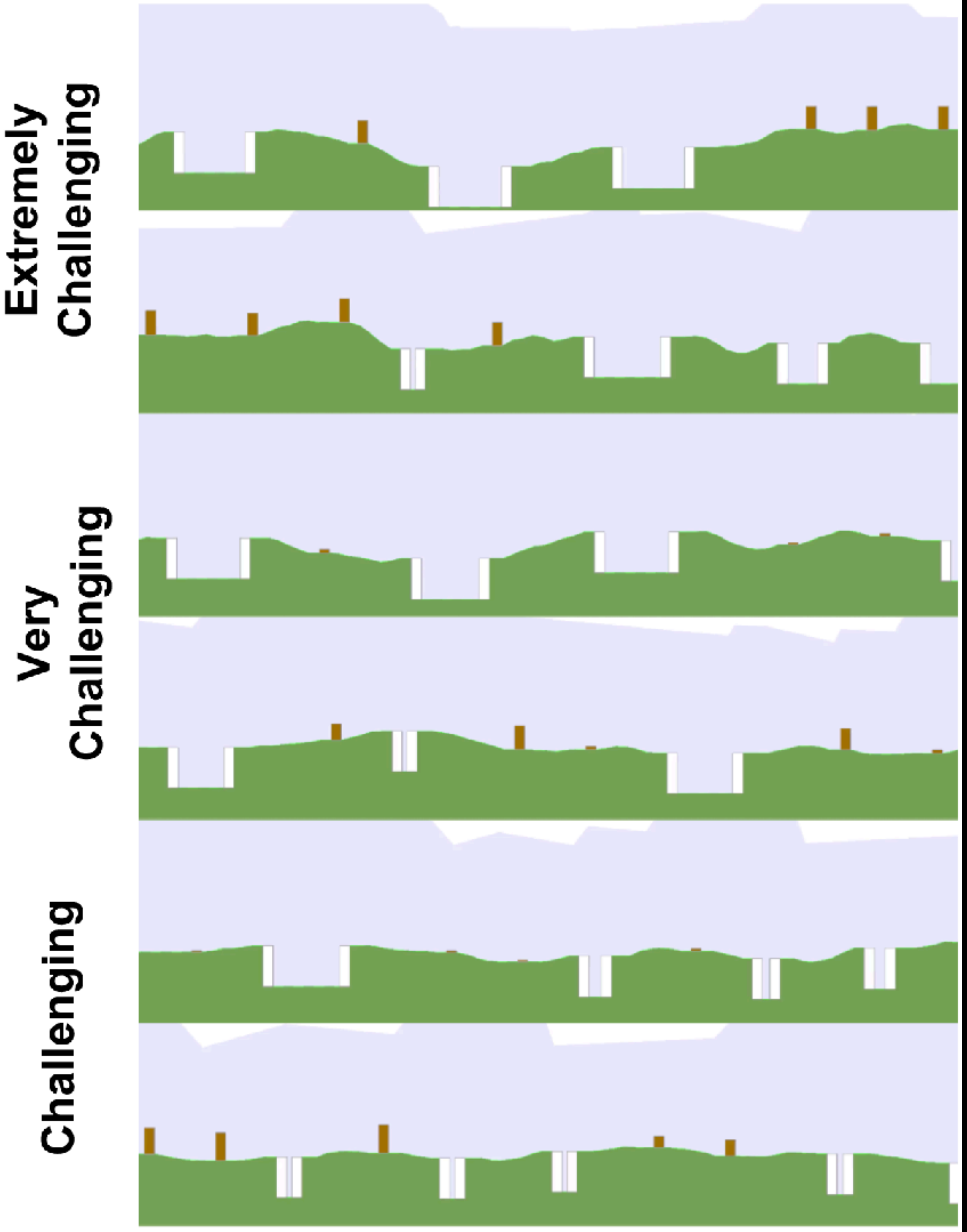
Methods

- 3-layer neural network optimized with ES (Salimans et al. 2017)
  - but any RL algorithm would work
- easy to distribute
- code: https://github.com/uber-research/poet

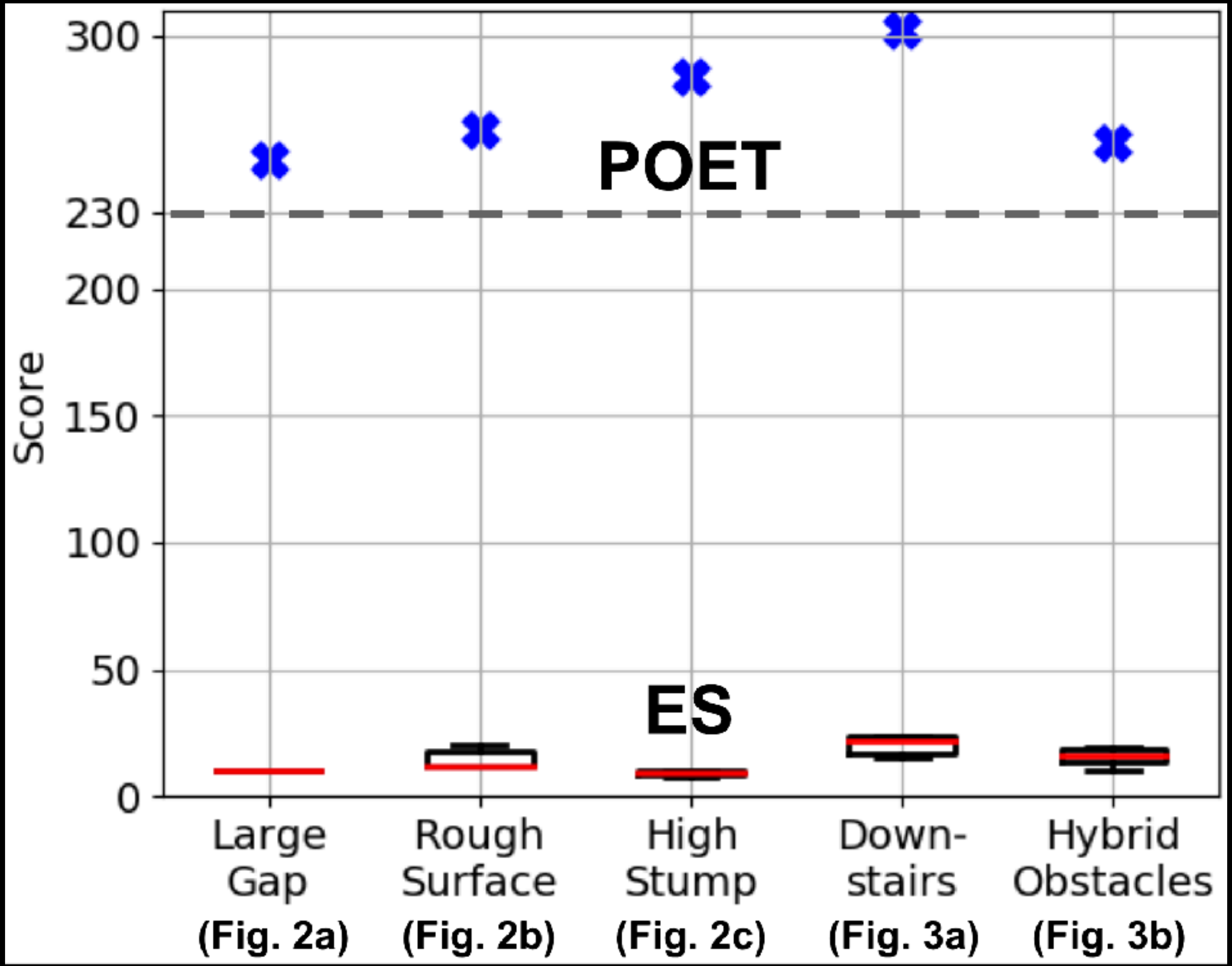

It starts off easy...

Another Challenging Environment



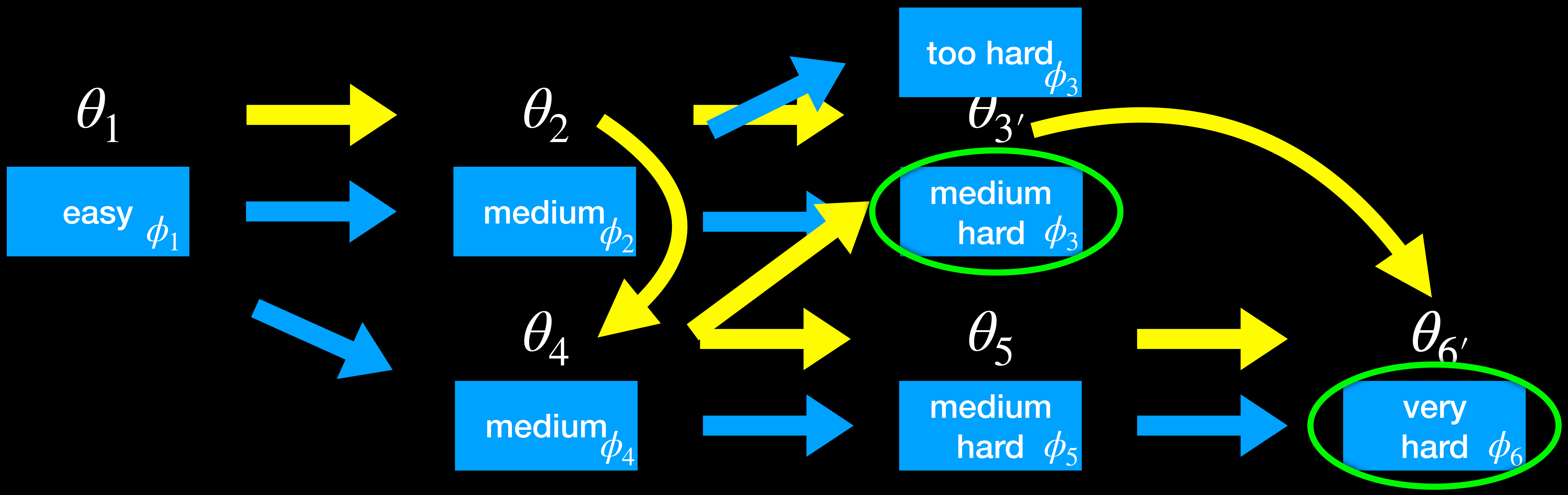

Extremely Challenging
Very Challenging
Challenging

Direct Optimization Fails



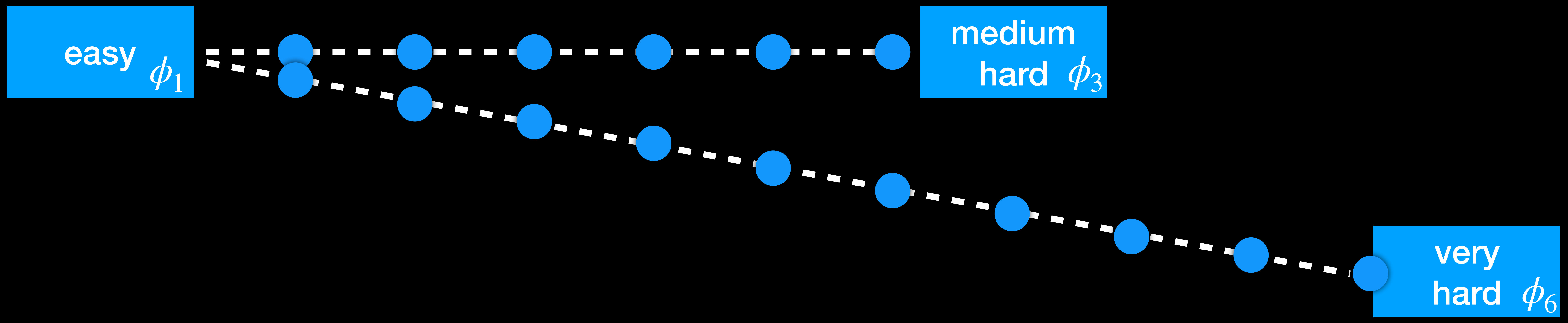
Direct Path Curriculum

- For each extremely challenging environment

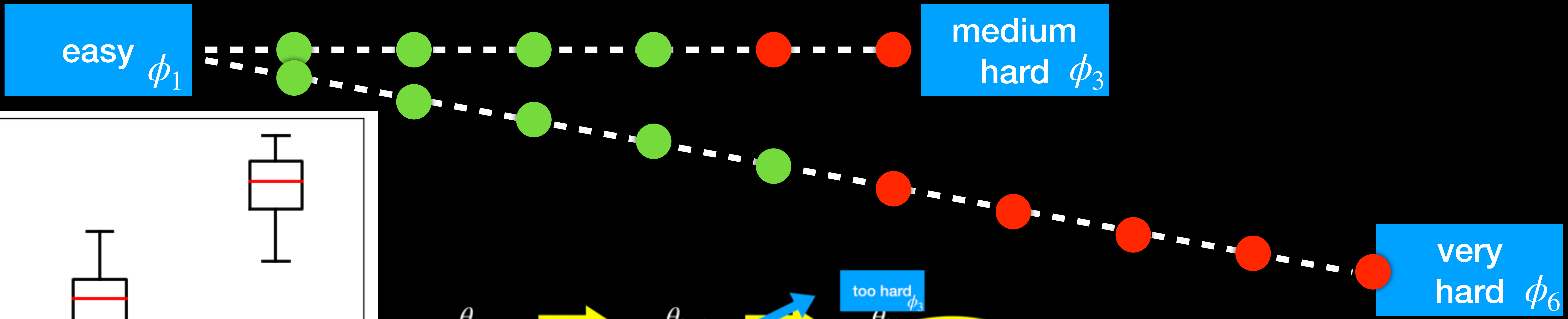


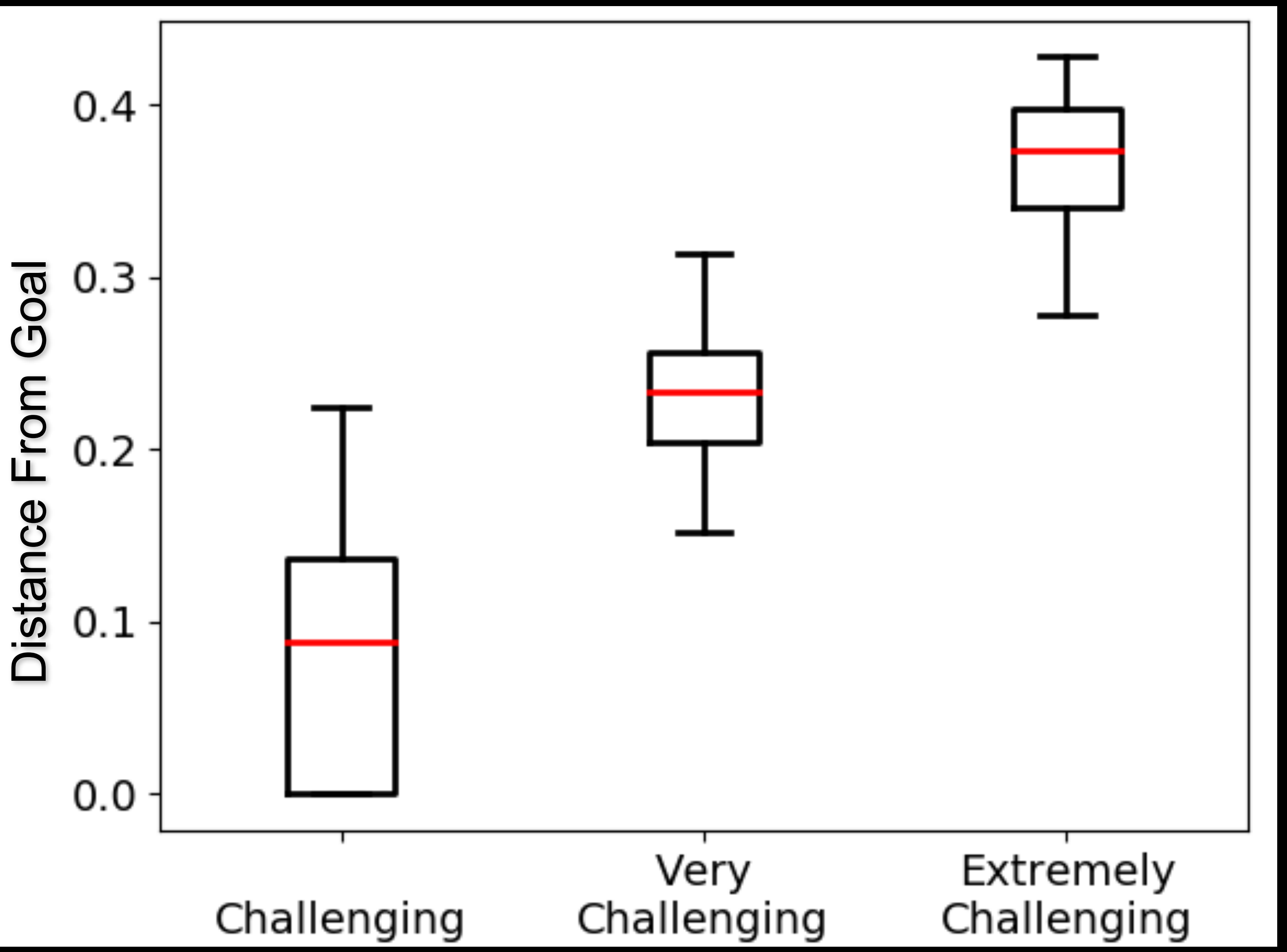
Direct Path Curriculum

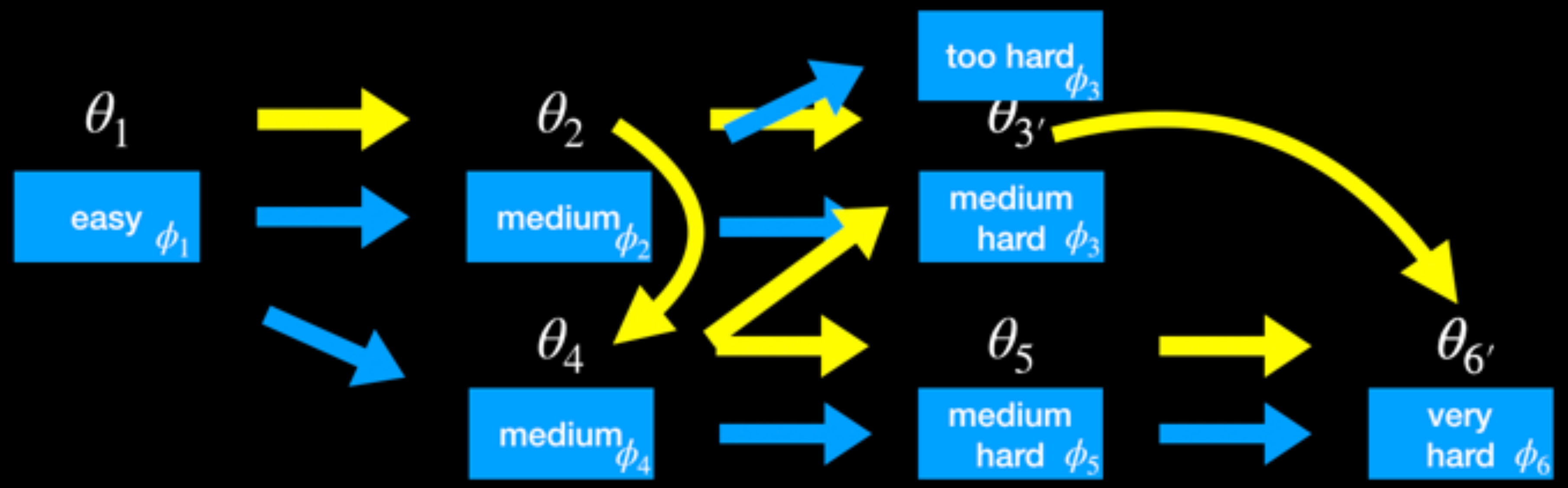
- For each extremely challenging environment
- Create intuitive curriculum (Bengio et al. 2009, Gomez et al. 1997, Karpathy et al. 2012)



Direct Path Curriculum Fails



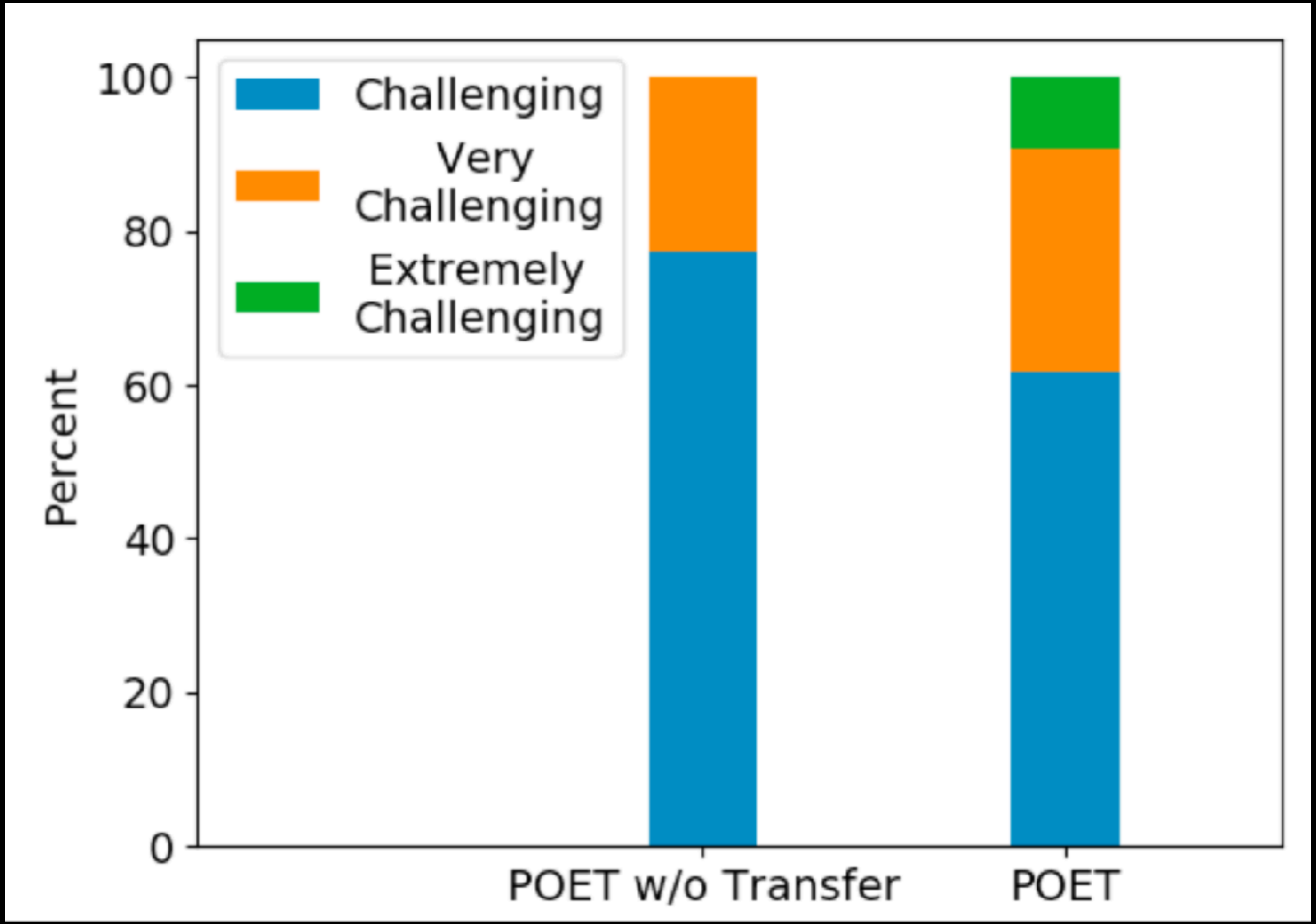




Goal Switching is Essential


298
207
304
311
349
309

Goal Switching Needed to Solve the Hardest Challenges

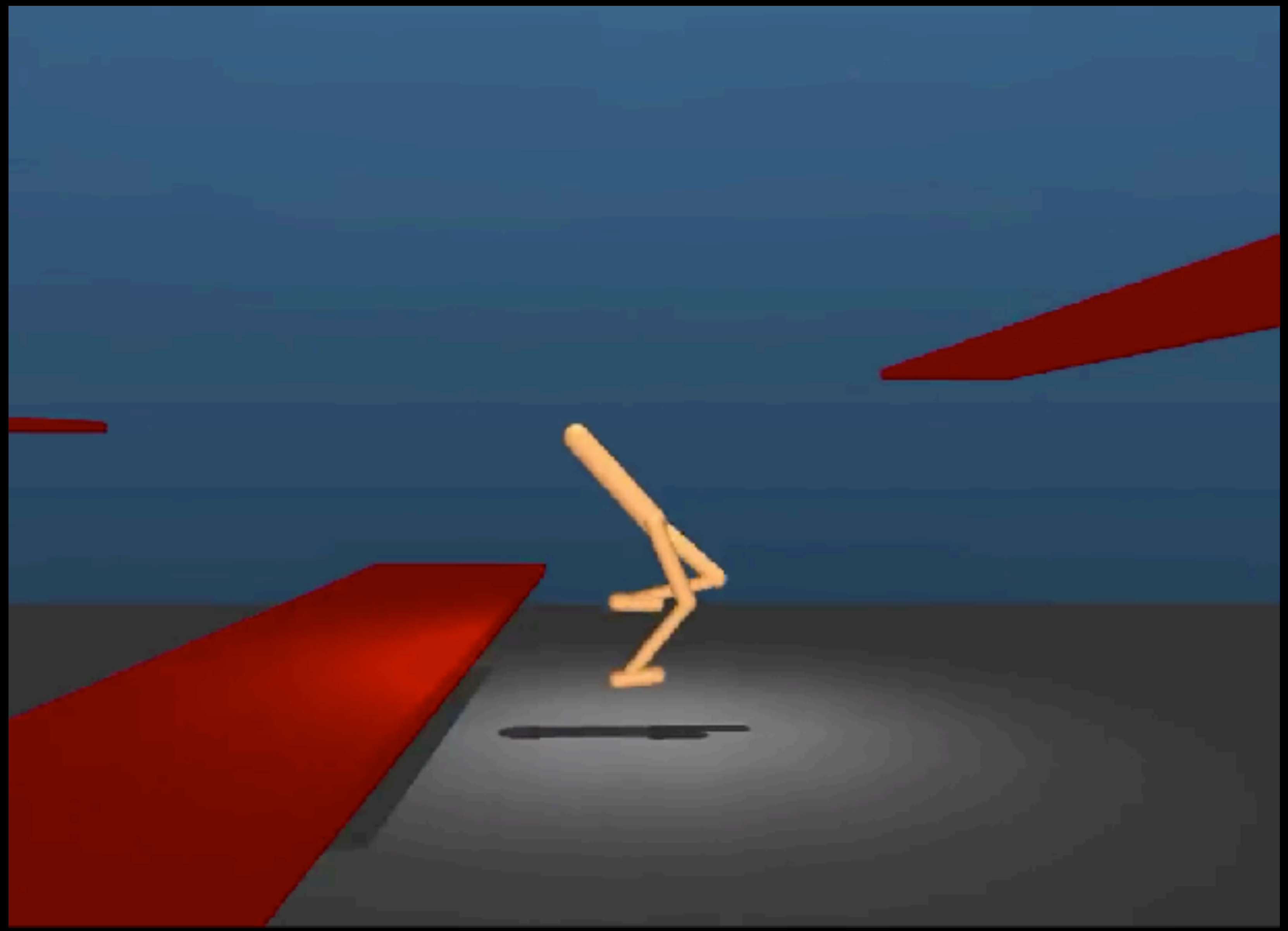


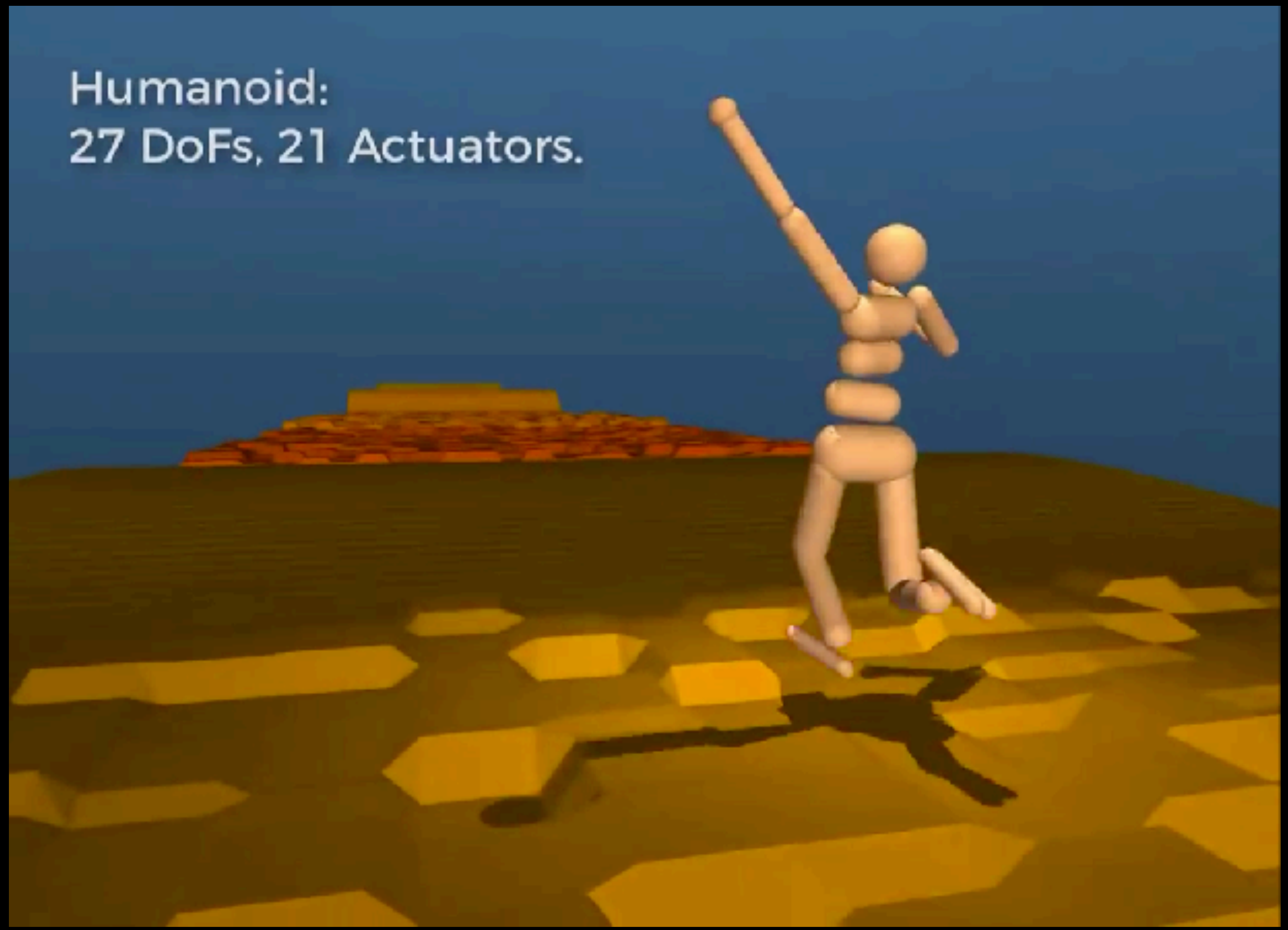
Work Related to Learning Curricula

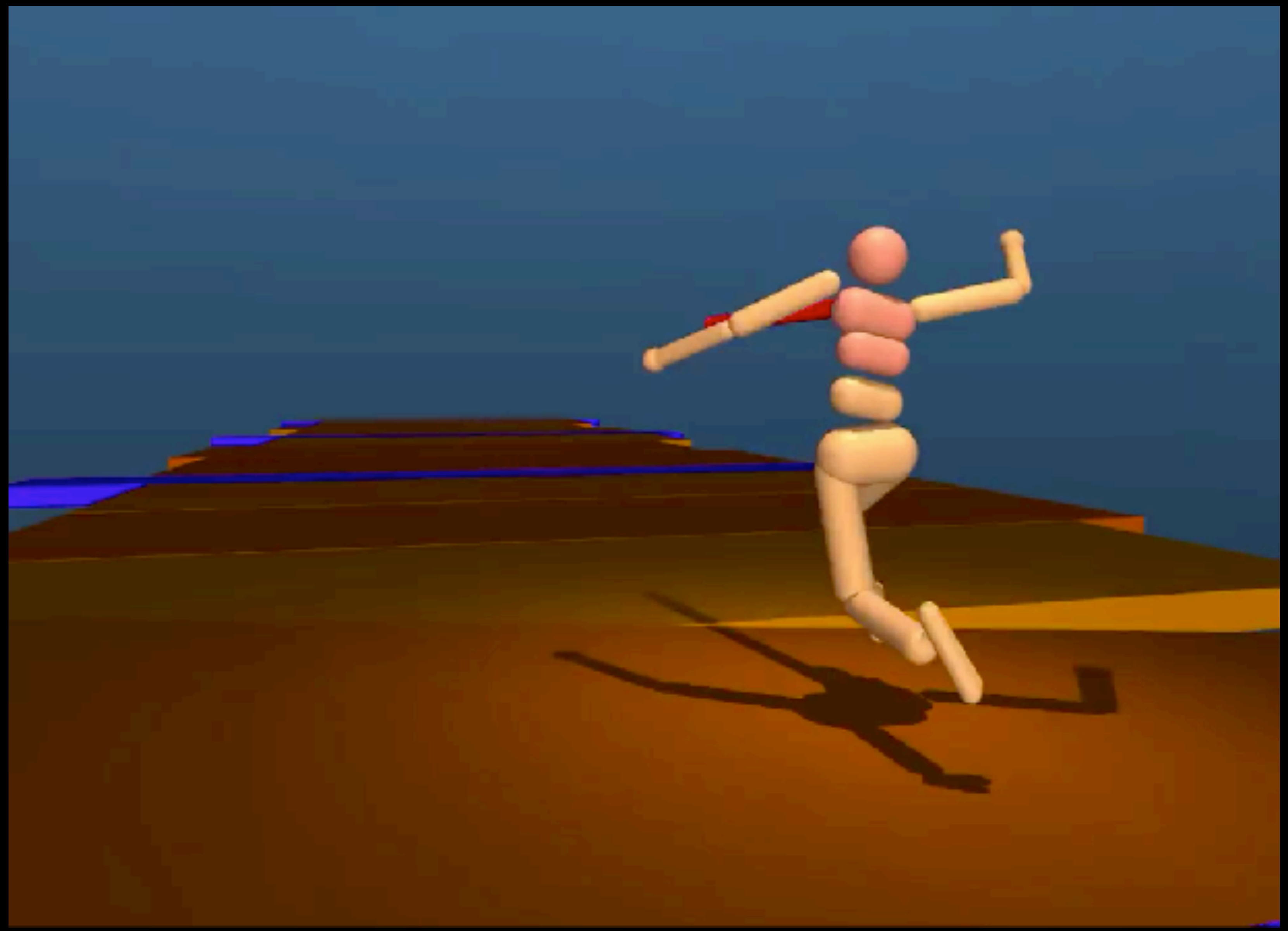
- See our paper for a full list + references https://arxiv.org/abs/1901.01753
  - Open-endedness
  - Procedural content generation
  - Quality diversity
  - Developmental robotics
  - Co-evolution/self-play
  - GANs
  - Minimal criterion co-evolution
  - Direct curriculum learning methods

Future Work

More complex environments (including encoding)

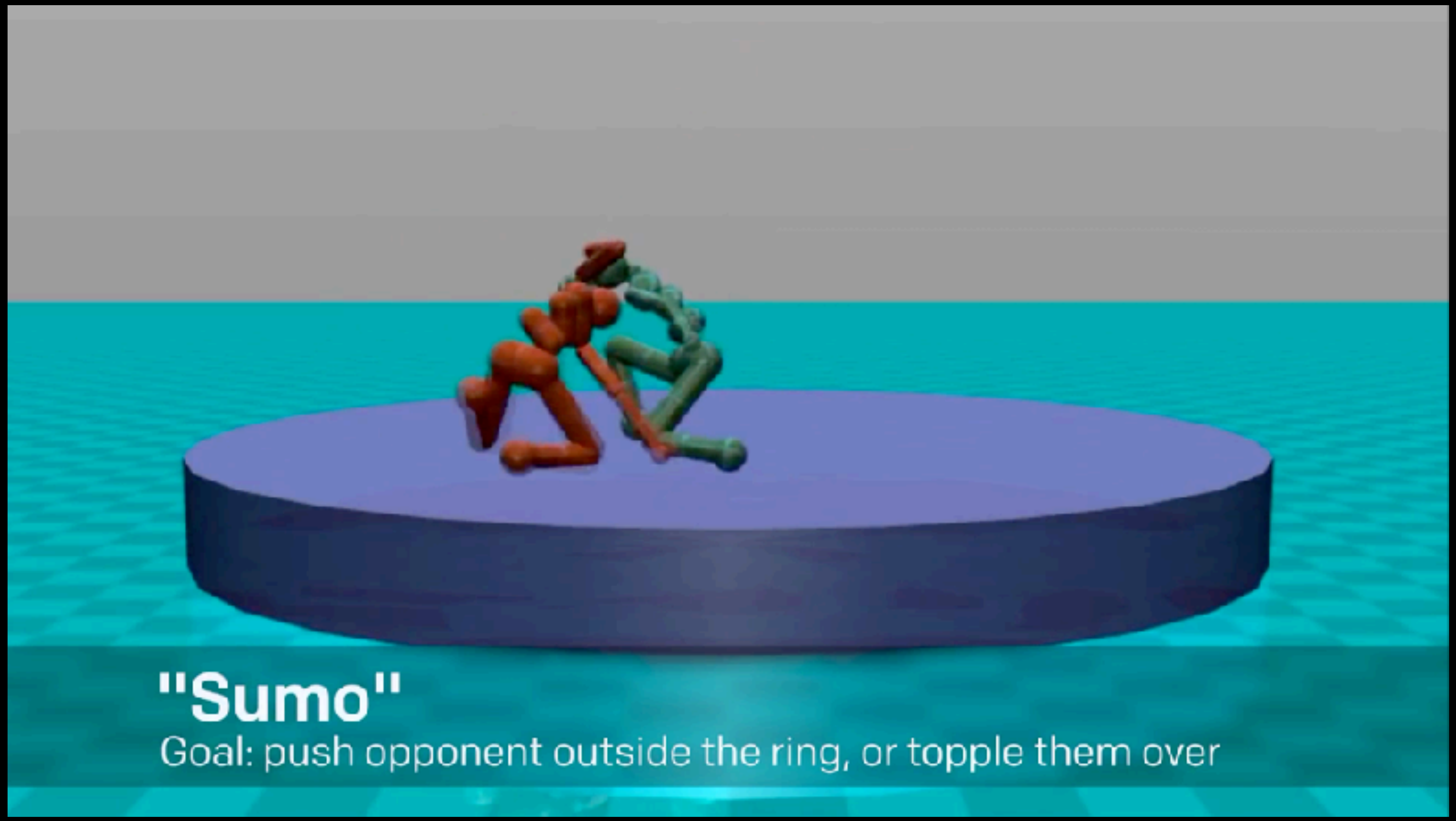






Heess et al. 2017

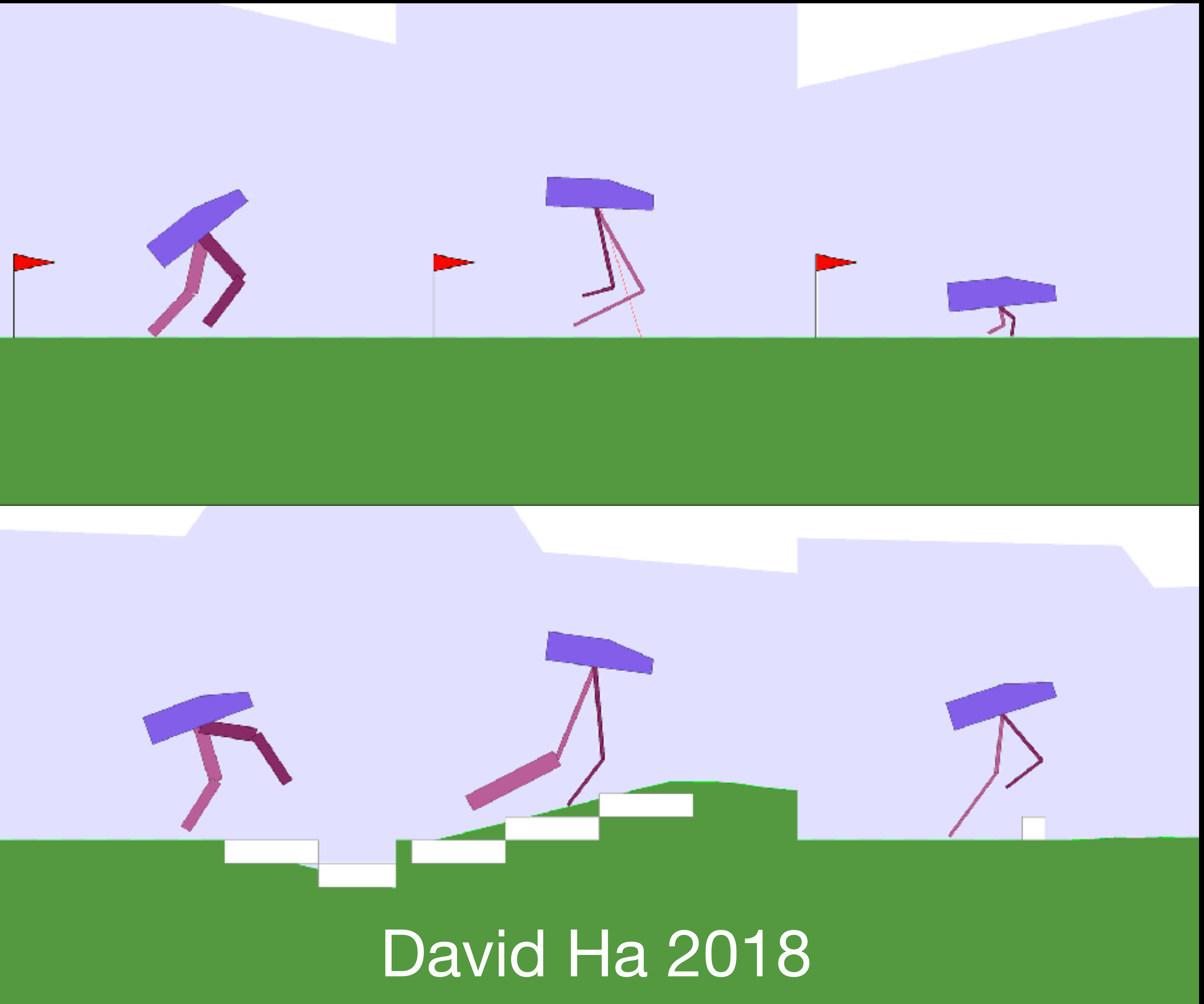
More complex environments (including encoding)



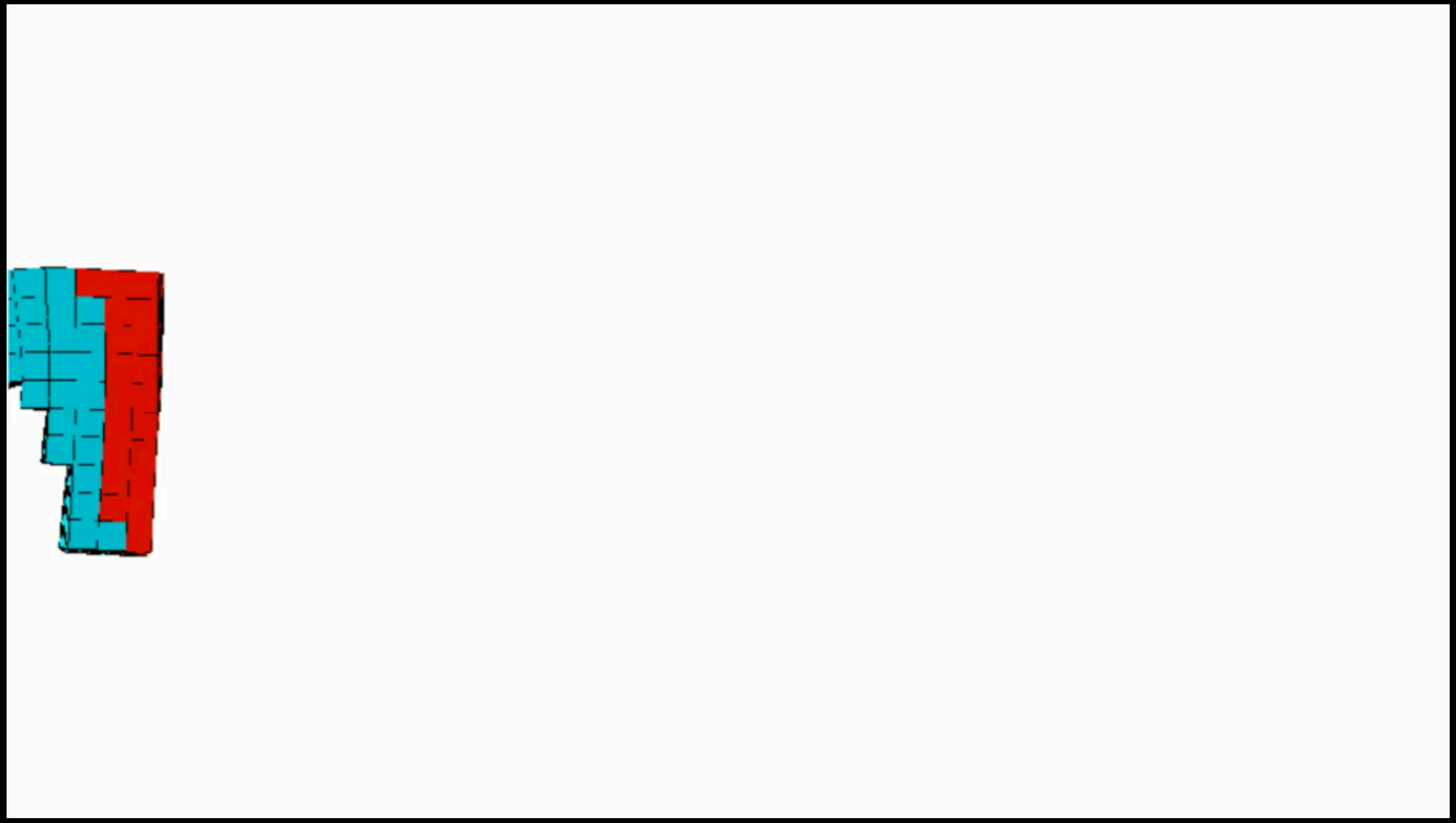
Bansal et al. 2017


"Assassin's Creed Odyssey" (2018)
Ubisoft
watchmojo

Optimize body



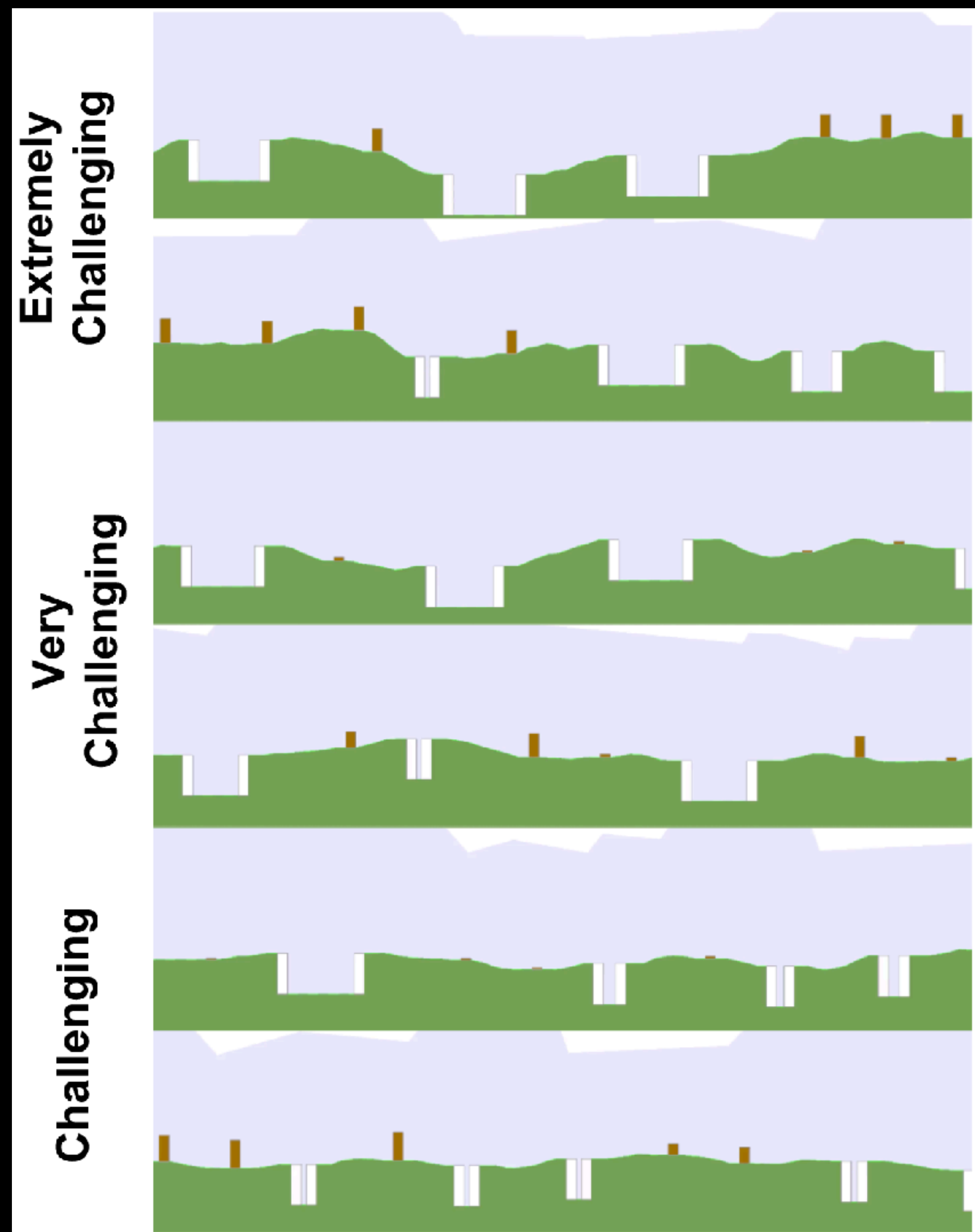
David Ha 2018



Cheney, MacCurdy, Clune, Lipson 2013

Future Work

- Generate reward function too
- More compute
  - wonderful property: more compute = more discoveries, complexity, skills, innovation, etc.
- Meta-Learning
  - requires distribution of tasks
  - requires humans to generate it (and do a good job)
  - could use POET-generated tasks instead



AI-Generating Algorithms

Clune 2019

Three Pillars

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

AI-Generating Algorithms

Clune 2019

- Learn as much as possible
- Bootstrap from simple to AGI
- Expensive outer loop
  - produces a sample-efficient, intelligent agent for inner loop
- We know it works
  - occurred on Earth

Automatically Generating Environments & Solutions

POET: Wang, Lehman, Clune, & Stanley 2019

- Invent effective curricula
  - key ingredients: create & collect stepping stones, harness goal switching
  - curricula are often very counterintuitive (e.g. harder tasks help solve simpler ones)
    - explains why goal-oriented attempts fail
  - hedges bets with multiple, overlapping curricula
- Endlessly innovates
- May be the only way to
  - solve ambitious problems
  - discover the full gamut of what is possible
- Captures spirit of open-ended innovation engines
  - Natural evolution, Cultural evolution (science, technology, art)
- Opens many exciting future research directions

For more, watch our ICML 2019 Tutorial on Population-Based Methods
https://youtu.be/g6HiuEnbwJE

Thanks!

Paired Open-Ended Trailblazer (POET)

Rui Wang

Joel Lehman

Jeff Clune*

Ken Stanley*

*Co-senior authors

UBER AI Labs

UNIVERSITY OF WYOMING

Automatically Generating Environments & Solutions

POET: Wang, Lehman, Clune, & Stanley 2019

- Invent effective curricula
  - key ingredients: create & collect stepping stones, harness goal switching
  - curricula are often very counterintuitive (e.g. harder tasks help solve simpler ones)
    - explains why goal-oriented attempts fail
  - hedges bets with multiple, overlapping curricula
- Endlessly innovates
- May be the only way to
  - solve ambitious problems
  - discover the full gamut of what is possible
- Captures spirit of open-ended innovation engines
  - Natural evolution, Cultural evolution (science, technology, art)
- Opens many exciting future research directions

For more, watch our ICML 2019 Tutorial on Population-Based Methods
https://youtu.be/g6HiuEnbwJE