Decentralized Neural Networks

Rich Sutton
Keen Technologies
University of Alberta
Alberta Machine Intelligence Institute
Reinforcement Learning and Artificial Intelligence Lab
Openmind Research Institute

Artificial intelligence research is properly ambitious

- AI researchers seek to understand intelligence well enough to create beings of greater intelligence than current humans
- Reaching this profound intellectual milestone will enrich our economies and challenge our societal institutions
  - It will be unprecedented and transformational, but also a continuation of trends that are thousands of years old
- People have always created tools and been changed by them; it’s what humans do
- The next big step is to understand ourselves
- This is a quest grand and glorious, and quintessentially human

My perspective

- The greatest impacts and advances in AI are still to come
  - If AI is a race, it's not a sprint. It's a marathon
- The creation of super-intelligent agents, and super-intelligent augmented humans, will be an unalloyed good for the world
- The path to intelligent agents runs through reinforcement learning
- The biggest bottleneck to ambitious AI is inadequate deep learning algorithms

Decentralized Neural Networks

Rich Sutton
Keen Technologies
University of Alberta
Alberta Machine Intelligence Institute
Reinforcement Learning and Artificial Intelligence Lab
Openmind Research Institute

Conclusions

- All is not well with today's deep learning and artificial neural networks
  - They forget catastrophically, lose plasticity, and collapse under extended training
- To achieve the full potential of DL and ANNs, something more is needed
  - We need an additional source of variation in under-utilized artificial neurons
  - We need to protect and preserve neurons whose variations are found useful
- To me, these needs strongly suggest that neurons should have the decentralized goal of getting other neurons to listen to them
- Though I can't demonstrate this yet

Outline

- The idea of neurons that have goals and want to connect and contribute
- New evidence of problems with conventional deep learning (Nature 2024)
    - Deep learning loses plasticity in continual supervised learning
    - Deep learning collapses with prolonged reinforcement learning
- These problems are solved by variation and selective survival (i.e., by decentralized goals)
- Normalization and step-size optimization can also help by enabling online streaming algorithms (Elsayed et. al 2024) and can be seen as decentralization

The definition that I will use in this talk:

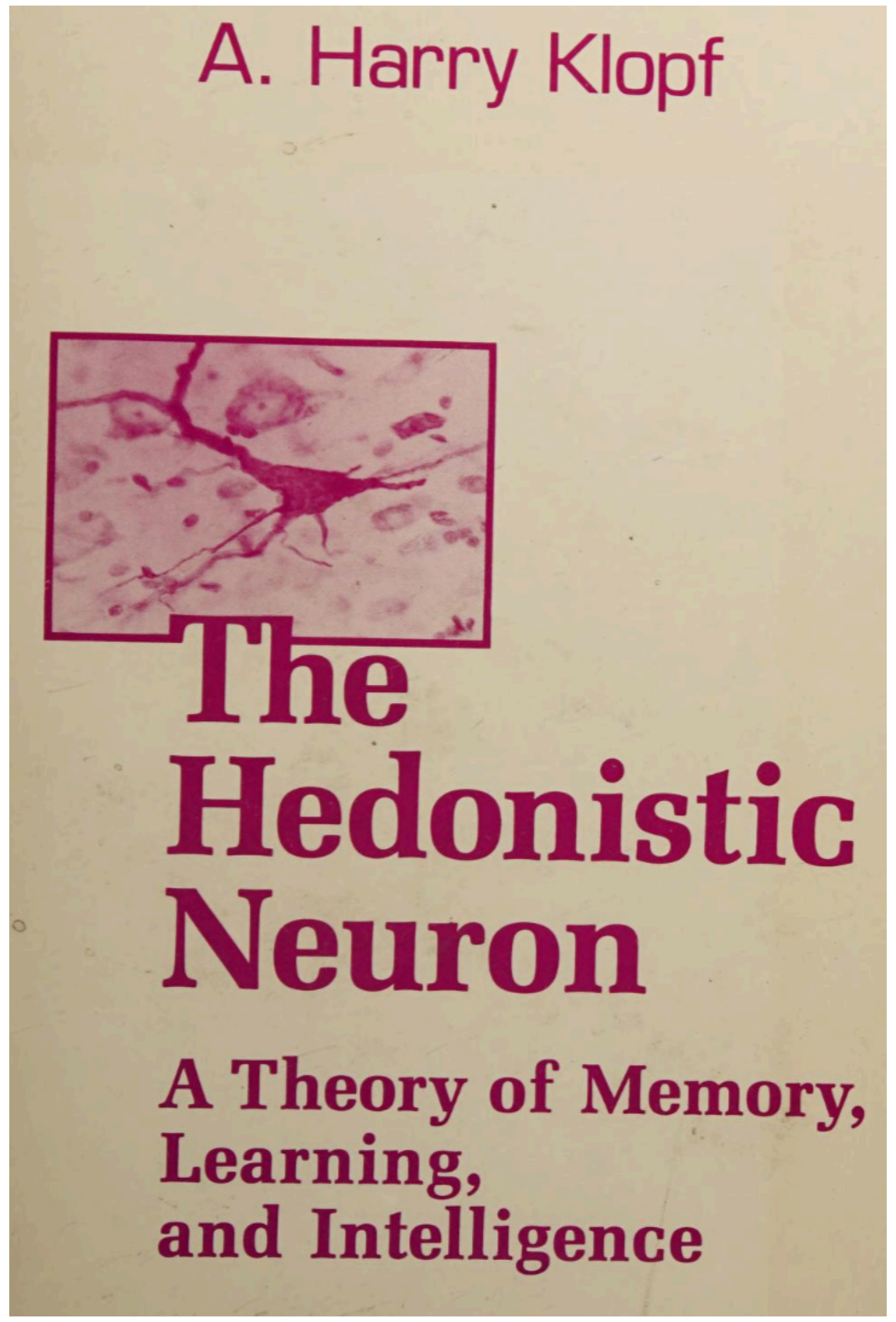
A *decentralized* neural network is one whose neurons seek their own goals distinct from the goals of the network as a whole

- For example, the overall network might seek to maximize reward, or to classify images as instructed by a training set
    - while individual neurons might have the goal
        - of providing signals that other neurons find useful
        - of being active at least 10% of the time
- A decentralized neural network is a “goal-seeking system made from goal-seeking components”

Modern reinforcement learning was originally conceived of as decentralized neural networks

"the grandfather of
modern reinforcement learning"

A. Harry Klopf (1941–1997)
Senior scientist with the
Avionics Directorate of the
Air Force Office of Scientific Research



1972, 1982

Klopf viewed neurons in a brain as goal-seeking agents, anologous to people in a society

- Each was a "hedonist" that sought to maximize a local analog of pleasure (reward)

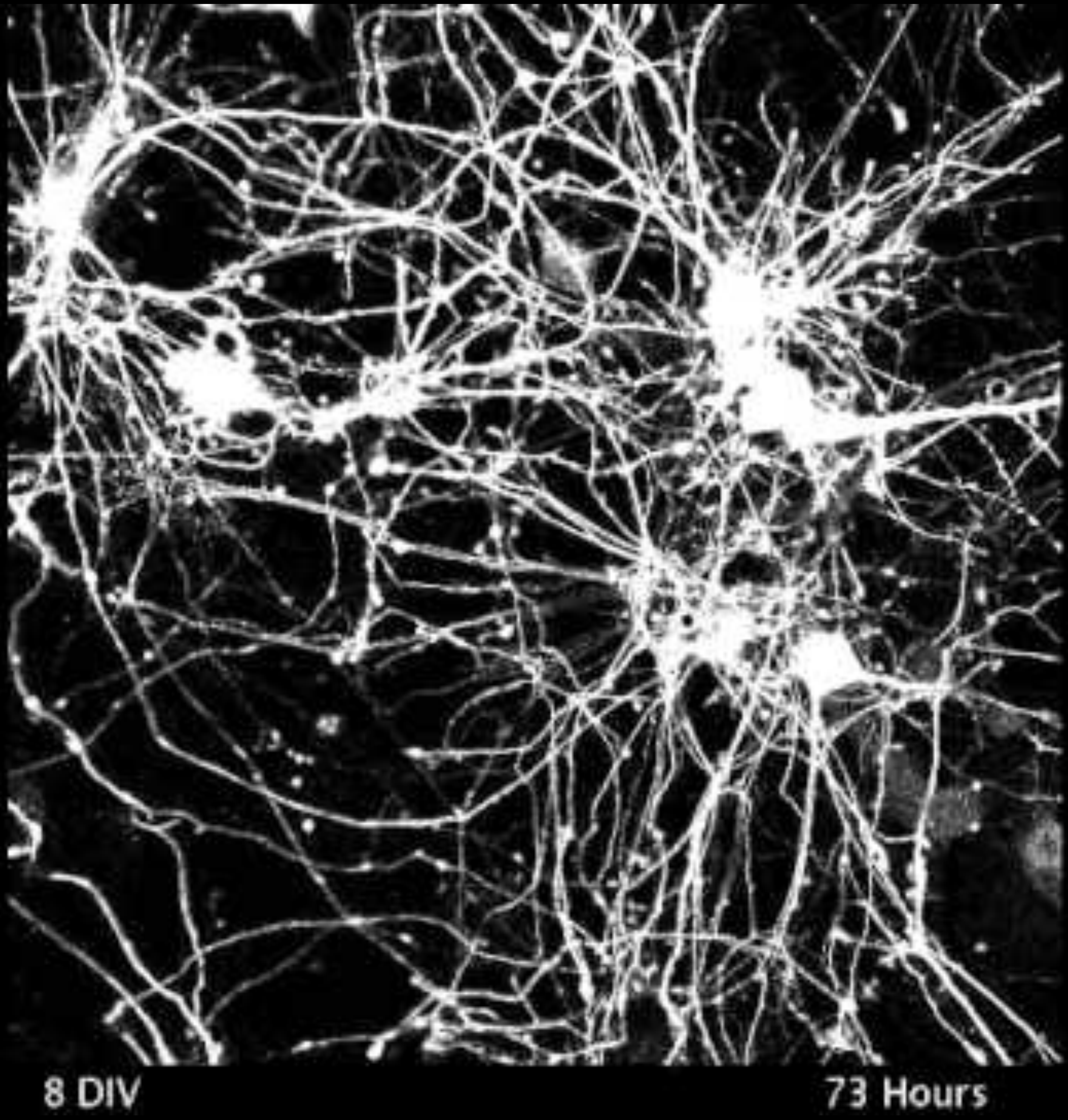
"Goal-seeking sytems from goal-seeking components"

- This decentralized perspective was otherwise absent from early work in cybernetics/neural networks

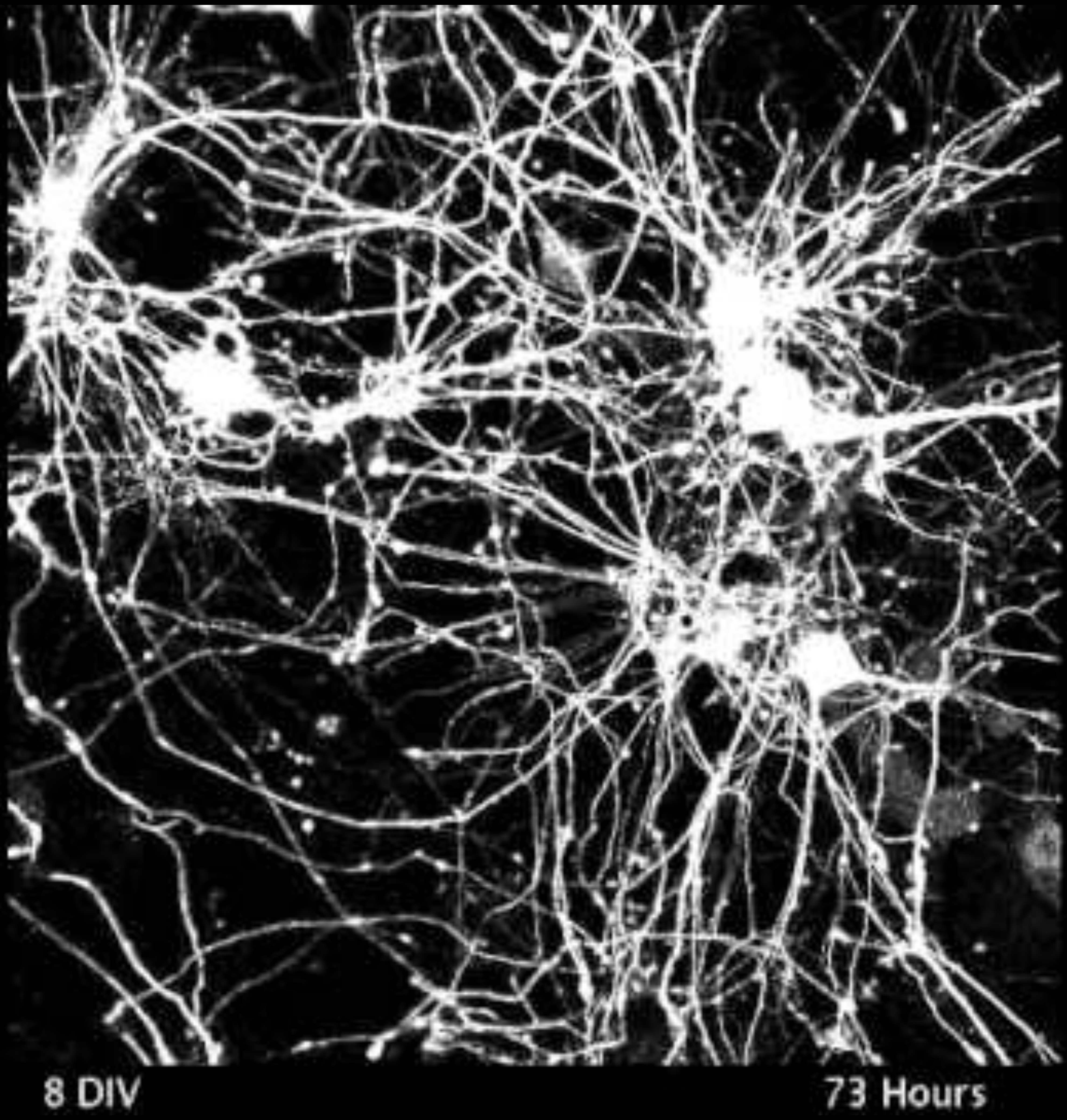
Klopf's ideas led directly to the reinforcement learning research of Sutton & Barto

- Klopf also enabled their Air Force funding

Neurons are active
and appear to seek out
connections to other neurons



Neurons are active
and appear to seek out
connections to other neurons



A decentralized neural network should adapt at 3 levels

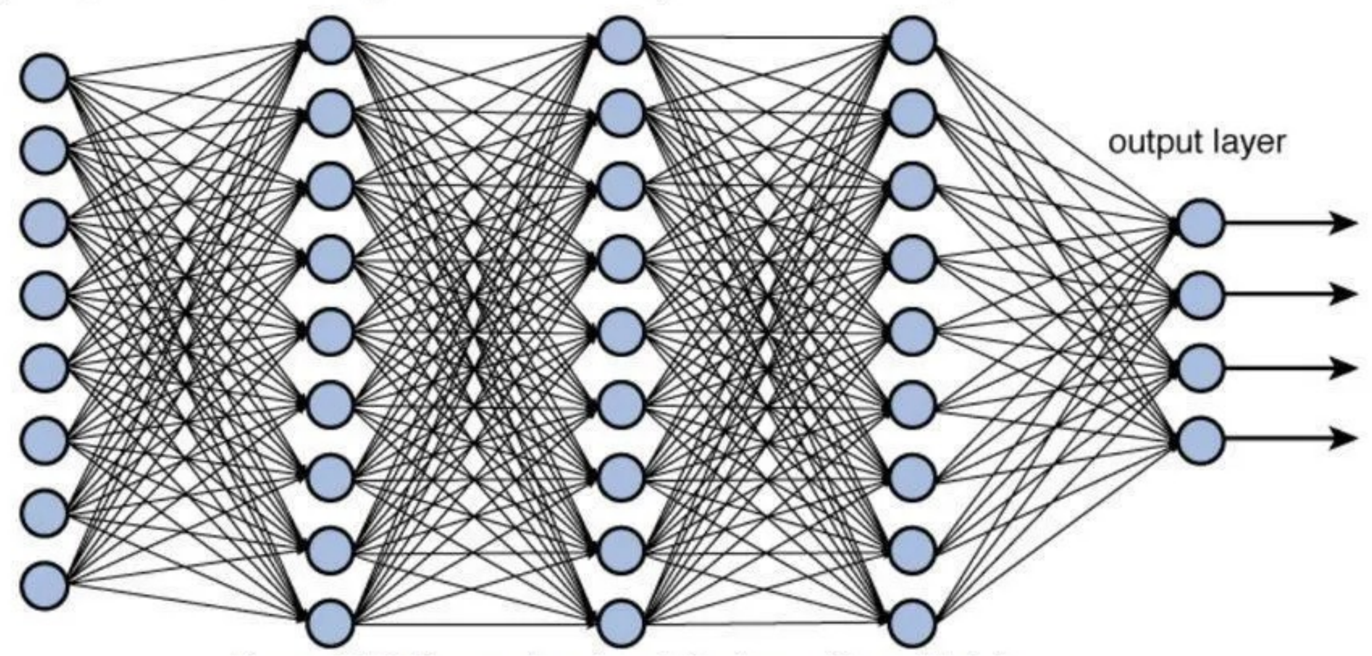
1. Adapting the wires and connections between artificial neurons

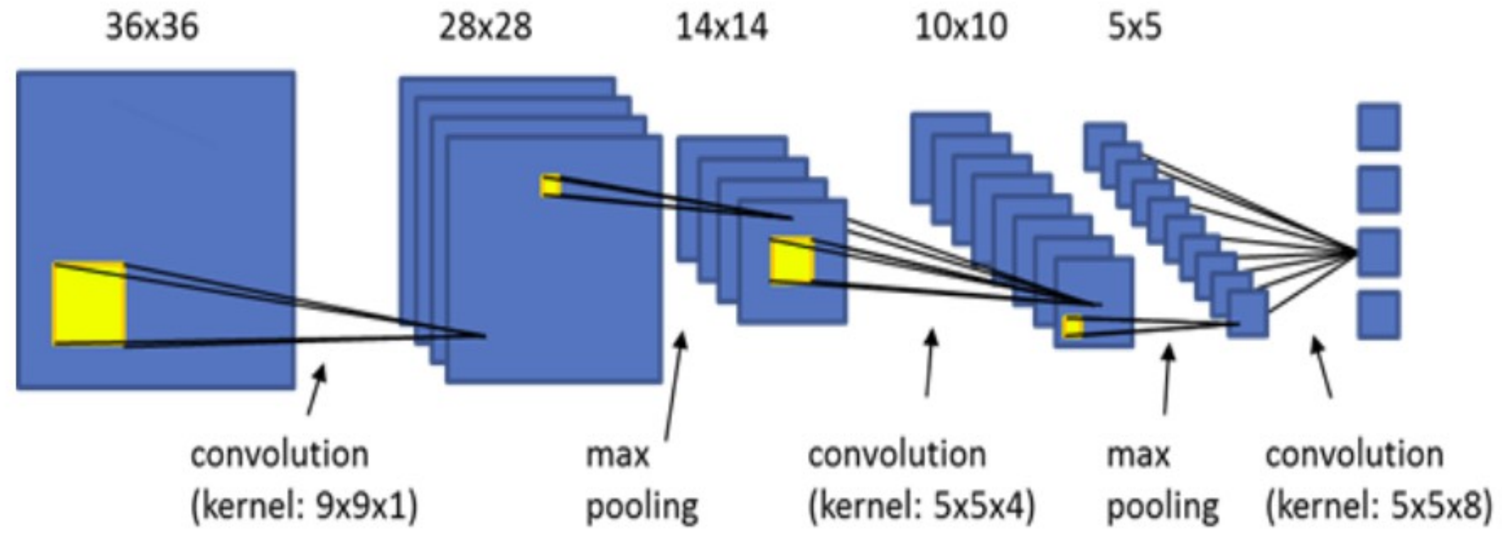
2. Adapting the weights

3. Adapting the step-size parameters

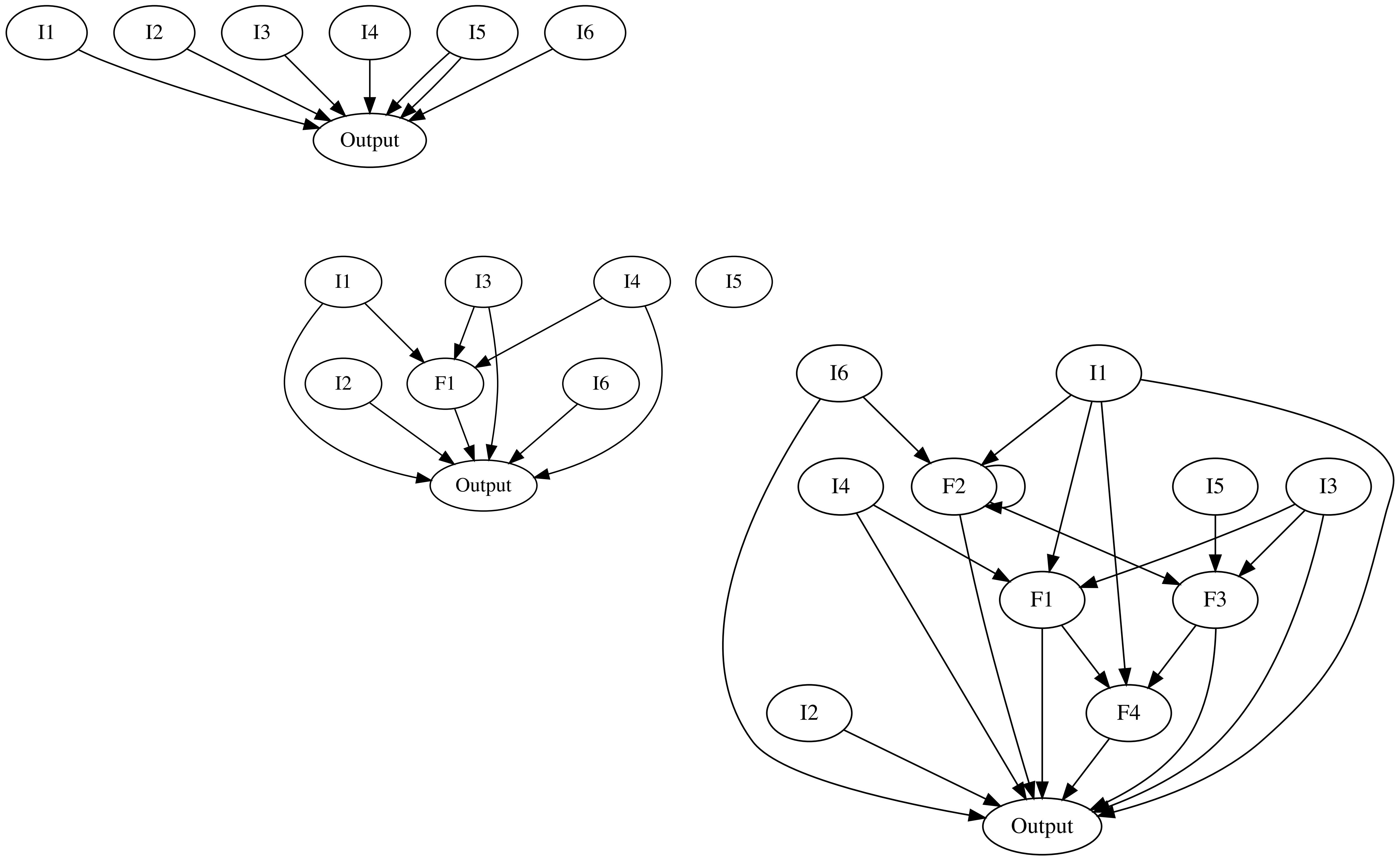
* Sutton, R.S., "Adapting Bias by Gradient Descent: An Incremental Version of Delta-Bar-Delta," ICML 1992.

Conventional DL networks have a fixed, designed structure





__Decentralized DL networks__ might be accumulated neuron by neuron



I will not be giving a specific algorithm for how the network is grown

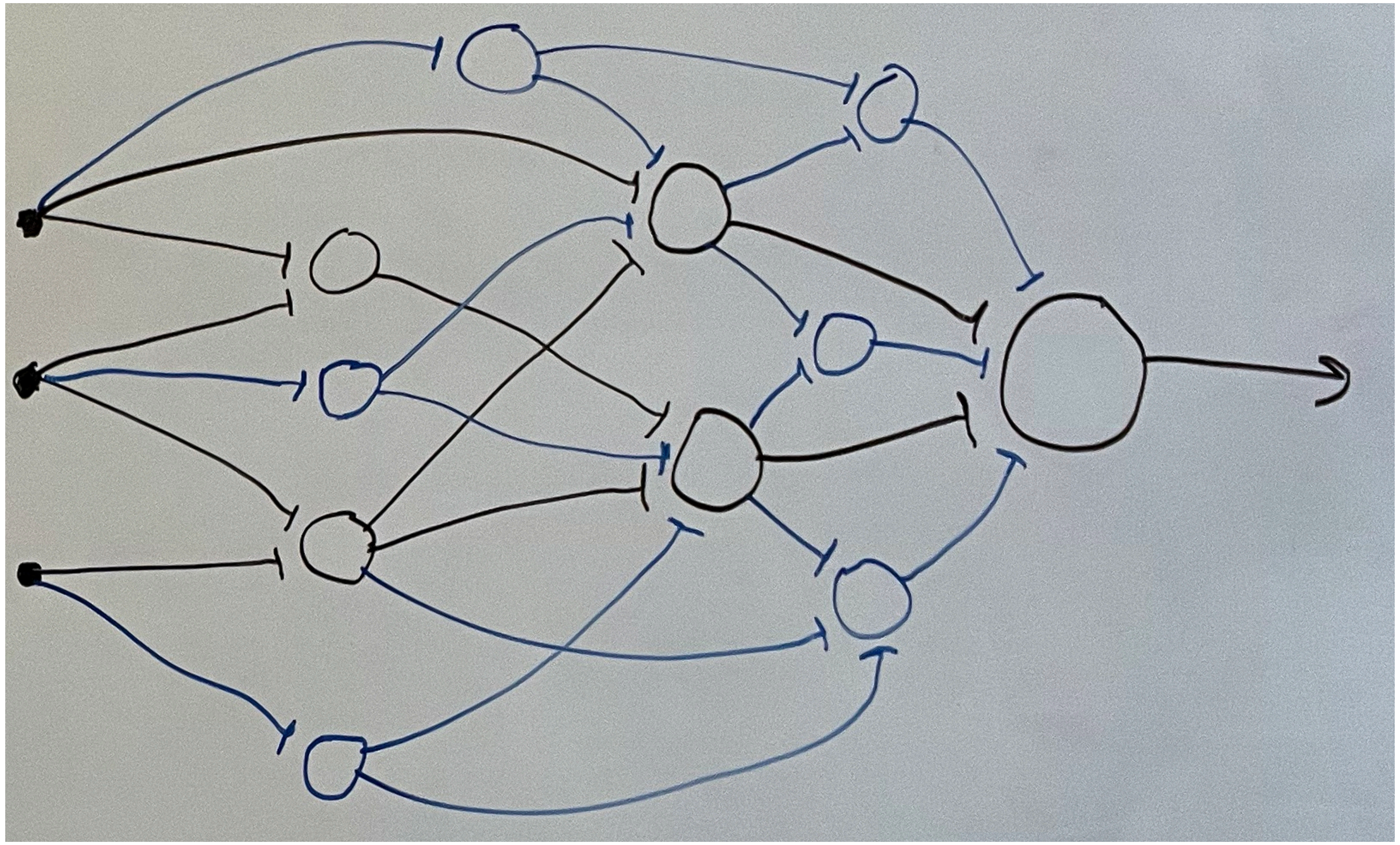
The first and most important algorithmic idea:

Distinguish the part of the network that has already been learned (the 'backbone') from the rest of the network (the 'fringe')

Preserve and protect the backbone; let the fringe explore

The *backbone* of a network

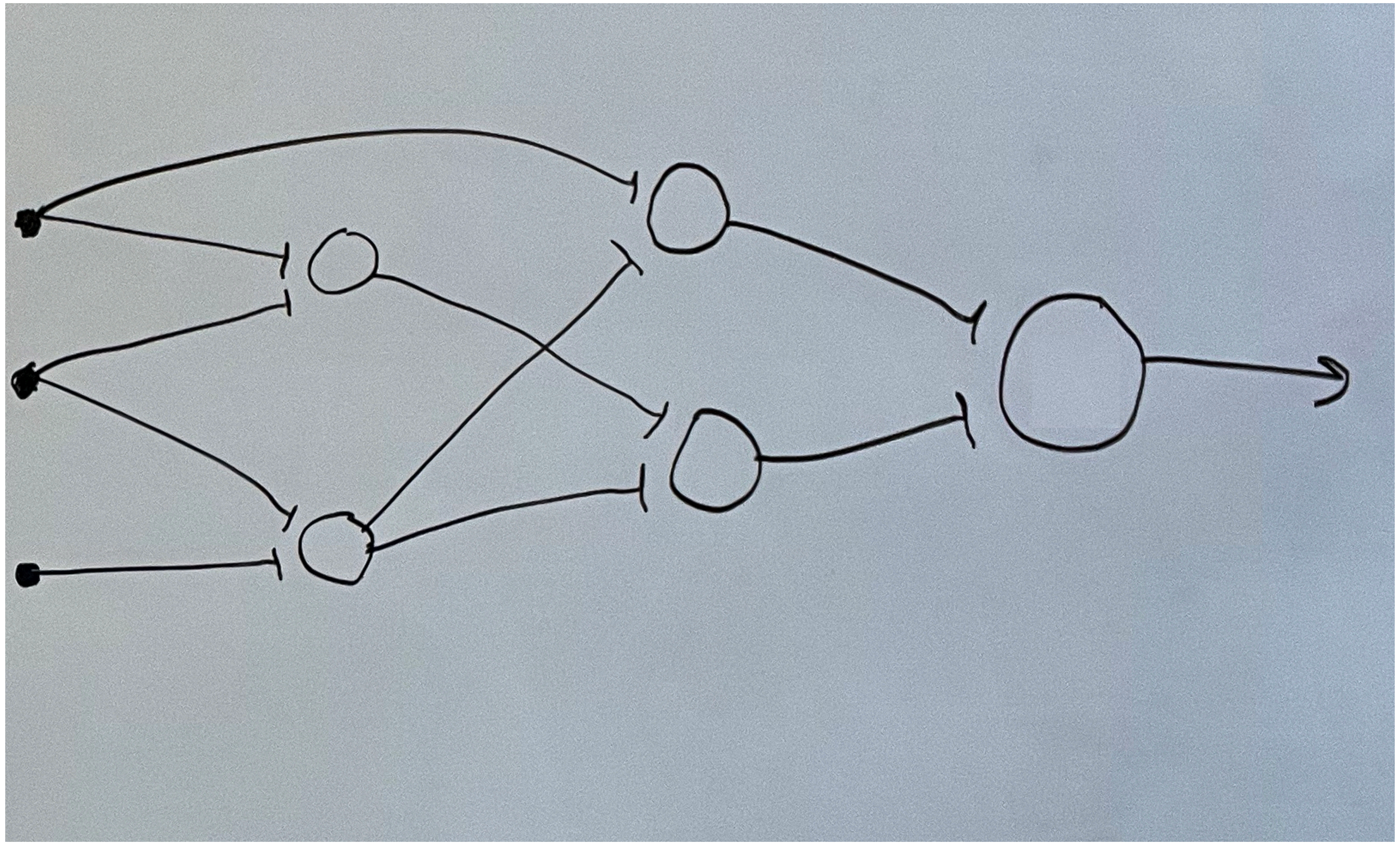
is the part that actually matters for its behavior



A full deep-learned network has many dead units—unused units that can be pruned away without changing the i/o function of the network

The *backbone* of a network

is the part that actually matters for its behavior



A full deep-learned network has many dead units—unused units that can be pruned away without changing the i/o function of the network

Decentralized NNs need multiple new algorithms

- Learning within the backbone
- Learning within the fringe
- Finding the backbone

Decentralized NNs need multiple new algorithms

- Learning within the backbone (backpropagation)
- Learning within the fringe
- Finding the backbone

Decentralized NNs need multiple new algorithms

- Learning within the backbone (backpropagation)
- Learning within the fringe (seek to be listened to)
- Finding the backbone

Learning within the fringe

- By definition, the gradients of the incoming weights $\frac{\partial E_t^2}{\partial w_{ij}}$ of fringe neurons j are always zero; backprop cannot be used on them
- Each fringe neuron has outgoing connections to successor neurons on the backbone that it hopes will listen to it
- Only the successor neurons can change those weights
- The fringe neuron can treat any increase in its outgoing weights as reward

Step-size optimization* is an integral part of learning on the backbone

- Controlling step sizes prevents catastrophic forgetting
    - and protects the backbone from the more-dynamic fringe
- If the fringe creates a useful neuron, the backbone will eventually incorporate it by increasing its step size and then its weight

* Sutton, R.S., "Adapting Bias by Gradient Descent: An Incremental Version of Delta-Bar-Delta," ICML 1992.

Outline

- The idea of neurons that have goals and want to connect and contribute
- New evidence of problems with conventional deep learning (Nature 2024)
  - Deep learning loses plasticity in continual supervised learning
  - Deep learning collapses with prolonged reinforcement learning
- These problems are solved by variation and selective survival (i.e., by decentralized goals)
- Normalization and step-size optimization can also help by enabling online streaming algorithms (Elsayed et. al 2024) and can be seen as decentralization

Early indications of problems with deep continual learning

- Catastrophic Forgetting (French, 1999; McCloskey & Cohen, 1989)
- Loss of Plasticity in early neural networks in the psych literature (Ellis & Ralph, 2000; Zevin & Seidenberg, 2002; Bonin et al., 2004)
- The failure of warm-starting (Ash & Adams, 2020)
- Primacy Bias and resetting in Deep RL (Nikishin et al., 2022)
- Capacity Loss in RL (Lyle et al, 2022)

But no one has previously done a thorough demonstration of Loss of Plasticity using modern deep learning methods

Deep learning loses plasticity in continual supervised learning

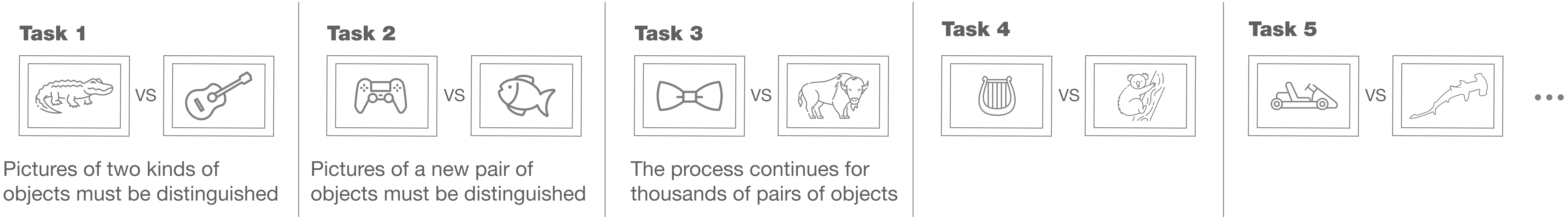
ImageNet — a classic deep-learning problem

- A database of millions of images labelled by nouns (classes)
- 1000 classes with 700 or more images
- Widely used in deep learning to classify images: image⇒class

Dohare, S., Hernandez-Garcia, J.F., Rahman, P., Lan, Q., Sutton, R.S., Mahmood, A.R. "Loss of plasticity in deep continual learning." *Nature 632*, pp. 768-774, August 22, 2024.

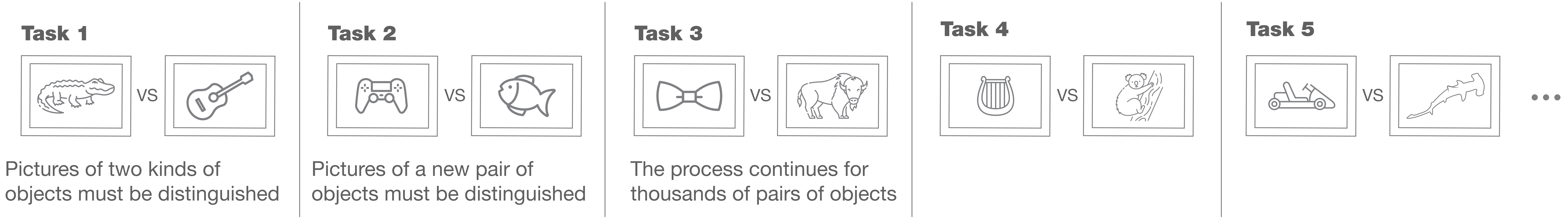
The *Continual* ImageNet Problem

- The classical ImageNet problem was minimally changed to make it continual
- Classes were taken in pairs to produce a sequence of binary classification tasks



The *Continual* ImageNet Problem

- The classical ImageNet problem was minimally changed to make it continual
- Classes were taken in pairs to produce a sequence of binary classification tasks



- Performance measure: %correct on test set (by argmax) at end of each task
- Averaged over 30 independent runs, varying class pairings, test sets

Network and Training Procedure (for ImageNet)

- All binary classification tasks shared the same network; both heads reset at task switch
- Standard neural network, though slightly narrow for ImageNet (bc. only 2 classes at a time)
  (3 convolution layers of 32/64/128 filters + 3 fully-interconnected layers of 128/128/2 artificial neurons)
- For each task, 12 batches of 100 examples, 250 epochs (passes through the data)
- Weights initialized by the standard Kaiming distribution, only once, before the first task
- Backpropagation with momentum on the cross-entropy loss, ReLU activations
- Many variations on the network and hyper-parameters were tested to obtain good and representative performance on the *first task*

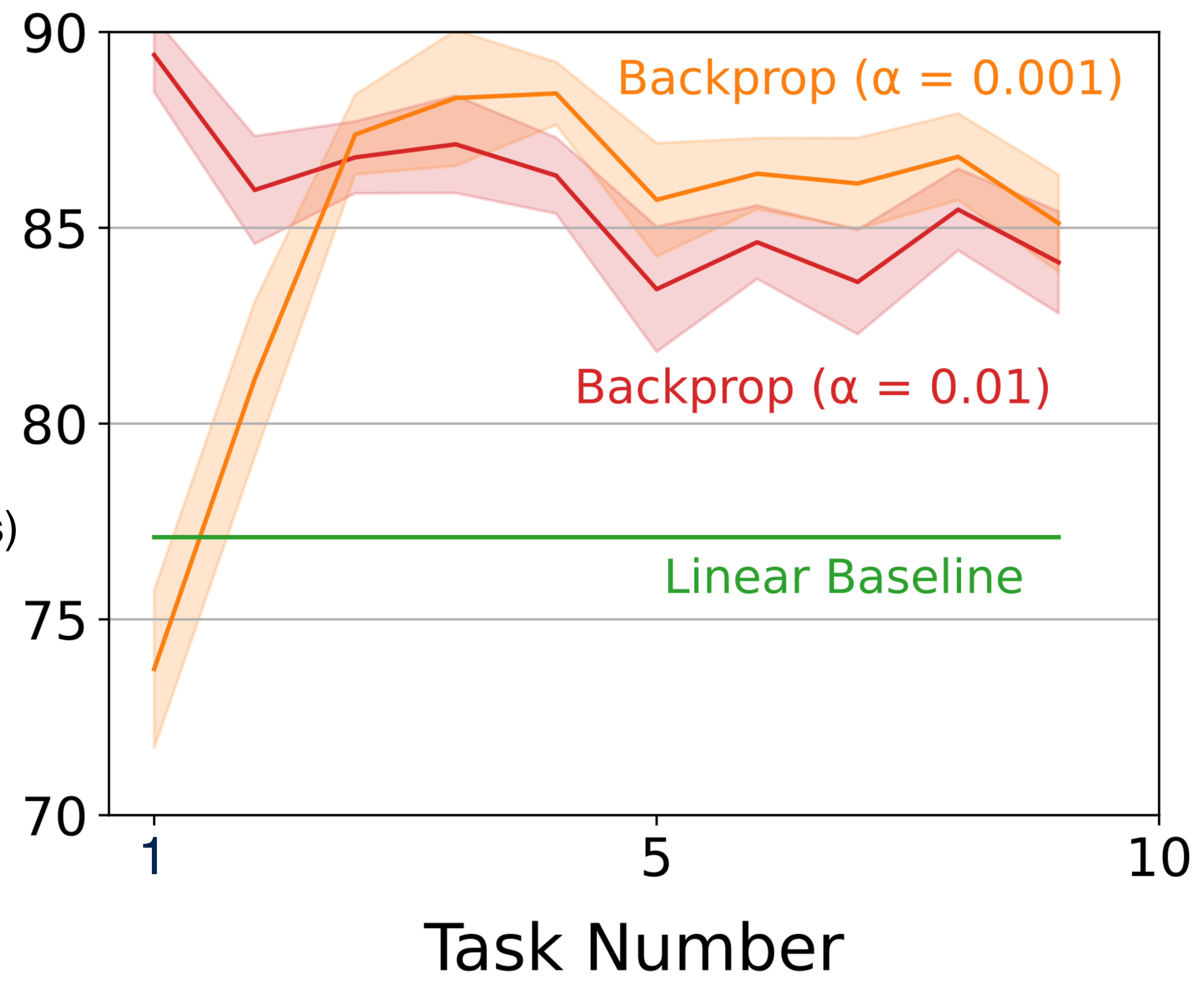
How will performance evolve over the sequence of tasks?

Will performance be better on the 1st task or the 2nd task? the 500th?

BackProp on Continual ImageNet (first 10 tasks)

%Correct
on Test Set

(at end of each task,
averaged over 30 runs)

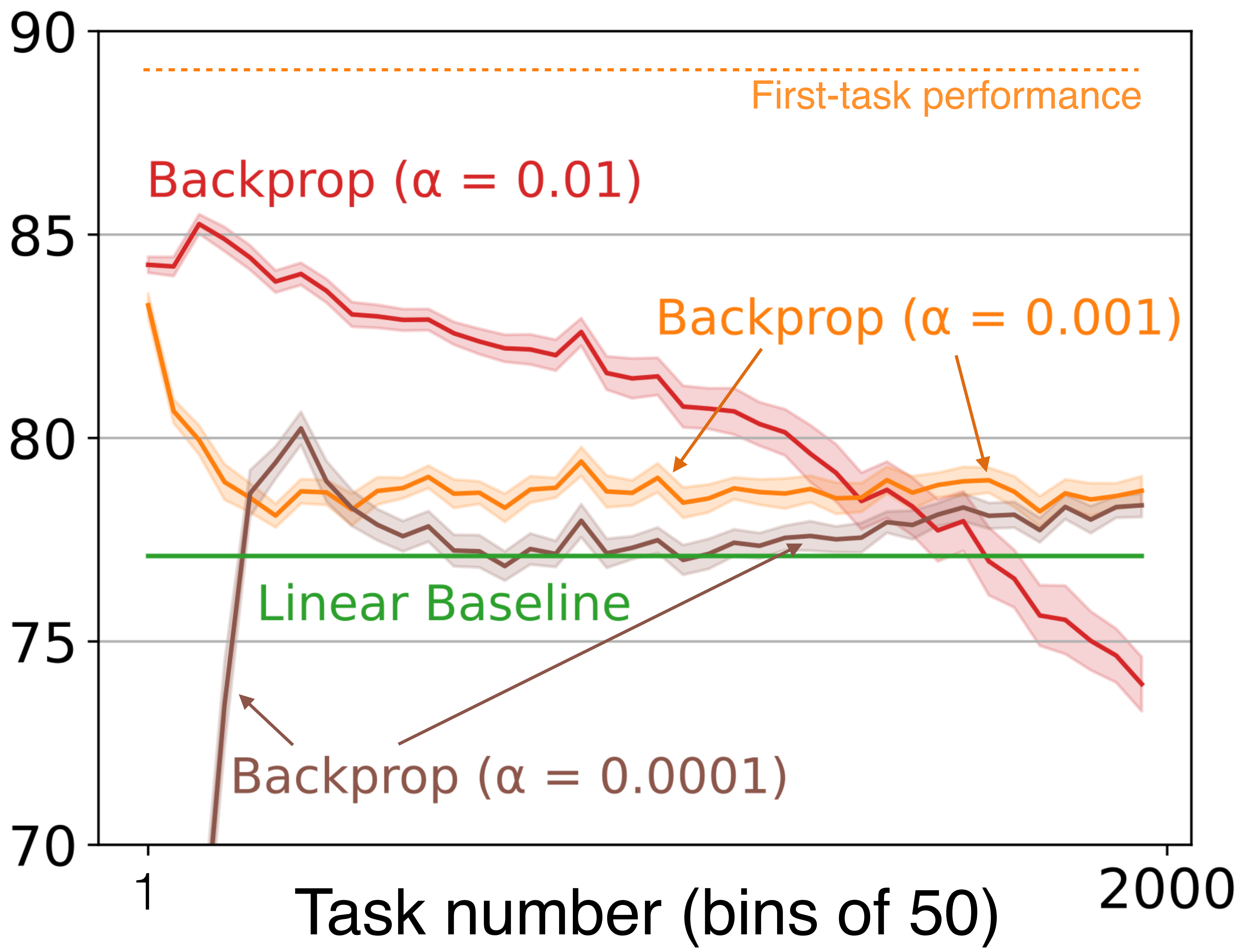


- Chance performance is 50%
- Best performance on first task is ≈89%
- Shaded region is one standard error
- Linear baseline is the performance of linear heads direct from pixels

Learning rate (plasticity) sometimes improves over early tasks, then…?

BackProp on Continual ImageNet (2000 tasks)

%Correct
on Test Set
(at end of each task, averaged over 30 runs)

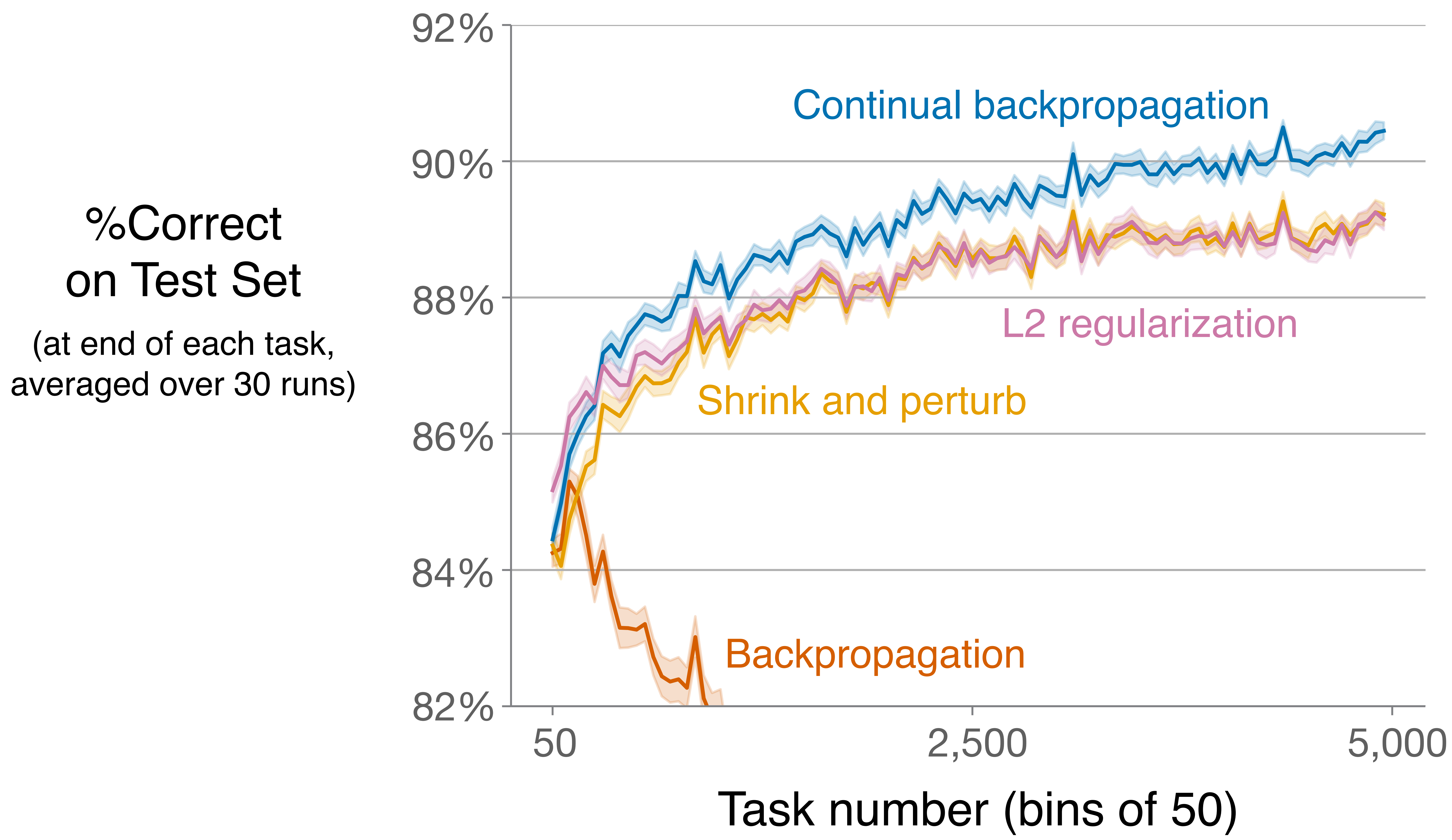


- This data is representative, the details depend on the details:
  - #epochs
  - step-sizes
  - network sizes
- Each line takes ≈24 hours to compute
- Most other variations of BackProp (Adam, Dropout, Batch norm) are worse

For good hyper-parameters, plasticity decreases across tasks, nearing the poor performance level of a one-layer (linear) network, or worse

BackProp shows "Catastrophic" Loss of Plasticity

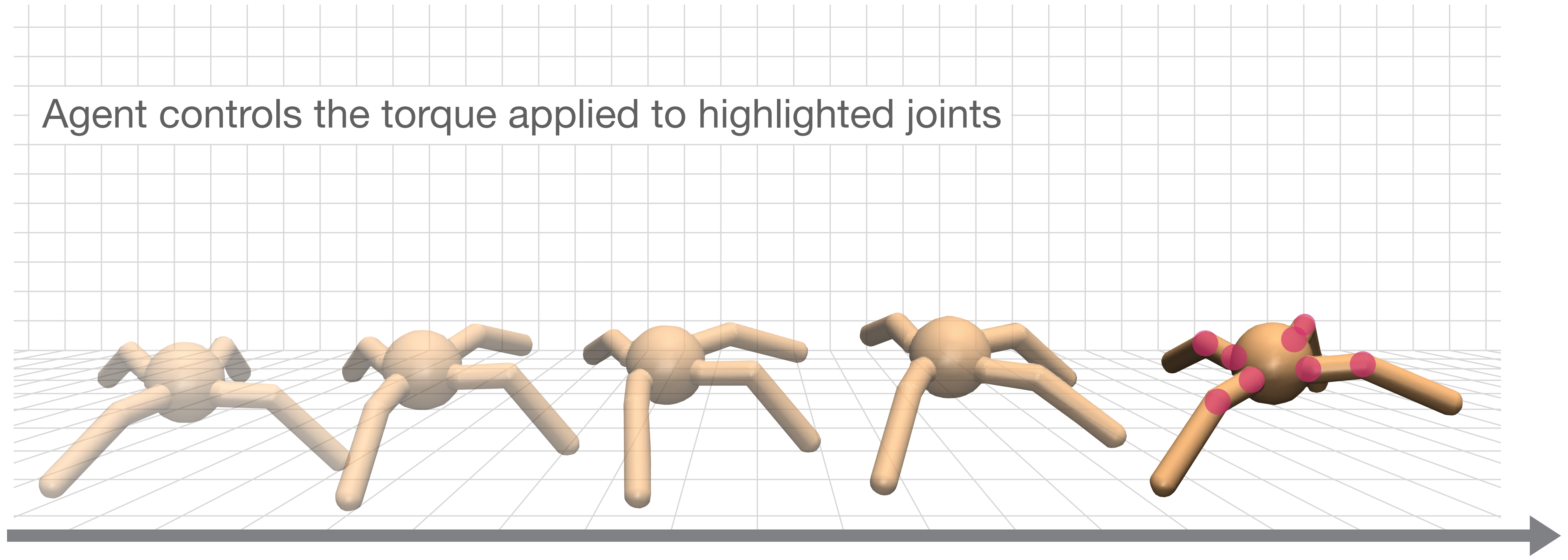
There are better algorithms on Continual ImageNet



- L2 regularization adds a penalty for large weights
- Shrink and Perturb is L2 reg. plus random variation of all weights
- Continual Backpropagation continually re-initializes a small fraction of units
  - otherwise its just like BackProp

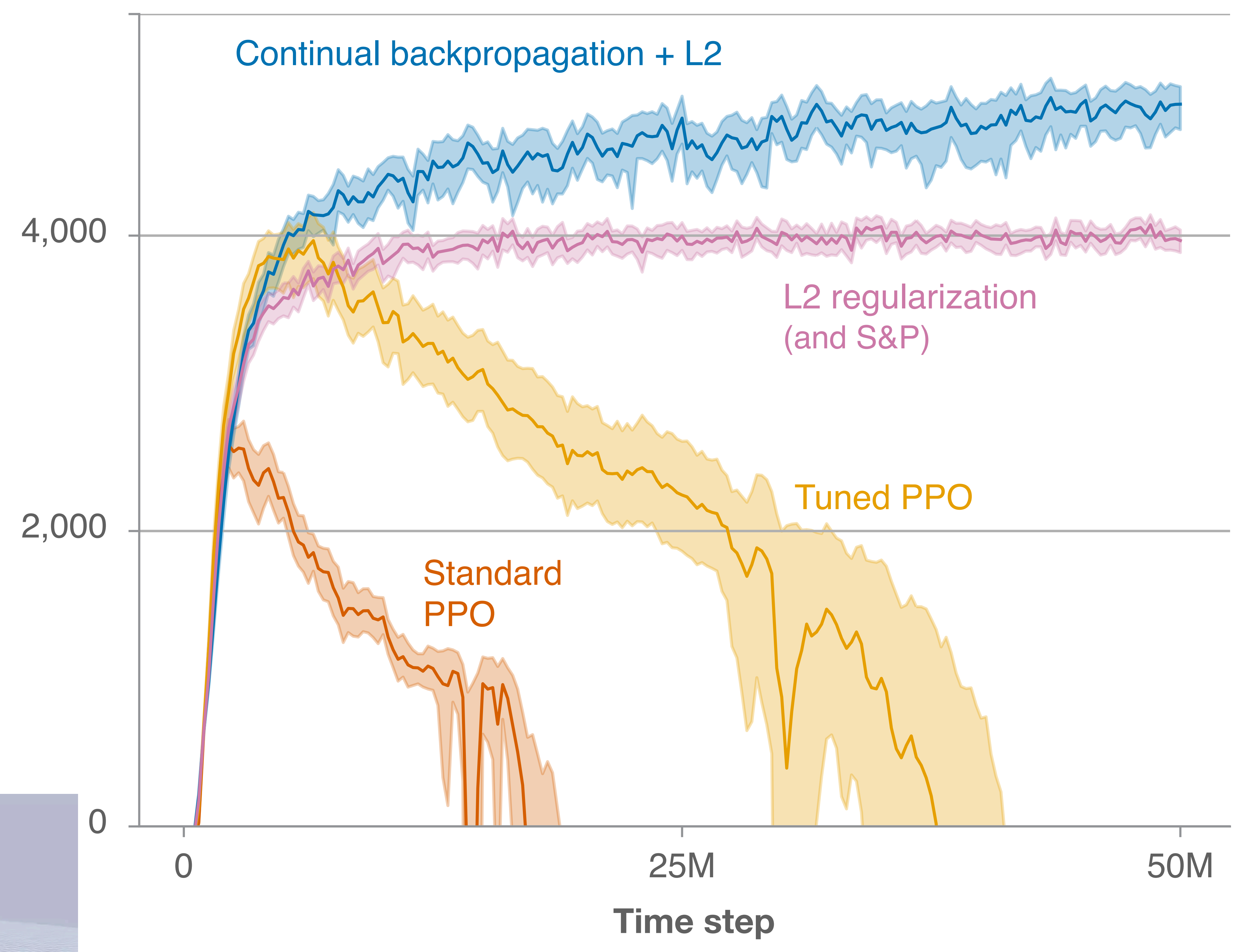
Deep learning collapses with prolonged reinforcement learning

a **Ant locomotion**

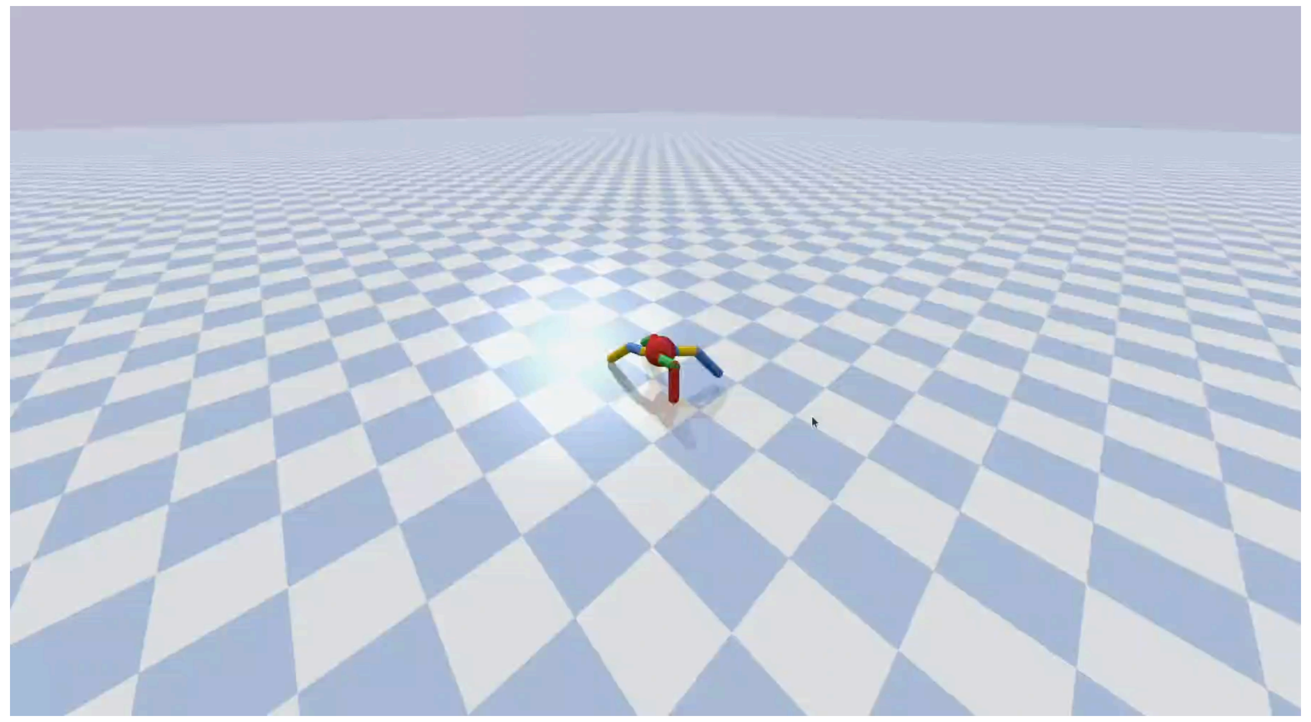


Agent is rewarded for foward motion and penalized if applied torque or contact forces are too large

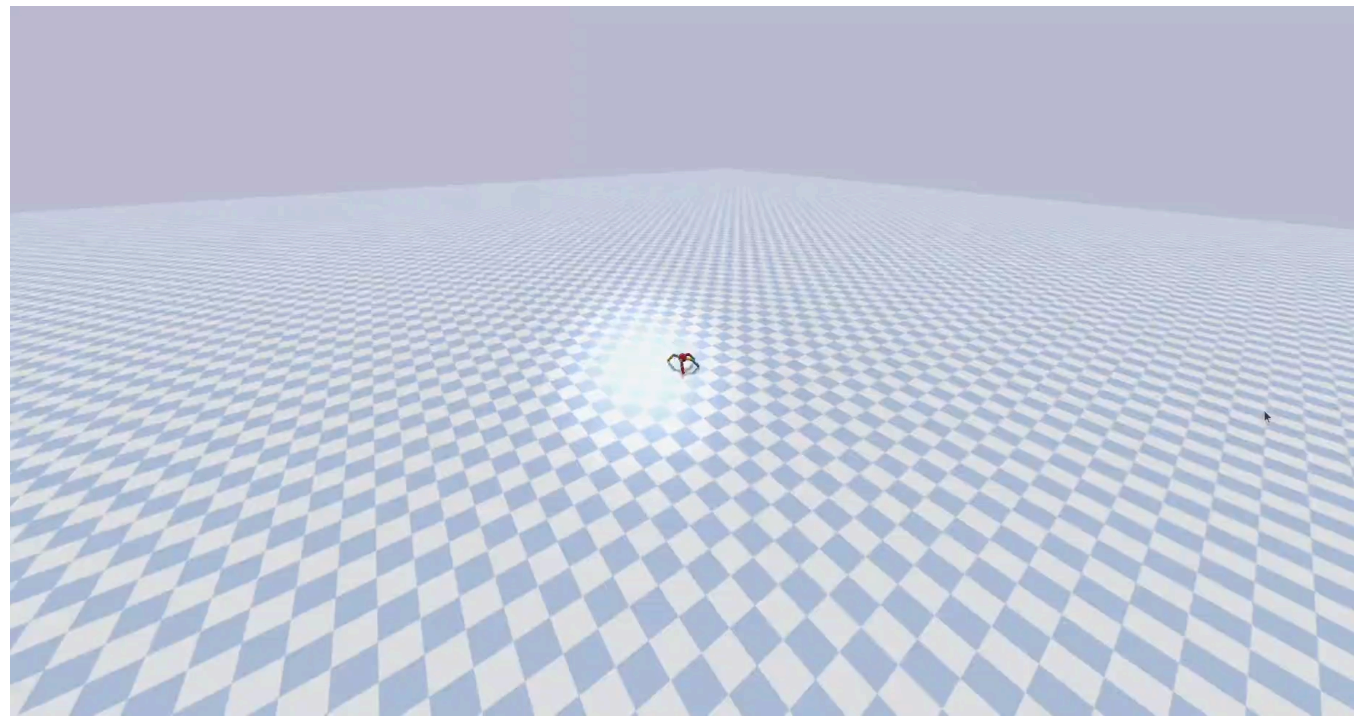
Reward per episode



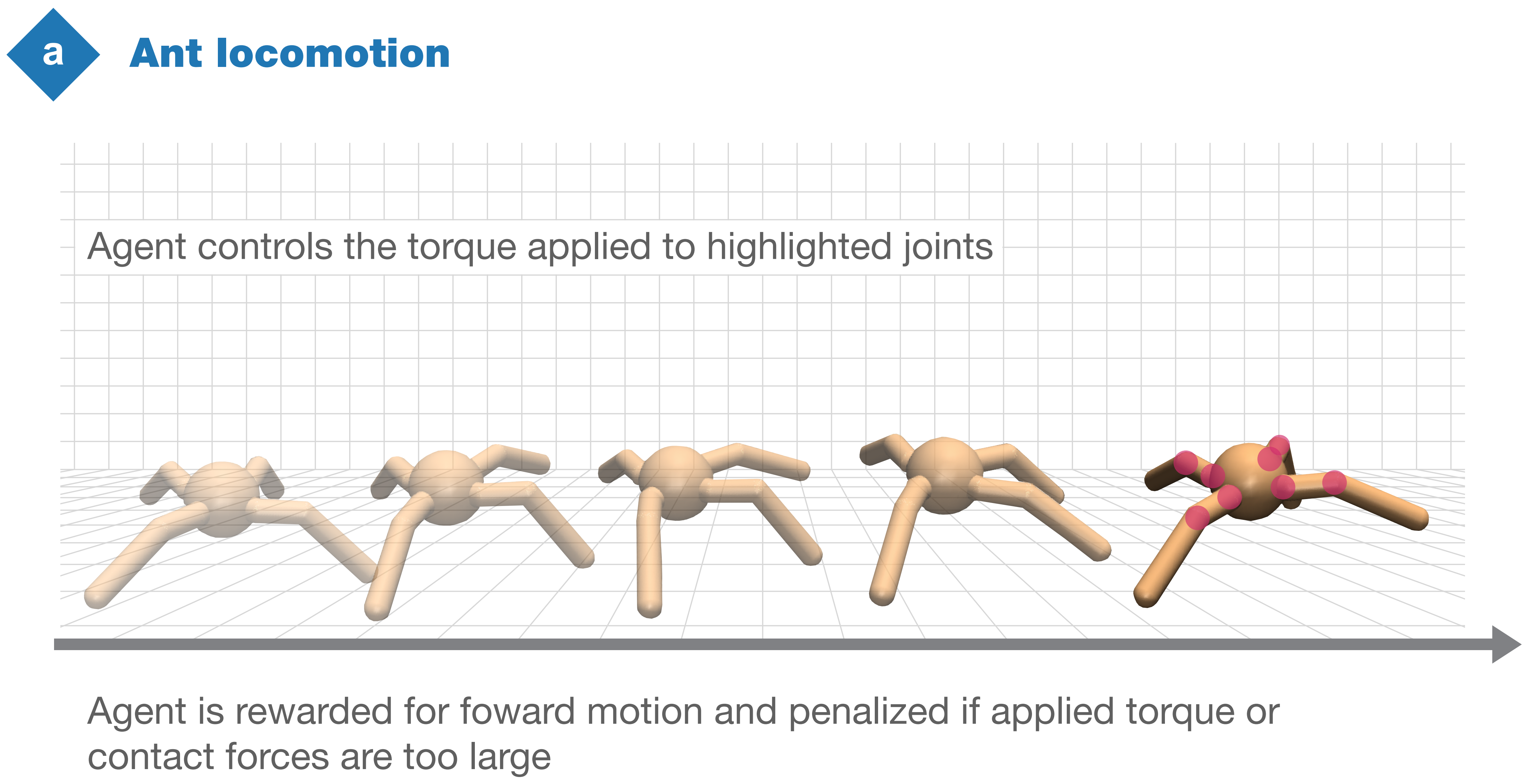
PPO

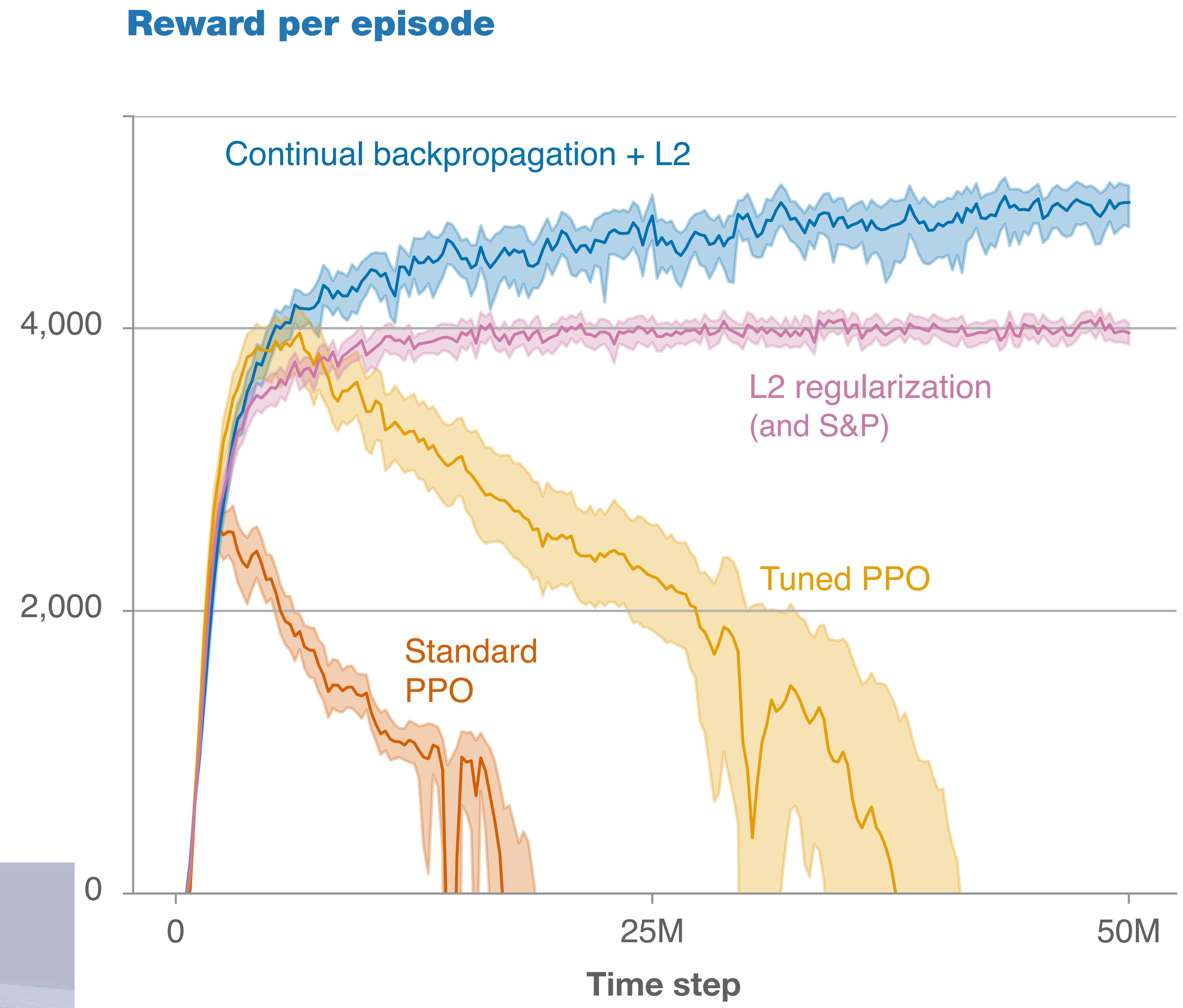


Continual PPO

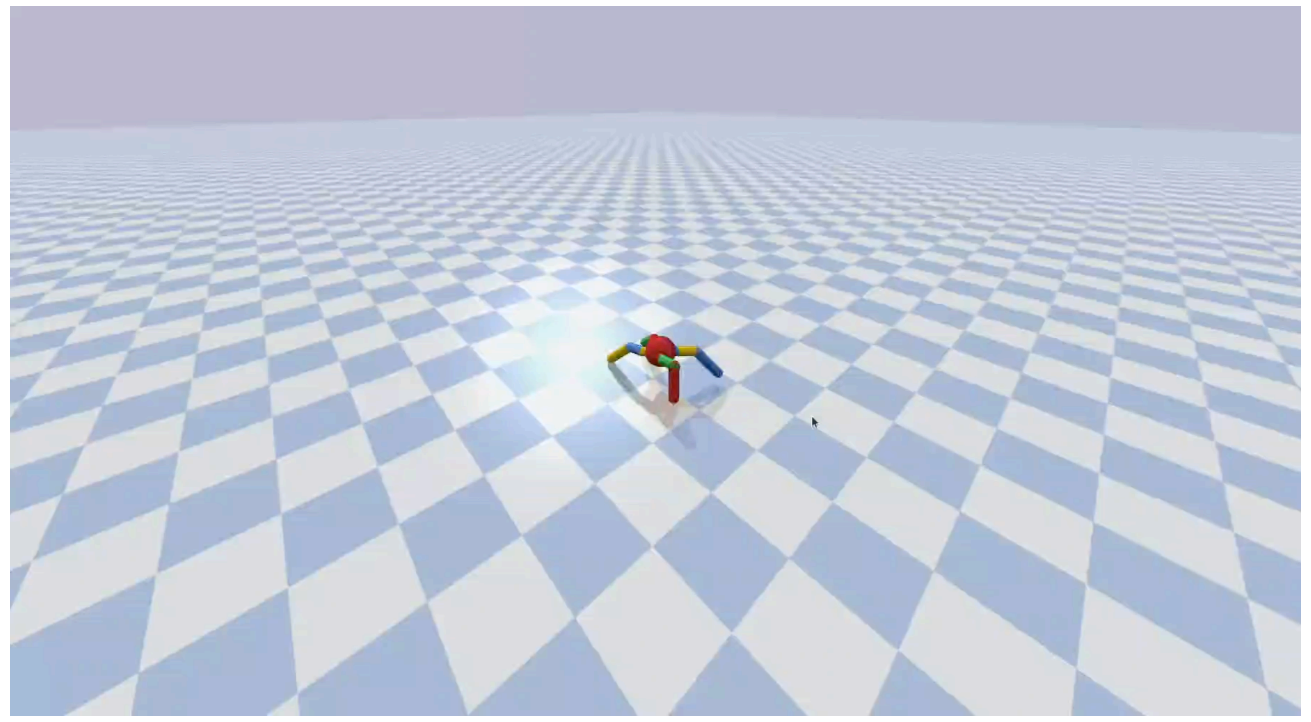


Deep learning collapses with prolonged reinforcement learning

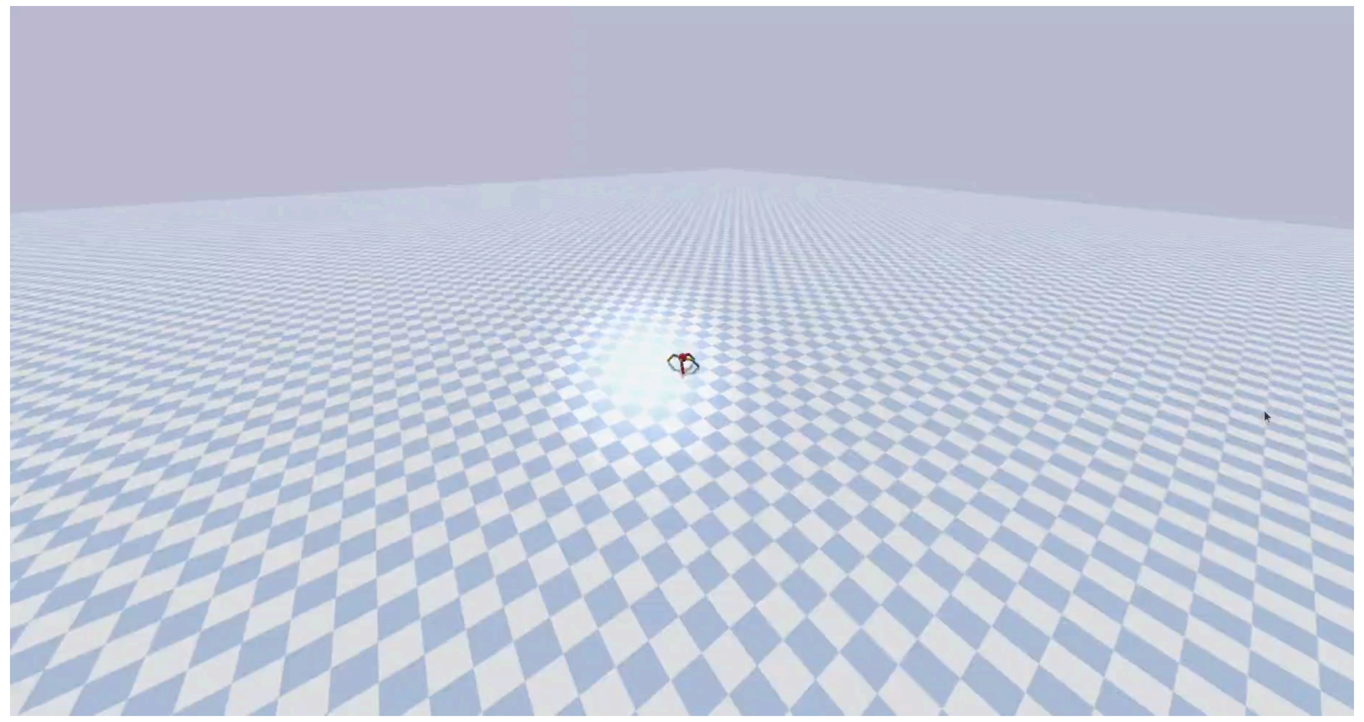




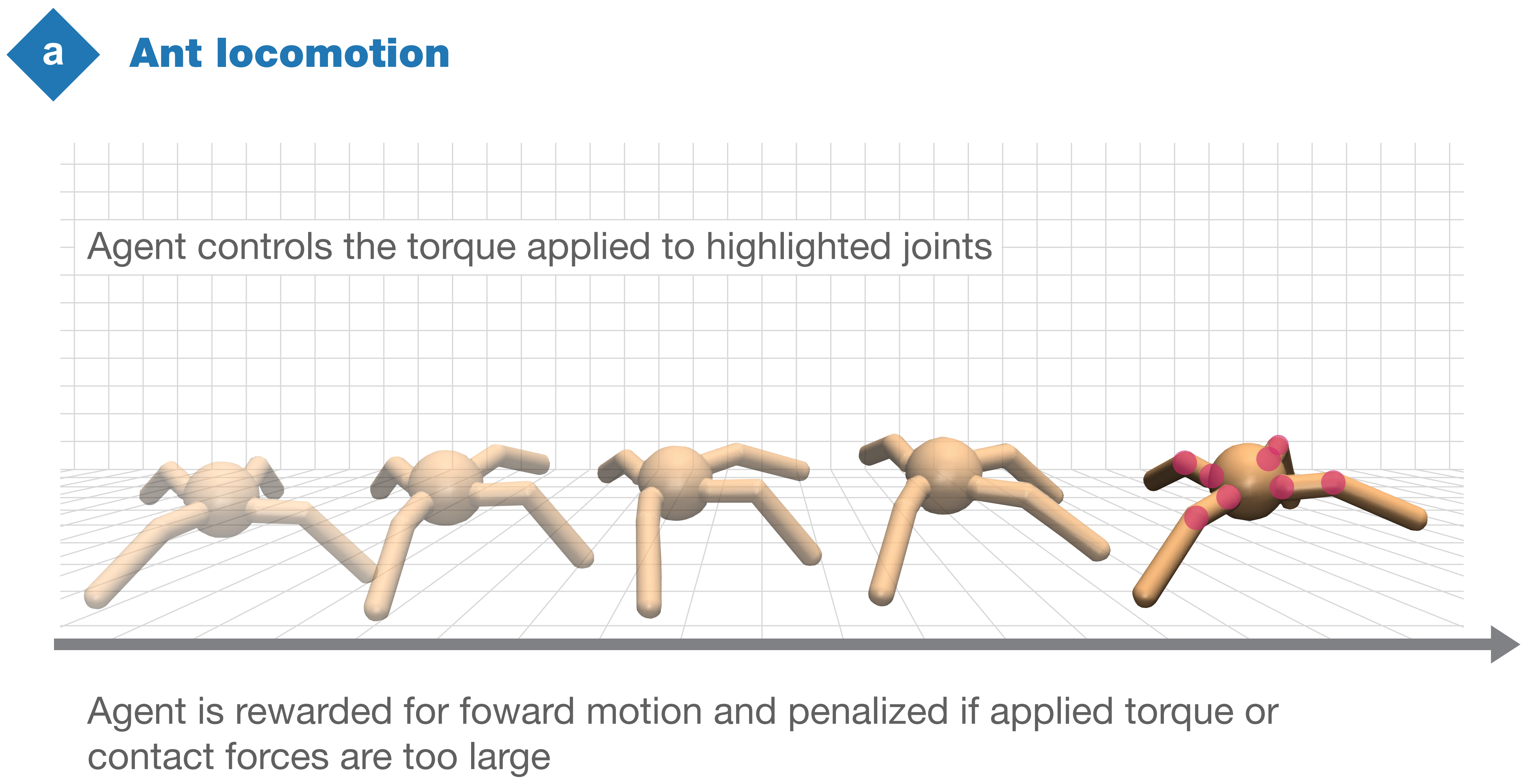
PPO

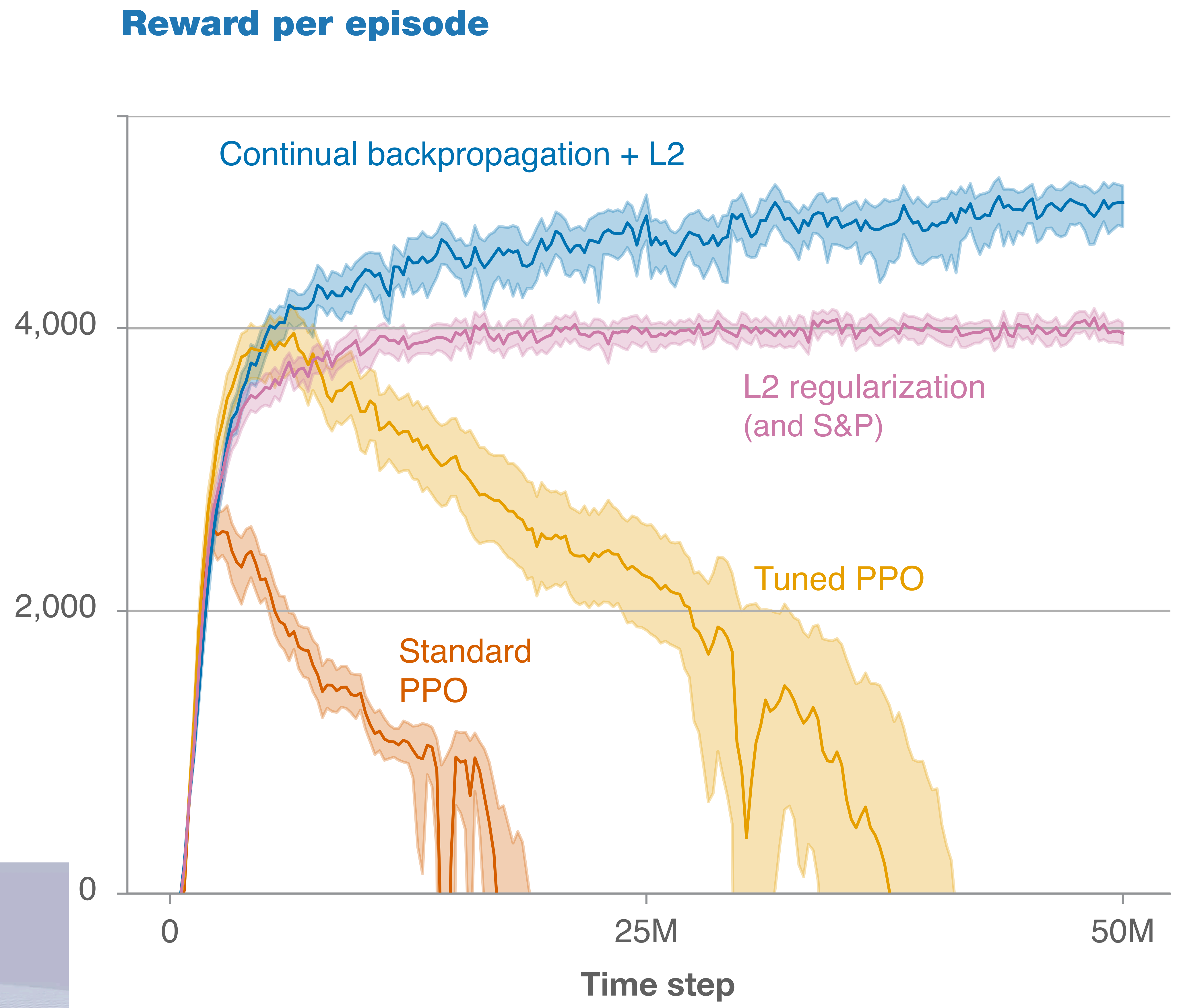


Continual PPO

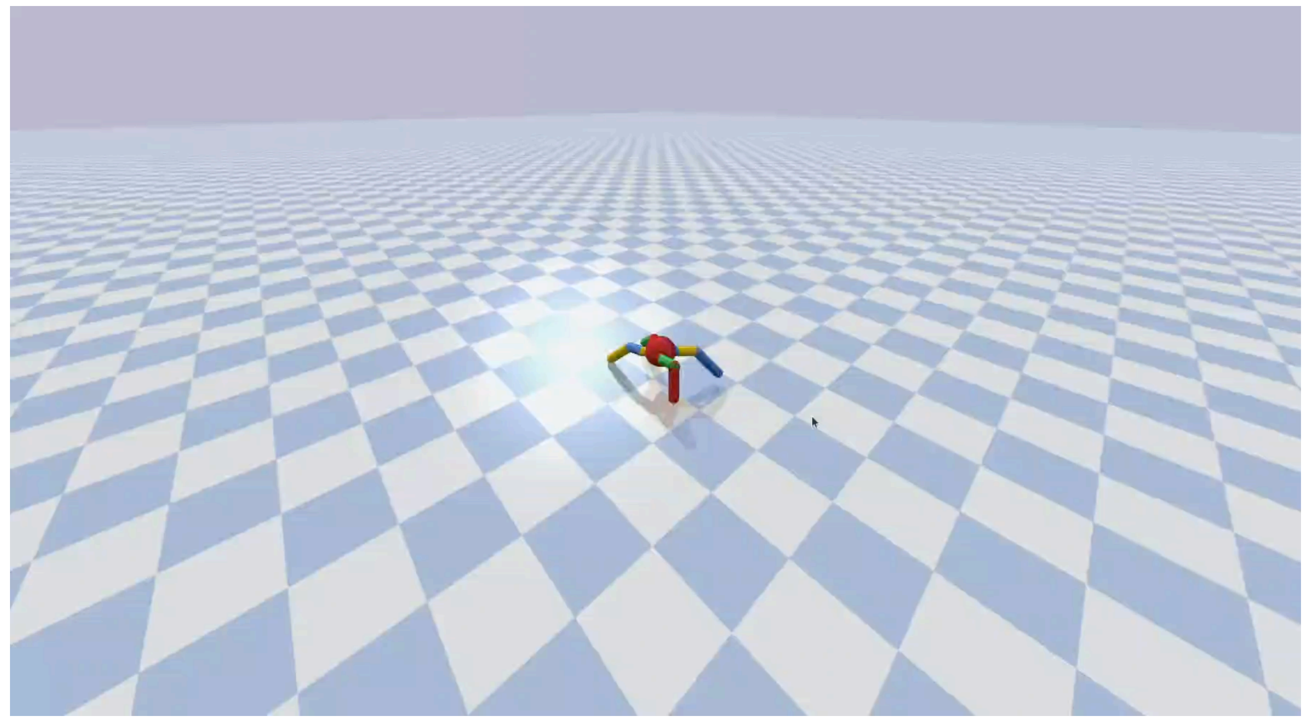


Deep learning collapses with prolonged reinforcement learning

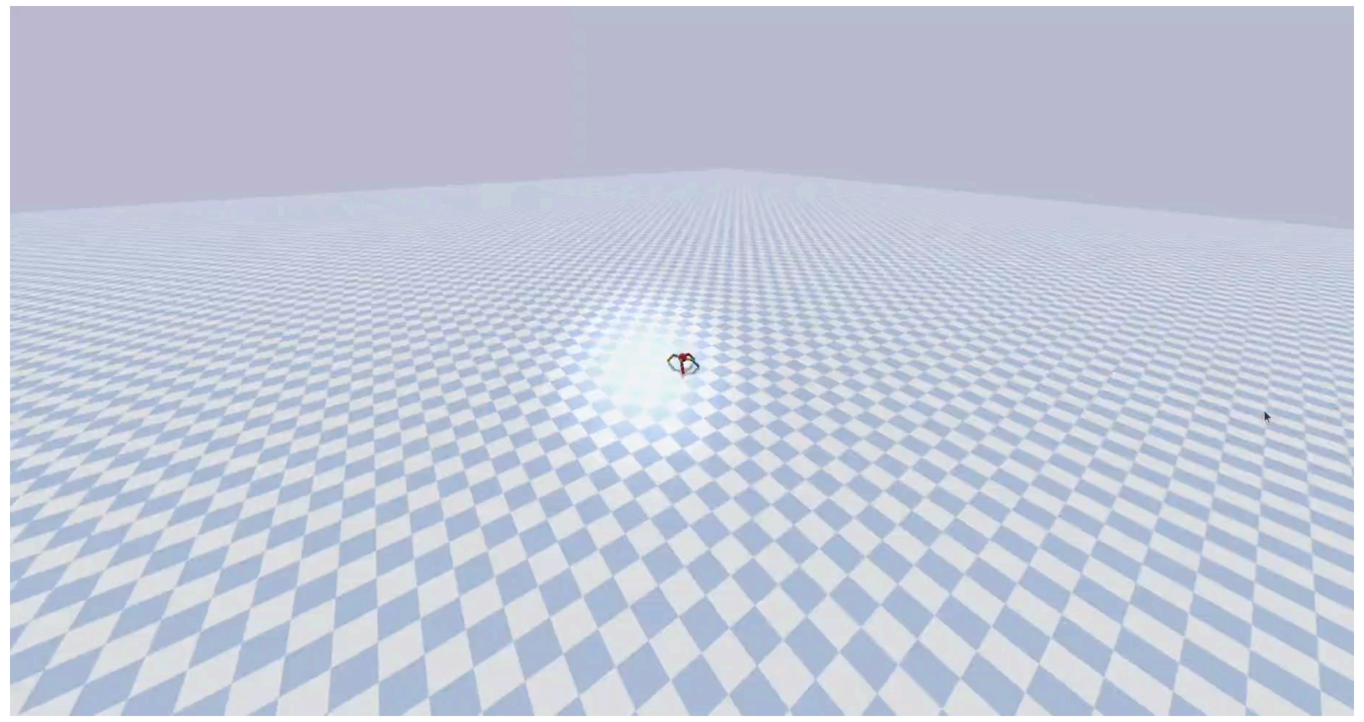




PPO

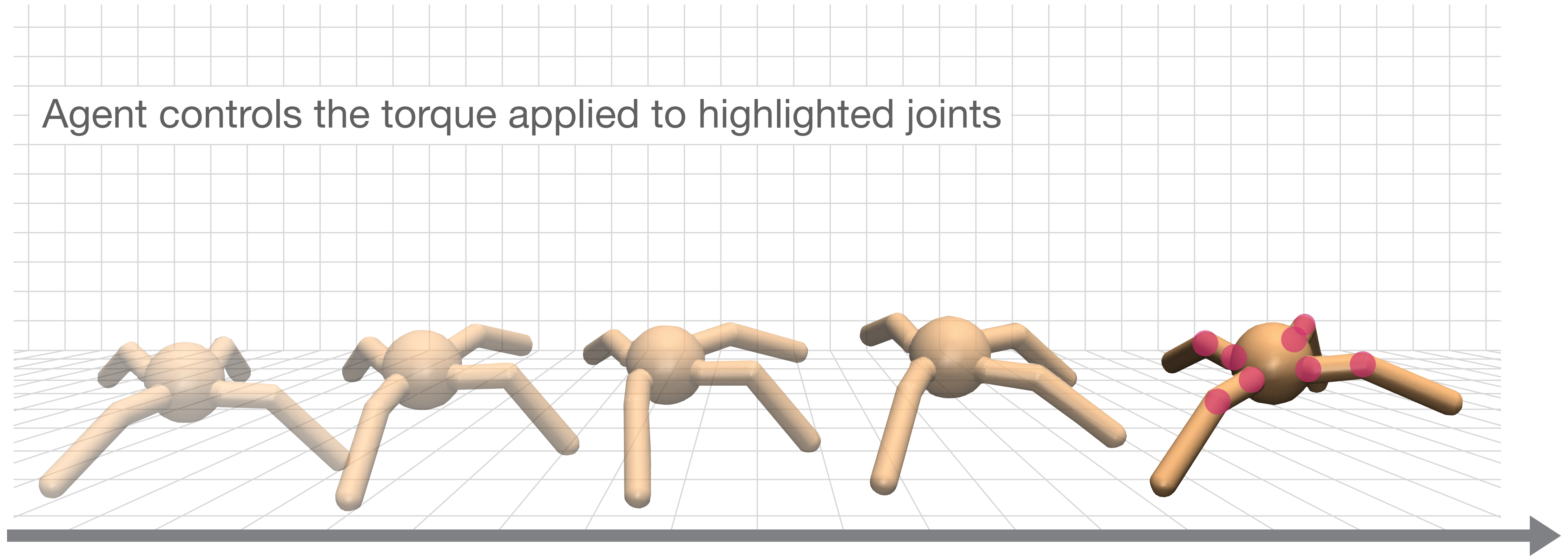


Continual PPO



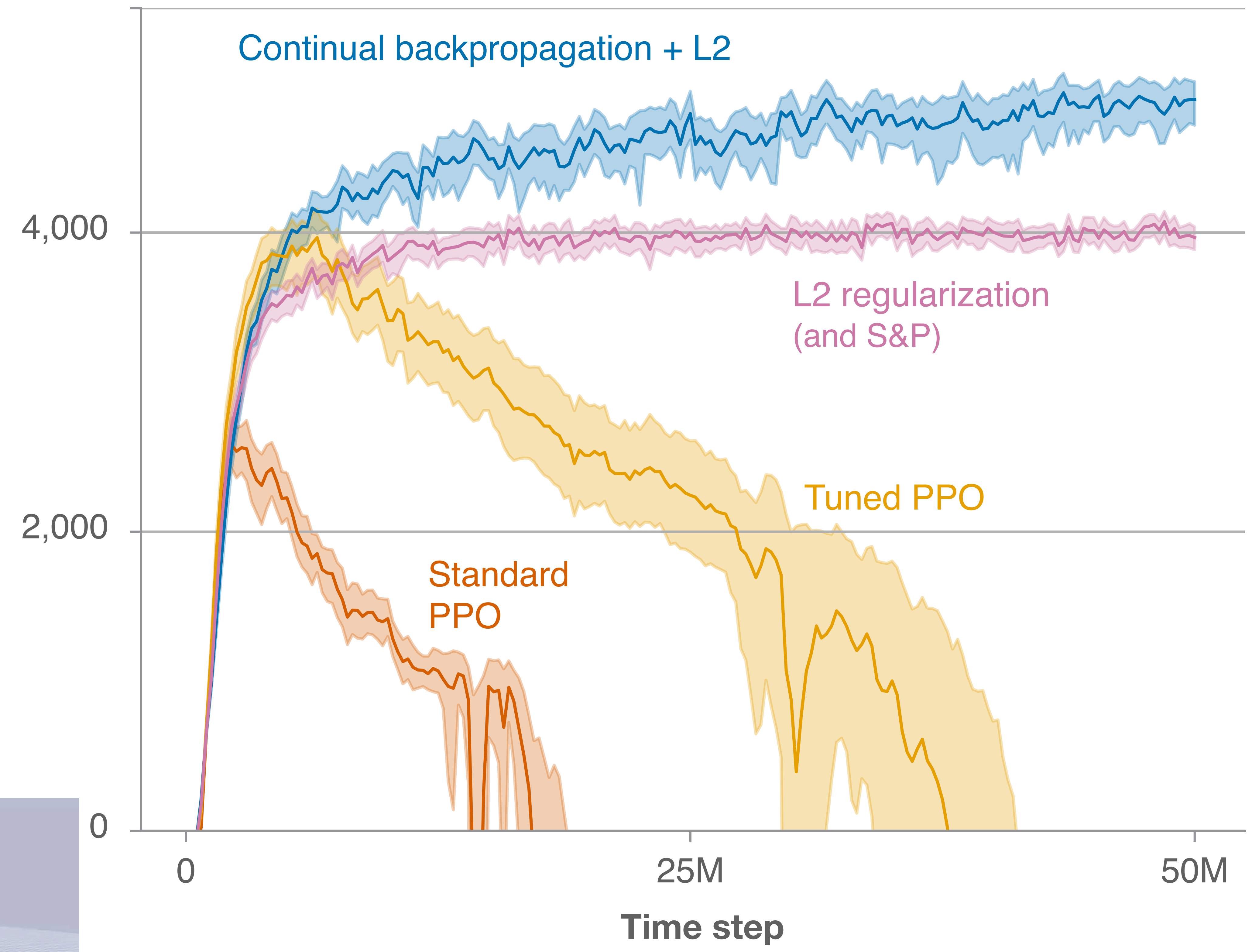
Deep learning collapses with prolonged reinforcement learning

a **Ant locomotion**

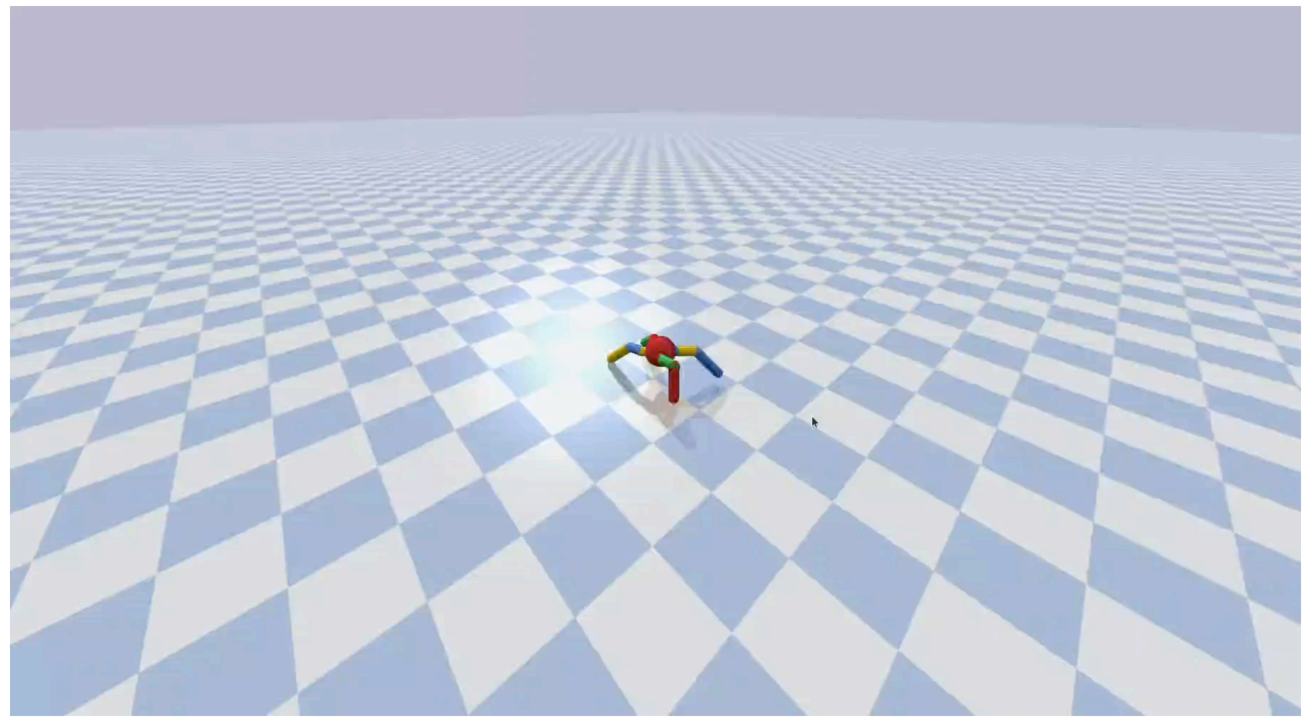


Agent is rewarded for foward motion and penalized if applied torque or contact forces are too large

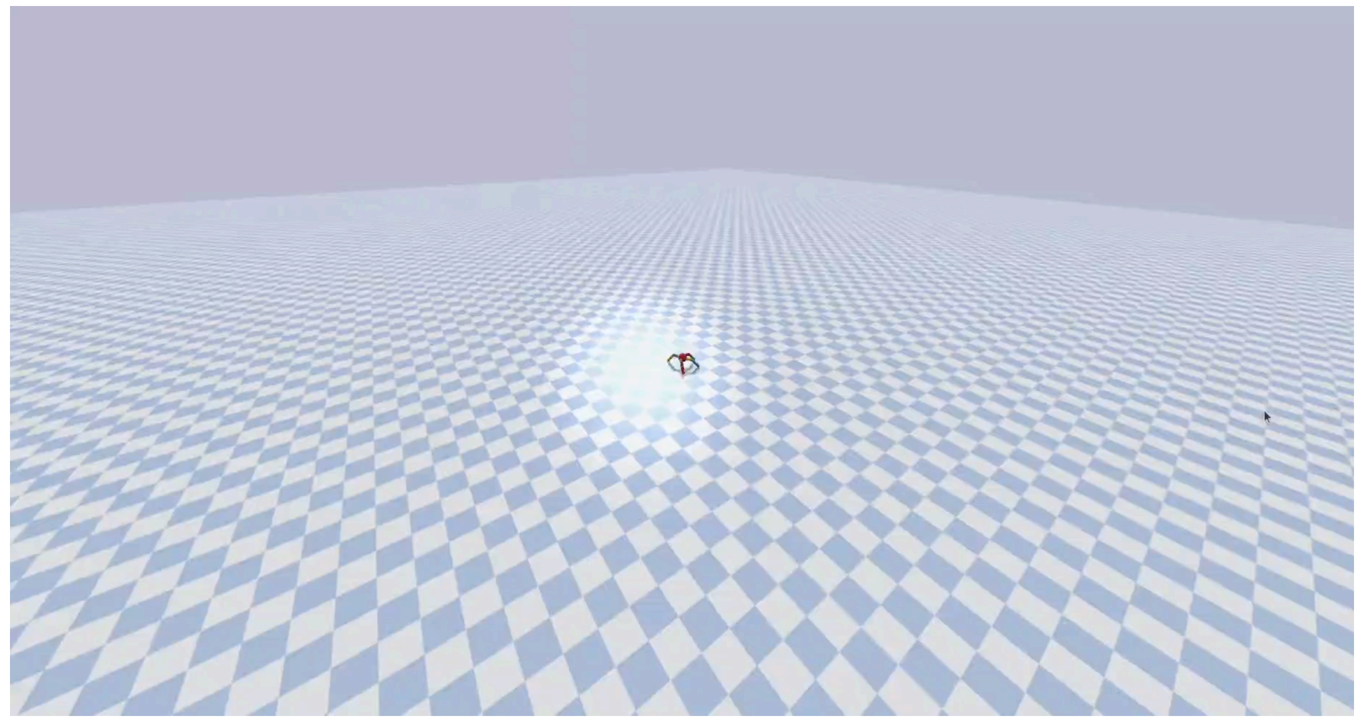
Reward per episode



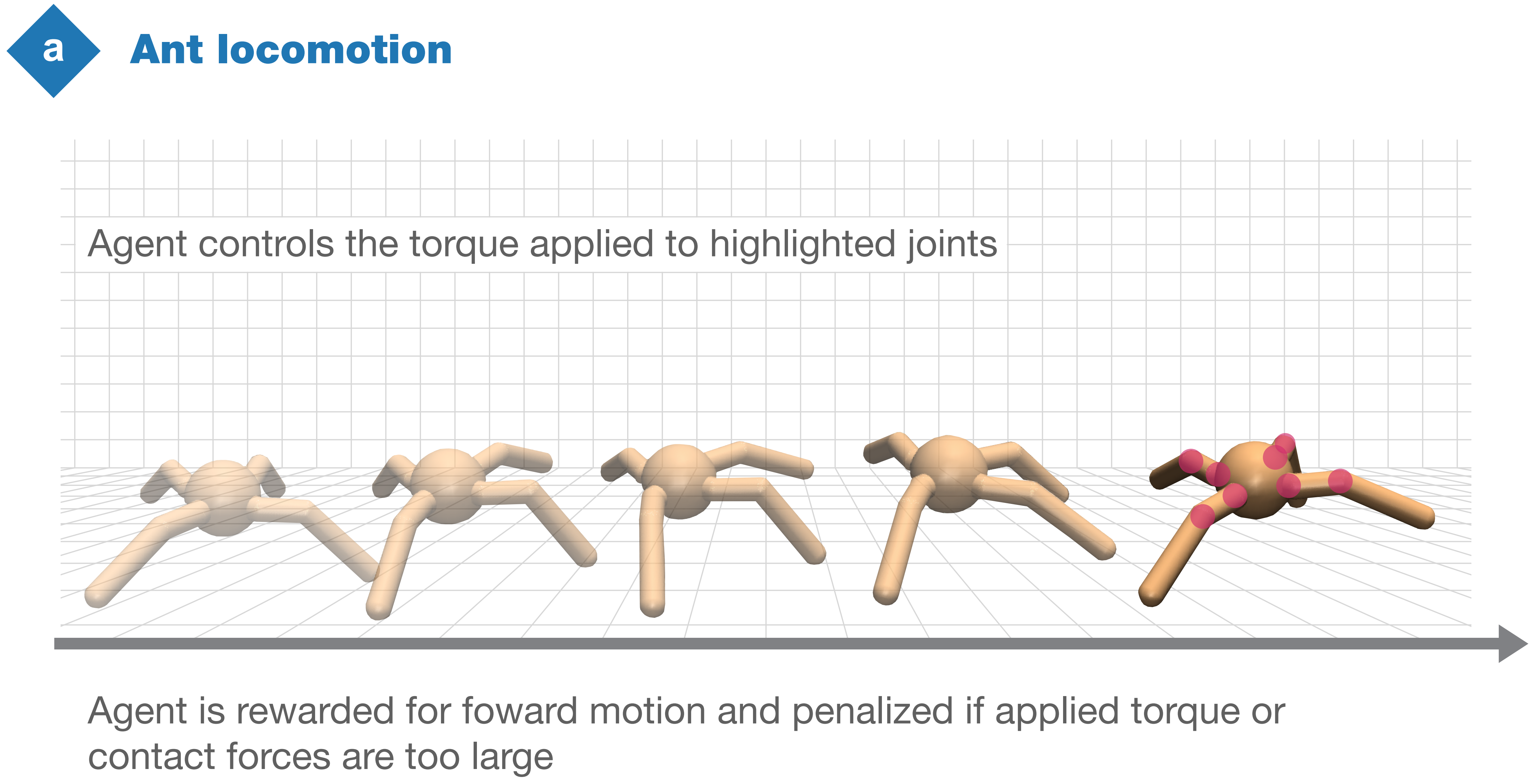
PPO

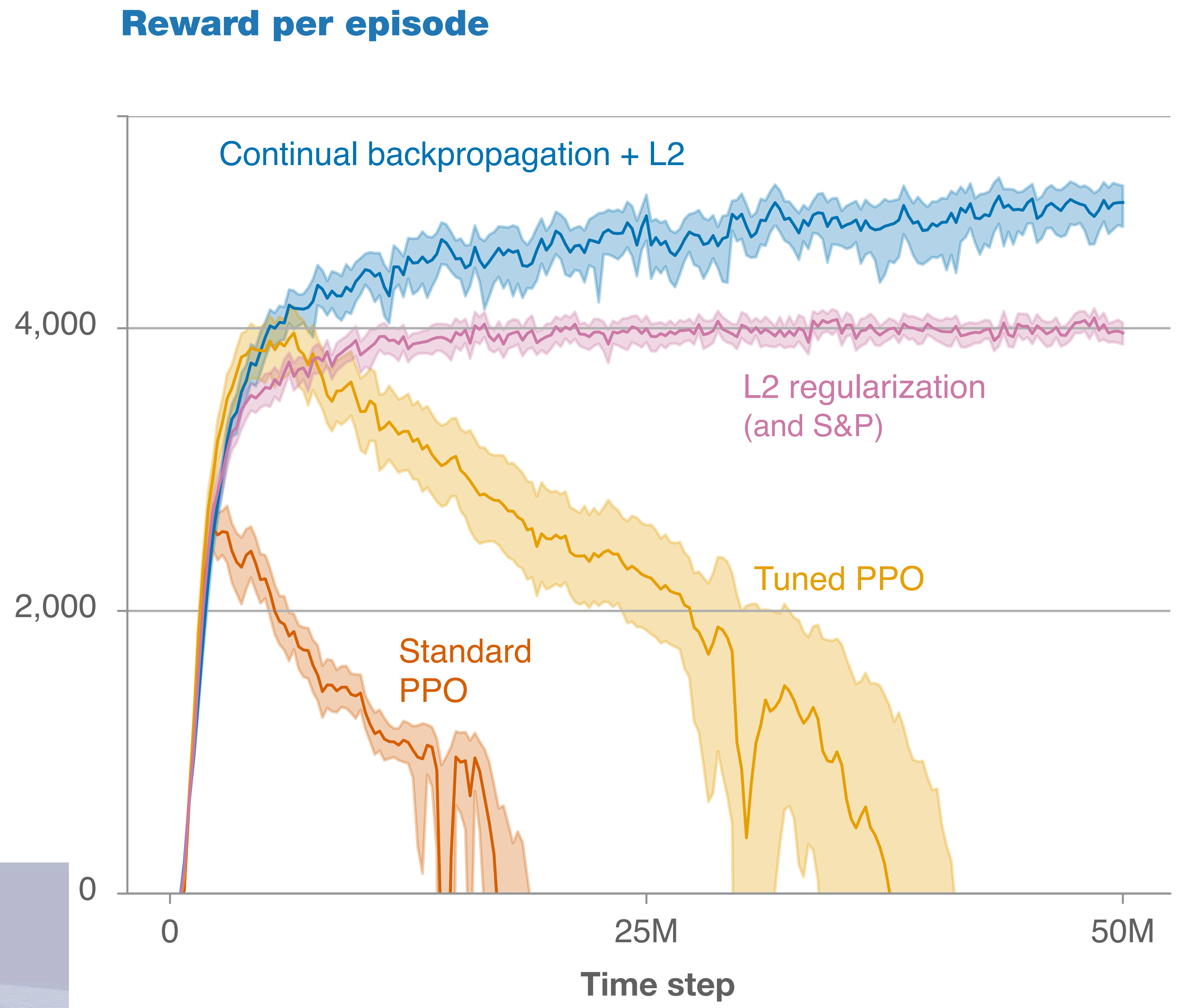


Continual PPO

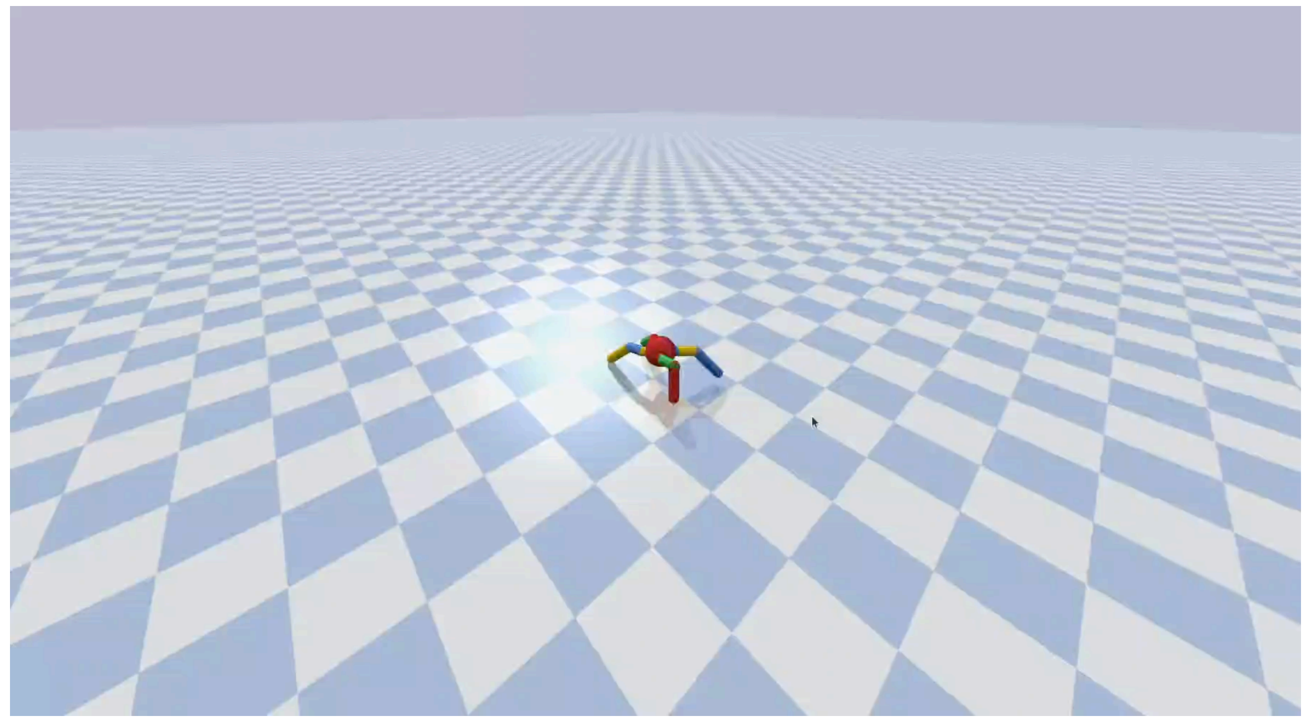


Deep learning collapses with prolonged reinforcement learning

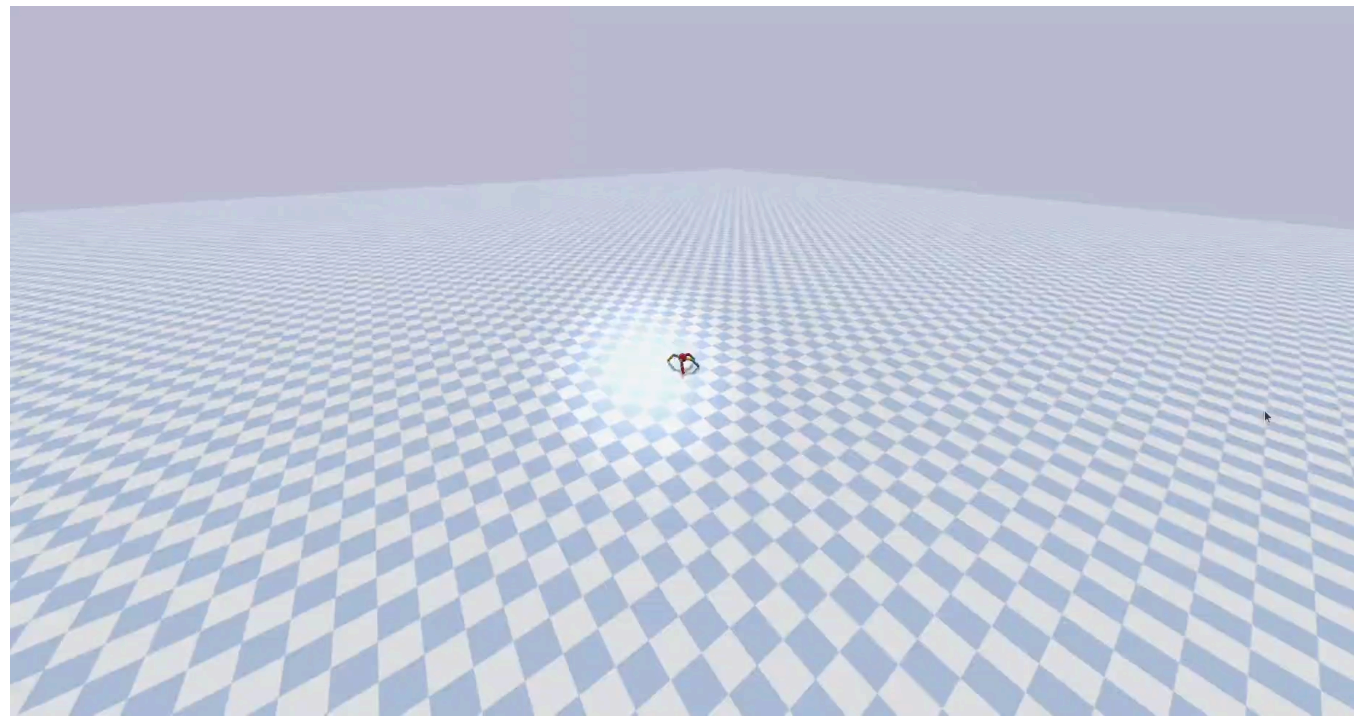




PPO



Continual PPO



Continual Backpropagation:

Stochastic Gradient Descent with Selective Reinitialization

- Just like backprop, except re-initializes a small fraction of the artificial neurons on every step

- Re-initialization is selective; the neurons are ranked by a notion of utility, and only the least useful are re-initialized

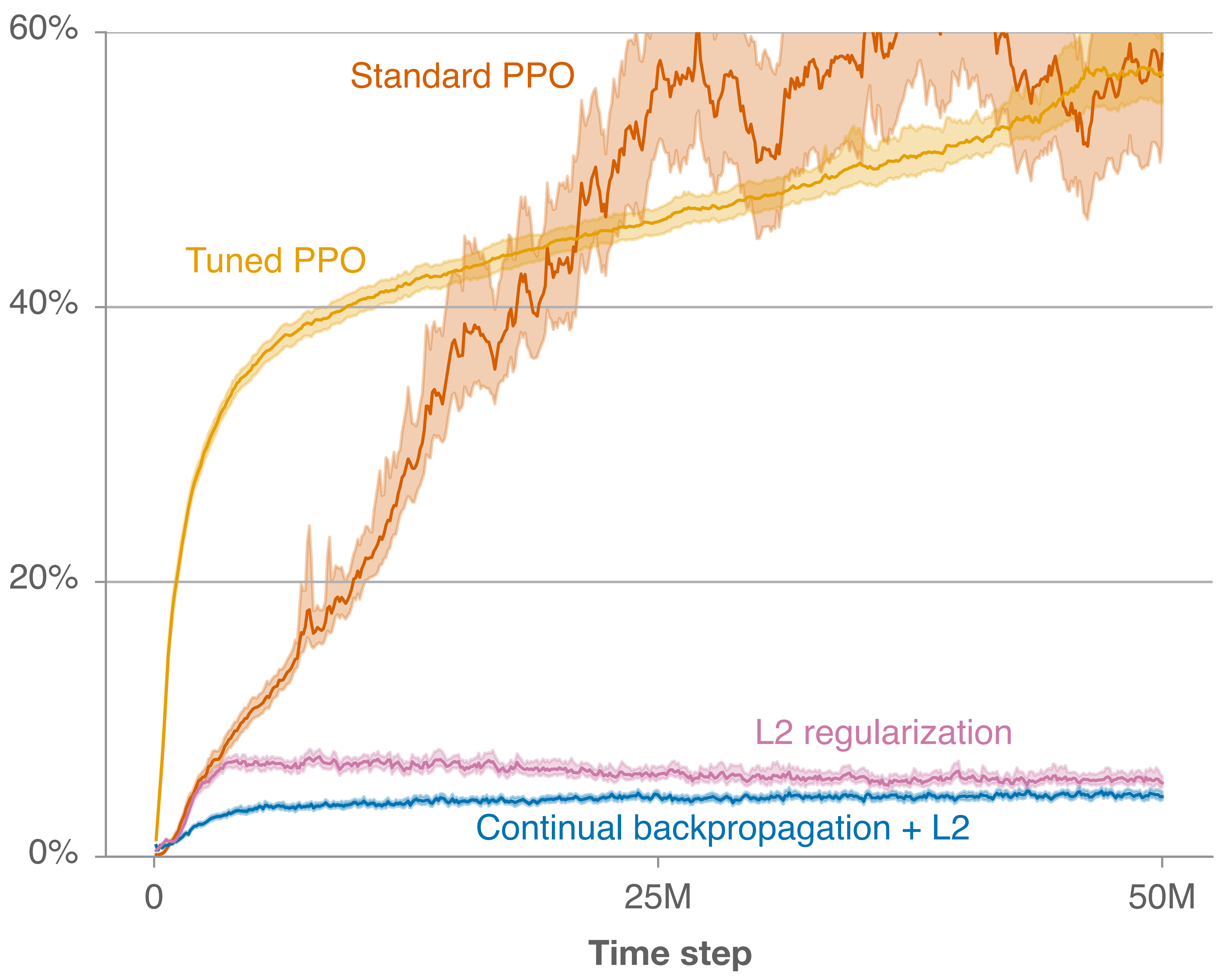
utility update for neuron i: $u_i^{t+1} \leftarrow \eta u_i^t + (1-\eta)|y_i^t| \sum_k |w_{ik}^t|$

activation of neuron i weight from i to k

- Neurons re-initialize until other neurons grow a weight from them; they "seek attention"; they have their own goal different from the network's

Why is deep learning failing?
Many of the artificial neurons become forever inactive

Percentage of dormant units (active <1% of the time) in ant locomotion



Most neurons go dormant with PPO

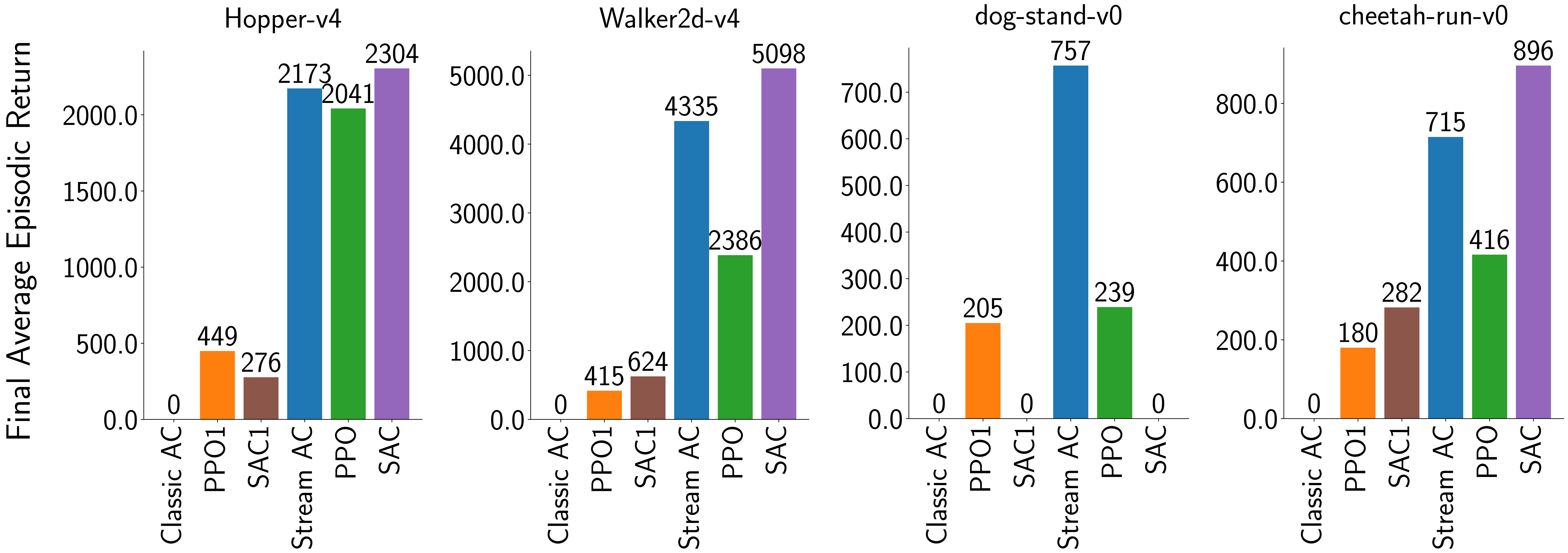
If neurons were taking responsibility for their own operation, then it would be easy for them to notice that they had gone dormant

i.e., decentralization would solve this problem

Outline

- The idea of neurons that have goals and want to connect and contribute
- New evidence of problems with conventional deep learning (Nature 2024)
  - Deep learning loses plasticity in continual supervised learning
  - Deep learning collapses with prolonged reinforcement learning
- These problems are solved by variation and selective survival (i.e., by decentralized goals)
- Normalization and step-size optimization can also help by enabling online streaming algorithms (Elsayed et. al 2024) and can be seen as decentralization

Streaming reinforcement learning algorithms* are competitive with batch algorithms for the first time



*Elsayed, M., Vasan, G., Mahmood, A. R. (2024) “Streaming deep reinforcement learning finally works,” arXiv:2410.14606

The innovations of the new streaming algorithms can be seen as following from the decentralized perspective

- The new streaming algorithms differ from prior attempts mainly in
  - signal normalization
  - constraining the step-size parameters to reasonable bounds
- Both are natural for an artificial neuron that takes responsibility for the conditioning of its local signals and learning processes

Conclusions

- All is not well with today's deep learning and artificial neural networks
  - They forget catastrophically, lose plasticity, and collapse under extended training
- To achieve the full potential of DL and ANNs, something more is needed
  - We need an additional source of variation in under-utilized artificial neurons
  - We need to protect and preserve neurons whose variations are found useful
- To me, these needs strongly suggest that neurons should have the decentralized goal of getting other neurons to listen to them
- Though I can't demonstrate this yet

Thank you for your attention

with thanks to:

Principal investigators:

Rich Sutton
Michael Bowling
Csaba Szepesvari
Dale Schuurmans
Patrick Pilarski
Martha White
Adam White
Matthew Taylor
Marlos Machado