

Structure in physical intelligence

Peter Battaglia



NeurIPS 2018 Workshop:

“Modeling the Physical World: Perception, Learning, and Control”

Montreal - December 7, 2018

How do you understand a scene?



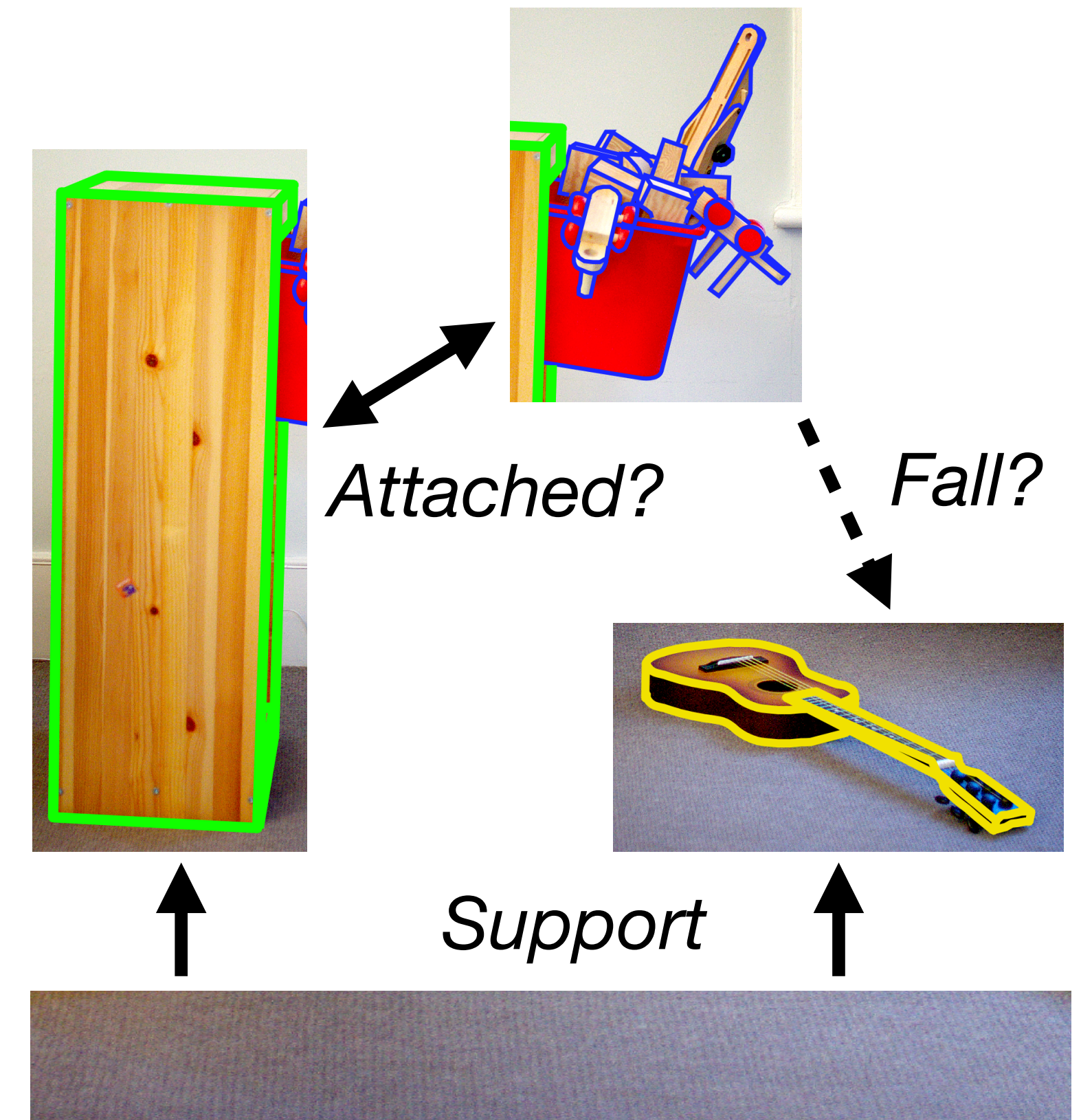
How do you understand a scene?

1. Parse it into physical objects and relations



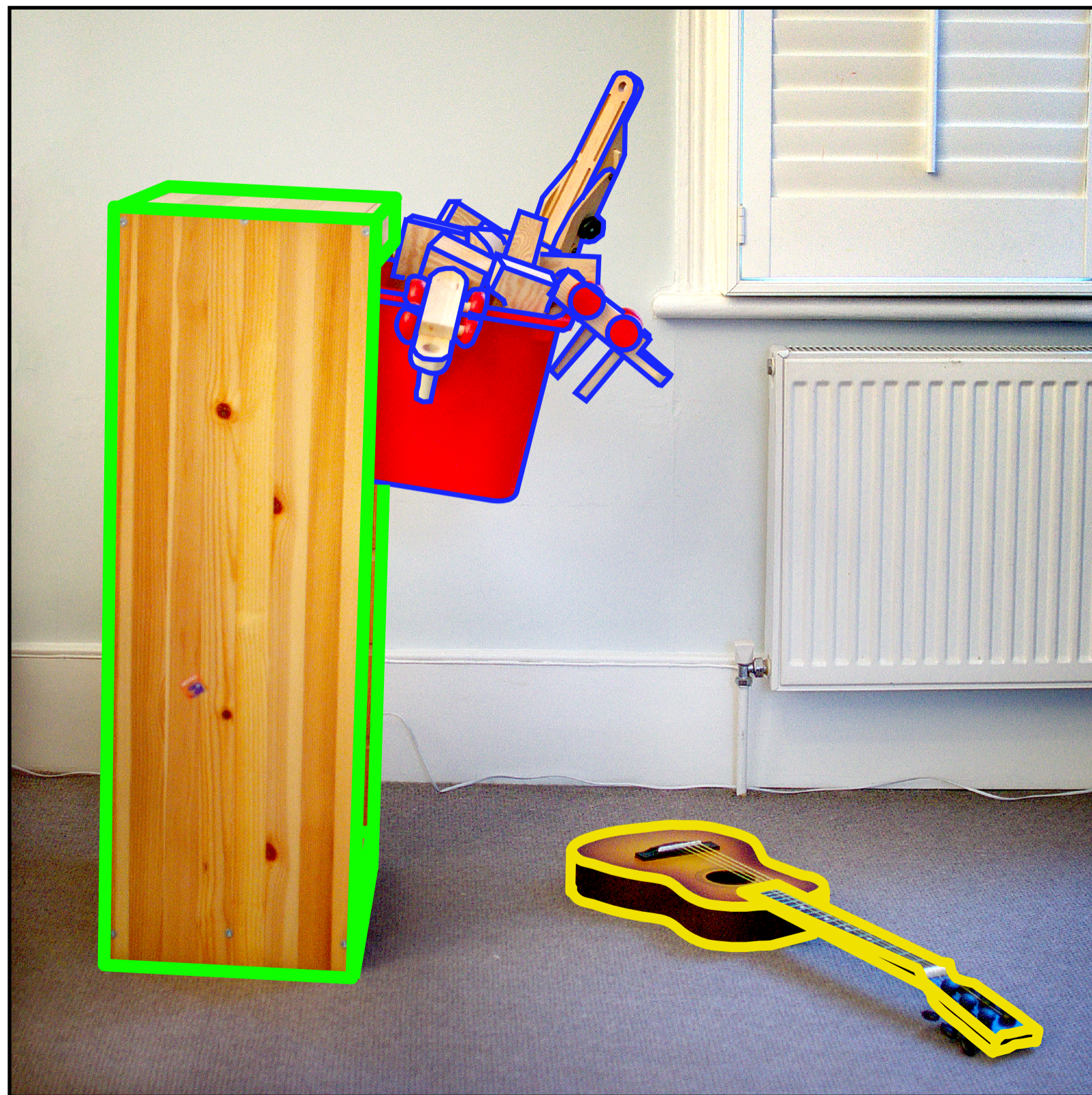
How do you understand a scene?

1. Parse it into physical objects and relations
2. Reason about the objects and their interactions

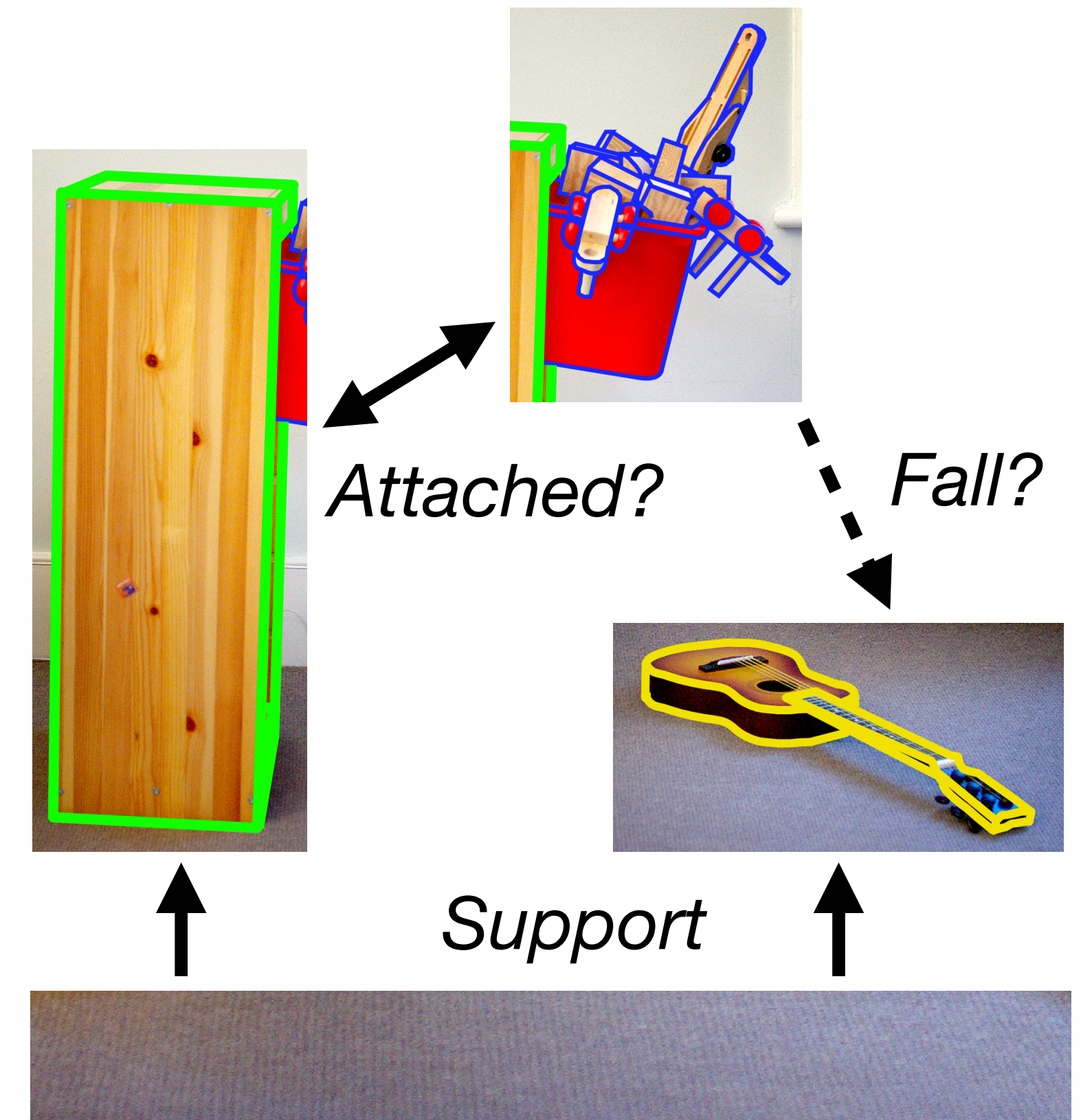


How do you understand a scene?

1. Parse it into physical objects and relations
2. Reason about the objects and their interactions



"Precarious"

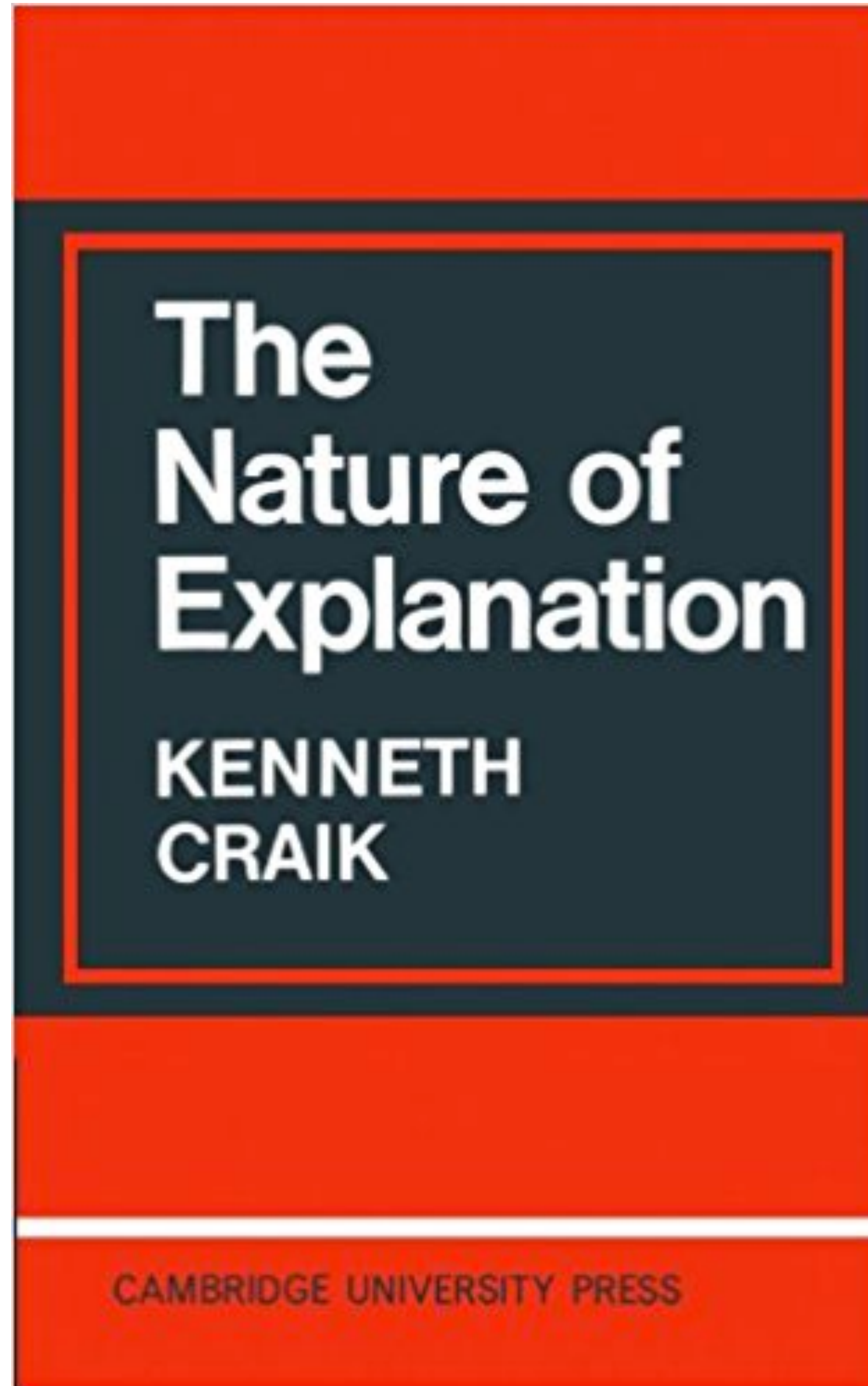


"Infinite use of finite means"

- von Humboldt, on the productivity of language

"Precarious"





Kenneth Craik, "The Nature of Explanation", 1943:

*"If the organism carries a '**small-scale model**' of external reality and of **its own possible actions** within its head, it is able to **try out various alternatives, conclude which is the best of them, react to future situations before they arise**, utilize the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it."* (pg 61)

*"This concept of '**thinghood**' is of fundamental importance for any **theory of thought**."* (pg 77)

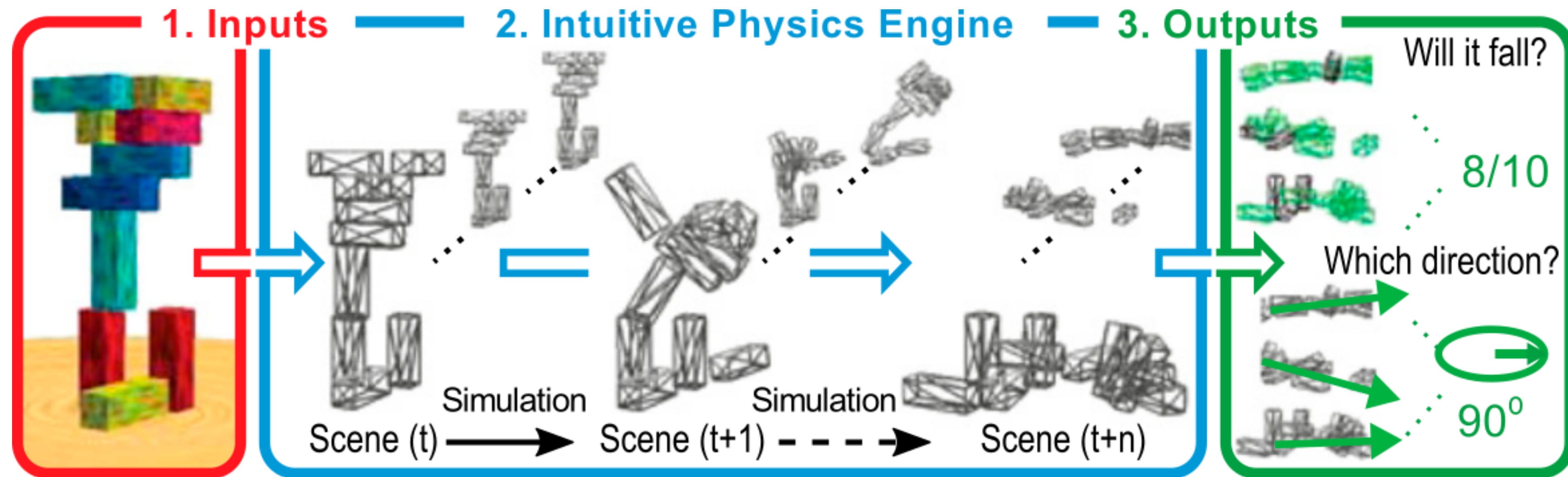
Claim: Human intelligence is structured

Founded on objects, relations, reasoning

- Objects and relations reflect *decisions* made by evolution, experience, and task demands about how to represent the world in an *efficient and useful way*
- Structure in our *core cognitive knowledge* evident very early in infancy (Spelke)
- *Model-building* over recognizing patterns (Tenenbaum)
- Combinatorial generalization via compositionality ("infinite use of finite means")

What is the mechanism of human intuitive physics?

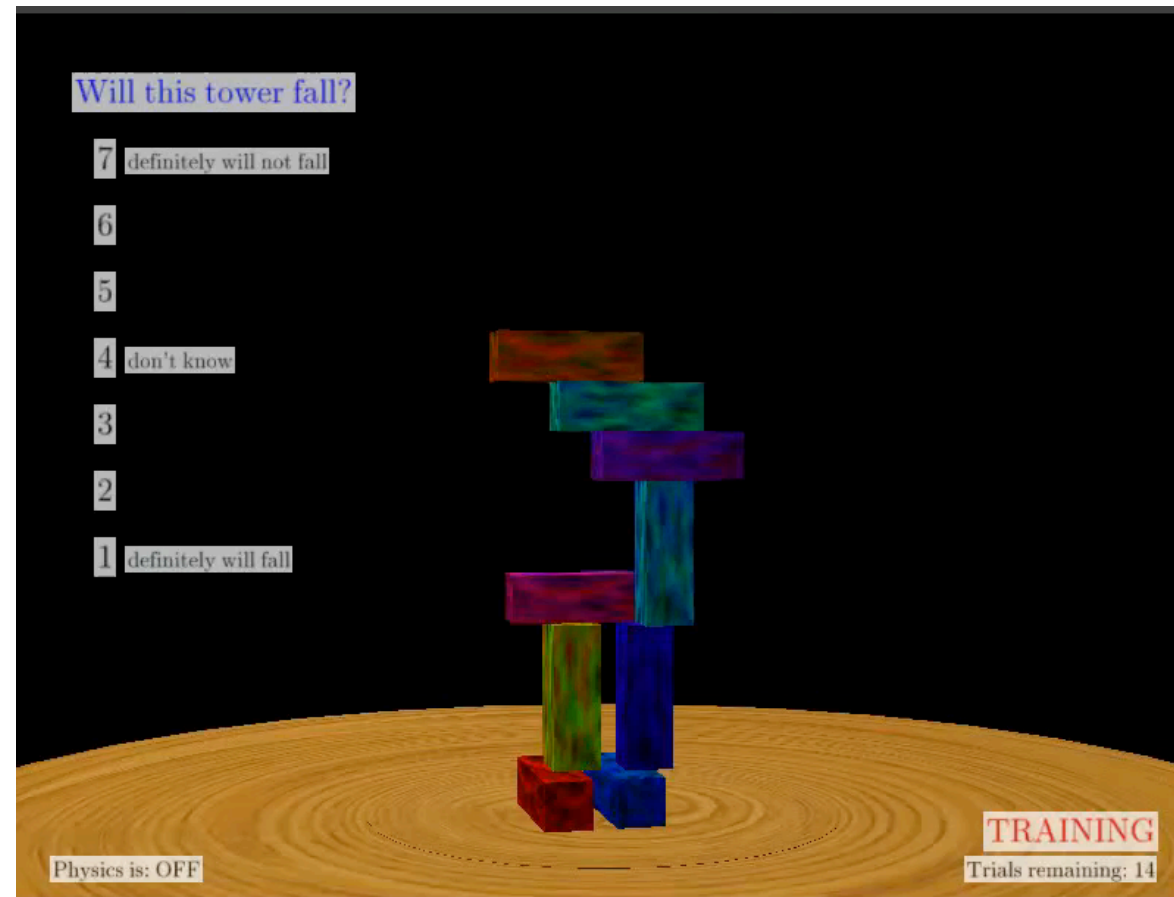
Intuitive Physics Engine: the "physics engine in the head"



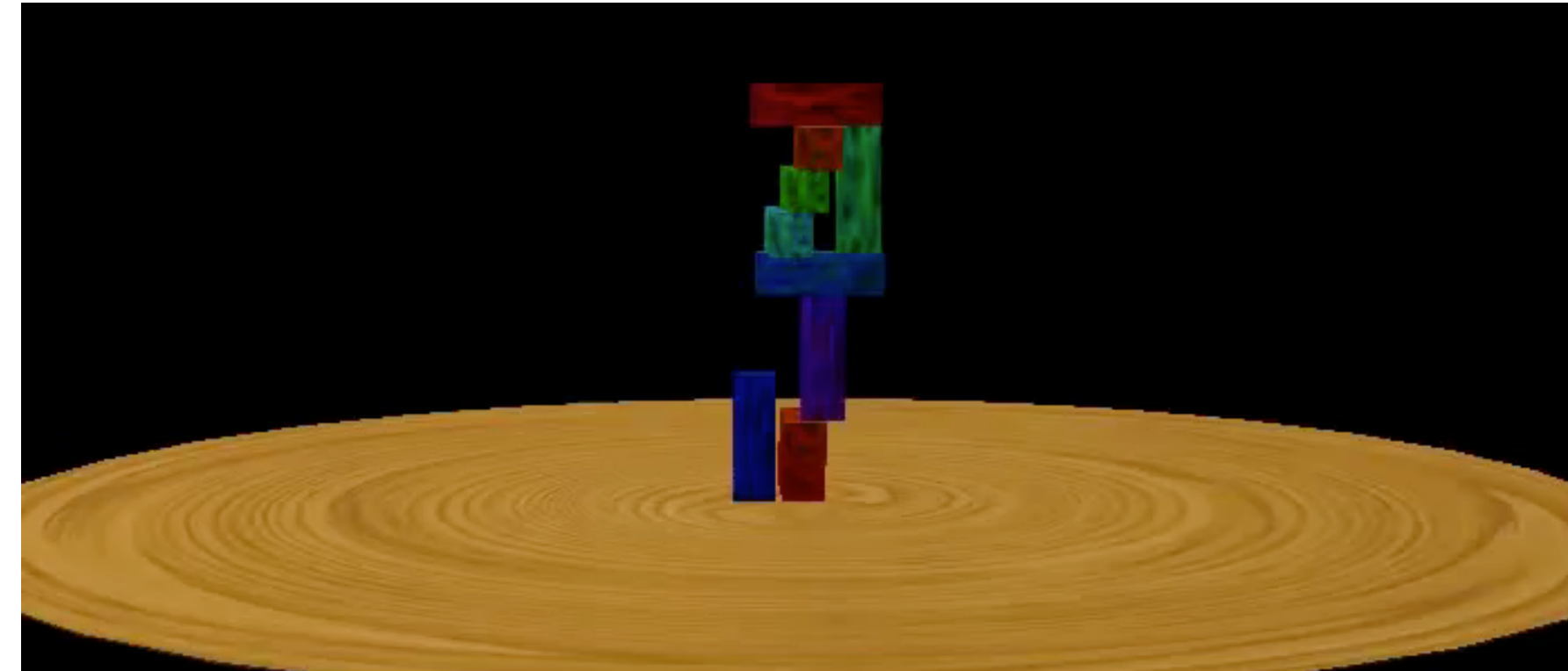
Battaglia, Hamrick, Tenenbaum, 2013, PNAS

Experiments: What will happen? Why?

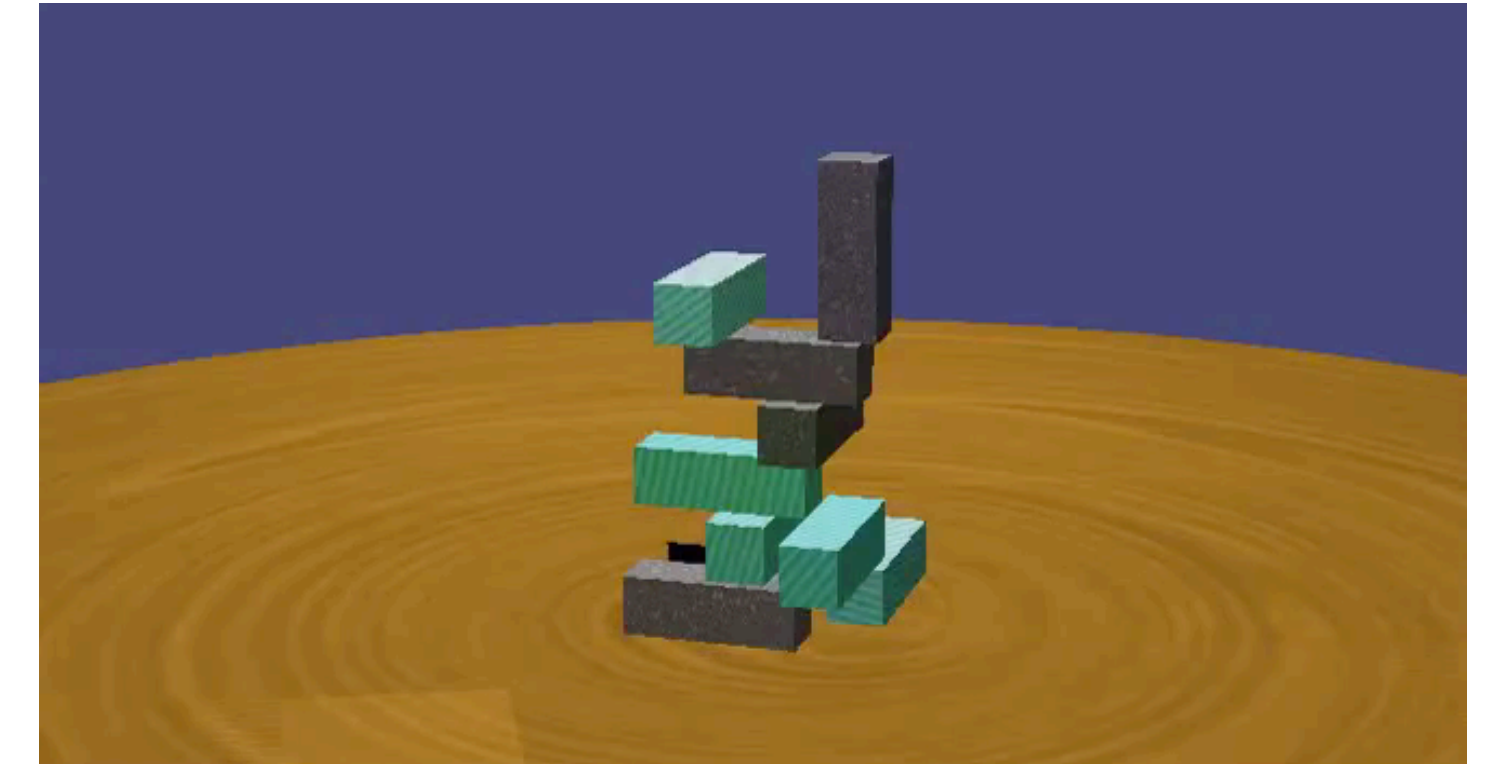
Will it fall?



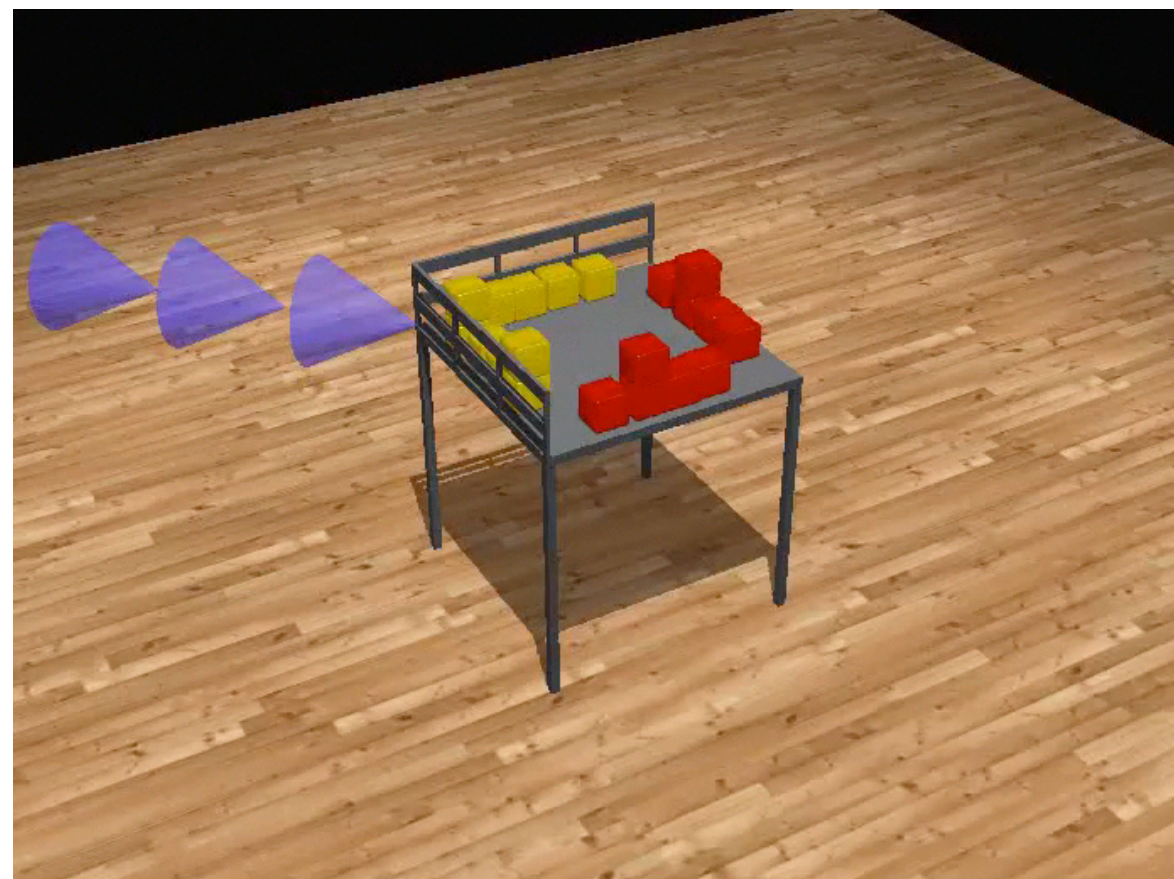
In which direction?



Different masses



Complex scenes



Infer the mass



Predict fluids



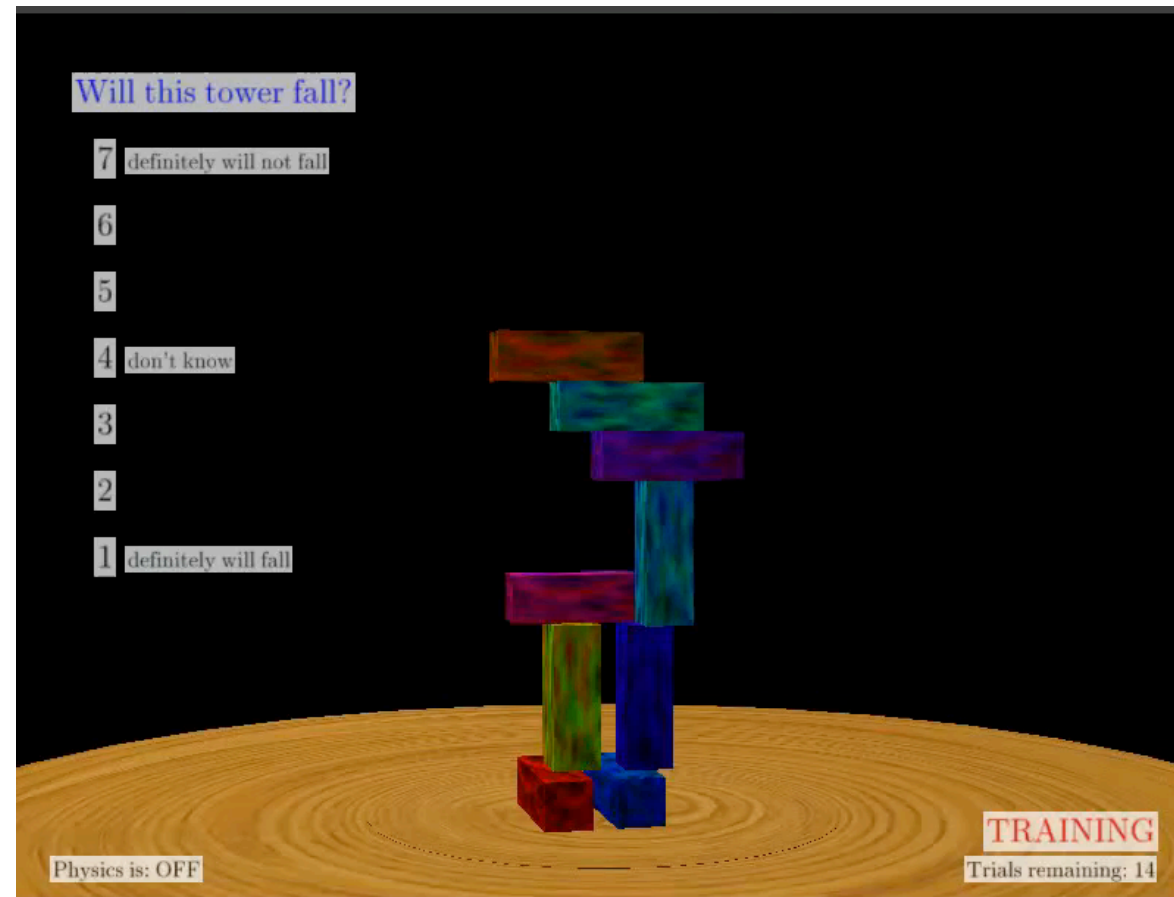
Battaglia et al., 2013

Hamrick et al., 2016

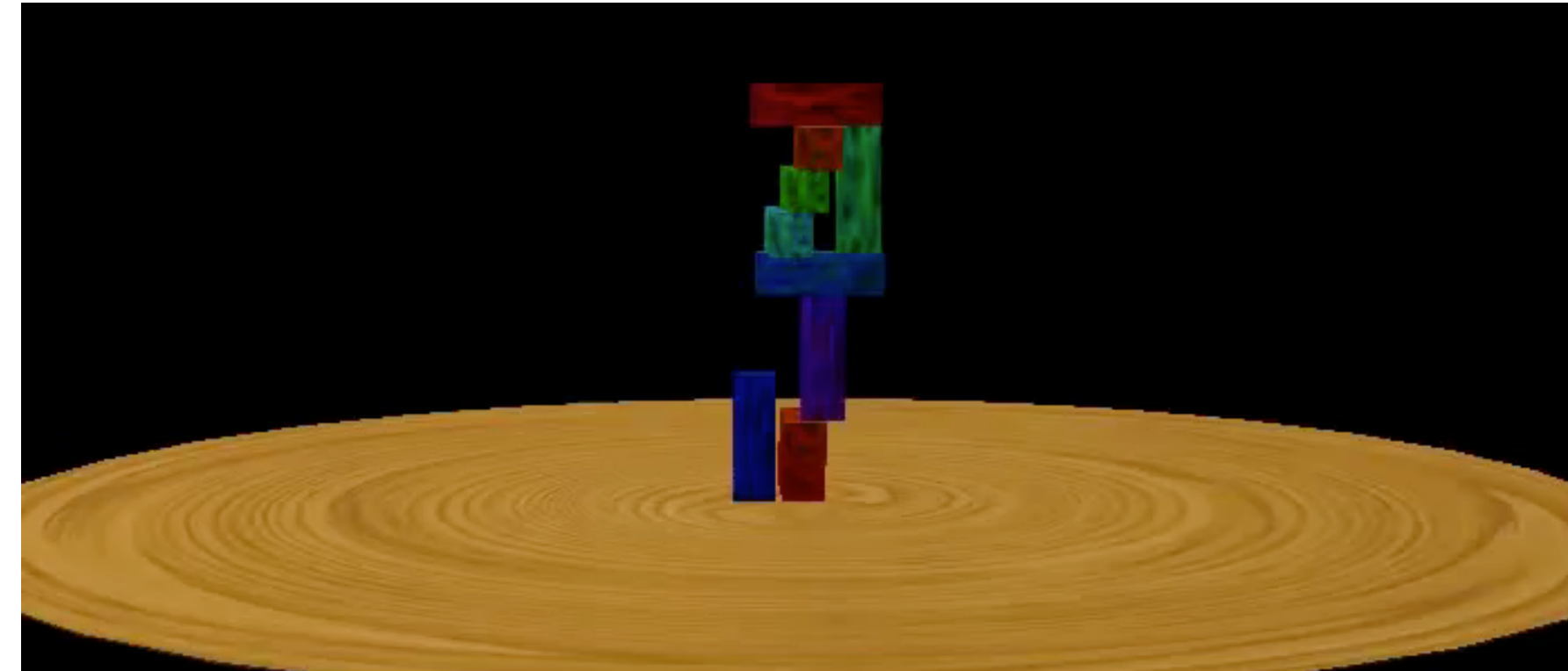
Bates et al., 2015, 2018

Experiments: What will happen? Why?

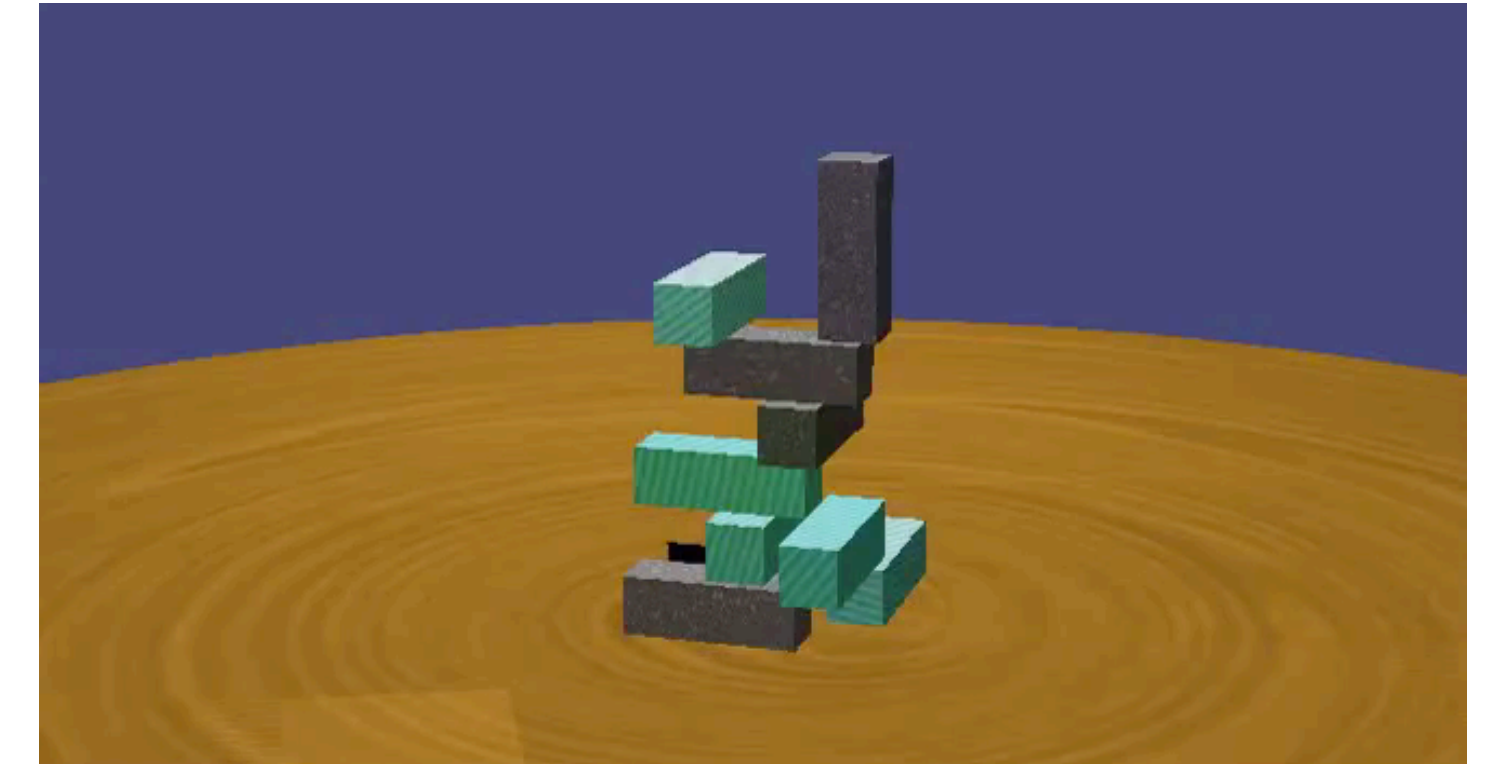
Will it fall?



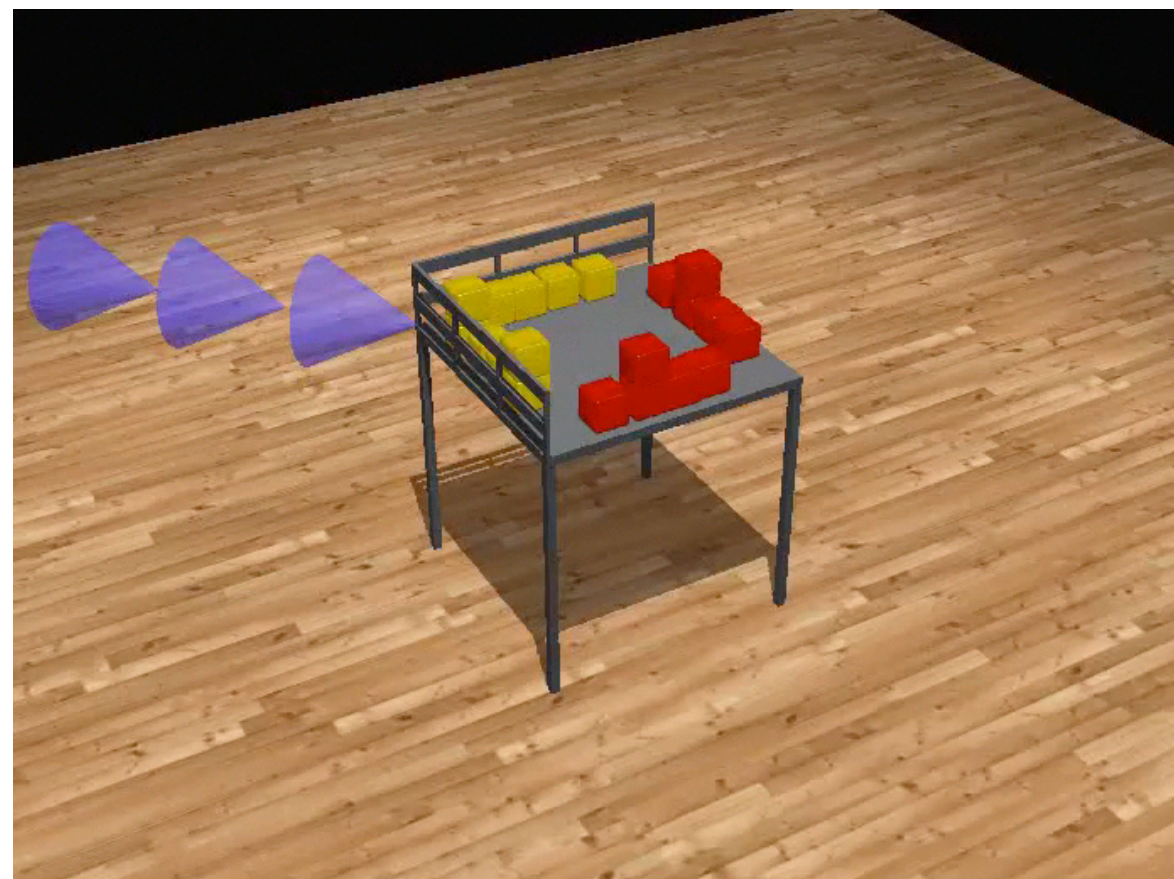
In which direction?



Different masses



Complex scenes



Infer the mass



Predict fluids



Battaglia et al., 2013

Hamrick et al., 2016

Bates et al., 2015, 2018

Message from cognition:

Humans use richly structured representations of objects and relations to reason about, and interact with, their everyday environment.

Message from cognition:

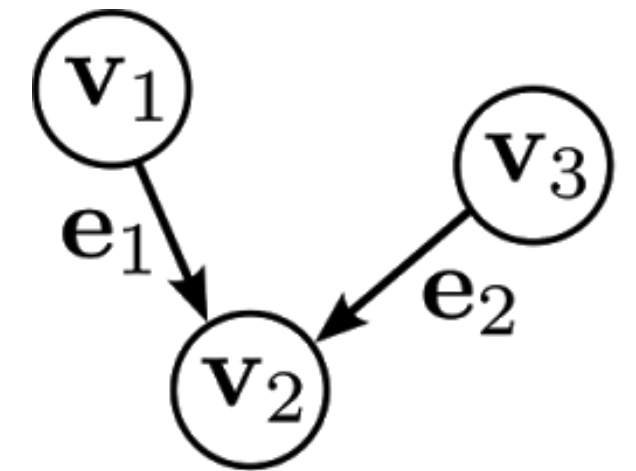
Humans use richly structured representations of objects and relations to reason about, and interact with, their everyday environment.

What insights does humans' structured intelligence offer AI?

We need better object- and relation-centric models in AI

A **graph** is a natural way to represent entities and their relations:

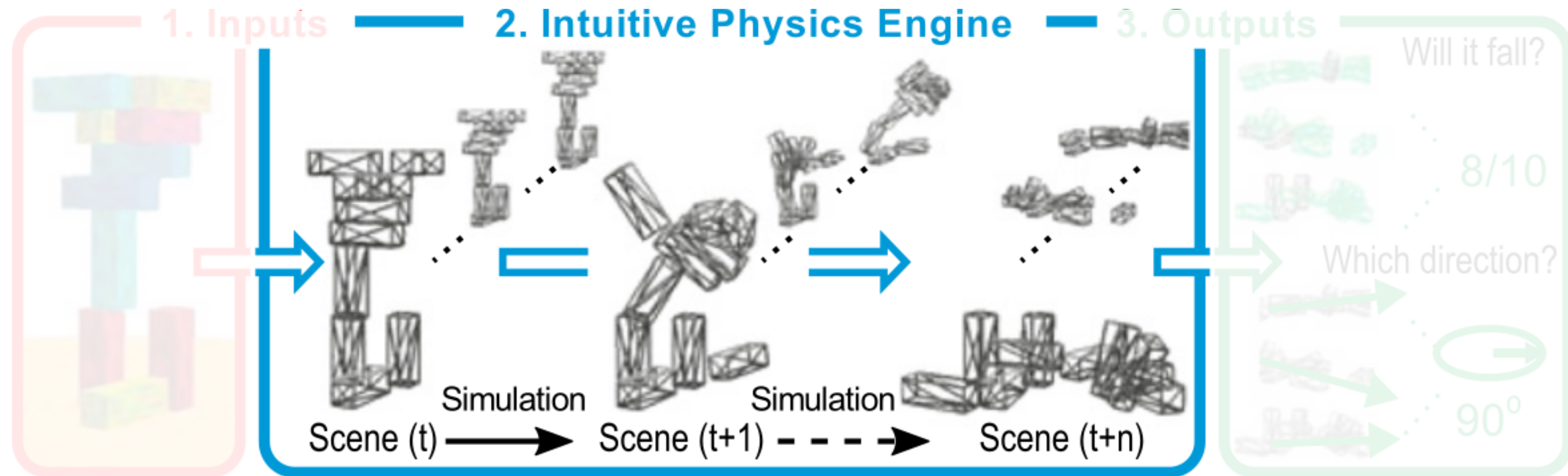
- “Nodes” correspond to entities, objects, events, etc.
- “Edges” correspond to their relations, interactions, transitions, etc.
- Inferences about entities and relations respect the graphical structure.



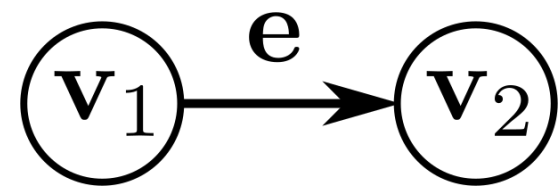
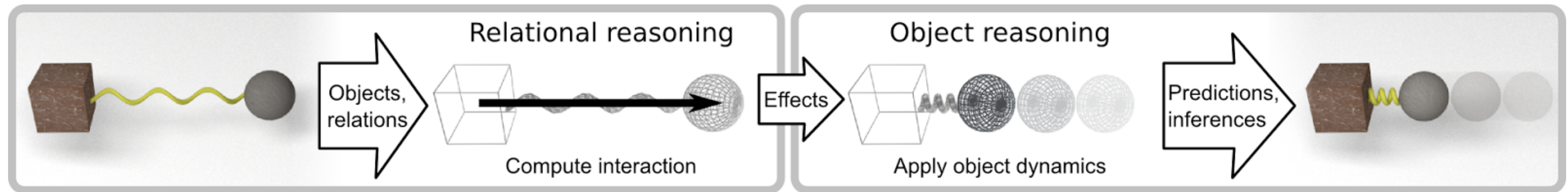
Graphs can capture data from many complex systems:

- Physical systems
- Scene graphs
- Social networks
- Linguistic structure
- Programs
- Search trees
- Communication networks
- Transportation networks
- Chemical structure
- Phylogenetic trees

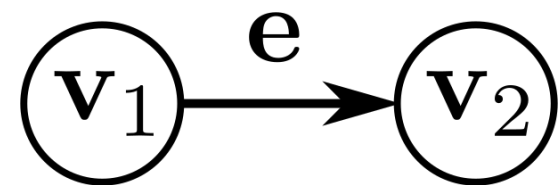
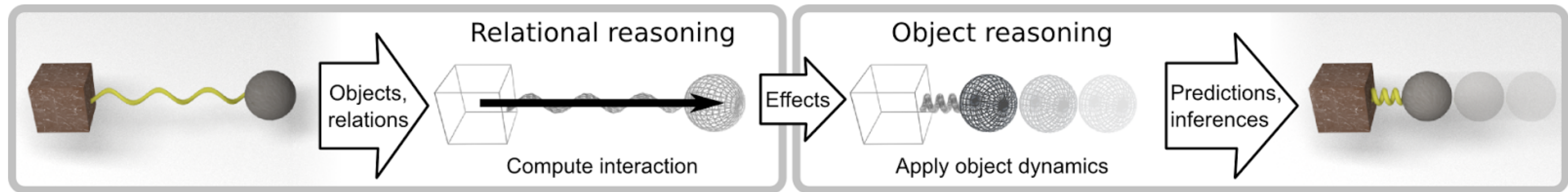
Intuitive physics as reasoning about graphs



Intuitive physics as reasoning about graphs



Interaction Network

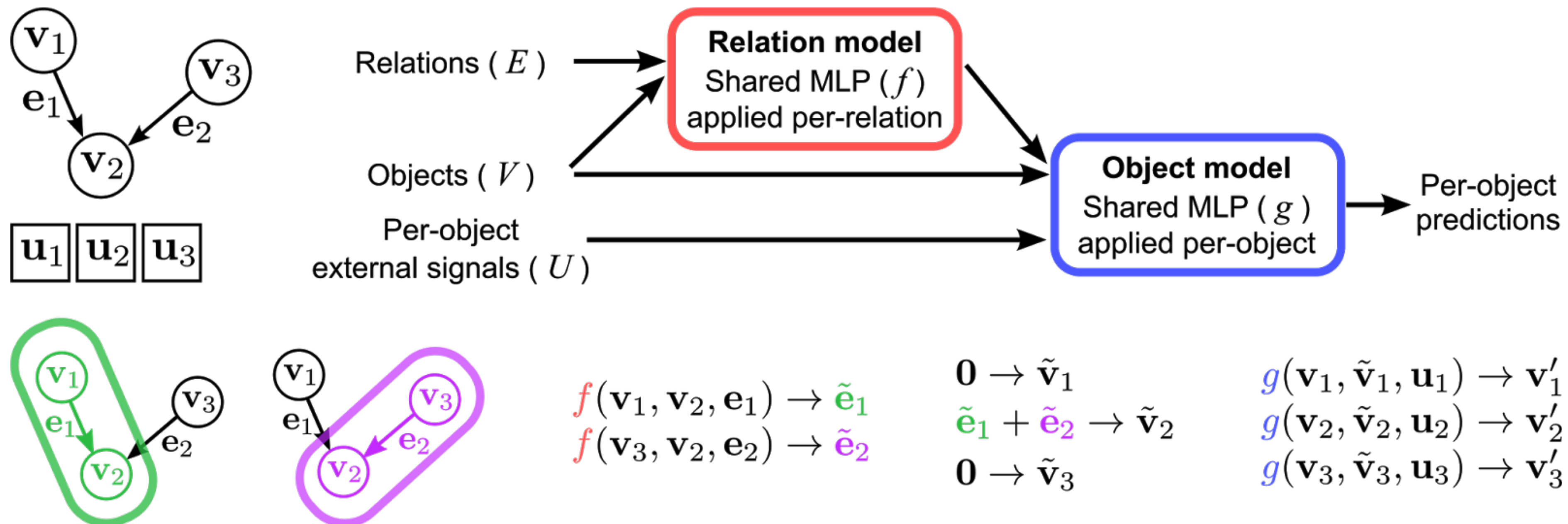
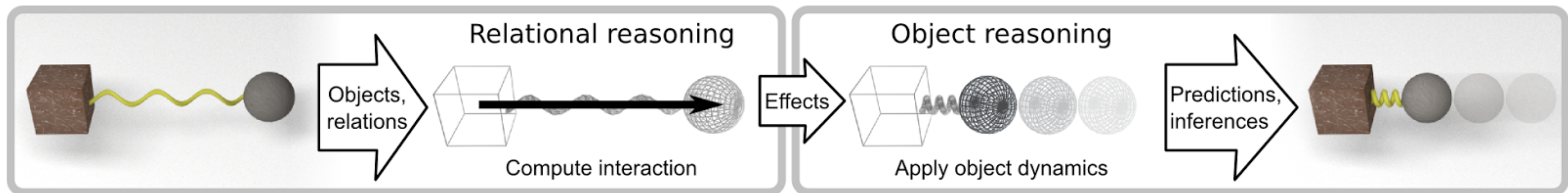


Strong relational inductive bias: Deep learning architecture which operates on graphs

Related to the broad family of "Graph Neural Networks" (Scarselli et al, 2009; Li et al, 2015) and "Message-Passing Neural Networks" (Gilmer et al., 2017).
Chang et al. (2016) also proposed a similar version in parallel.

Battaglia et al., 2016, NeurIPS

Interaction Network

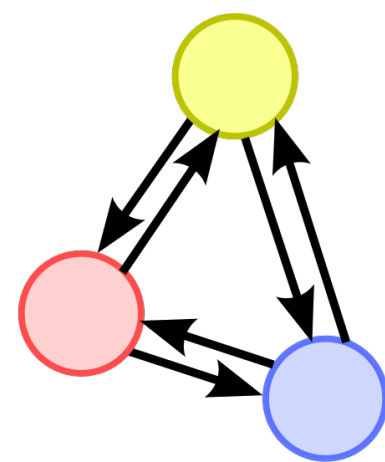
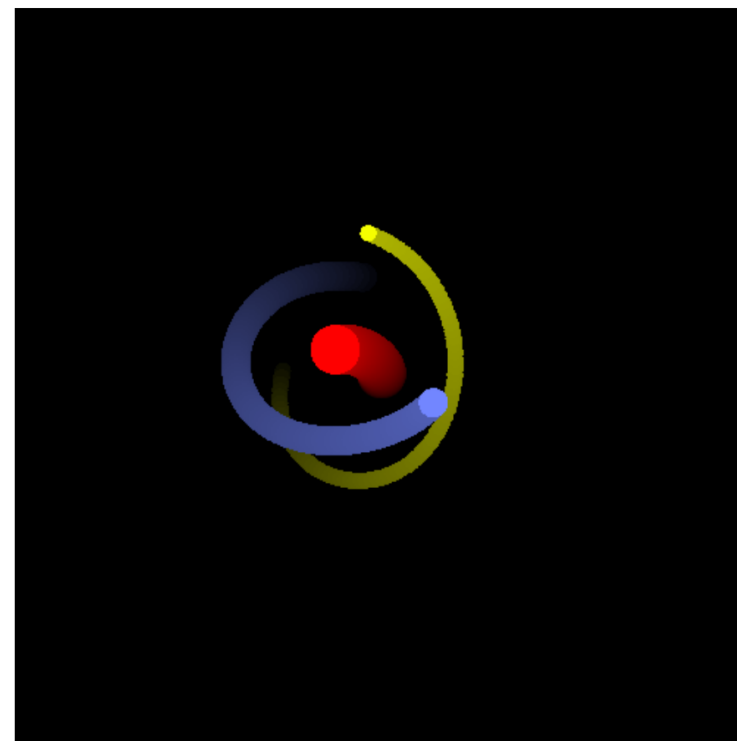


Battaglia et al., 2016, NeurIPS

Interaction Network

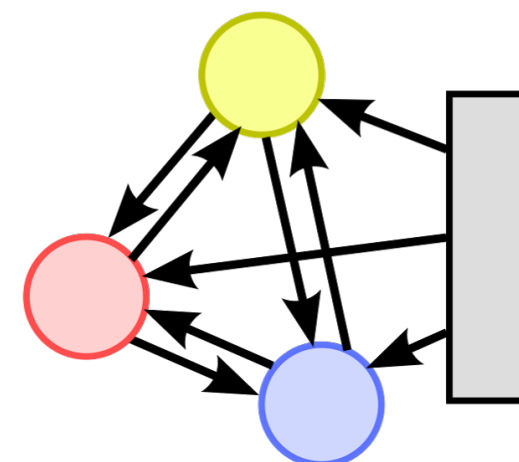
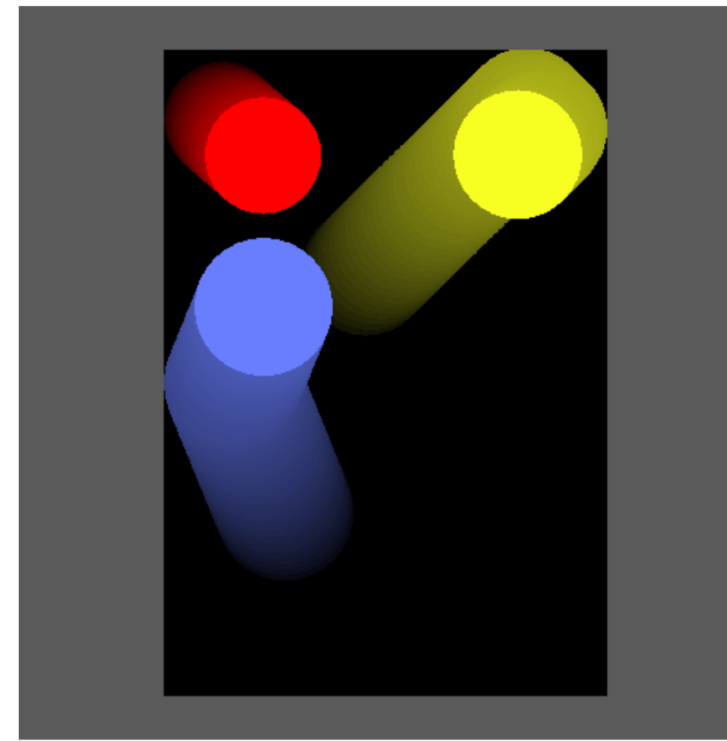
Can learn a general-purpose physics engine, simulating future states from initial ones

n-body



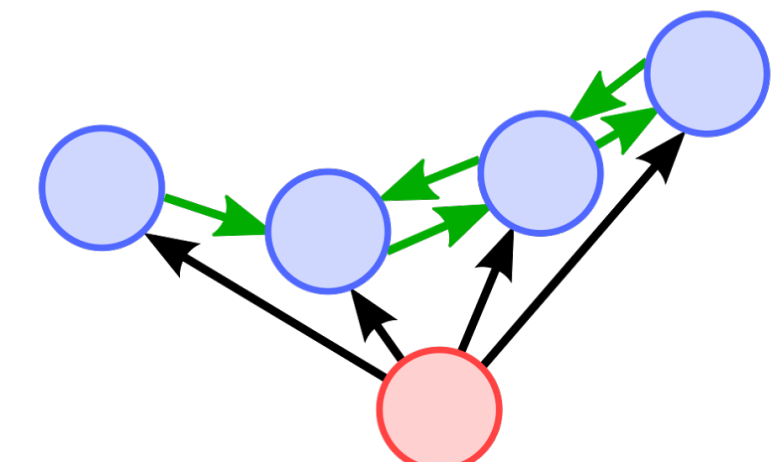
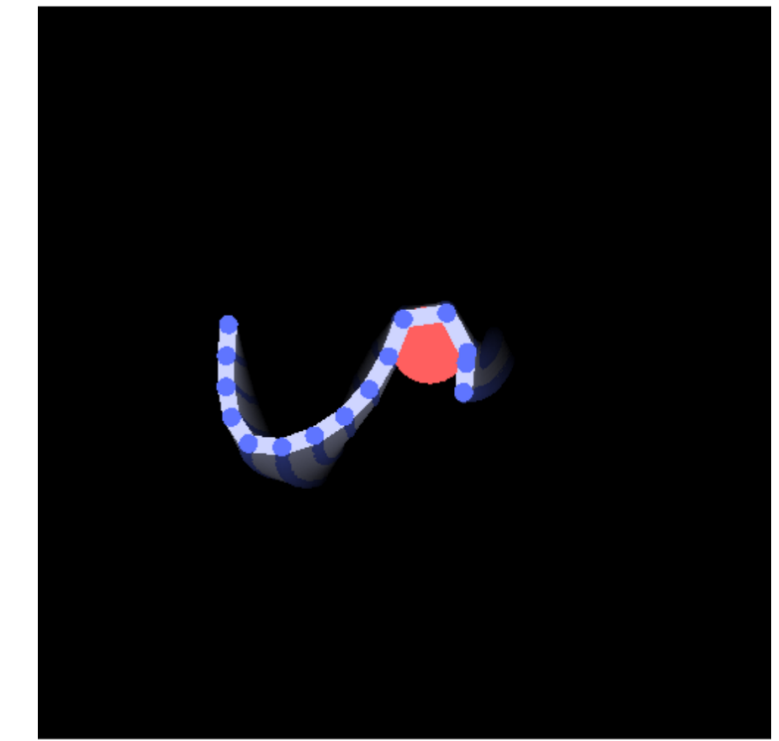
Gravitational forces

Balls



Rigid collisions between
walls and balls

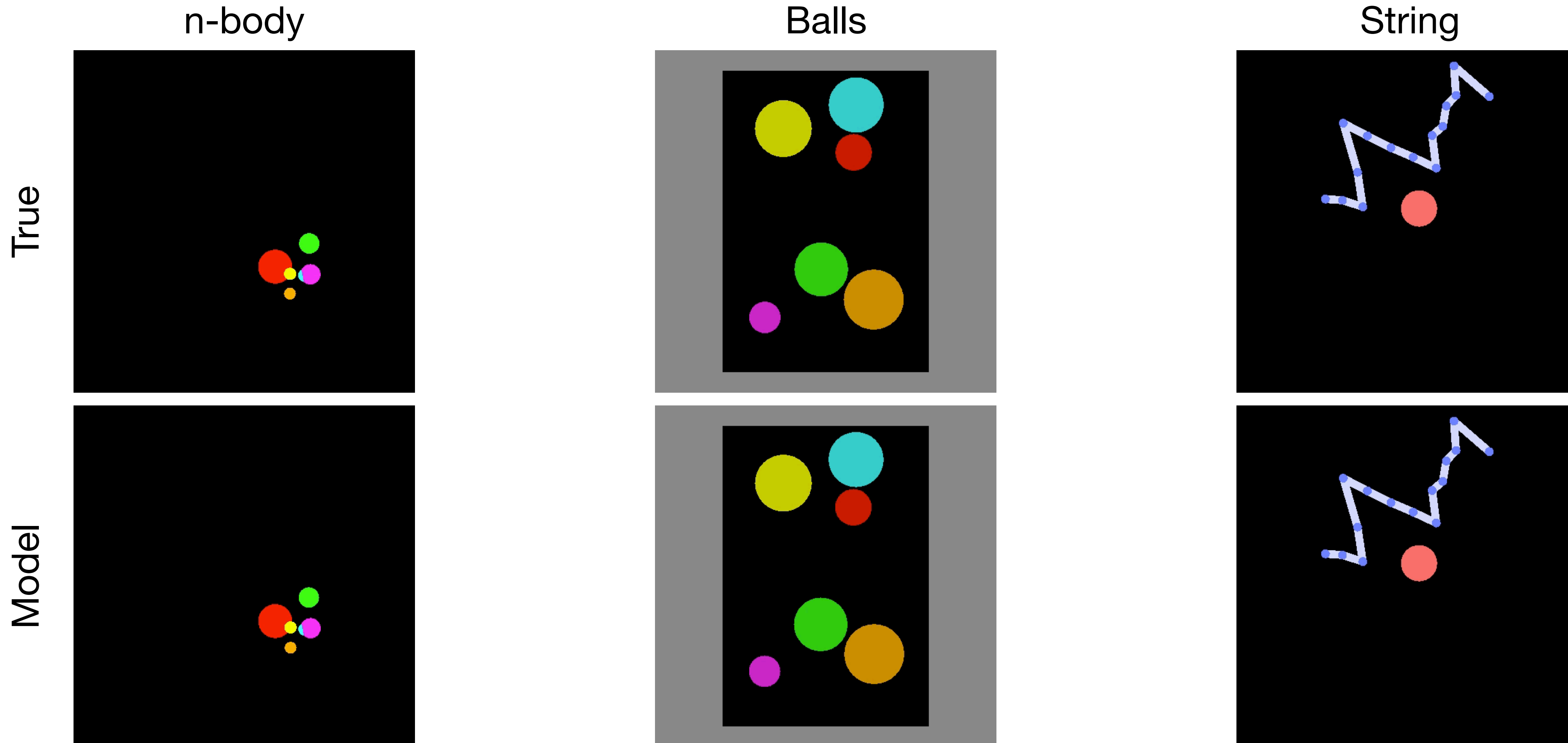
String



Springs and rigid collisions

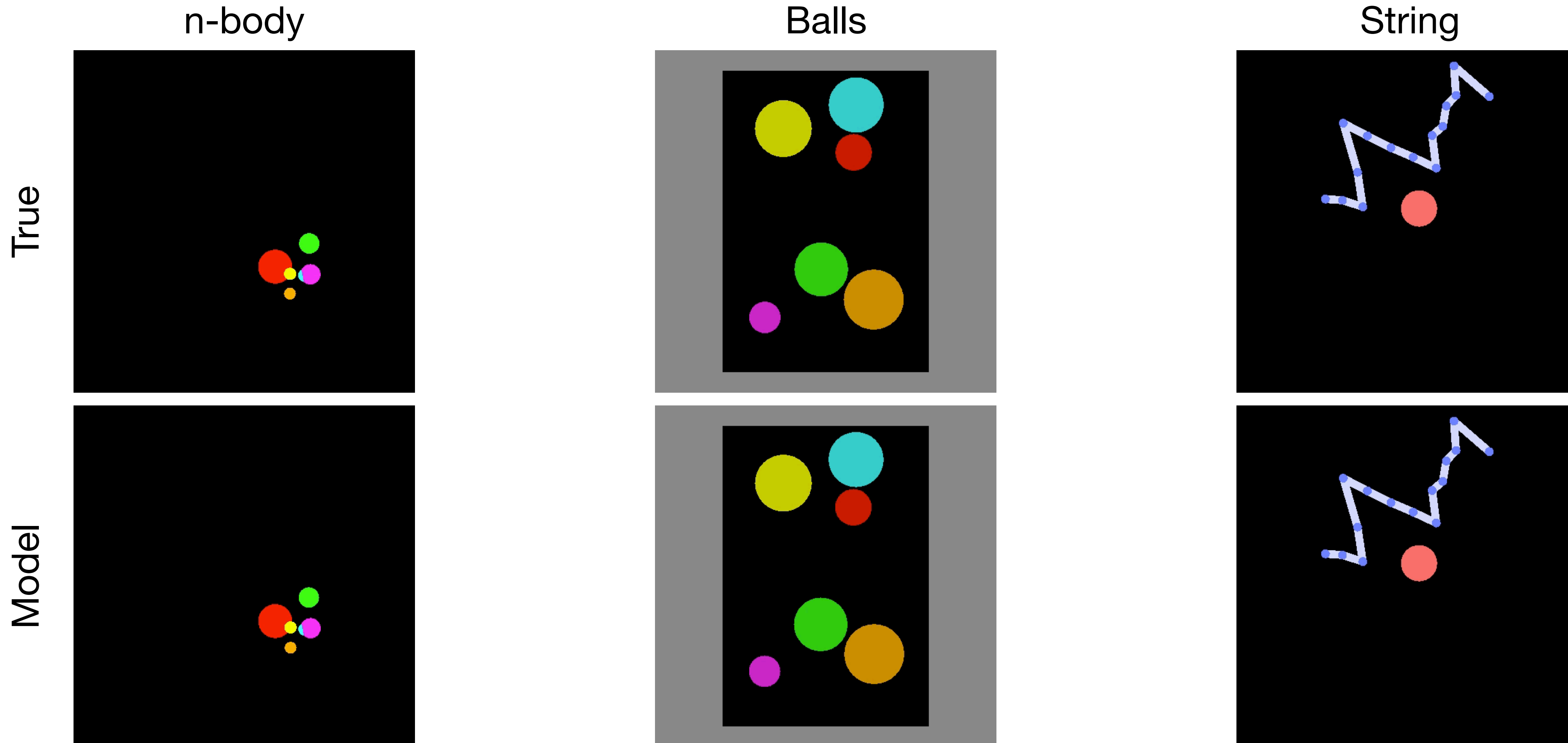
Battaglia et al., 2016, NeurIPS

1000-step rollouts from 1-step supervised training



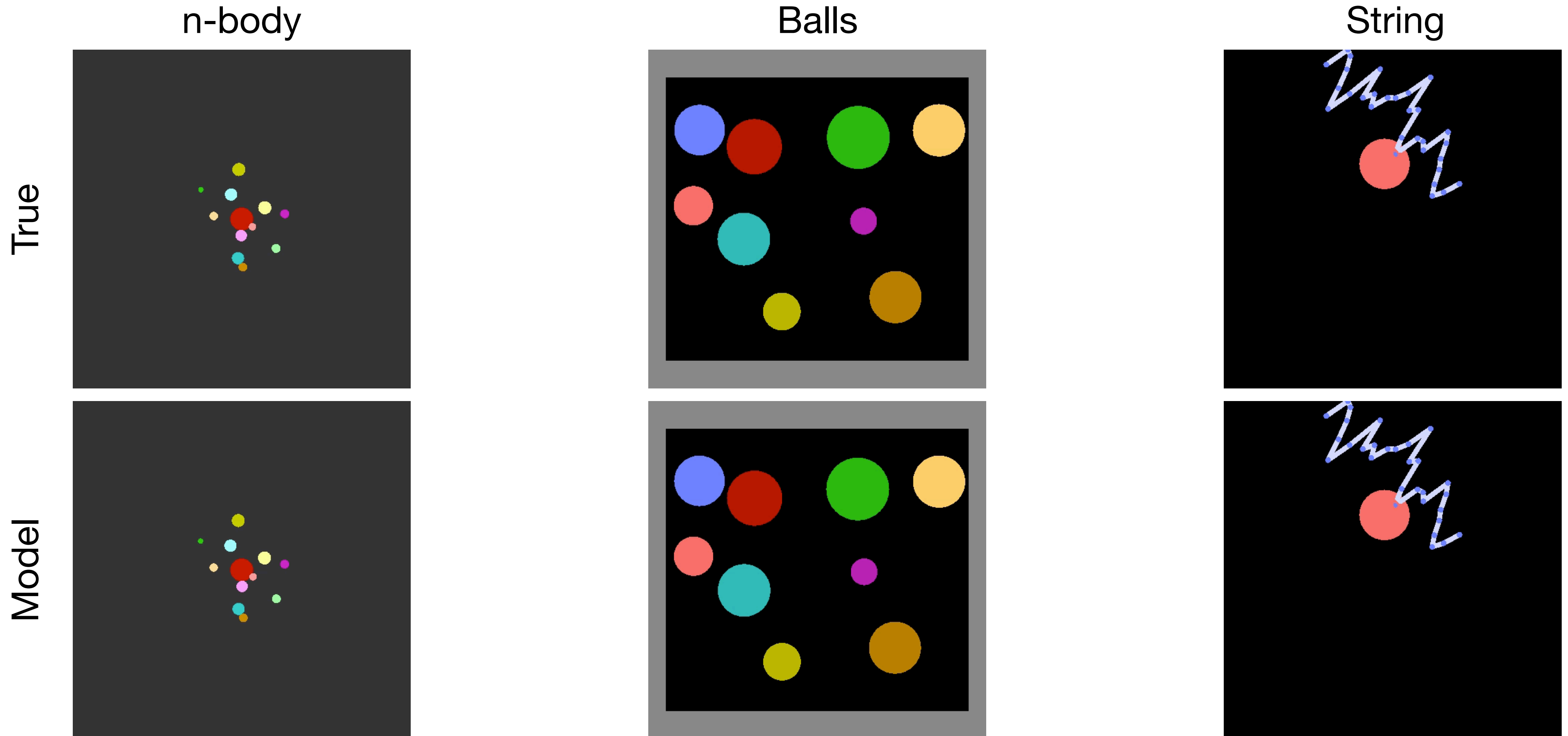
Battaglia et al., 2016, NeurIPS

1000-step rollouts from 1-step supervised training



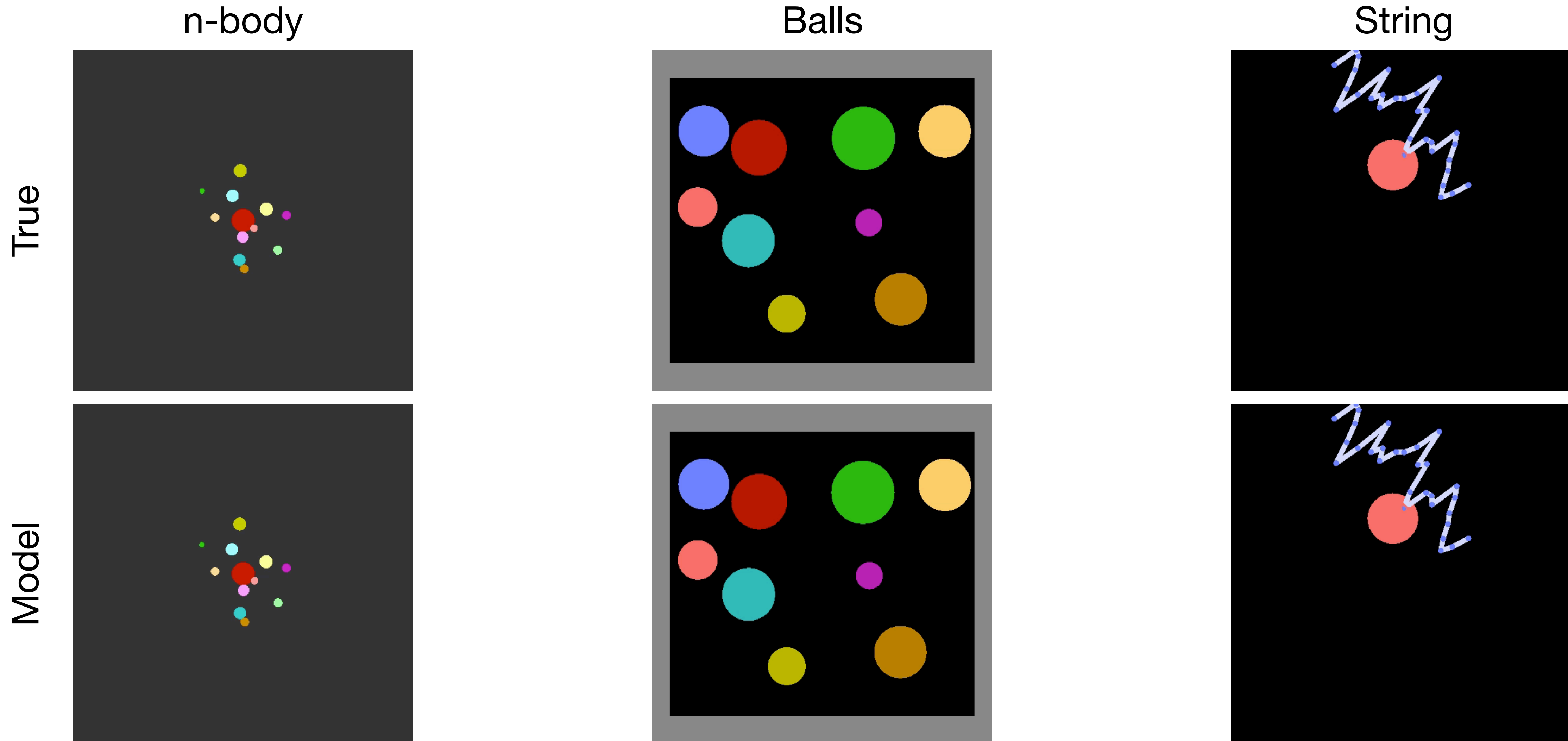
Battaglia et al., 2016, NeurIPS

Zero-shot generalization to larger systems



Battaglia et al., 2016, NeurIPS

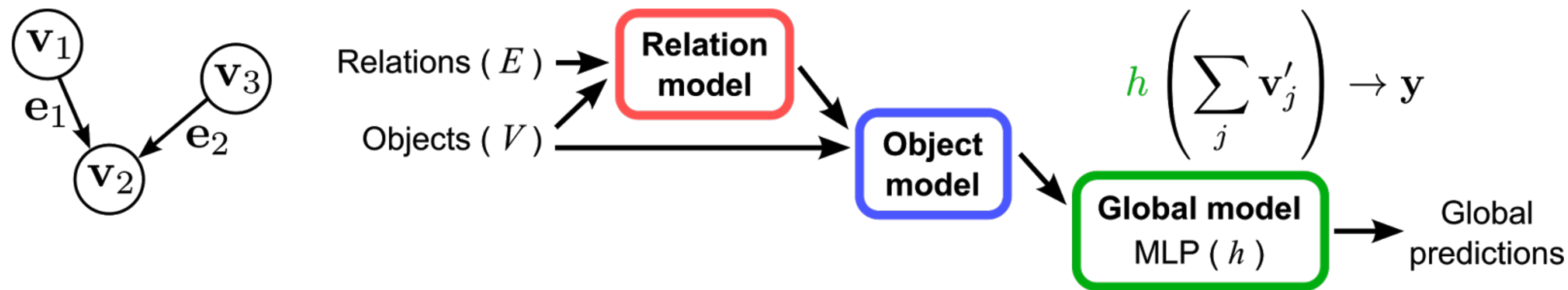
Zero-shot generalization to larger systems



Battaglia et al., 2016, NeurIPS

Interaction Network for system-level predictions

A "global model" can be added, which aggregates the per-object outputs to make predictions.

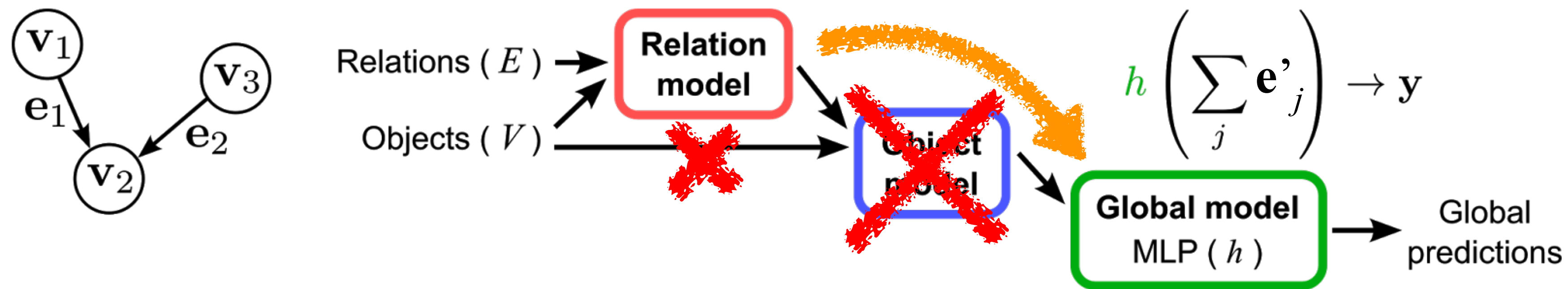


Can be trained to predict potential energy of a system, outperforming MLP baselines

Battaglia et al., 2016, NeurIPS

Relation Network

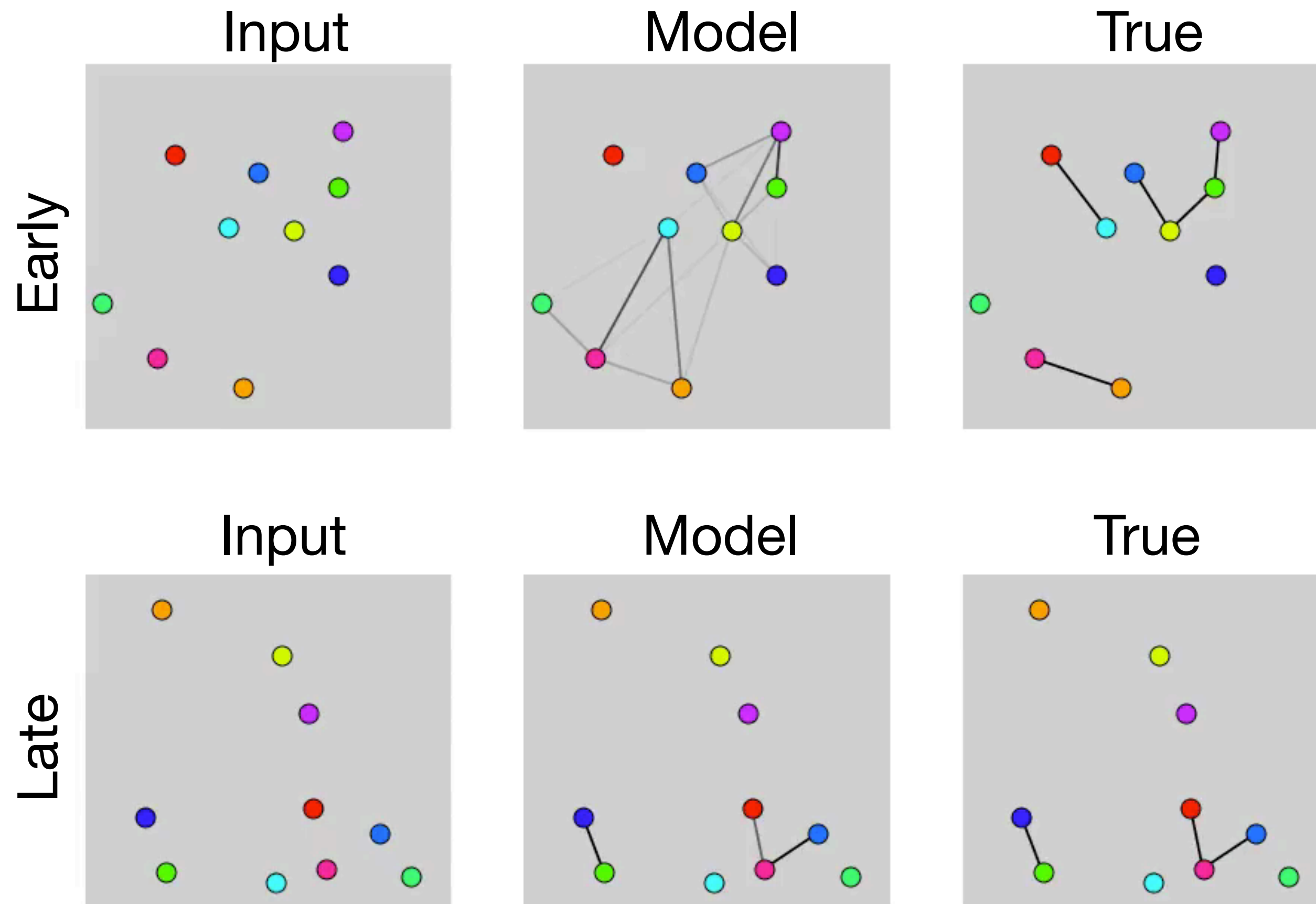
Remove “object model” and predict global outputs only using “relation model”’s output



Raposo et al., 2017, ICLR workshop;
Santoro et al., 2017, NeurIPS

Relation Networks can infer relations in dot motion

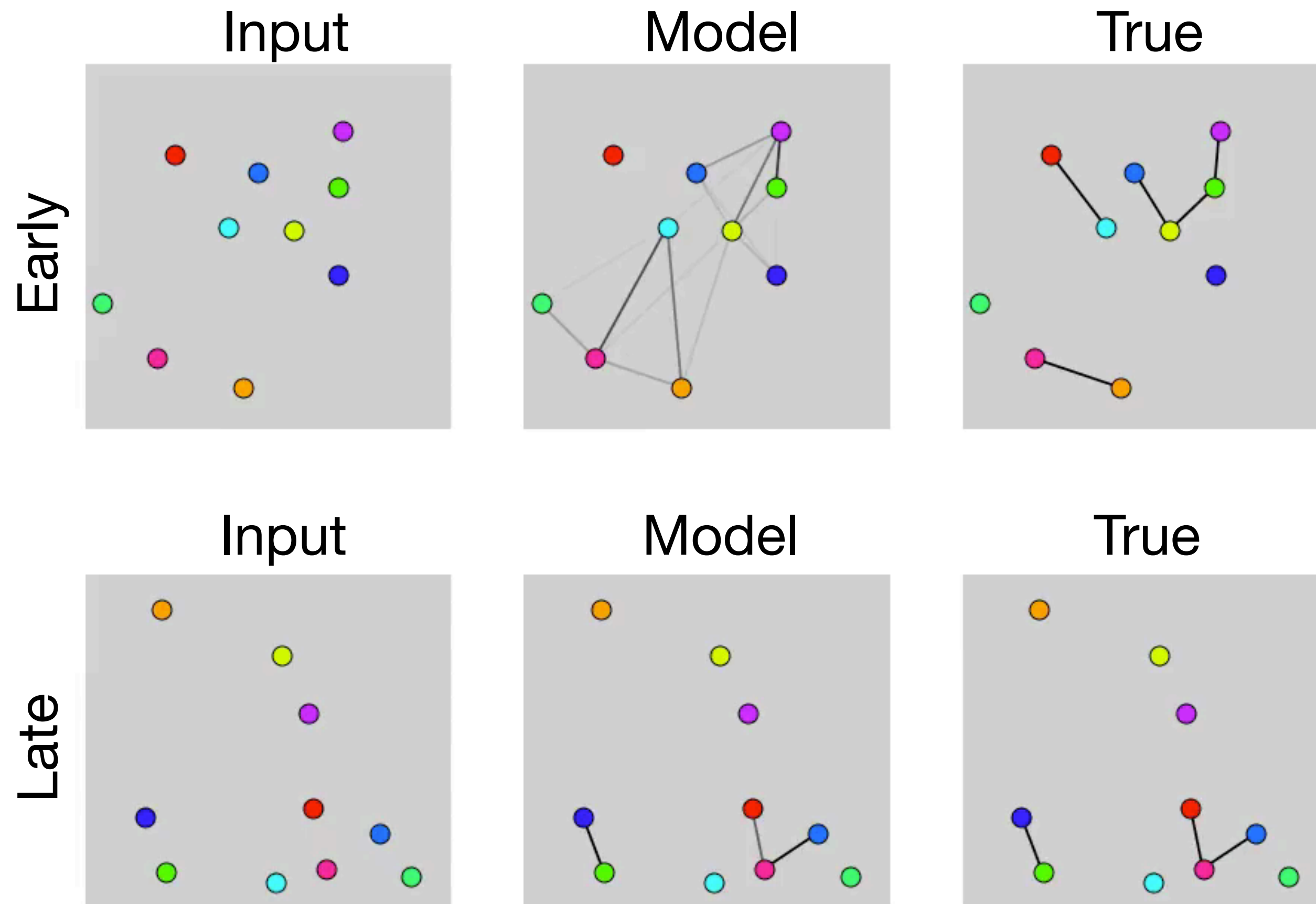
Trained on mass-spring systems



Santoro et al., 2017, NeurIPS

Relation Networks can infer relations in dot motion

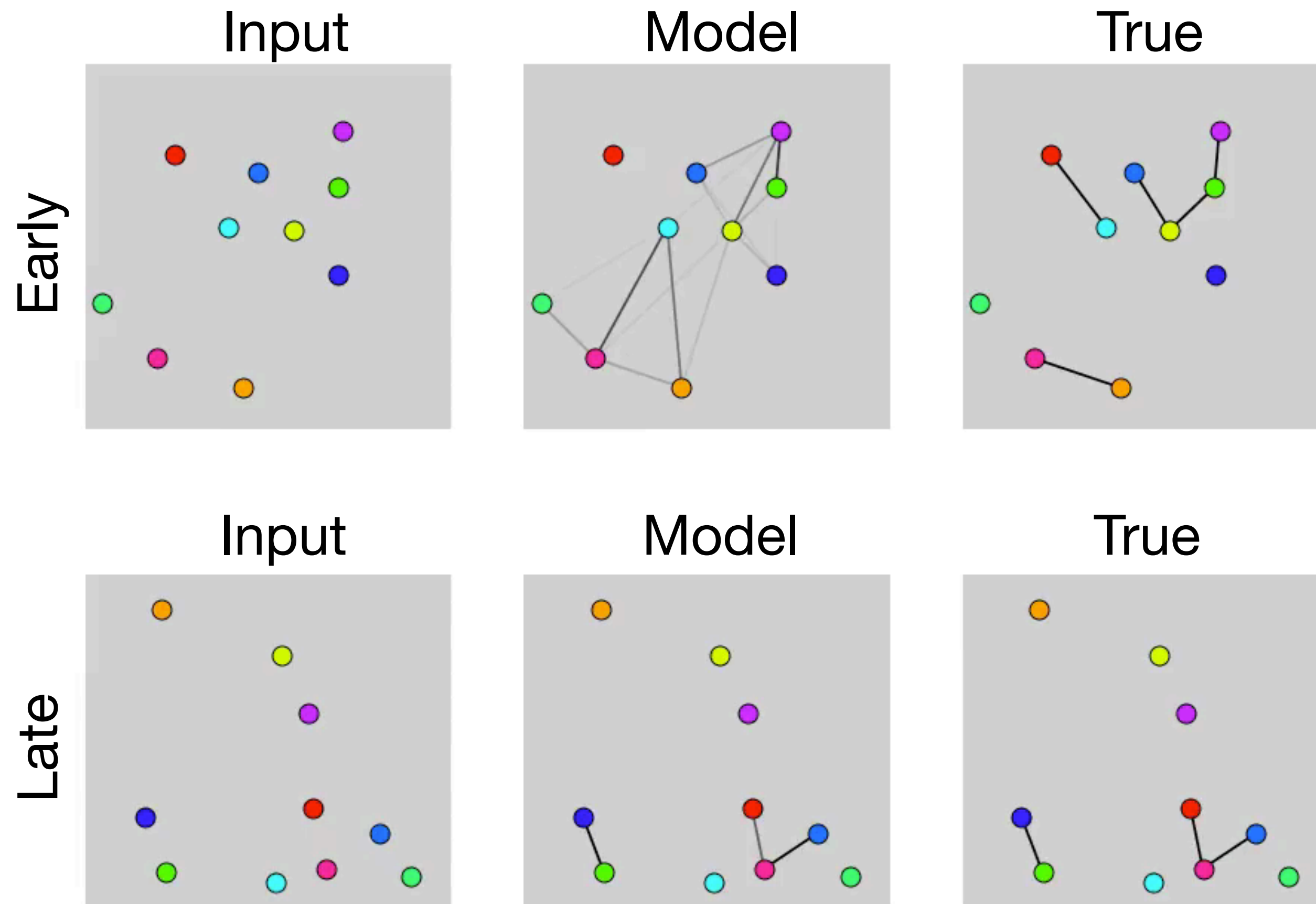
Trained on mass-spring systems



Santoro et al., 2017, NeurIPS

Relation Networks can infer relations in dot motion

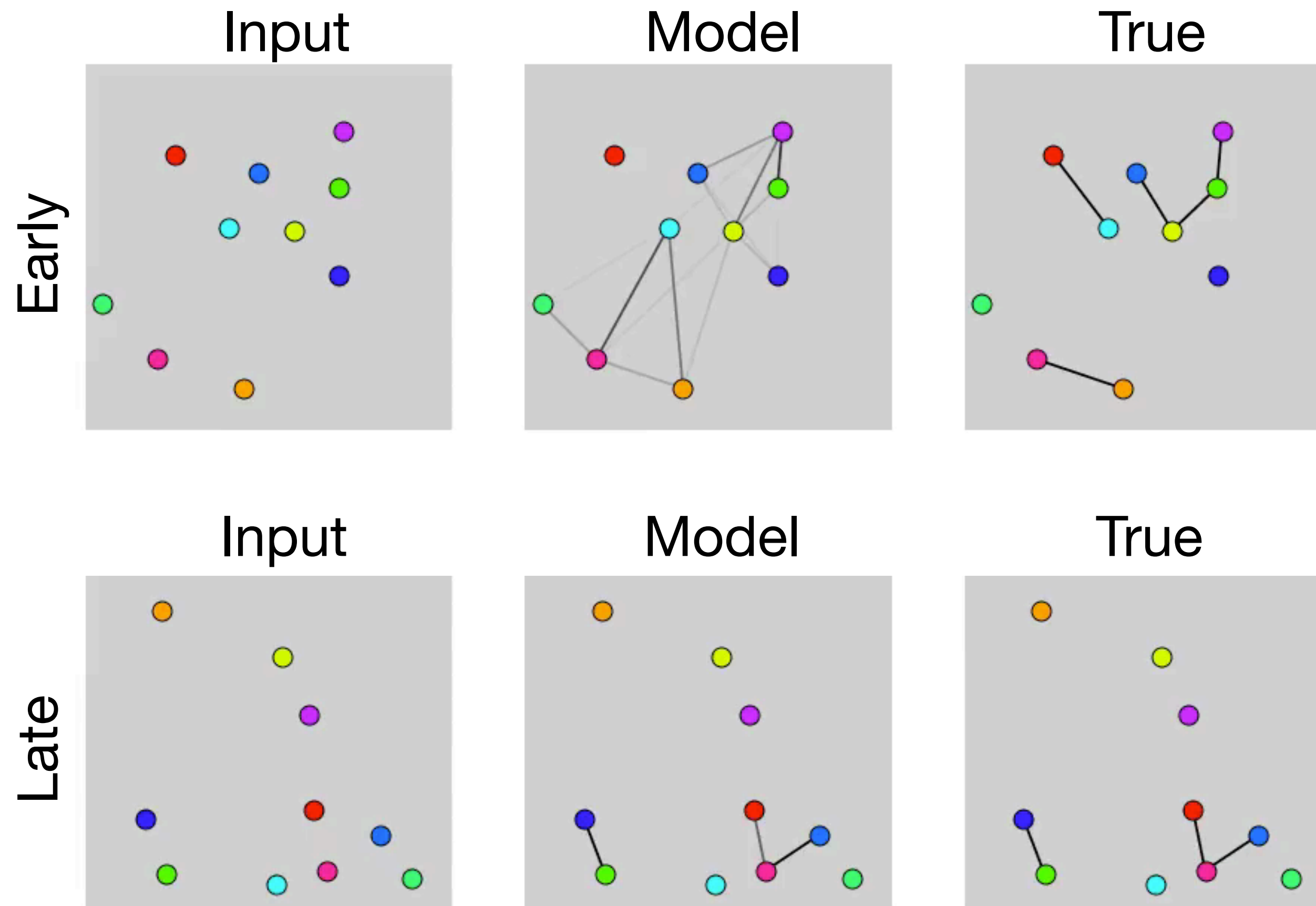
Trained on mass-spring systems



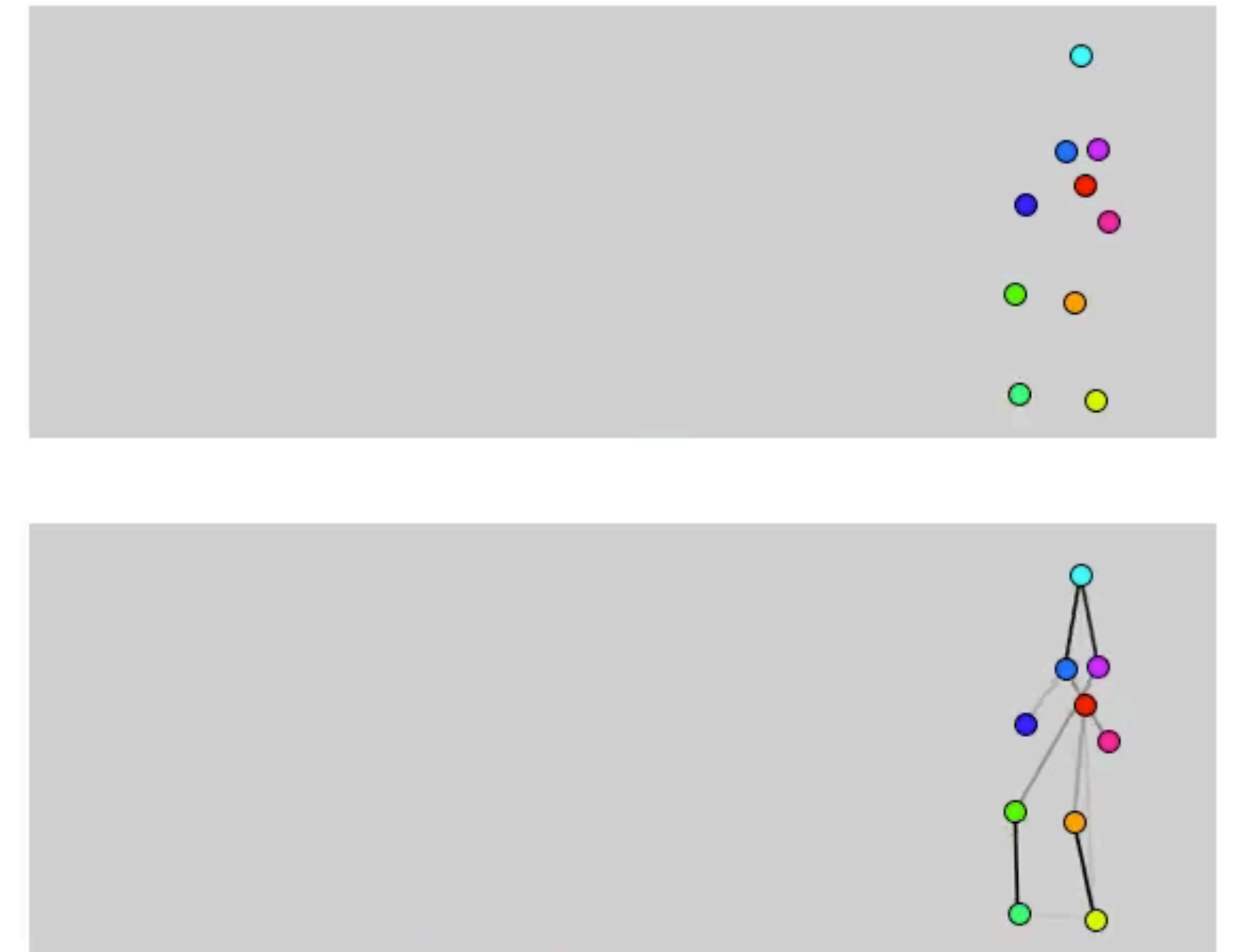
Santoro et al., 2017, NeurIPS

Relation Networks can infer relations in dot motion

Trained on mass-spring systems



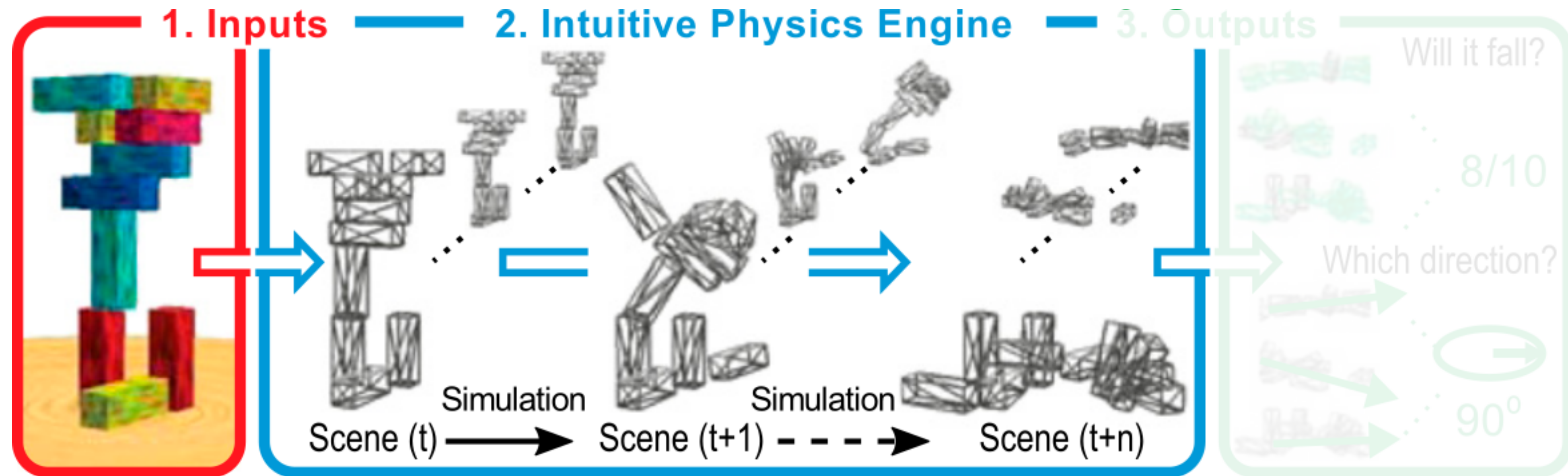
Generalizes to point-light walkers



Santoro et al., 2017, NeurIPS

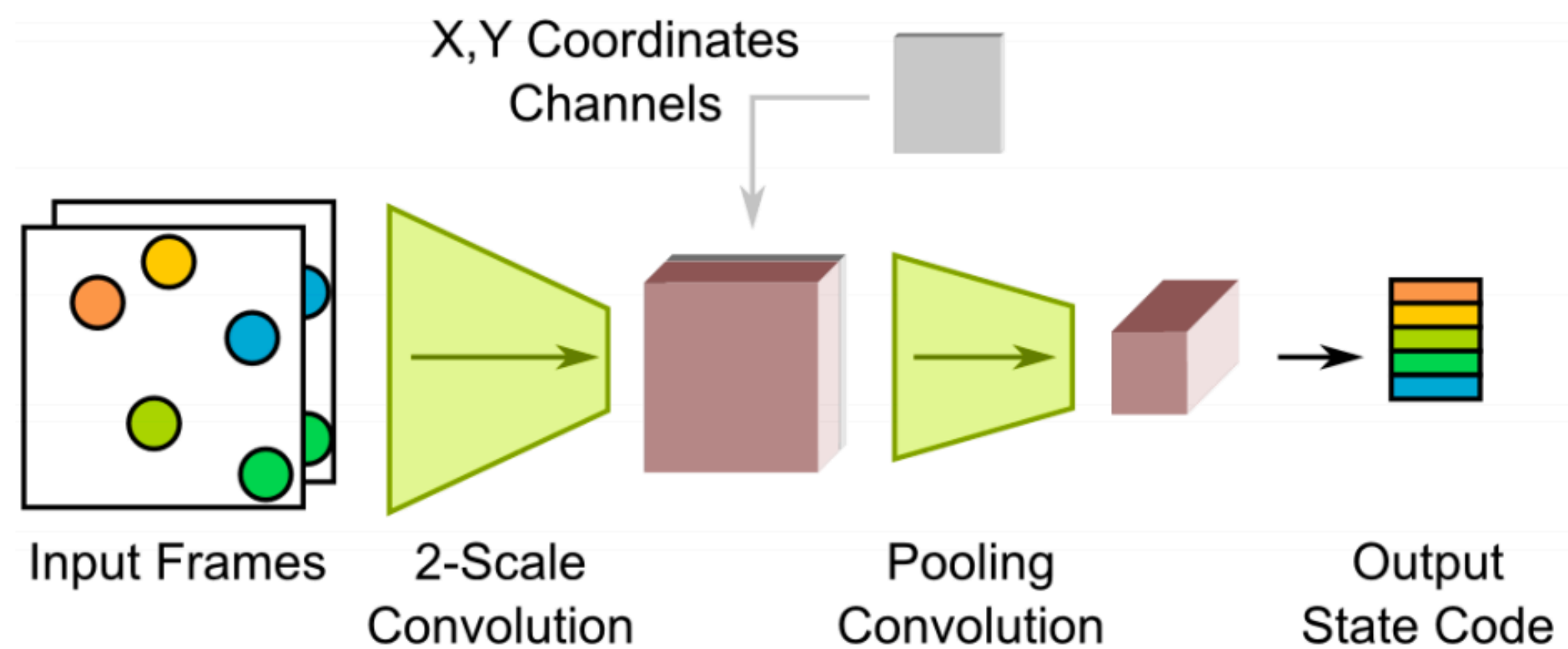
"Visual interaction network"

An interaction network augmented with a learnable perception system

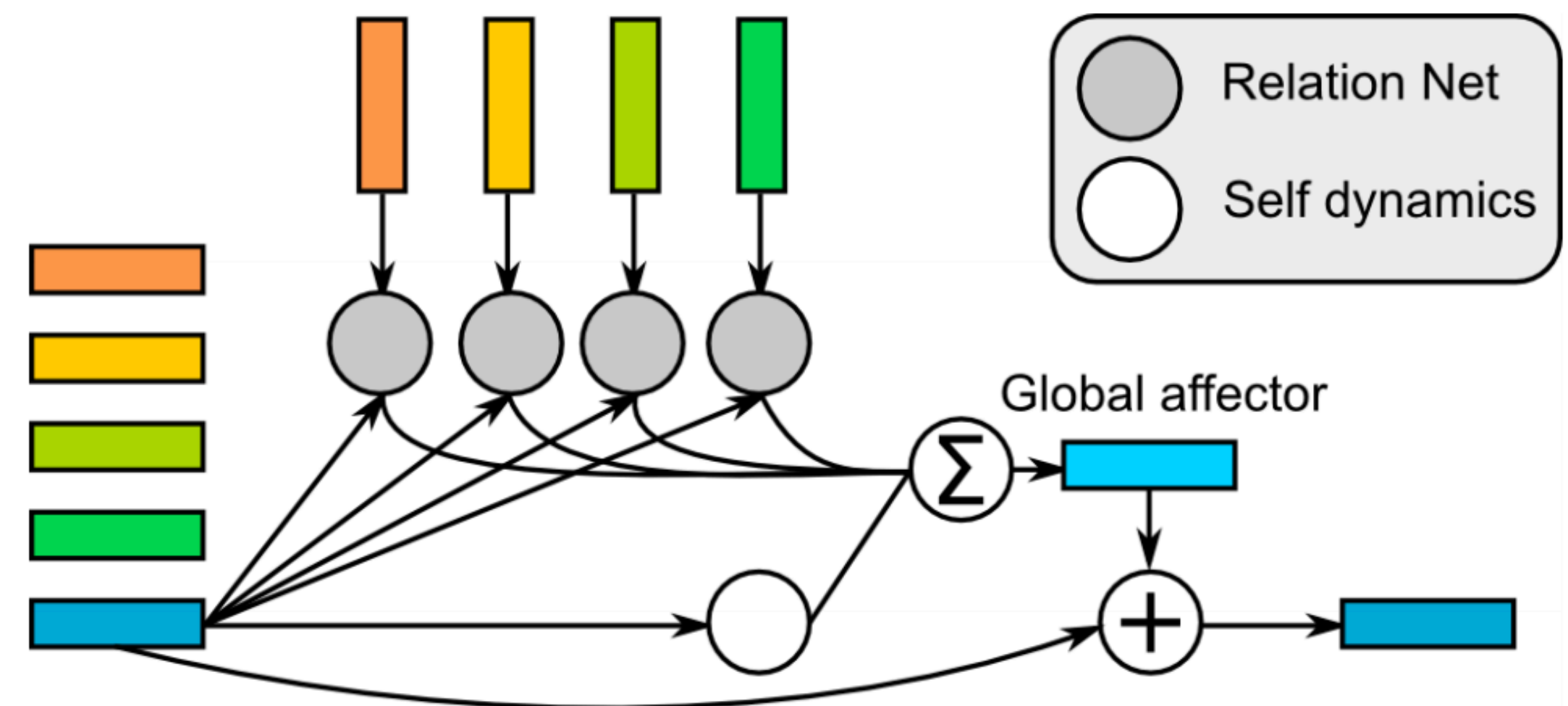


"Visual interaction network"

Multi-frame encoder (conv net-based)



Interaction network



Watters et al., 2017, NeurIPS

"Visual interaction network"

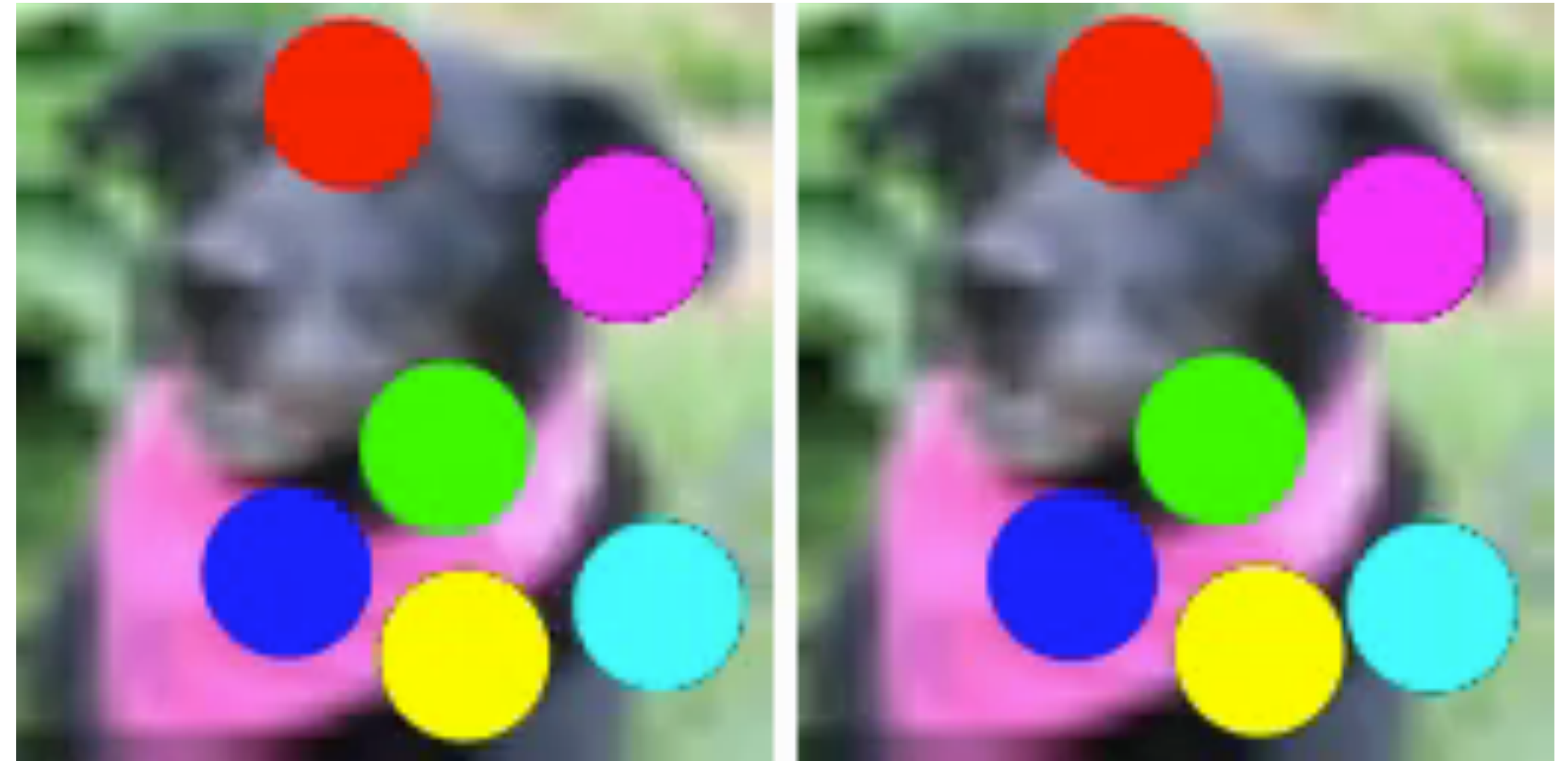
Mass-springs



True

Model

Bouncing balls



True

Model

Can even predict invisible objects, inferred from how they affect visible ones

Watters et al., 2017, NeurIPS

"Visual interaction network"

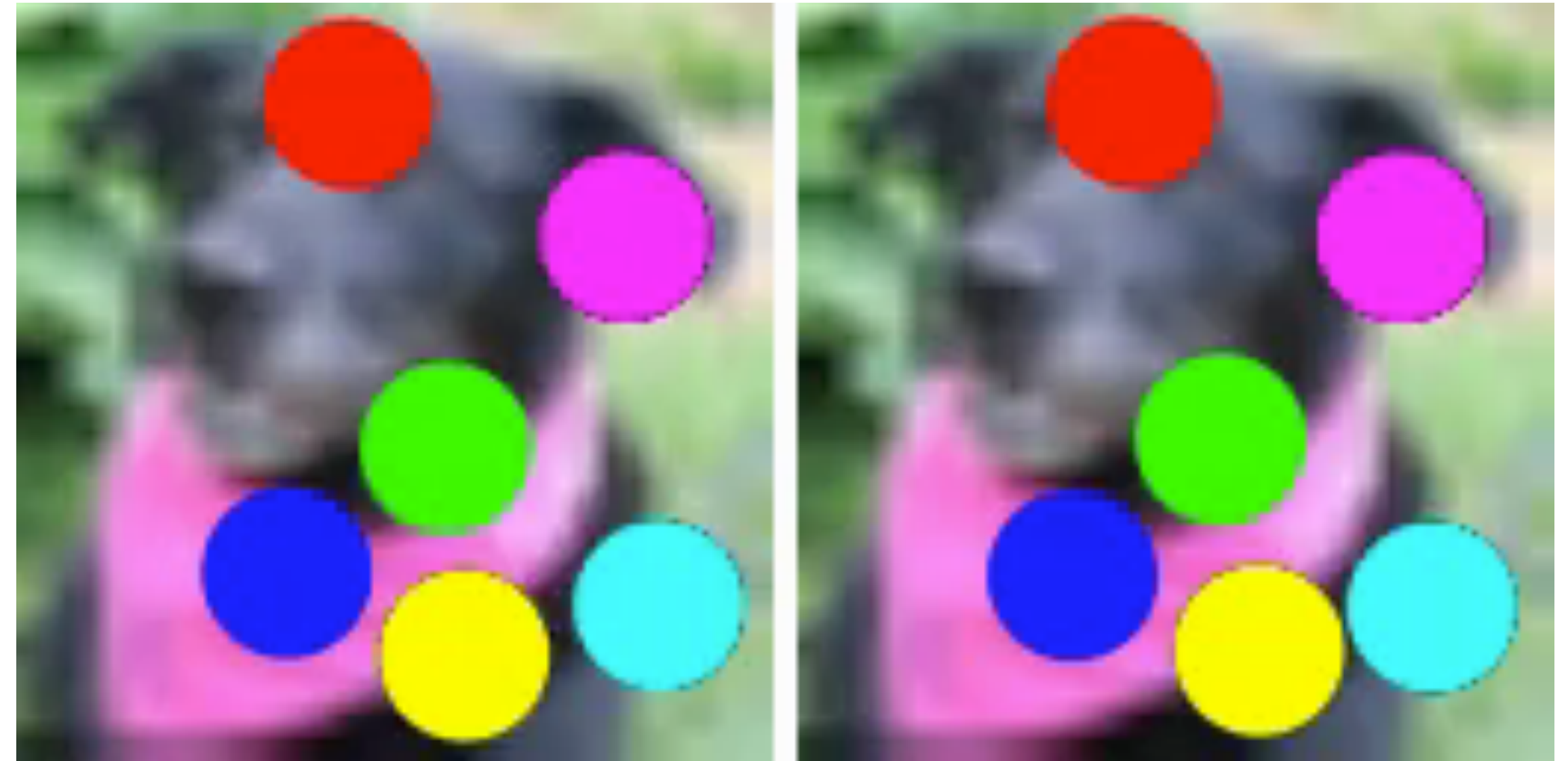
Mass-springs



True

Model

Bouncing balls



True

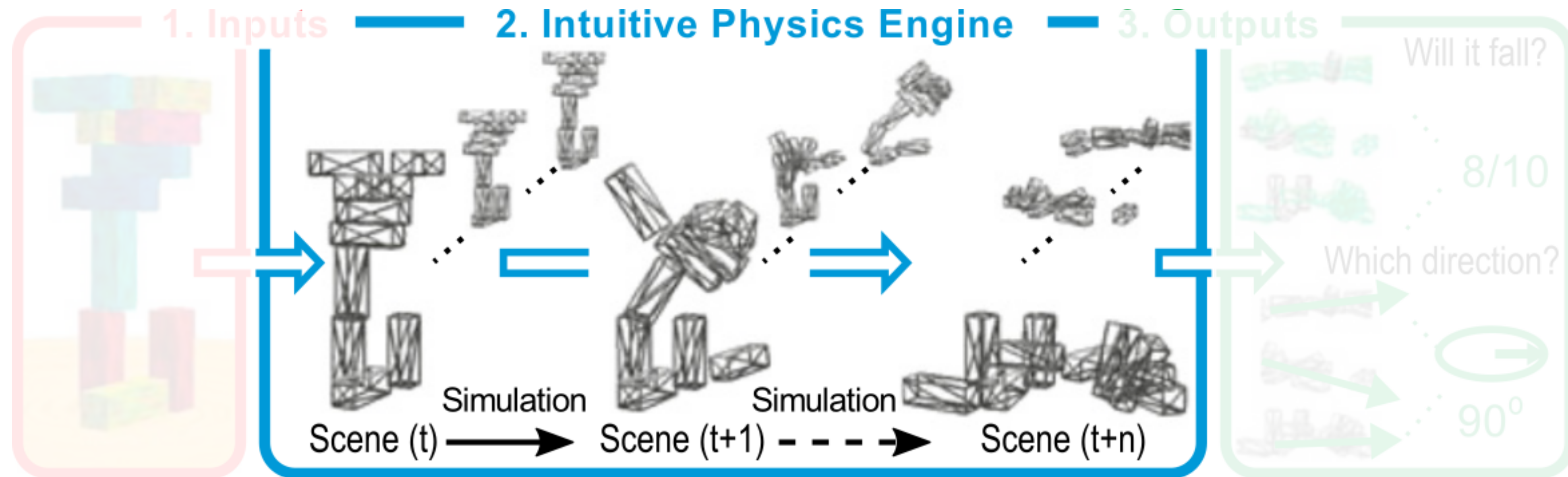
Model

Can even predict invisible objects, inferred from how they affect visible ones

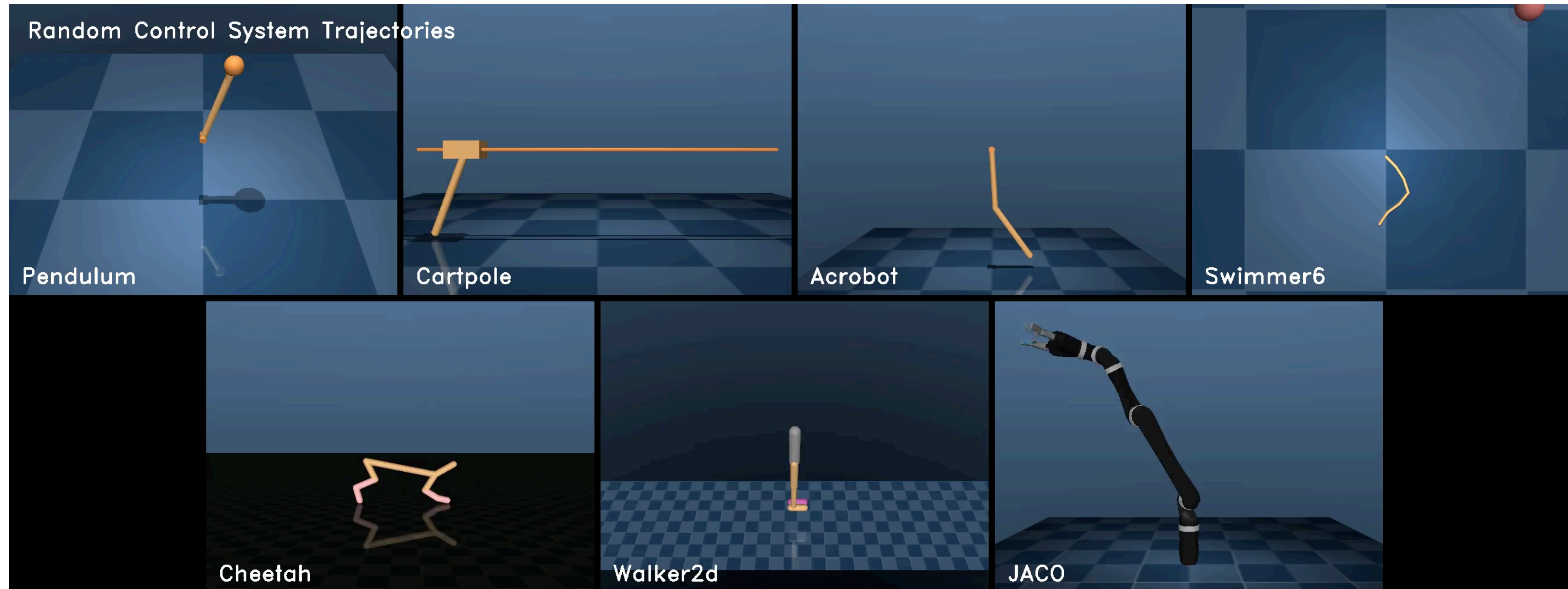
Watters et al., 2017, NeurIPS

Learning to simulate more complex robotic systems

Alvaro Sanchez-Gonzalez, Nicolas Heess, Tobi Springenberg,
Josh Merel, Martin Riedmiller, Raia Hadsell, Peter Battaglia
ICML, 2018



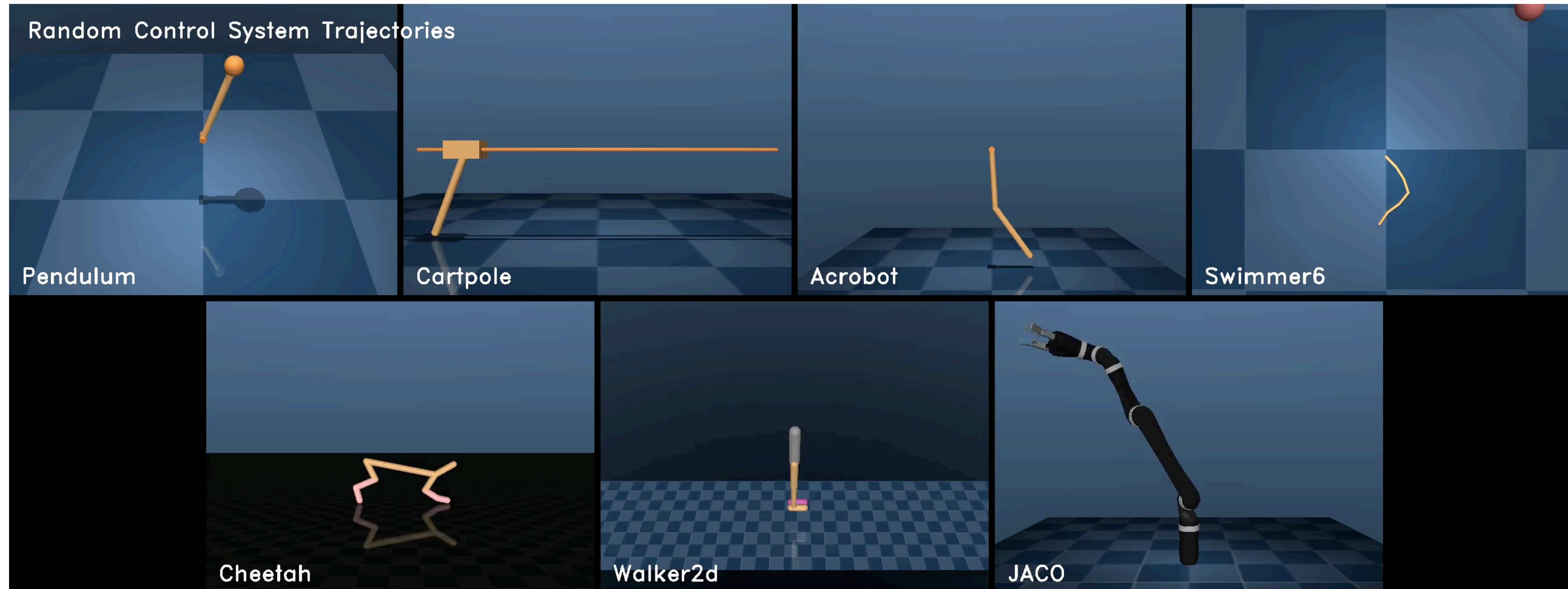
Systems: "DeepMind Control Suite" (Mujoco) & real JACO



JACO Arm

DeepMind Control Suite (Tassa et al., 2018)

Systems: "DeepMind Control Suite" (Mujoco) & real JACO



JACO Arm

DeepMind Control Suite (Tassa et al., 2018)

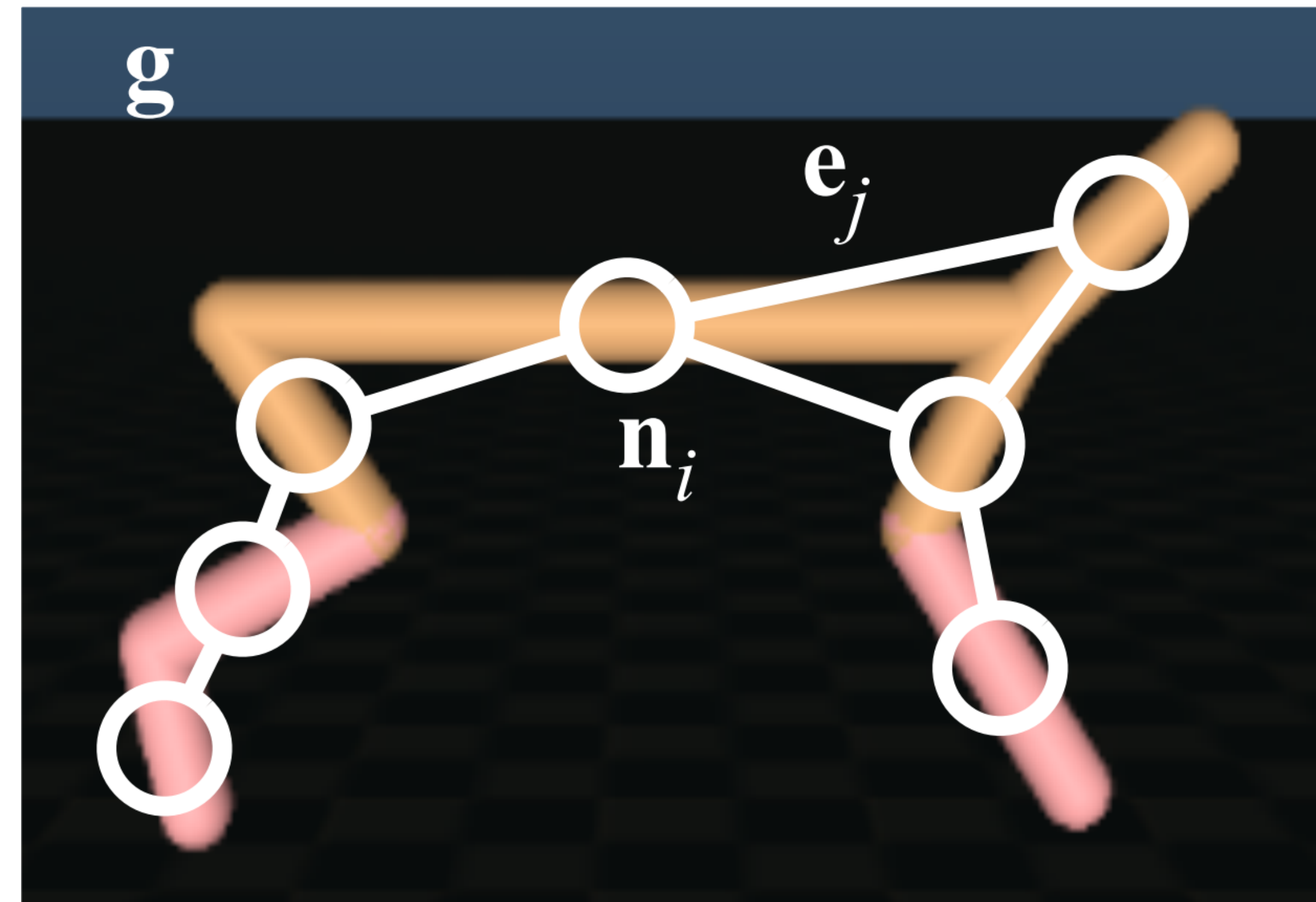
Kinematic tree of the actuated system as a graph

Representing physical system as a graph:

- Bodies $\rightarrow$ Nodes
- Joints $\rightarrow$ Edges
- Global properties

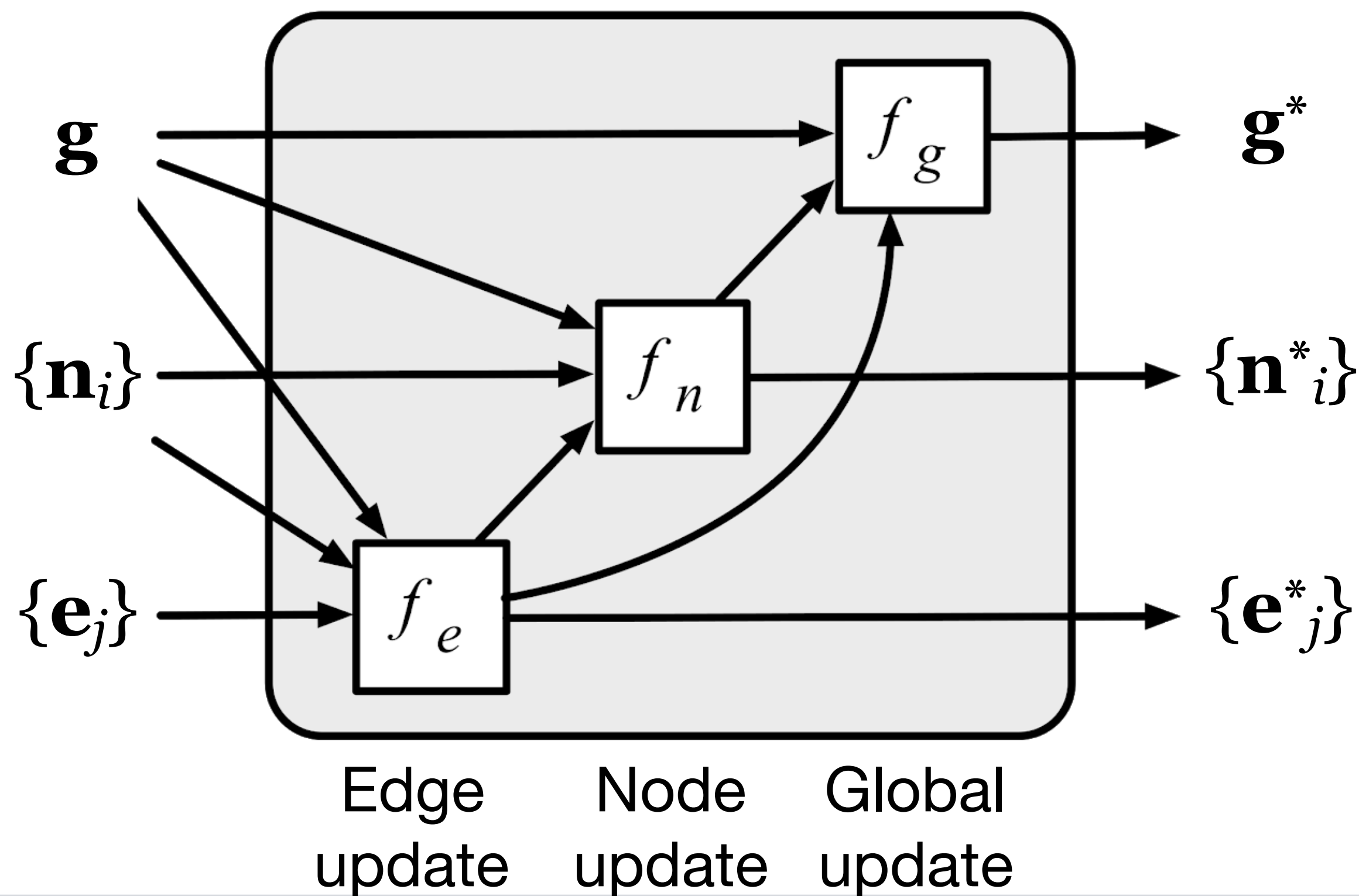
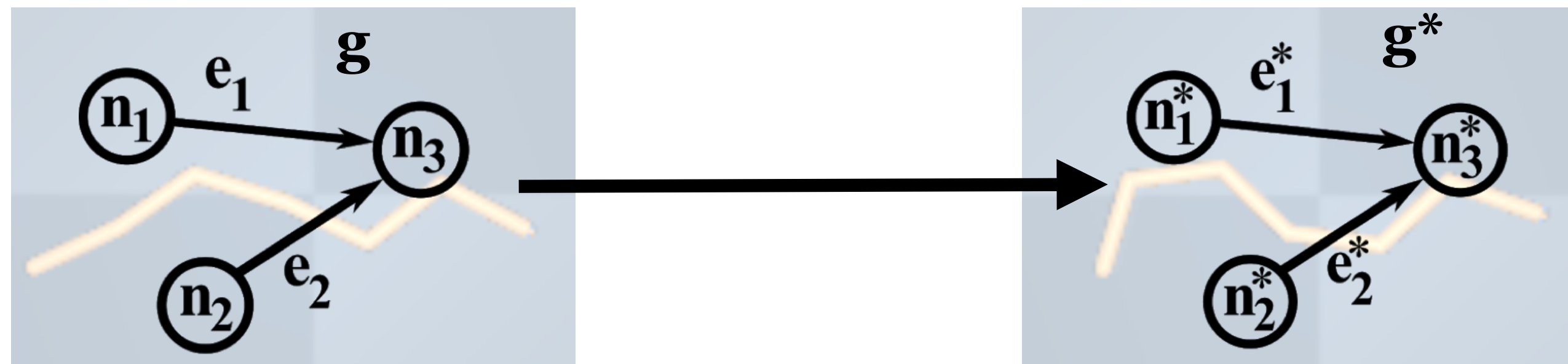
Similar representation to:

- Interaction Networks (Battaglia et al. 2016)
- NerveNet (Wang et al. 2018) (graph-structured policy, rather than model)

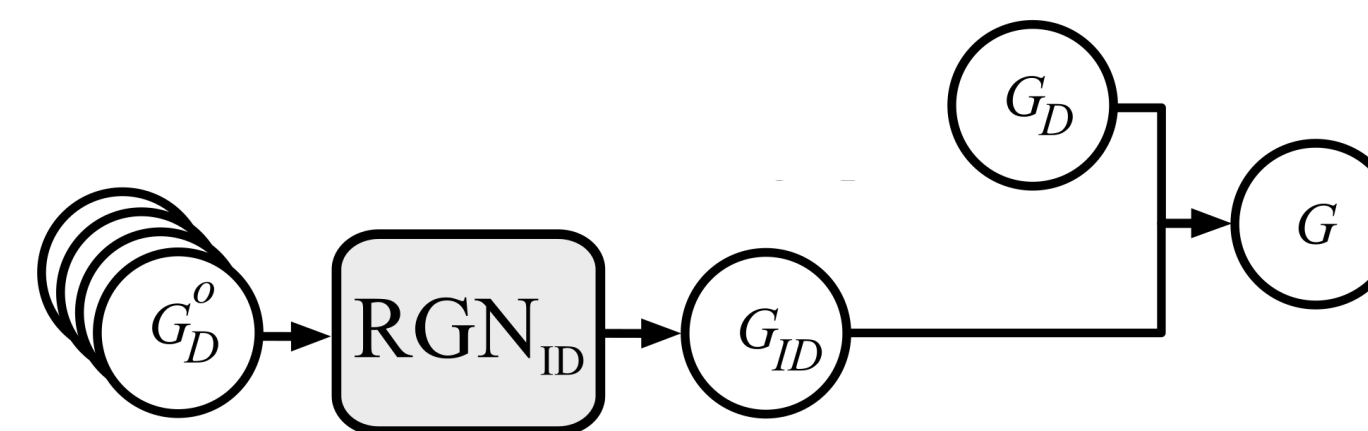
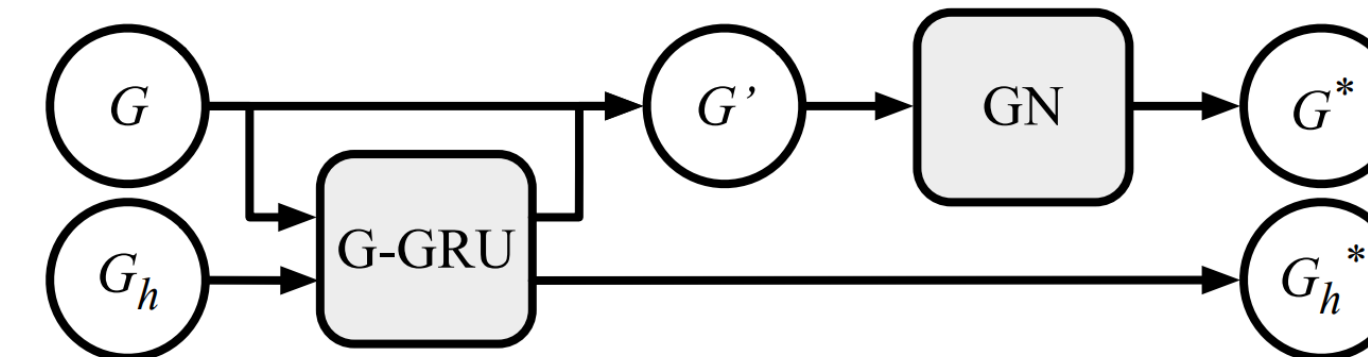
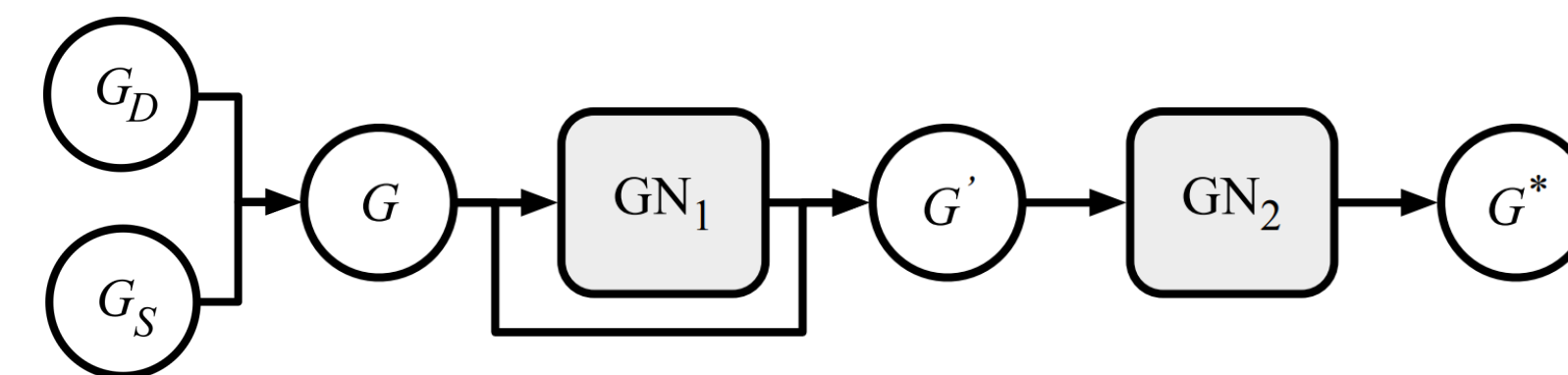


Graph Network (GN)

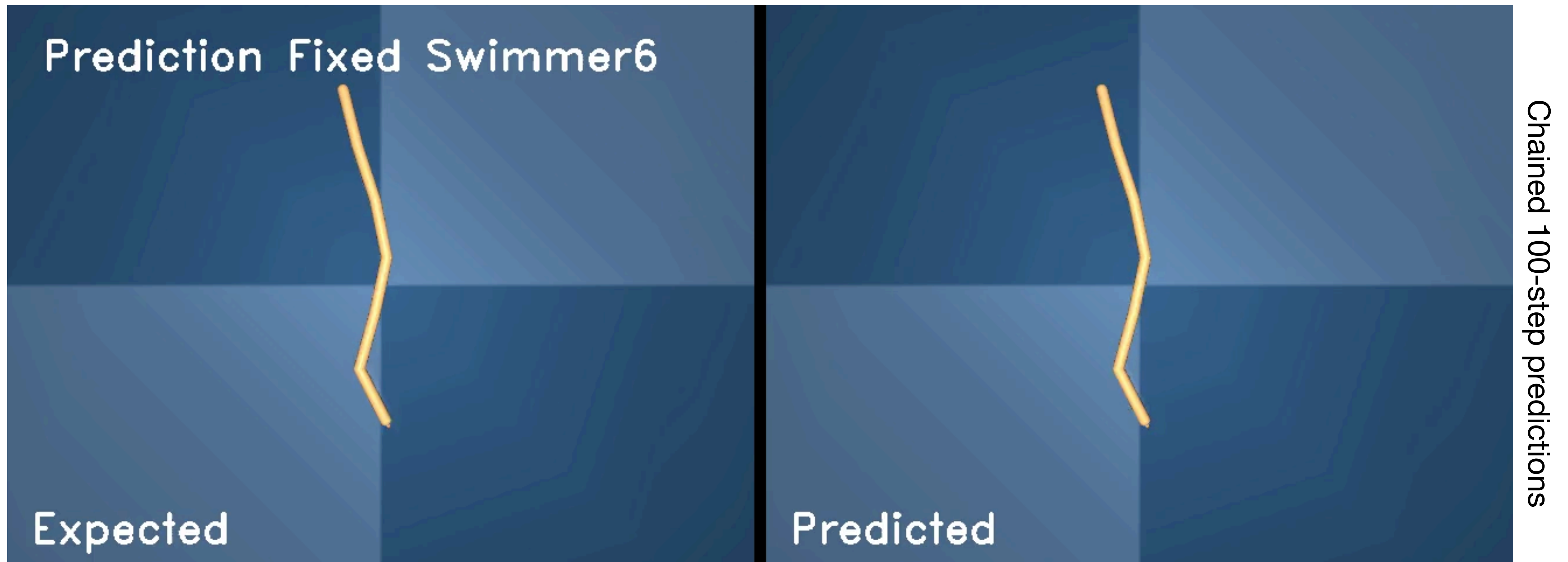
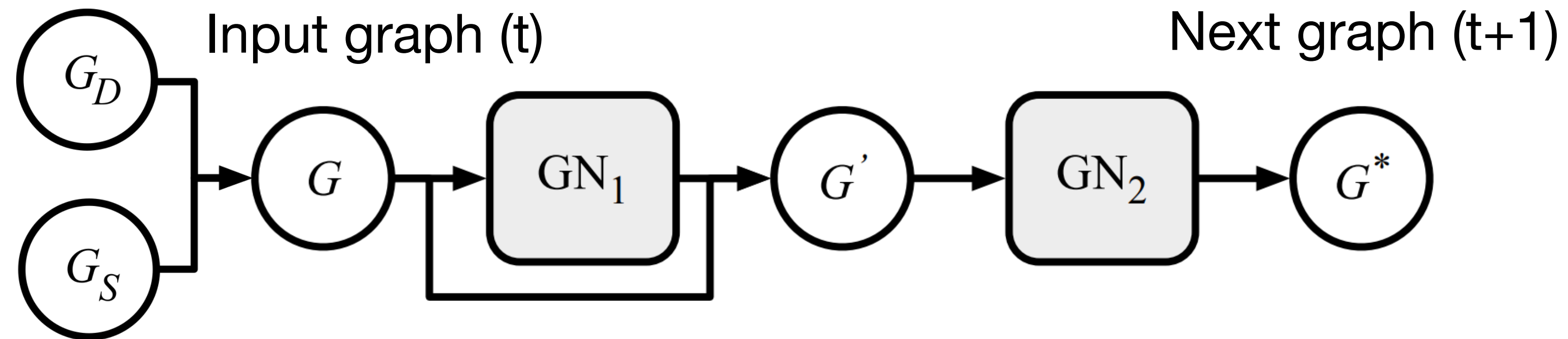
Battaglia et al., 2018



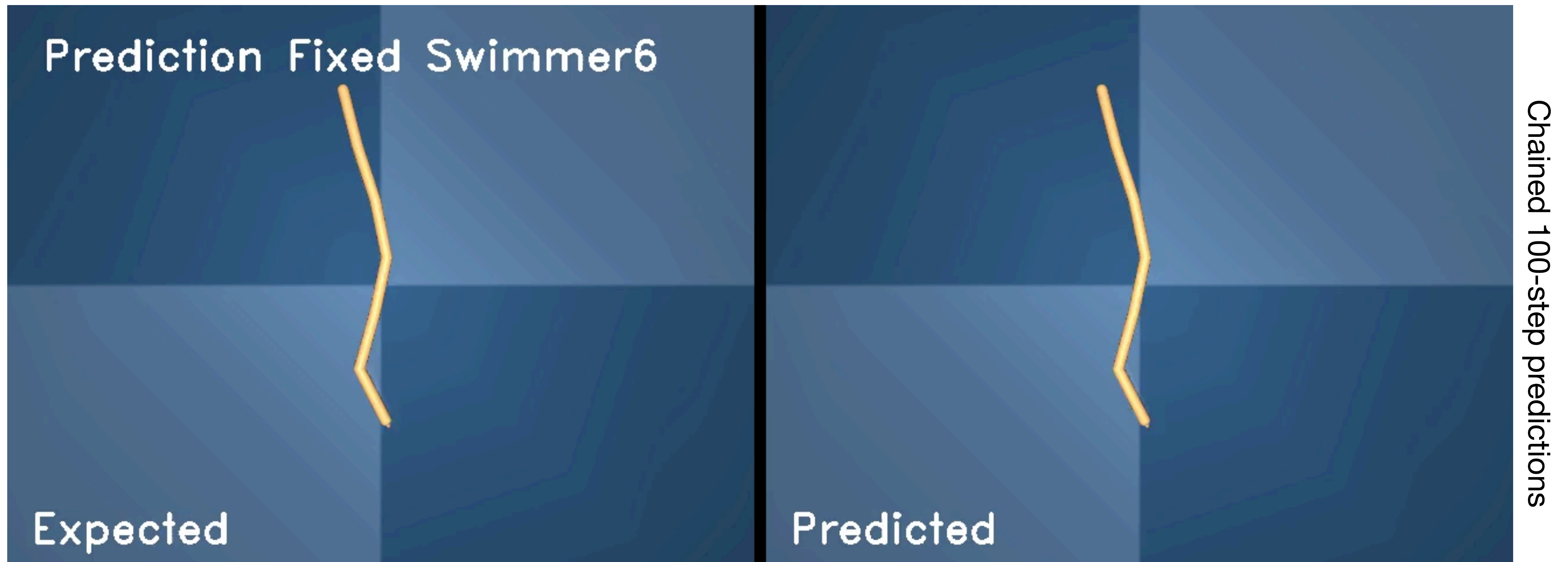
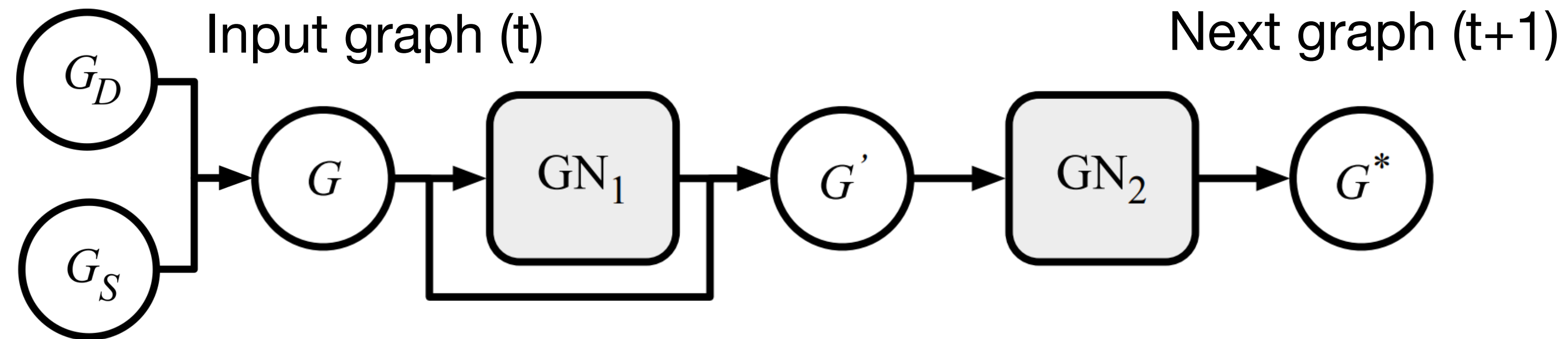
Graph-to-graph, modular block design



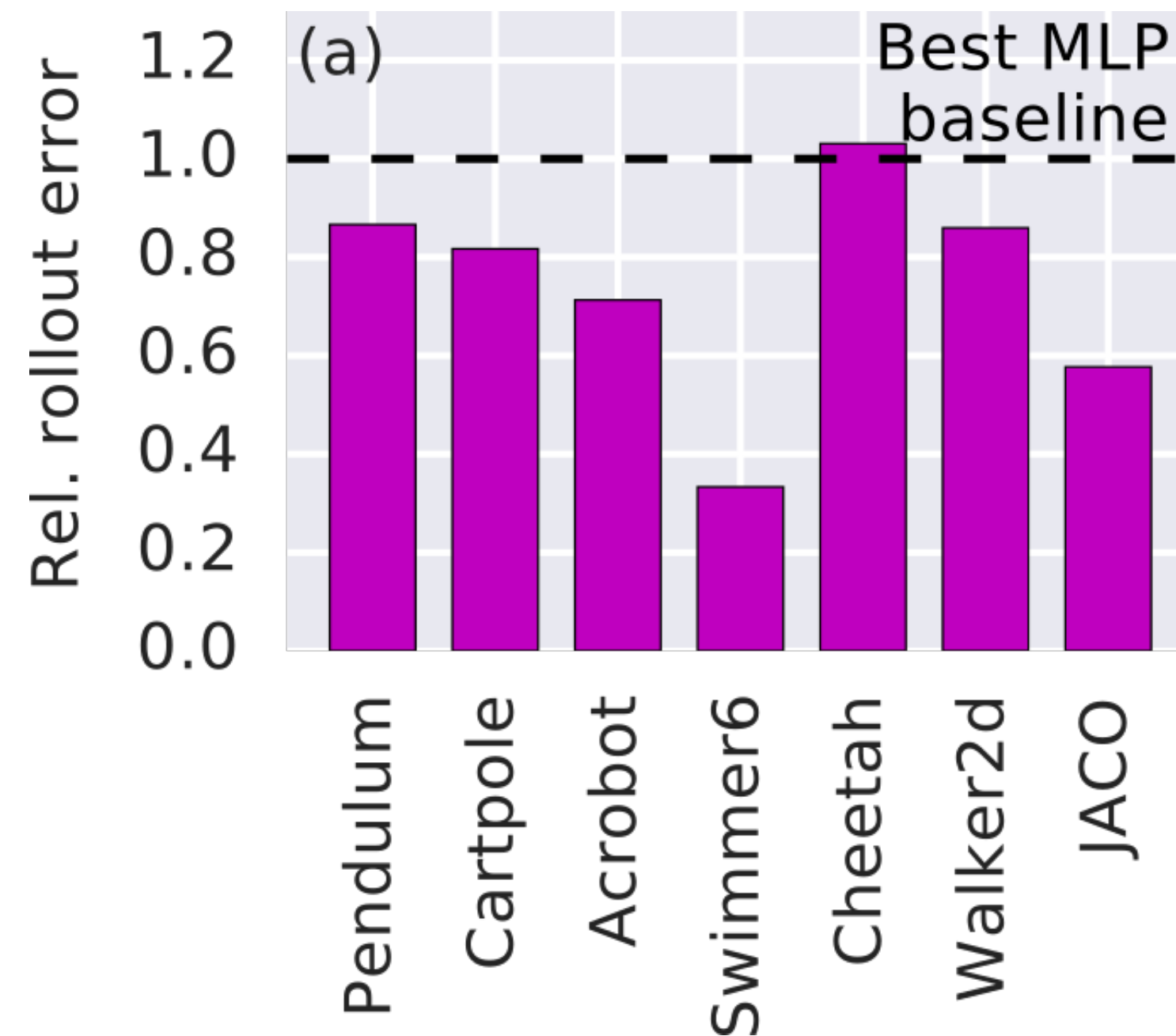
Forward model: supervised, 1-step training w/ random control inputs



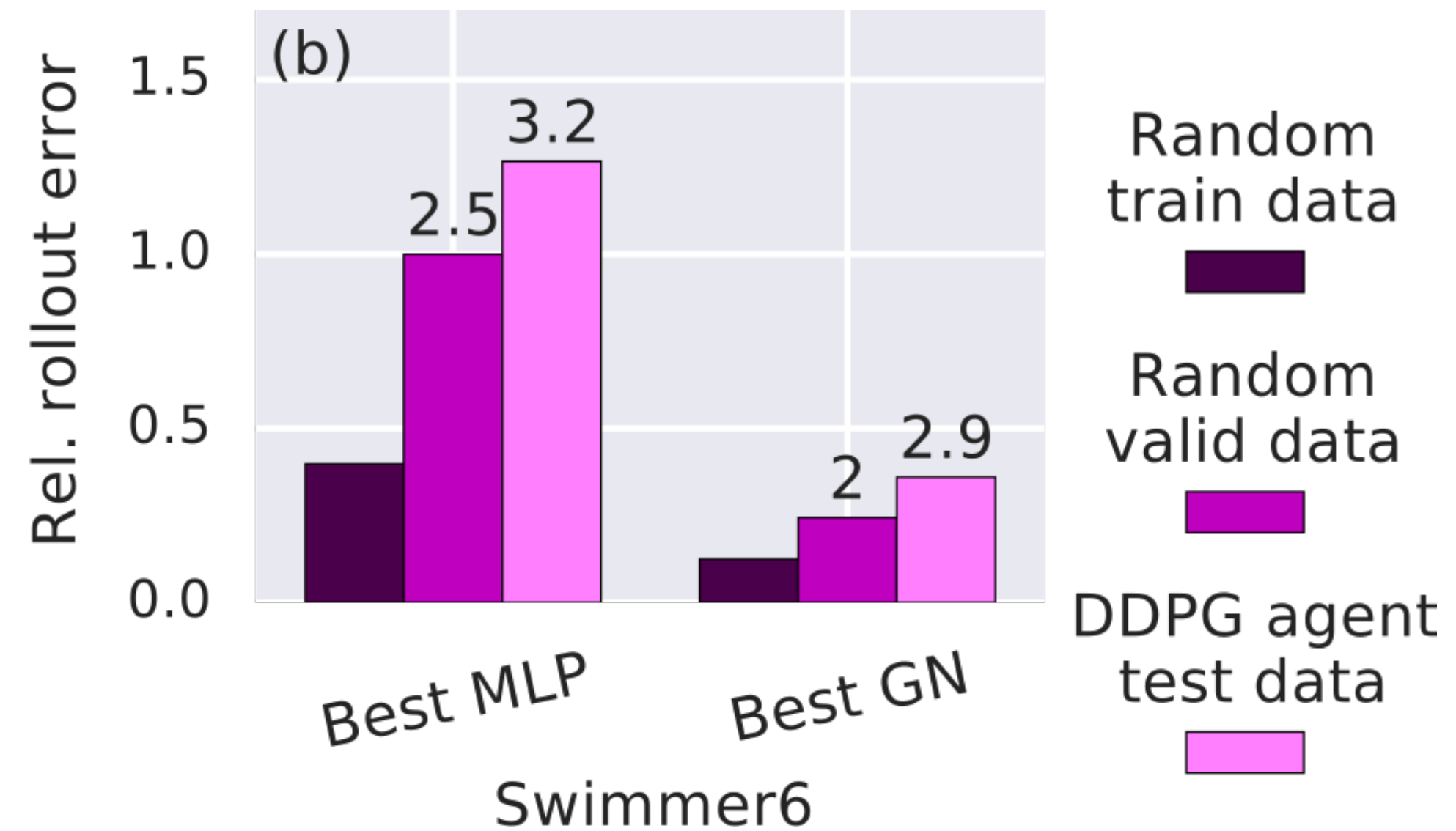
Forward model: supervised, 1-step training w/ random control inputs



Results: Graph Net (GN) vs MLP forward models



More repeated structure:
Better performance over MLP

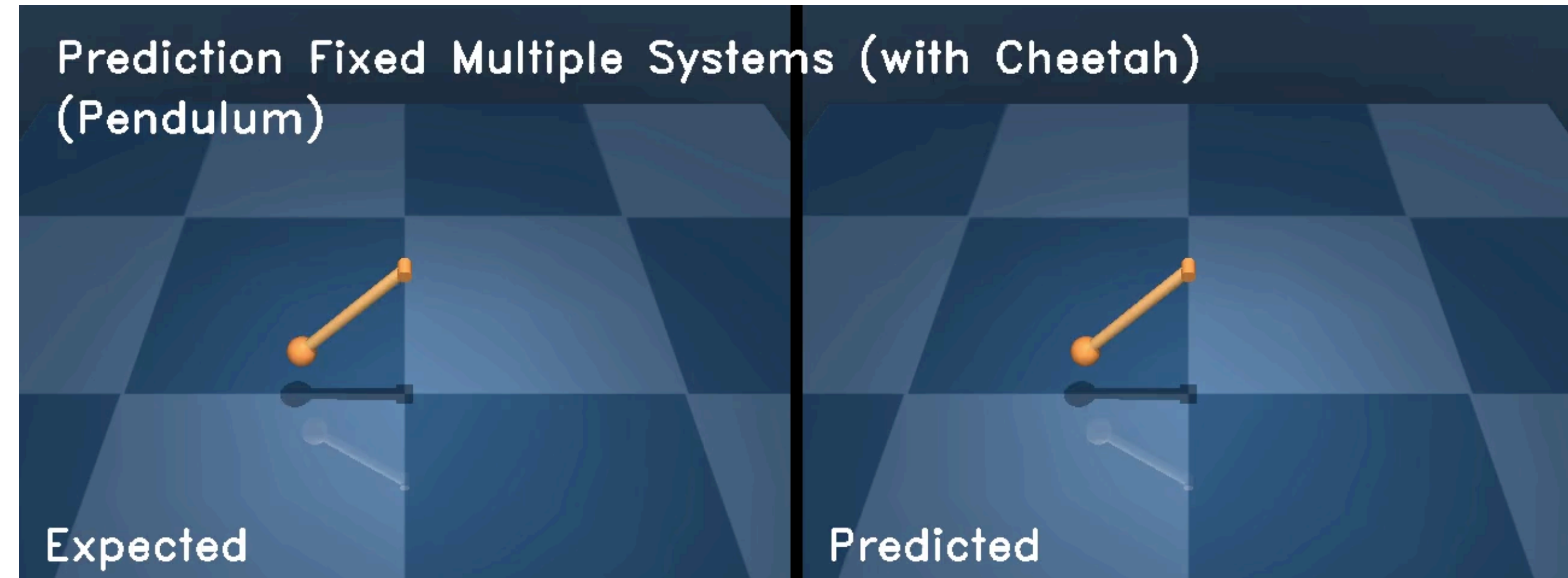


Better test generalization,
within and outside of the
training distribution

GN forward model: Multiple systems & zero-shot generalization

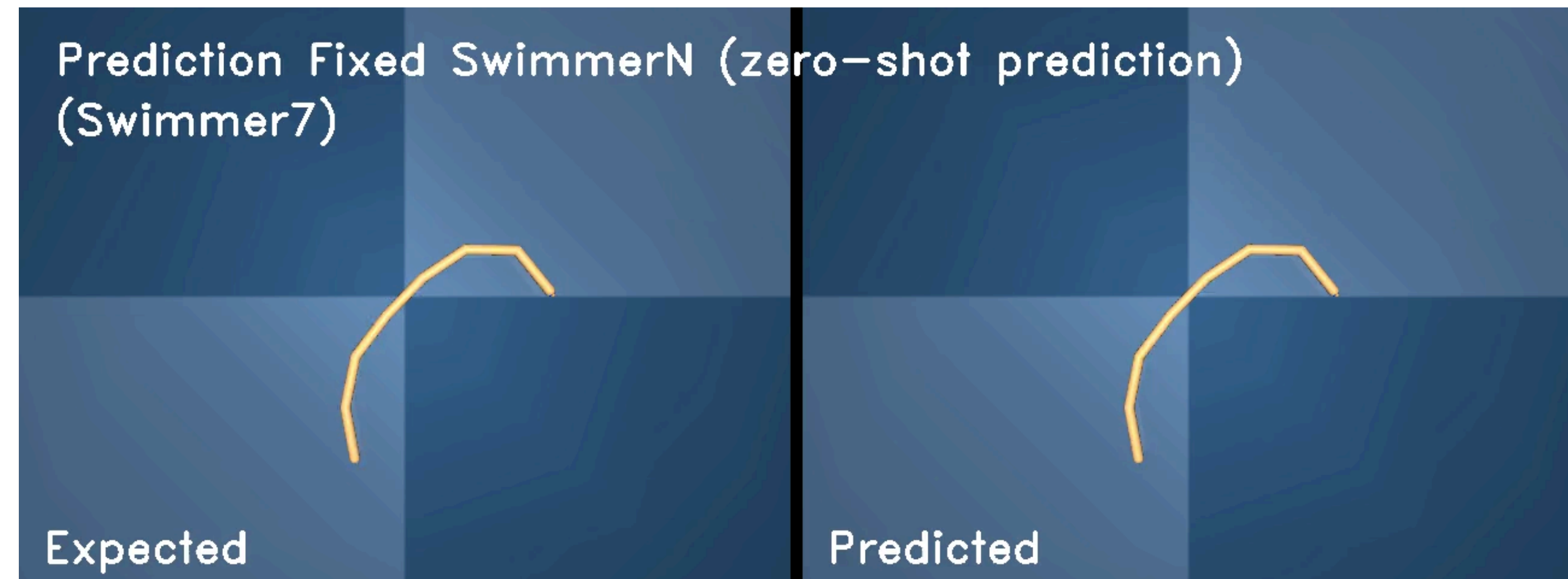
Single model trained:

- Pendulum, Cartpole, Acrobot, Swimmer6 & Cheetah



Zero-shot generalization: Swimmer

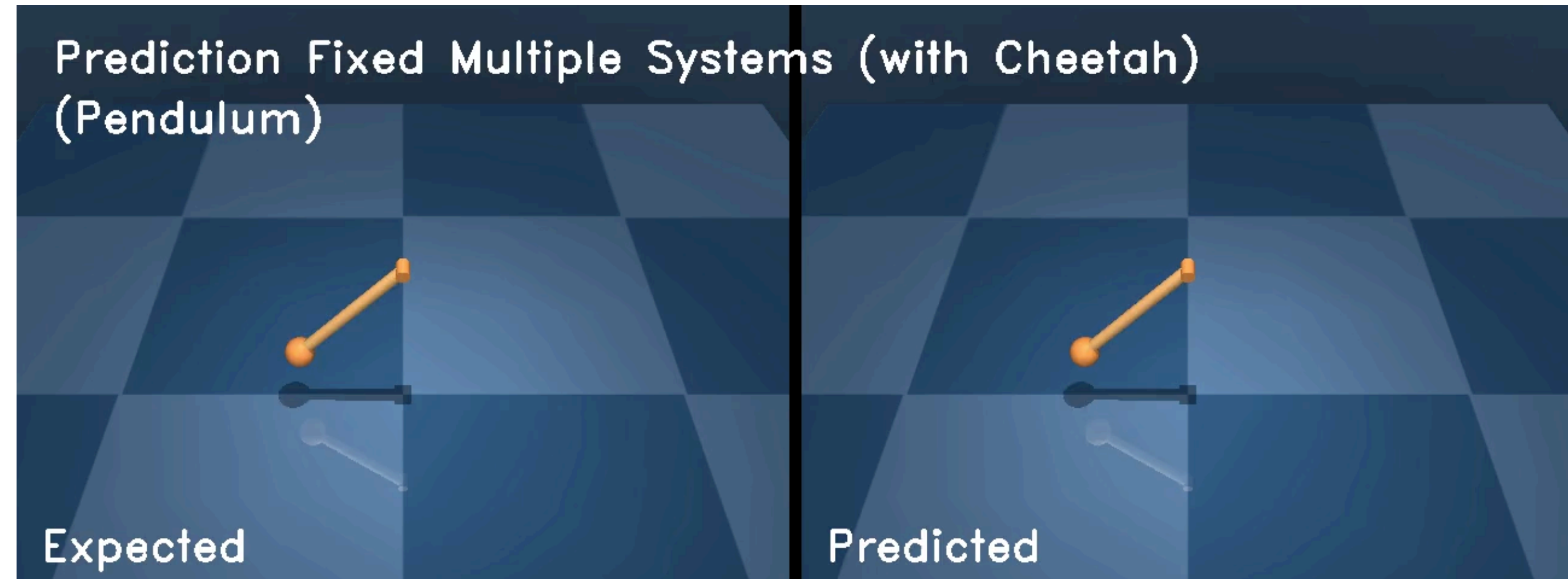
- # training links: {3, 4, 5, 6, -, 8, 9, -, -, ...}
- # testing links: {-, -, -, -, 7, -, -, 10-14}



GN forward model: Multiple systems & zero-shot generalization

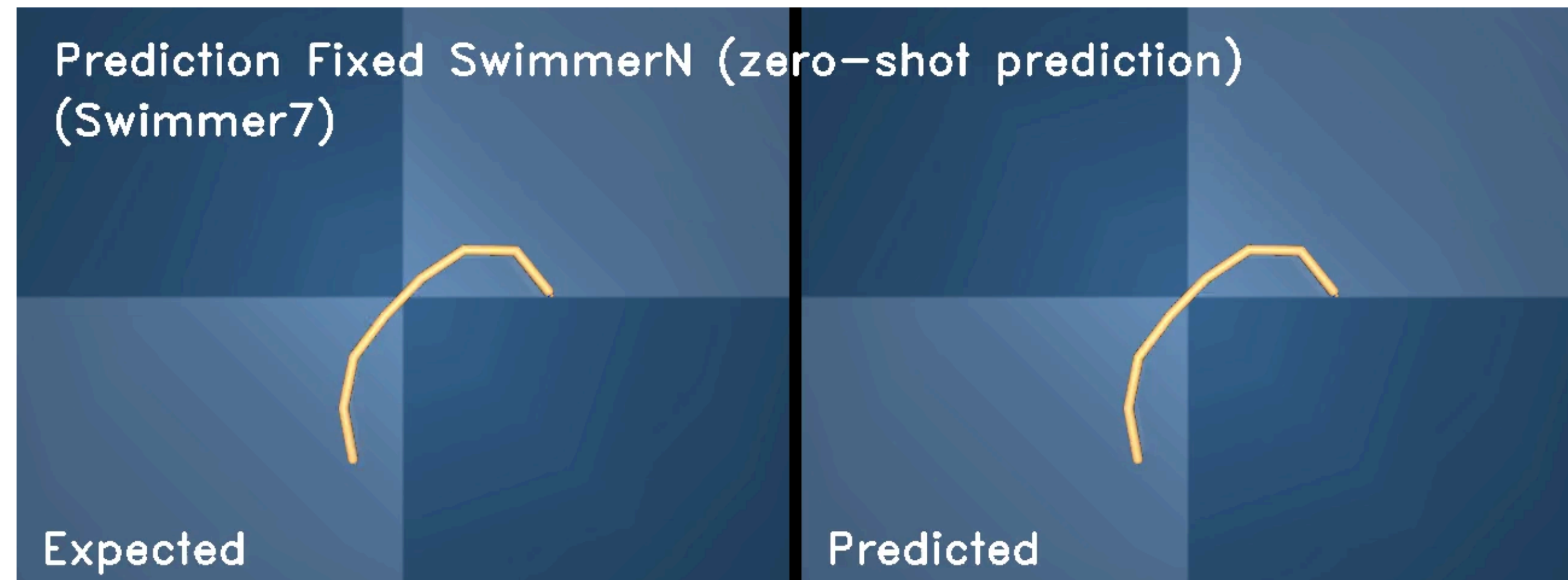
Single model trained:

- Pendulum, Cartpole, Acrobot, Swimmer6 & Cheetah

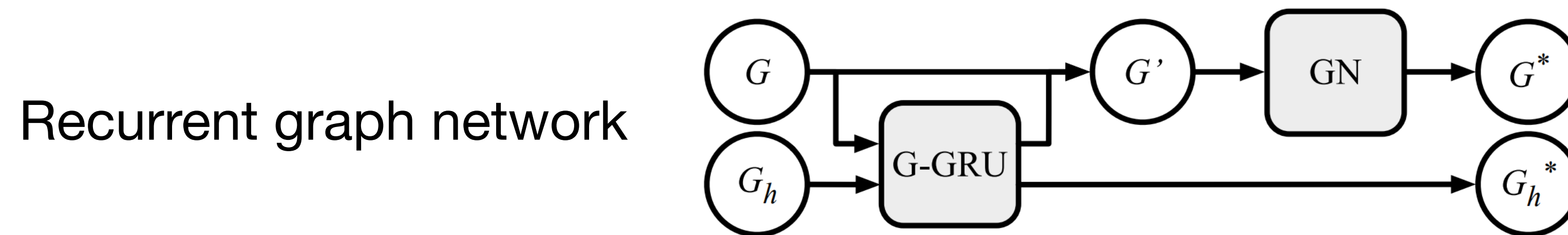


Zero-shot generalization: Swimmer

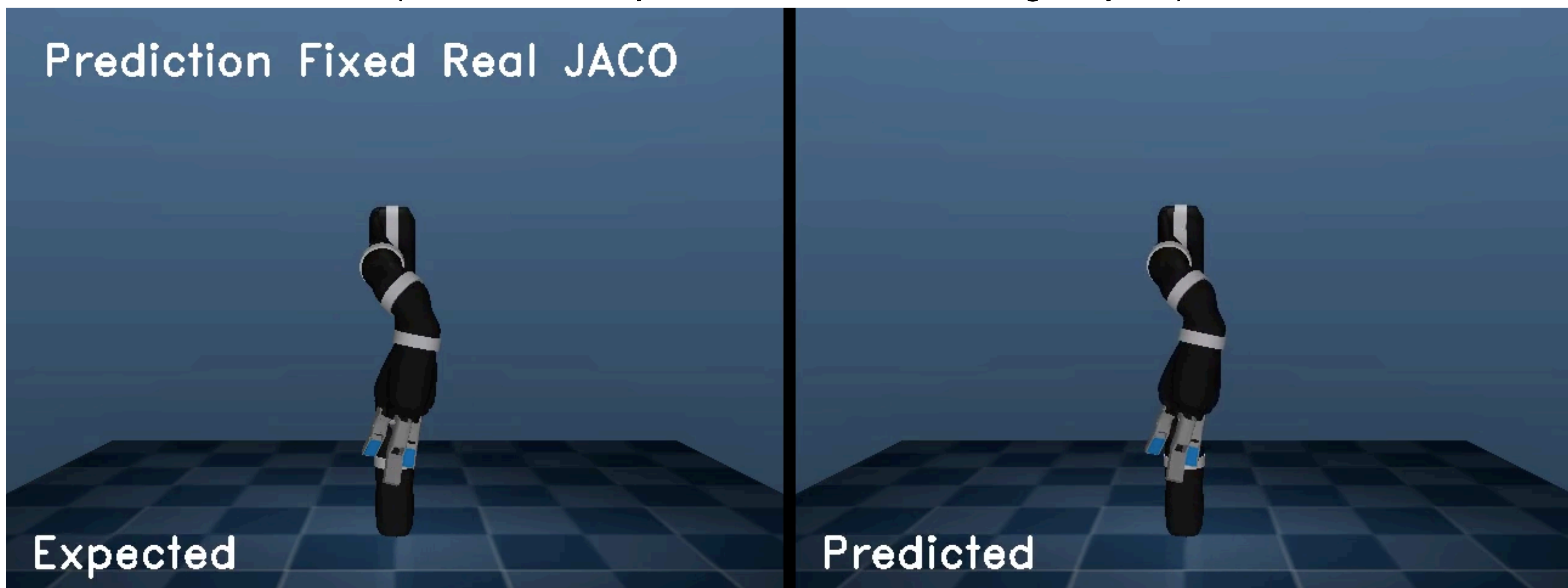
- # training links: {3, 4, 5, 6, -, 8, 9, -, -, ...}
- # testing links: {-, -, -, -, 7, -, -, 10-14}



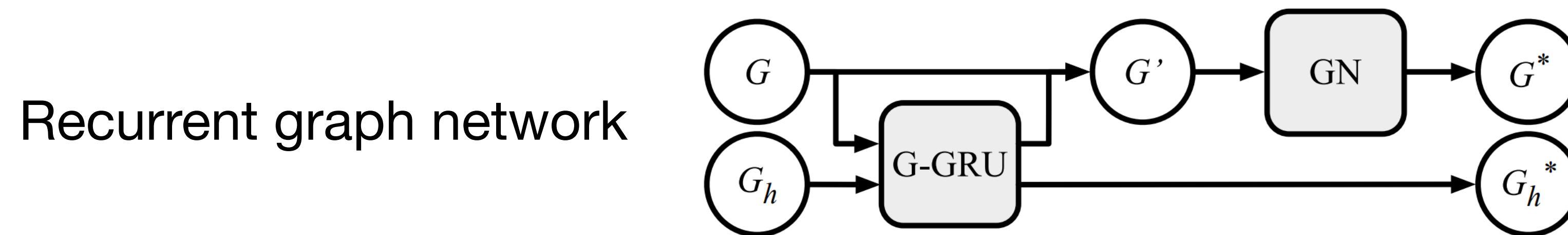
GN forward model: Real JACO data



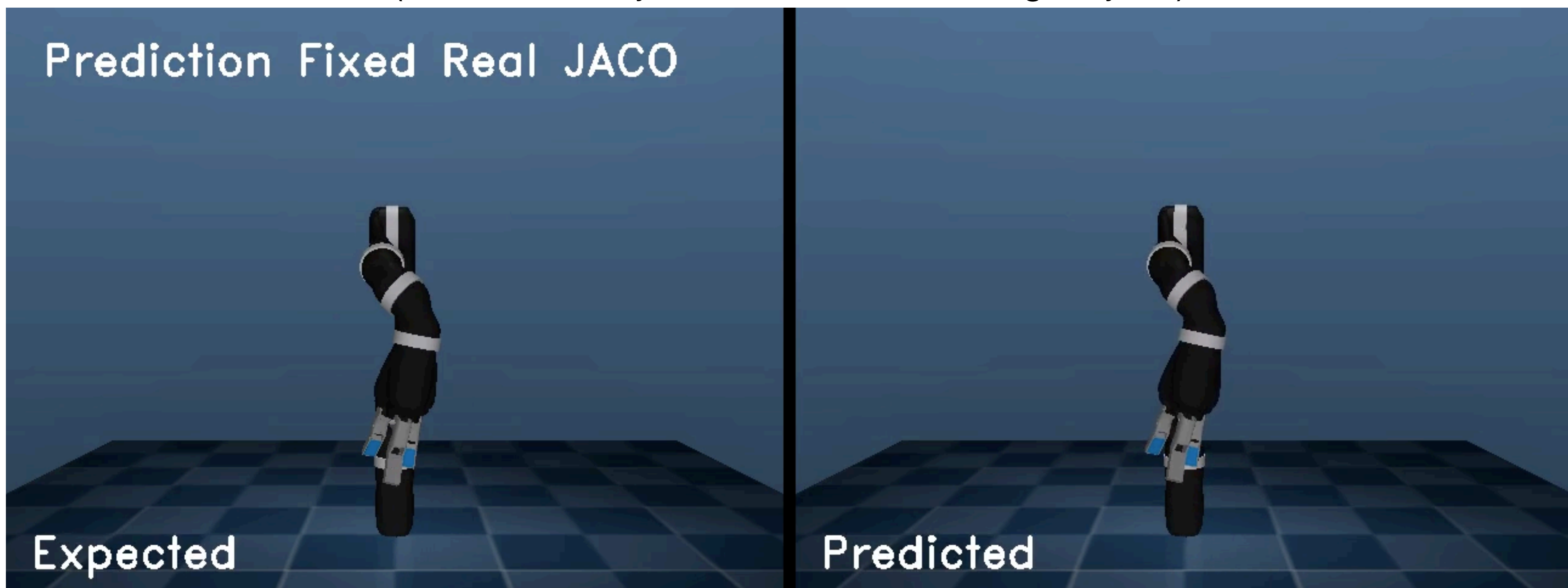
(Real JACO trajectories, rendered using Mujoco)



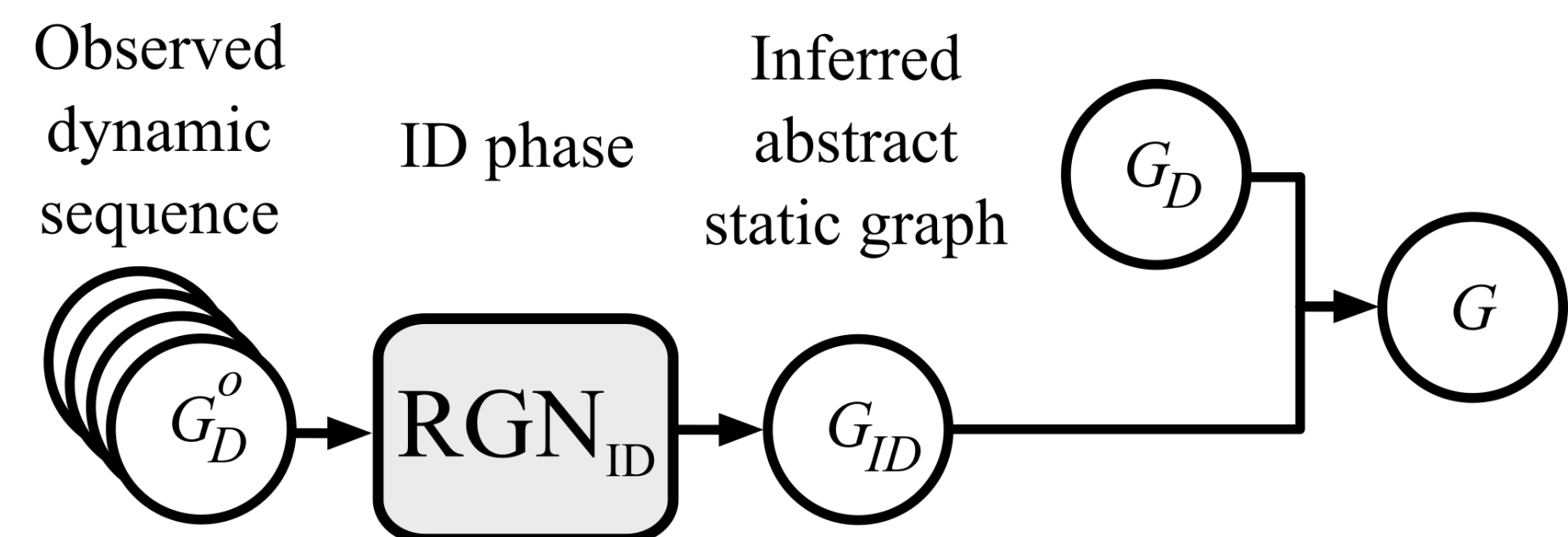
GN forward model: Real JACO data



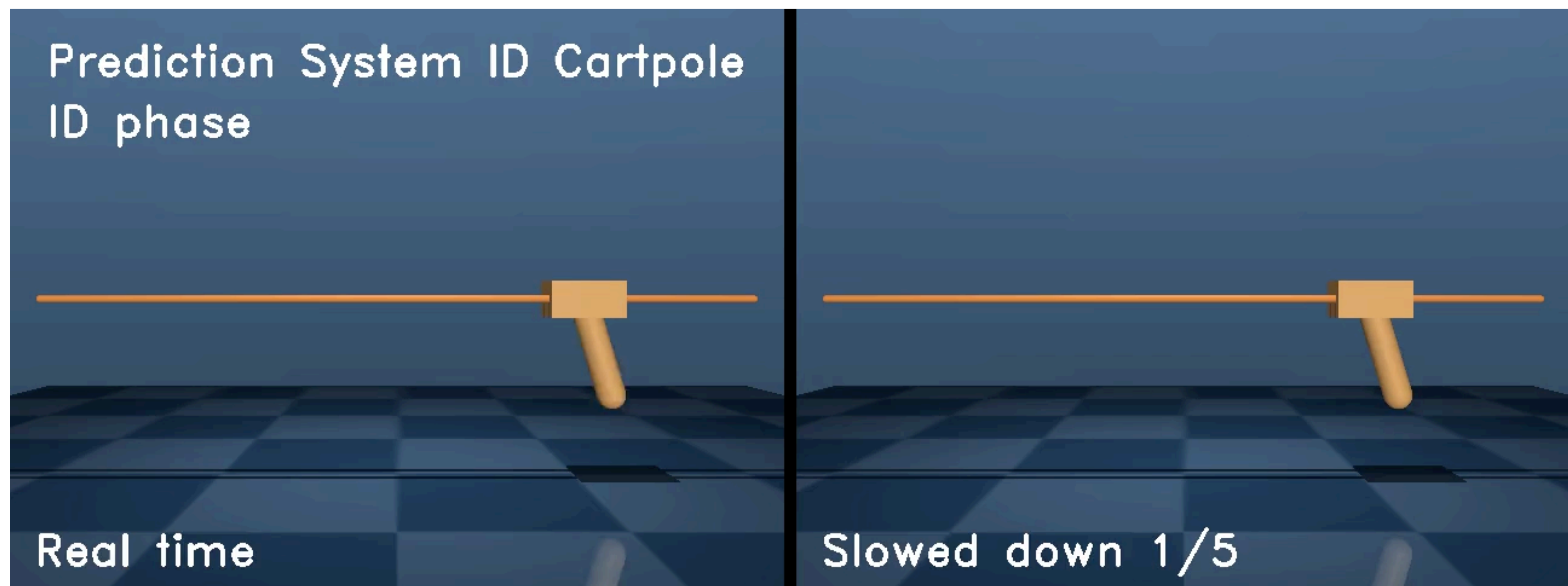
(Real JACO trajectories, rendered using Mujoco)



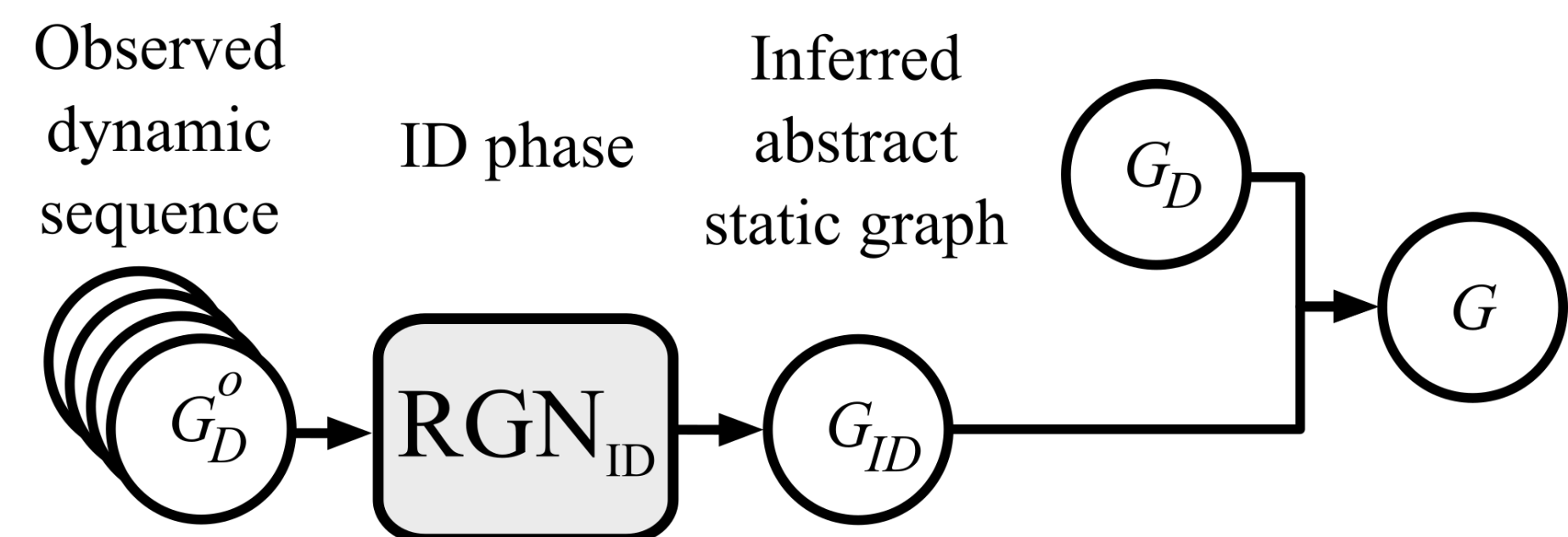
System identification: GN-based inference, under diagnostic control inputs



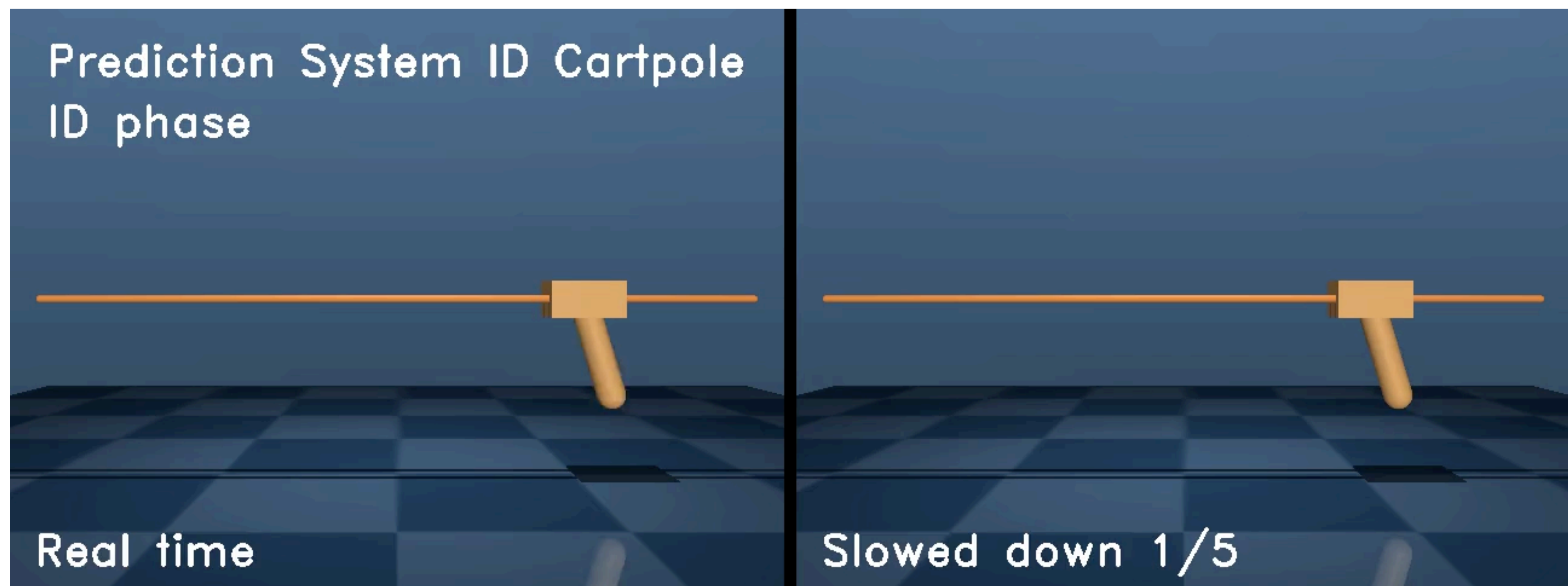
Unobserved system parameters (e.g. mass, length) are implicitly inferred



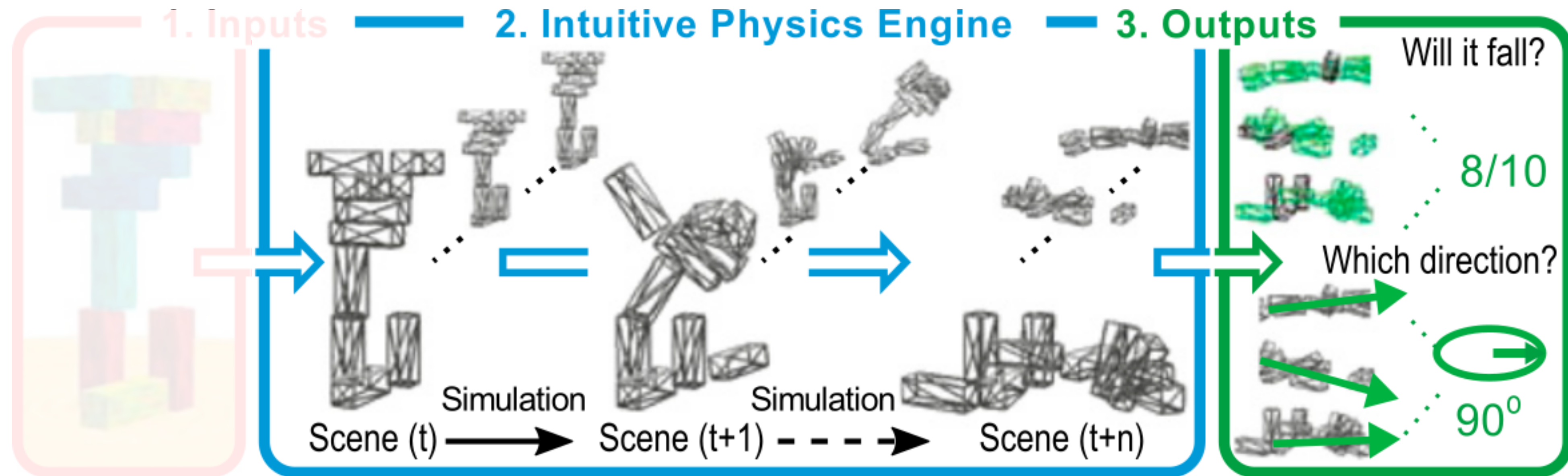
System identification: GN-based inference, under diagnostic control inputs



Unobserved system parameters (e.g. mass, length) are implicitly inferred

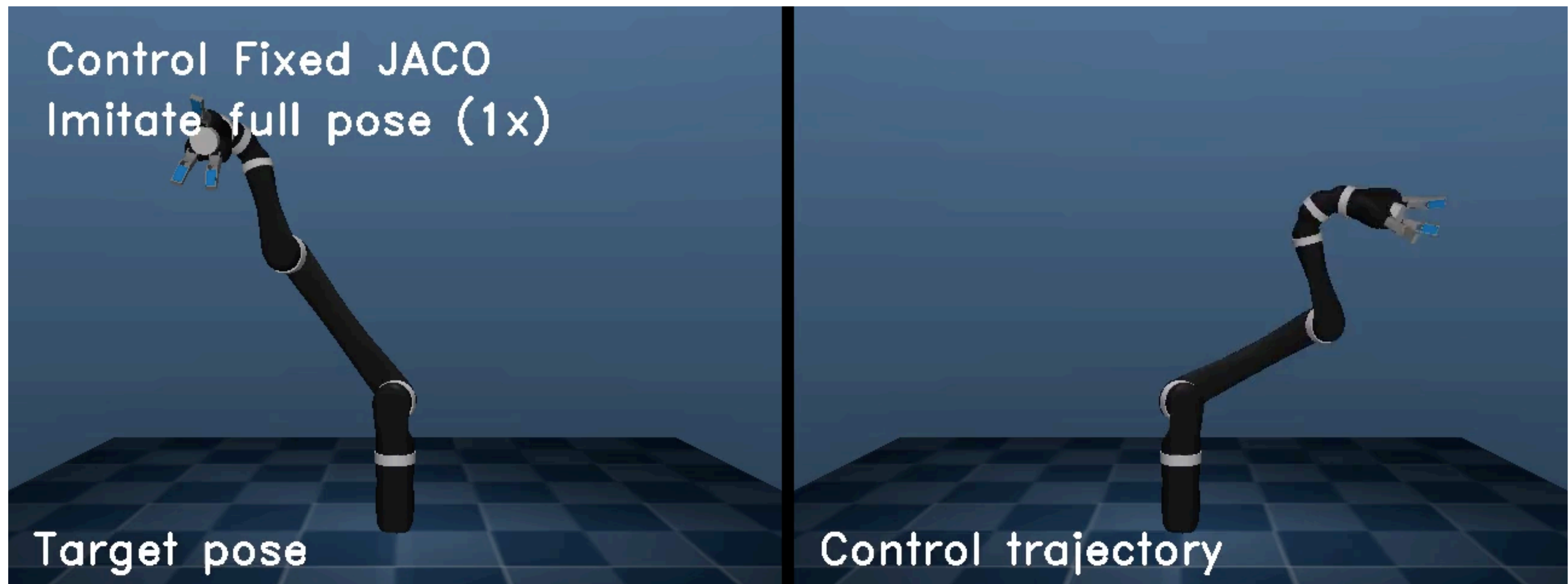


Using learned models for control



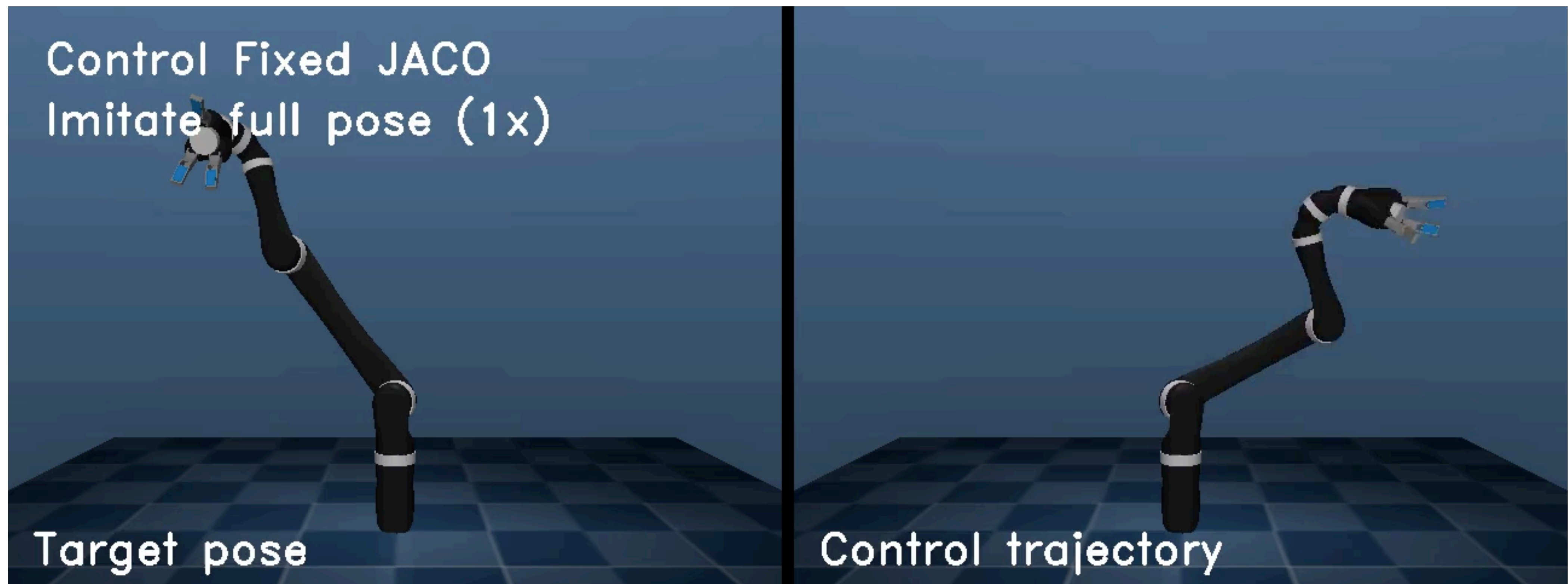
Control: Model-based planning

Trajectory optimization: the GN-based forward model is differentiable, so we can backpropagate through it, and find a sequence of actions that maximize reward

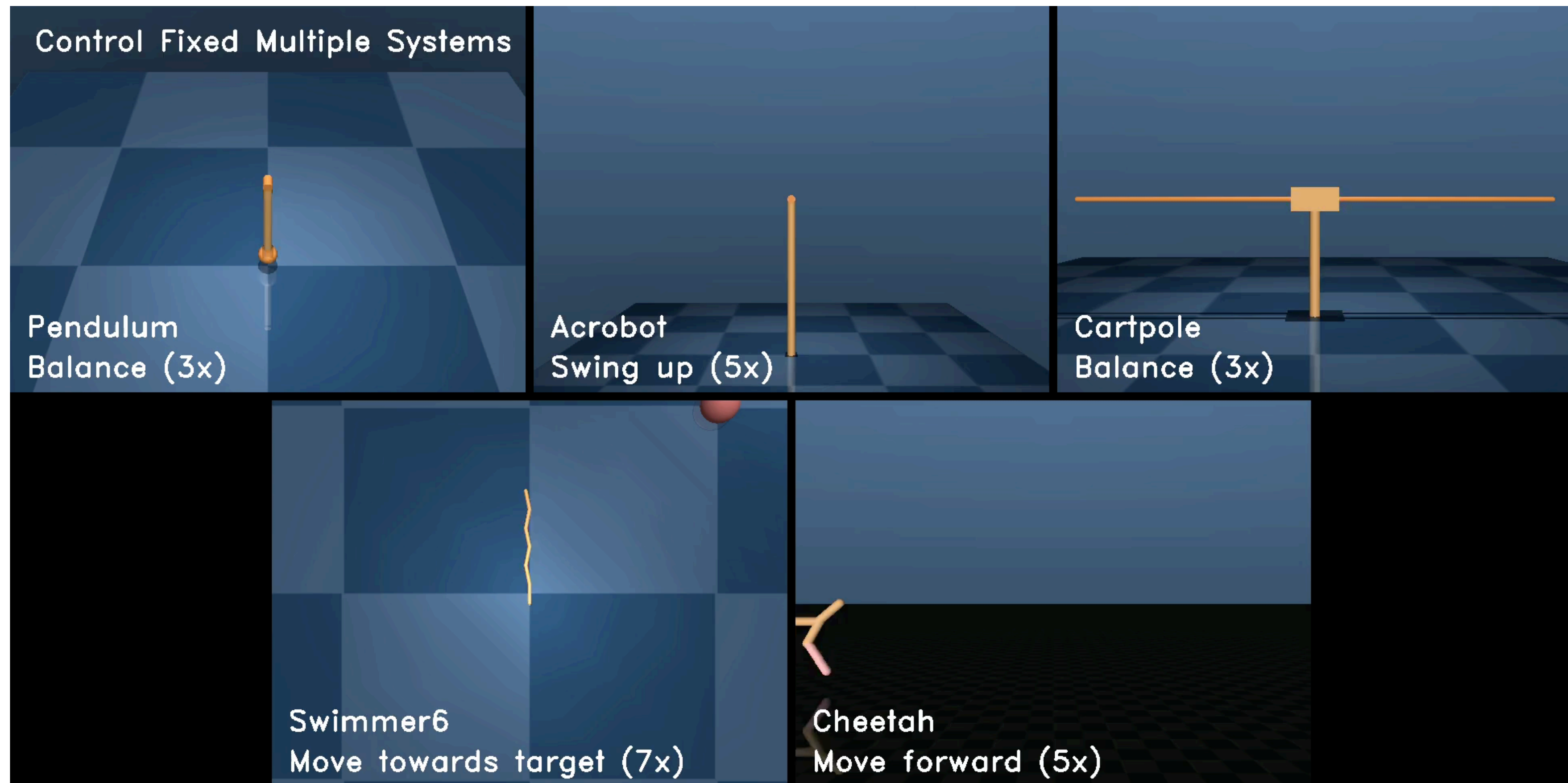


Control: Model-based planning

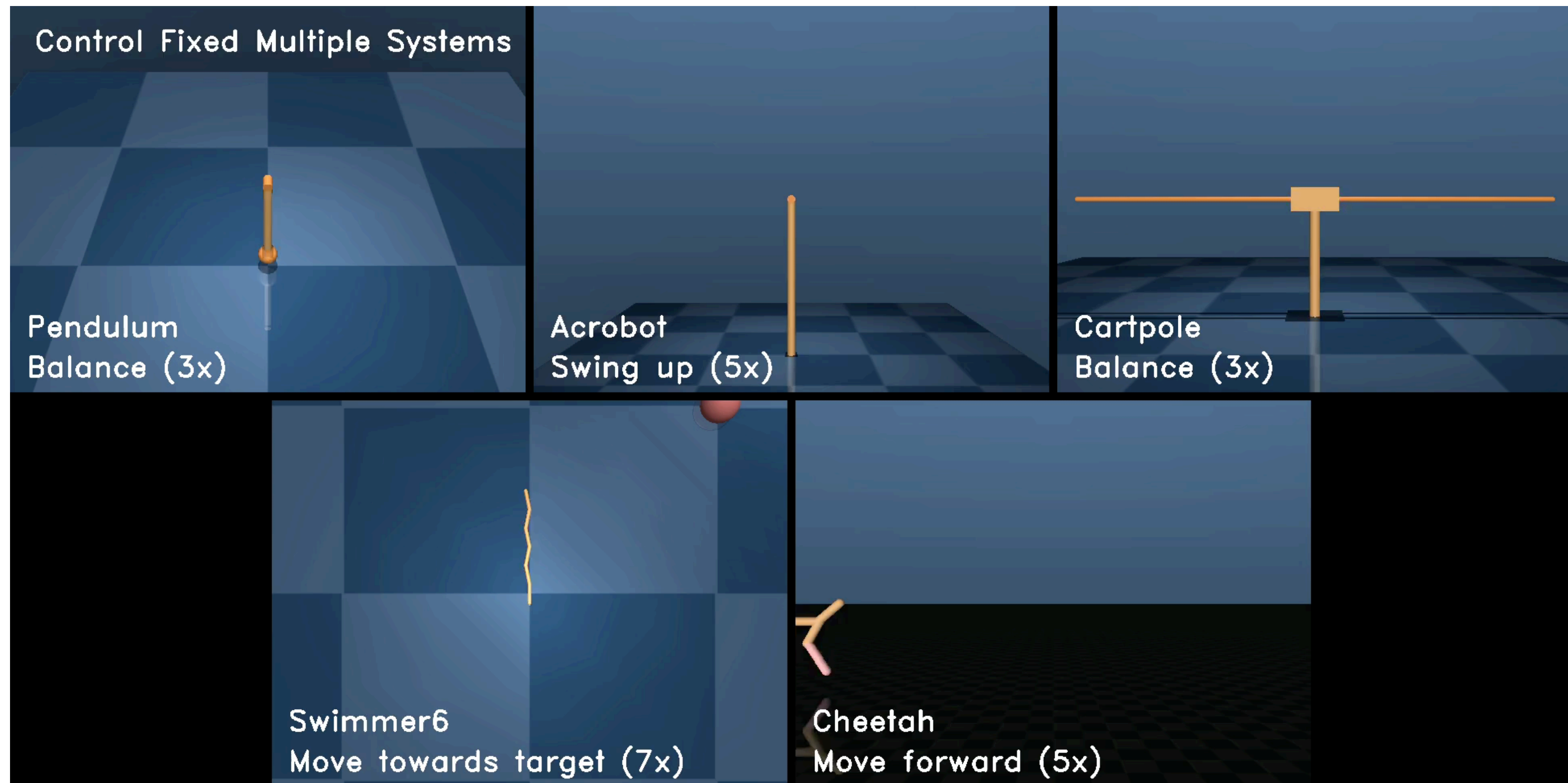
Trajectory optimization: the GN-based forward model is differentiable, so we can backpropagate through it, and find a sequence of actions that maximize reward



Control: Multiple systems via a single model



Control: Multiple systems via a single model



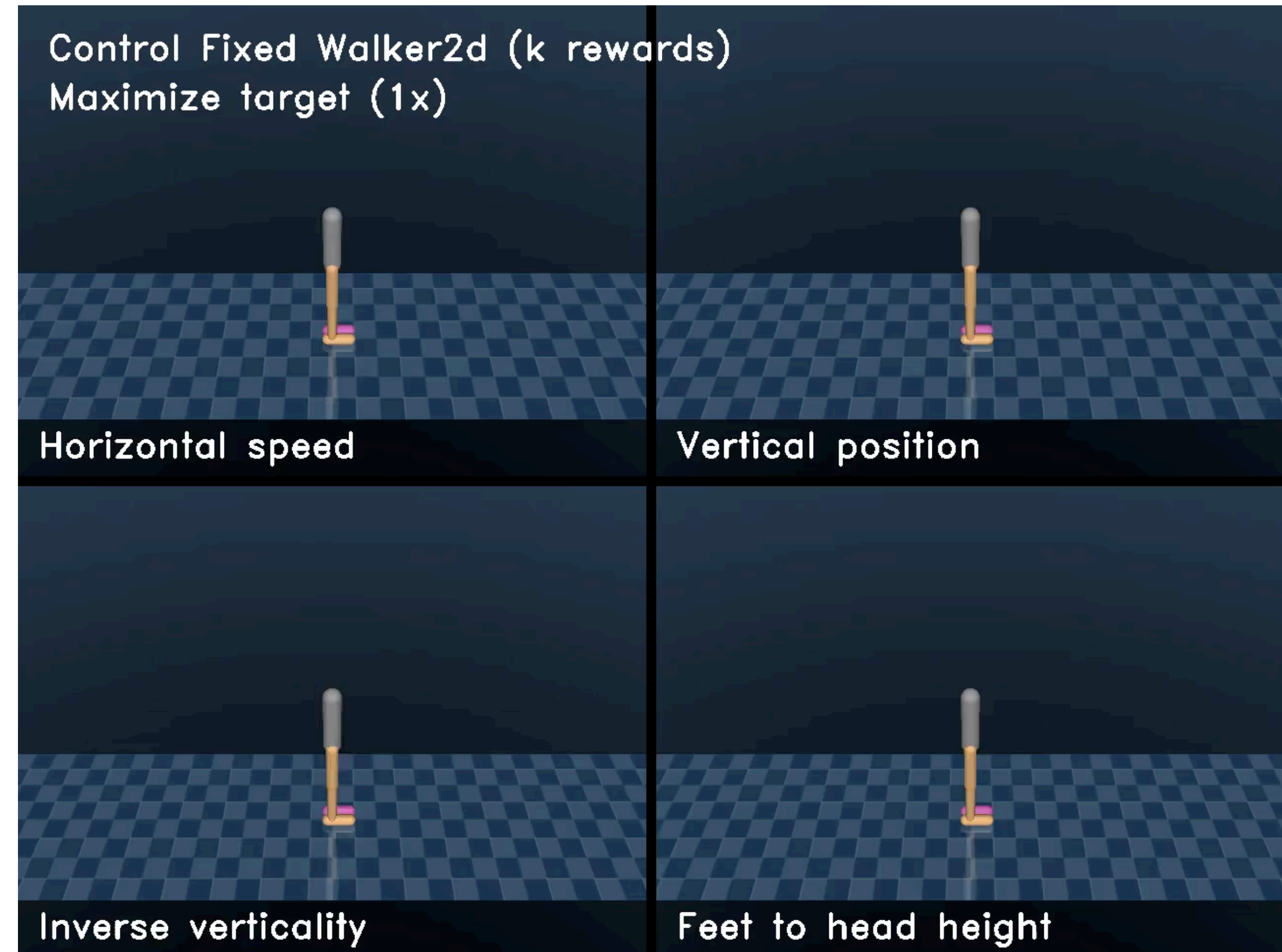
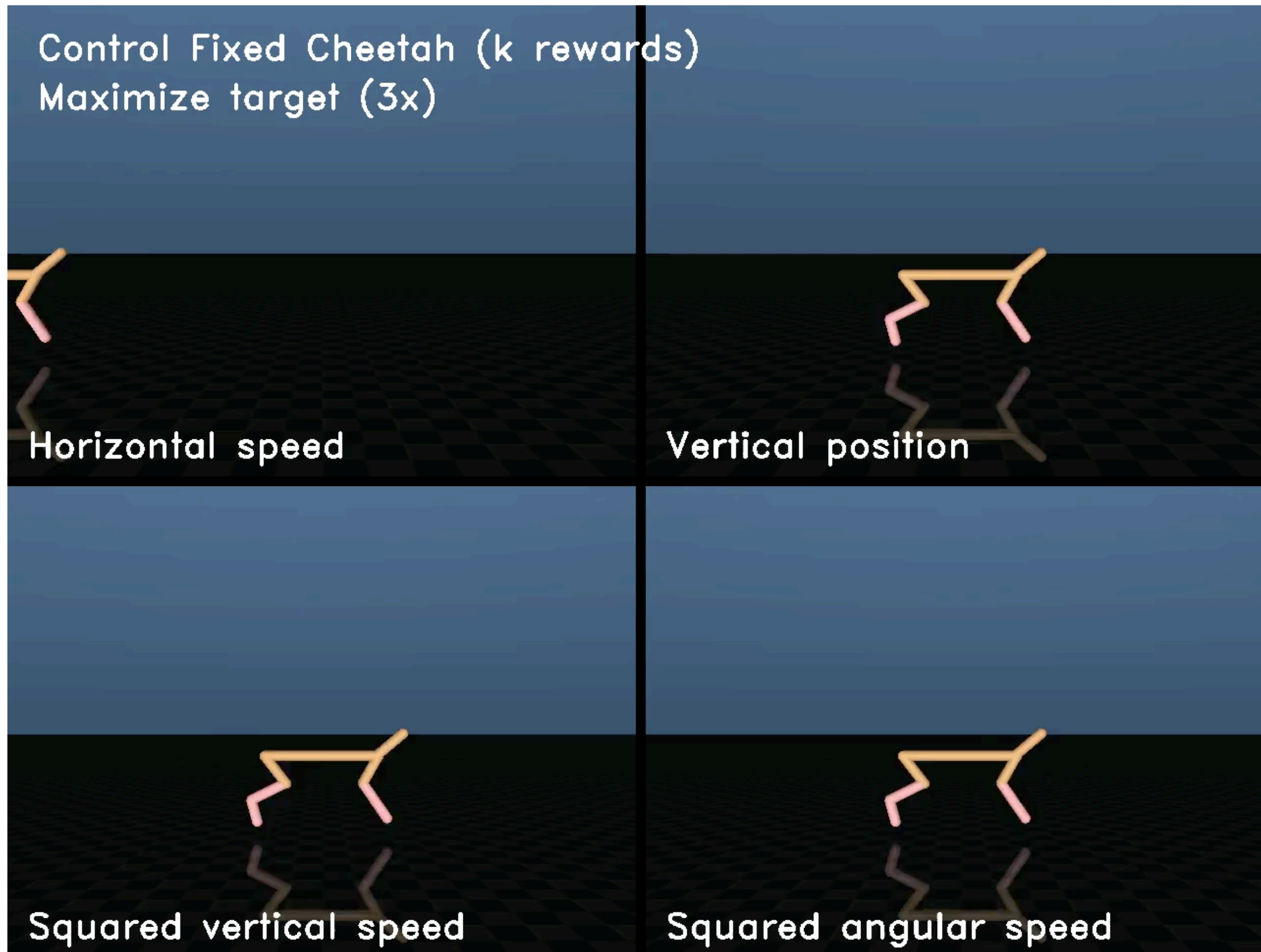
Control: Zero-shot control



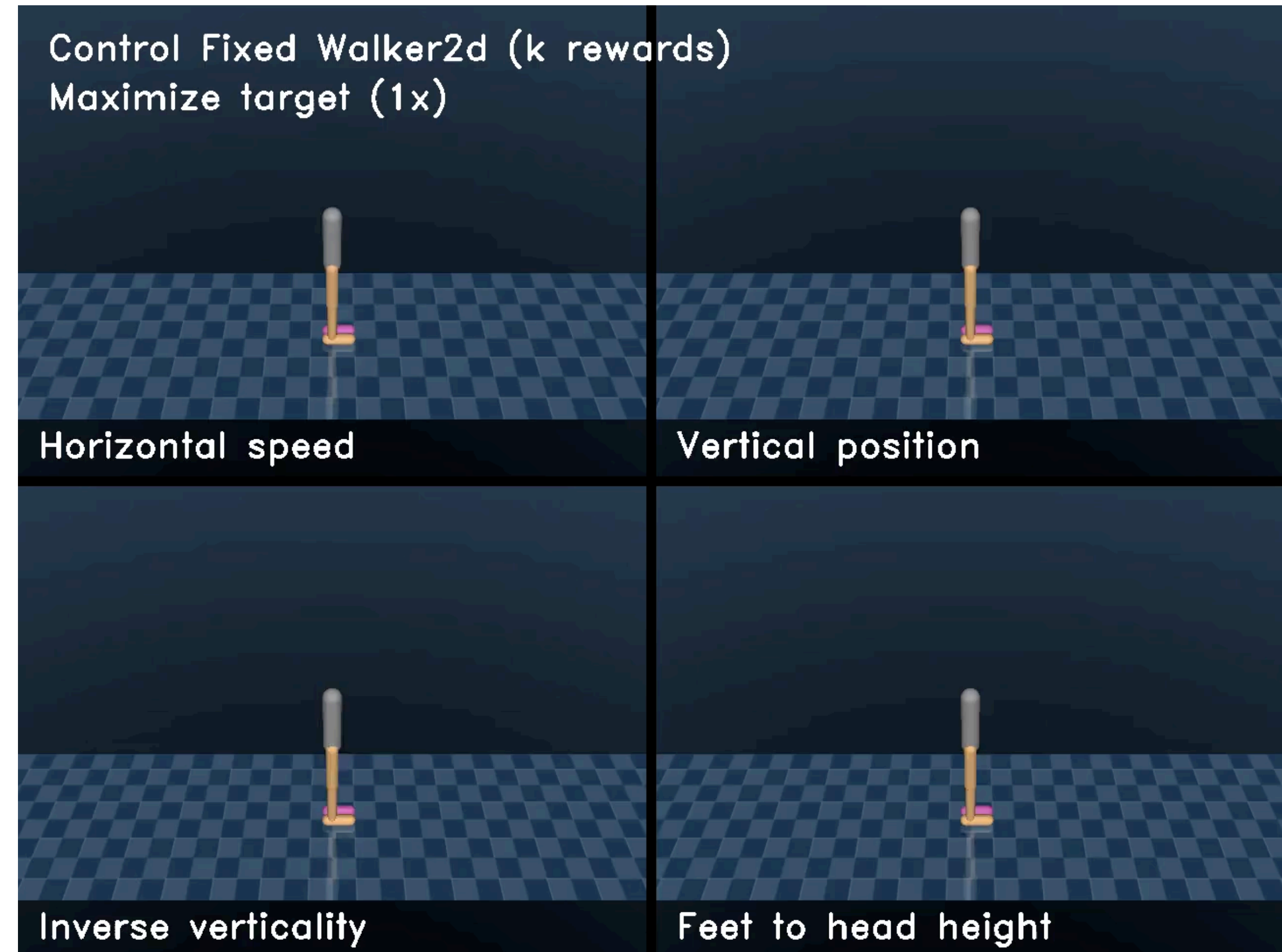
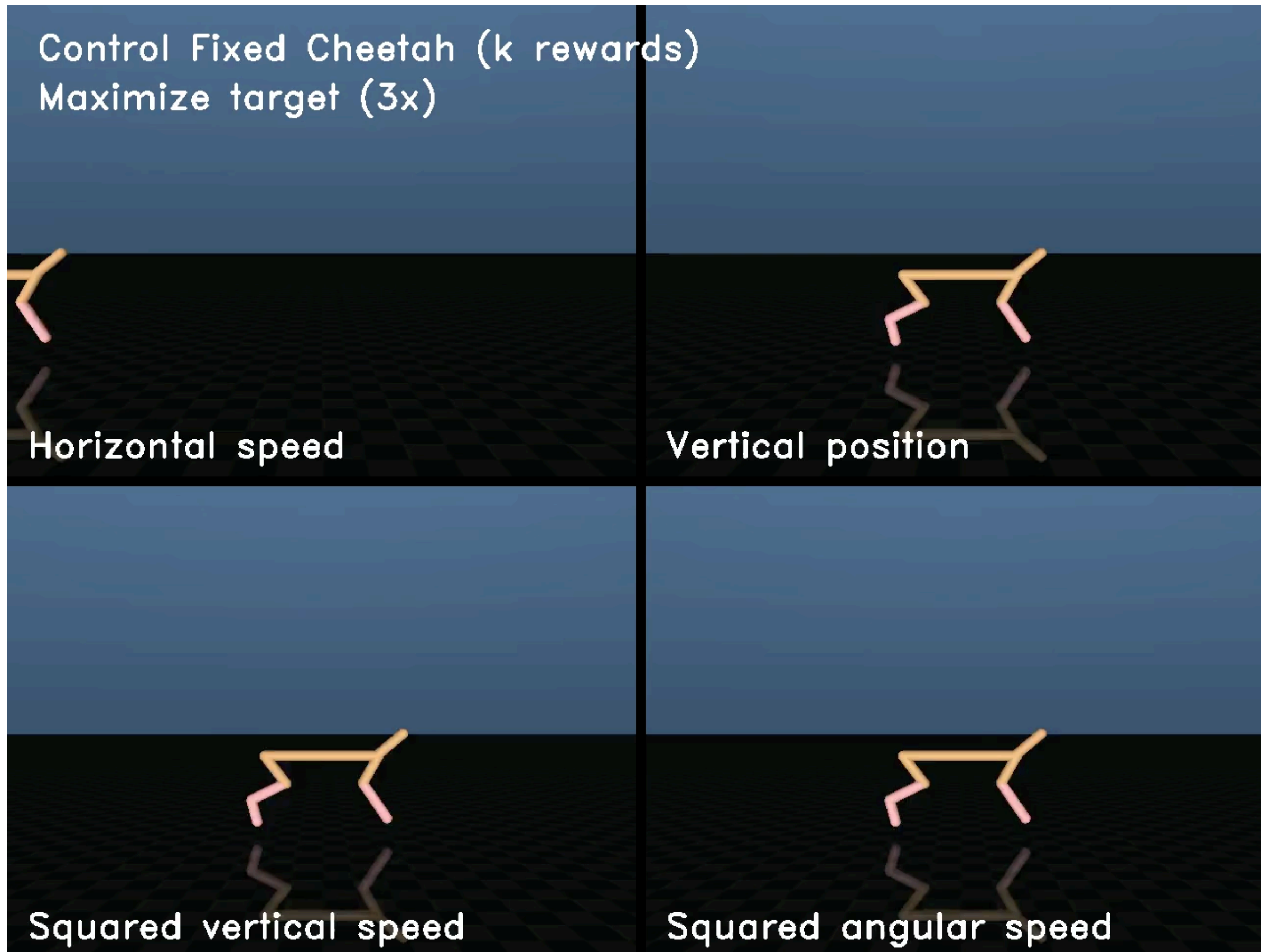
Control: Zero-shot control



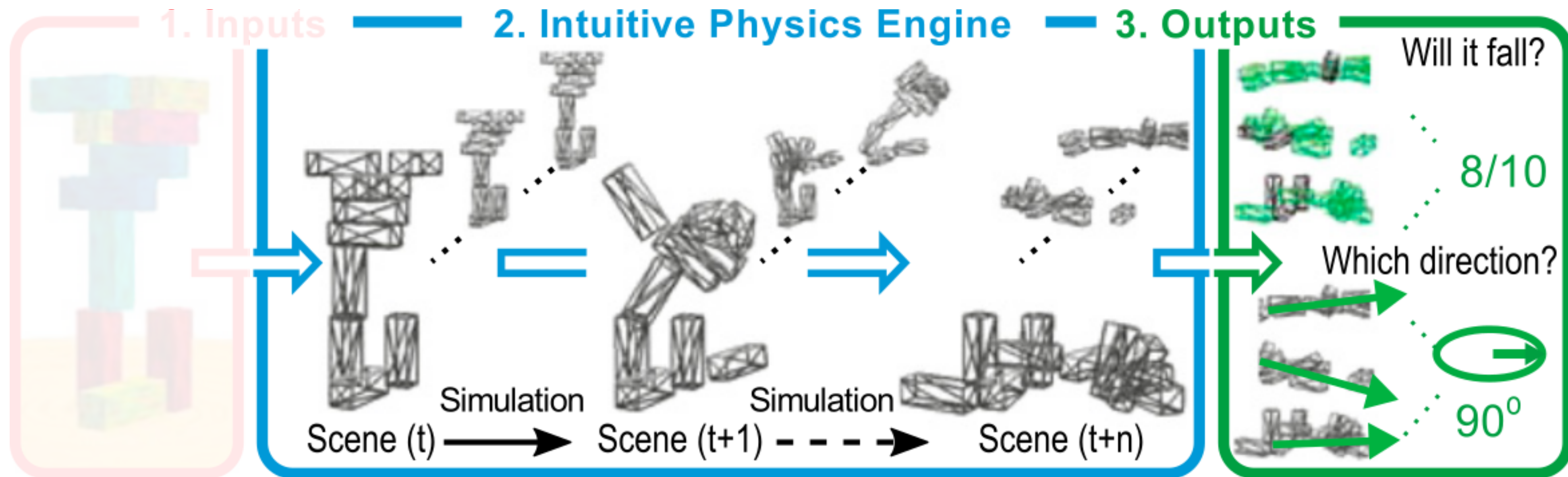
Control: Multiple reward functions



Control: Multiple reward functions

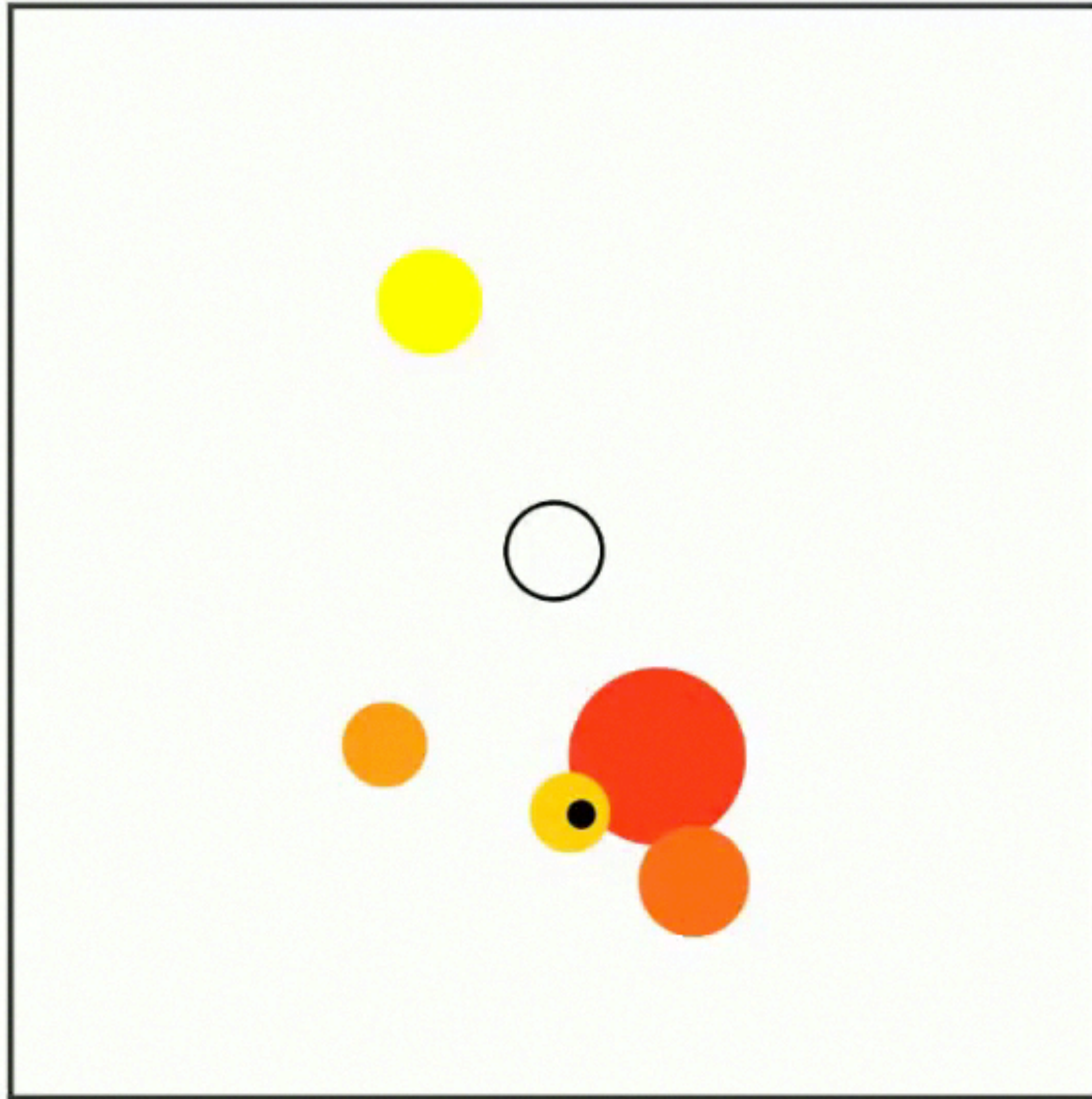


Learning to use mental simulation



Learning to use mental simulation

"Imagination-based metacontroller"



"Spaceship task":

- Navigate to your home planet by choosing a force vector
- Challenging because the planets exert gravity

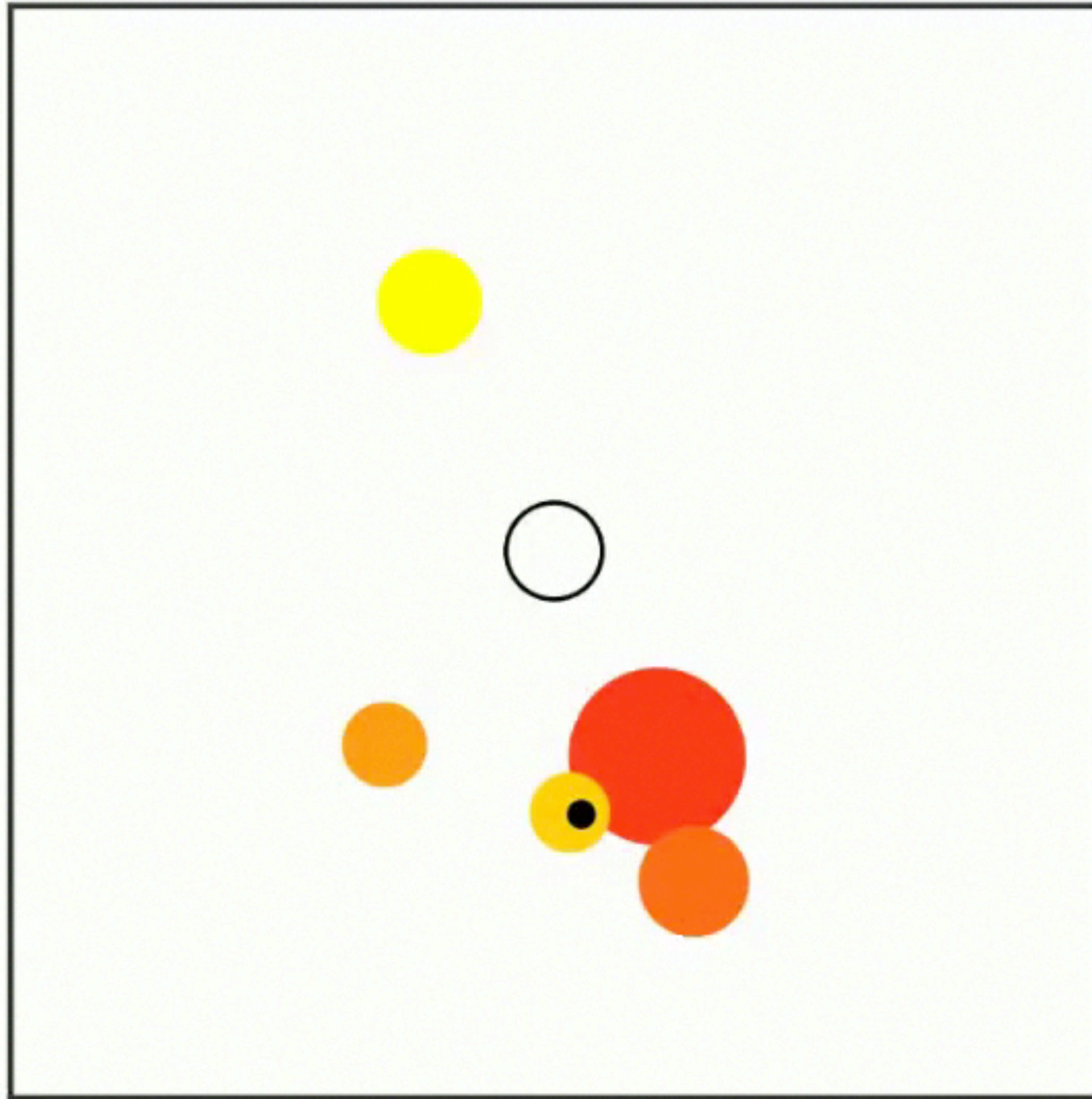
The agent learns 3 components:

1. Action policy (via stochastic value gradients (Heess et al. 2015))
2. GN-based forward model (via supervised 1-step training)
3. Internal strategy for using imagination to test potential actions before selecting one to execute (via REINFORCE)

Hamrick et al., 2017, ICLR

Learning to use mental simulation

"Imagination-based metacontroller"



"Spaceship task":

- Navigate to your home planet by choosing a force vector
- Challenging because the planets exert gravity

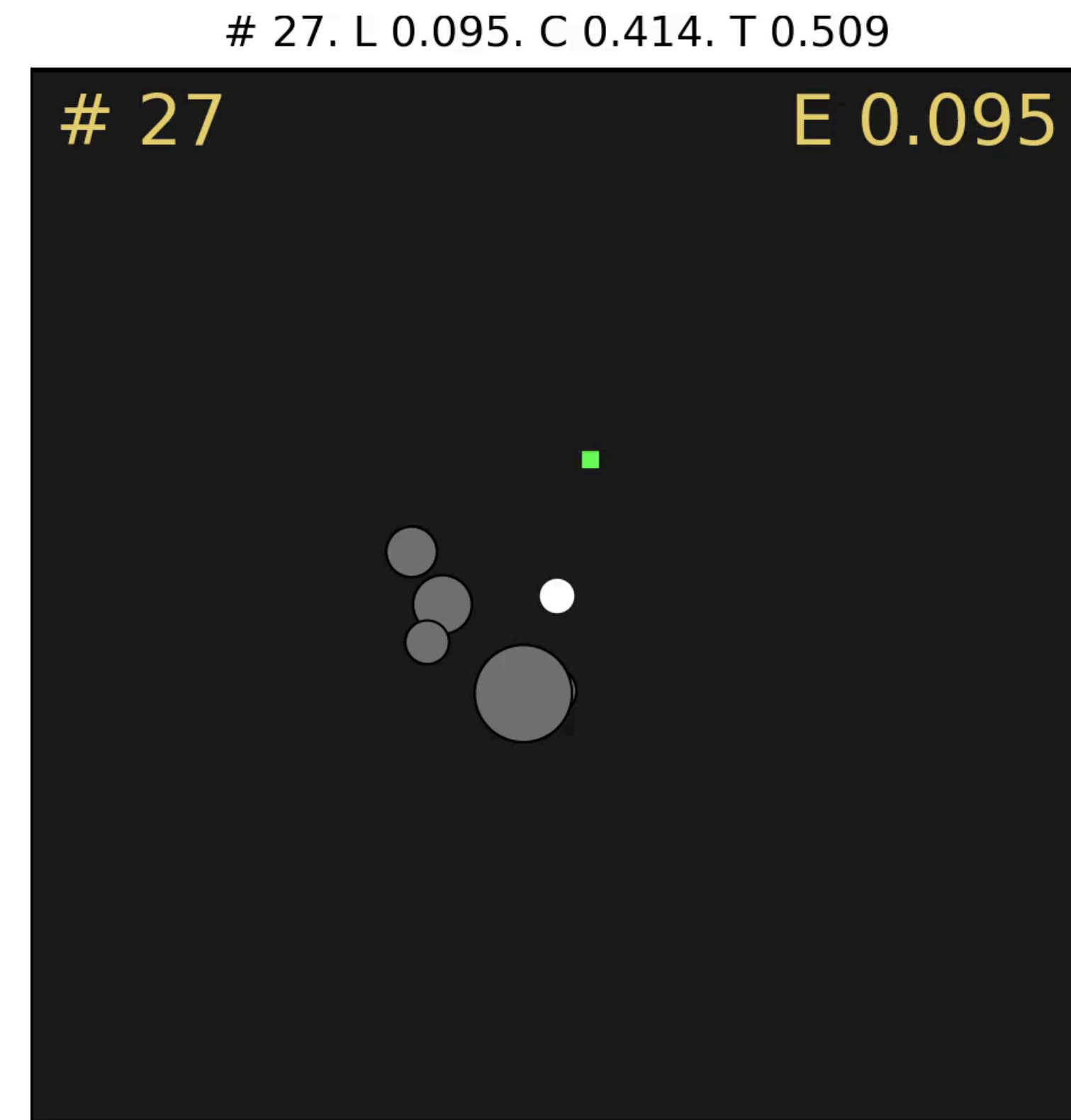
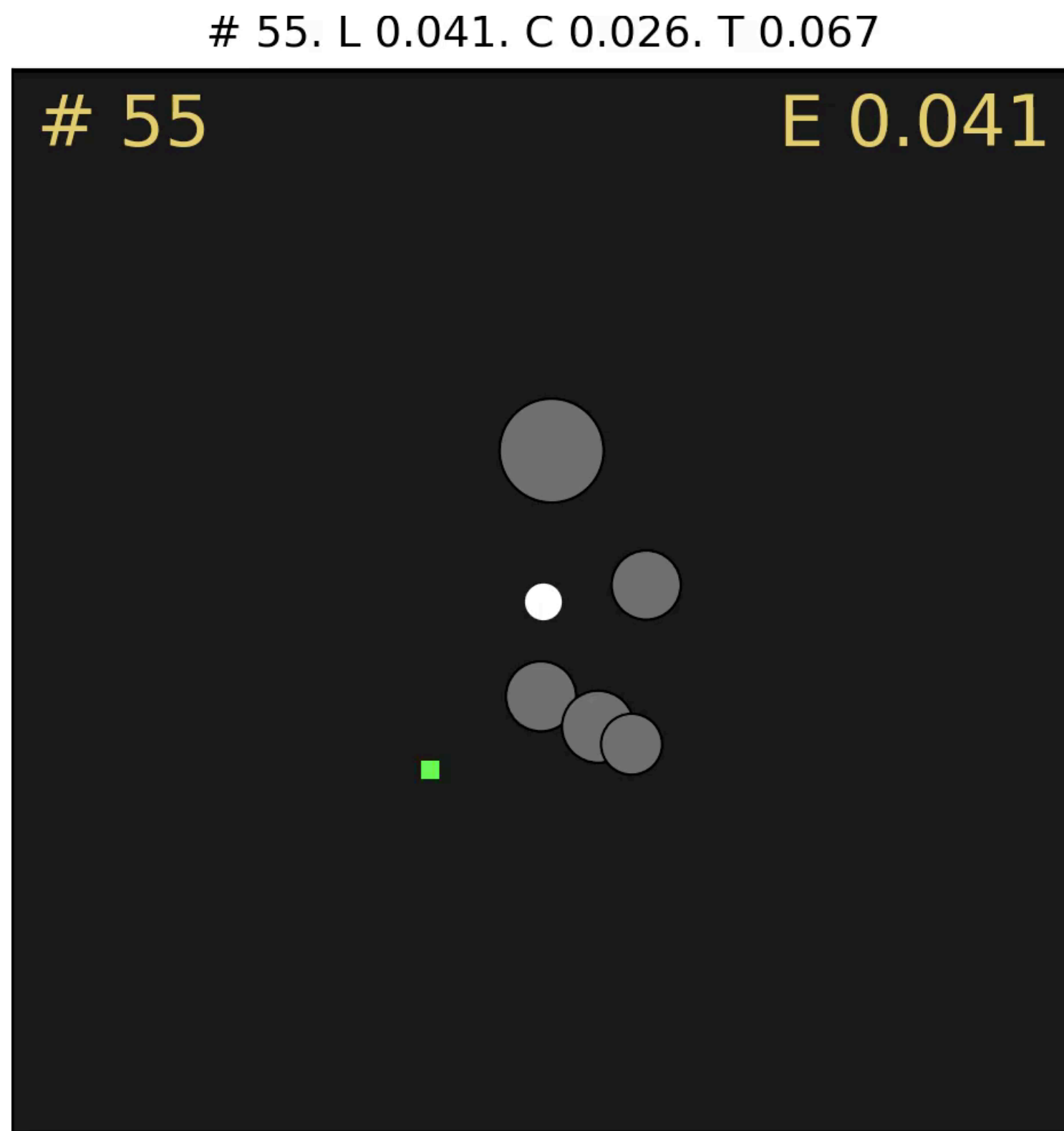
The agent learns 3 components:

1. Action policy (via stochastic value gradients (Heess et al. 2015))
2. GN-based forward model (via supervised 1-step training)
3. Internal strategy for using imagination to test potential actions before selecting one to execute (via REINFORCE)

Hamrick et al., 2017, ICLR

Learning to use mental simulation

"Imagination-based planner"

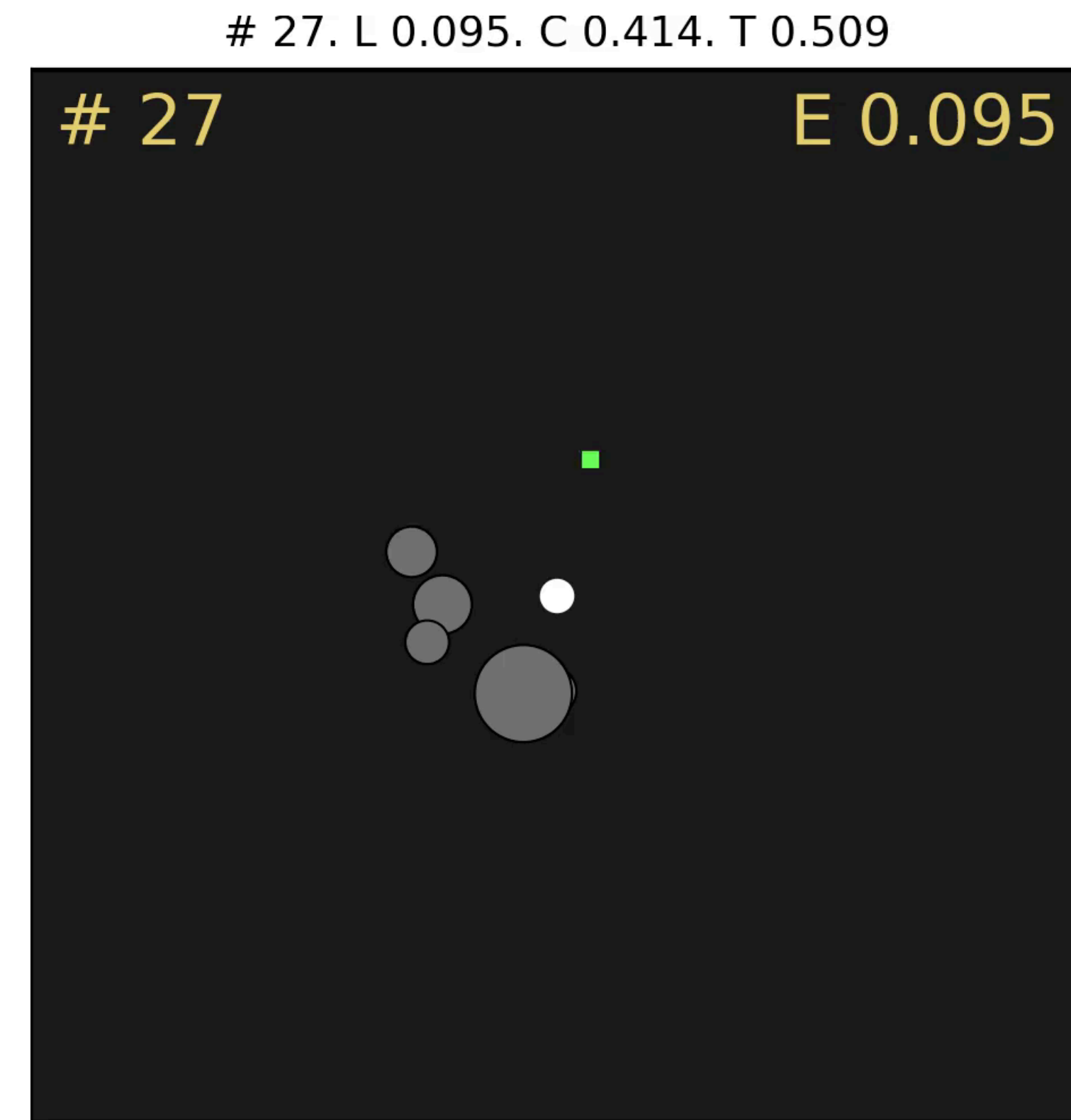
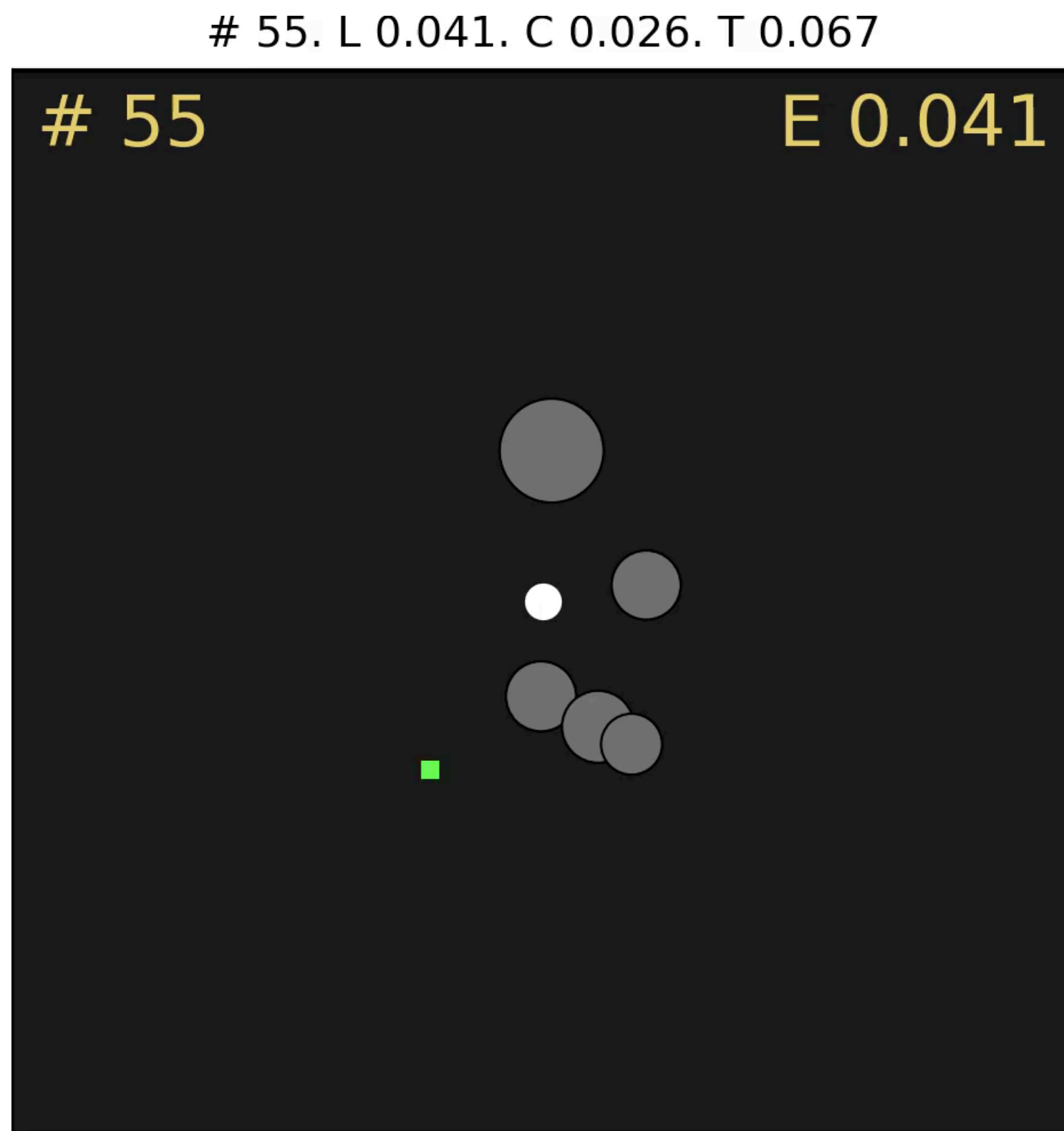


- Red: real actions
- Blue: 1 step of imagination
- Green: 2+ steps of imagination

Pascanu et al., 2017, arXiv

Learning to use mental simulation

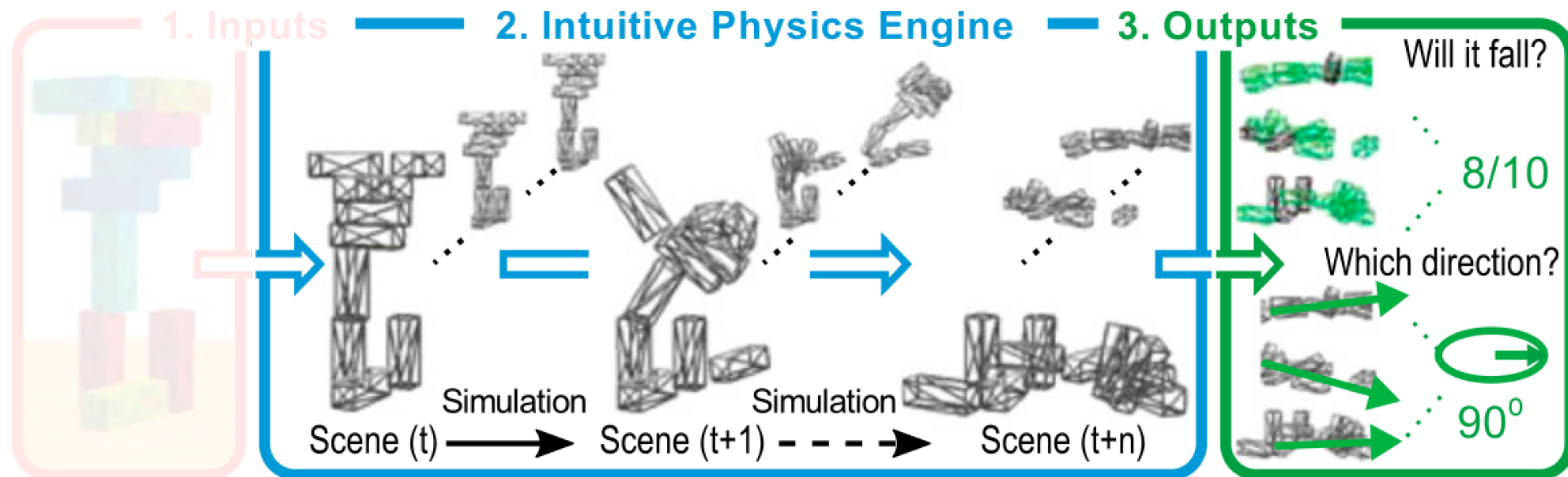
"Imagination-based planner"



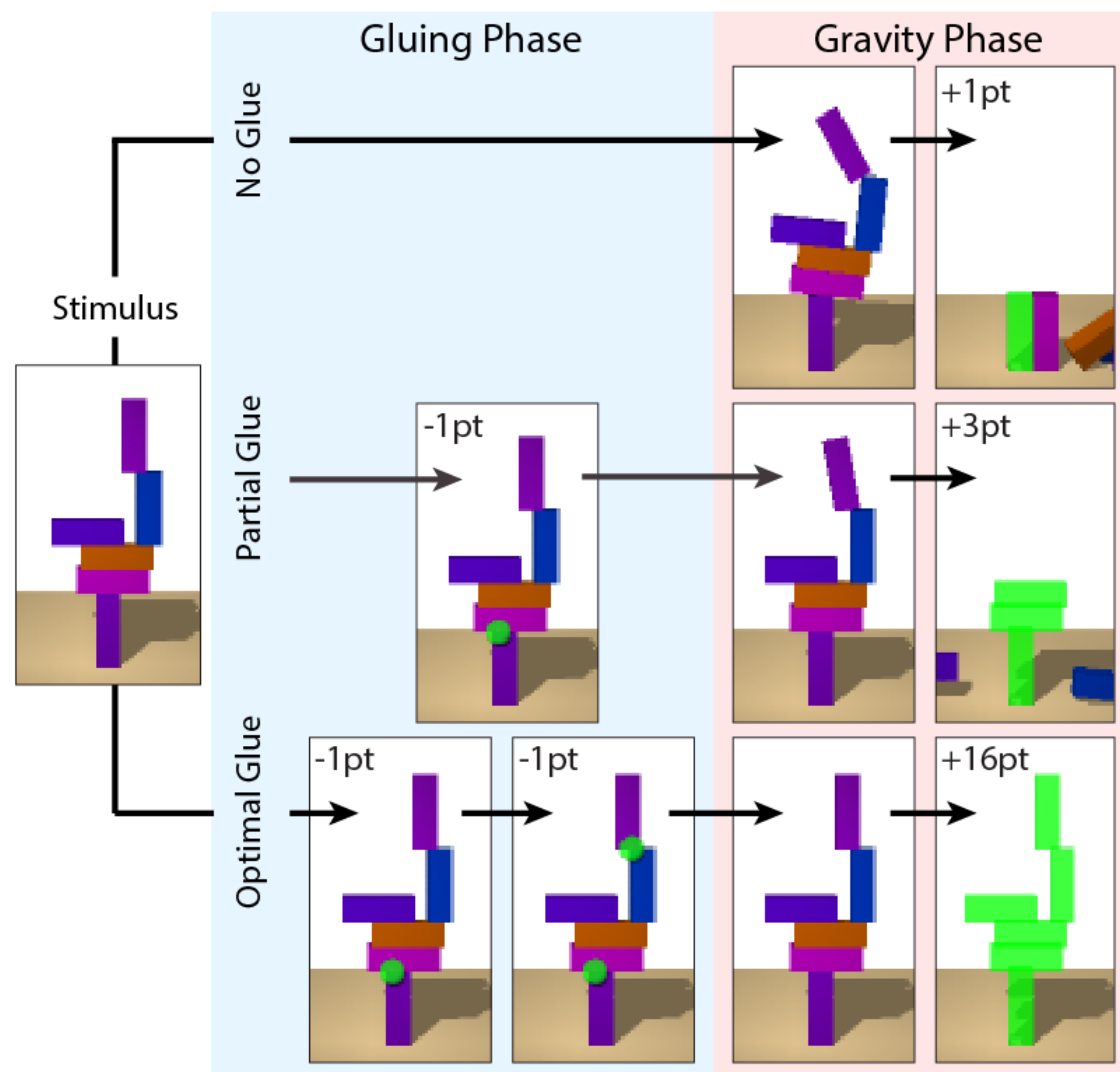
- Red: real actions
- Blue: 1 step of imagination
- Green: 2+ steps of imagination

Pascanu et al., 2017, arXiv

Graph-structured model-free policies



Graph-structured model-free policies for physical construction in humans and AI



Jess Hamrick, Kelsey Allen,
Victor Bapst, Tina Zhu, Kevin McKee,
Josh Tenenbaum, Peter Battaglia
Proc Cog Sci, 2018

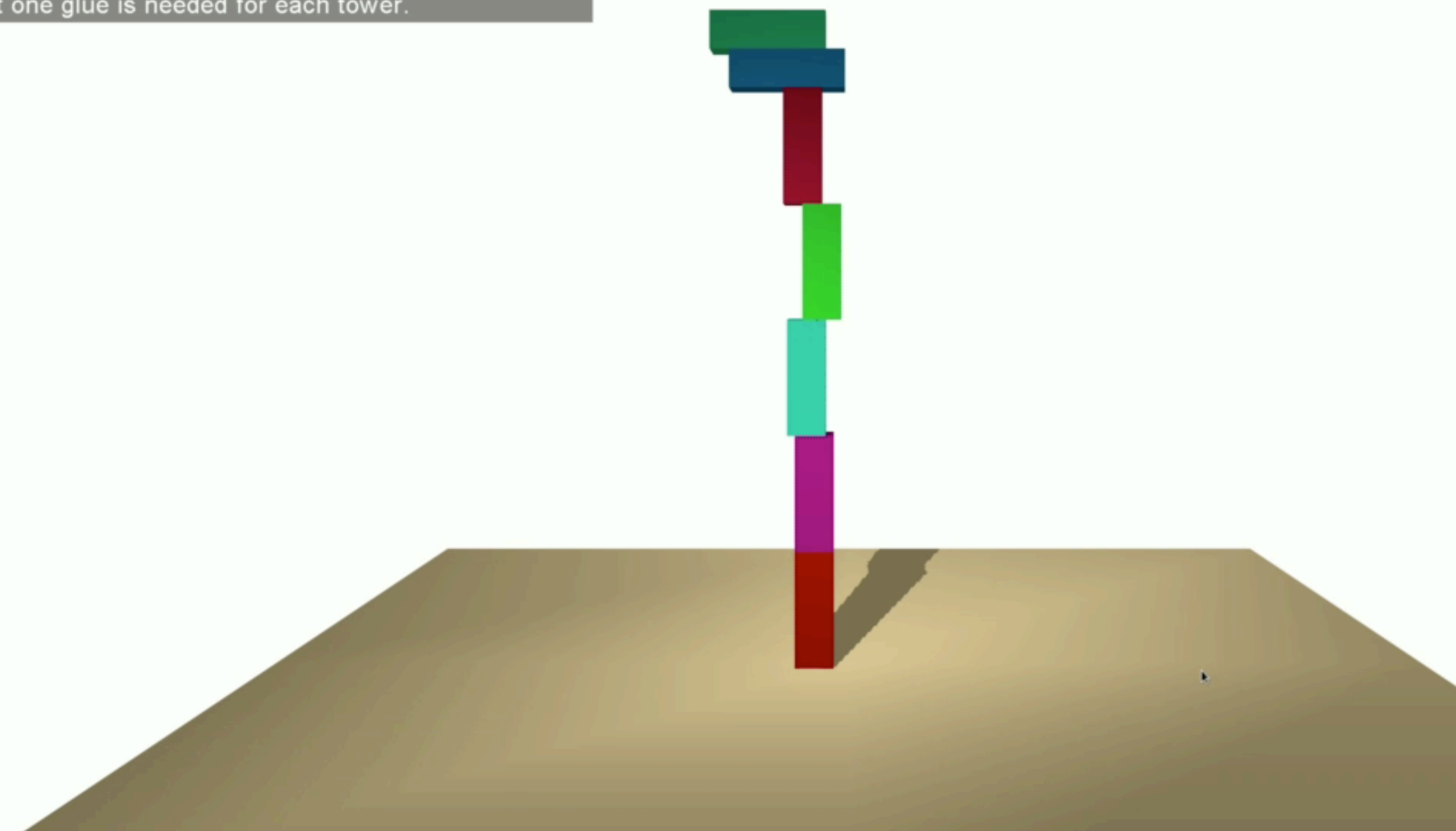
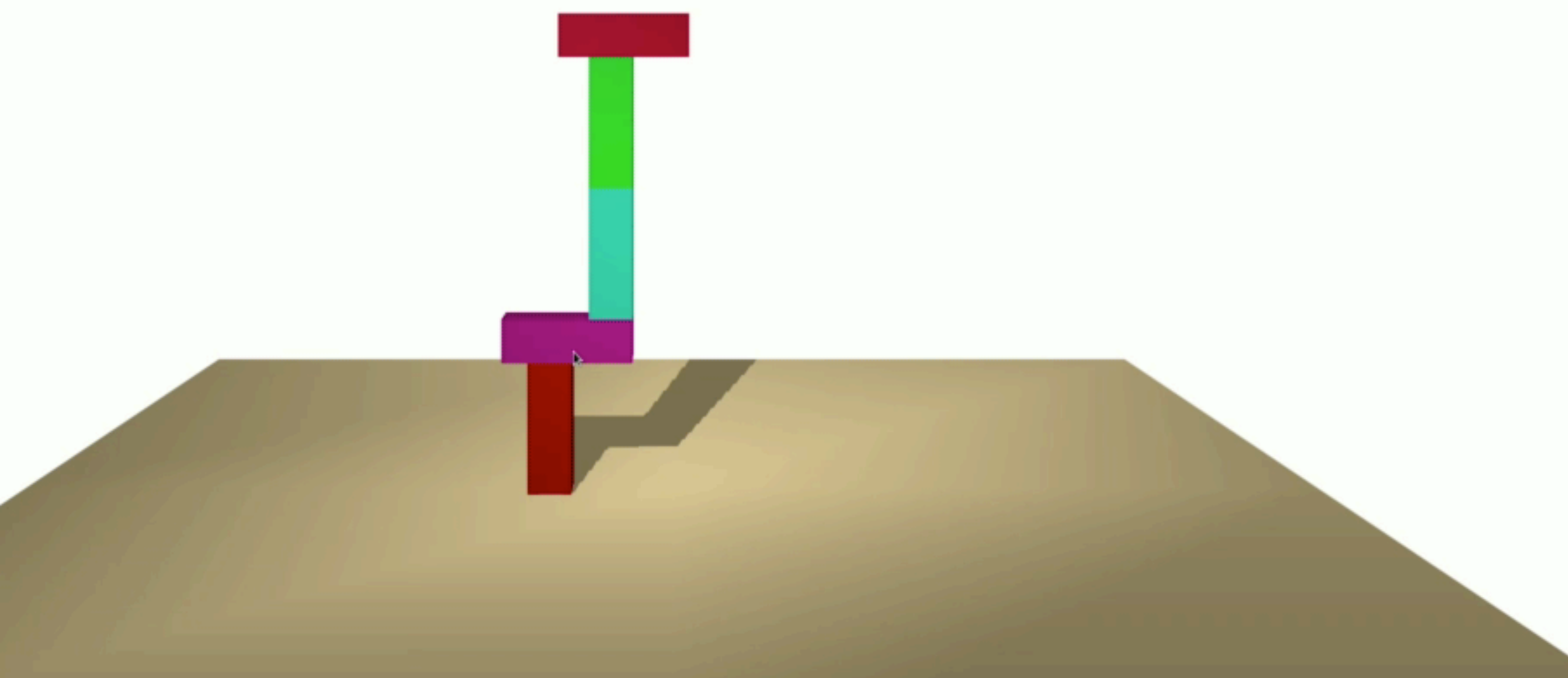
The "glue task"

Goal: Glue blocks together to make the tower stable, using the minimum amount of glue.

Structured model-free policies for physical construction in humans and AI

le)
to select it.
floor) to glue them together.
to the tower.
that doesn't fall.
ed costs 1pt.
to keep the tower stable,
or each tower.

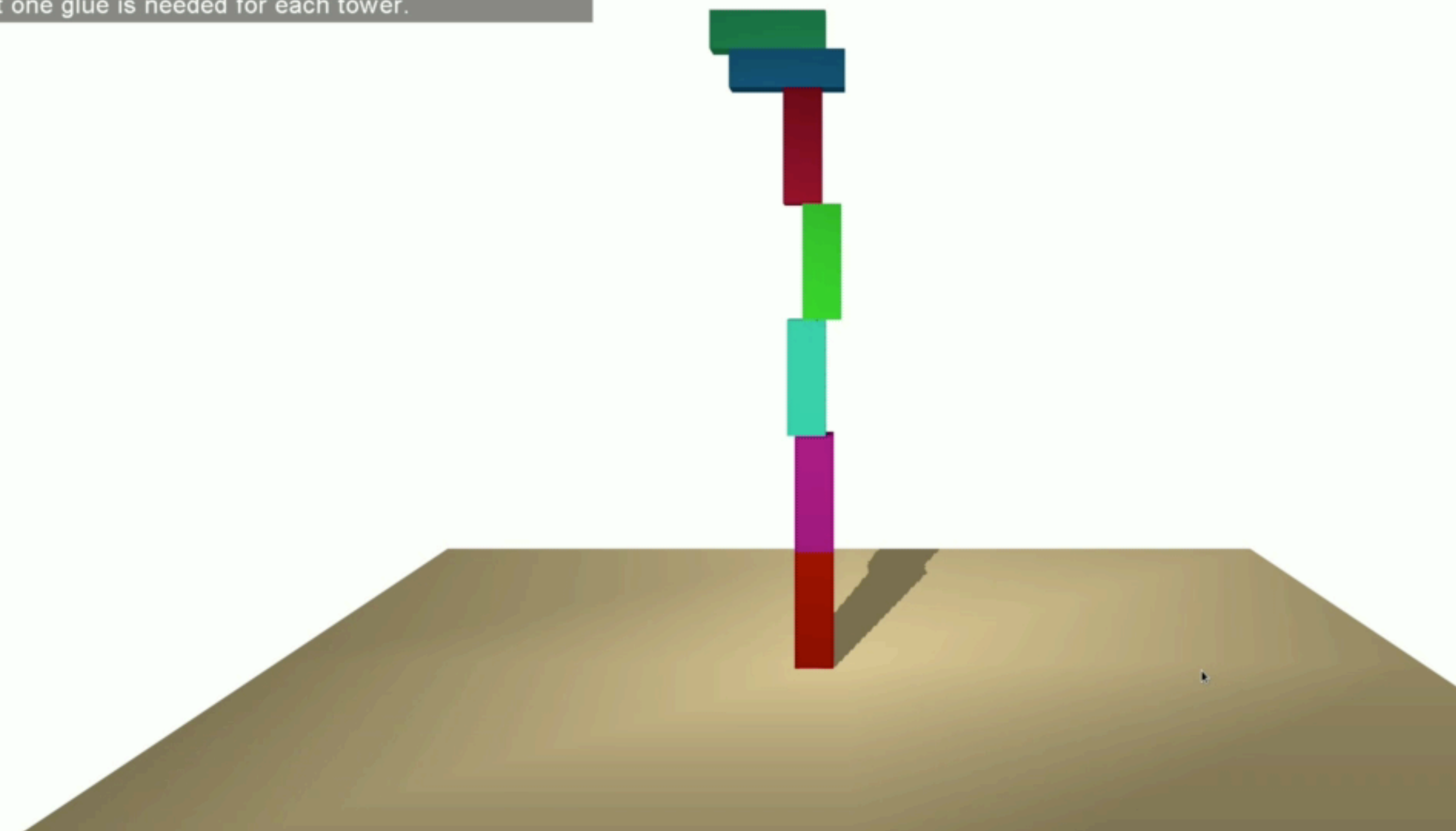
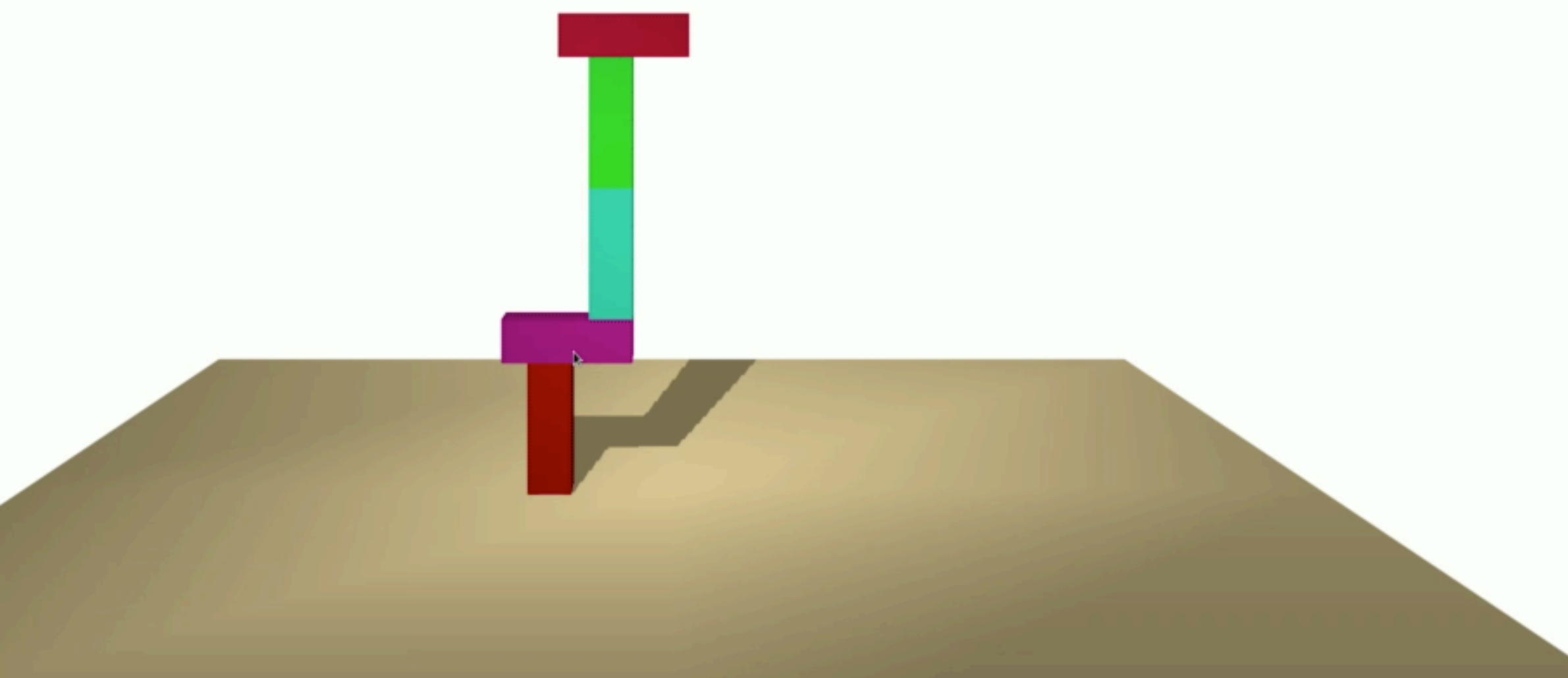
Mode	Instructions (press 'h' to show/hide)
Trial	1. Click on a block (or the floor) to select it.
Current	2. Click on another block (or the floor) to glue them together.
Total po	3. Press enter to apply gravity to the tower.
	4. You earn 1pt for each block that doesn't fall.
	5. Each pair of blocks that is glued costs 1pt.
	6. If you use the minimum glue to keep the tower stable, you earn a 10pt bonus.
	7. At least one glue is needed for each tower.



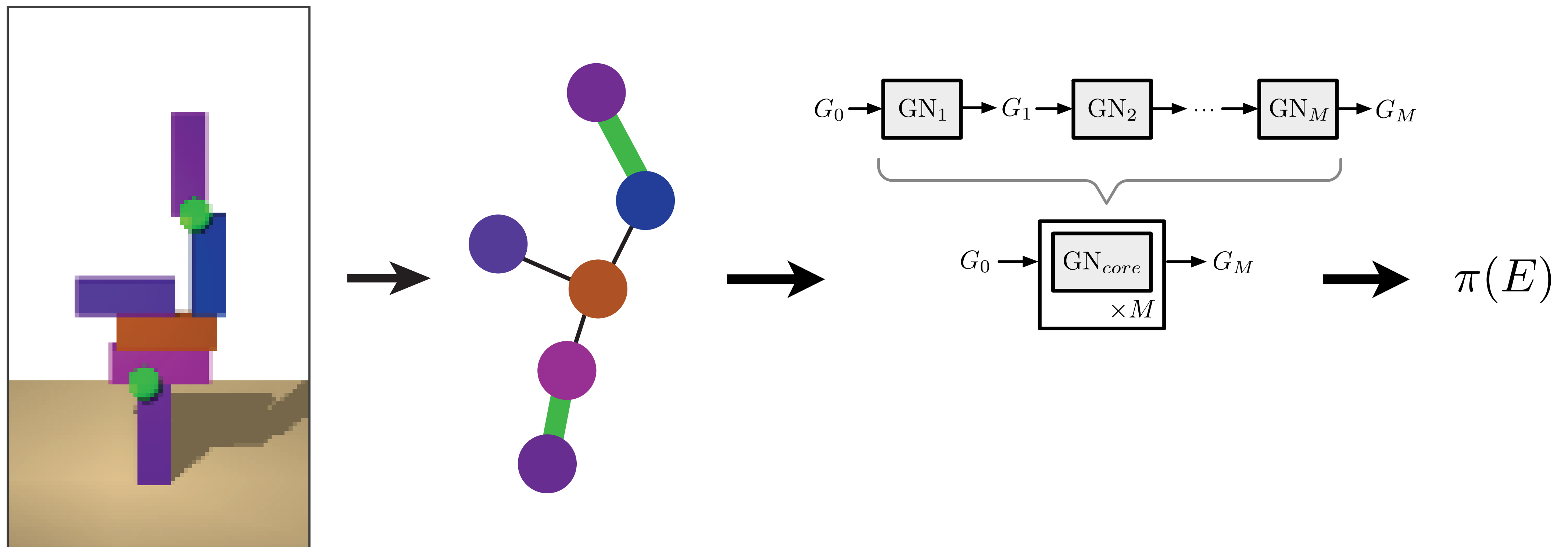
Structured model-free policies for physical construction in humans and AI

le)
to select it.
floor) to glue them together.
to the tower.
that doesn't fall.
ed costs 1pt.
to keep the tower stable,
or each tower.

Mode	Instructions (press 'h' to show/hide)
Trial	1. Click on a block (or the floor) to select it.
Current	2. Click on another block (or the floor) to glue them together.
Total po	3. Press enter to apply gravity to the tower.
	4. You earn 1pt for each block that doesn't fall.
	5. Each pair of blocks that is glued costs 1pt.
	6. If you use the minimum glue to keep the tower stable, you earn a 10pt bonus.
	7. At least one glue is needed for each tower.



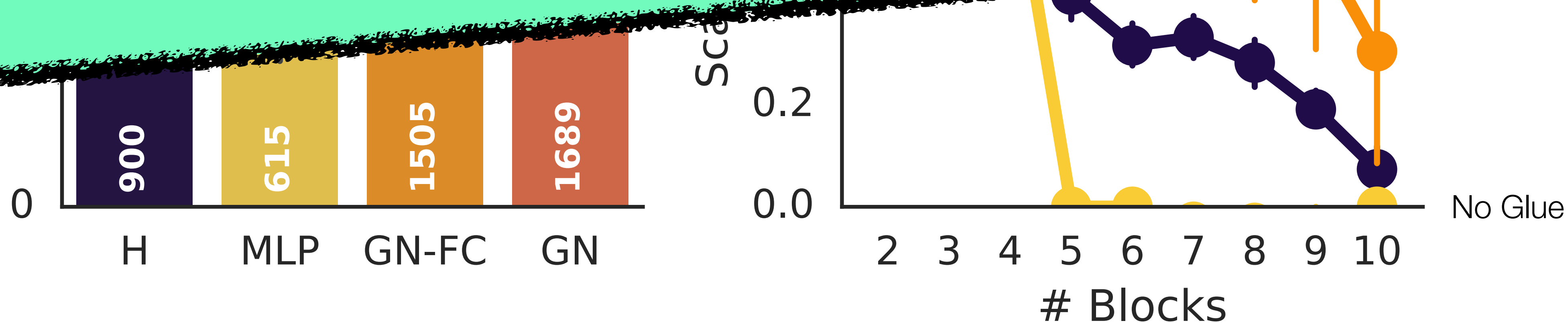
Structured model-free policies for physical construction in humans and AI



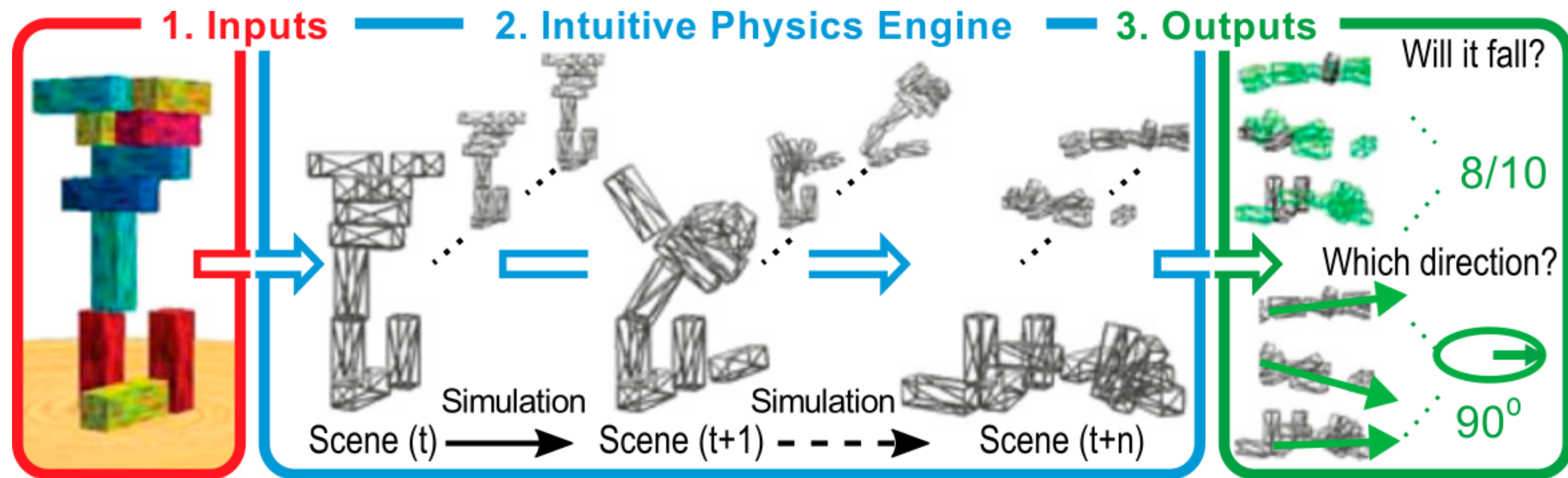
Structured model-free policies for physical construction in humans and AI

(a)

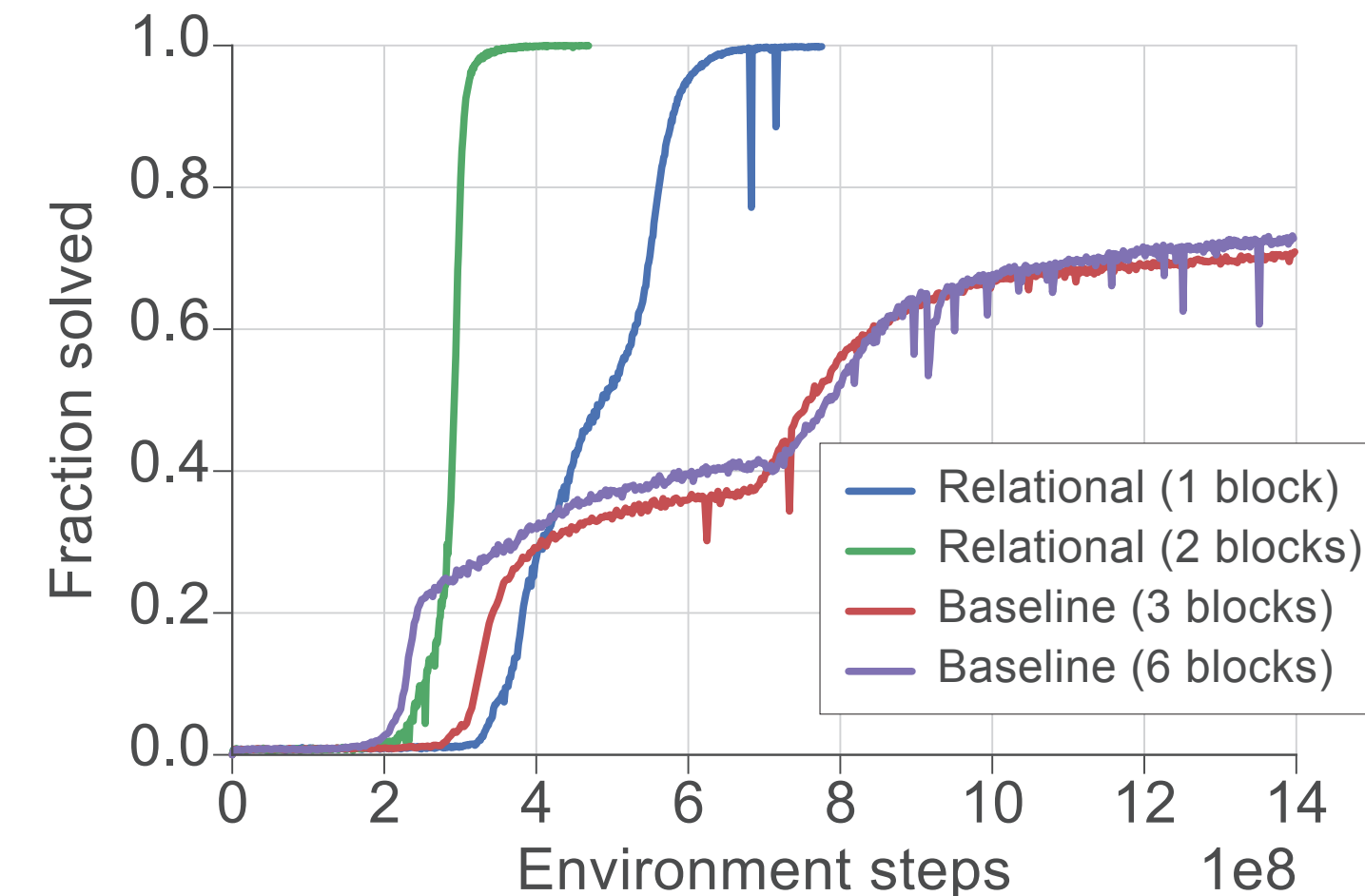
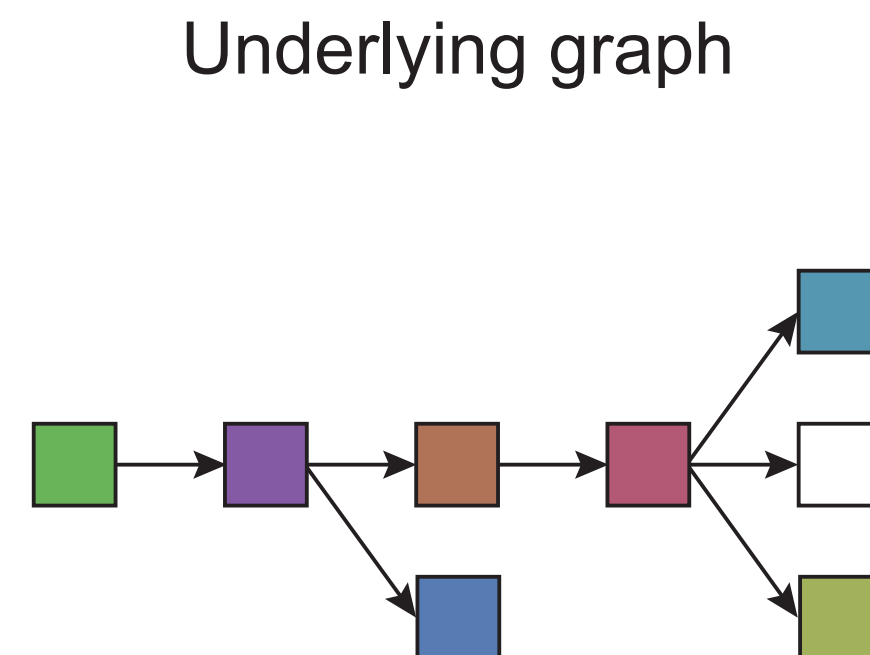
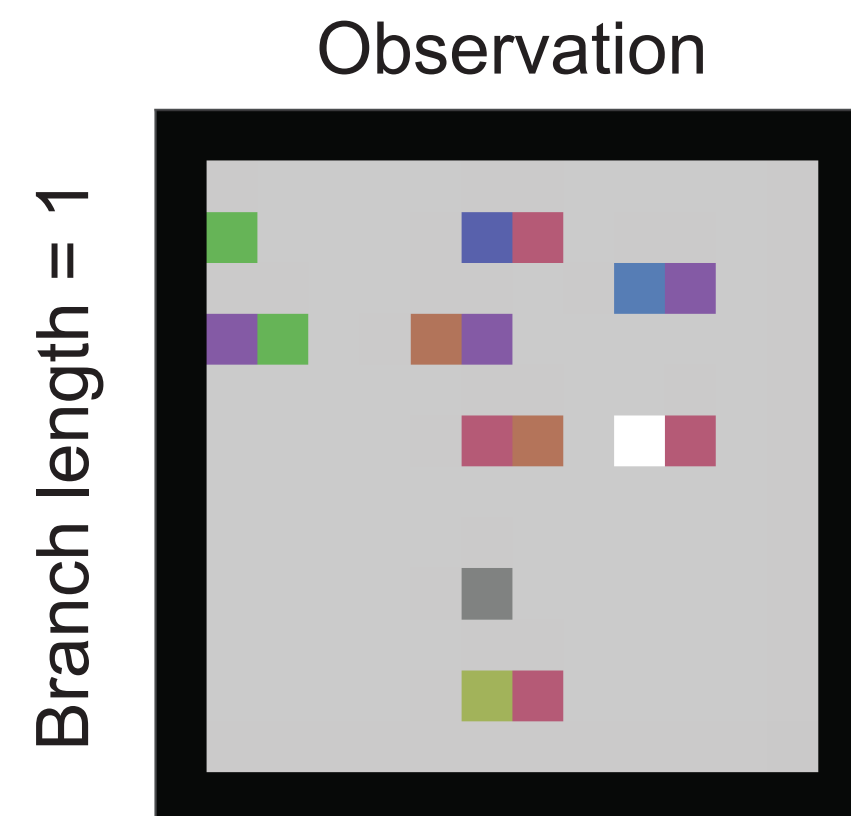
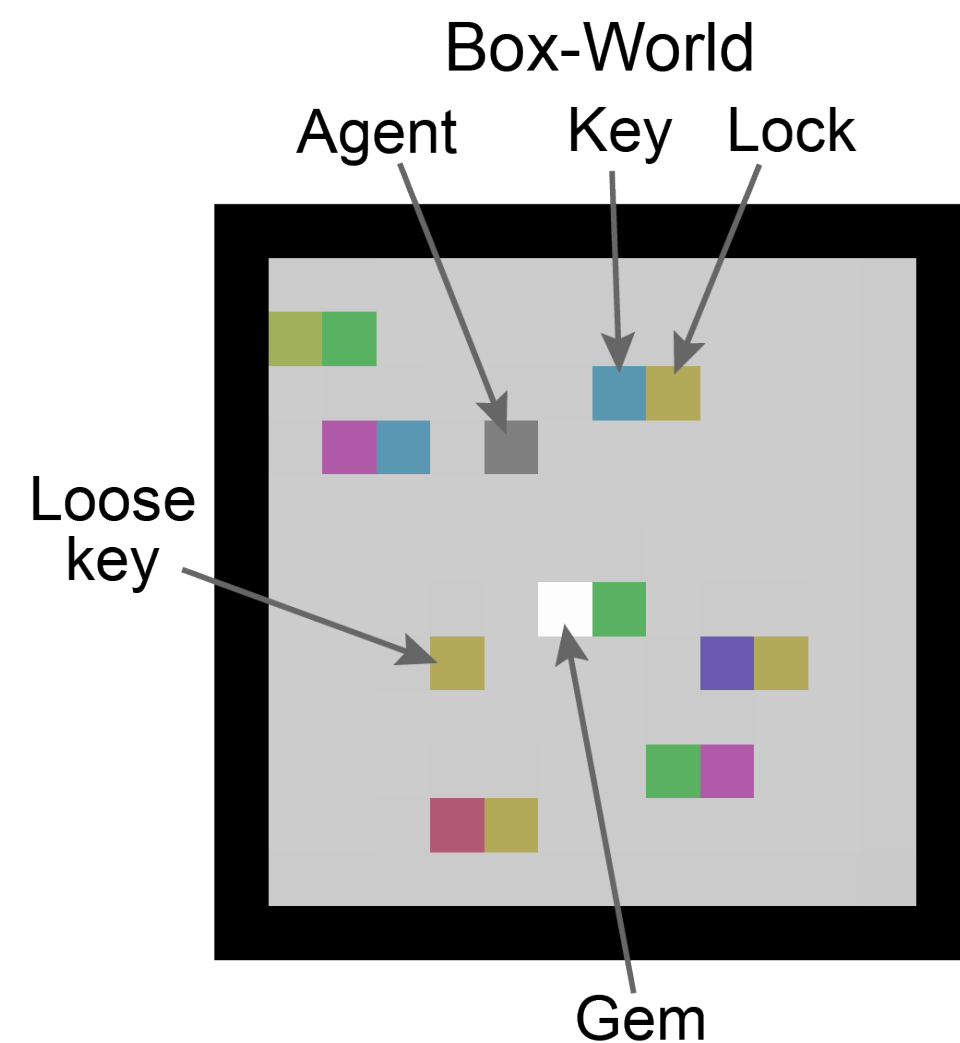
For the full story and results:
see Jess and Kelsey's poster to my right



Graph-structured representations for model-free RL

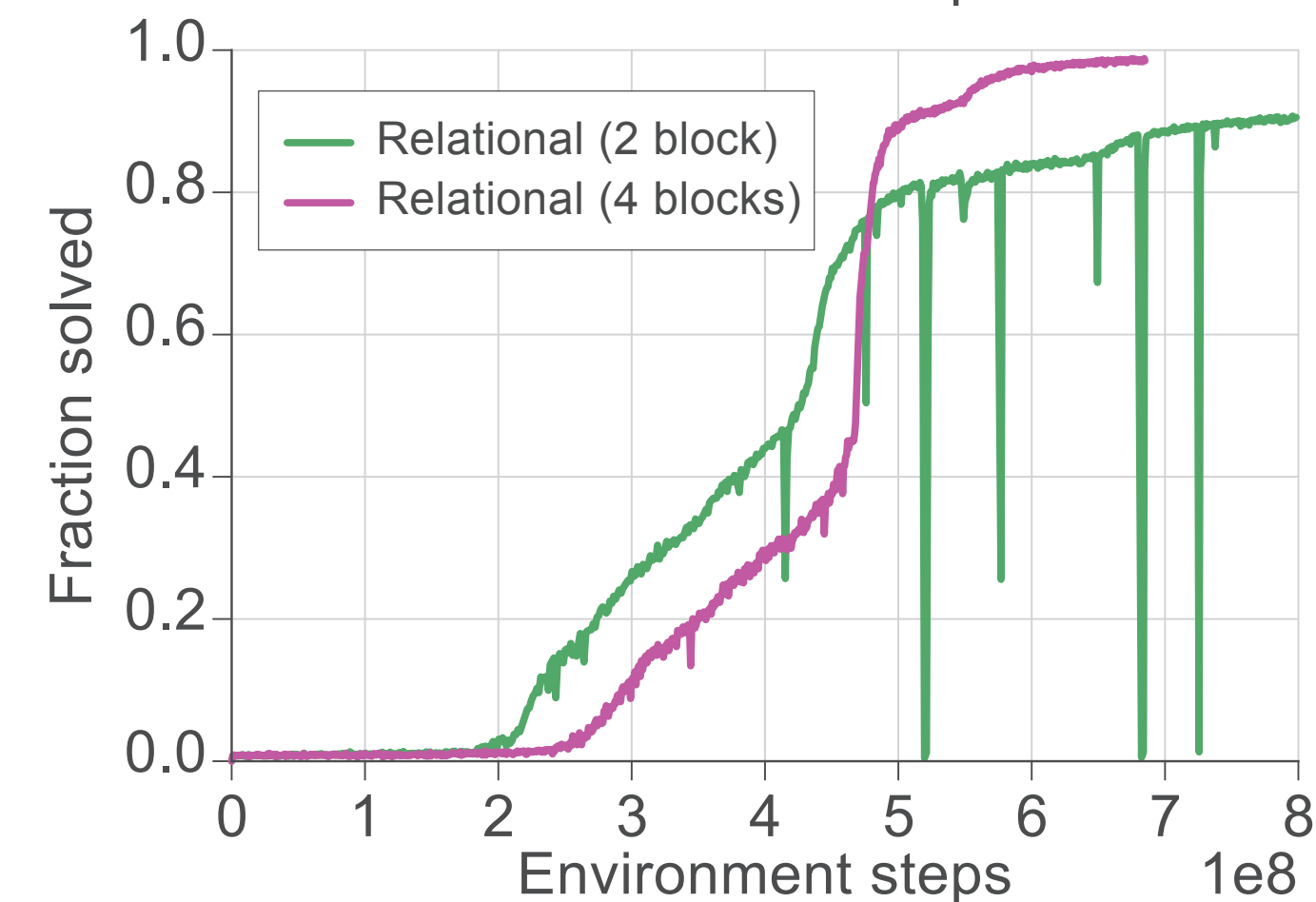
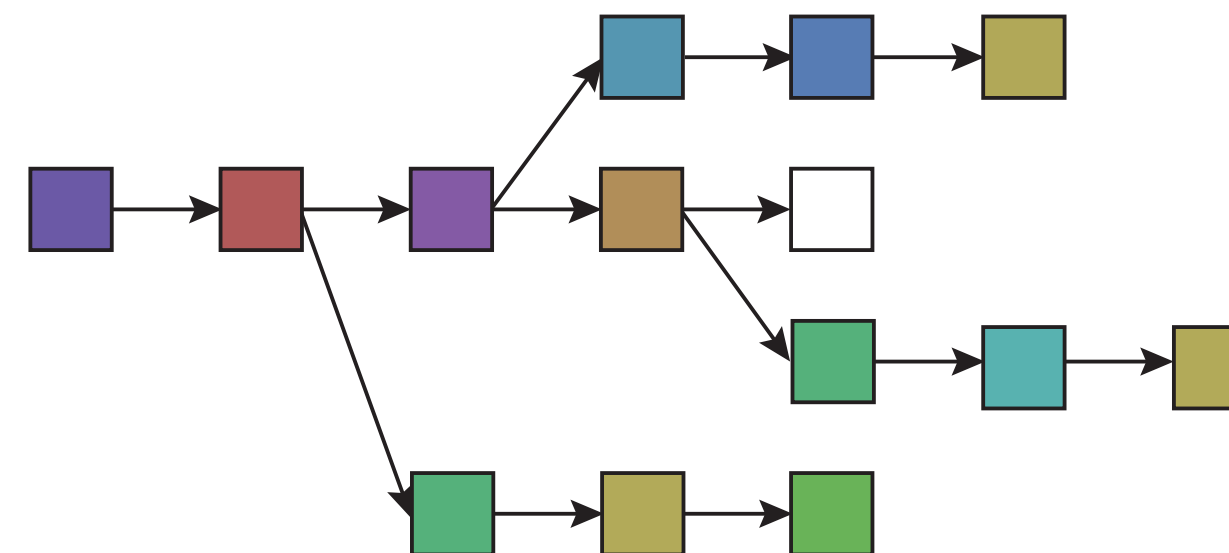
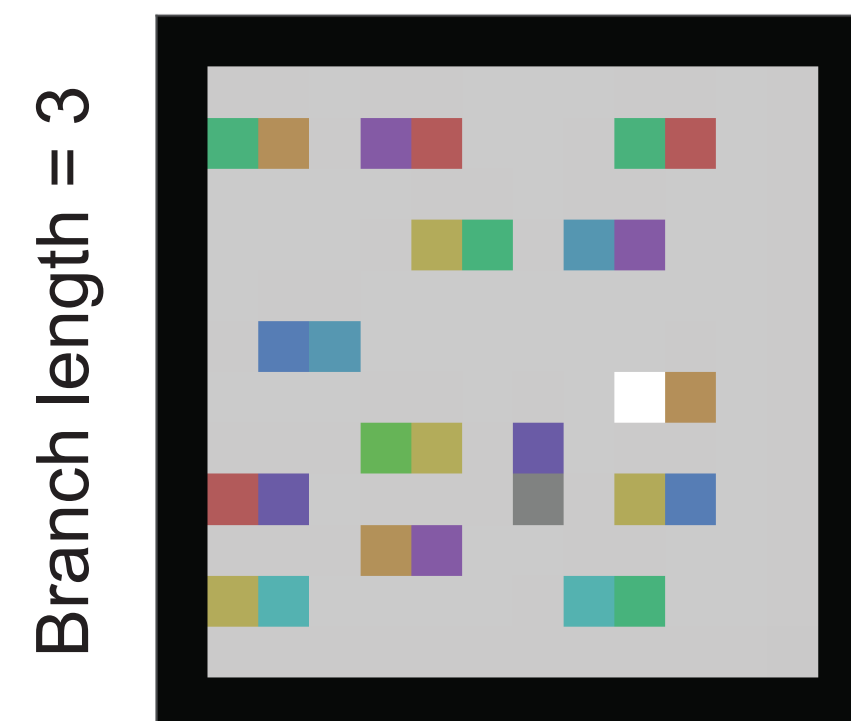


Relational deep reinforcement learning

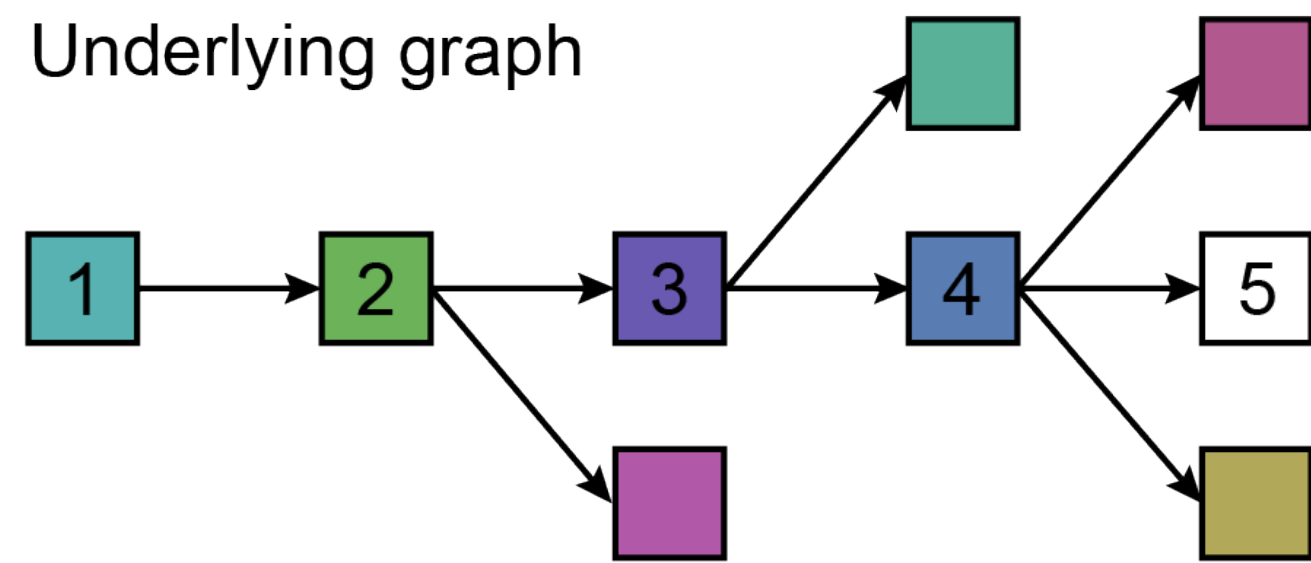


Box-World:

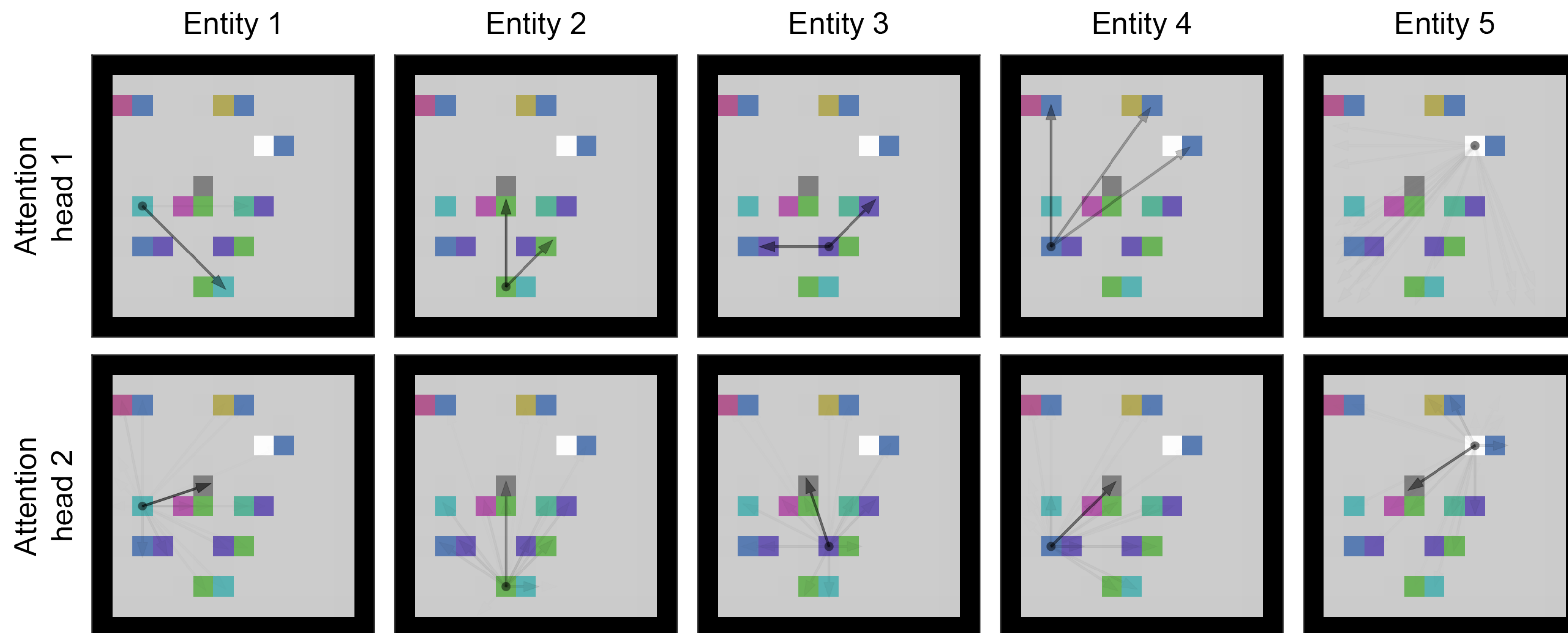
- Acquire gem (white) by opening a sequence of locked boxes
- Model-free (A2C) with self-attention/GN state representation, and message-passing



Visualizing the learned representations



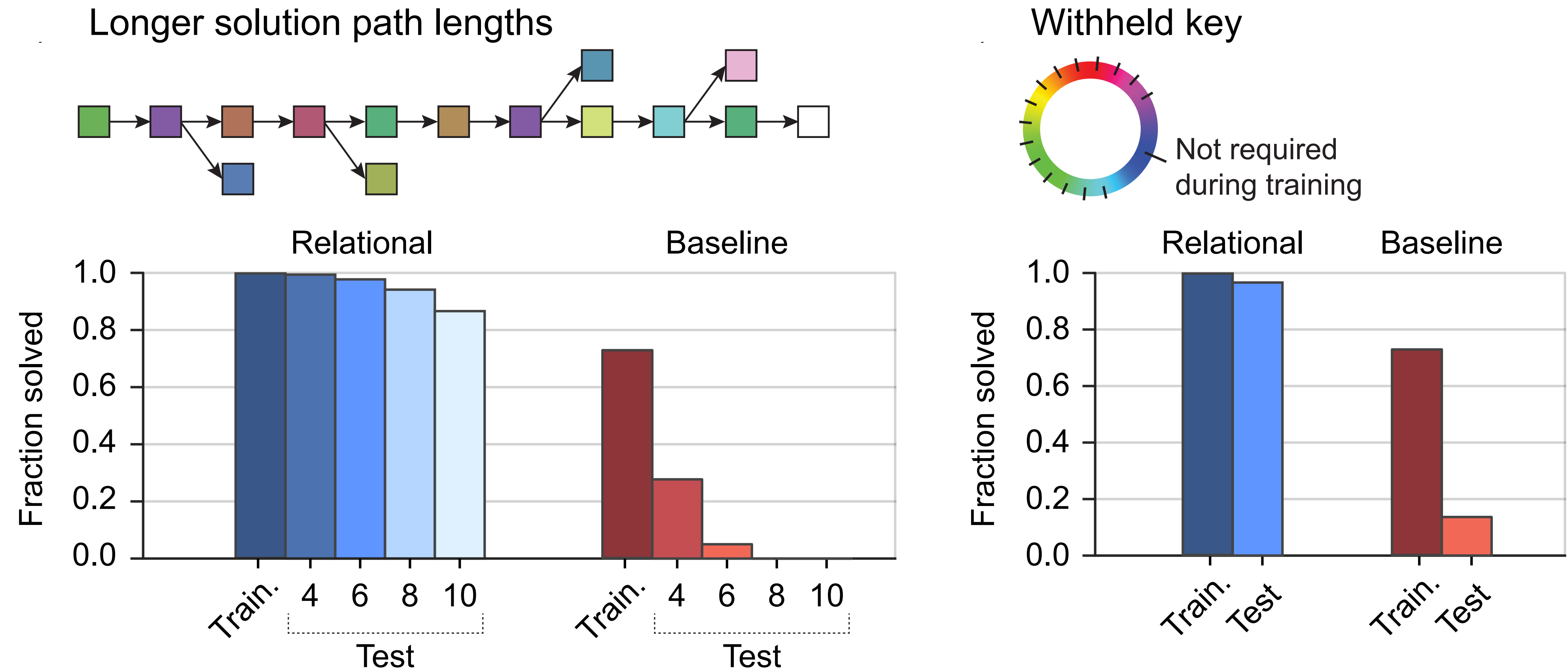
Highest attention values per head



1. Keys $\longleftrightarrow$ locks they can open

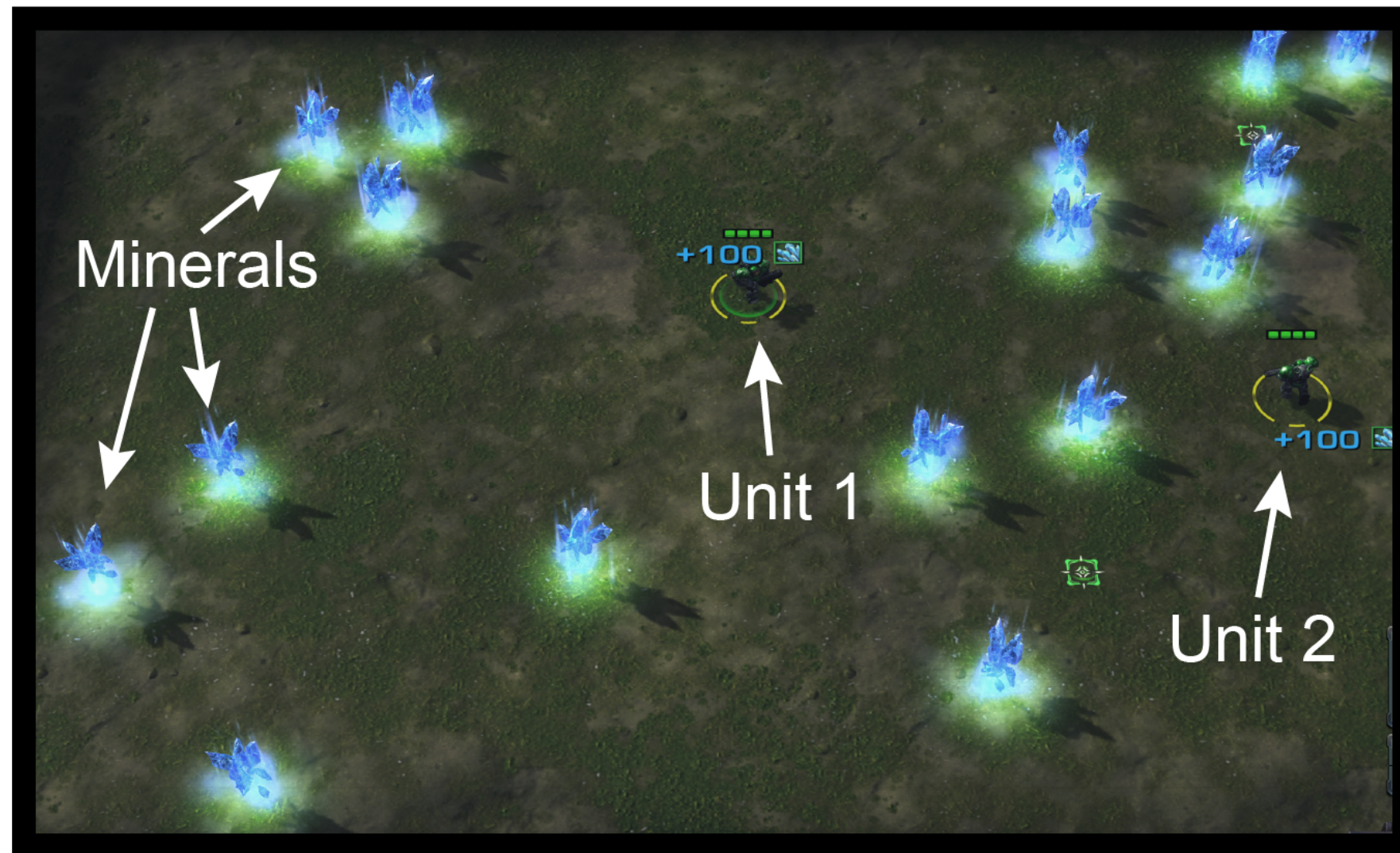
2. Avatar $\longleftrightarrow$ keys

Generalization to longer solutions and withheld keys

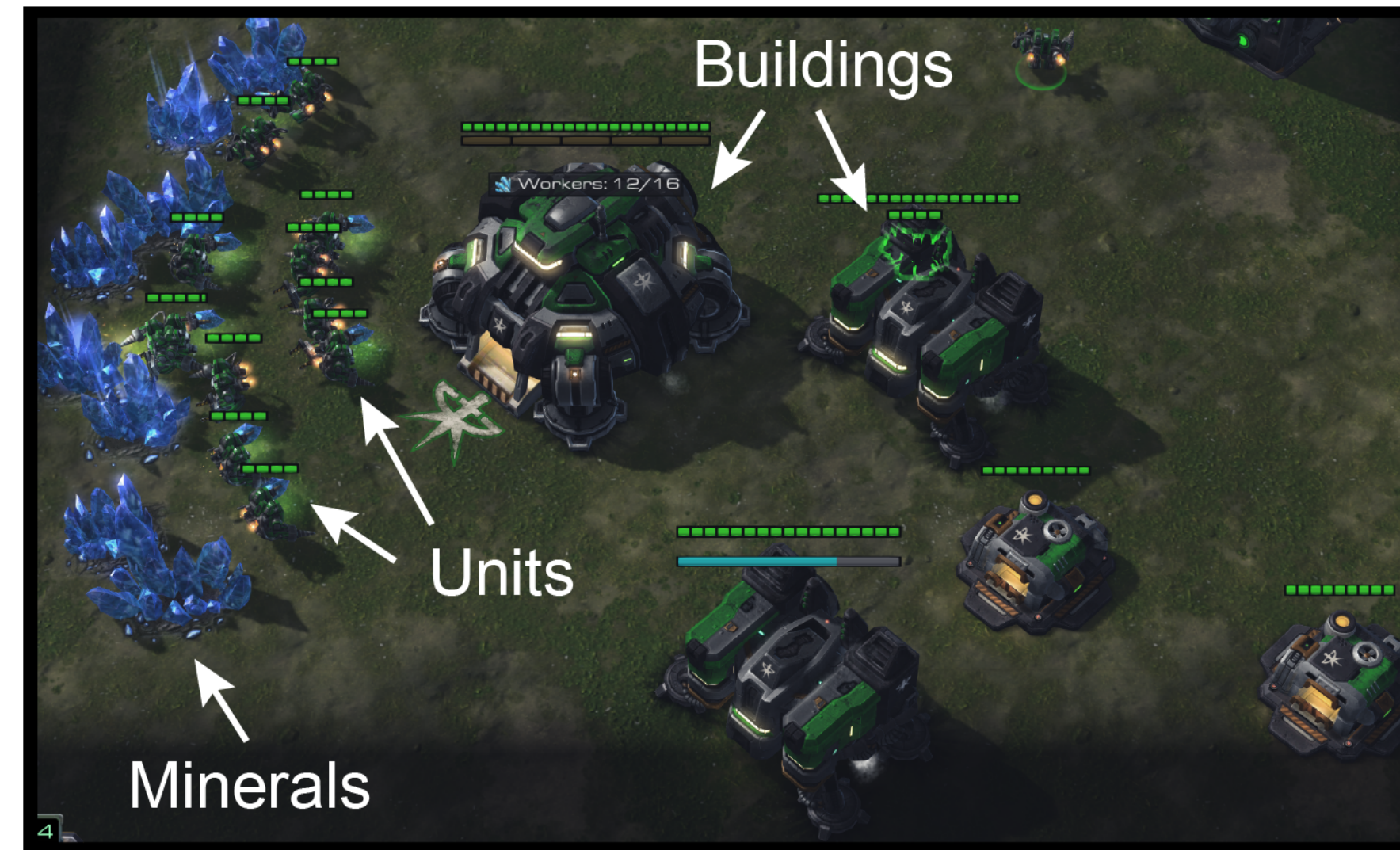


Scales up to StarCraft II mini-games

Collect Mineral Shards

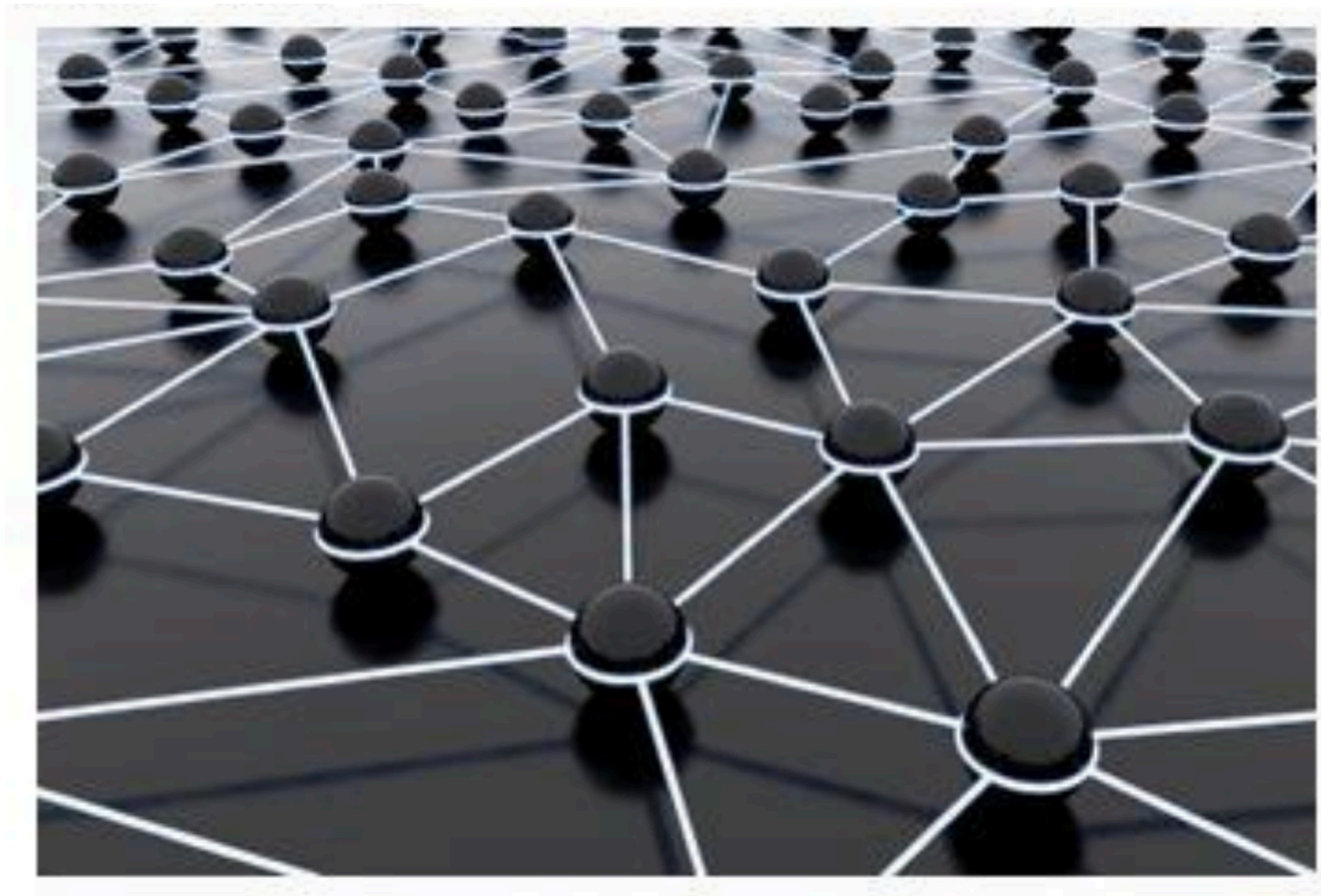


Build Marines



- State-of-the-art in 6 of 7 StarCraft II mini-games
- Beats grandmaster-level in 4 of 7

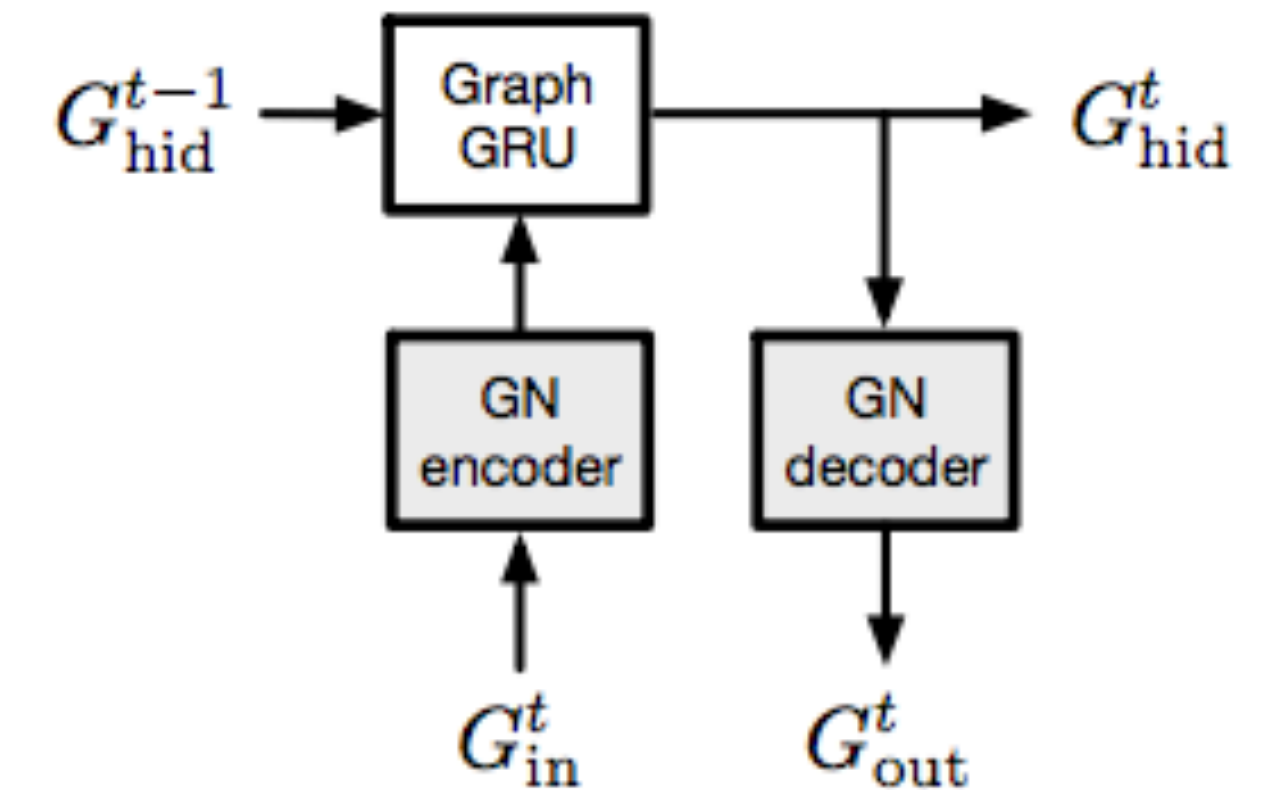
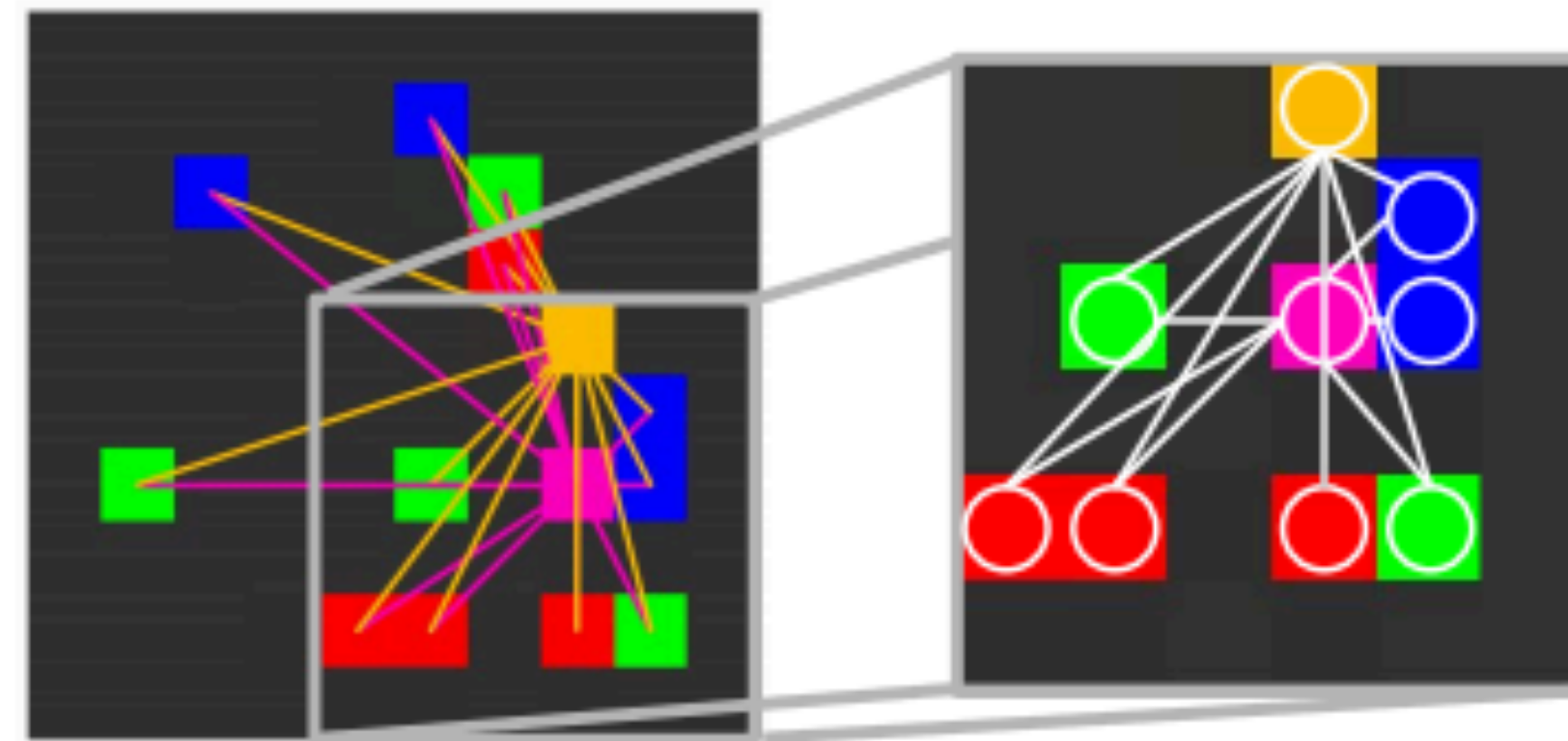
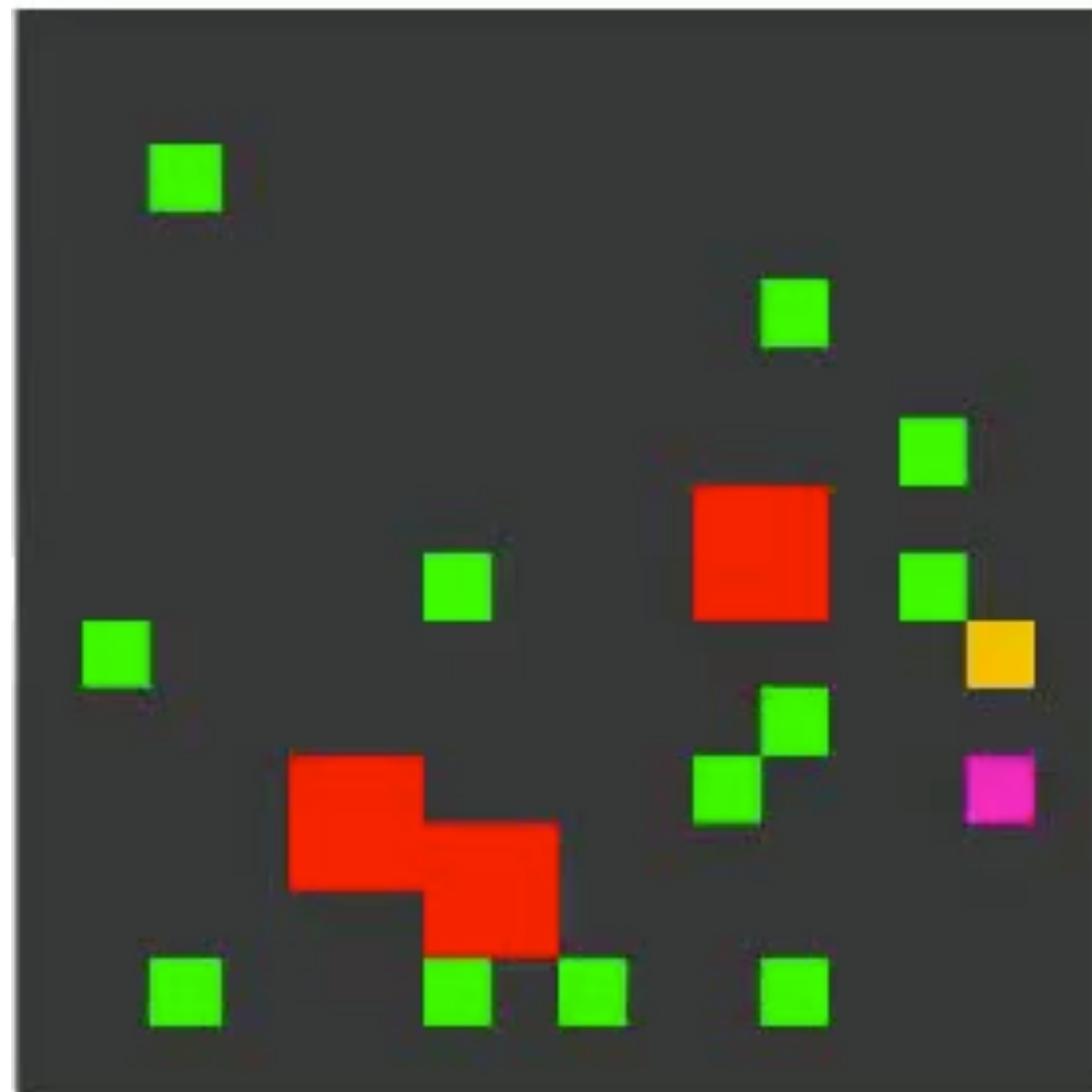
Structured models in multi-agent RL



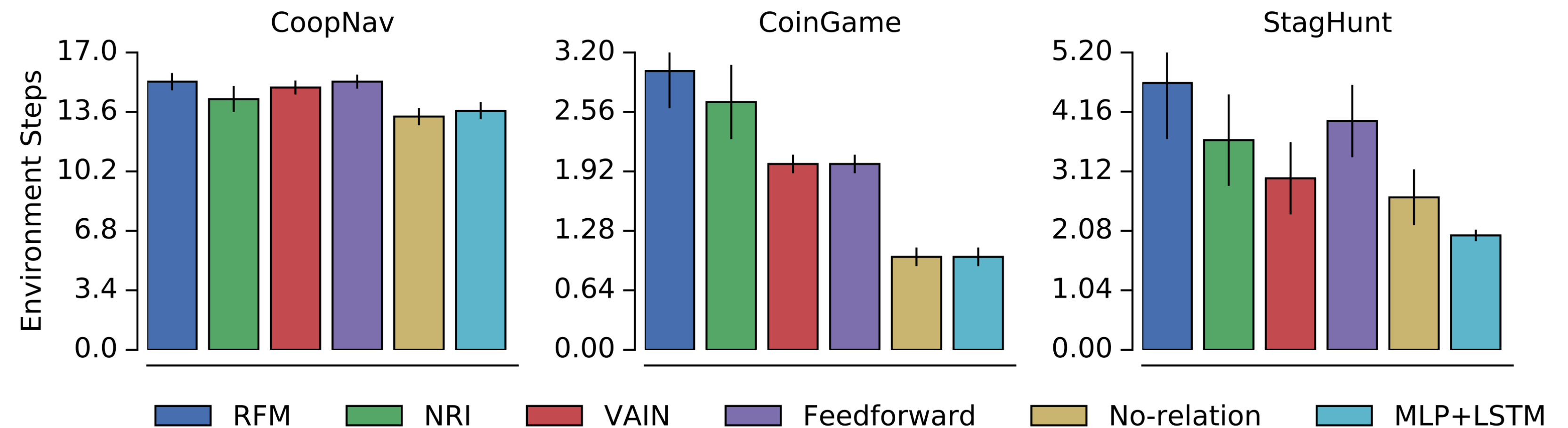
Relational forward models for multi-agent RL

Stag hunt

Step 0 with last actions



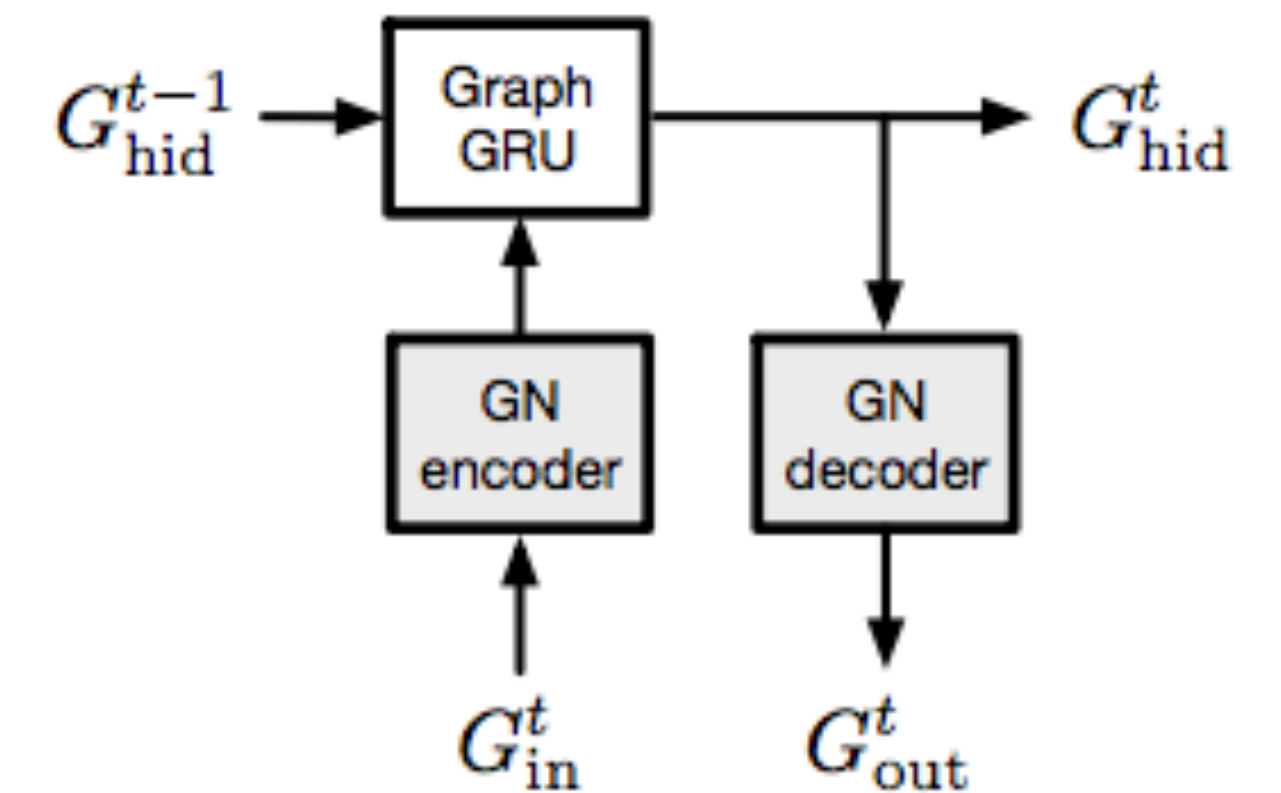
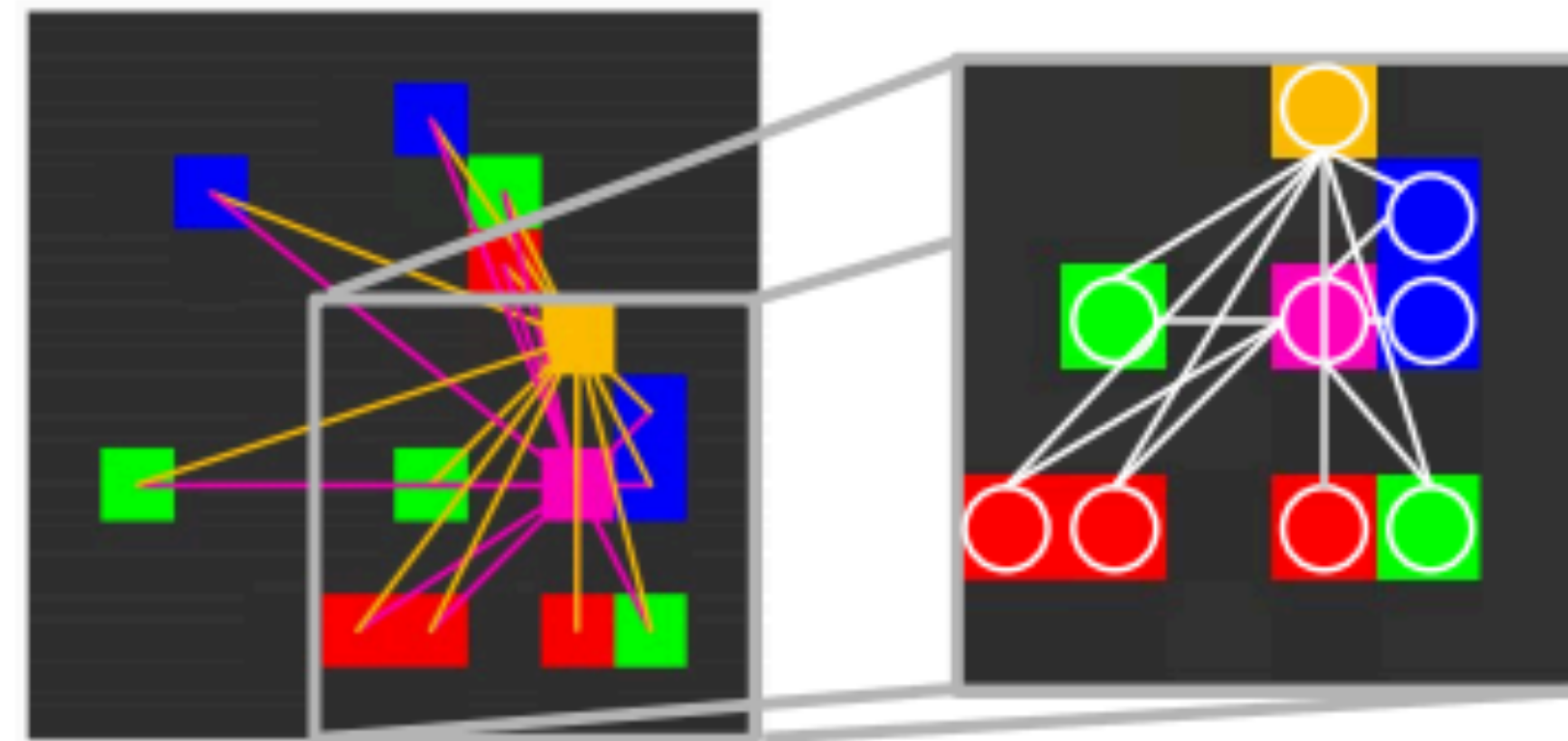
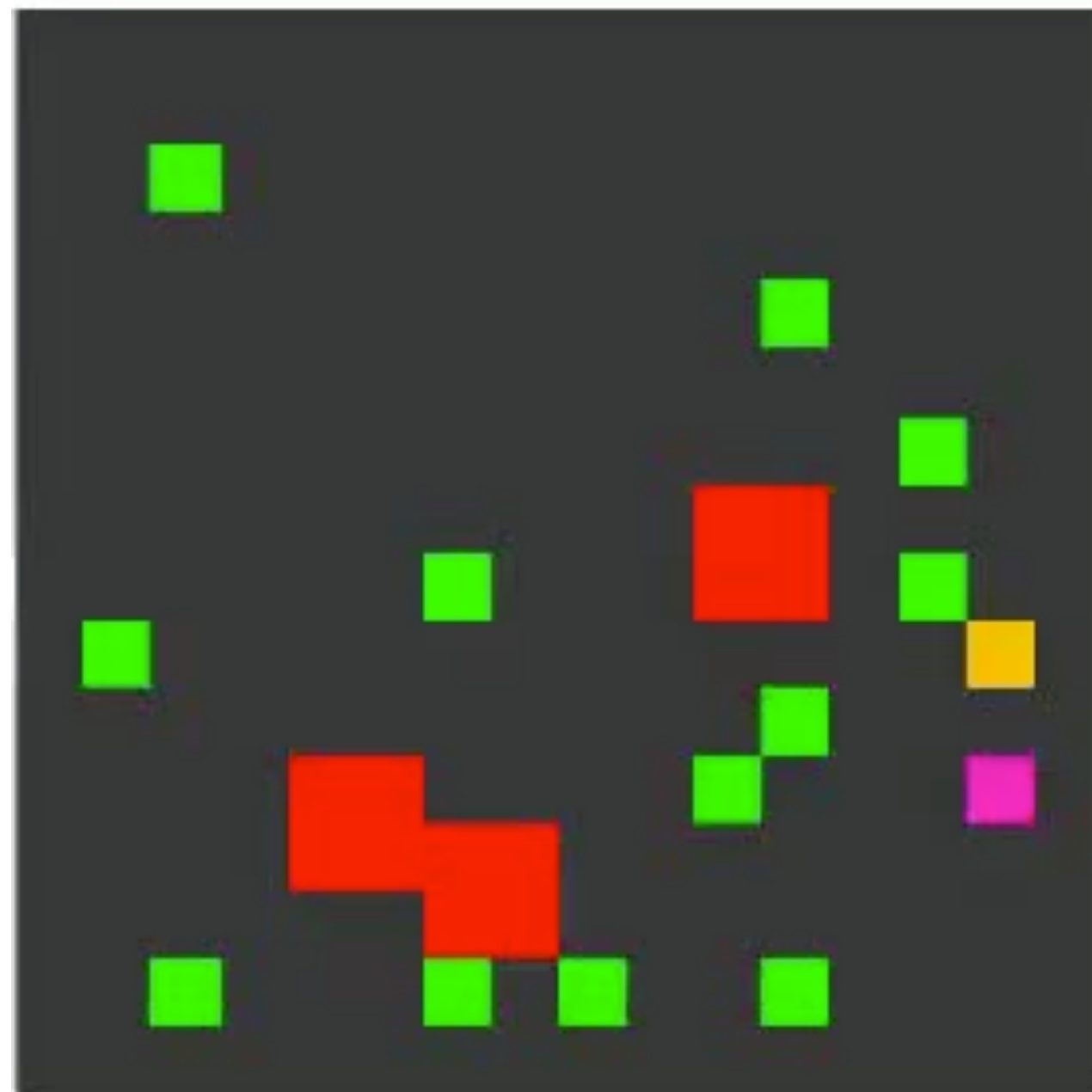
Forward prediction performance



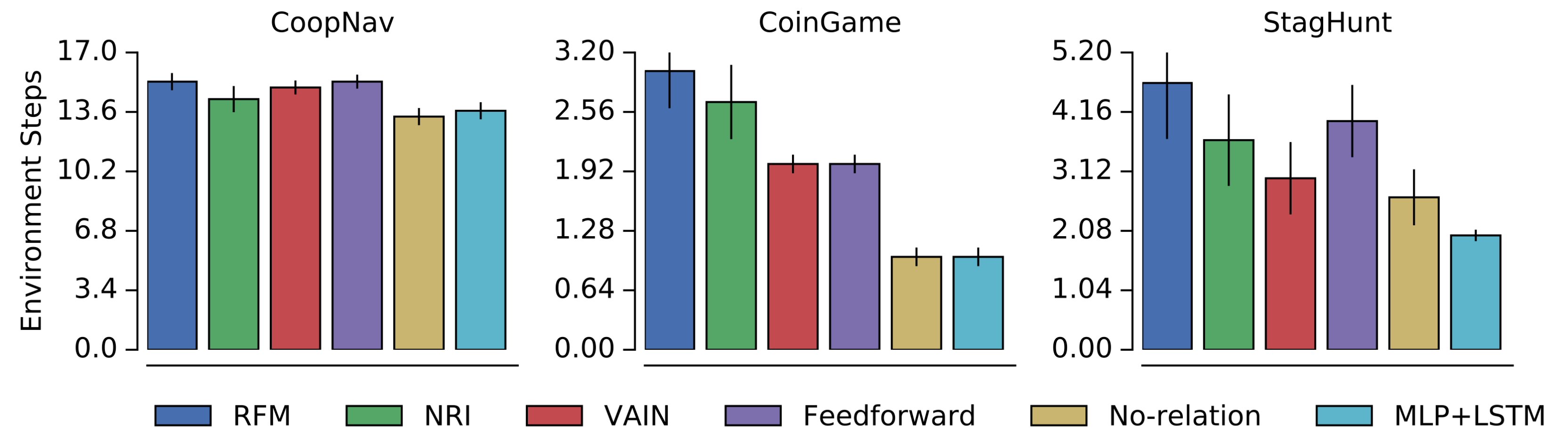
Relational forward models for multi-agent RL

Stag hunt

Step 0 with last actions

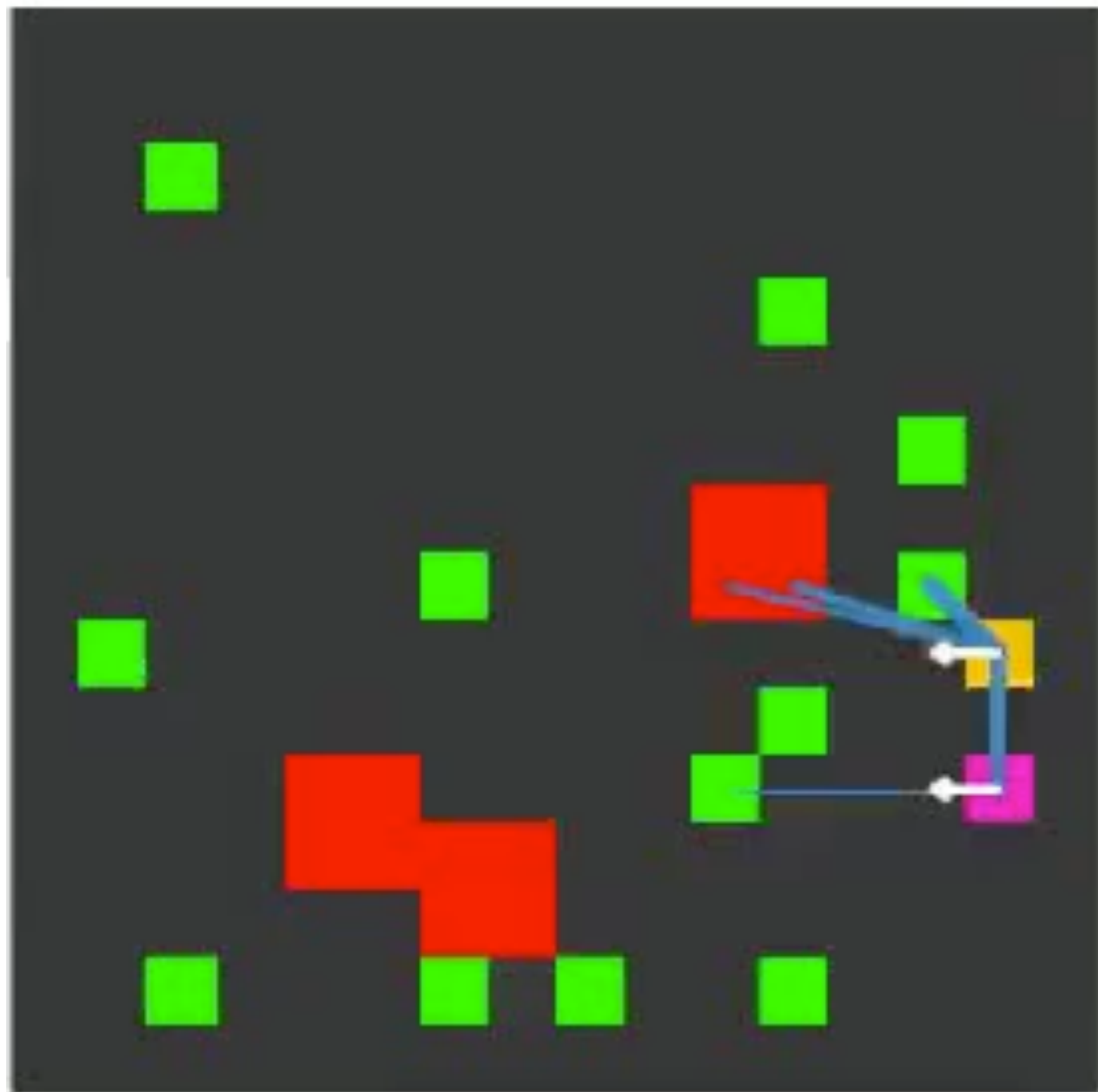


Forward prediction performance

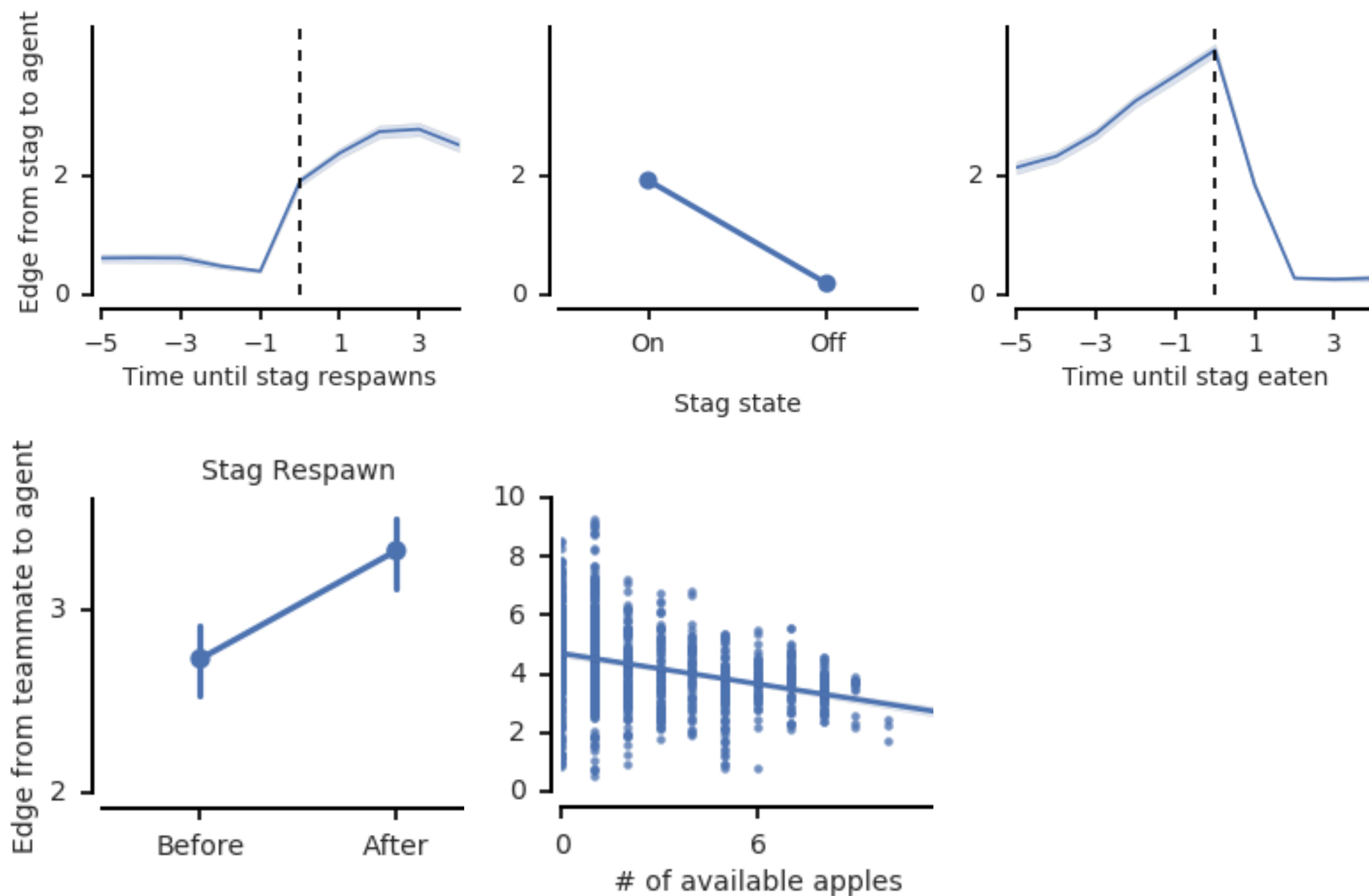


Interpretable learned representations

Predicted actions and top 5 edges

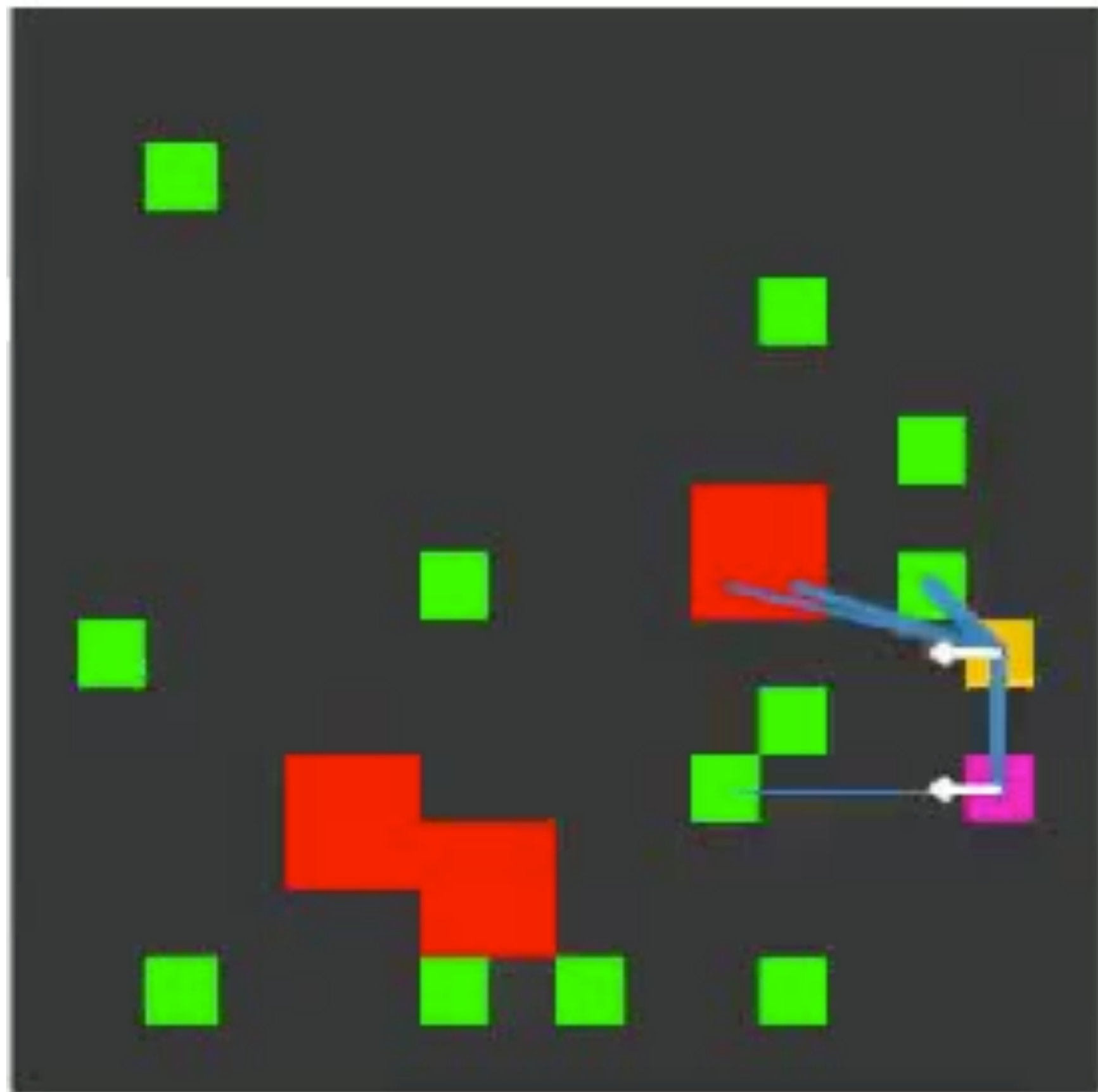


Magnitude of edge/“message” vector is meaningful

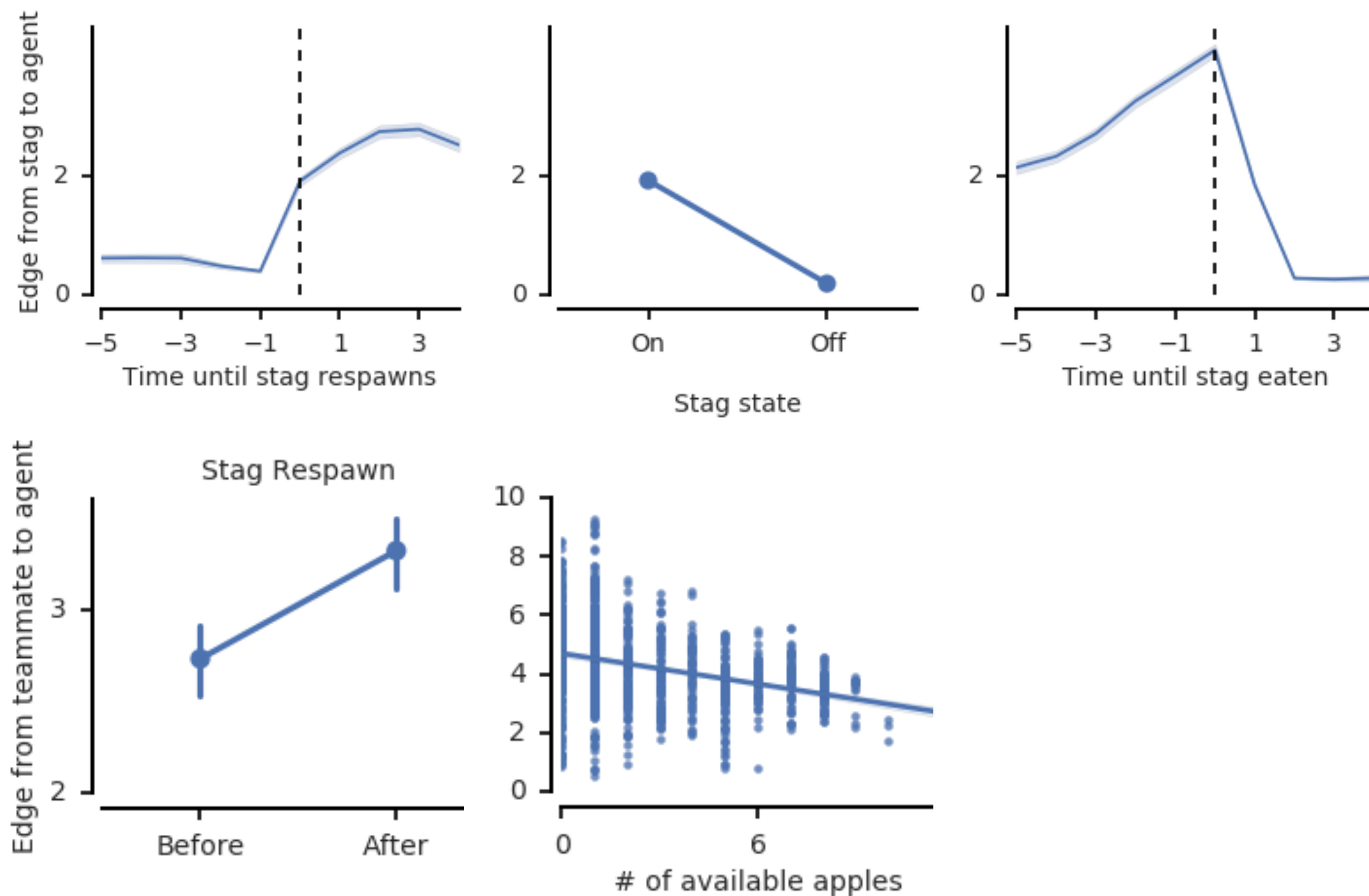


Interpretable learned representations

Predicted actions and top 5 edges



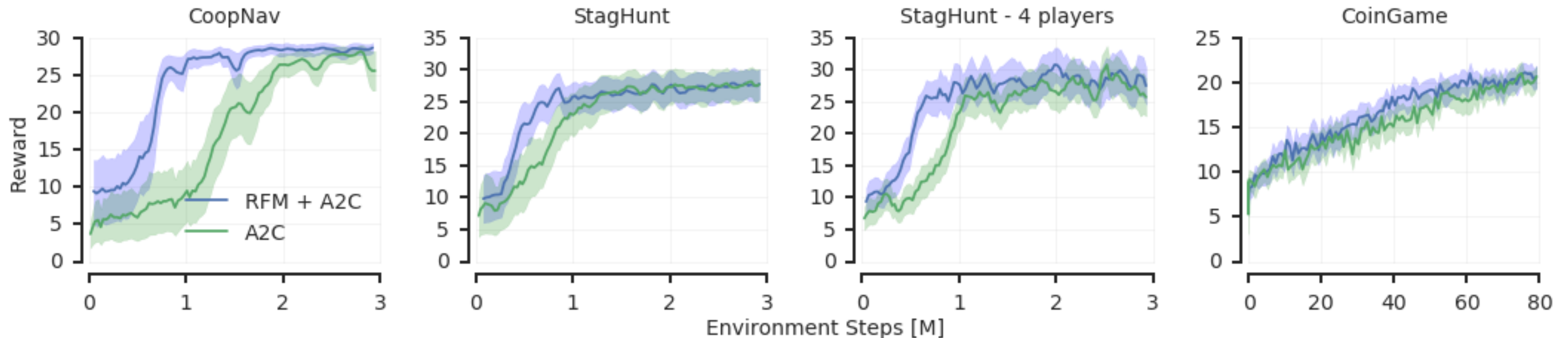
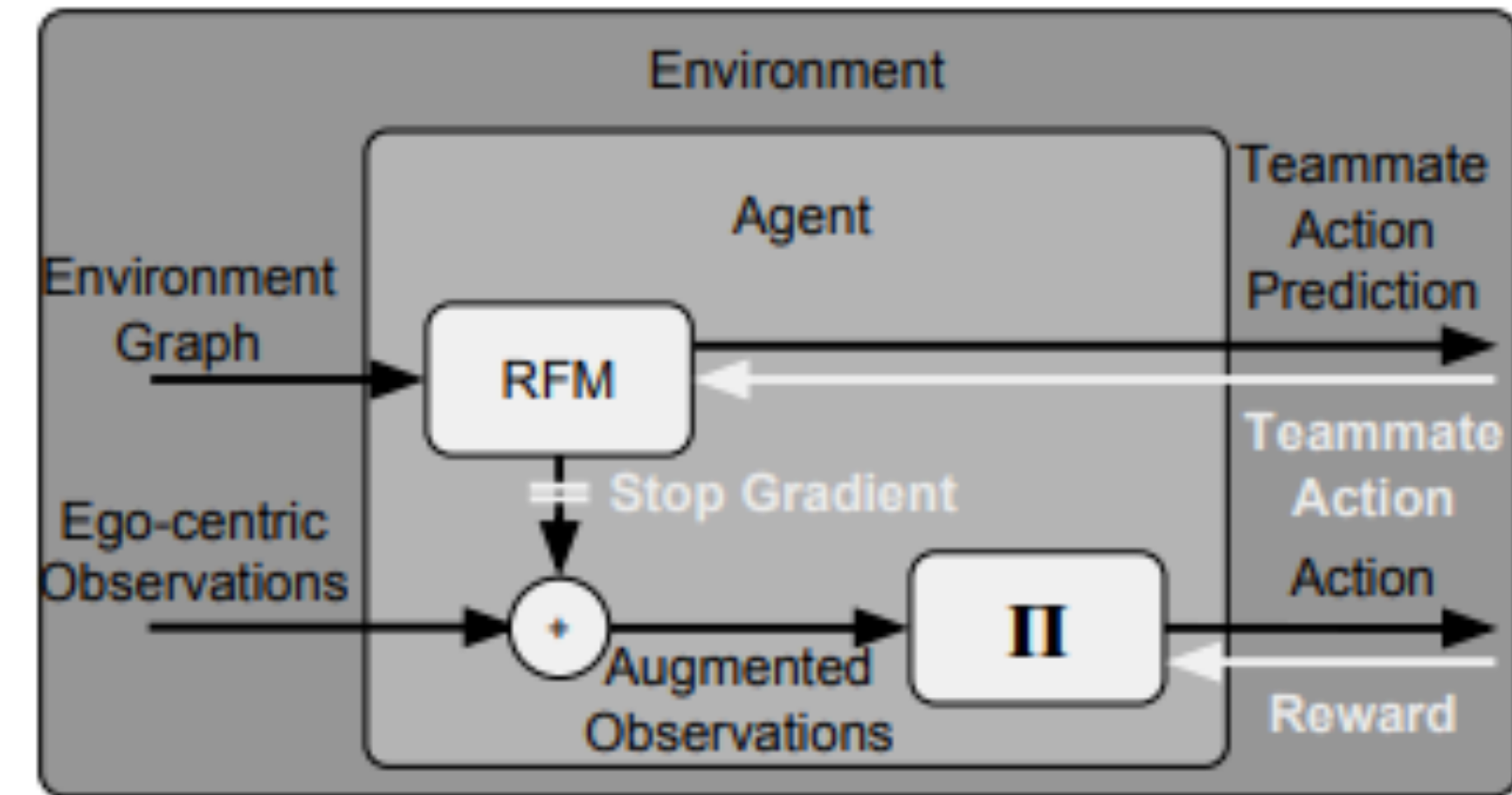
Magnitude of edge/“message” vector is meaningful



Agents learn faster with model-augmented observations

1. Train a set of agents to perform each game.
2. Train an RFM to predict future actions.
3. Train a new agent, whose observations are augmented with message magnitudes (crude “theory of mind”?).

The new agent (blue curve) trains faster in all environments.



Conclusions

Human use richly structured generative knowledge

- Combinatorial generalization: “Infinite use of finite means”
- Object- and relation-centric representations
- Structured mental simulation

Conclusions

Human use richly structured generative knowledge

- Combinatorial generalization: “Infinite use of finite means”
- Object- and relation-centric representations
- Structured mental simulation

Graph Networks: strong relational inductive bias

- Naturally support combinatorial generalization via compositional sharing
- Graph-structured representations and policies
- Open-source library: github.com/deepmind/graph_nets (with demos, including physics!)

Reject false choices

- Nature **vs** Nurture
- Structure **vs** Flexibility
- Symbolic **vs** Connectionist
- Hand-engineered **vs** End-to-end

Reject false choices

- Nature ***and*** Nurture
- Structure ***and*** Flexibility
- Symbolic ***and*** Connectionist
- Hand-engineered ***and*** End-to-end

The “bias versus variance trade-off” is real—
however the emphasis shouldn’t be on “***versus***”, but rather on “***trade-off***”.

Reject false choices

- Nature *and* Nurture
- Structure *and* Flexibility
- Symbolic *and* Connectionist
- Hand-engineered *and* End-to-end

The “bias versus variance trade-off” is real—
however the emphasis shouldn’t be on “**versus**”, but rather on “**trade-off**”.

Biology doesn’t choose between nature **versus** nurture. It uses nature and nurture **jointly**, to build wholes which are greater than the sums of their parts.

There’s great promise in synthesizing new techniques by drawing on the full AI toolkit and marrying the best approaches from today with those which were essential during times when data and computation were at a premium.

Key collaborators

Jess Hamrick
Josh Tenenbaum
Chris Bates
Razvan Pascanu
Nick Watters
Daniel Zoran
Theo Weber
Andy Ballard
Nicholas Heess
Yujia Li
Oriol Vinyals

Kelsey Allen
Alvaro Sanchez
Victor Bapst
Vinicius Zambaldi
David Raposo
Adam Santoro
Mateusz Malinowski
Andrea Tacchetti
Francis Song



Key collaborators

Jess Hamrick
Josh Tenenbaum
Chris Bates
Razvan Pascanu
Nick Watters
Daniel Zoran
Theo Weber
Andy Ballard
Nicholas Heess
Yujia Li
Oriol Vinyals

Kelsey Allen
Alvaro Sanchez
Victor Bapst
Vinicius Zambaldi
David Raposo
Adam Santoro
Mateusz Malinowski
Andrea Tacchetti
Francis Song



References

[Battaglia et al., 2013, PNAS](#)
[Hamrick et al., 2016, Cognition](#)
[Bates et al., 2015, Proc Cog Sci](#)
[Battaglia et al., 2016, NeurIPS](#)
[Watters et al., 2017, NeurIPS](#)
[Raposo et al., 2017, ICLR workshop](#)
[Santoro et al., 2017, NeurIPS](#)
[Sanchez-Gonzalez et al., 2018, ICML](#)
[Hamrick et al., 2017, ICLR](#)
[Pascanu et al., 2017, arXiv](#)
[Hamrick et al., 2018, Proc Cog Sci](#)
[Zambaldi et al., 2018, arXiv/under review](#)
[Tacchetti et al., 2018, arXiv/under review](#)
[Battaglia et al. 2018 arXiv](#)