Towards Practical Multi-object Manipulation using Relational Reinforcement Learning

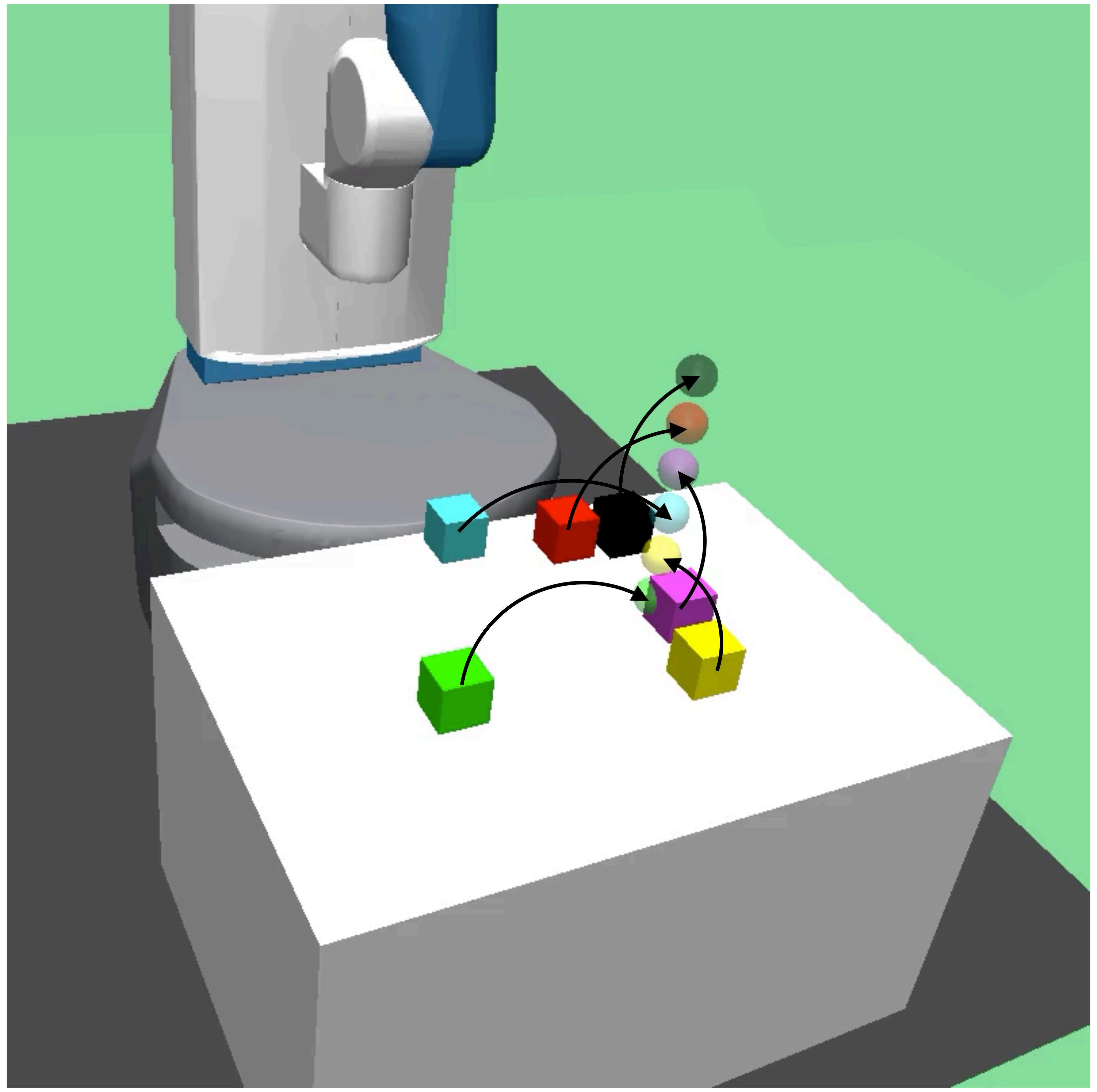
Richard Li, Allan Jabri, Trevor Darrell, Pulkit Agrawal

Massachusetts Institute of Technology

Berkeley UNIVERSITY OF CALIFORNIA

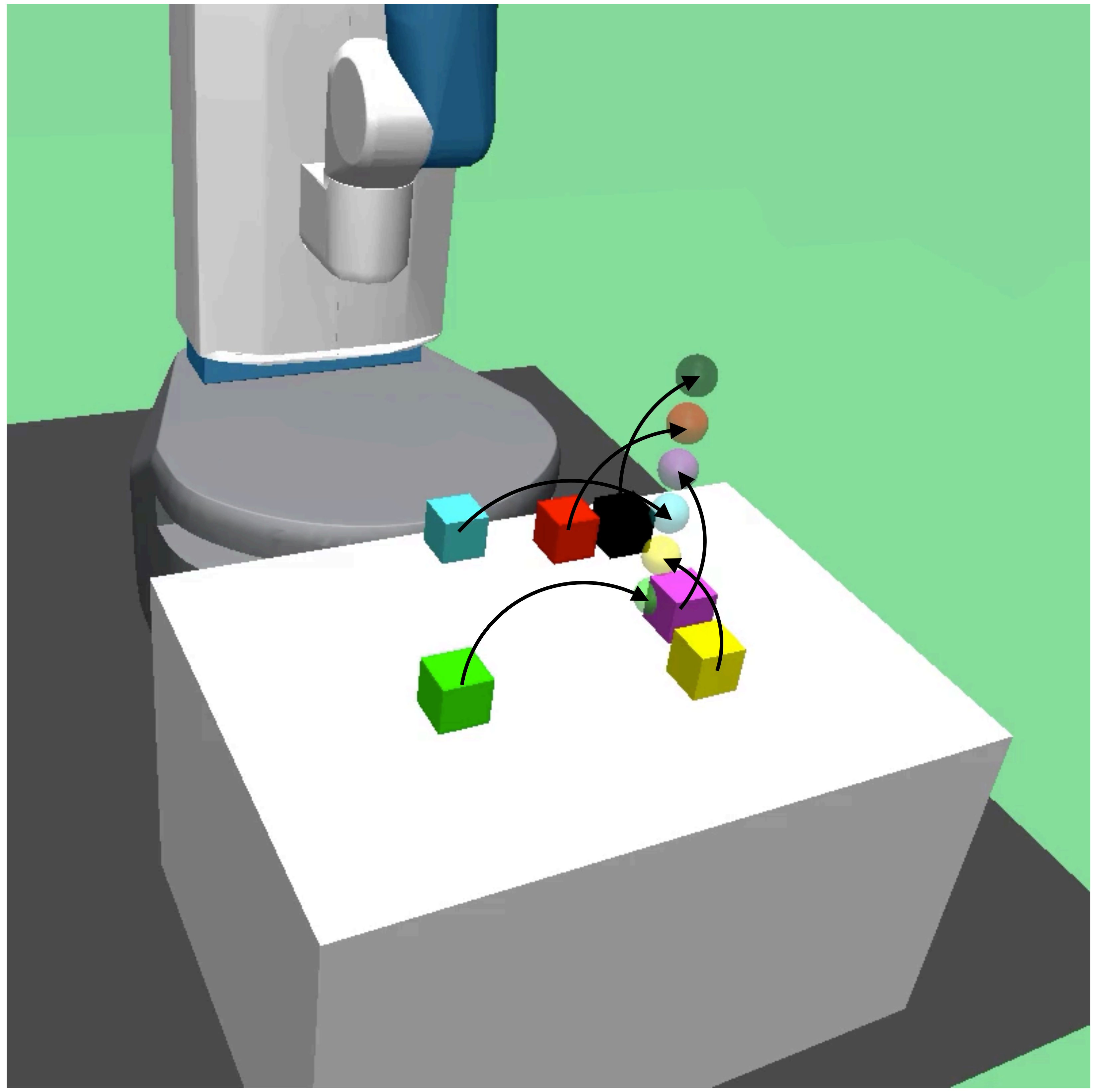
ICRA 2020

Block stacking task in reinforcement learning



+1 reward per block close to goal

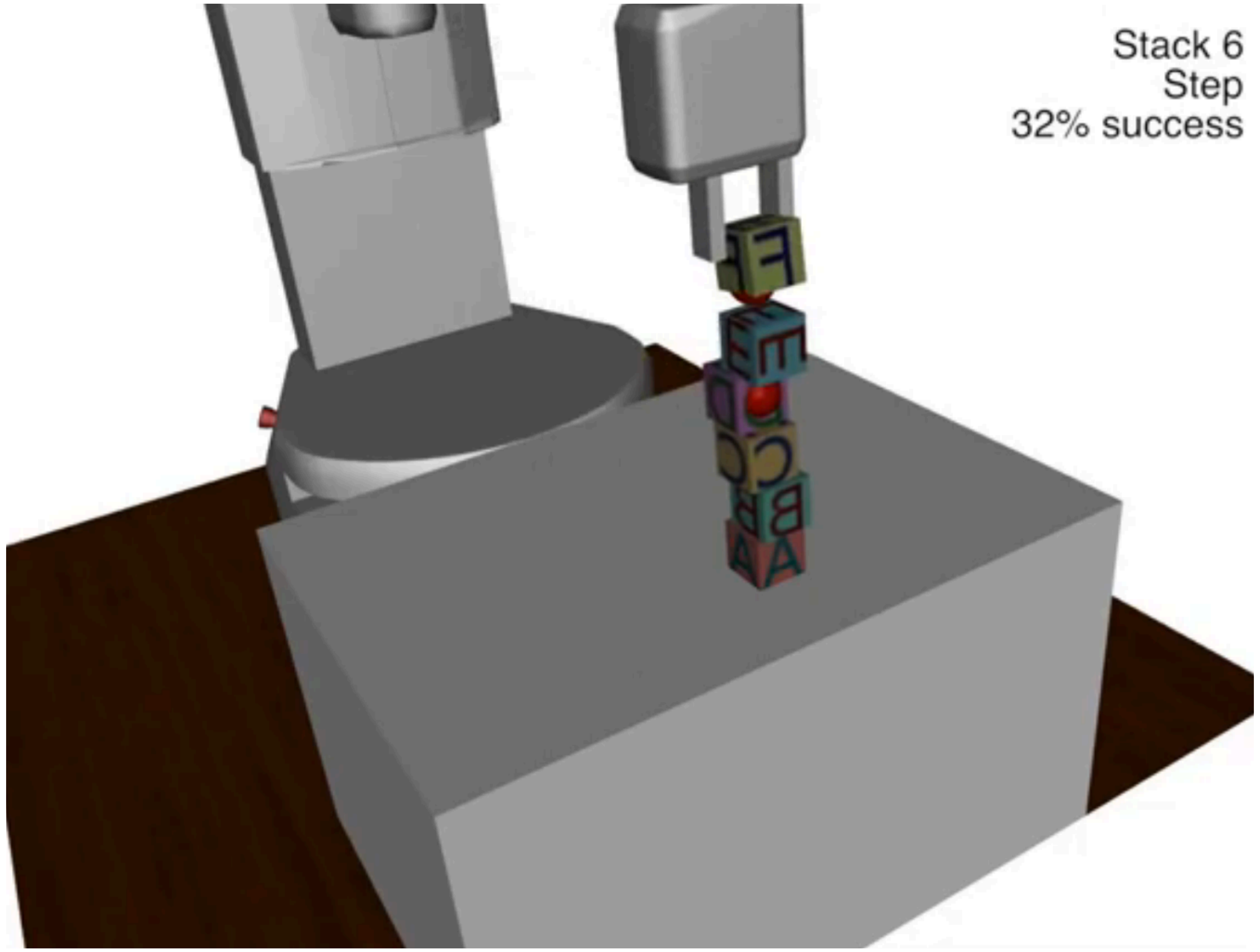
Block stacking task in reinforcement learning



- Difficult exploration
  - Sparse reward
  - Long horizon
  - Multiple objects

+1 reward per block close to goal

Demonstration-guided RL



Nair '18

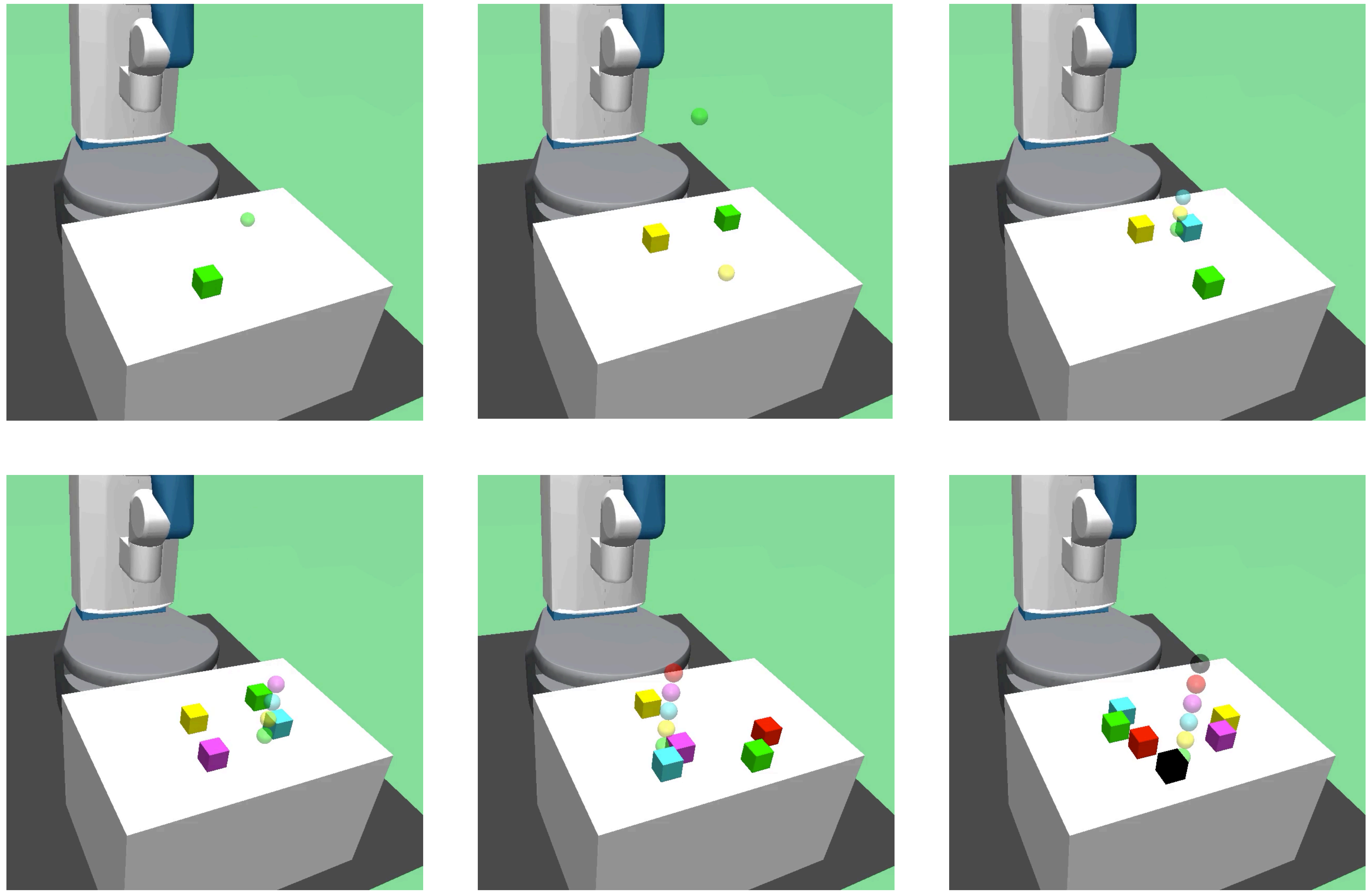
- Behavioral cloning loss

Comparison of results

Stacking 6 Blocks

	Nair ’18 (SOTA)	Ours
Success rate	32%	**75% (↑43%)**
Environment steps	2300M	**30M (↓77x)**

Curriculum over object cardinality

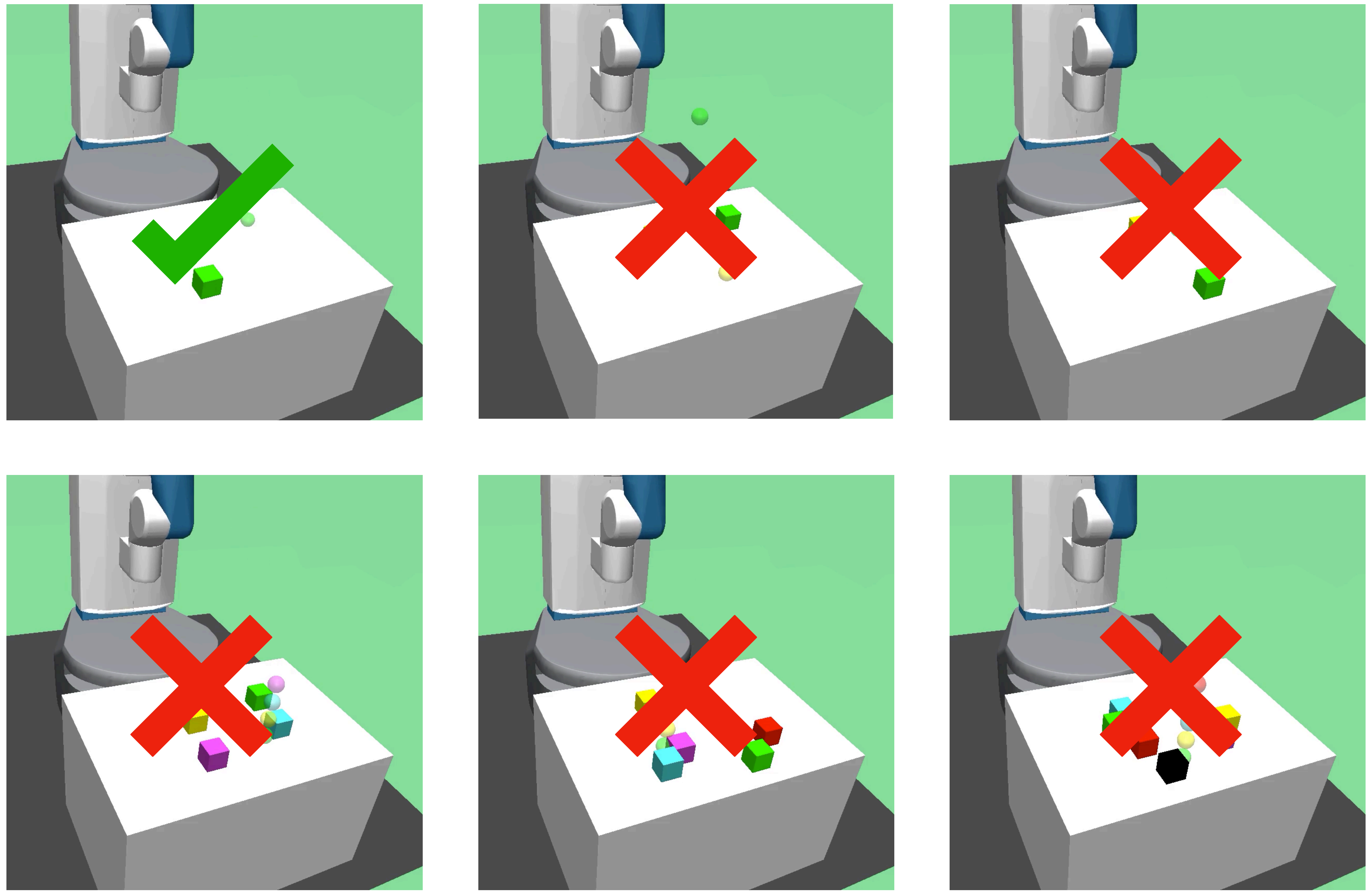


- Start with 1 block, sequentially transfer to harder tasks
- Soft Actor Critic + Hindsight Experience Replay

NerveNet (Wang ’18)

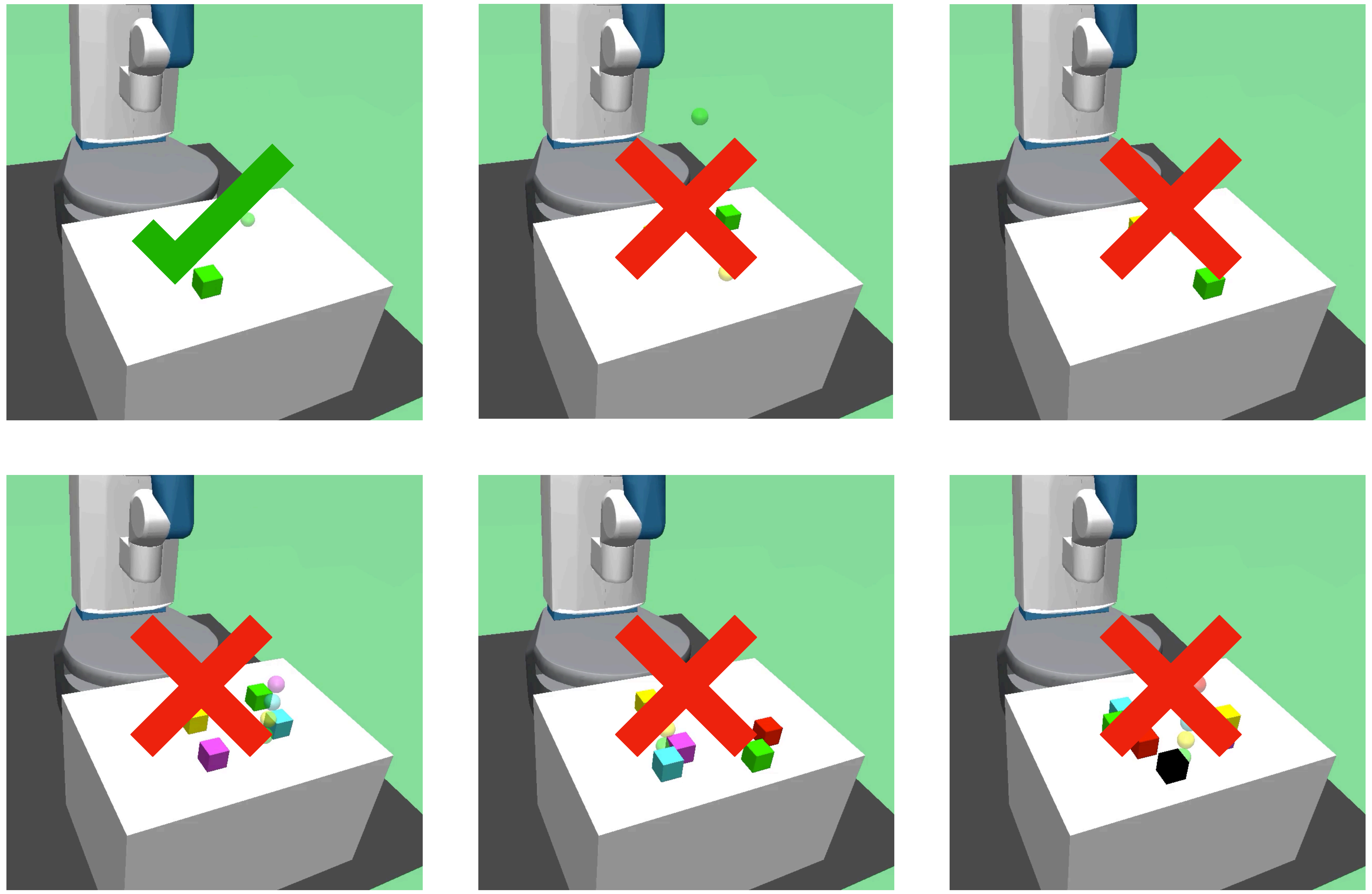
Self-Assembling Morphologies (Pathak ’19)

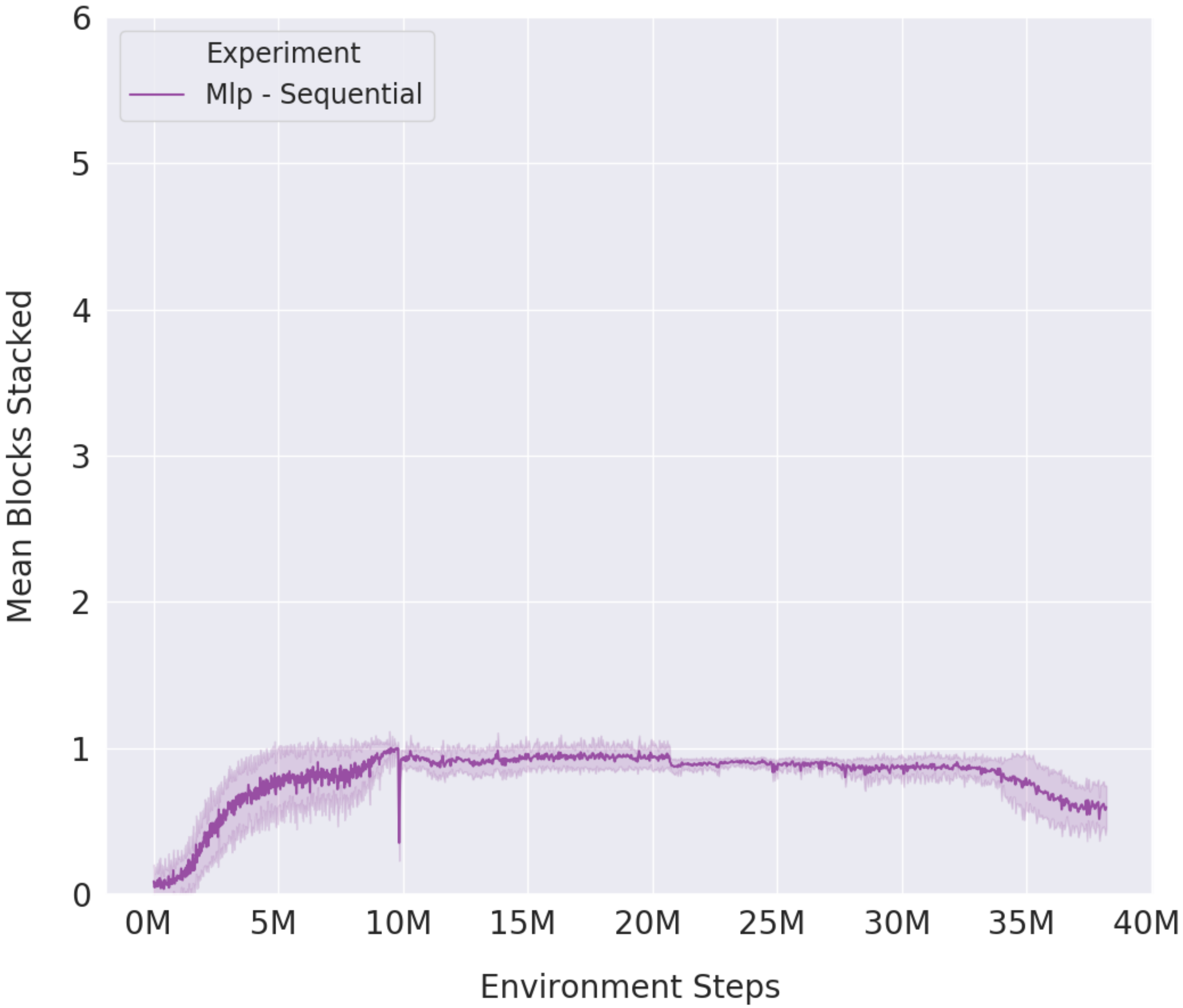
Curriculum with Multi-layer Perceptron (MLP)



- Start with 1 block, sequentially transfer to harder tasks
- Soft Actor Critic + Hindsight Experience Replay

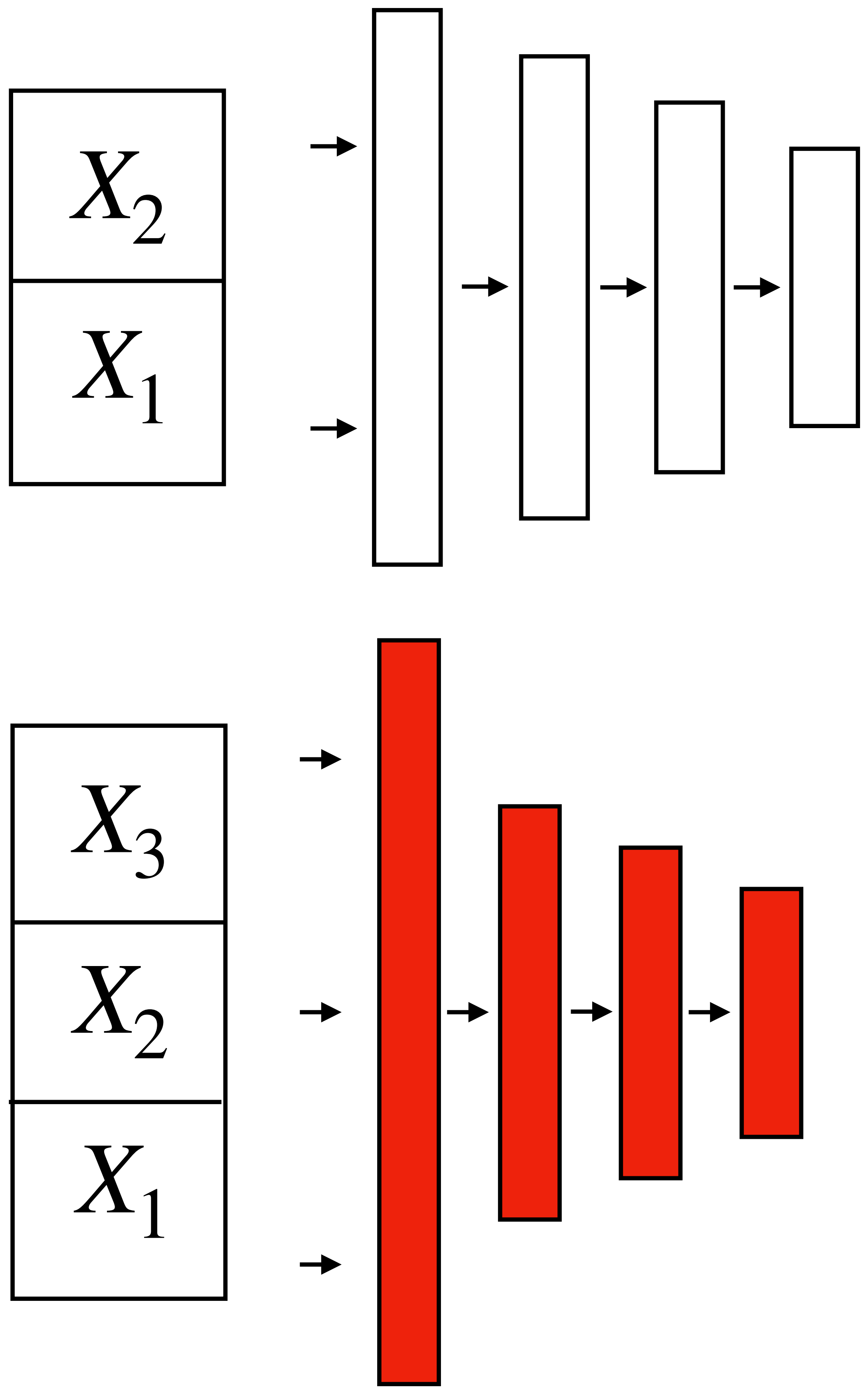
Curriculum with Multi-layer Perceptron (MLP)



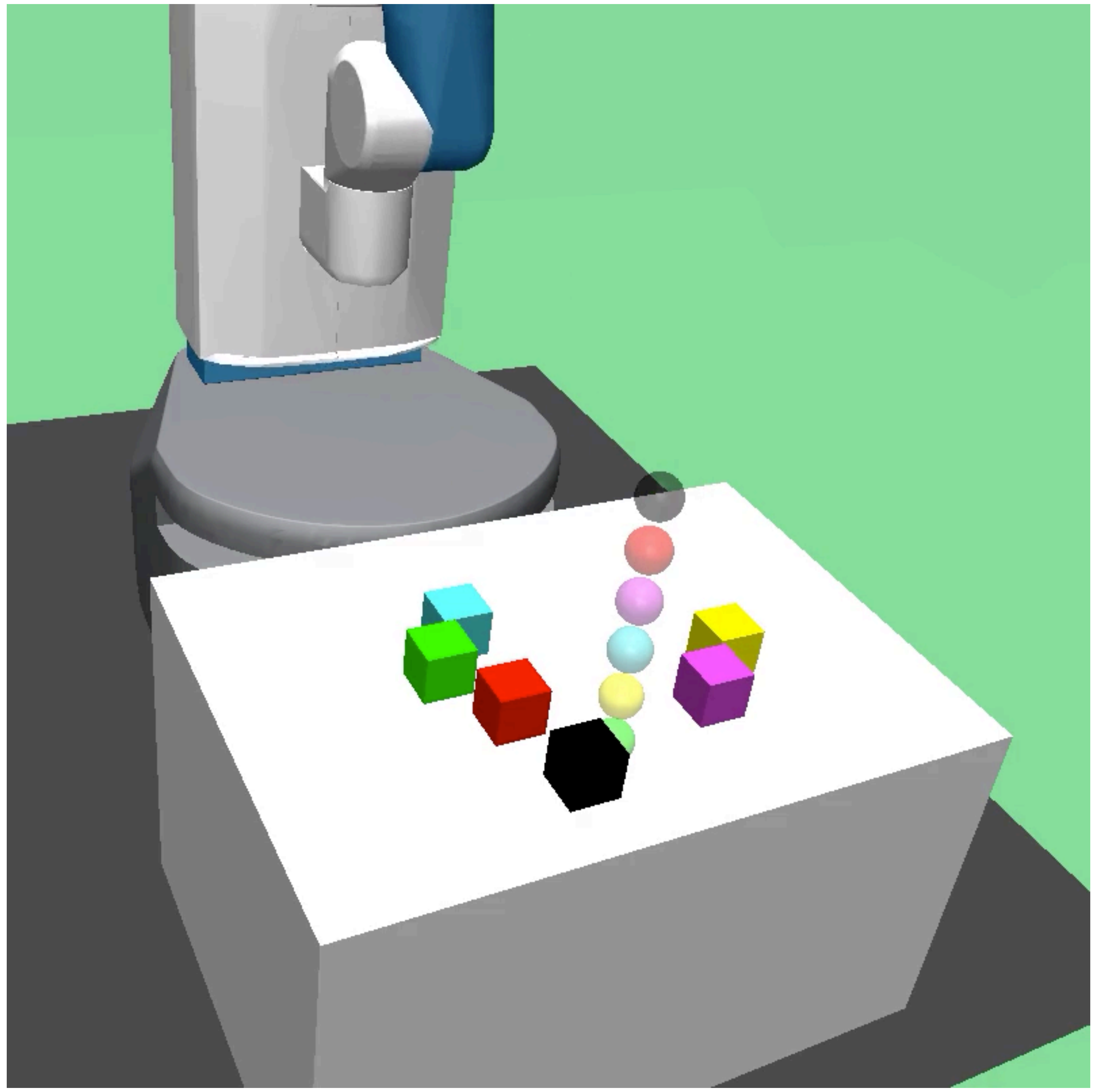


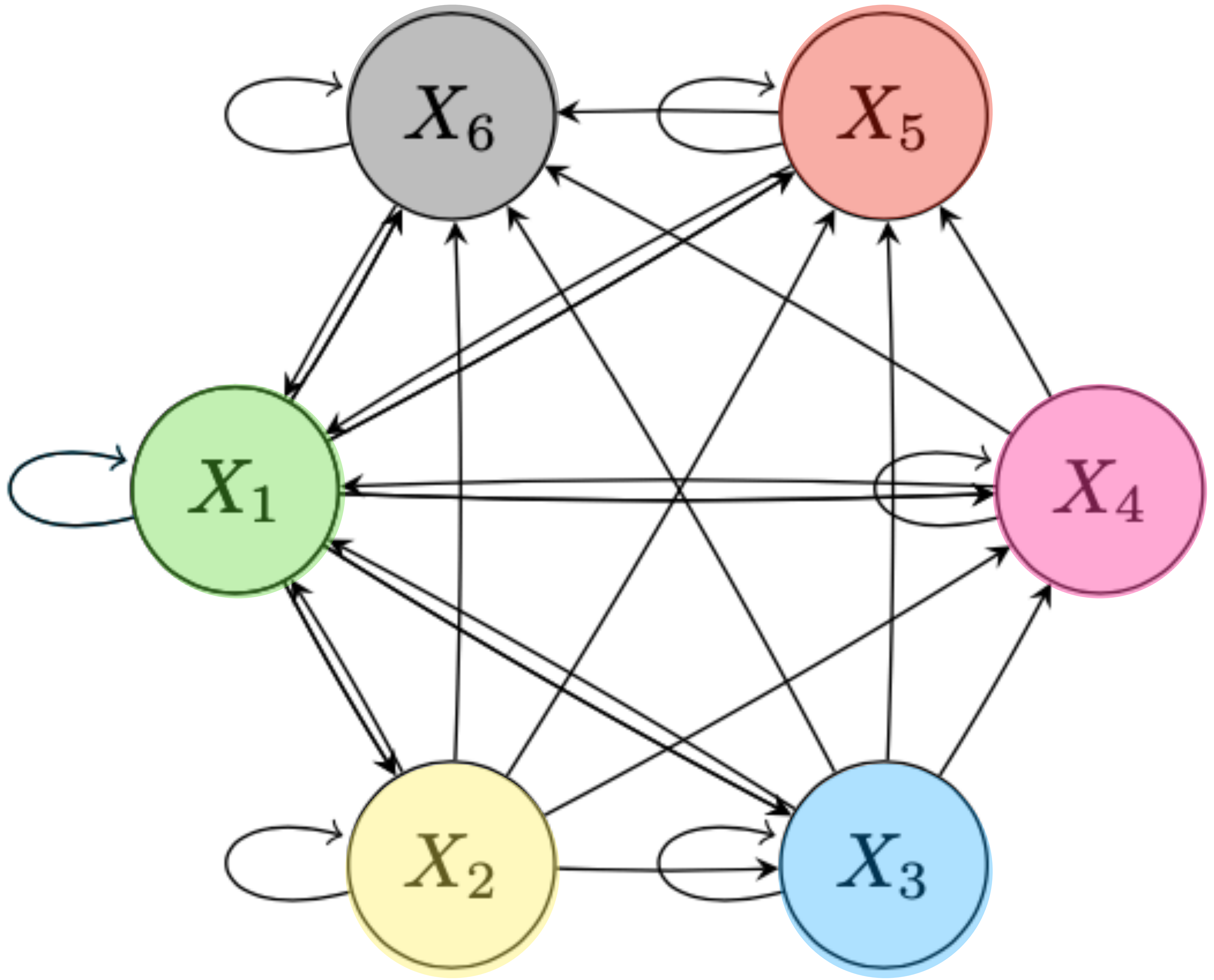
Problem: Discrete domain shift

2 blocks → 3 blocks



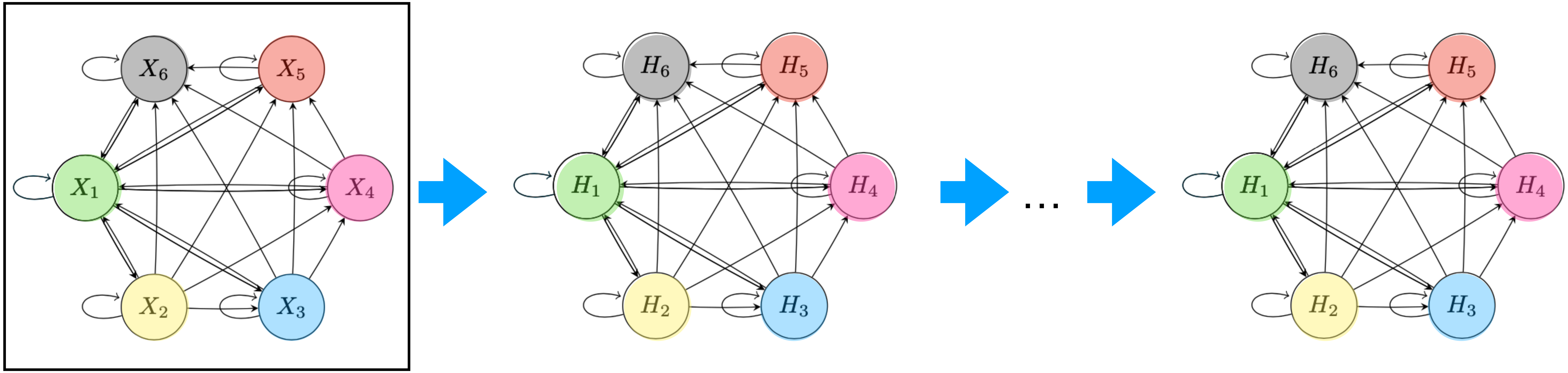
Swap MLP for Graph Neural Network (GNN)





Interaction Networks (Battaglia '16)

Attention-based GNN



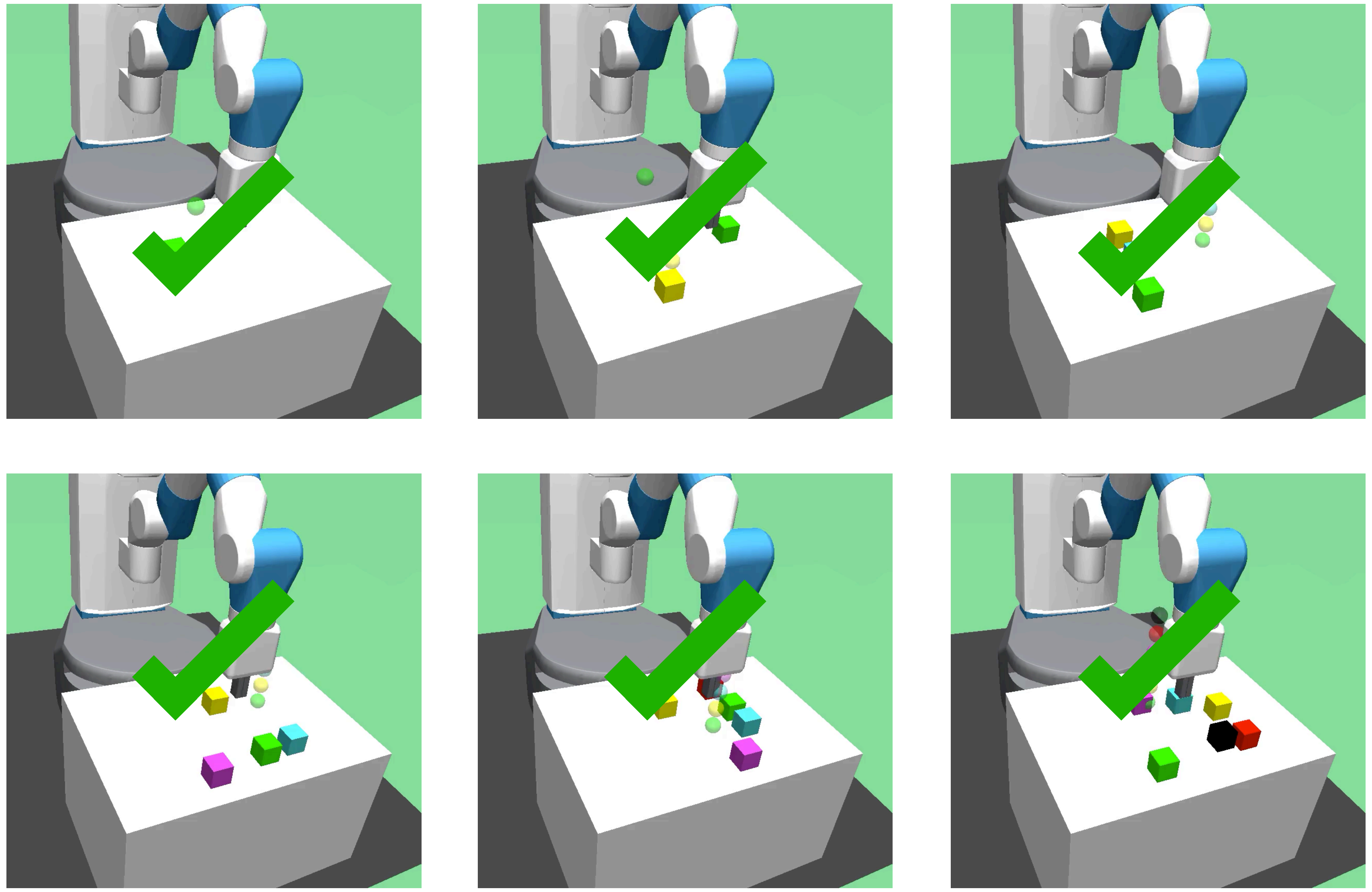
q_i: query
k_i: key
m_i: message / value

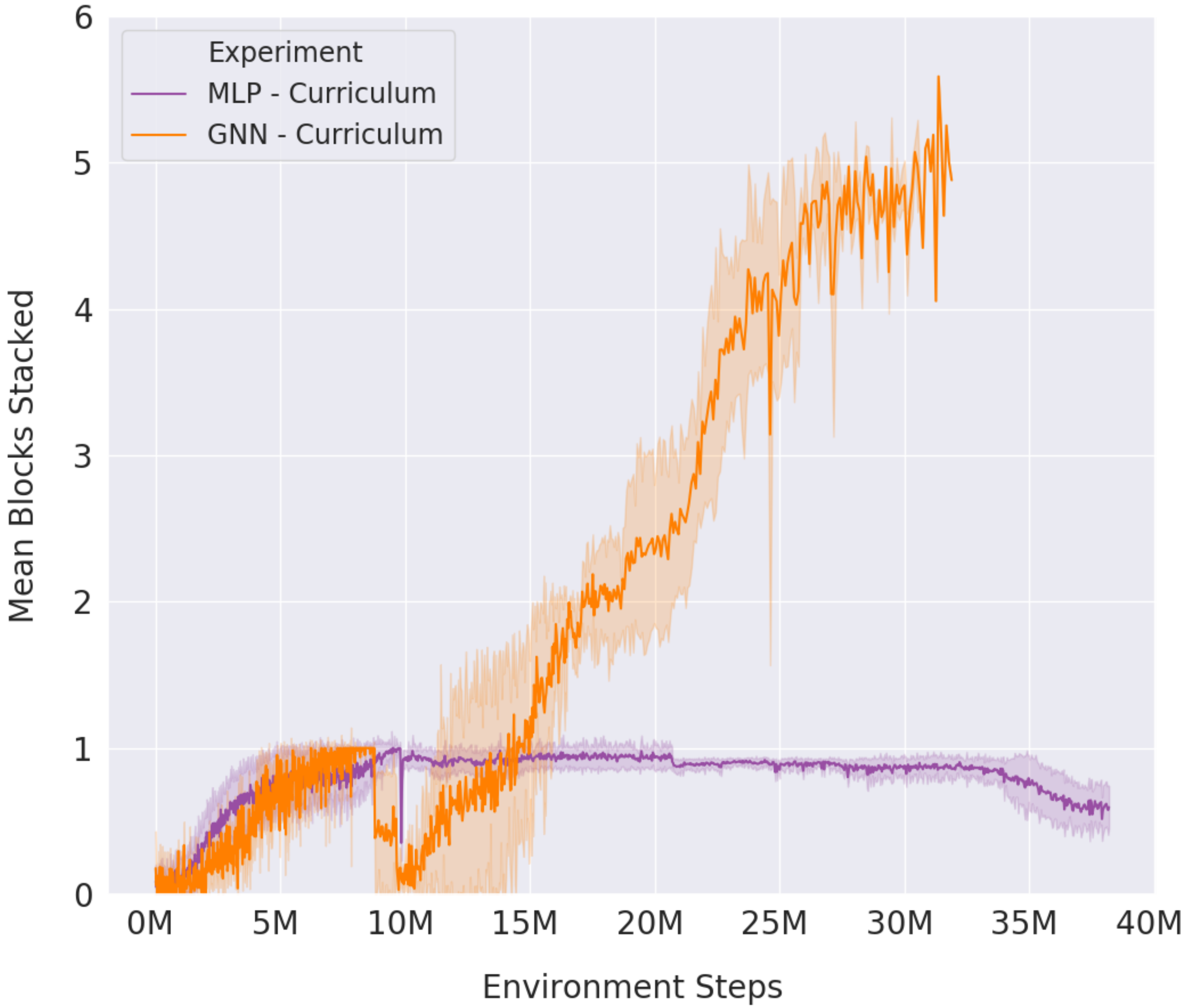
MHDPA/Self-attention (Vaswani '17)

Relational Deep RL (Zambaldi '19)

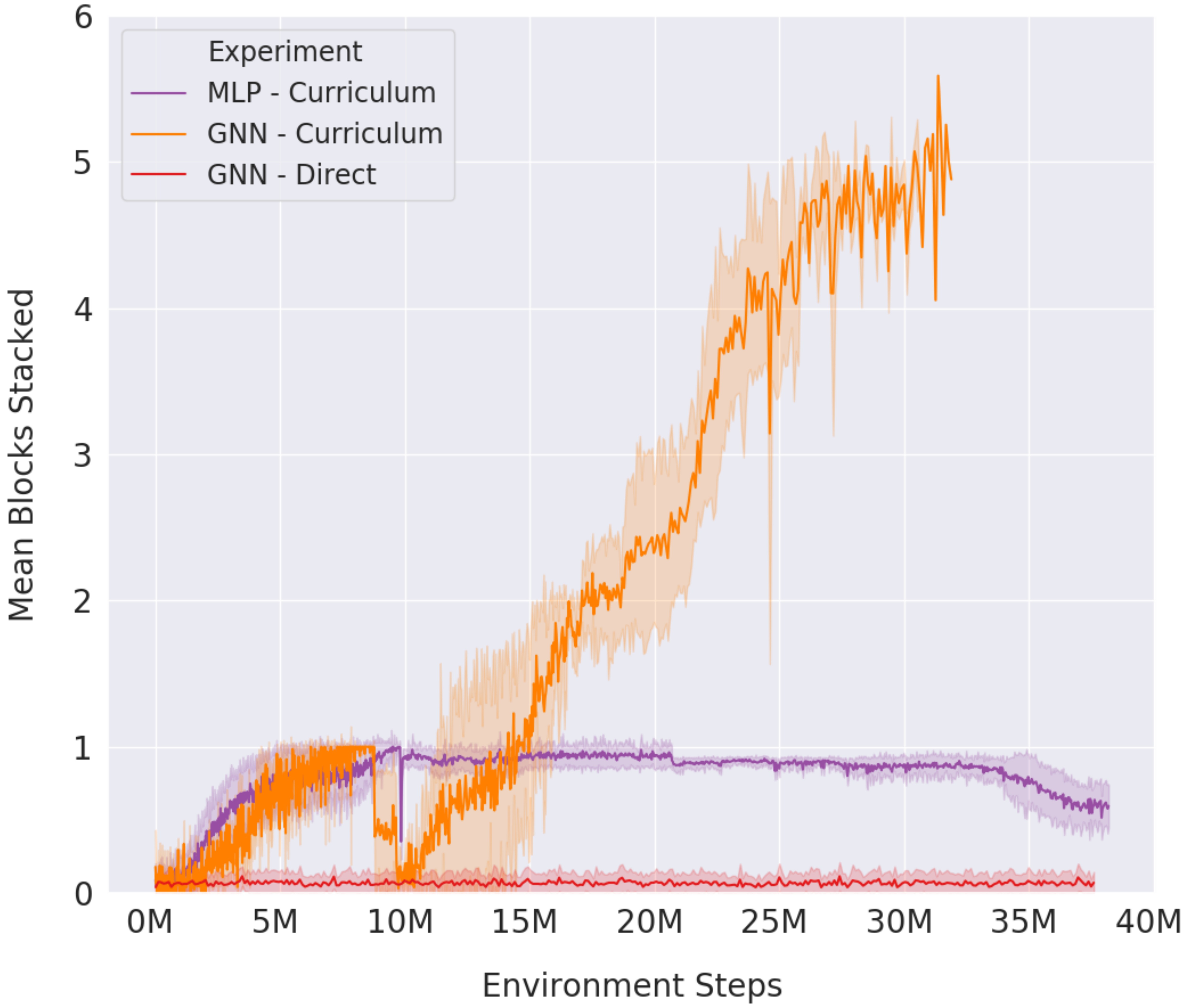
Curriculum with attention-based GNN

(Our method)

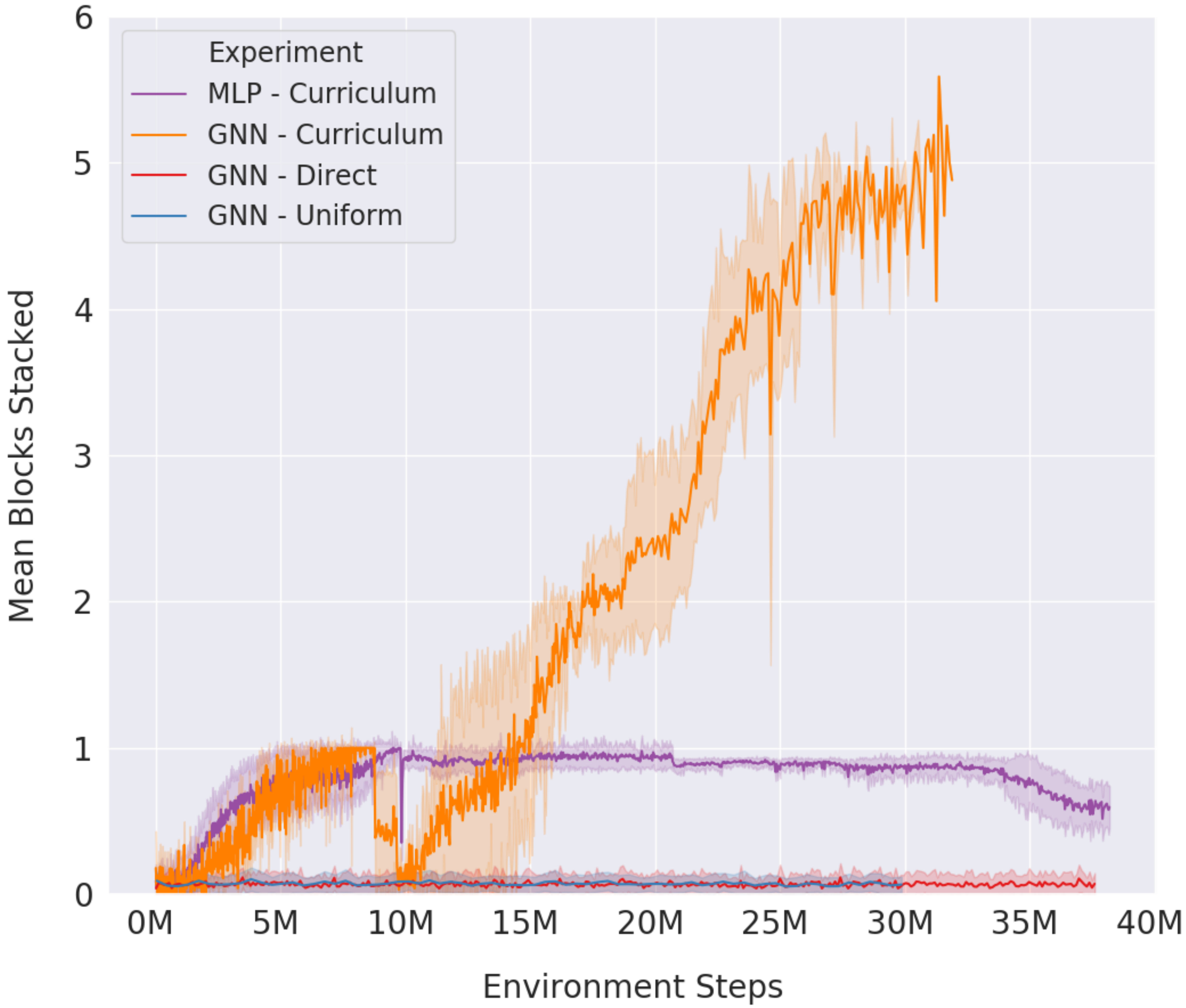




Comparison of GNN with/without Curriculum



Comparison of GNN with/without Curriculum

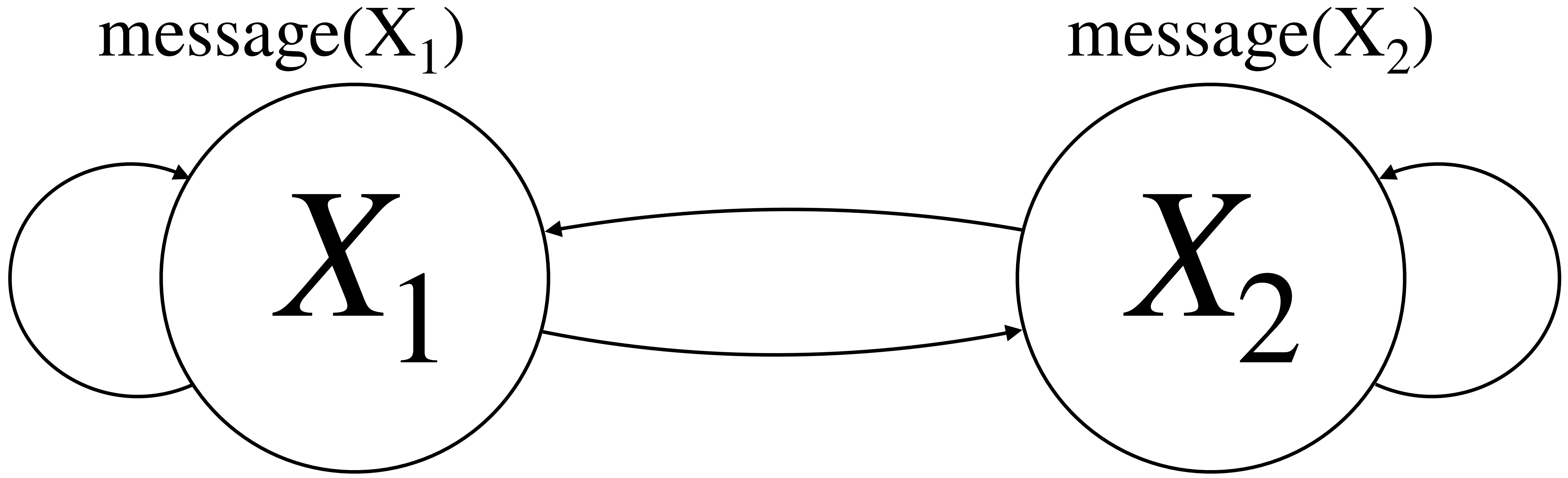


Why does the GNN transfer to more objects?

Local functions can be applied inductively to arbitrary # of nodes

Why does the GNN transfer to more objects?

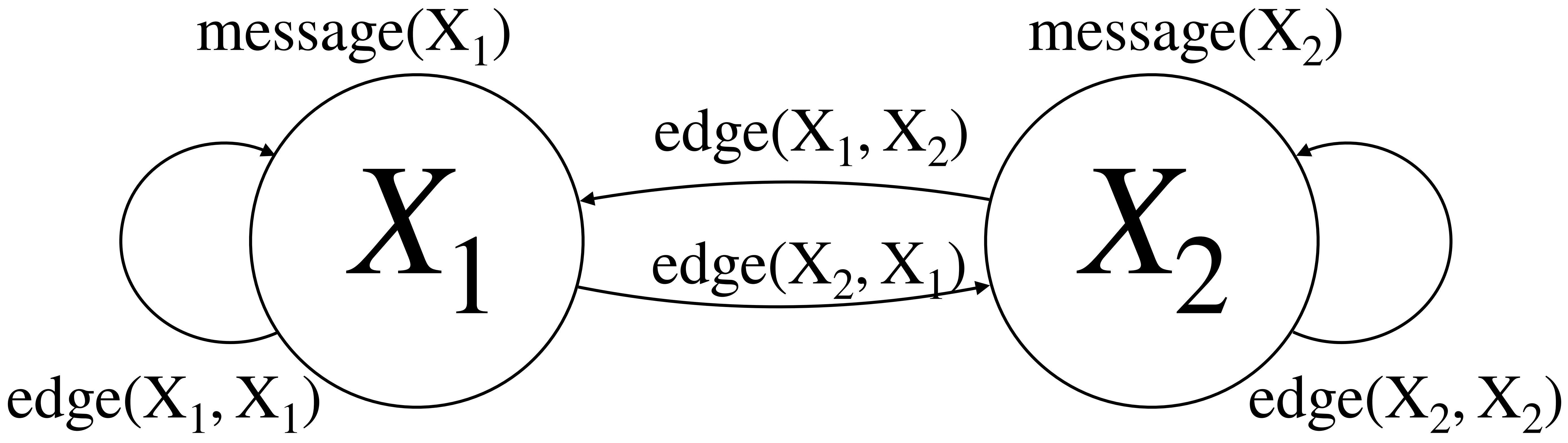
Local functions can be applied inductively to arbitrary # of nodes



message(v_i)

Why does the GNN transfer to more objects?

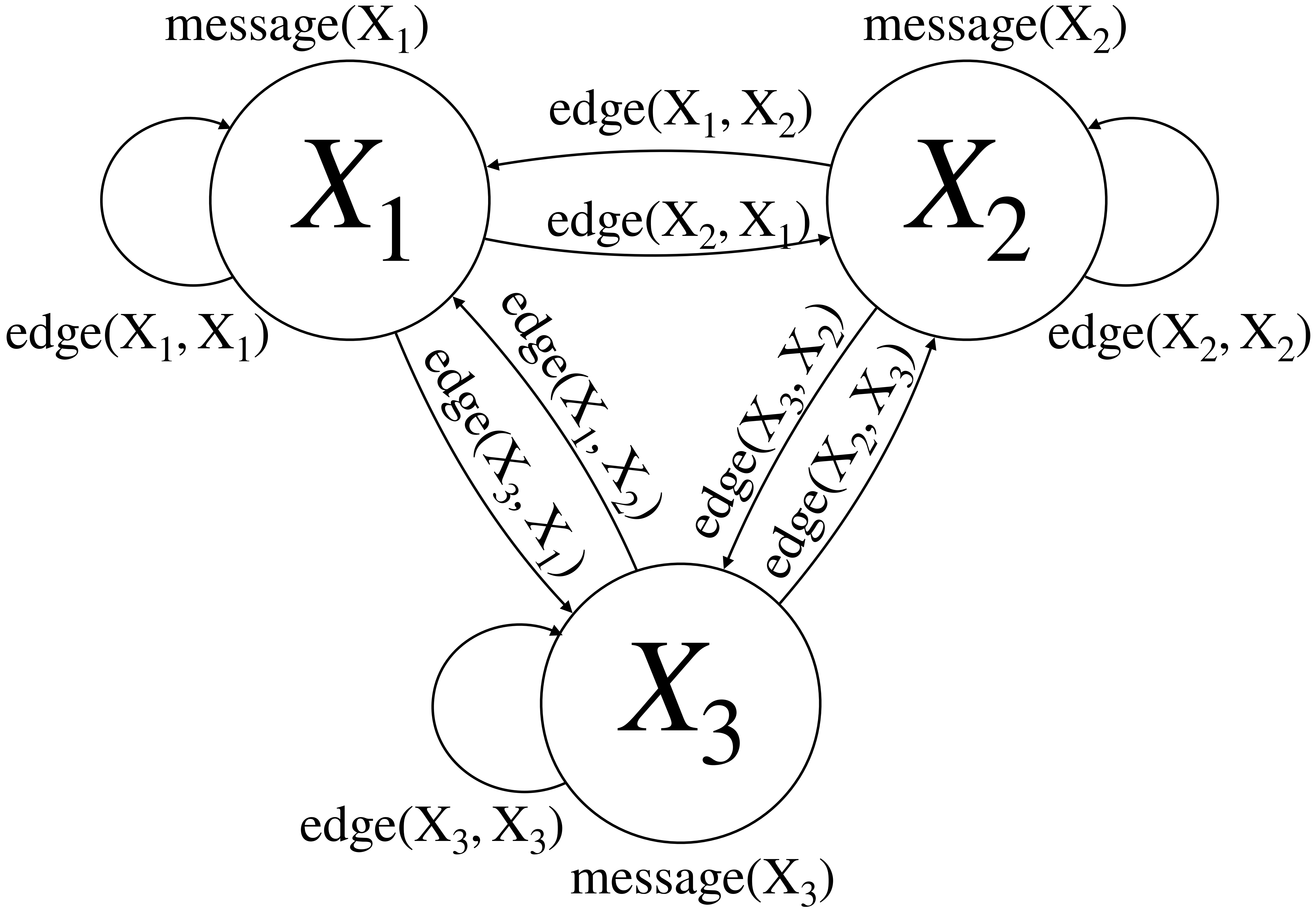
Local functions can be applied inductively to arbitrary # of nodes



message(v_i)
edge(v_i, v_j)

Why does the GNN transfer to more objects?

Local functions can be applied inductively to arbitrary # of nodes

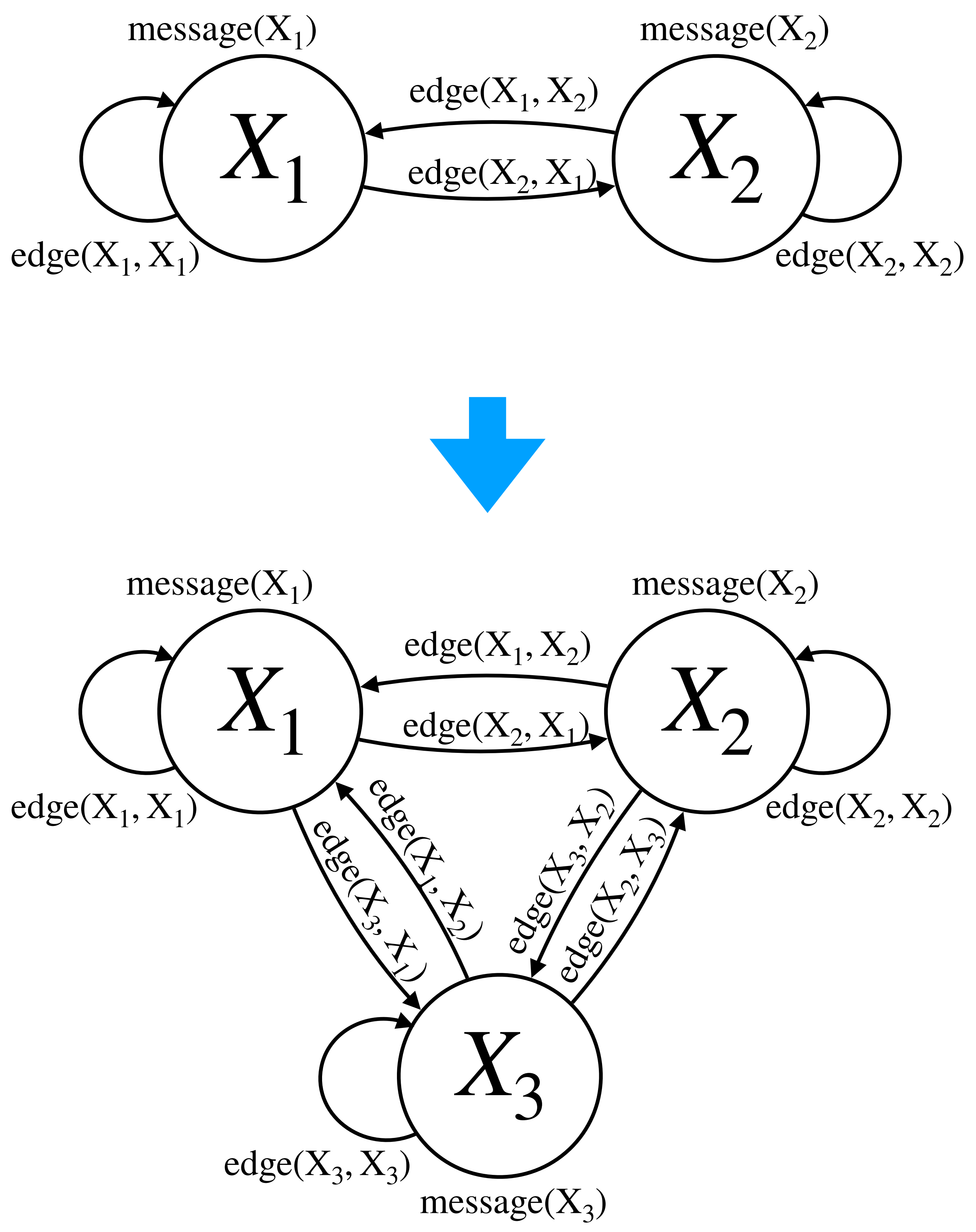


message(v_i)
edge(v_i, v_j)

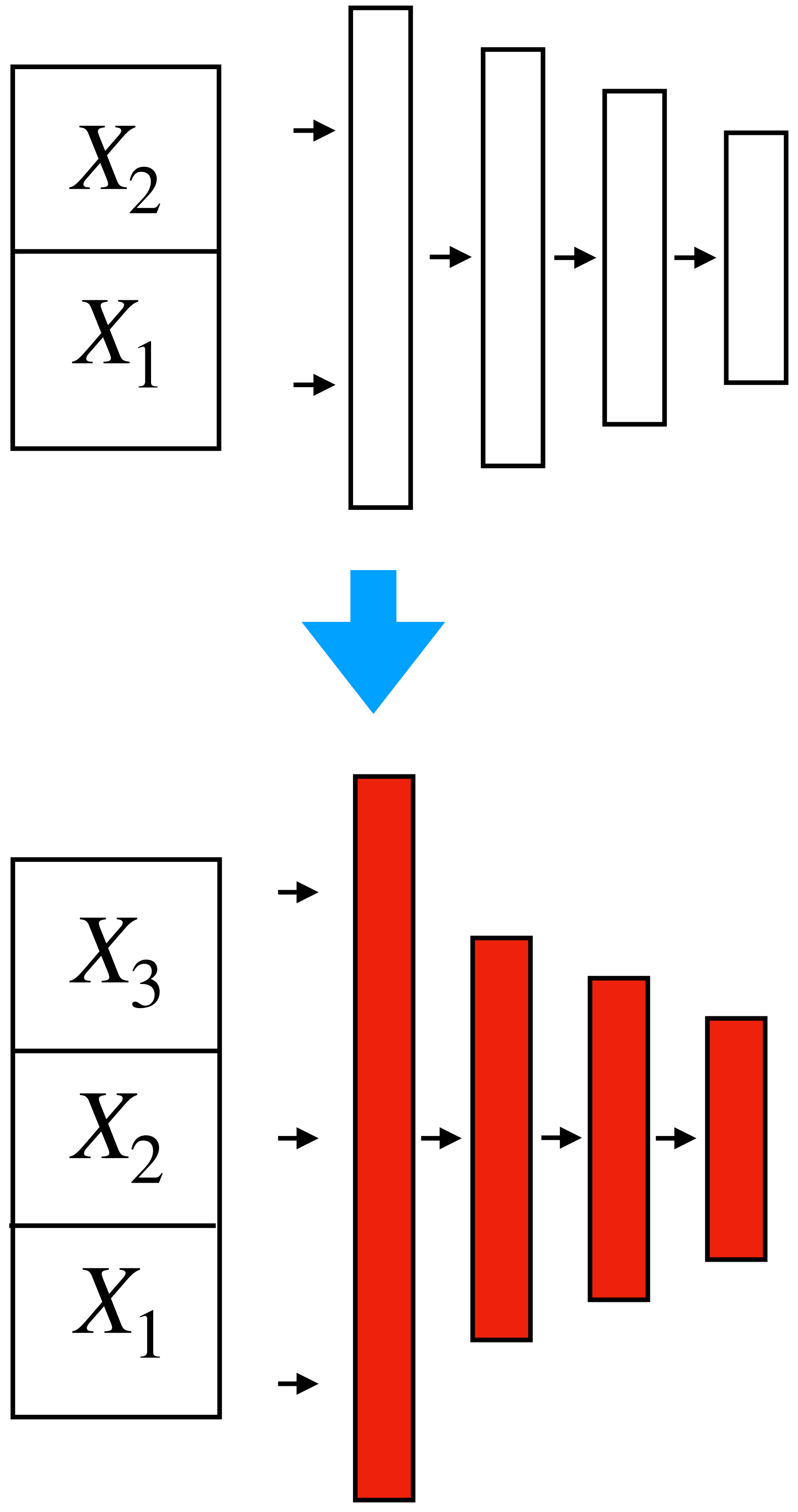
Why does the GNN transfer to more objects?

Local functions can be applied inductively to arbitrary # of nodes

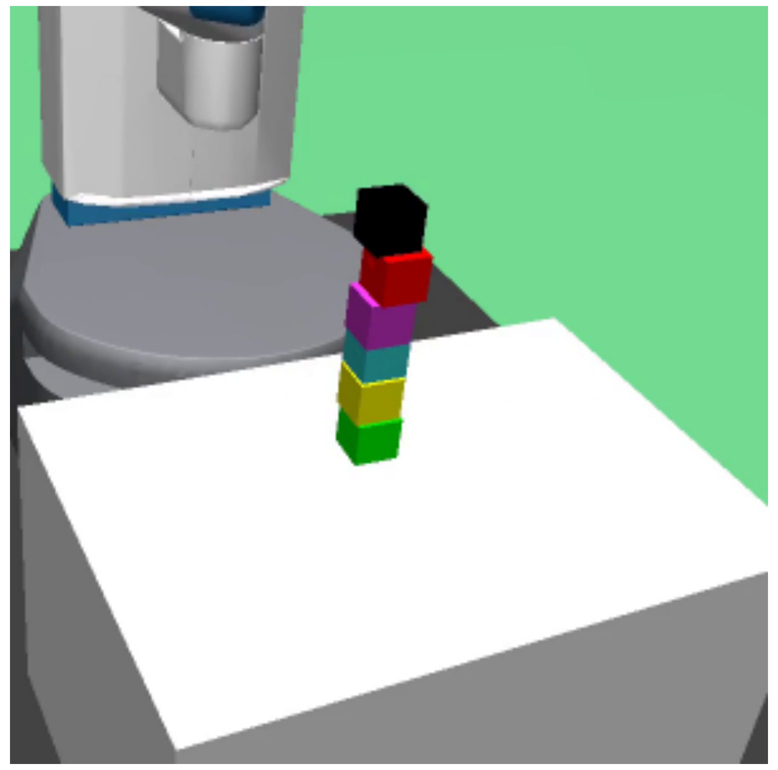
GNN



MLP



Zero-shot generalization



training
single tower 6

Emergent strategies

Failure modes

Conclusions

- Curriculum learning + graph architecture = drastic improvement in sample efficiency!
- Generalization to novel goals
- Learn efficient strategies that are hard to manually design

Website: http://richardrl.github.io/relational-rl