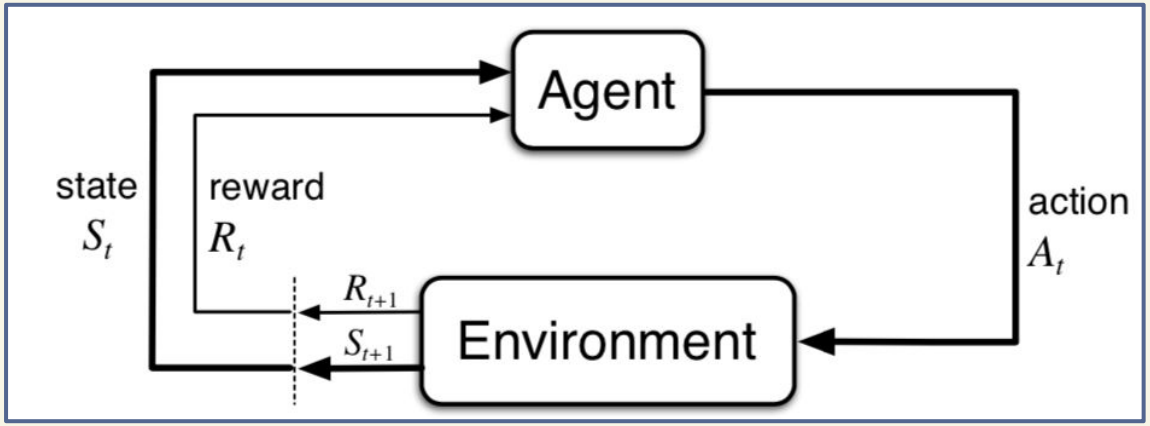
Learning Transferable Skills in Complex 3D Scenarios via Deep Reinforcement Learning

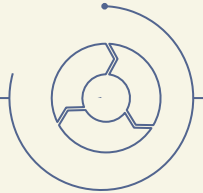
Lim You Rong
Supervised by Associate Professor Bo An

Markov Decision Process

- State
- Actions
- Reward
- Transition Probability
- Discount Factor



Policy Iteration



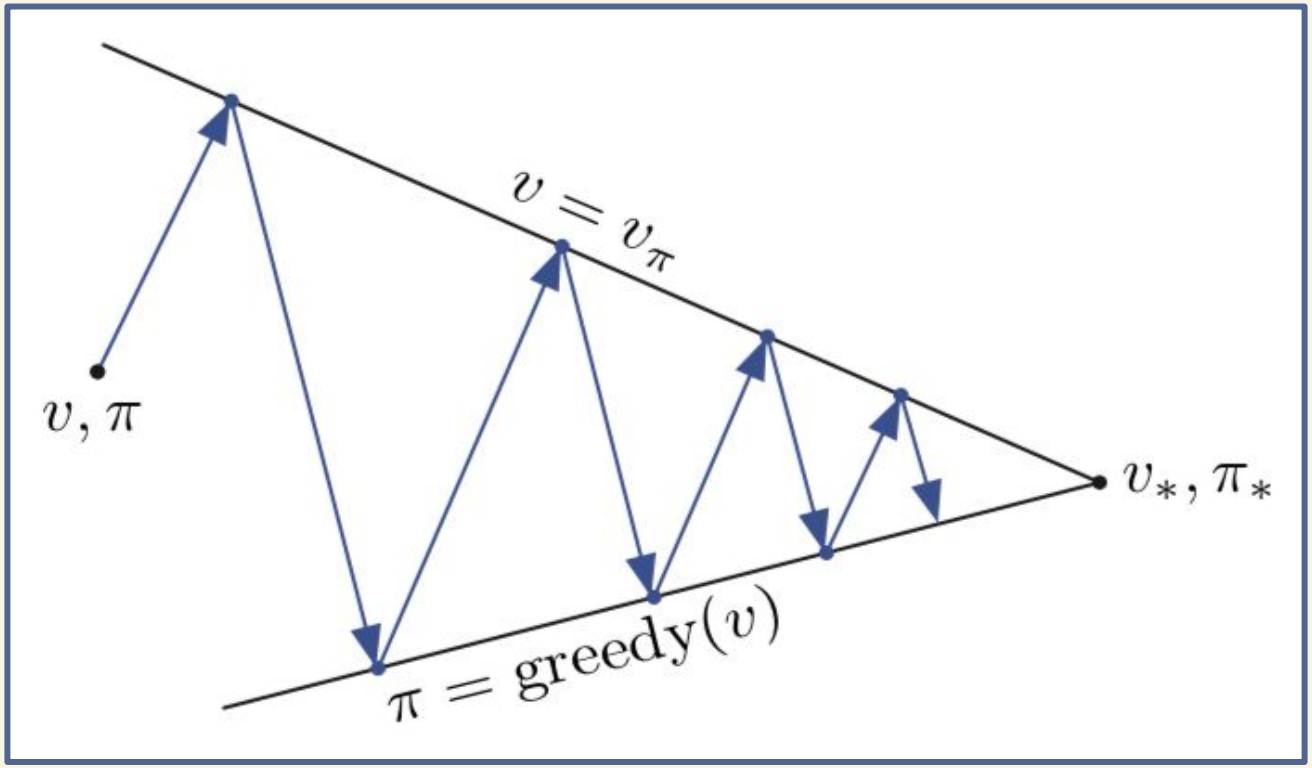
Value Function

Estimate value for all actions at all states

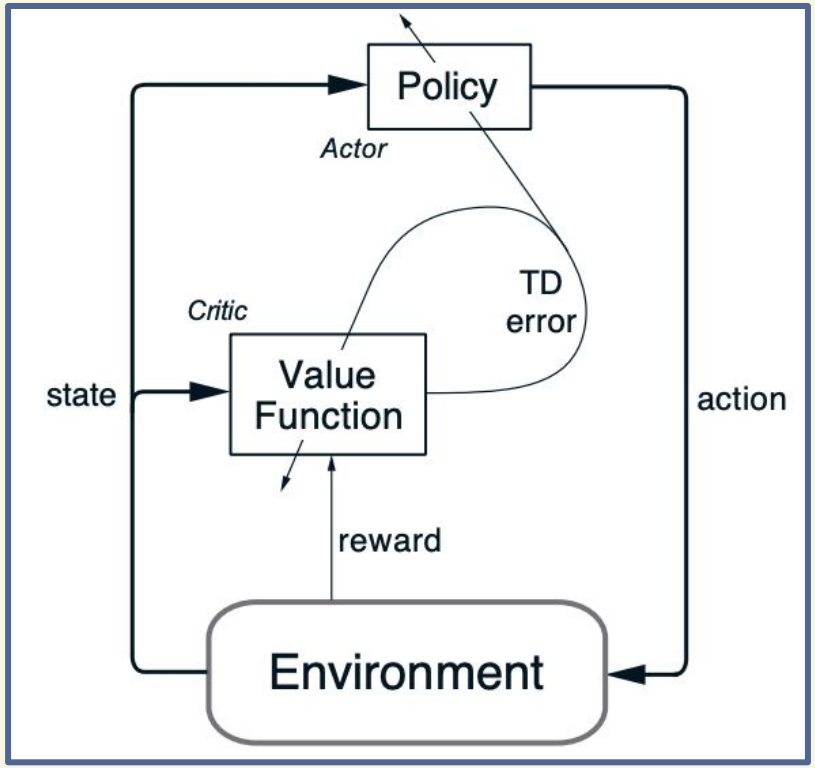
Policy Gradient

Optimize the policy directly using Gradient Ascend

Policy Iteration
(Exploration vs Exploitation)

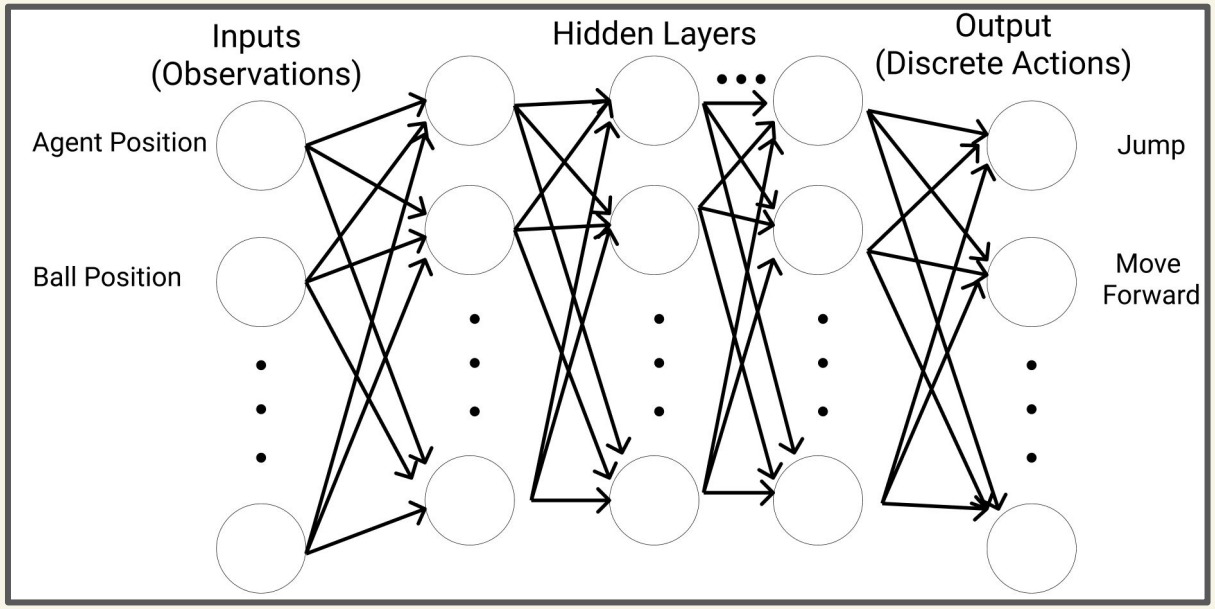


Actor Critic Method



Deep Reinforcement Learning

- Specialised model based off training data
- Memory efficient using function approximation



Problem & Challenges

Large State Space

Complexity Yet Learnable

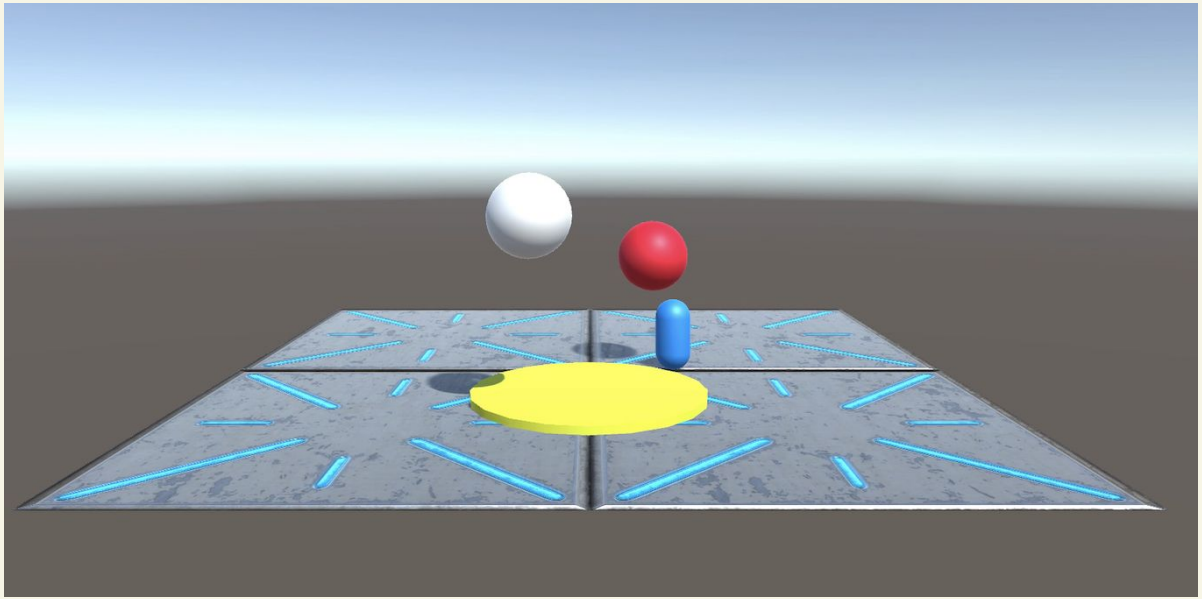
Transferability Of Skills

Algorithms

	Soft Actor Critic (SAC)	Proximal Policy Optimization (PPO)
Policy Design	Off Policy	On Policy
Key Feature	Replay Buffer & Entropy Regularization	Clipping Function using Trust Region
Strengths	Sample Efficiency	Training Stability

Problem Introduction

- Agent - Blue
- Ball - Red
- Scoring Area - Yellow
- Target - White



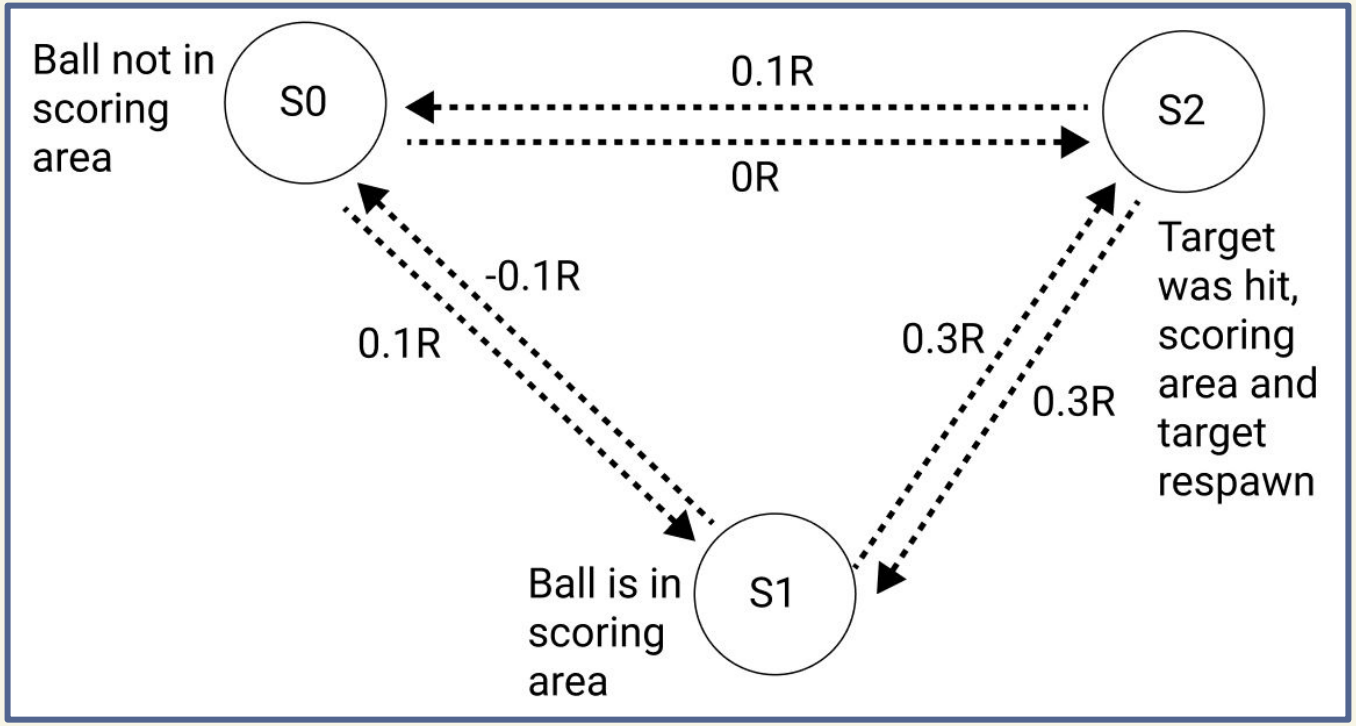
Sparse Reward Environment

Lack appropriate feedback for improvement

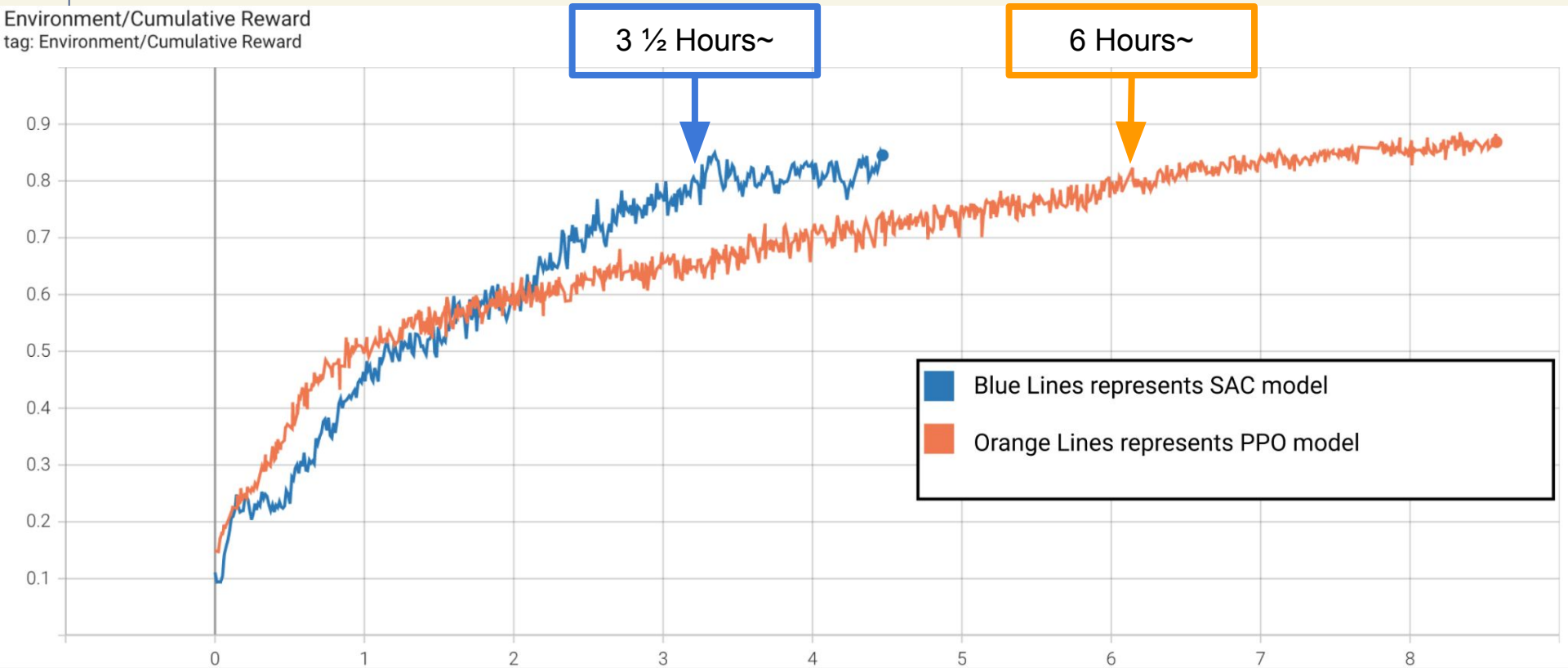
Time and resource intensive

$$r_{\mathcal{M}}(\mathbf{s}, \mathbf{a}) = \begin{cases} \delta_{\mathbf{s}_g}(\mathbf{s}) & \text{if } d(\mathbf{s}, \mathbf{s}_g) \leq \epsilon \\ 0 & \text{else,} \end{cases}$$

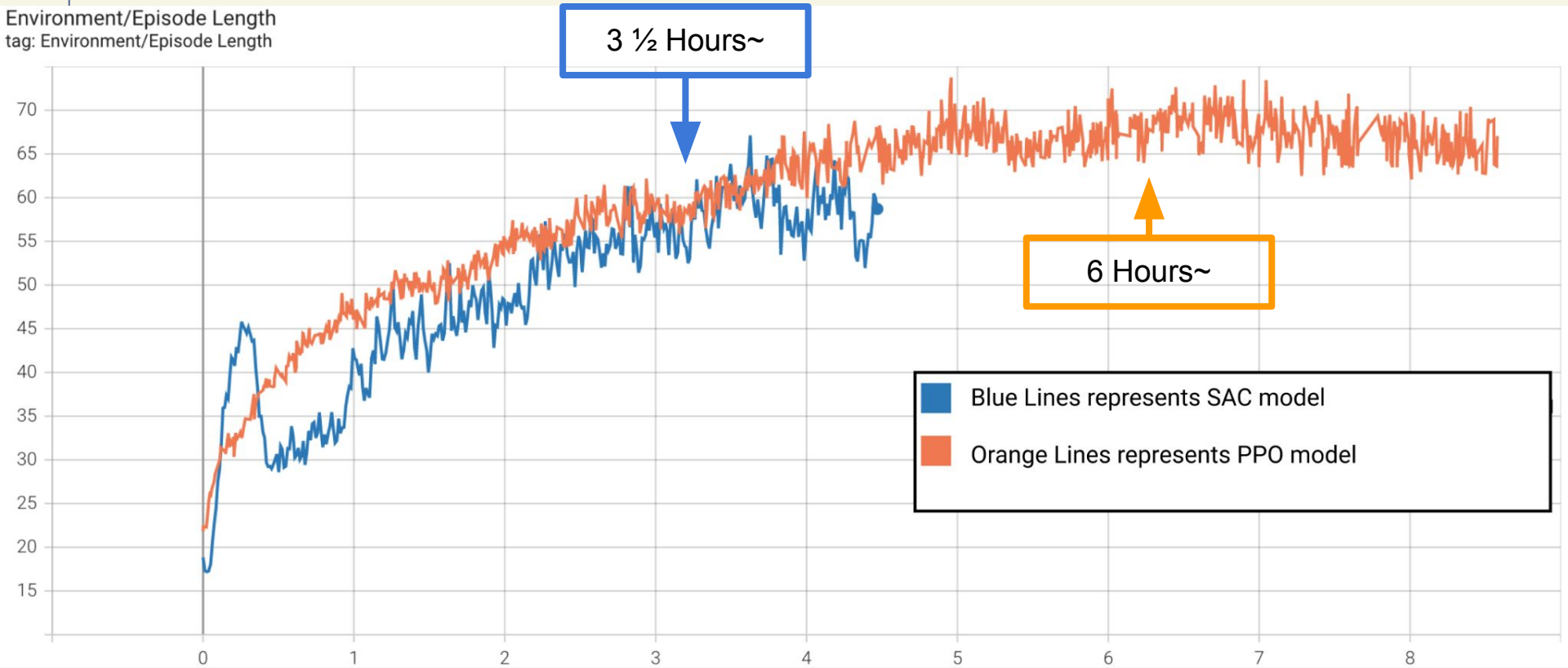
Reward System



Reward Shaping (PPO vs SAC)



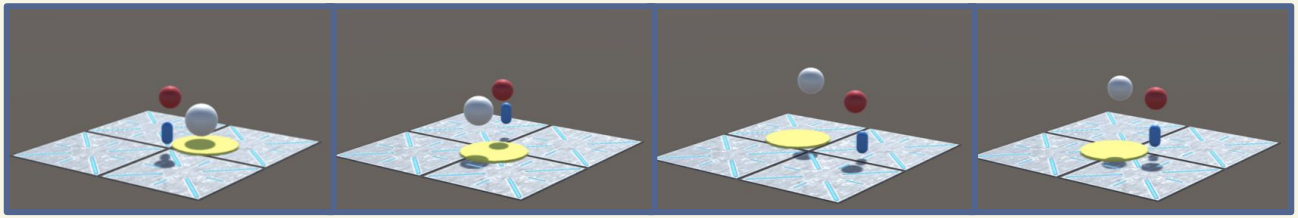
Reward Shaping (PPO vs SAC)

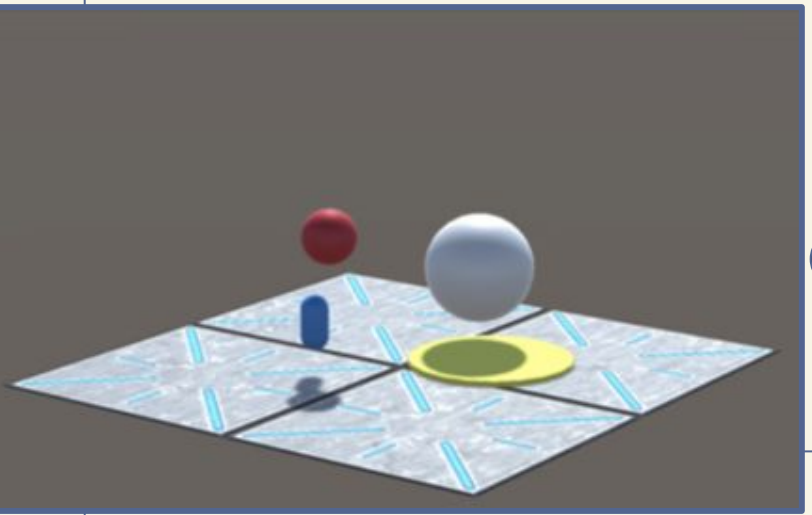


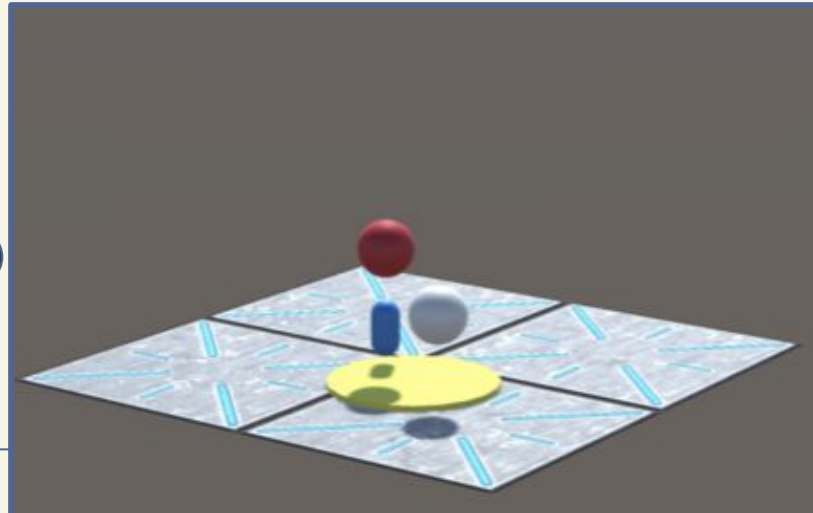
Curriculum Learning

Reduced Difficulty

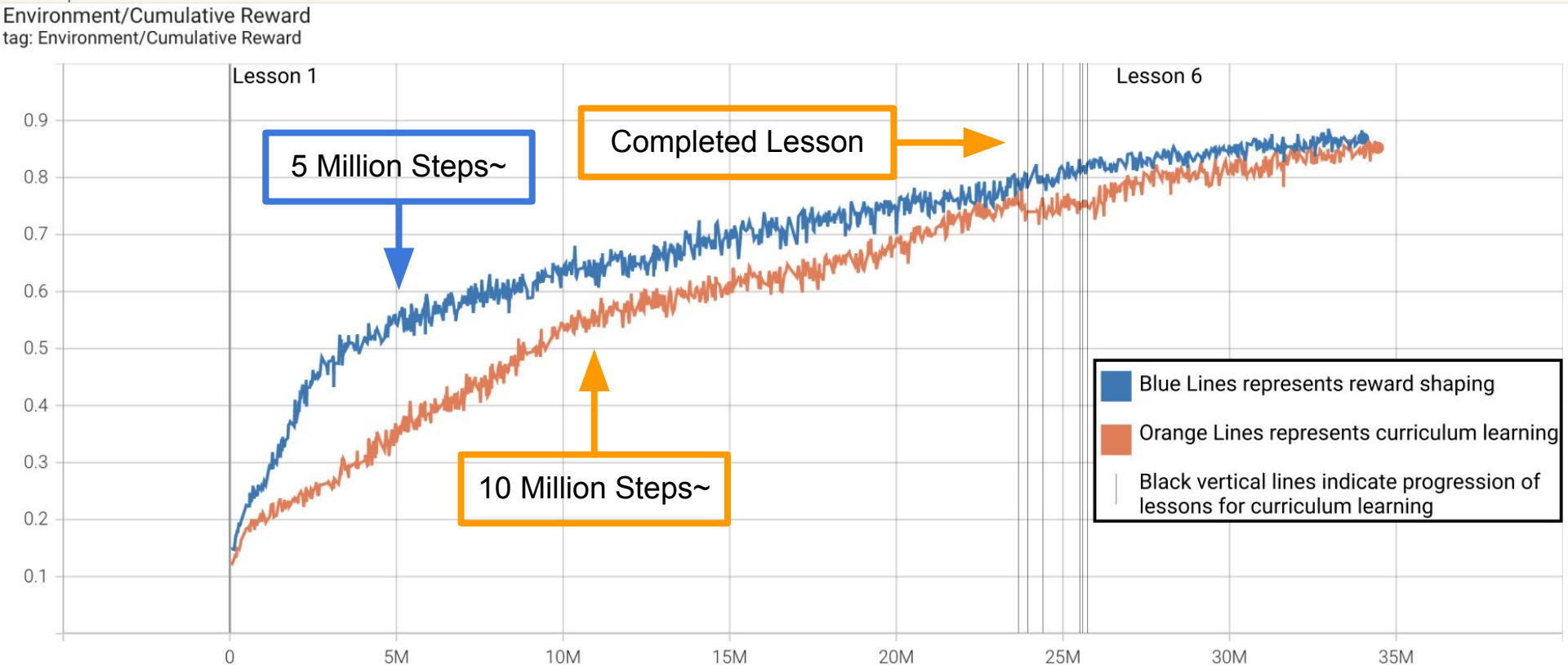
Original Difficulty



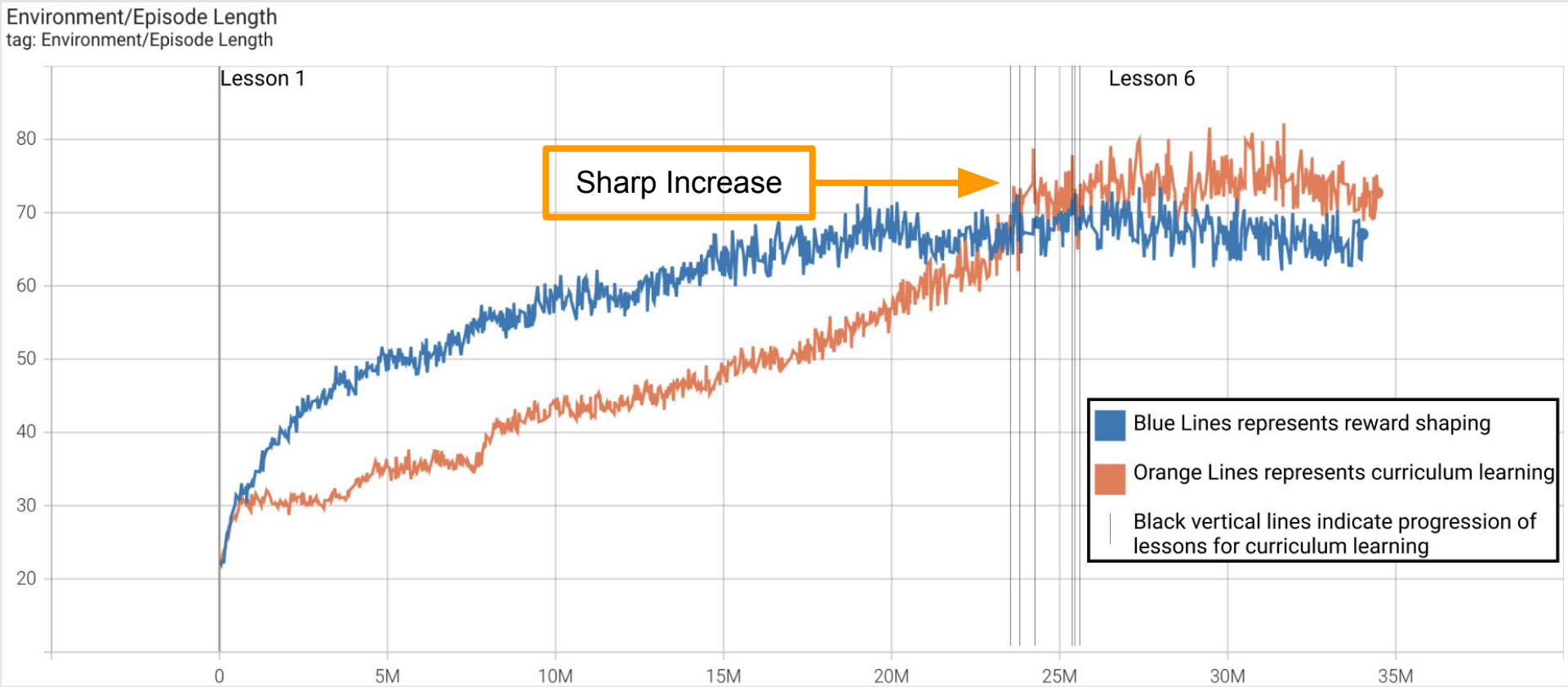




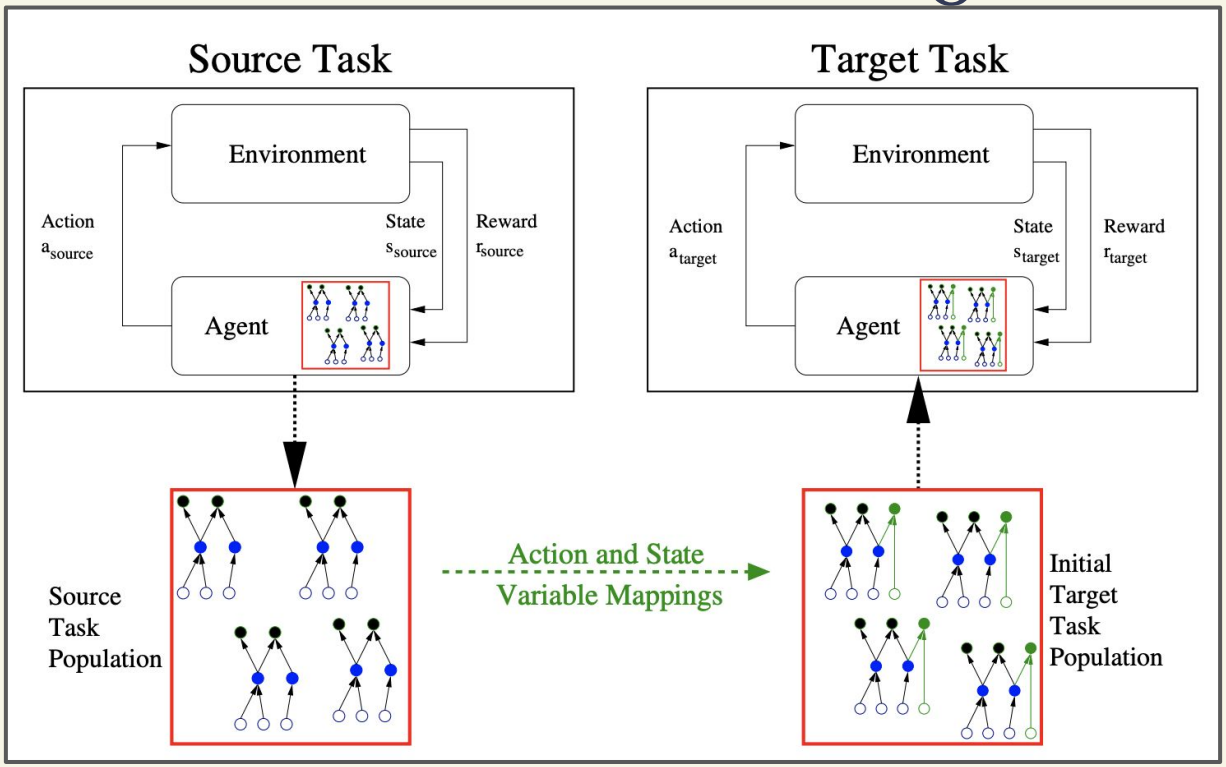
Curriculum Learning



Curriculum Learning

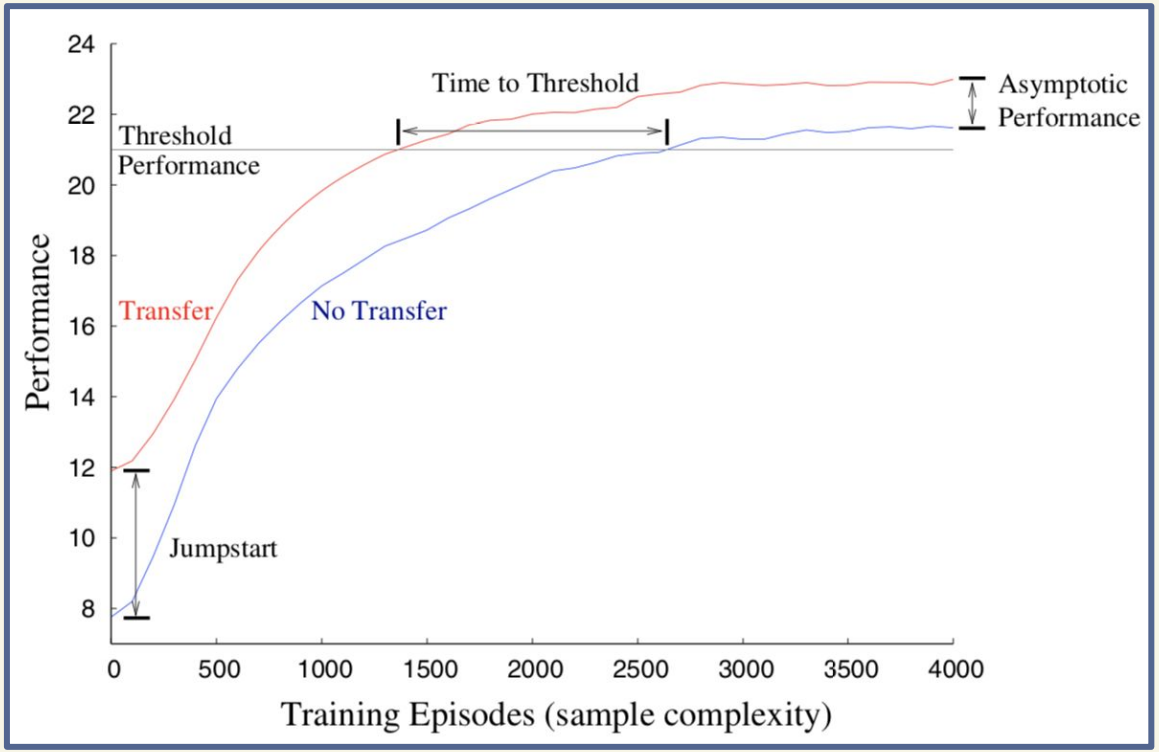


Transfer Learning

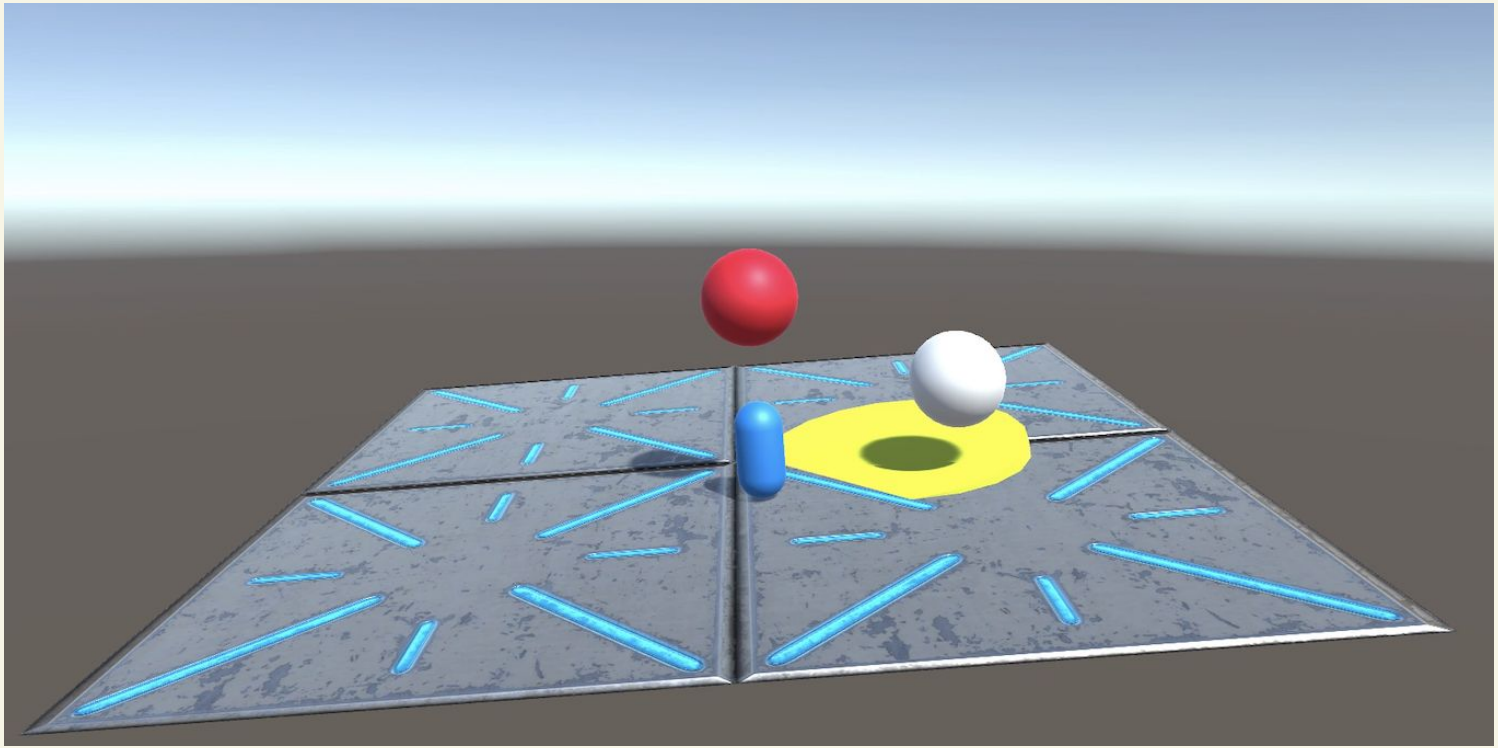


Baseline

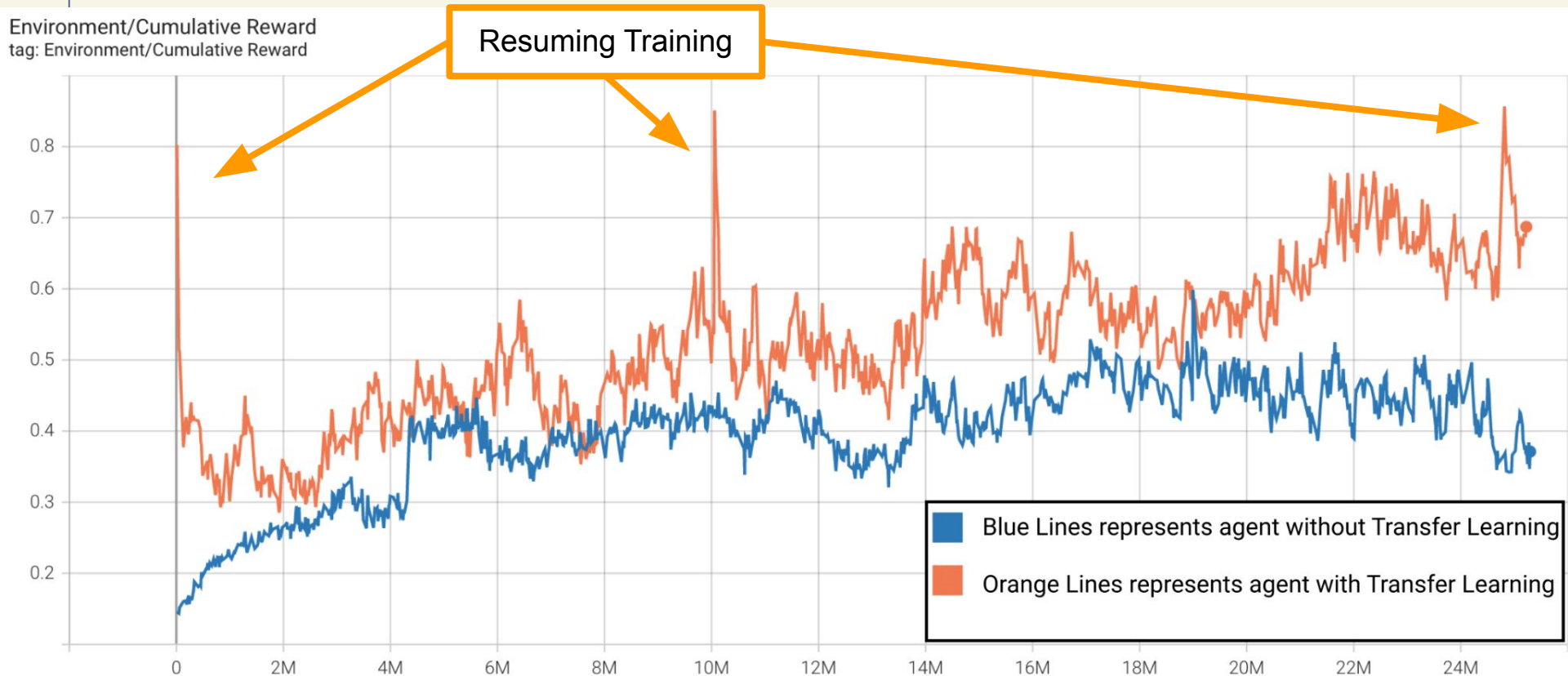
- Threshold performance
- Jump Start
- Min Episode Length above Threshold



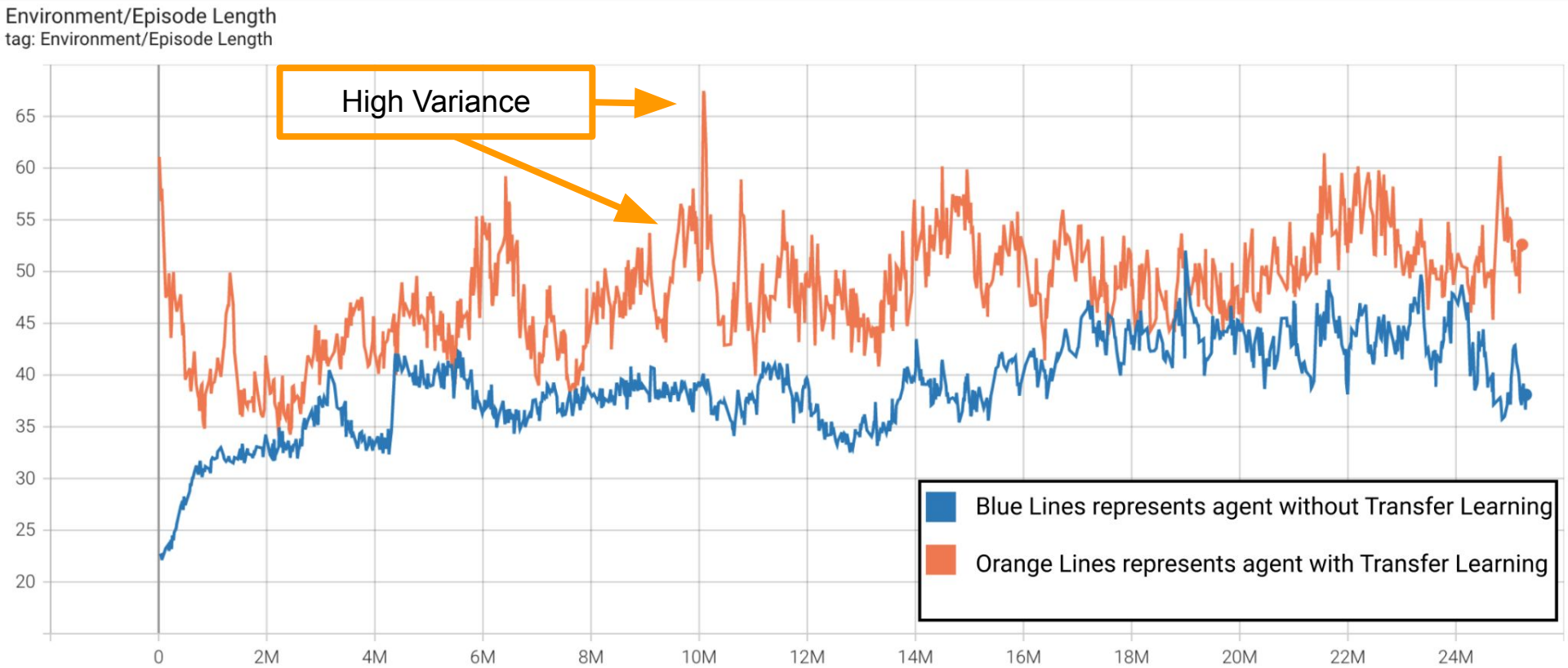
Transfer for sloped surface



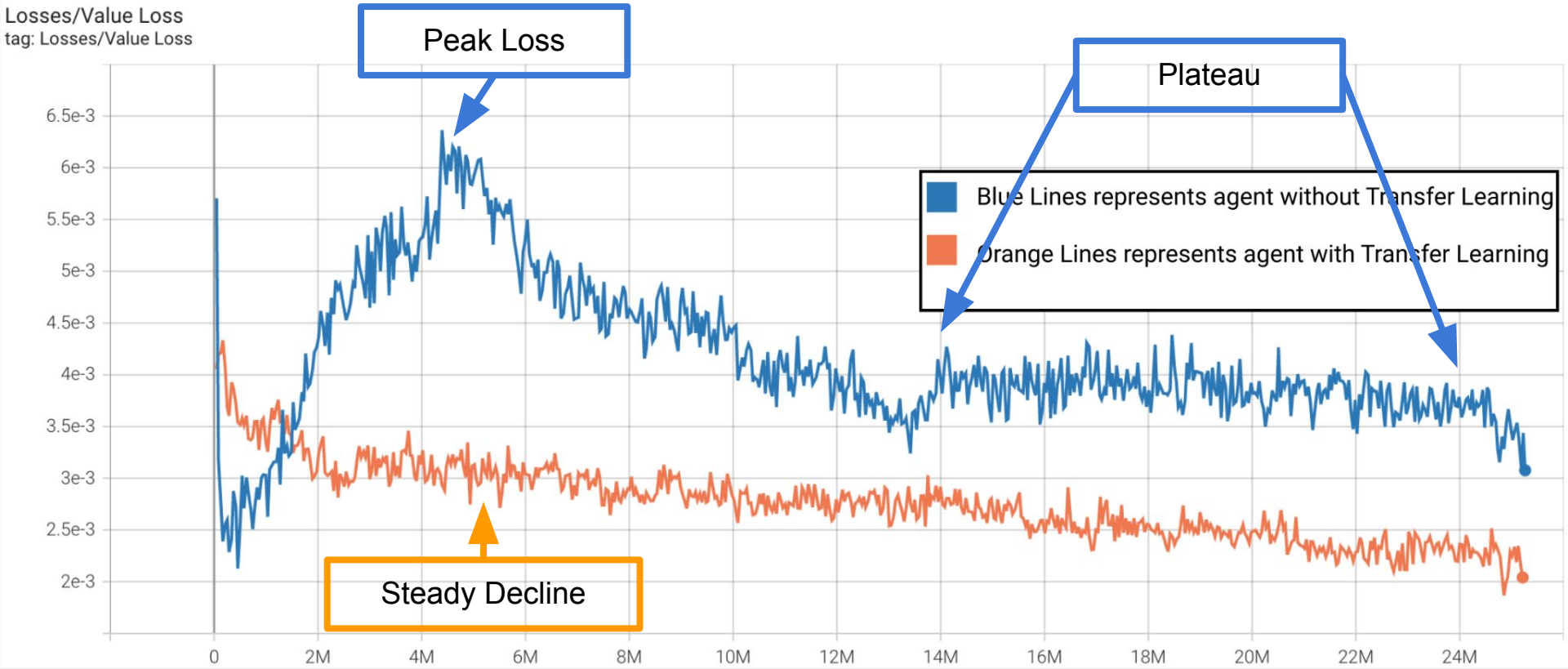
Transfer for sloped surface



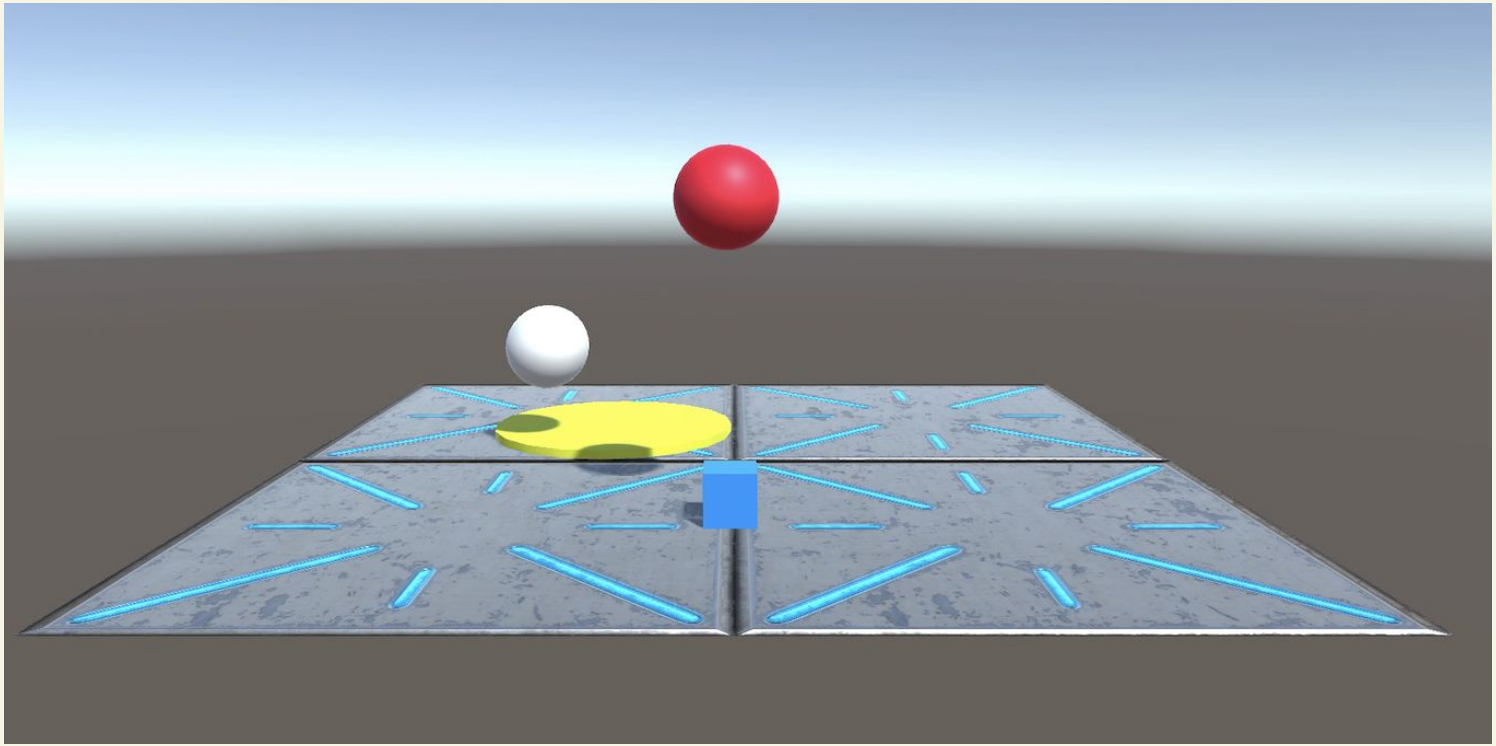
Transfer for sloped surface



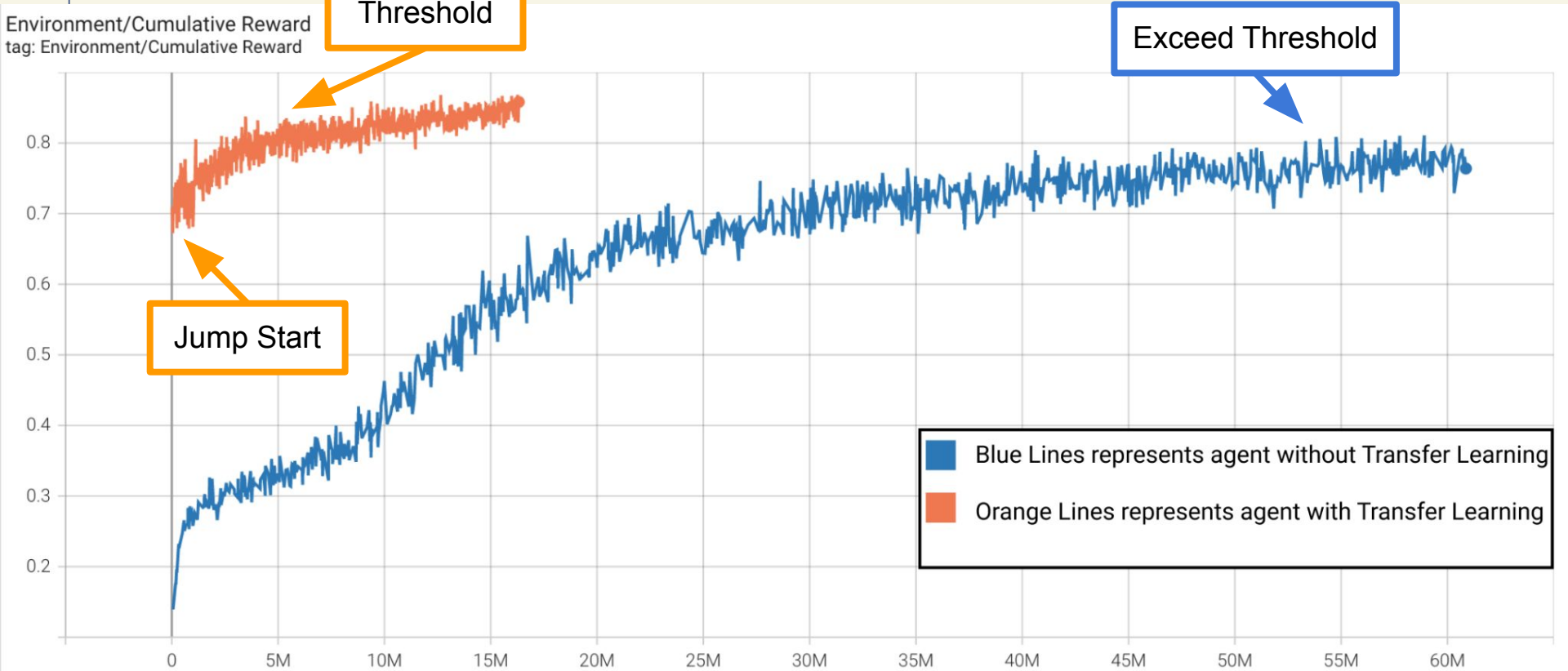
Transfer for sloped surface



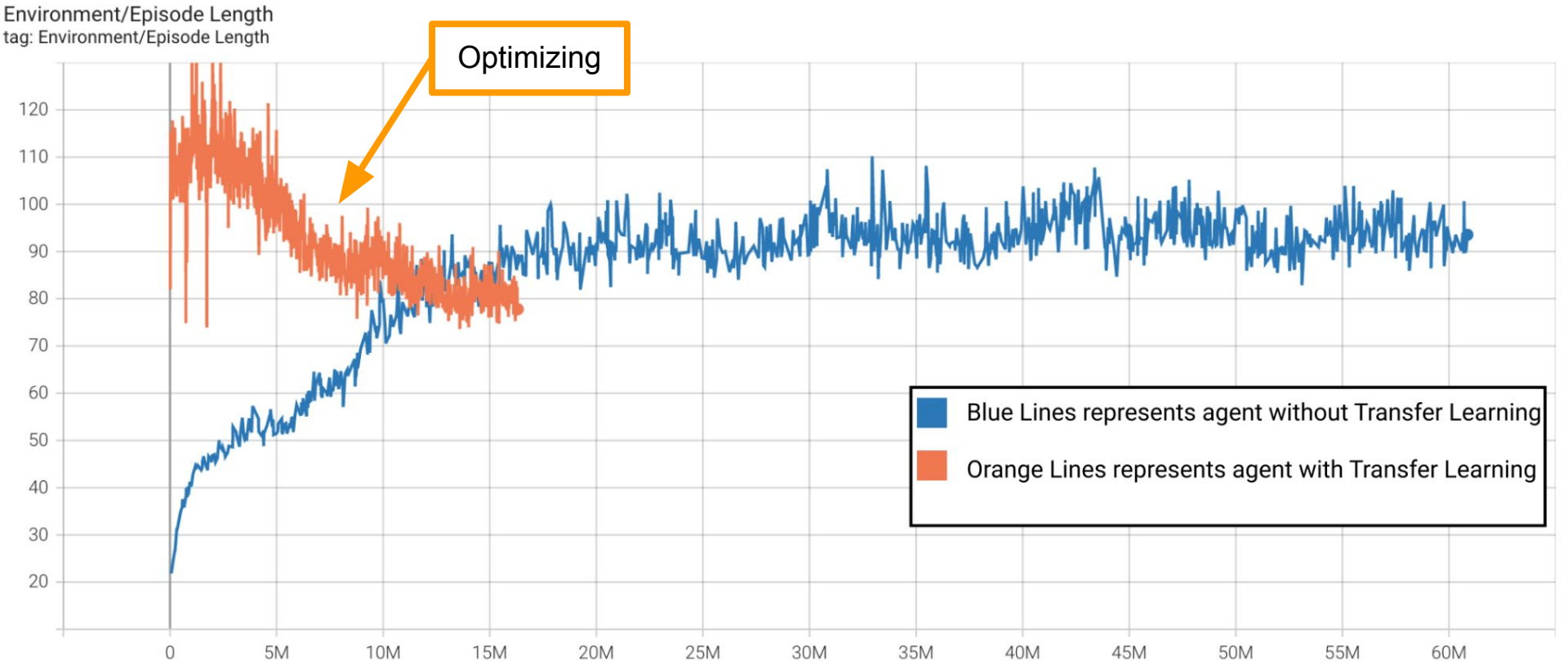
Transfer For Cuboid Agent



Transfer For Cuboid Agent



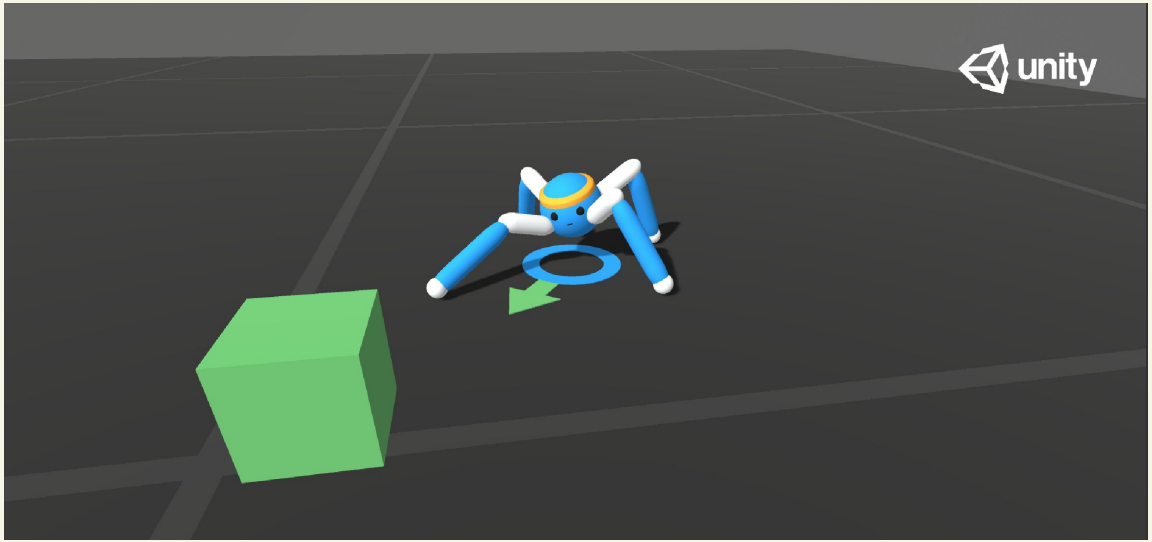
Transfer For Cuboid Agent



“Transferring skills **is possible** in complex Sparse Reward scenarios trained with Deep Reinforcement Learning.”

Future Work

- Transfer to agent with different movement scheme



Failed Strategies

Pure DRL

Reverse Curriculum

Non-Linear Transfer

THANKS

Do you have any questions?

youremail@freepik.com
+91 620 421 838
yourcompany.com

References

Images from
R. S. Sutton and A. G. Barto, Reinforcement learning : an introduction / Richard S. Sutton and Andrew G. Barto., Second edition. Cambridge, Massachusetts: The MIT Press, 2018.

M. Taylor, Transfer in Reinforcement Learning Domains [electronic resource] / by Matthew Taylor., 1st ed. 2009. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. doi: 10.1007/978-3-642-01882-4.

https://github.com/Unity-Technologies/ml-agents/blob/main/docs/Learning-Environment-Examples.md