Deep Reinforcement Learning in a Handful of Trials with Probabilistic Dynamics Models

Kurtland Chua, Roberto Calandra, Rowan McAllister, Sergey Levine

University of California, Berkeley

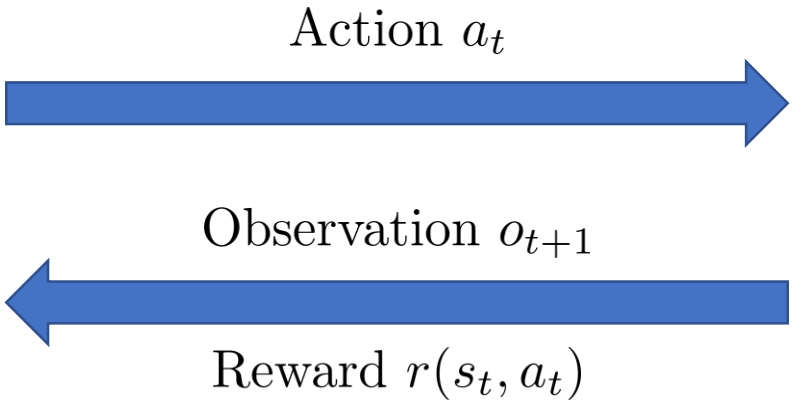
What is Reinforcement Learning?

What is Reinforcement Learning?

What is Reinforcement Learning?

Action a_t

What is Reinforcement Learning?



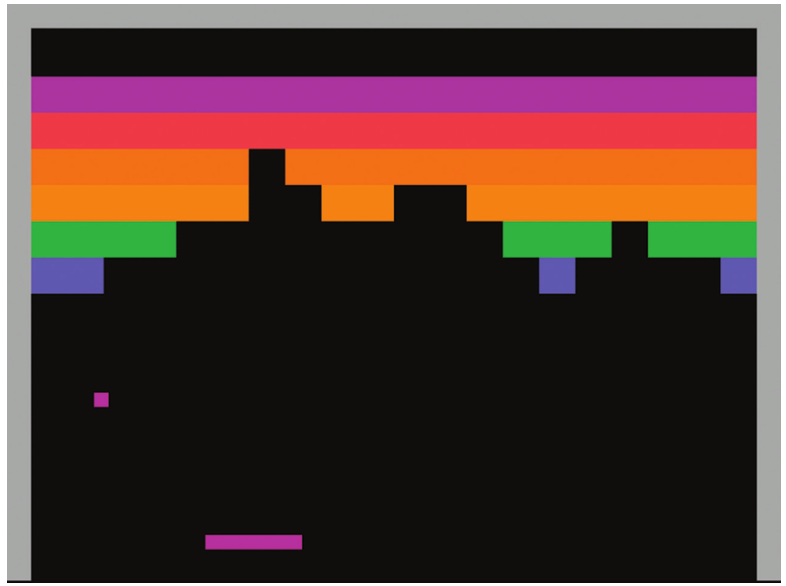
What is Reinforcement Learning?

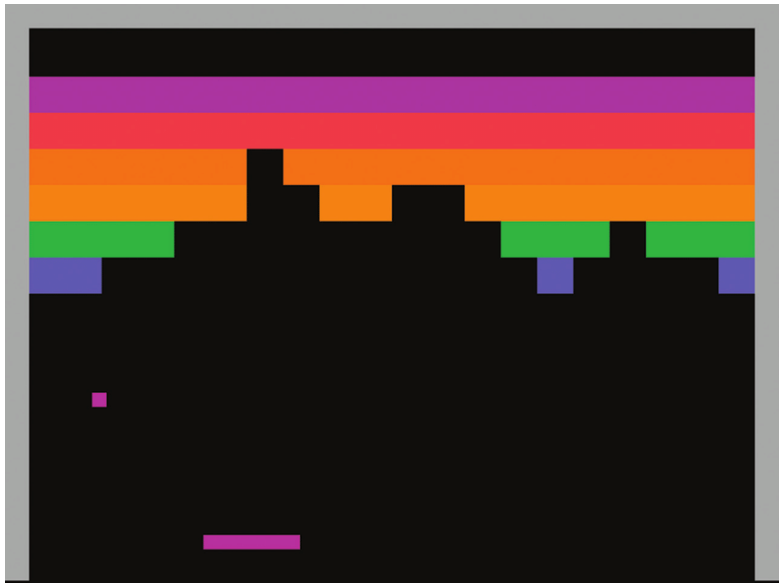
Action a_t

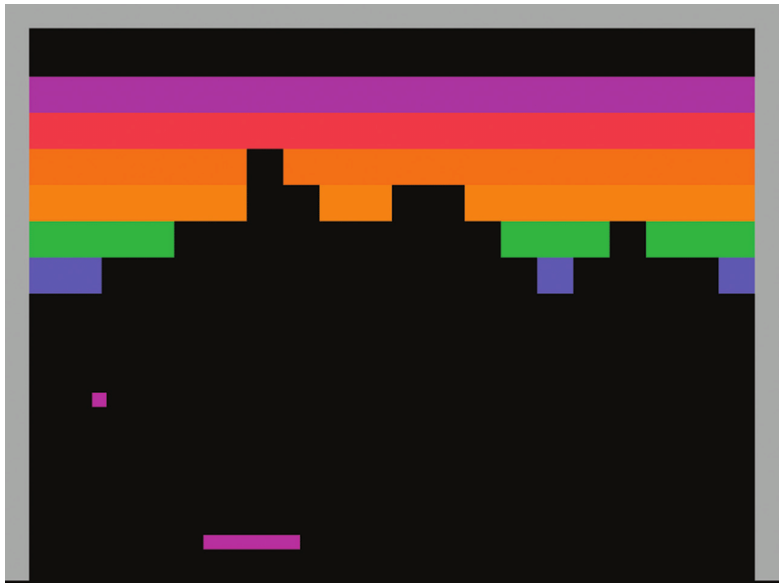
Observation o_{t+1}

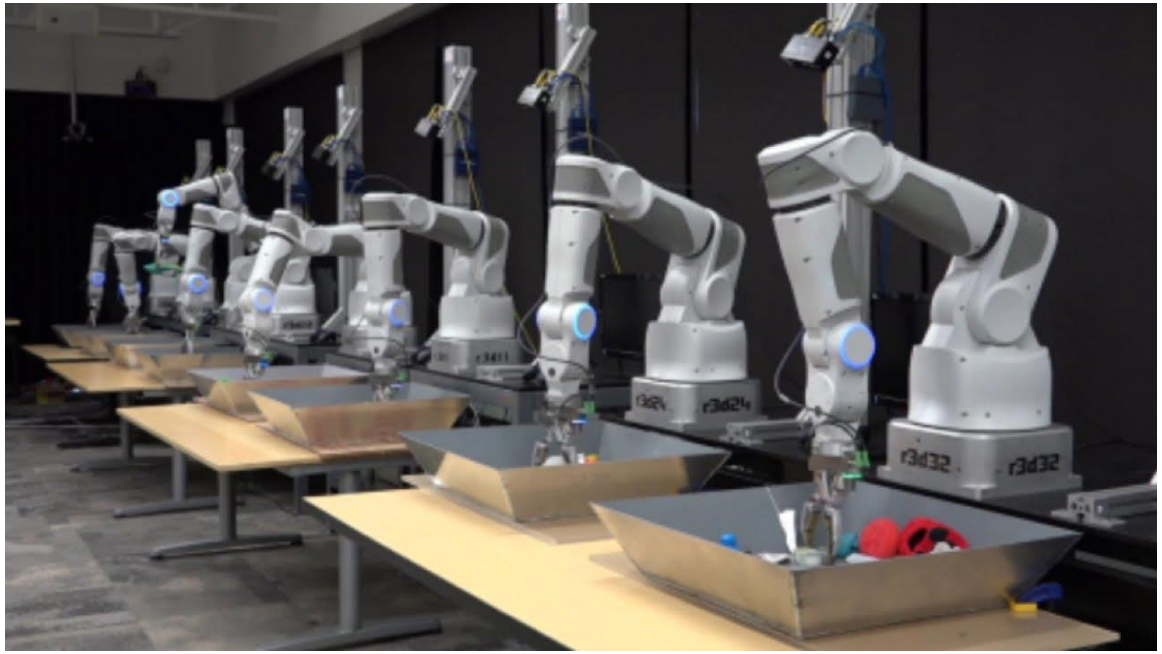
Reward $r(s_t, a_t)$

$$\max_{a_1,\ldots,a_{T-1}} \sum_{i=1}^{T-1} r(s_t, a_t)$$





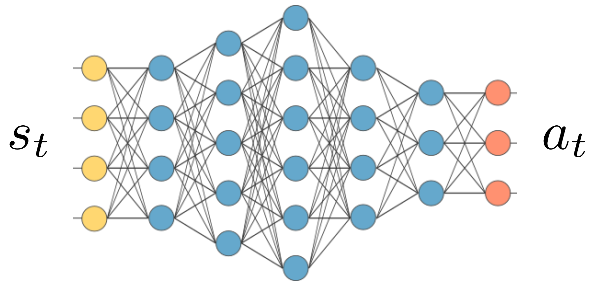




Current Approaches to Reinforcement Learning

Current Approaches to Reinforcement Learning

Policy Gradients

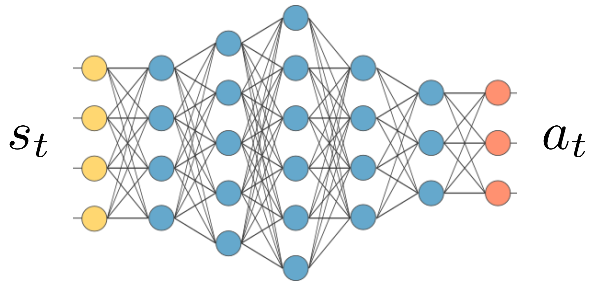


PPO, TRPO

(Schulman et al. ‘15, Schulman et al. ‘17)

Current Approaches to Reinforcement Learning

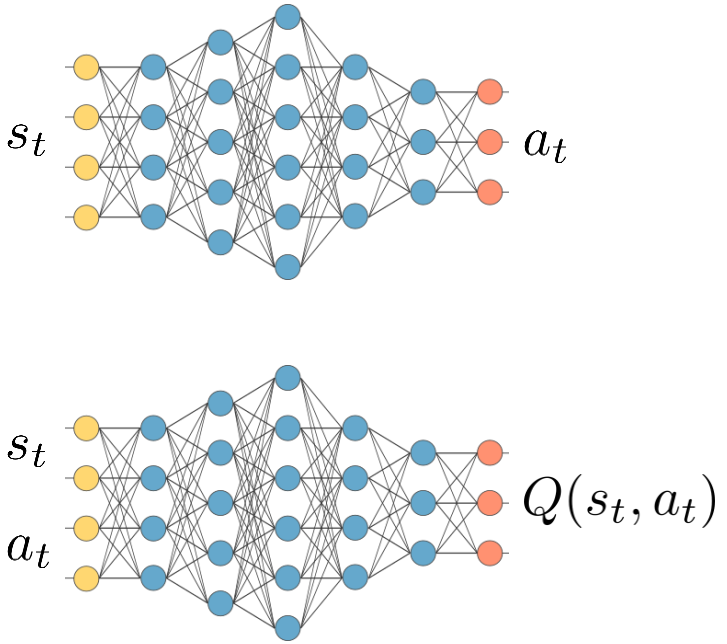
Policy Gradients



PPO, TRPO

(Schulman et al. ‘15, Schulman et al. ‘17)

Actor-Critic

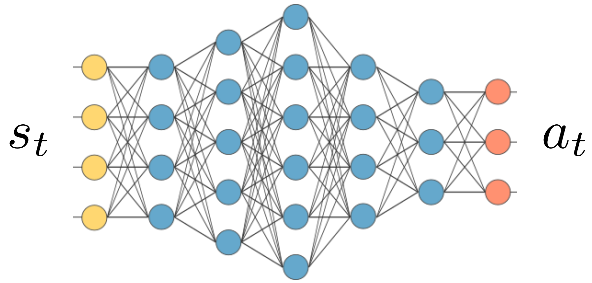


SAC, DDPG

(Haarnoja et al., Lillicrap et al.)

Current Approaches to Reinforcement Learning

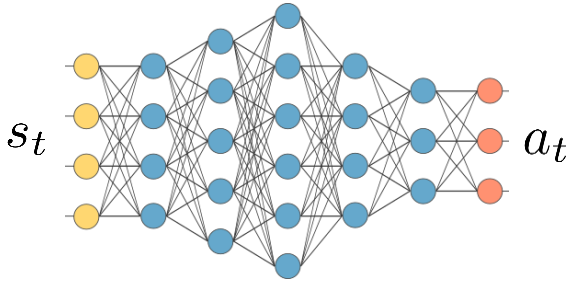
Policy Gradients

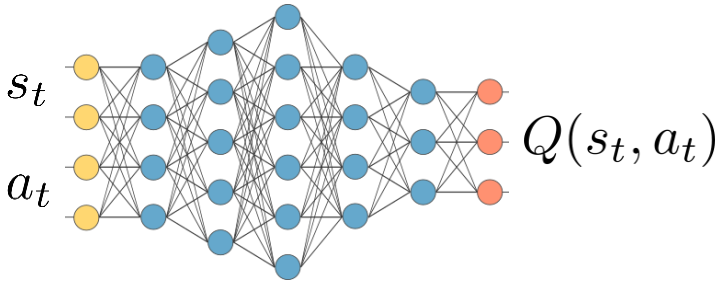


PPO, TRPO

(Schulman et al. ‘15, Schulman et al. ‘17)

Actor-Critic

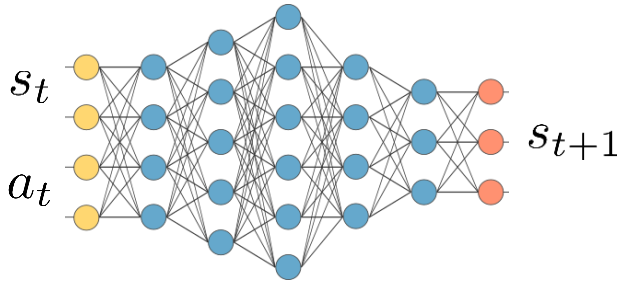




SAC, DDPG

(Haarnoja et al., Lillicrap et al.)

Model-based RL

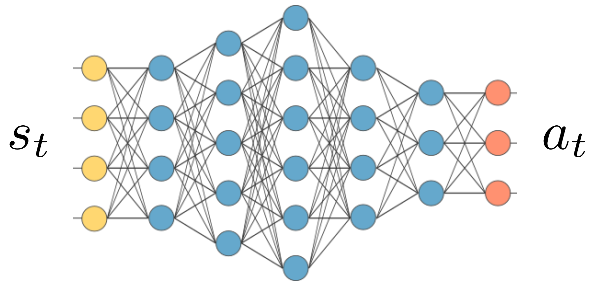


MBMF

(Nagabandi et al.)

Current Approaches to Reinforcement Learning

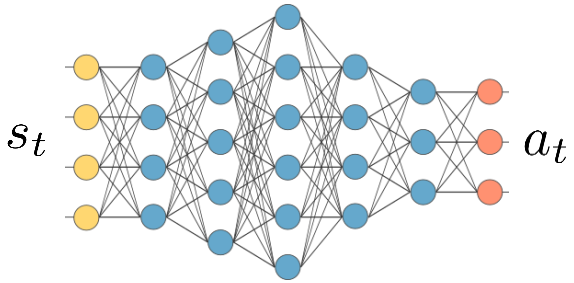
Policy Gradients

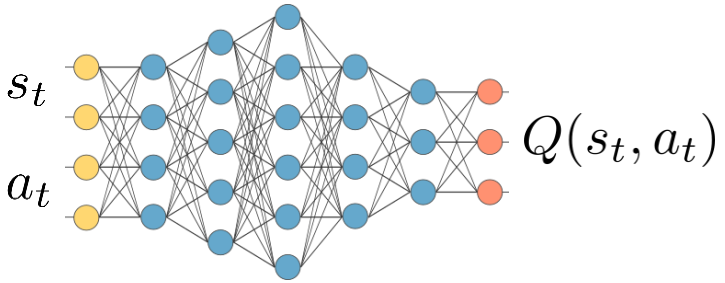


PPO, TRPO

(Schulman et al. ‘15, Schulman et al. ‘17)

Actor-Critic

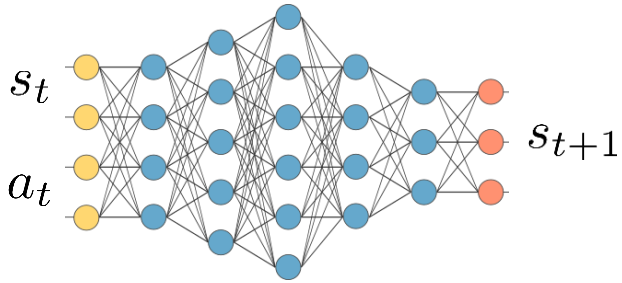




SAC, DDPG

(Haarnoja et al., Lillicrap et al.)

Model-based RL



MBMF

(Nagabandi et al.)

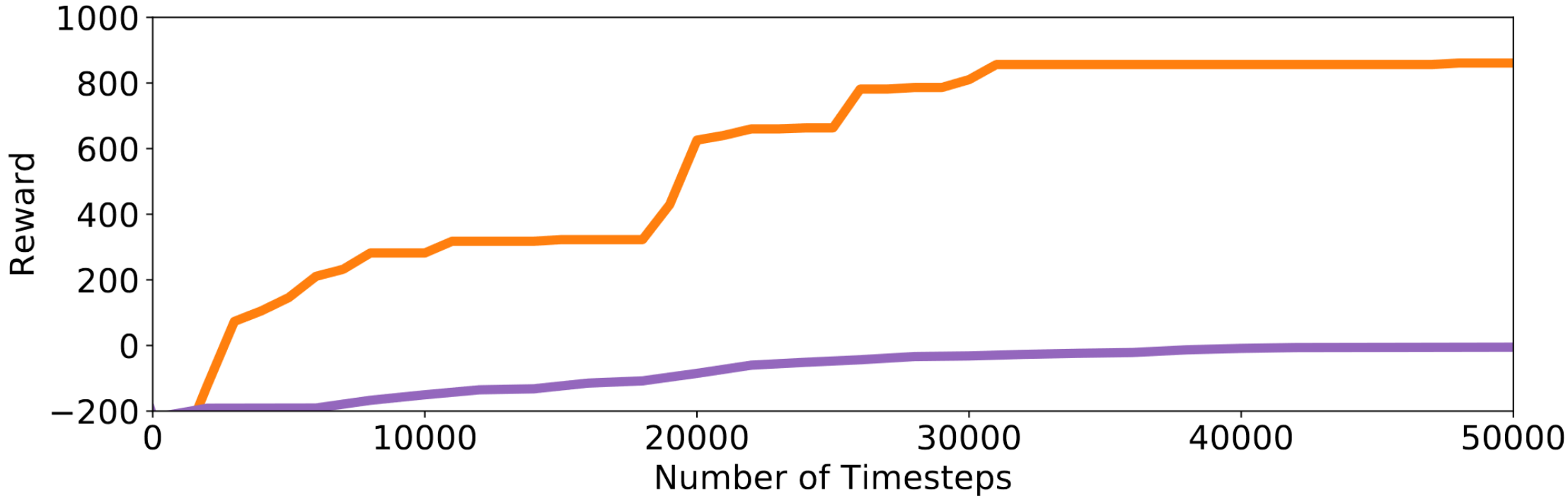
Least Efficient — Most efficient

Data Efficiency

Comparative Performance on HalfCheetah

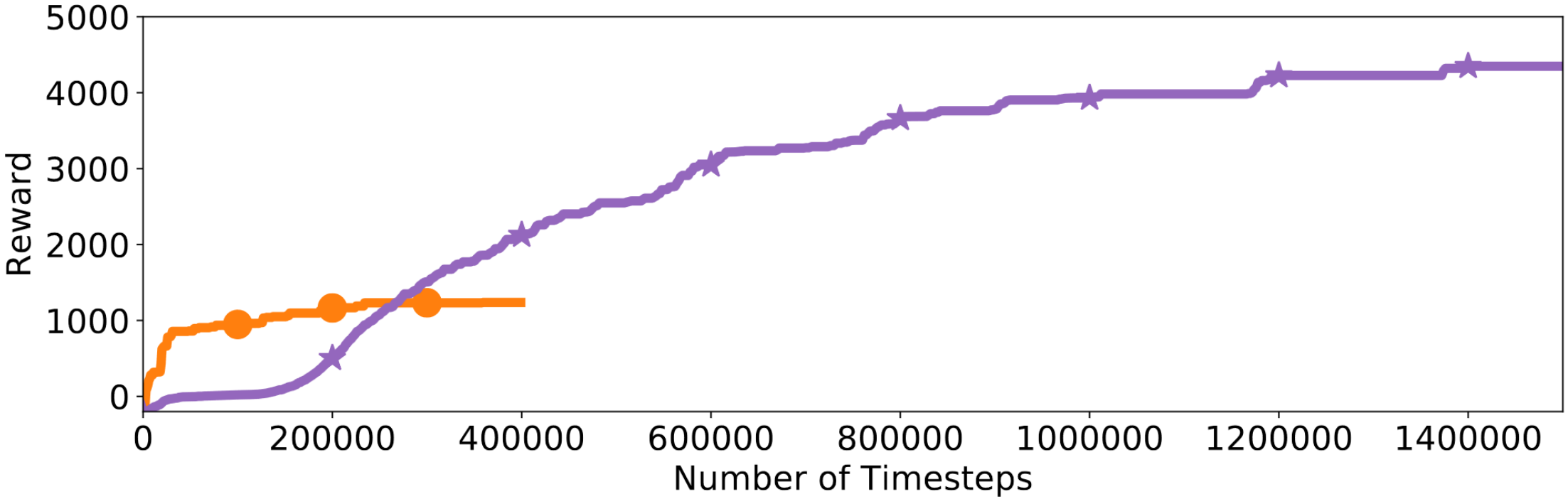
Comparative Performance on HalfCheetah

PPO [Nagabandi et al. 2017]



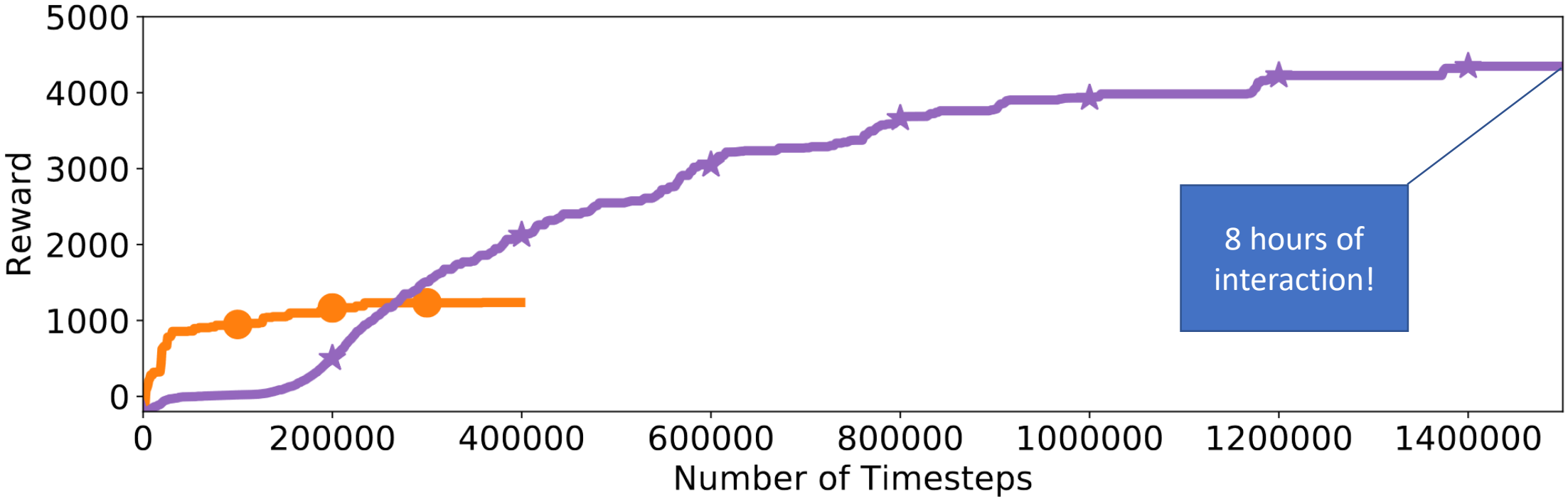
Comparative Performance on HalfCheetah

PPO [Nagabandi et al. 2017]



Comparative Performance on HalfCheetah

PPO [Nagabandi et al. 2017]

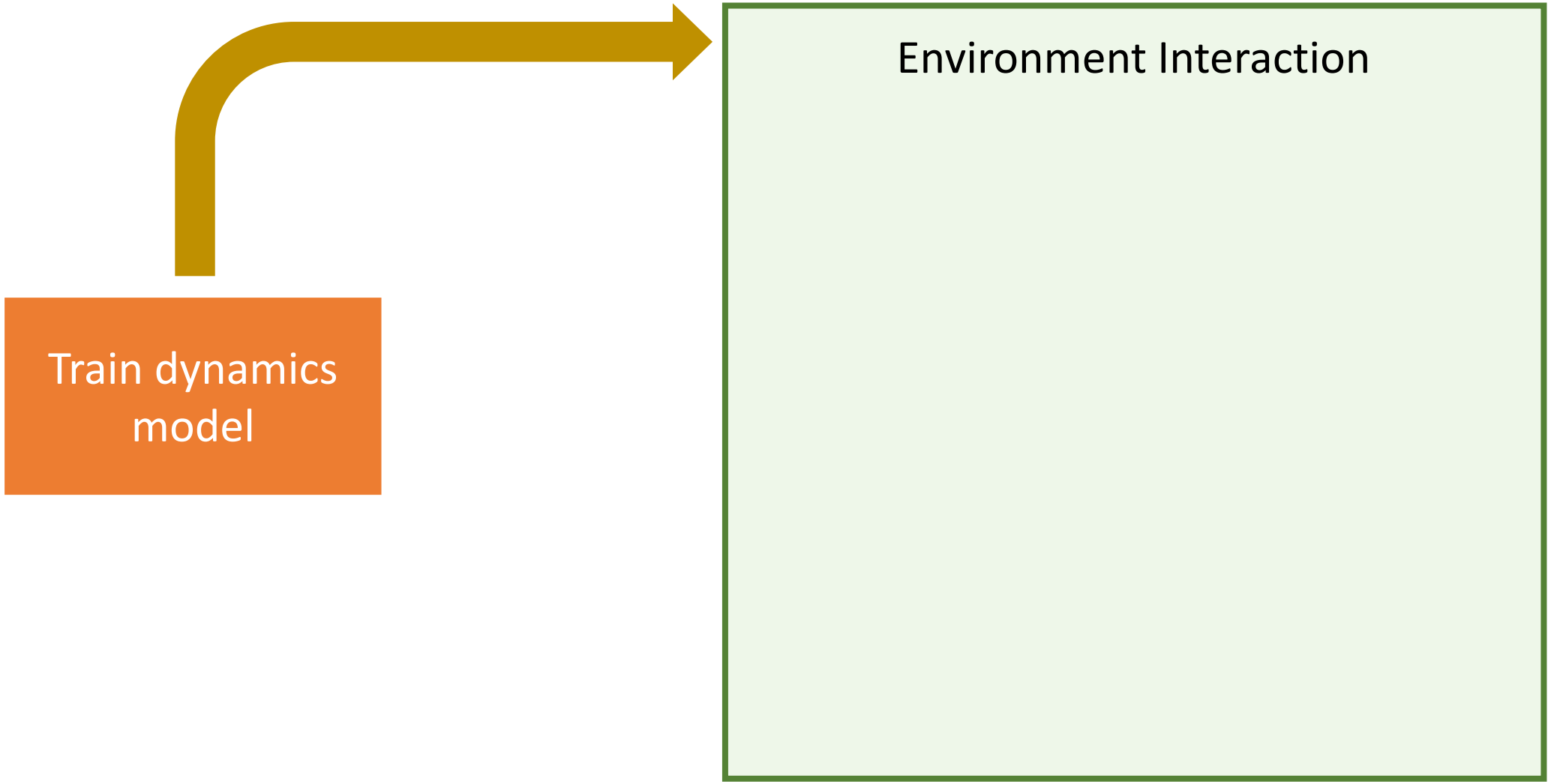


Model-based Reinforcement Learning

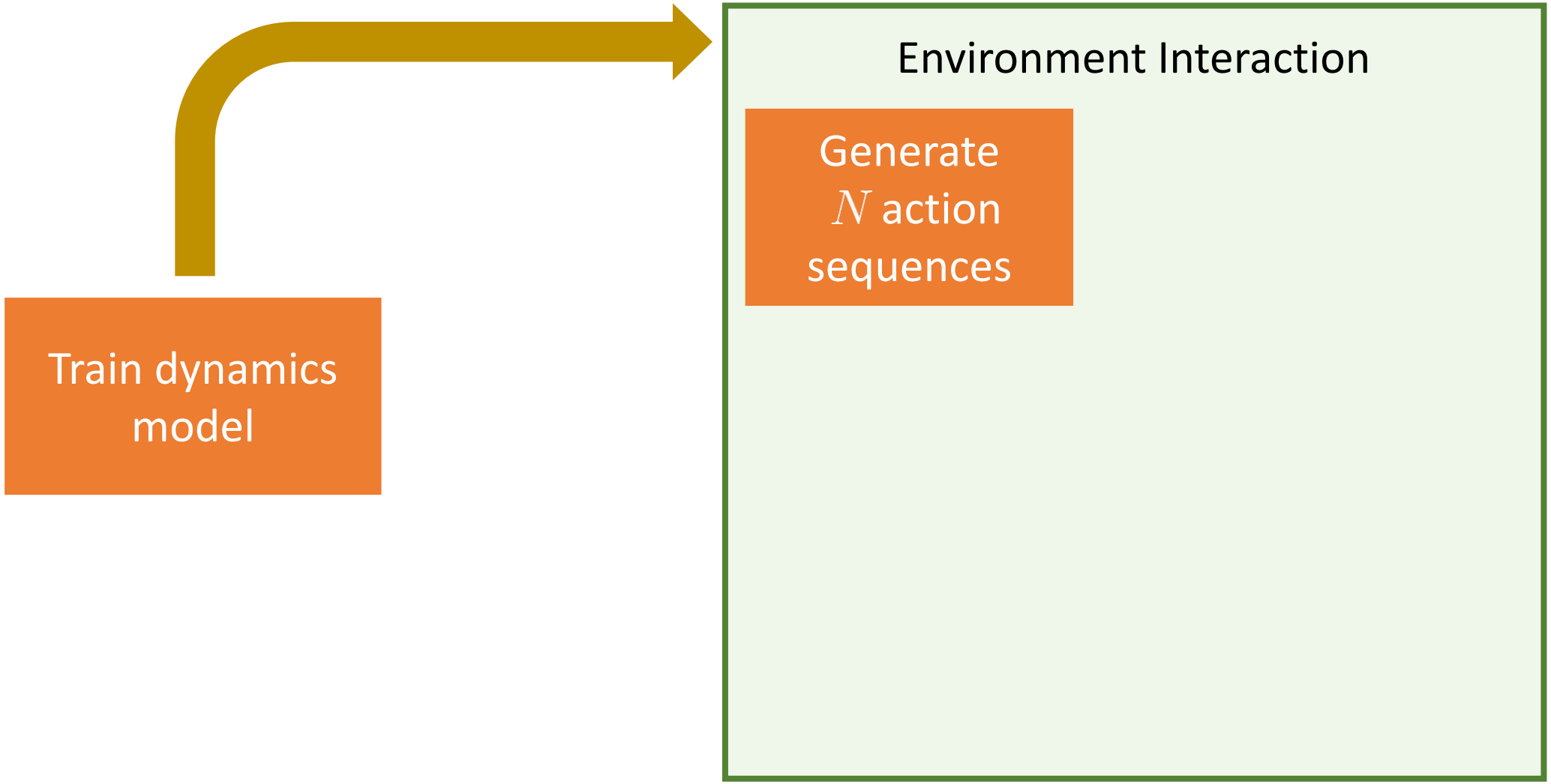
Model-based Reinforcement Learning

Train dynamics model

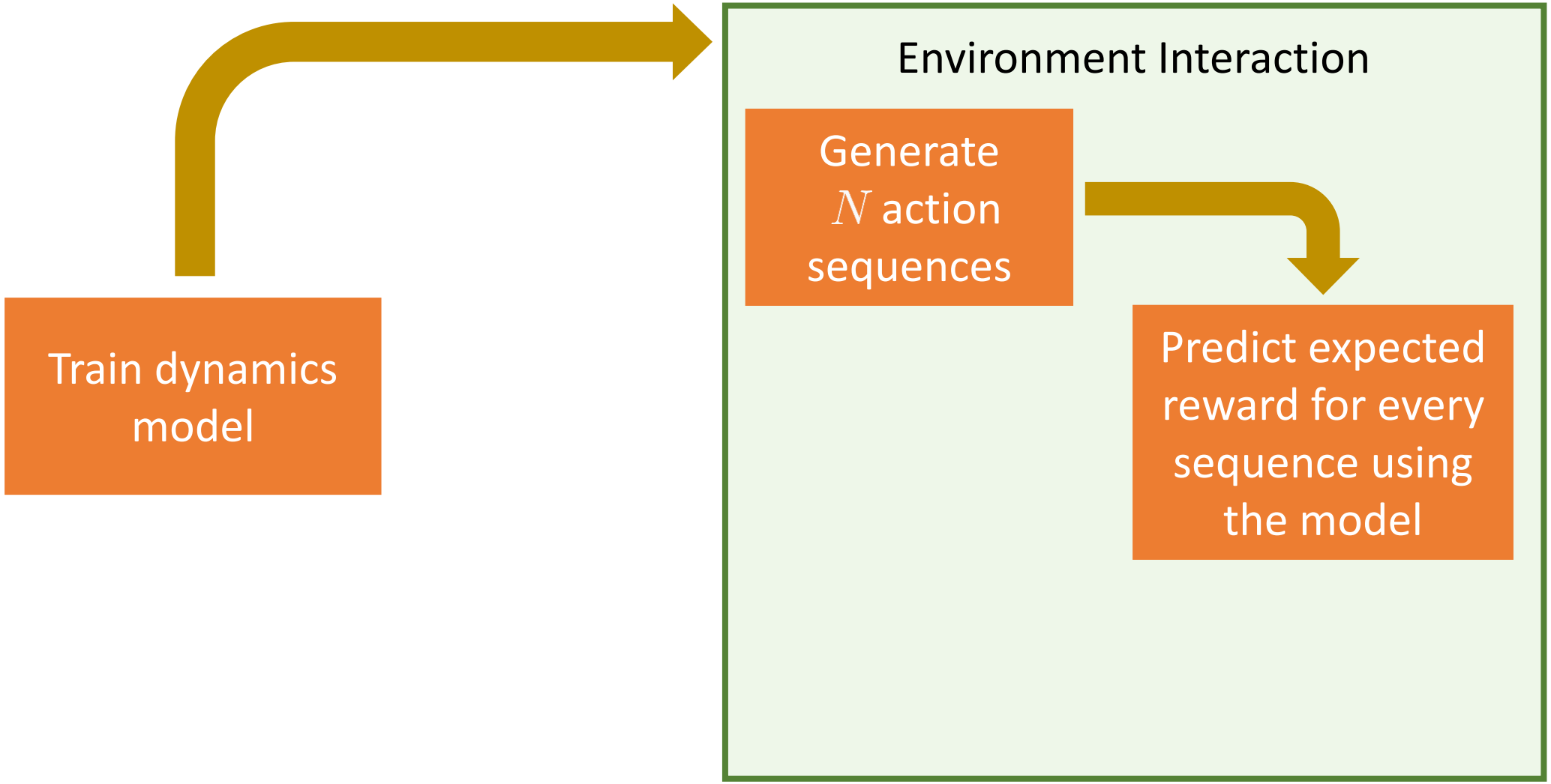
Model-based Reinforcement Learning



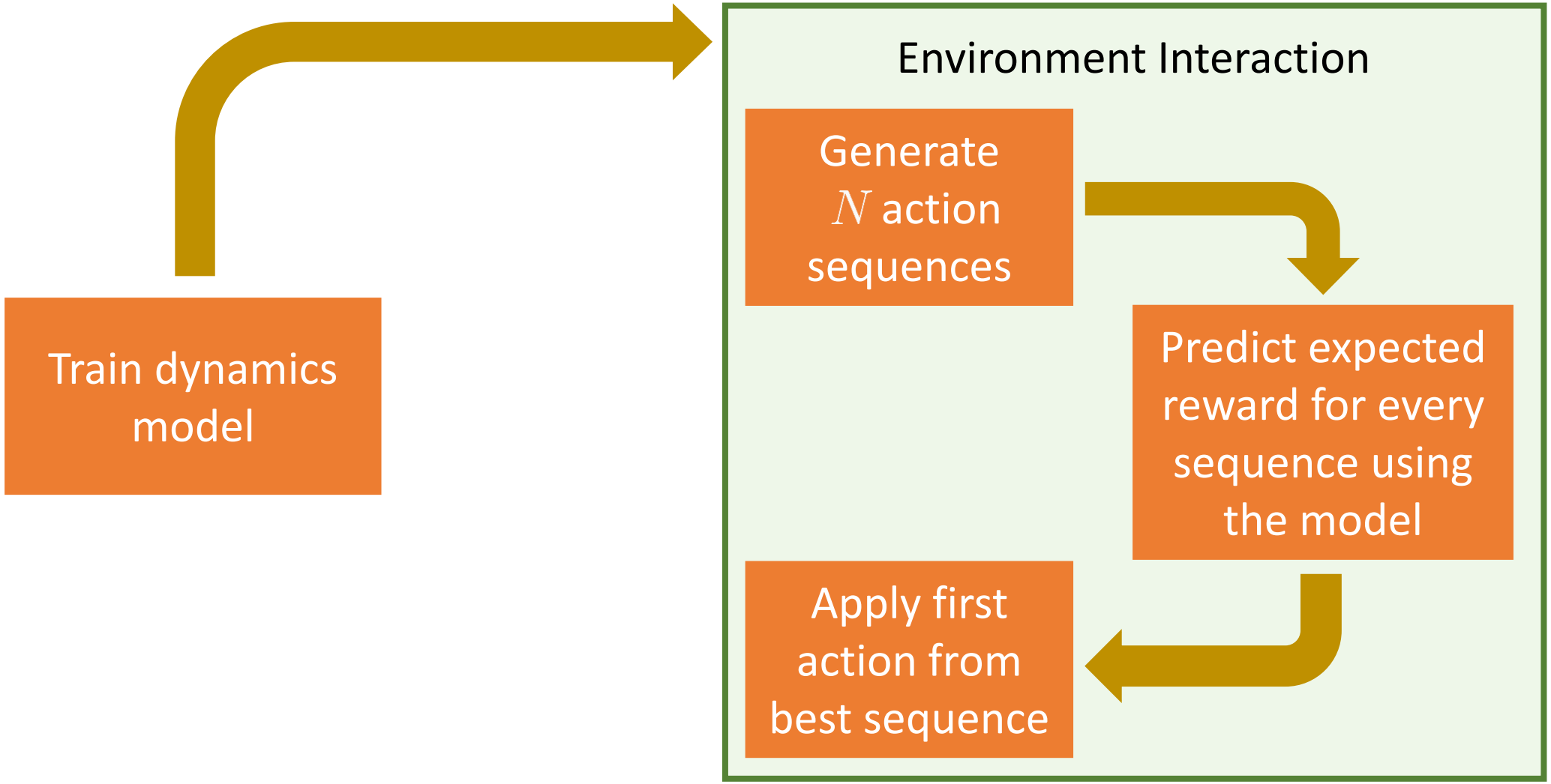
Model-based Reinforcement Learning



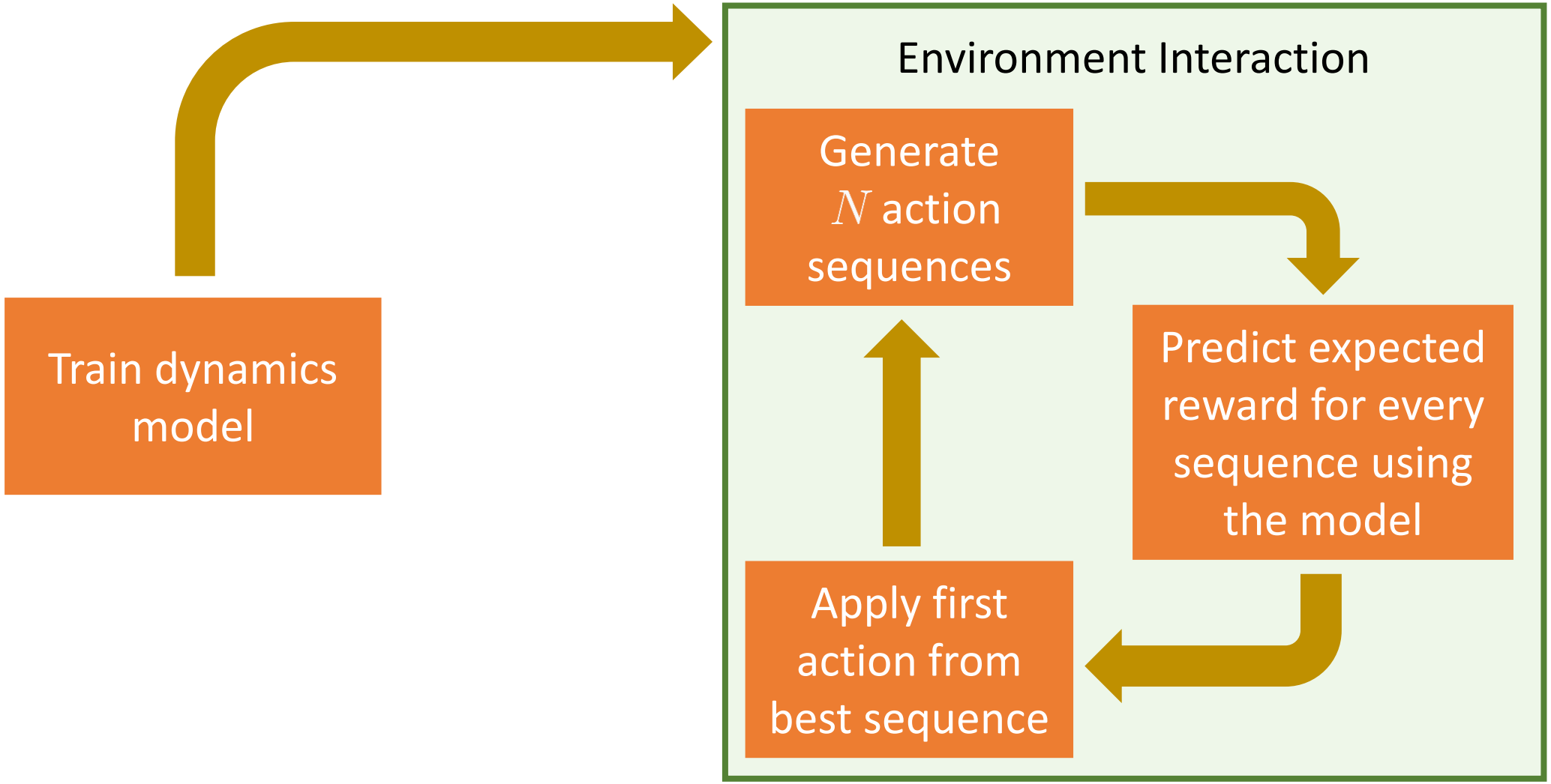
Model-based Reinforcement Learning



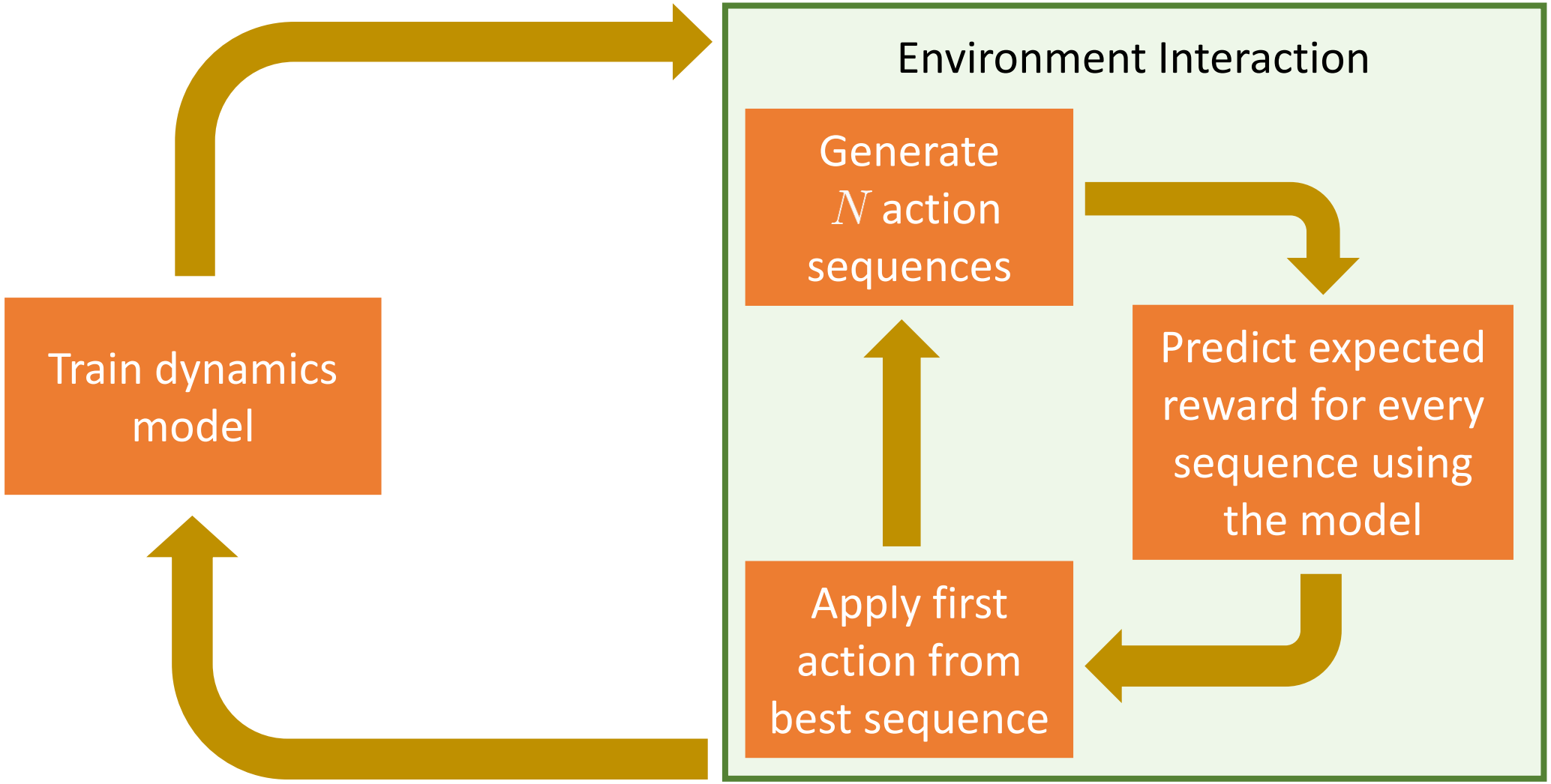
Model-based Reinforcement Learning



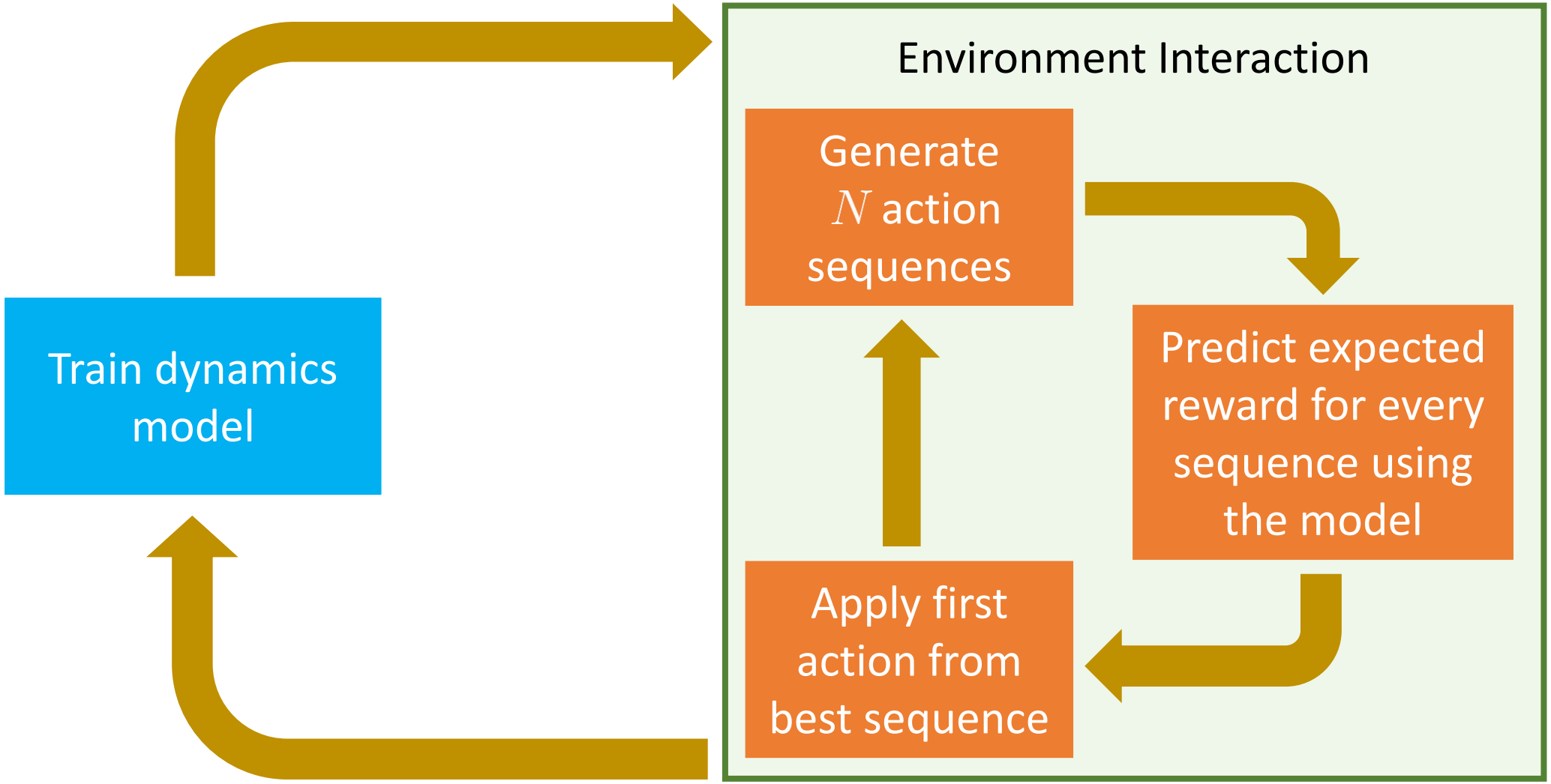
Model-based Reinforcement Learning



Model-based Reinforcement Learning

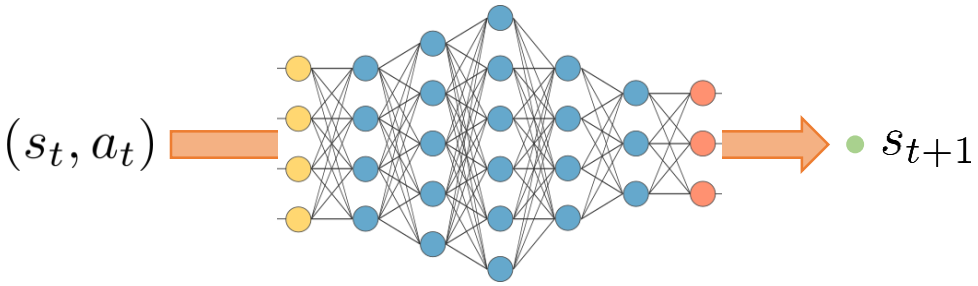


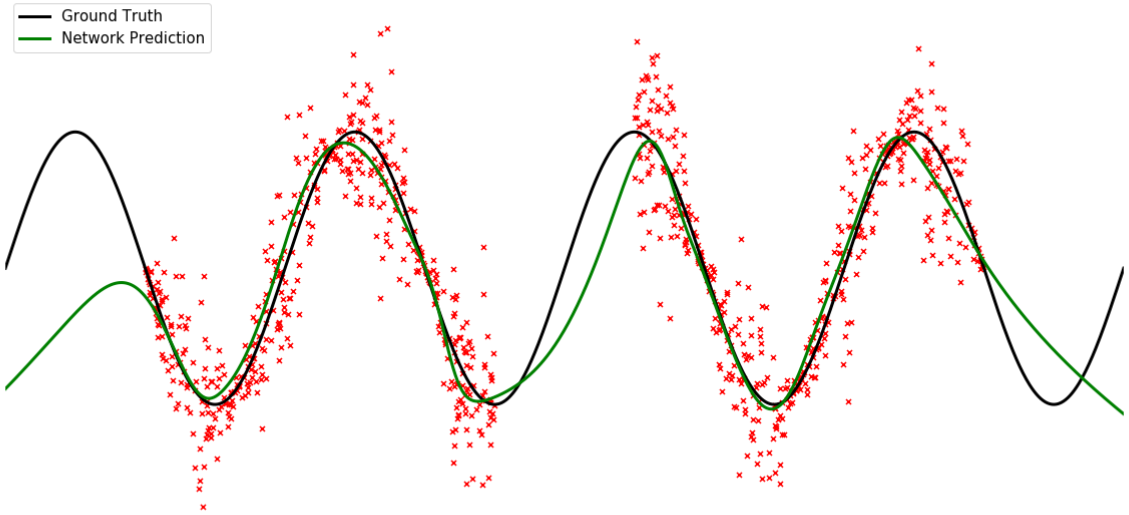
Model-based Reinforcement Learning



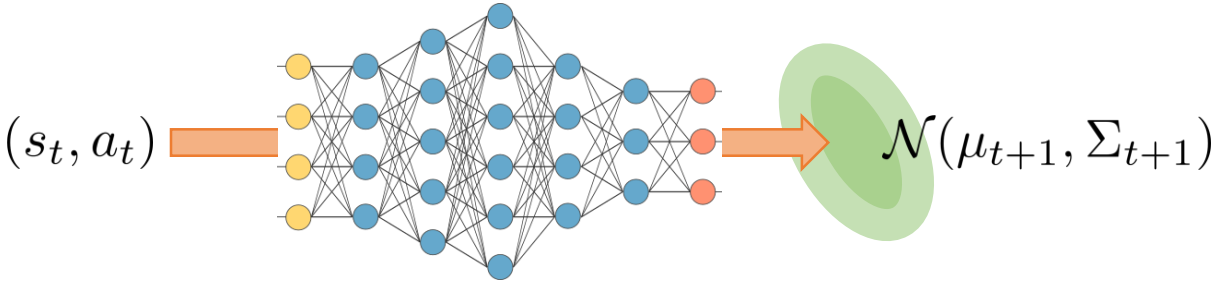
Deterministic Neural Nets as Models

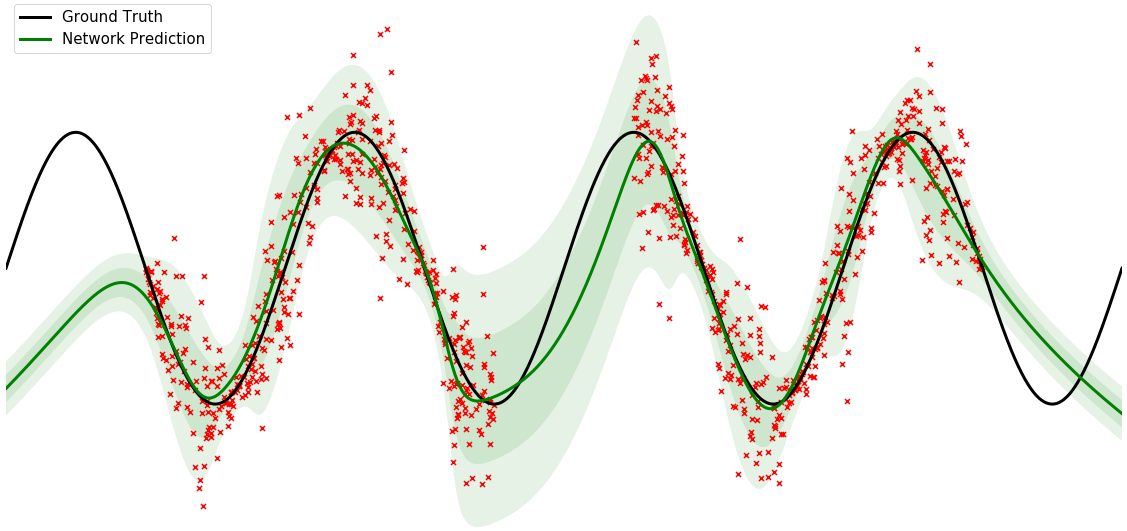
Deterministic Neural Nets as Models



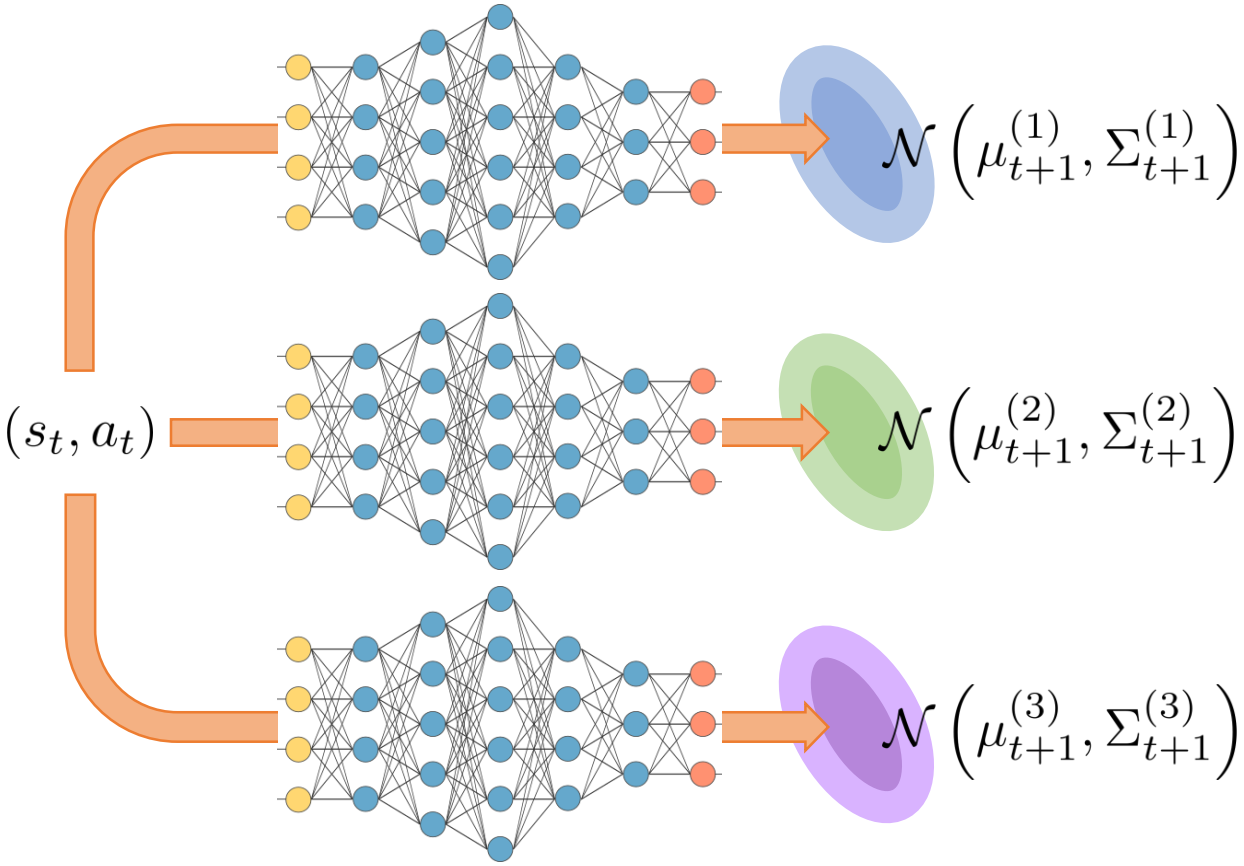


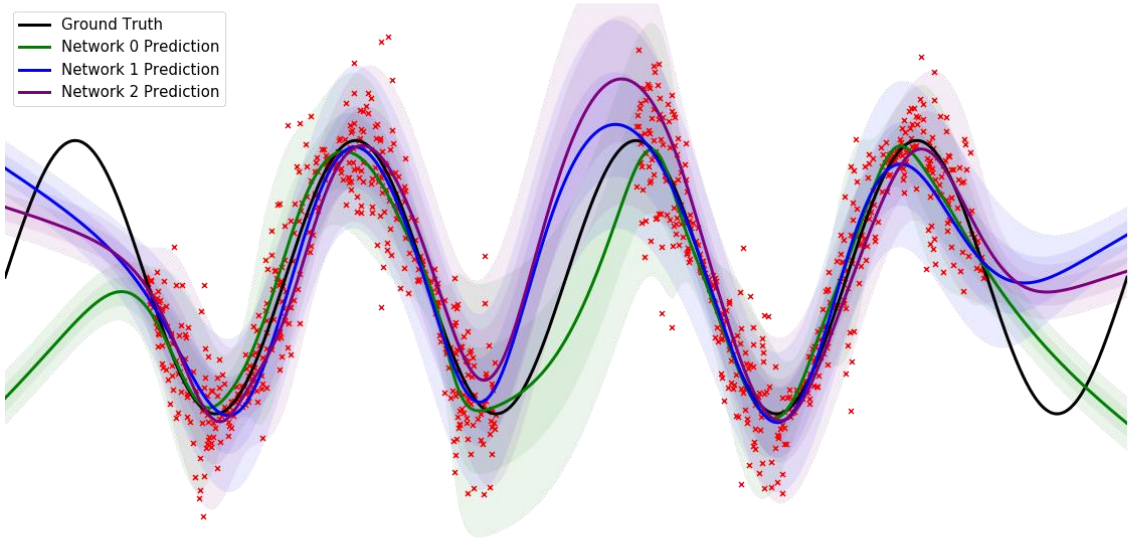
Probabilistic Neural Nets as Models



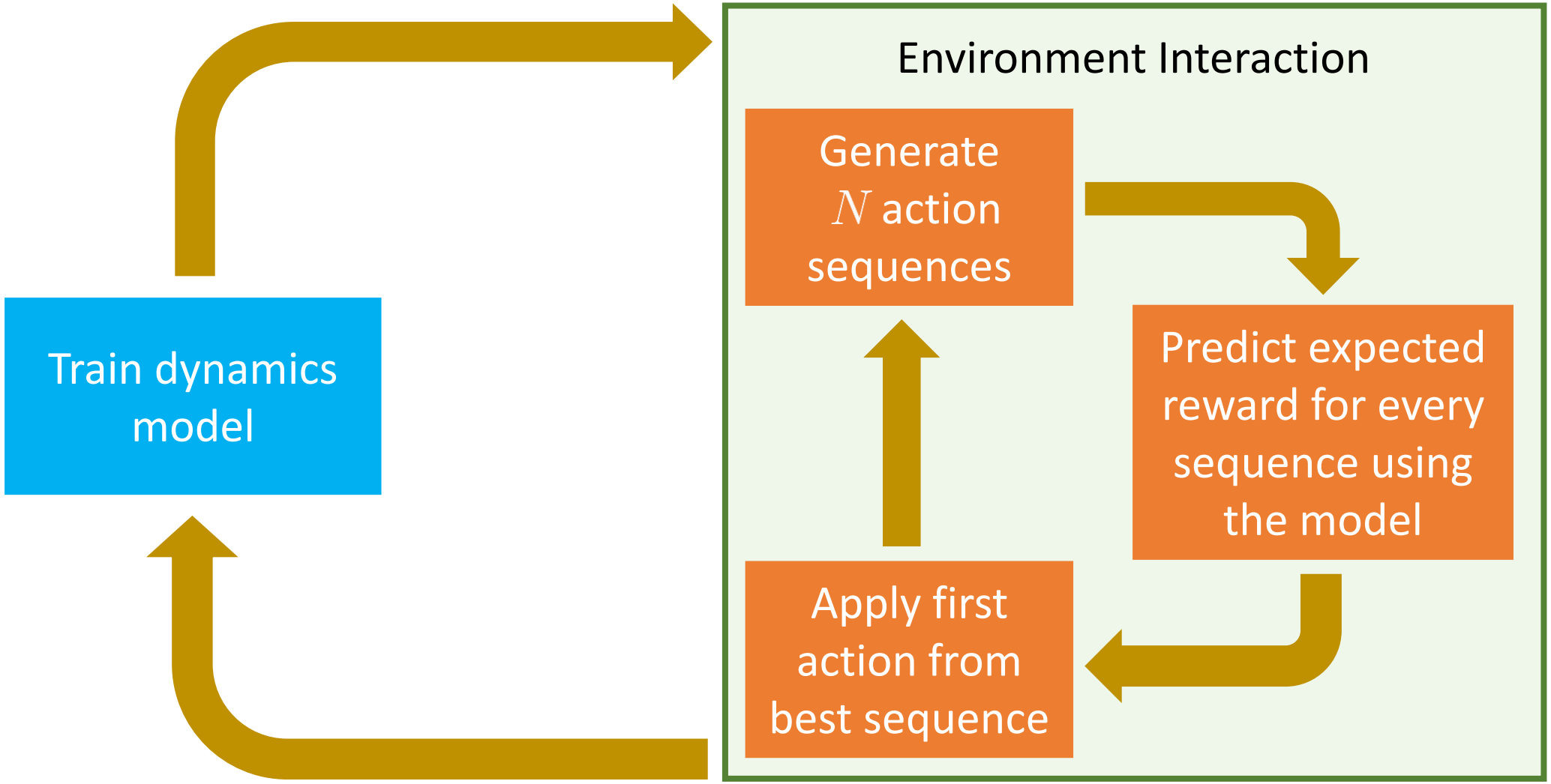


Probabilistic Ensembles as Models

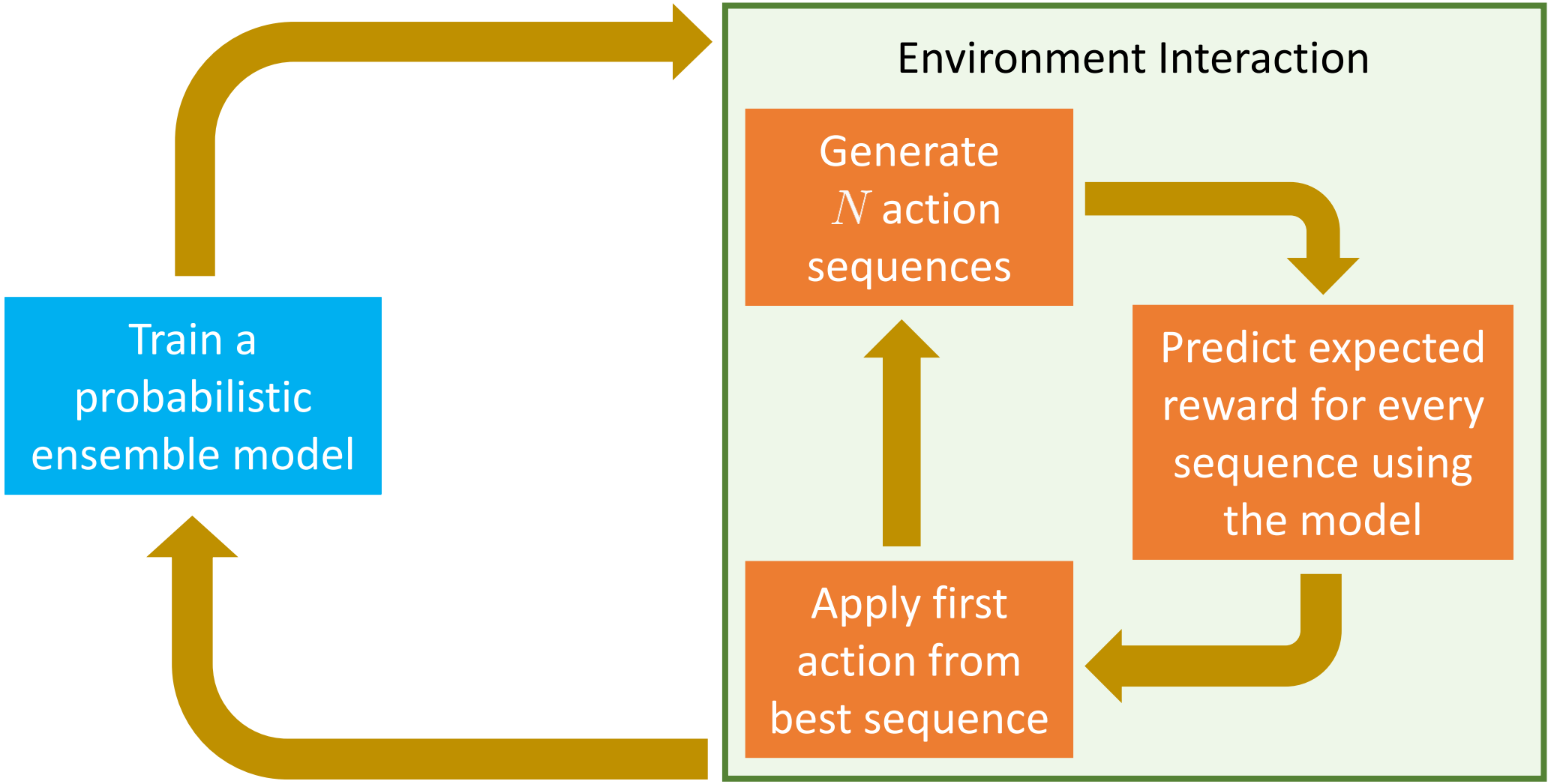




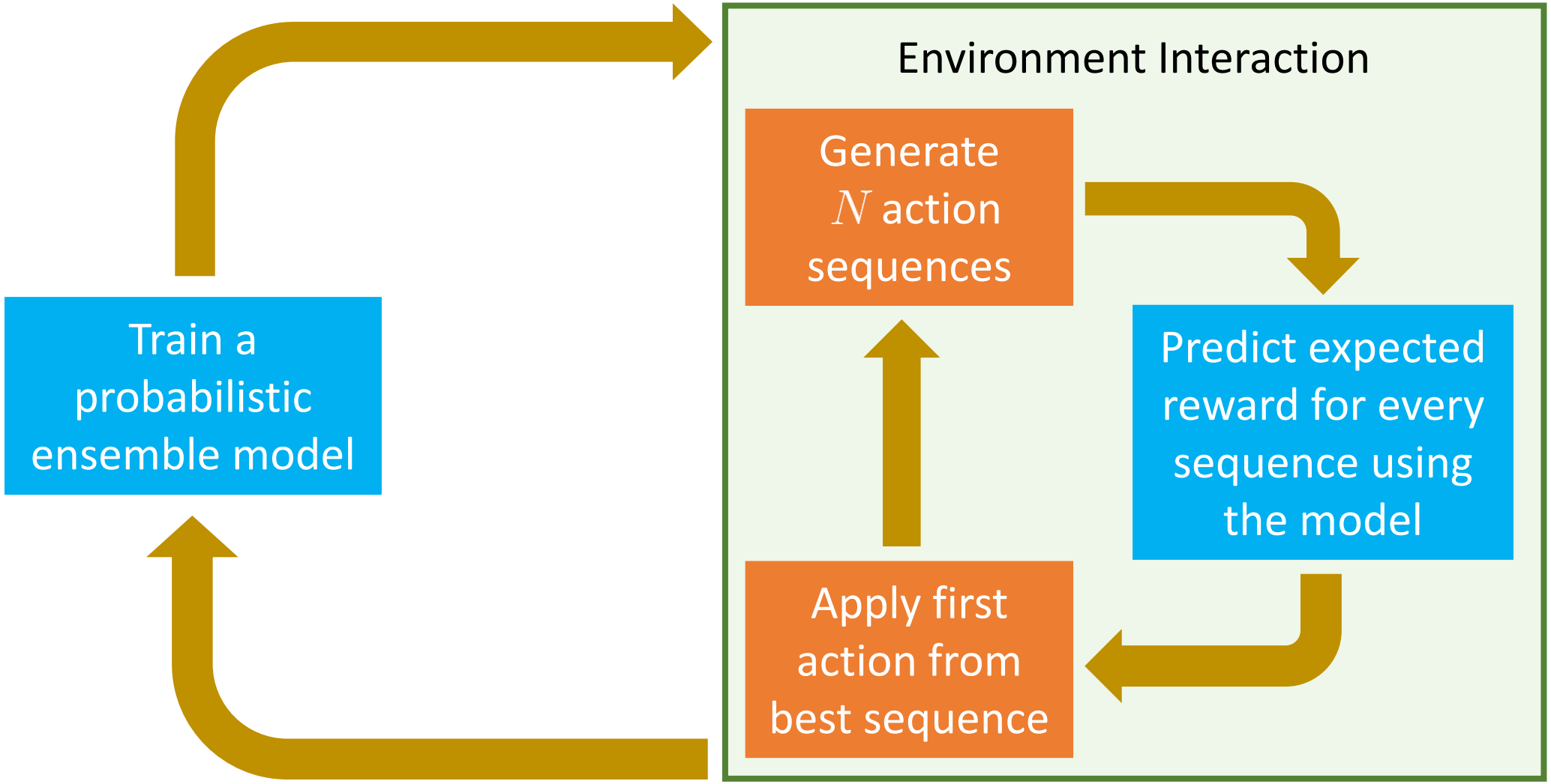
Model-based Reinforcement Learning



Model-based Reinforcement Learning



Model-based Reinforcement Learning



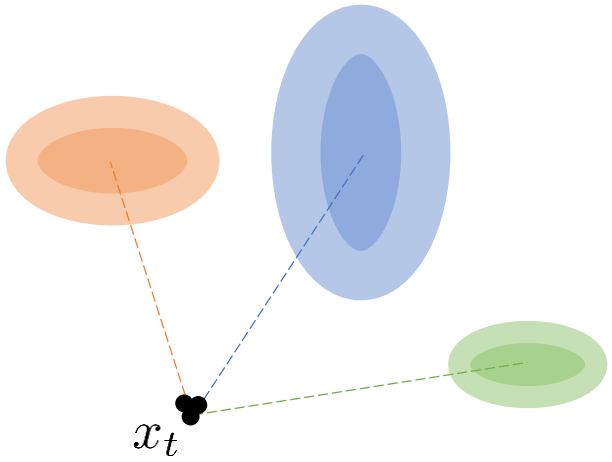
Trajectory Sampling for State Propagation

Trajectory Sampling for State Propagation

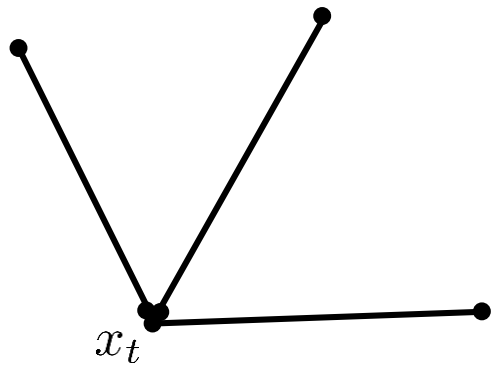
x_t

Trajectory Sampling for State Propagation

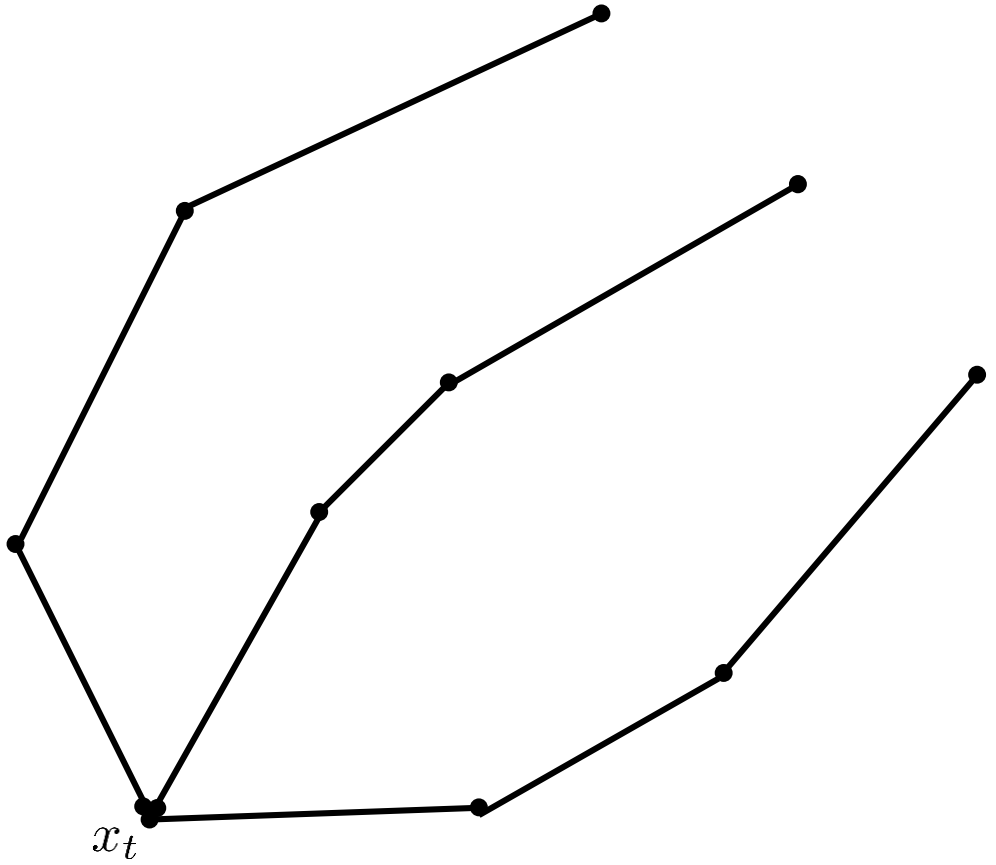
Trajectory Sampling for State Propagation



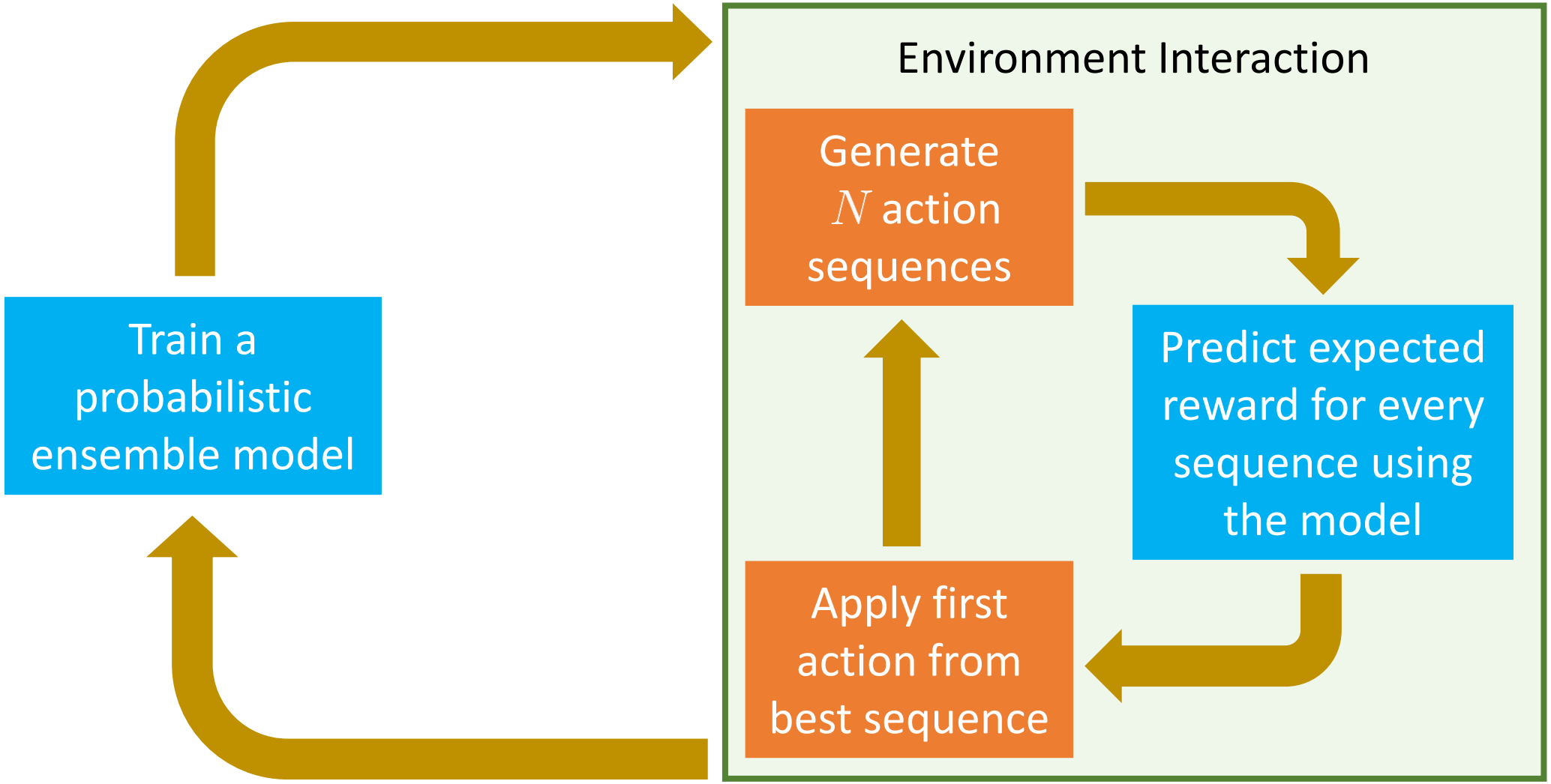
Trajectory Sampling for State Propagation



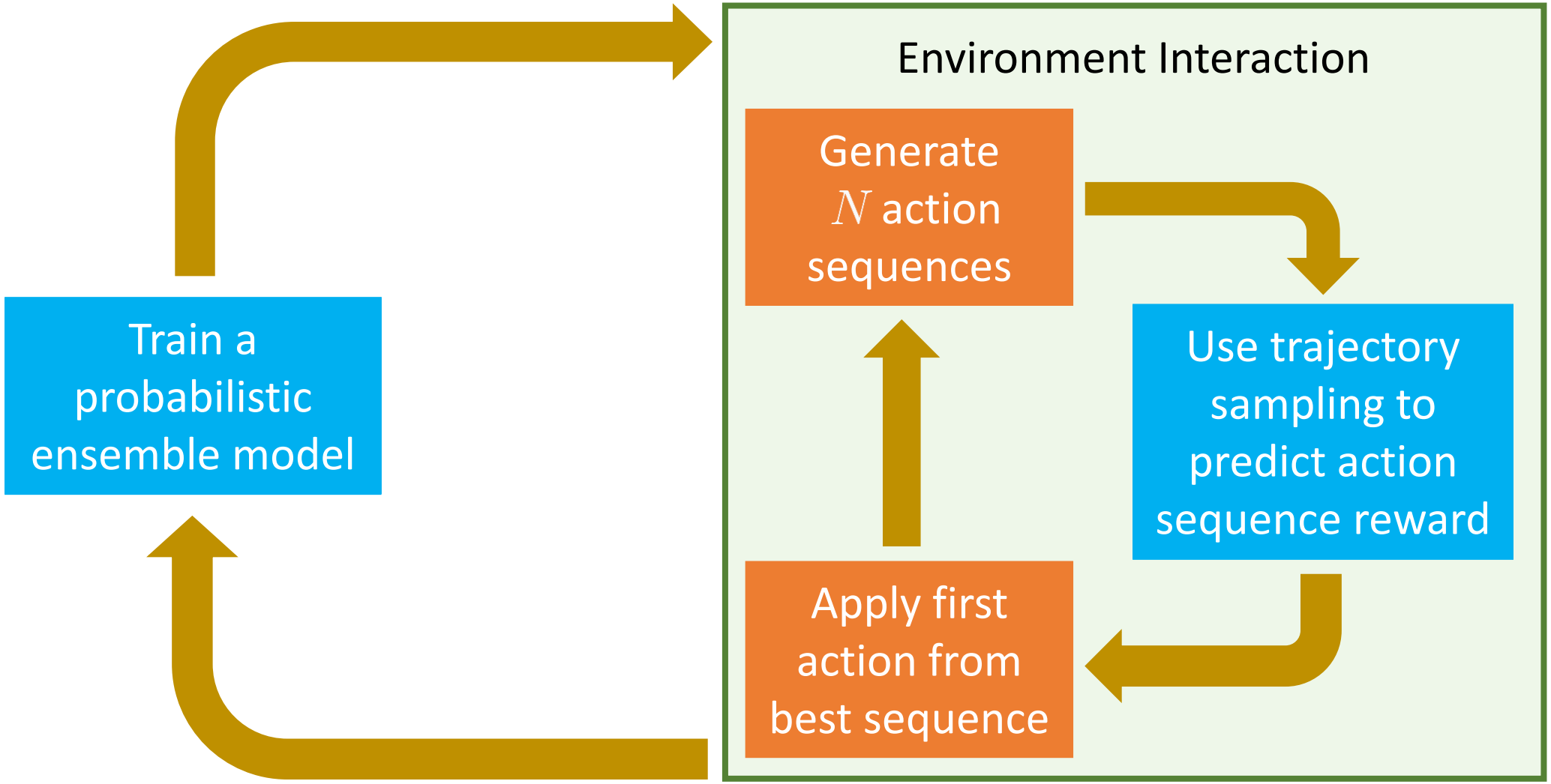
Trajectory Sampling for State Propagation



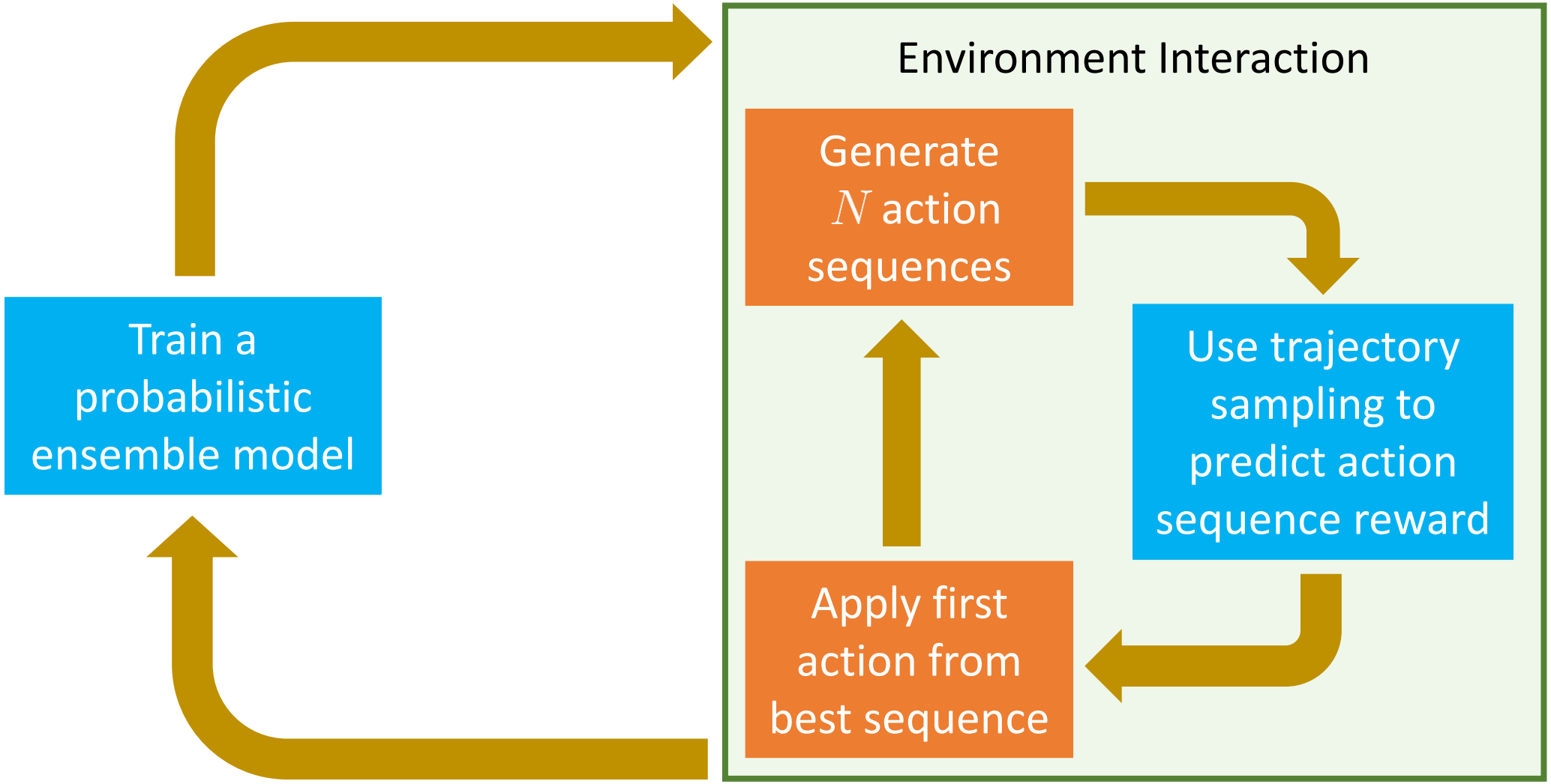
Model-based Reinforcement Learning



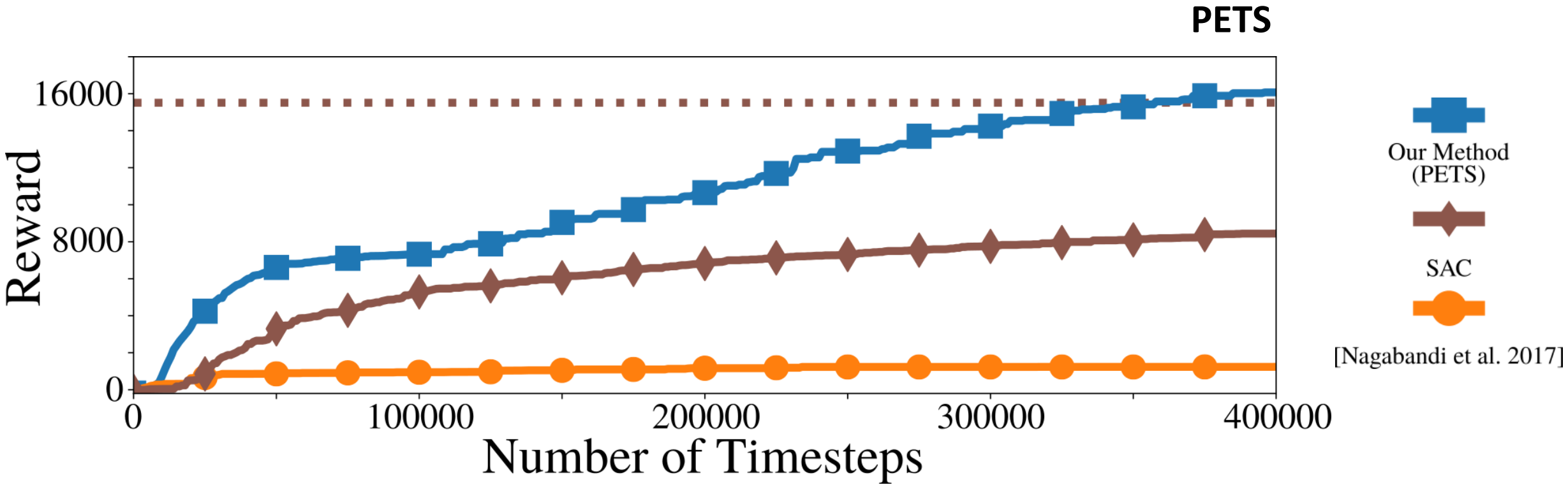
Model-based Reinforcement Learning



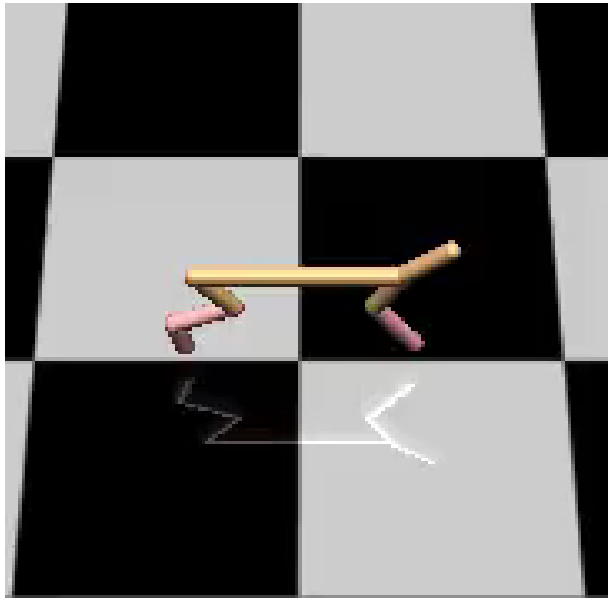
Our Method: Probabilistic Ensembles with Trajectory Sampling (PETS)



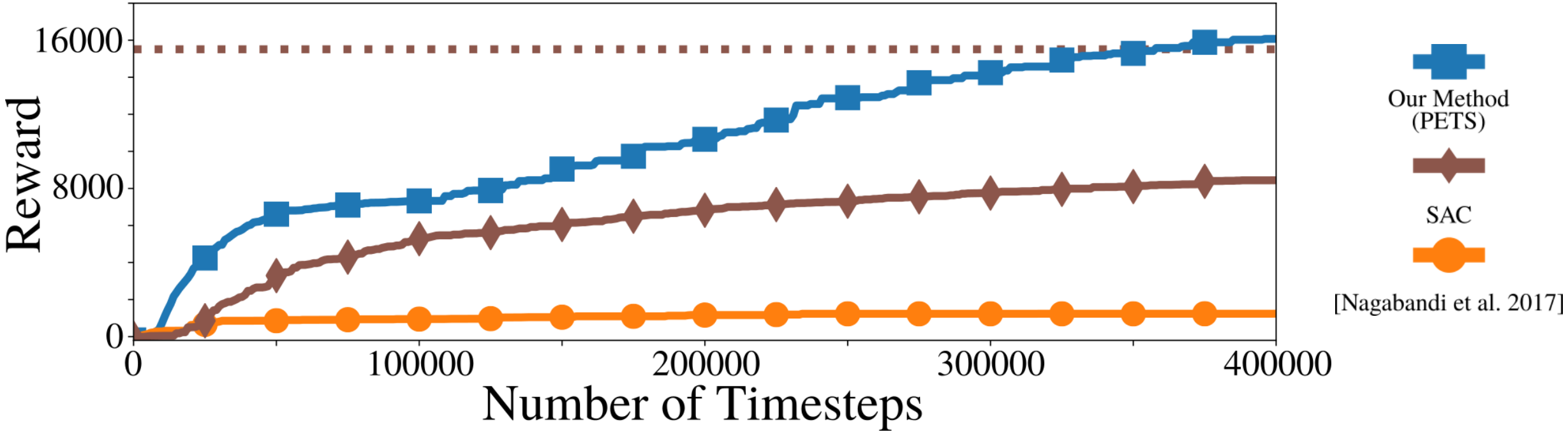
Experimental Results



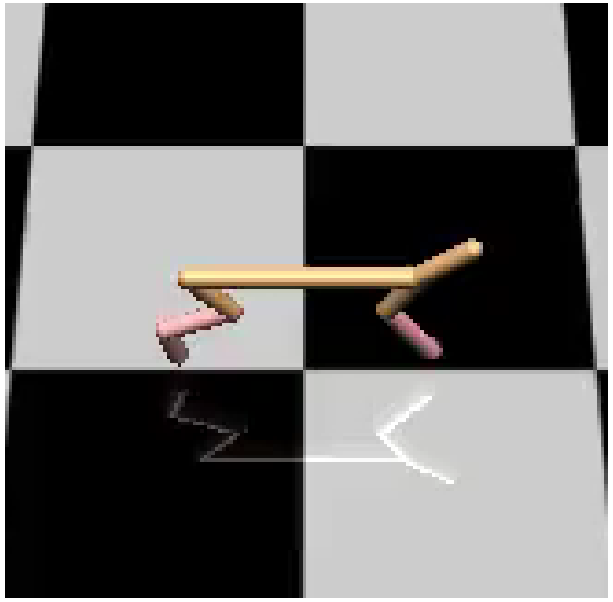
Experimental Results



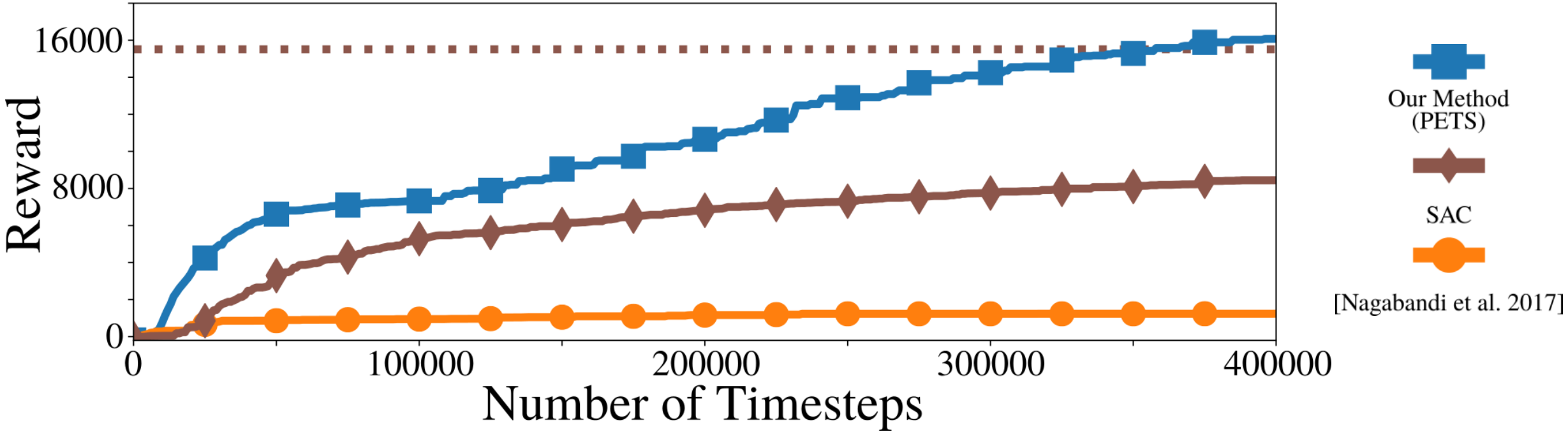
PETS



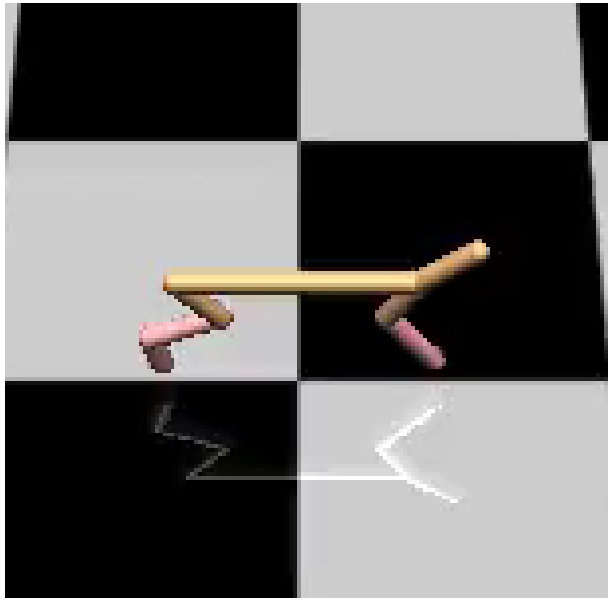
Experimental Results



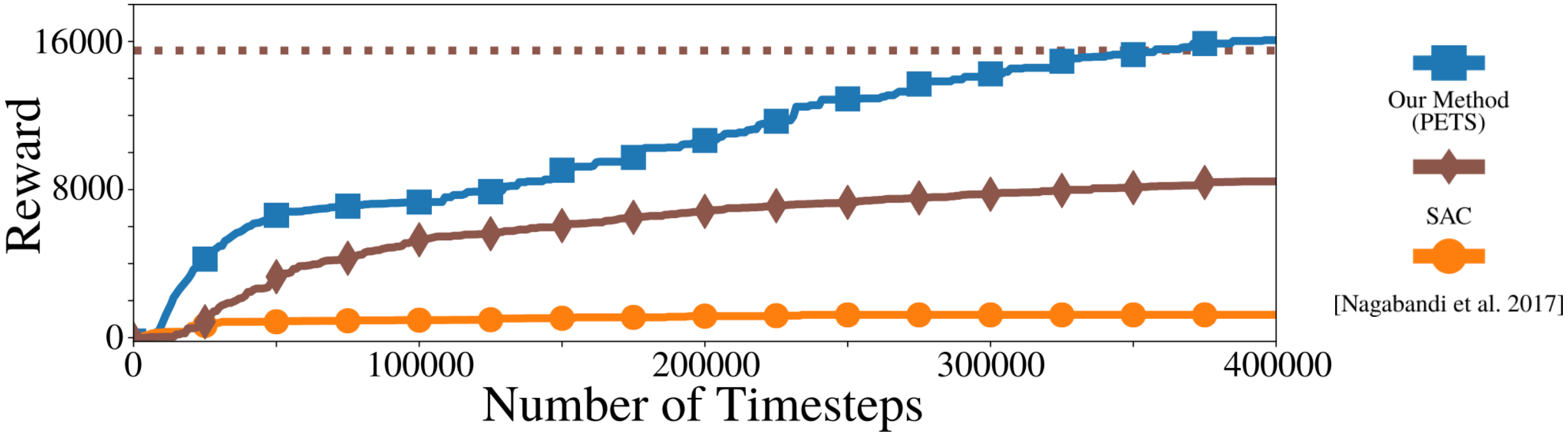
PETS



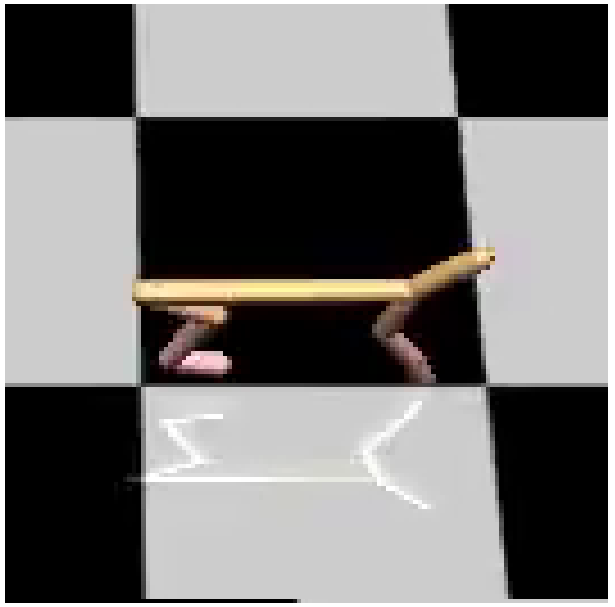
Experimental Results



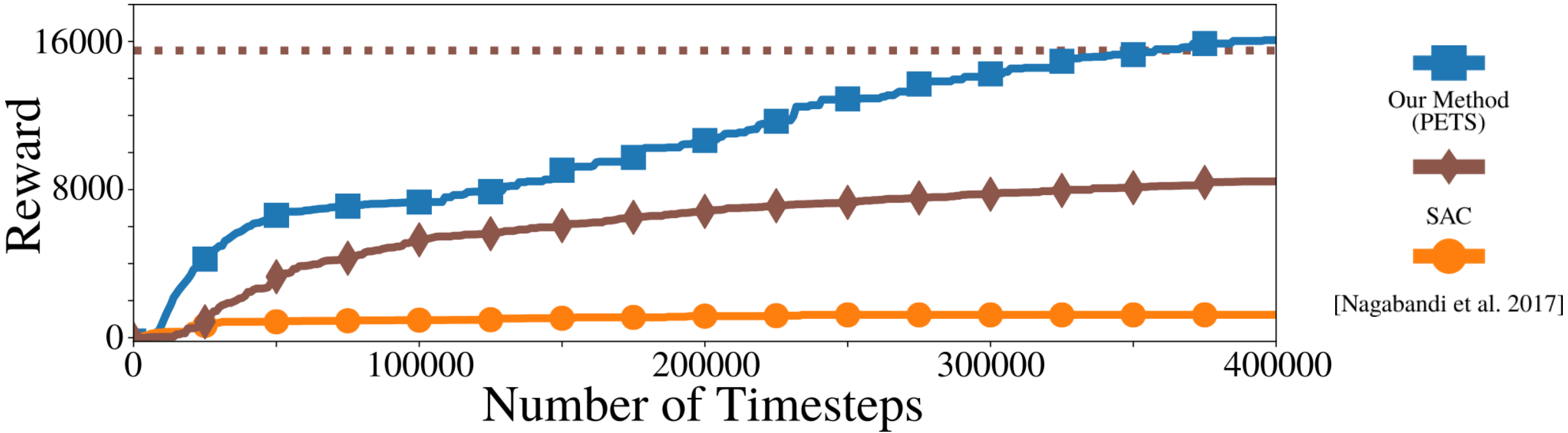
PETS



Experimental Results



PETS



Why use PETS?

Why use PETS?

- ✓Data efficiency

Why use PETS?

- ✓Data efficiency
- ✓Competitive asymptotic performance

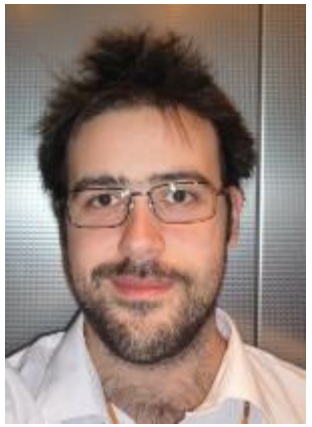
Why use PETS?

- ✓Data efficiency
- ✓Competitive asymptotic performance
- ✓Ease of implementation

Deep Reinforcement Learning in a Handful of Trials with Probabilistic Dynamics Models

Code: https://github.com/kchua/handful-of-trials
Website: https://sites.google.com/view/drl-in-a-handful-of-trials
Paper: https://arxiv.org/abs/1805.12114


Kurtland Chua


Roberto Calandra


Rowan McAllister


Sergey Levine