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

Kurtland Chua, Roberto Calandra, Rowan McAllister, Sergey Levine

University of California, Berkeley
How Long Does Learning Take?

~50 million frames

[Mnih et al. 2015]

~800,000 grasp attempts

~21 million games

[Silver et al. 2017]

[Levine et al. 2017]
Can we speed this up?
Model-Based Reinforcement Learning

- Train Dynamics Model
- Optimize Policy
- Execute Policy
Comparative Performance on HalfCheetah

![Graph showing performance comparison over time]
Comparative Performance on HalfCheetah
Deterministic Neural Nets as Models
Deterministic Neural Nets as Models
Deterministic Neural Nets as Models

\[(s_t, a_t) \rightarrow s_{t+1}\]
Deterministic Neural Nets as Models

\[(s_t, a_t) \rightarrow s_{t+1}\]
Deterministic Neural Nets as Models
Probabilistic Neural Nets as Models

\((s_t, a_t) \rightarrow p(s_{t+1})\)
Probabilistic Ensembles as Models
Probabilistic Ensembles as Models
Trajectory Sampling for State Propagation
Trajectory Sampling for State Propagation

$S_t$
Trajectory Sampling for State Propagation

$S_t$
Trajectory Sampling for State Propagation
Trajectory Sampling for State Propagation
Trajectory Sampling for State Propagation

$S_t$
Trajectory Sampling for State Propagation
Trajectory Sampling for State Propagation
Trajectory Sampling for State Propagation
Trajectory Sampling for State Propagation
Experimental Results

![Experimental Results Image]

- **Reward vs Interaction Time**

  - **Our Method**
  - **SOTA Model-Based (Nagabandi et al. 2017)**
  - **SOTA Model-Free (Haarnoja et al. 2018)**
  - **SOTA Model-Free at convergence**

The graph shows the reward over time for different methods, indicating how each method performs in terms of reward gain as interaction time increases. The x-axis represents interaction time in minutes, and the y-axis represents reward. The legend at the bottom of the graph provides a key for interpreting the different data series.
Deep Reinforcement Learning in a Handful of Trials using Probabilistic Dynamics Models

Poster #165

Code:  https://github.com/kchua/handful-of-trials
Website:  https://sites.google.com/view/drl-in-a-handful-of-trials

✓ Data efficient
✓ Competitive asymptotic performance
✓ Easy to implement

Kurtland Chua  Roberto Calandra  Rowan McAllister  Sergey Levine