# Learning from Demonstrations via Capability-Aware Goal Sampling

Yuanlin Duan, Yuning Wang, Wenjie Qiu and He Zhu

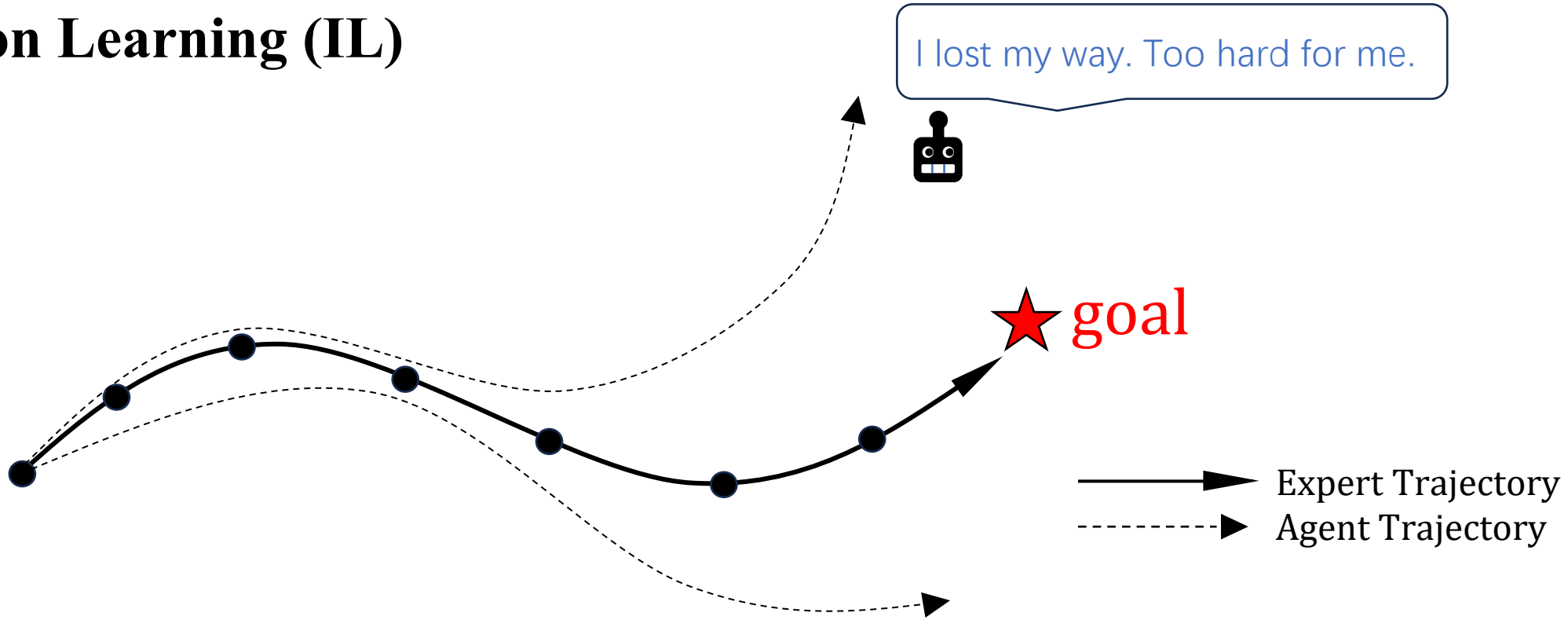
**Cago Github**

RUTGERS THE STATE UNIVERSITY OF NEW JERSEY
NEURAL INFORMATION PROCESSING SYSTEMS

# Problem

***Previous Learning from demonstrations methods have severe limitations.***

- **Imitation Learning (IL)**

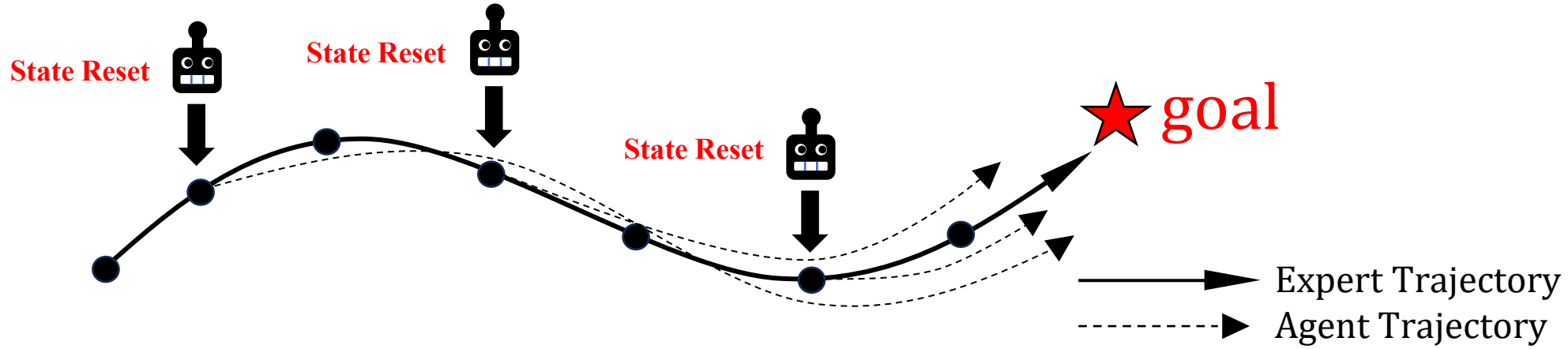



**IL Problem**:

1. Compounding errors
2. Distribution flat matching without considering the agent's evolving capabilities.

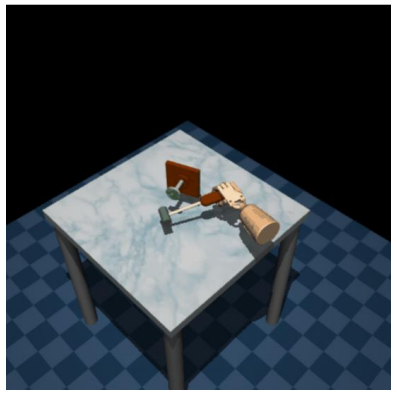
# Problem

- **Curriculum Learning (CL) & State Reset**

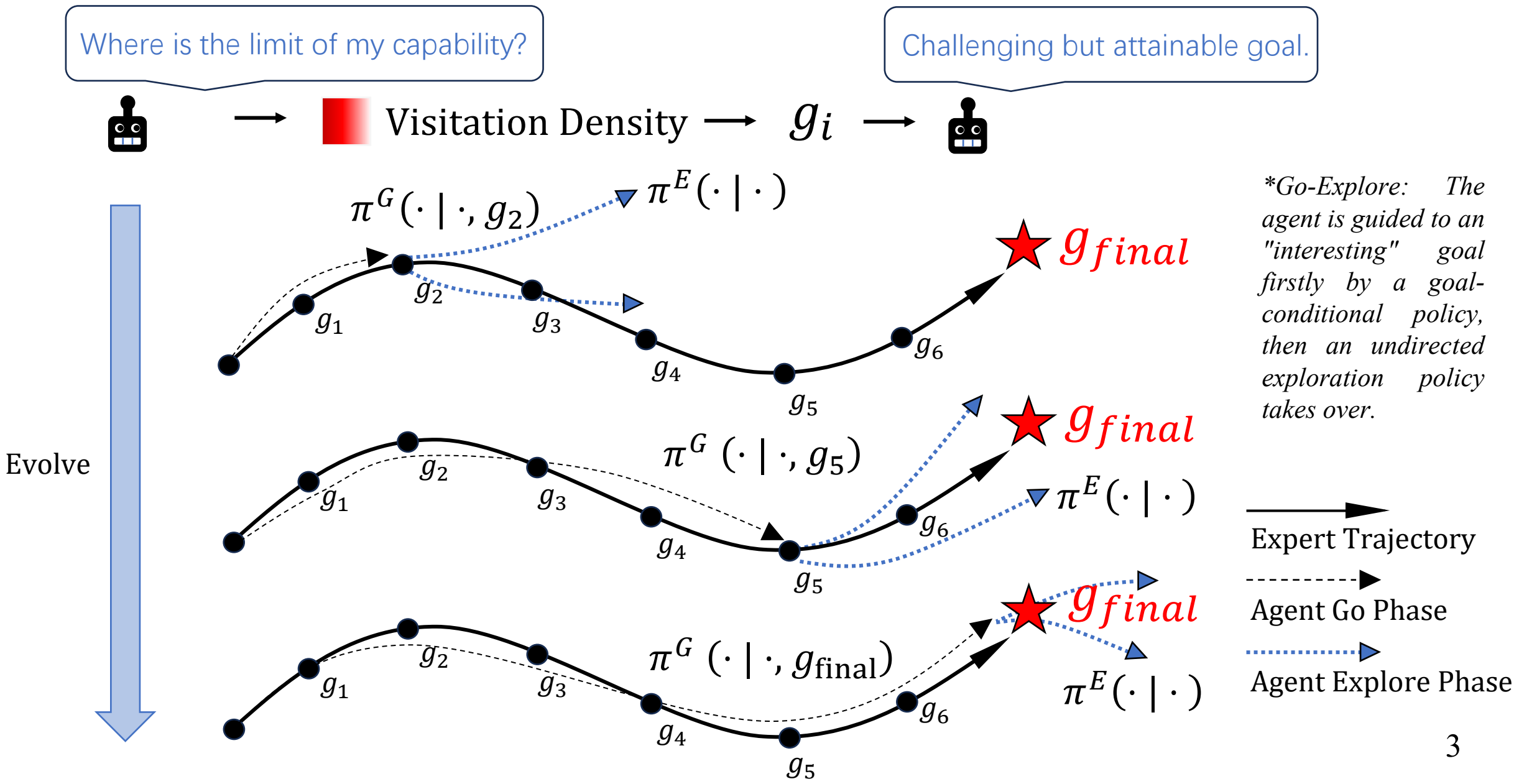



**CL & State Reset Problem**:

Impractical in real-world settings due to challenges in replicating physical conditions like joint velocities and angular momentum.

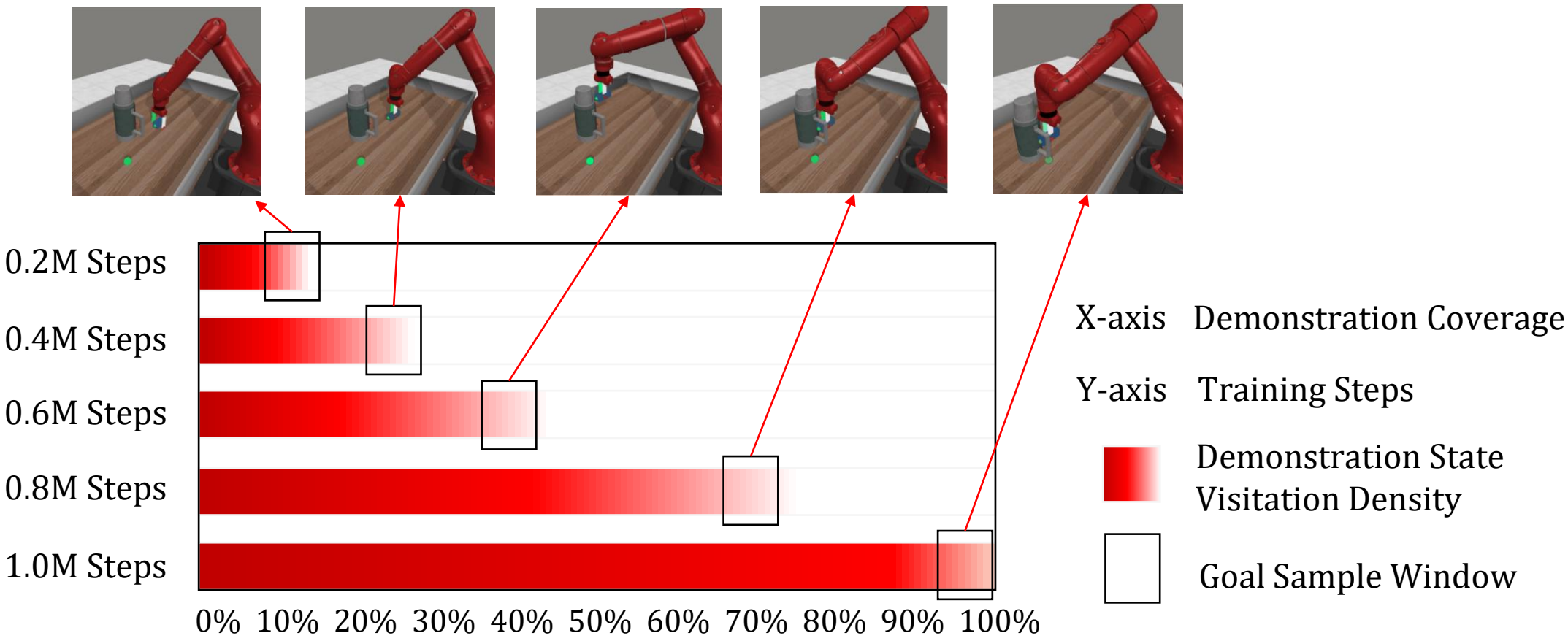


# Illustration of the Cago Framework (our method)



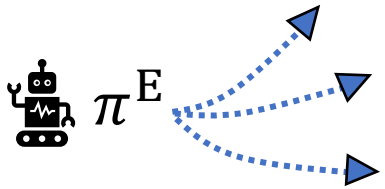
# Method - Cago

- **Goal Sampling Progress for $\pi^G$**



# Method - Cago

- **Behavior-Clone Explorer $\pi^E$**

$\mathcal{D}_{demo}$ —Behavior Clone→  $\pi^{\mathrm{E}}$

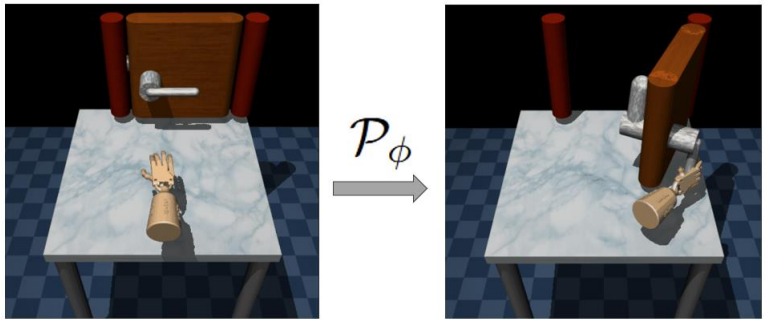
$\pi^{\mathrm{E}}$ outputs a stochastic action distribution

Exploration ←Balance→ Imitation

- **Evaluation of $\pi^G\left(\cdot \mid \cdot, \mathcal{P}_\phi(\cdot)\right)$ under random seeds via Goal Predictor $\mathcal{P}_\phi$**

Define: $\mathcal{P}_\phi : s \mapsto \hat{g}, \quad$ where $\hat{g} = \mathcal{P}_\phi(s)$

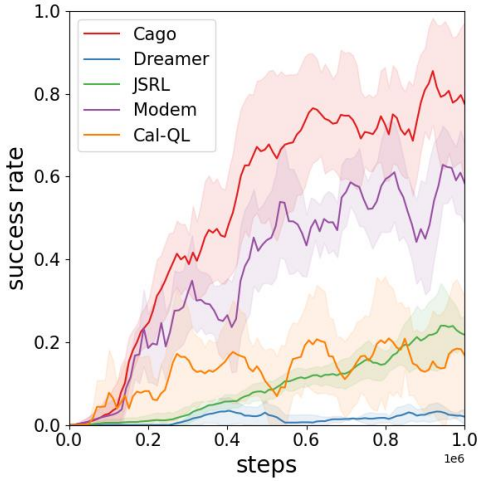
$$\min_\phi E_{\left(\tau^{(i)}=s_0^{(i)},\ldots,s_L^{(i)}\right)\sim\mathcal{D}_{demo}} \left|\mathcal{P}_\phi\left(s_t^{(i)}\right) - s_L^{(i)}\right|_2^2$$



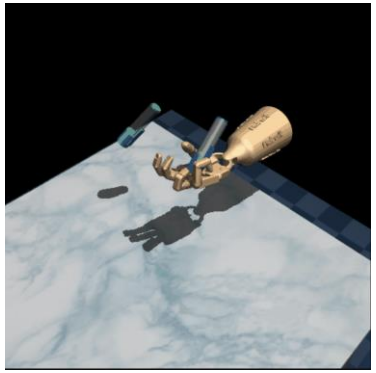
# Experiments

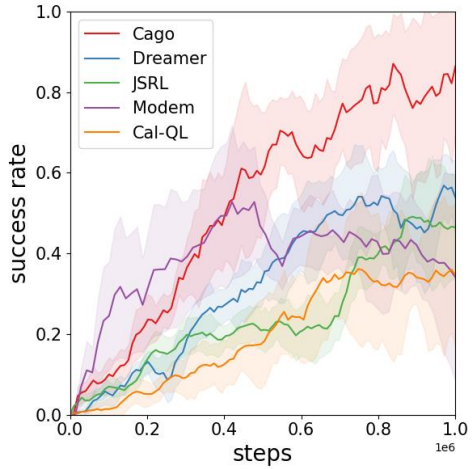
- **Subset of Test Environments**

**MetaWorld-Disassemble**

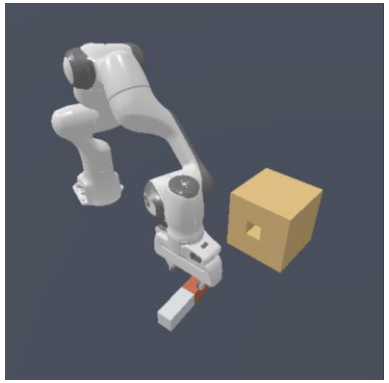


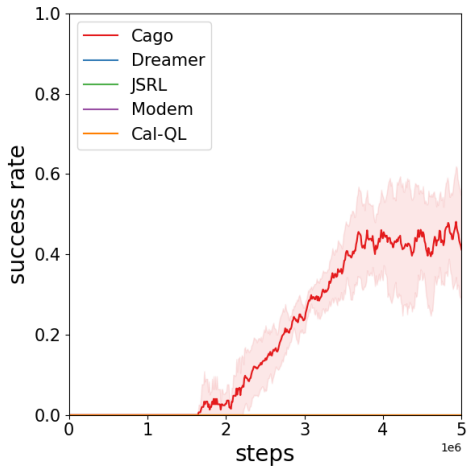

**Adroit-Pen**






**Maniskill-PegInsertion**

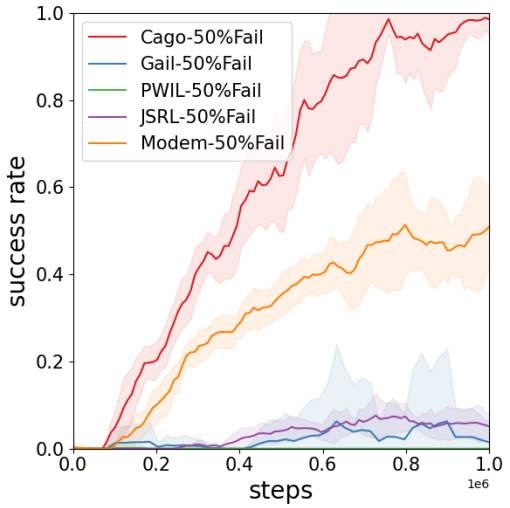




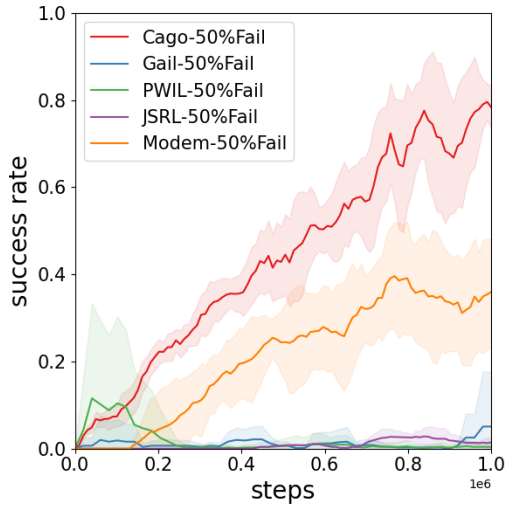
# Experiments

- **Robustness to Suboptimal and Failed Demonstrations**

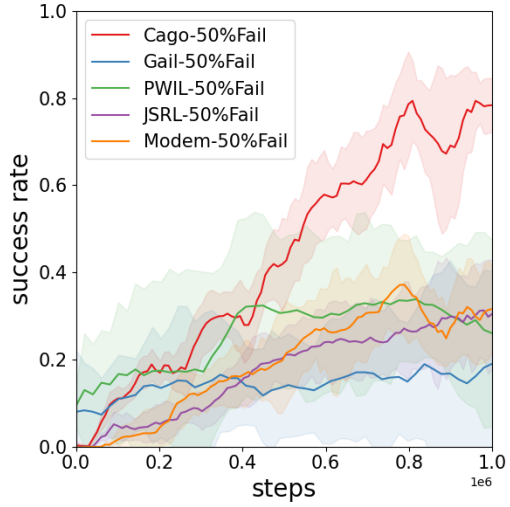
| Environment | Original Demo | Missing Obs | Noisy Actions | Random Actions |
|---|---|---|---|---|
| Stick-Push | 0.99 | 0.99 | 0.97 | 0.96 |
| Disassemble | 0.80 | 0.77 | 0.88 | 0.96 |
| Adroit-Pen | 0.82 | 0.88 | – | – |




Stick-Push




Disassemble




Adroit-Pen

# Learning from Demonstrations via Capability-Aware Goal Sampling

Yuanlin Duan, Yuning Wang, Wenjie Qiu and He Zhu

**Cago Github**

RUTGERS THE STATE UNIVERSITY OF NEW JERSEY
NEURAL INFORMATION PROCESSING SYSTEMS