

Consistent Zero-Shot Imitation with Contrastive Goal Inference

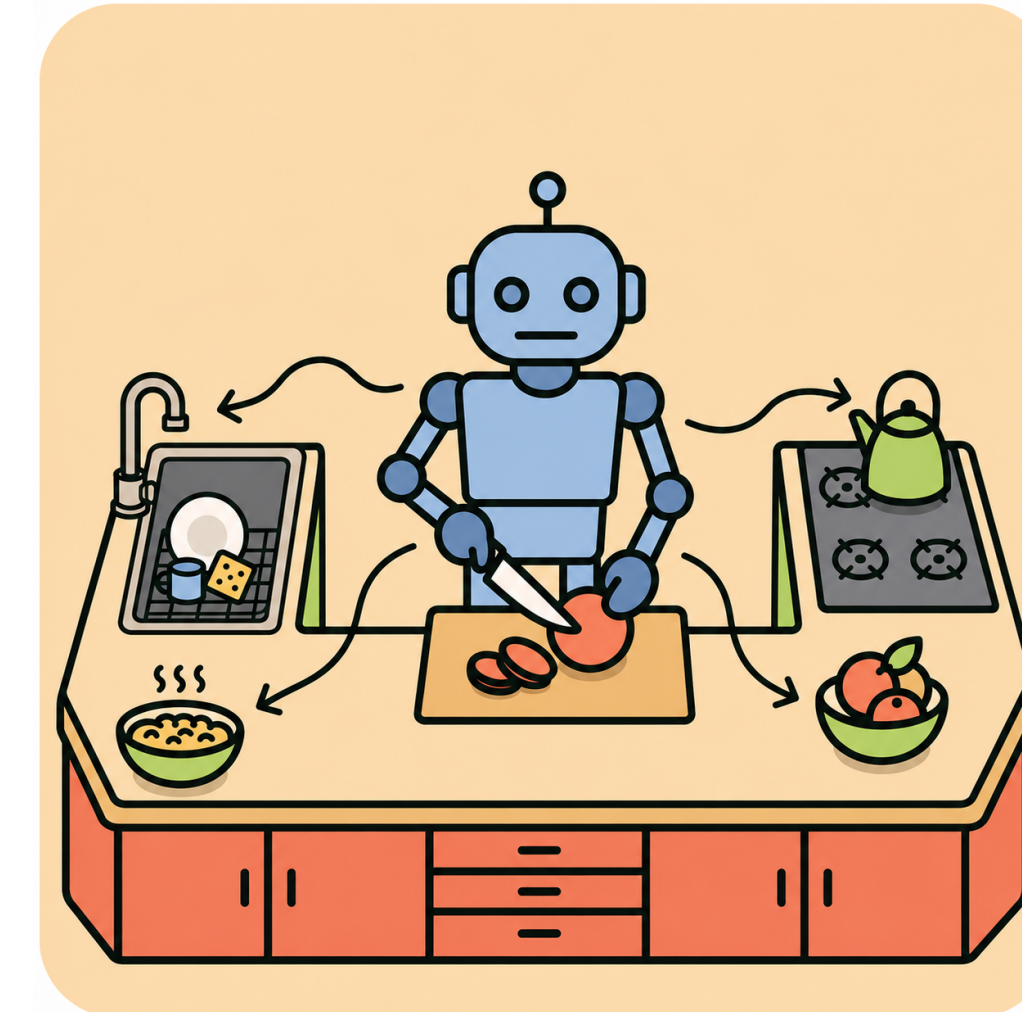
Kathryn Wantlin, Chongyi Zheng, Benjamin Eysenbach
Princeton University



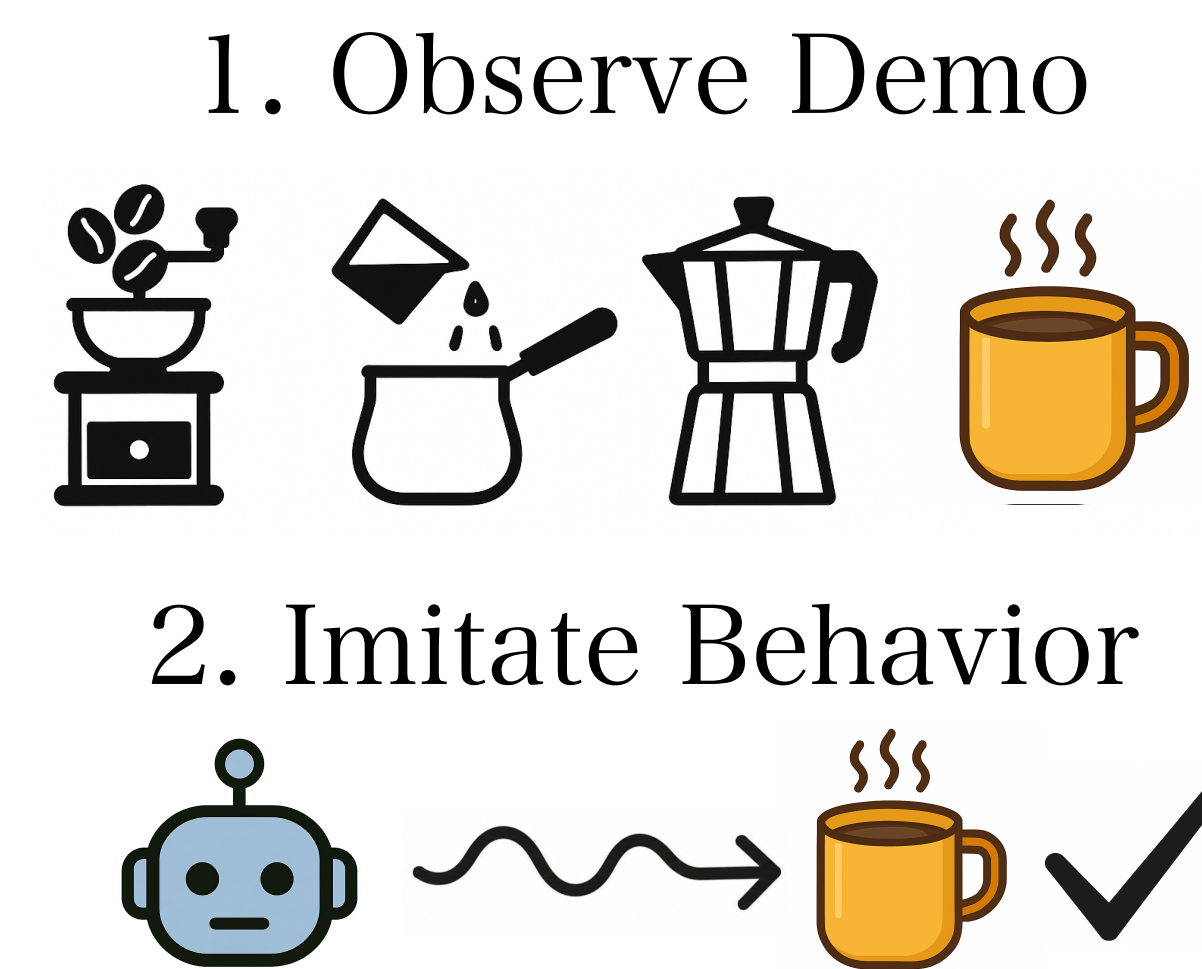
Motivation

We investigate whether *purely self-directed exploration & task execution* enable an agent to instantly imitate new behaviors. During training, our agent proposes and practices reaching its own goals in a multi-task environment, and at test time must imitate a demonstration *zero-shot, with low regret*.

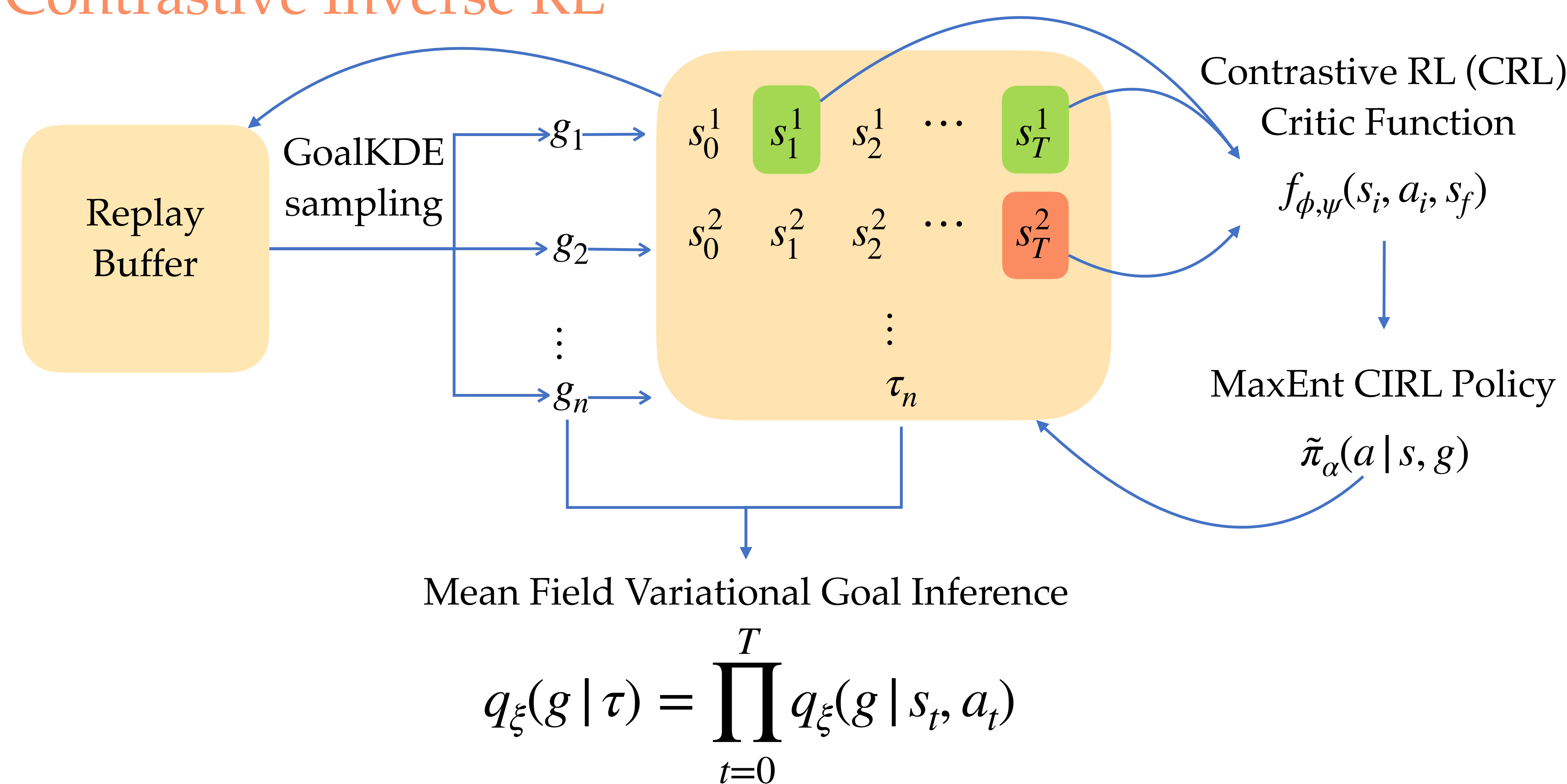
Train Time



Test Time



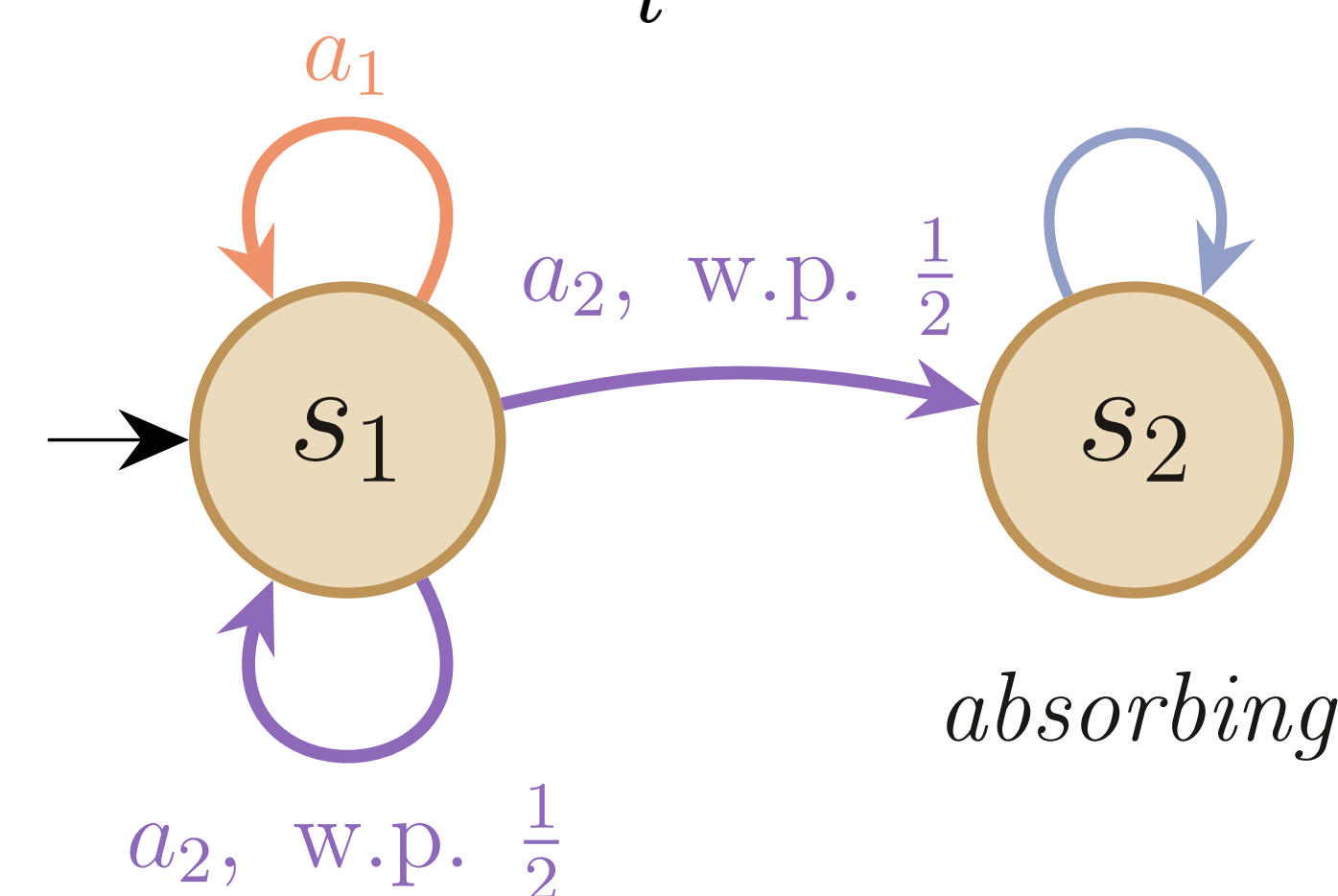
Contrastive Inverse RL



CIRL, unlike structurally similar methods, produces *theoretically consistent imitation*.

FB Representation

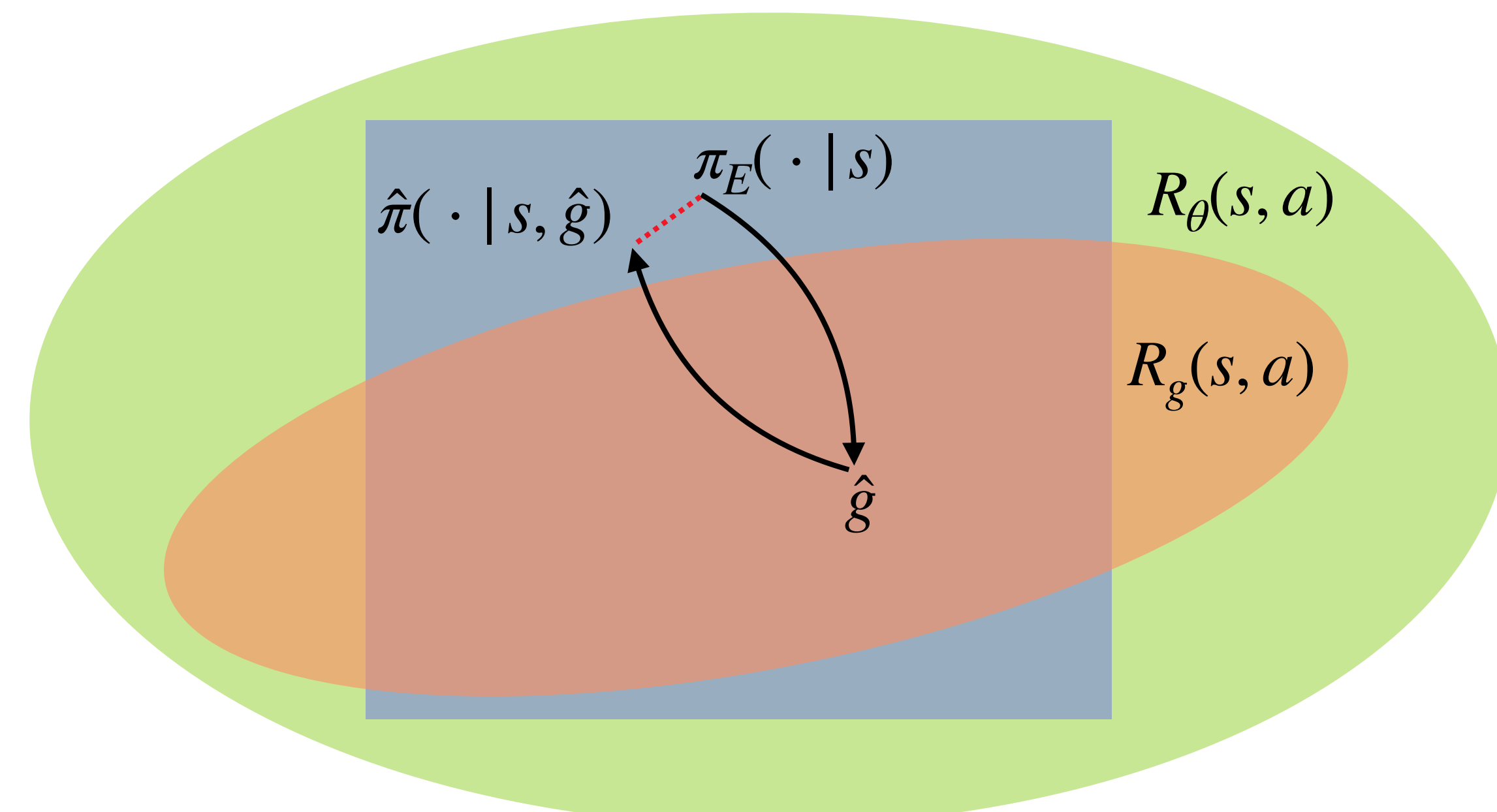
$$g = \sum_t \mathbb{1}(s_t)$$



$$\pi_E(s) = a_2 \rightarrow g = s_1 \quad \text{✗}$$

CIRL

$$\arg \max_{\theta} \mathbb{E}_{p_{\pi_E}(\tau)} [\log p_{\theta}(\tau)] \rightarrow \arg \max_{\xi} \mathbb{E}_{p_{\theta^*}(\tau, g)} [\log q_{\xi}(g | \tau)] \quad \text{✓}$$



Related Work

Reward-Free Zero-Shot RL

FB (Touati & Ollivier, 2021)

HILP (Park et al., 2024)

Fast Imitation with BFM

(Pirota et al., 2024)

Faulty assumption that distribution matching is sufficient for inferring the expert's policy.

Unsupervised Skill Learning

VIC (Gregor et al., 2016)

DIAYN (Eysenbach et al., 2019)

LSD (Park et al., 2022)

MISL (Zheng et al., 2024)

Not designed for zero-shot imitation from demonstrations.

Successor Representations

SR (Dayan, 1993)

SF (Barreto et al., 2016)

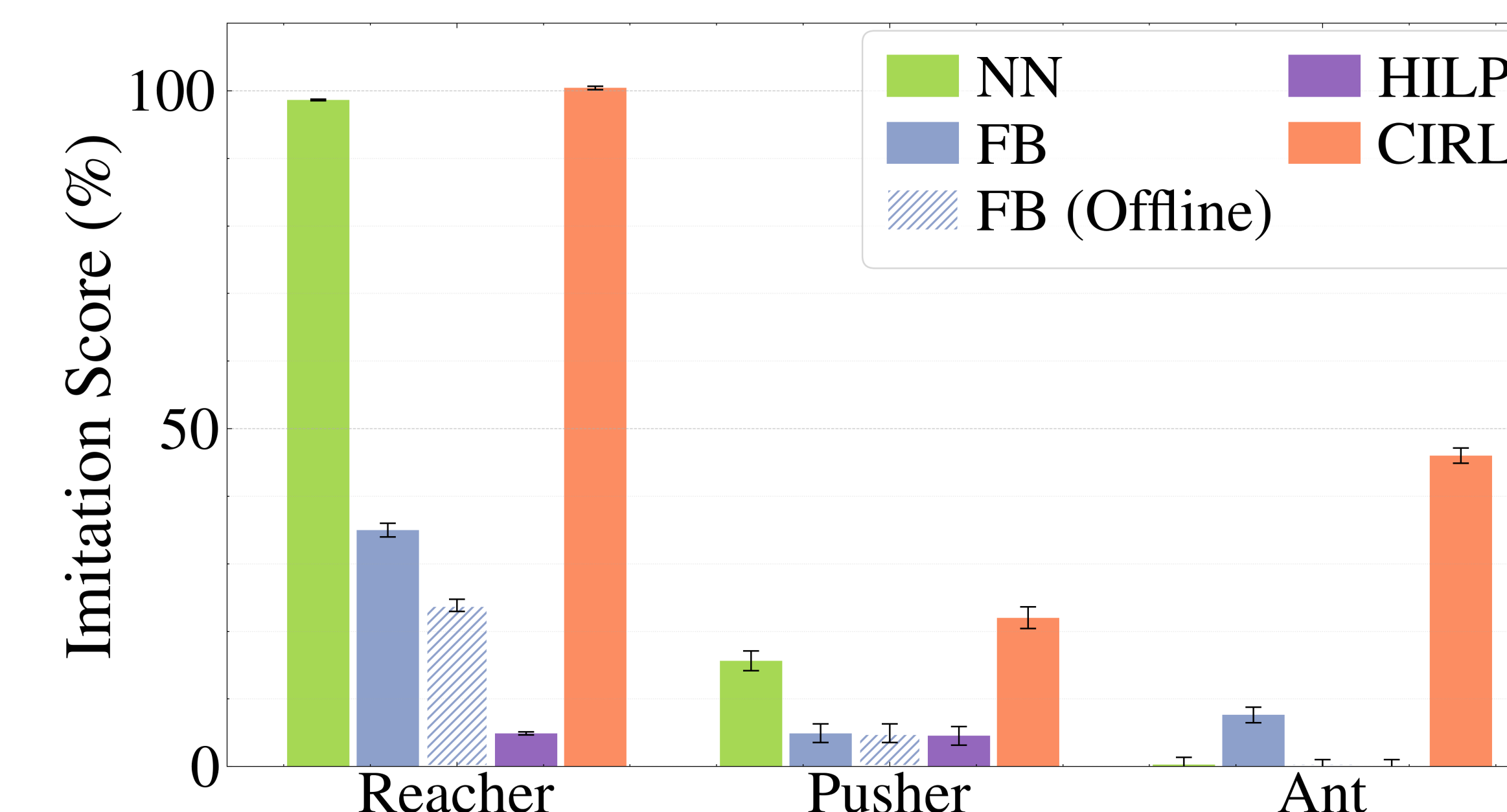
USFA (Borsa et al., 2018)

FB (Touati et al., 2021)

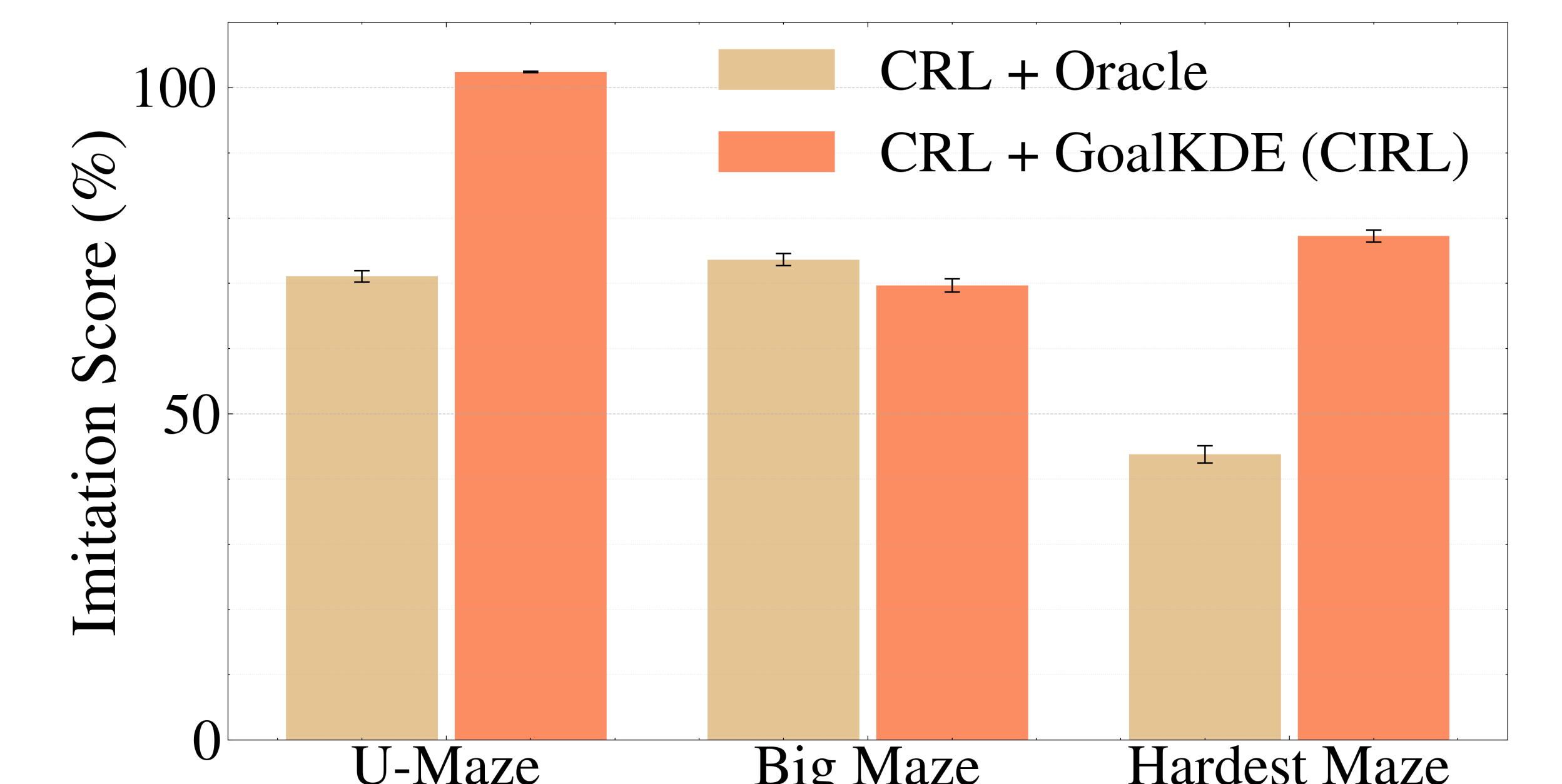
Learn universal reps for all rewards, ignoring benefits of a focus on goal-reaching rewards.

Experiments demonstrate that goals are an **efficient & expressive** summary of behavior, allowing us to reduce the IRL problem to a goal inference problem.

CIRL outperforms a suite of baselines on the JaxGCRL benchmark.



GoalKDE's exploration rivals oracle sampling and gracefully scales to larger state spaces.



CIRL effectively captures non-goal-conditioned reward functions, using the *expressivity of augmented goal spaces*.

