Benchmarking Constraint Inference in Inverse Reinforcement Learning

Guiliang Liu[123], Yudong Luo[23], Ashish Gaurav[23], Kasra Rezaee[4], Pascal Poupart[23].

[1]The Chinese University of Hong Kong, Shenzhen

[2]University of Waterloo, [3]Vector Institute, & [4]Huawei

Presented at ICLR 2023
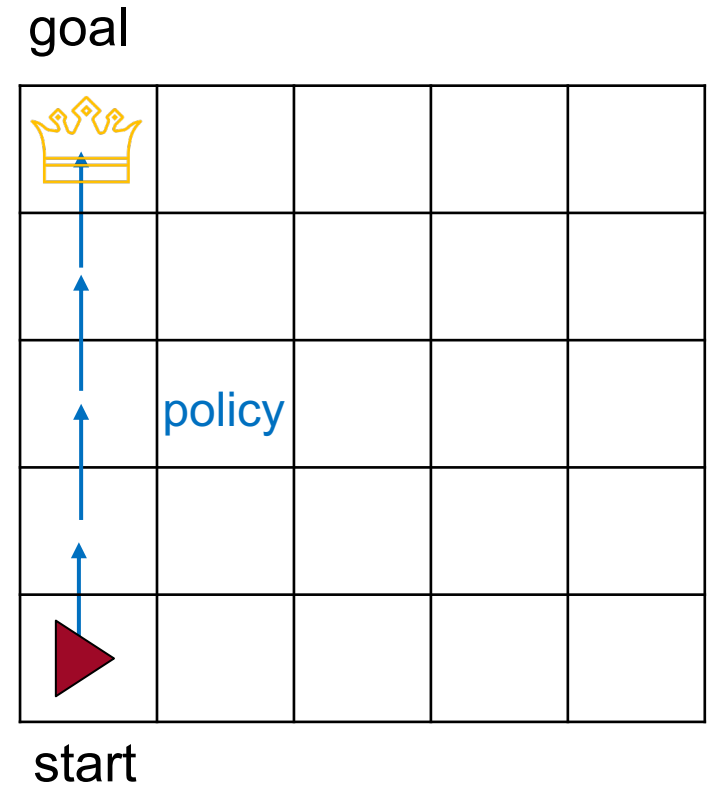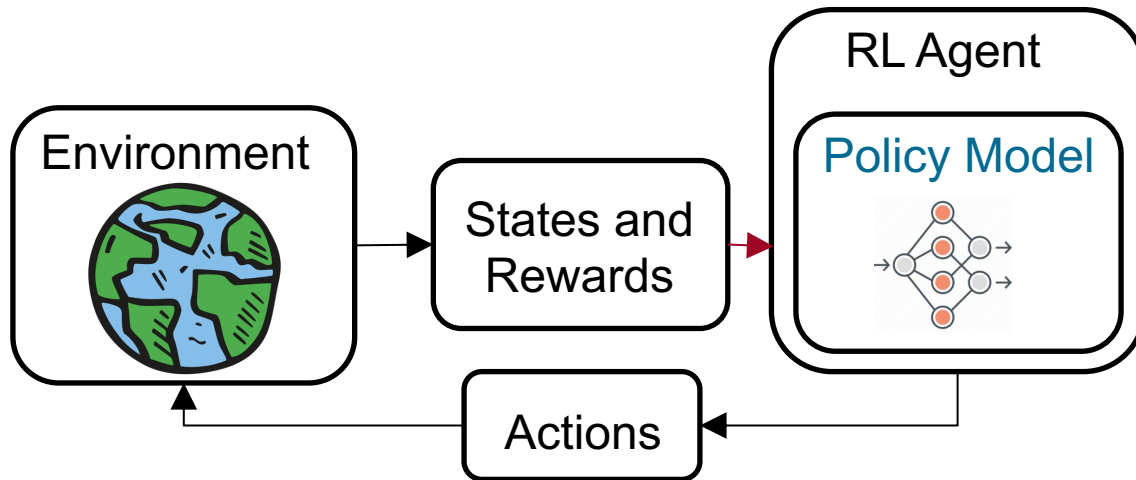
International Conference on Learning Representations (ICLR)
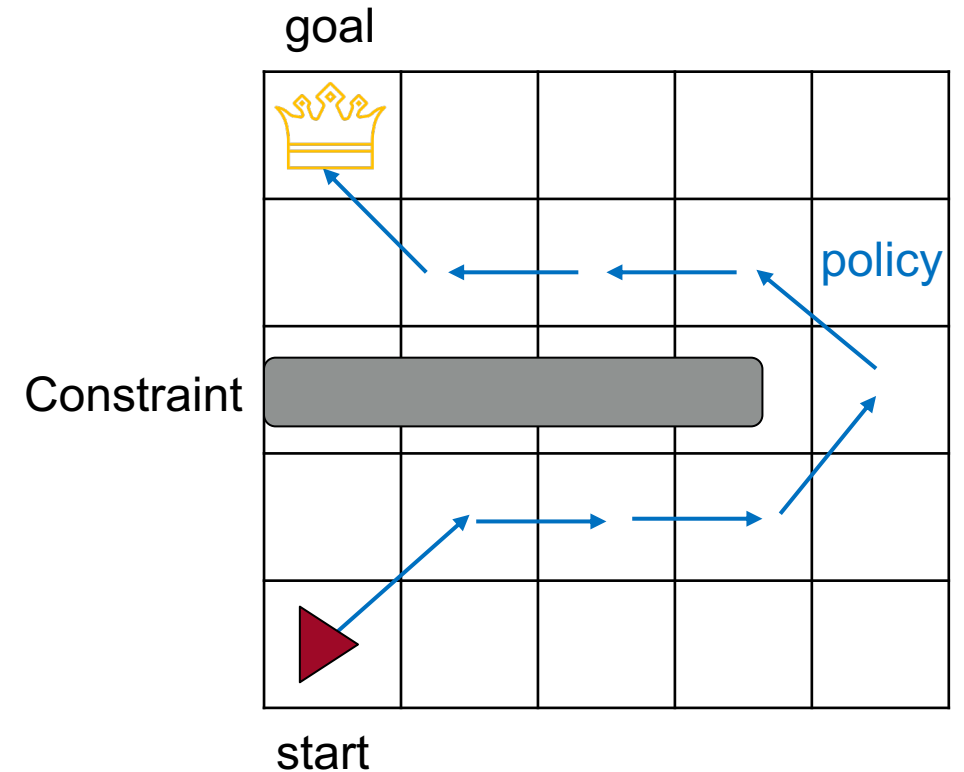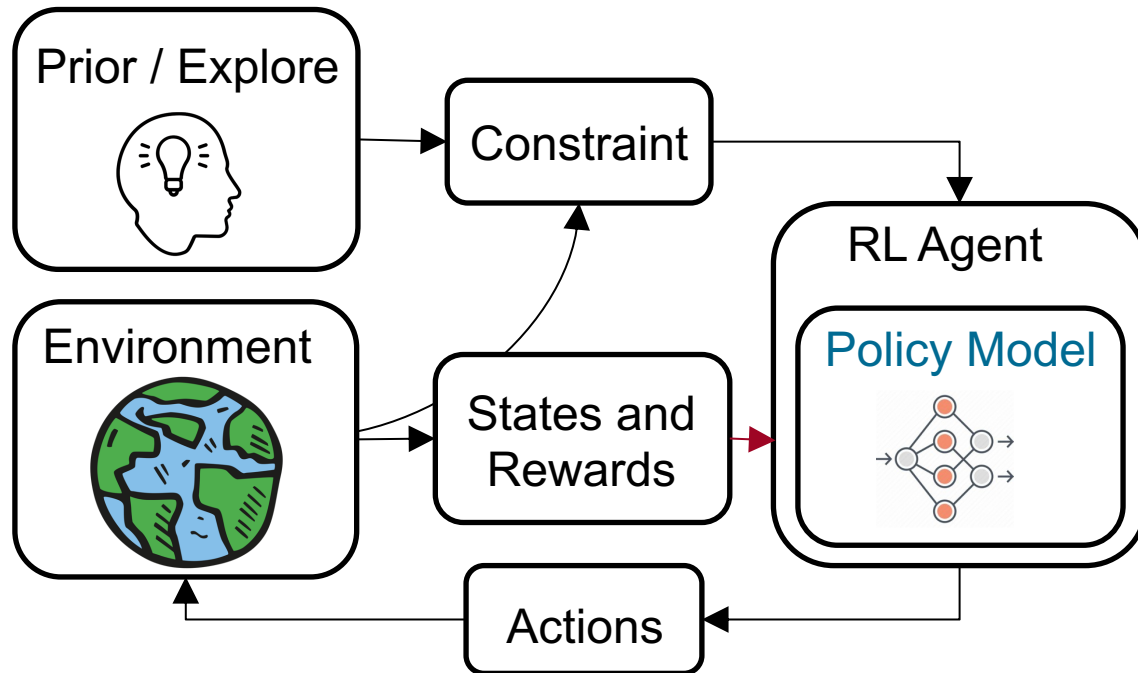
Introduction

Reinforcement Learning (RL)

- Optimize the control policy in the environment.

- E.g., general RL tasks (e.g., Go, Atari, Mujoco).

- What if having constraints?
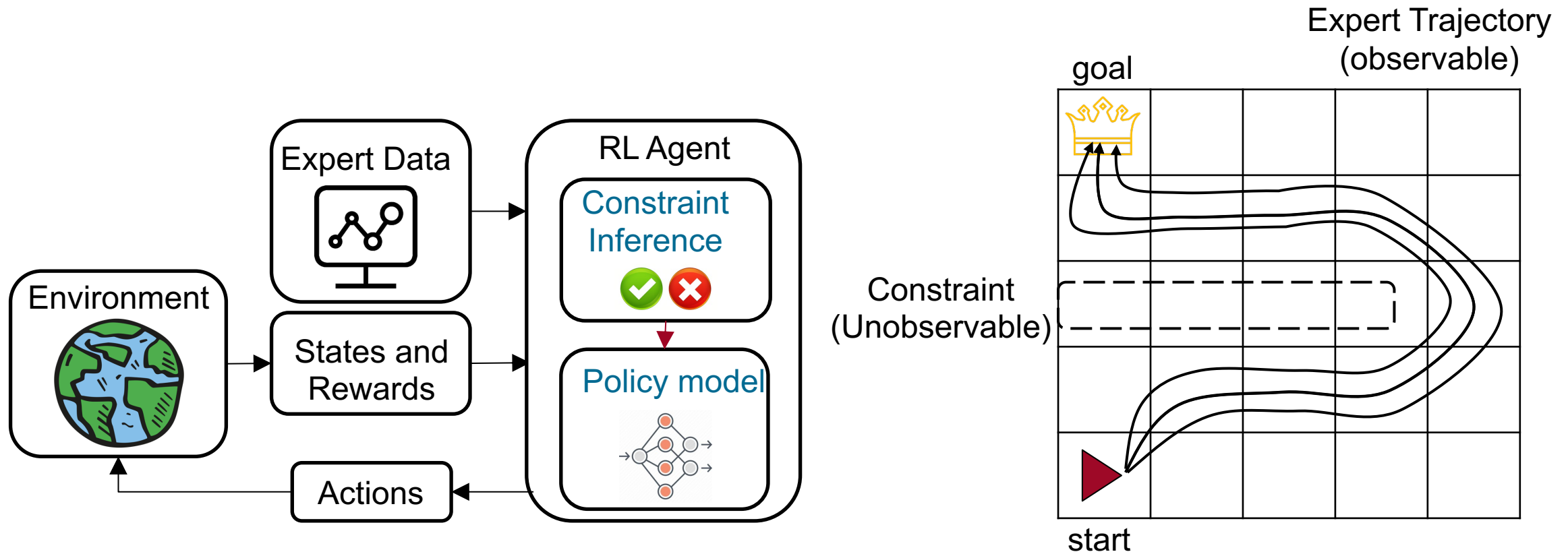
Constrained Reinforcement Learning (CRL)

- Optimize the control policy under a set of constraints.

- E.g., Safe Reinforcement Learning (e.g., Auto-driving).

- What if constraint not known ?

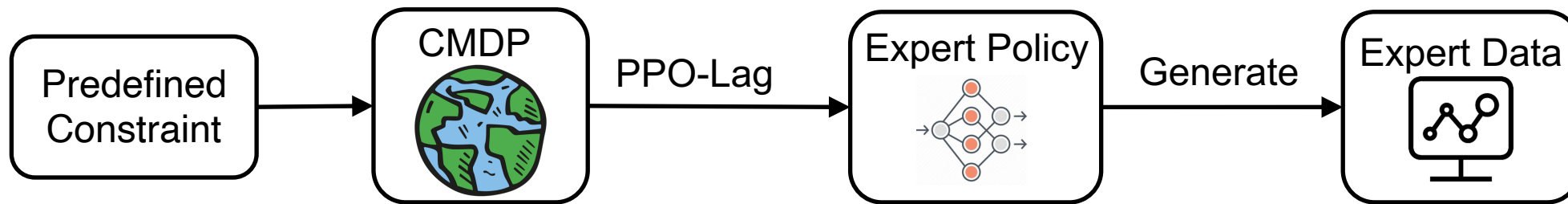## Inverse Constrained Reinforcement Learning (ICRL)

- Policy Optimization + Constrain Inference.

Benchmarks

Demonstration Dataset:

- Add predefine constraints to the environment to construct a Constrained MDP (CMDP).

- Trained expert agent with Proximal Policy Optimization Lagrange (PPO-Lag) algorithm.

- Generate the expert dataset with the expert agent.

Baseline methods:

- Binary Classifier Constraint Learning (BC2L).

- Generative Adversarial Constraint Learning (GACL).

- Maximum Entropy Constraint Learning (MECL).

- Variational Inverse Constrained Reinforcement Learning (VICRL).
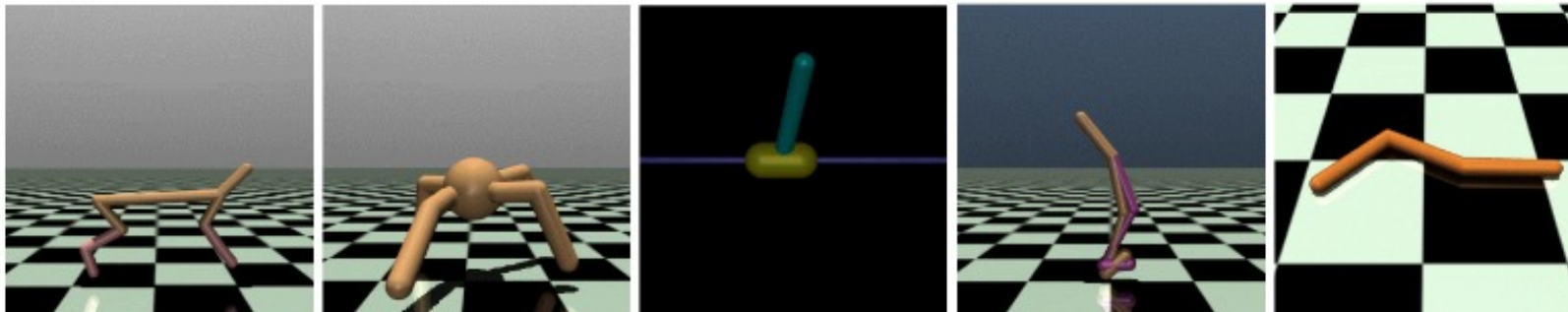
Environment Settings:

- Based on the MuJoCo Game engine.

- Incorporating predefined constraints into the environment.

Table 1: The virtual and realistic environments in our benchmark.

Type	Name	Dynamics	Obs. Dim.	Act. Dim.	Constraints
	Blocked Half-cheetah	Deterministic	18	6	X-Coordinate $\geq$ -3
	Blocked Ant	Deterministic	113	8	X-Coordinate $\geq$ -3
Virtual	Biased Pendulumn	Deterministic	4	1	X-Coordinate $\geq$ -0.015
	Blocked Walker	Deterministic	18	6	X-Coordinate $\geq$ -3
	Blocked Swimmer	Deterministic	10	2	X-Coordinate $\leq$ 0.5

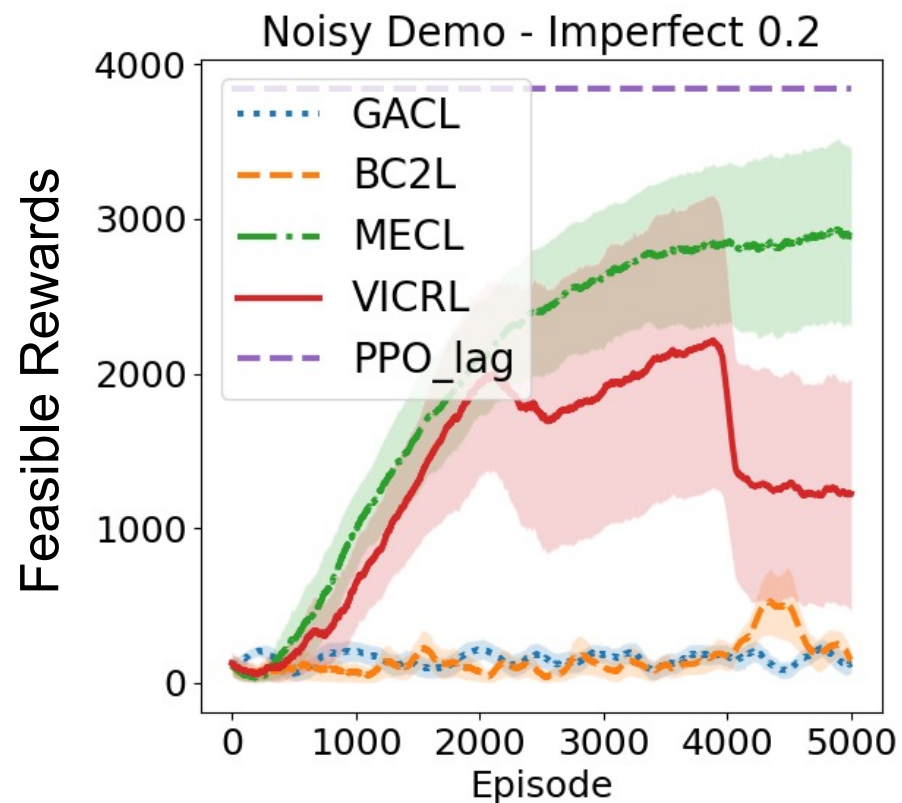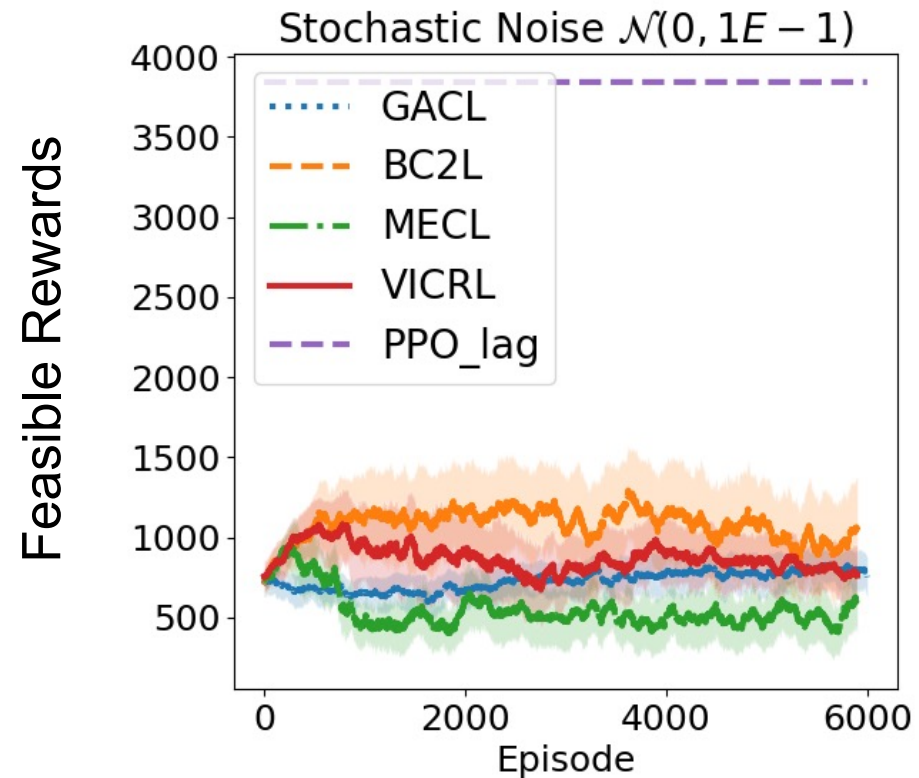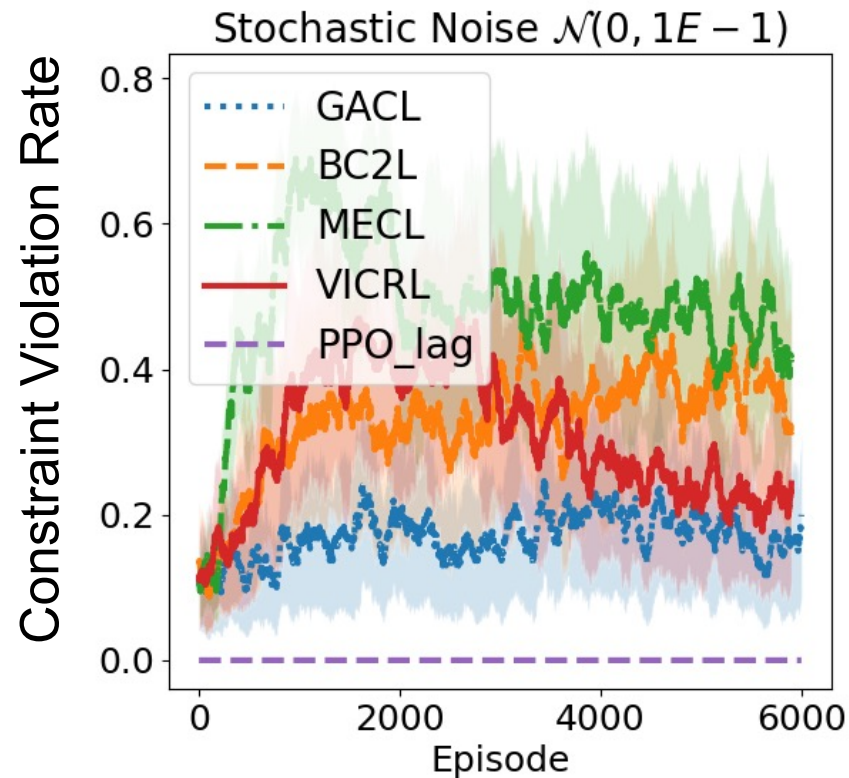*How well do the algorithms perform in the virtual environment?*
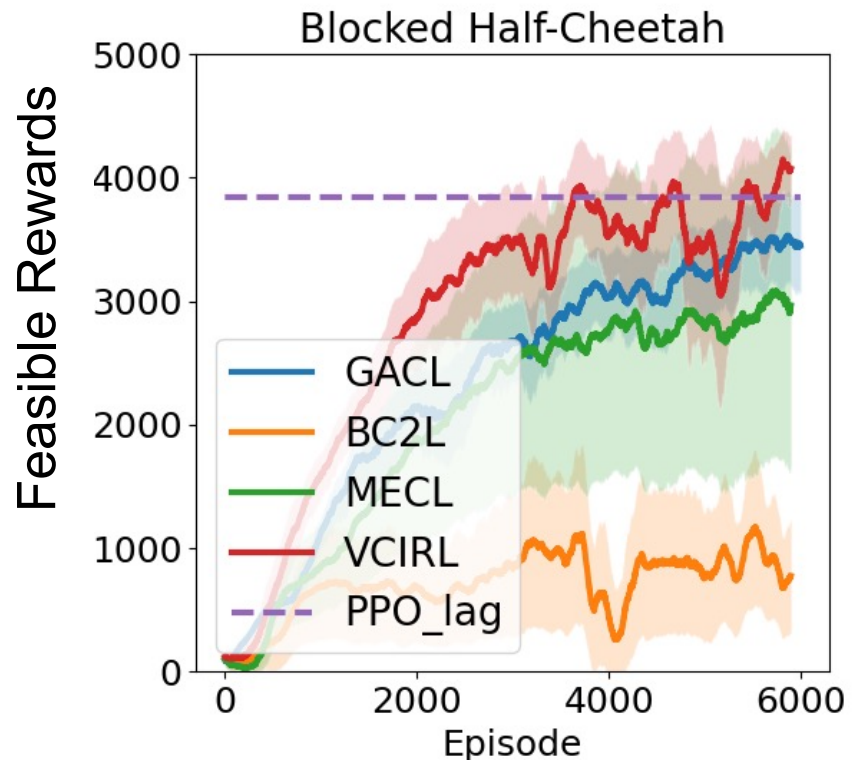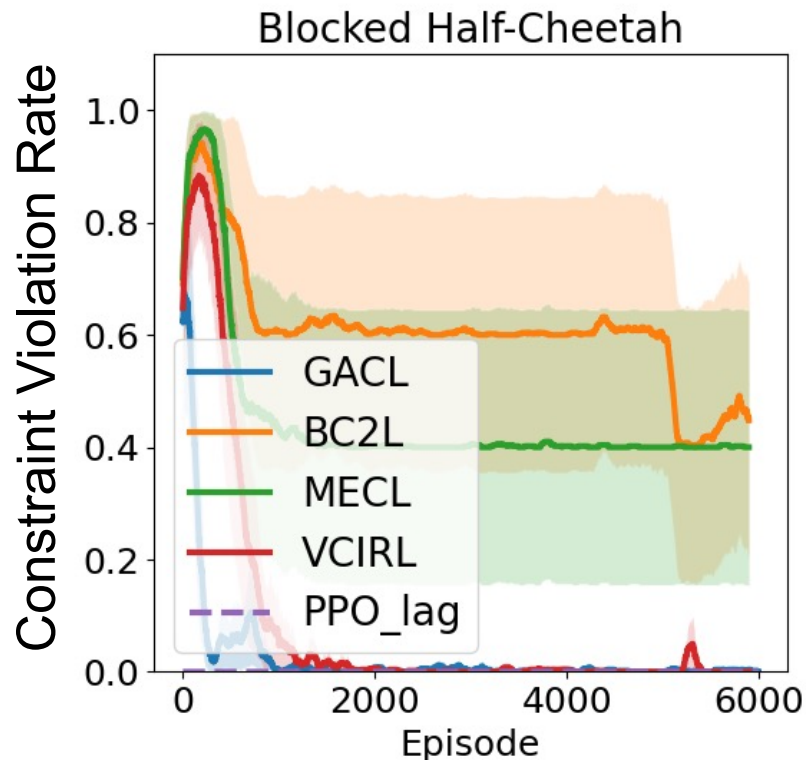
How well do the algorithms perform **when the expert demonstrations may violate the true underlying constraint?**
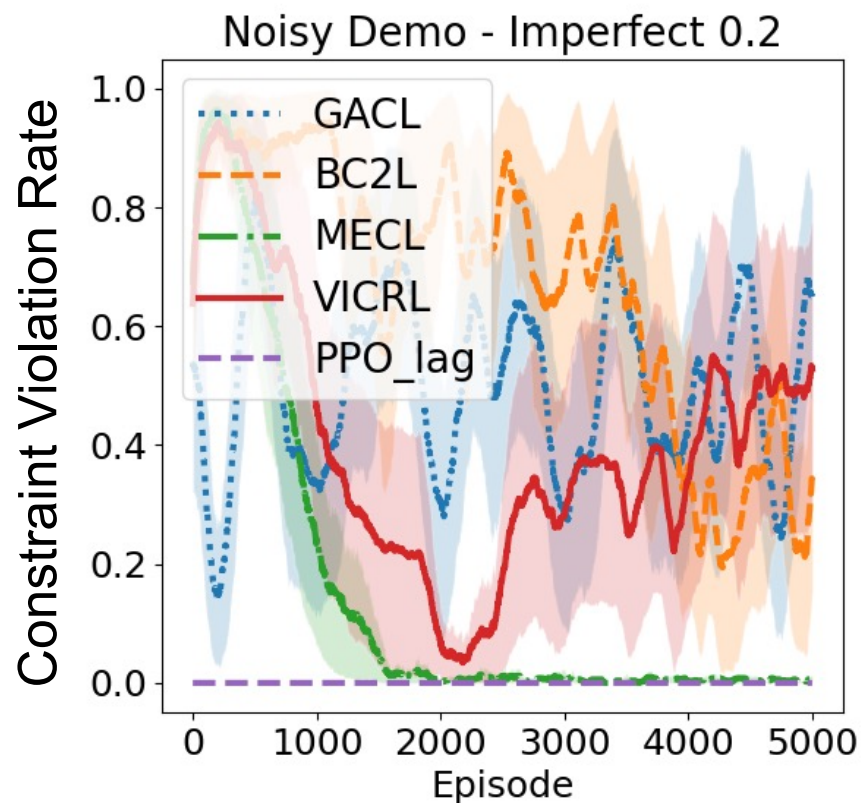
Virtual Environment

*How well do ICRL algorithms perform **in stochastic environments?***

Environment Settings:

- Based on the common road game engine.

- Incorporating predefined constraints into the environment.

Table 3: The constraints for realistic environments.

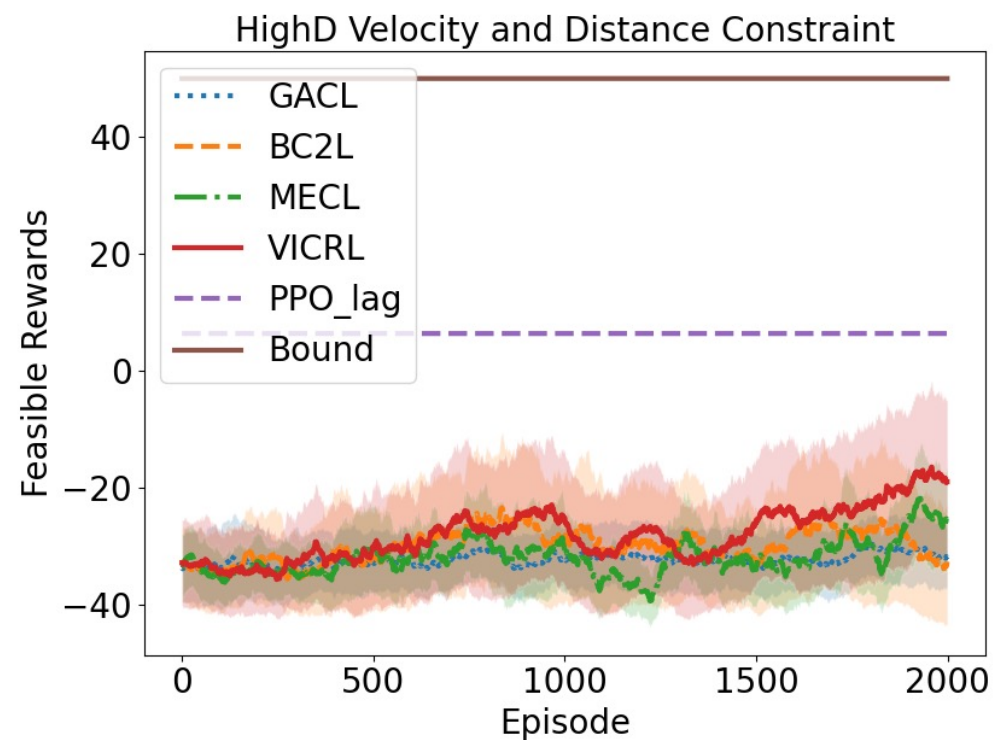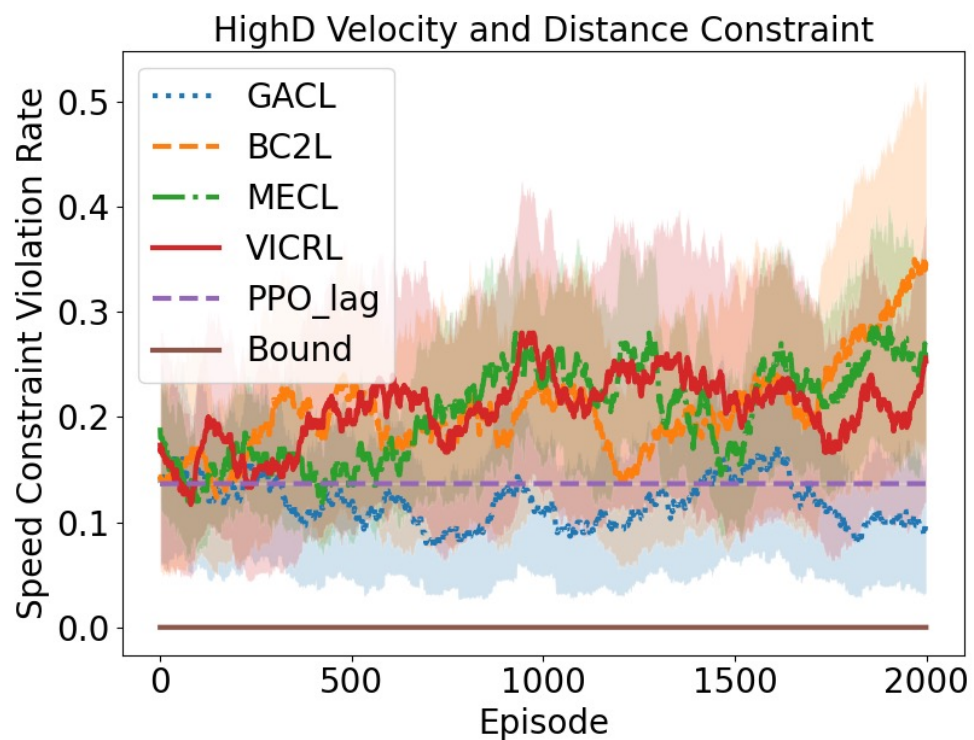| Type	Name	Dynamics	Obs. Dim.	Act. Dim.	Constraints
Realistic	HighD Velocity Constraint	Stochastic	76	2	Car Velocity $\leq$ 40 m/s
	HighD Distance Constraint	Stochastic	76	2	Car Distance $\geq$ 20 m

Realistic Environment

How well do the algorithms perform in the realistic environment?
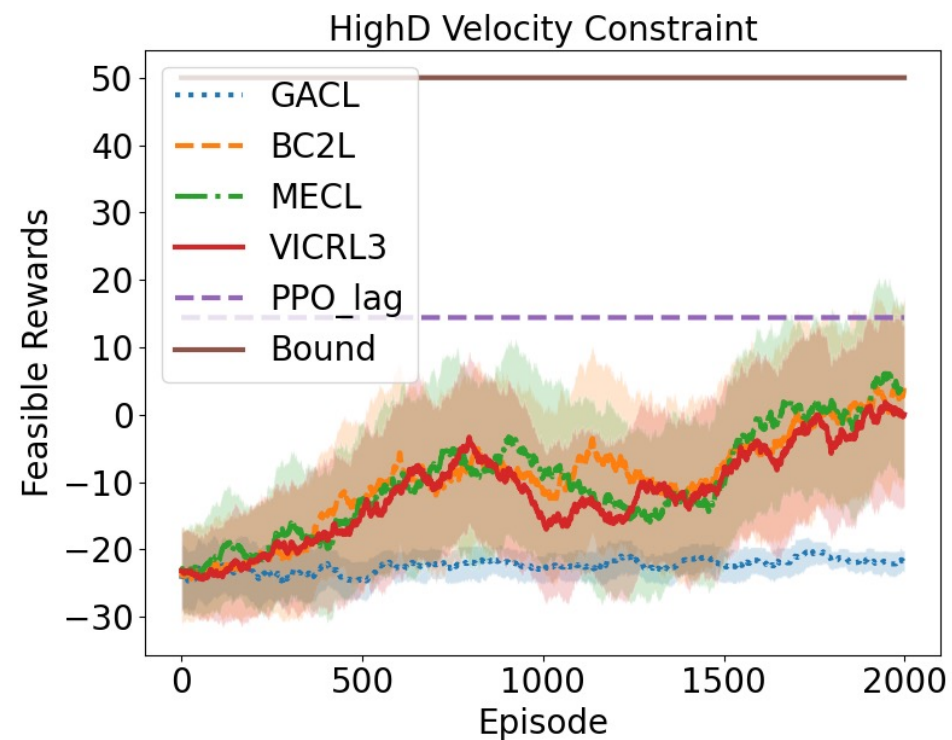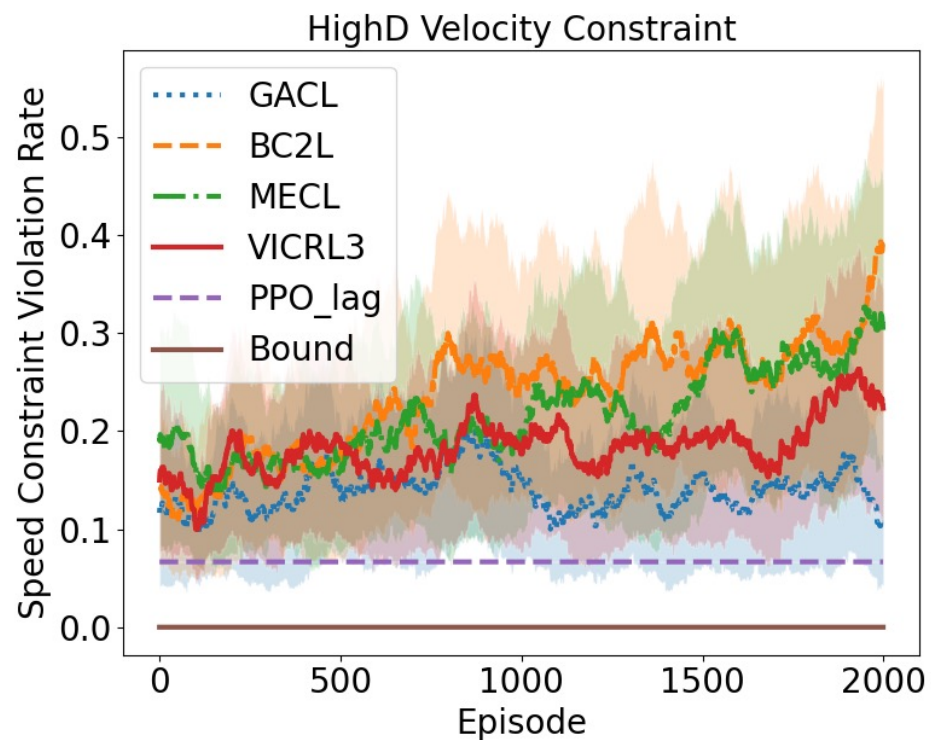
Realistic Environment

How well do ICRL algorithms work in terms of recovering multiple constraints?

Question and Answering (Q&A)