Inverse Constrained Reinforcement Learning

Usman Anwar, Shehryar Malik, Alireza Aghasi, Ali Ahmed

Introduction & Motivation

- Constraints are physically possible actions that should be avoided, for example, driving the car above the speed limit or breaking a traffic signals are examples of constraint violations.
- Focus of this work: automated learning of constraints.

Contributions

- Provide a **model free** constraint learning method for **high dimensional, continuous setting**.
- Empirically show that **learned constraints transfer** to agents with different dynamics and morphologies.

Training Objective

Neural Network based soft parametrization of indicator set over constrained trajectories.

$$\frac{1}{N}\sum_{i=1}^{N}\nabla_\theta \log \zeta_\theta(\tau^{(i)}) - \frac{1}{M}\sum_{j=1}^{M}\nabla_\theta \log \zeta_\theta(\hat{\tau}^{(j)}) - \delta \sum_{\tau \sim \{\mathcal{D}, \mathcal{S}\}} |1 - \zeta_\theta(\tau)|$$

Samples From Expert

Samples From Training Agent

Objective can be loosely interpreted as trying to match average soft cost for both expert and RL agent.

Regularizer: Puts Penalty On Over Constraining,

$$\omega(s_t, a_t) = \frac{\zeta_\theta(s_t, a_t)}{\zeta_{\bar{\theta}}(s_t, a_t)}$$

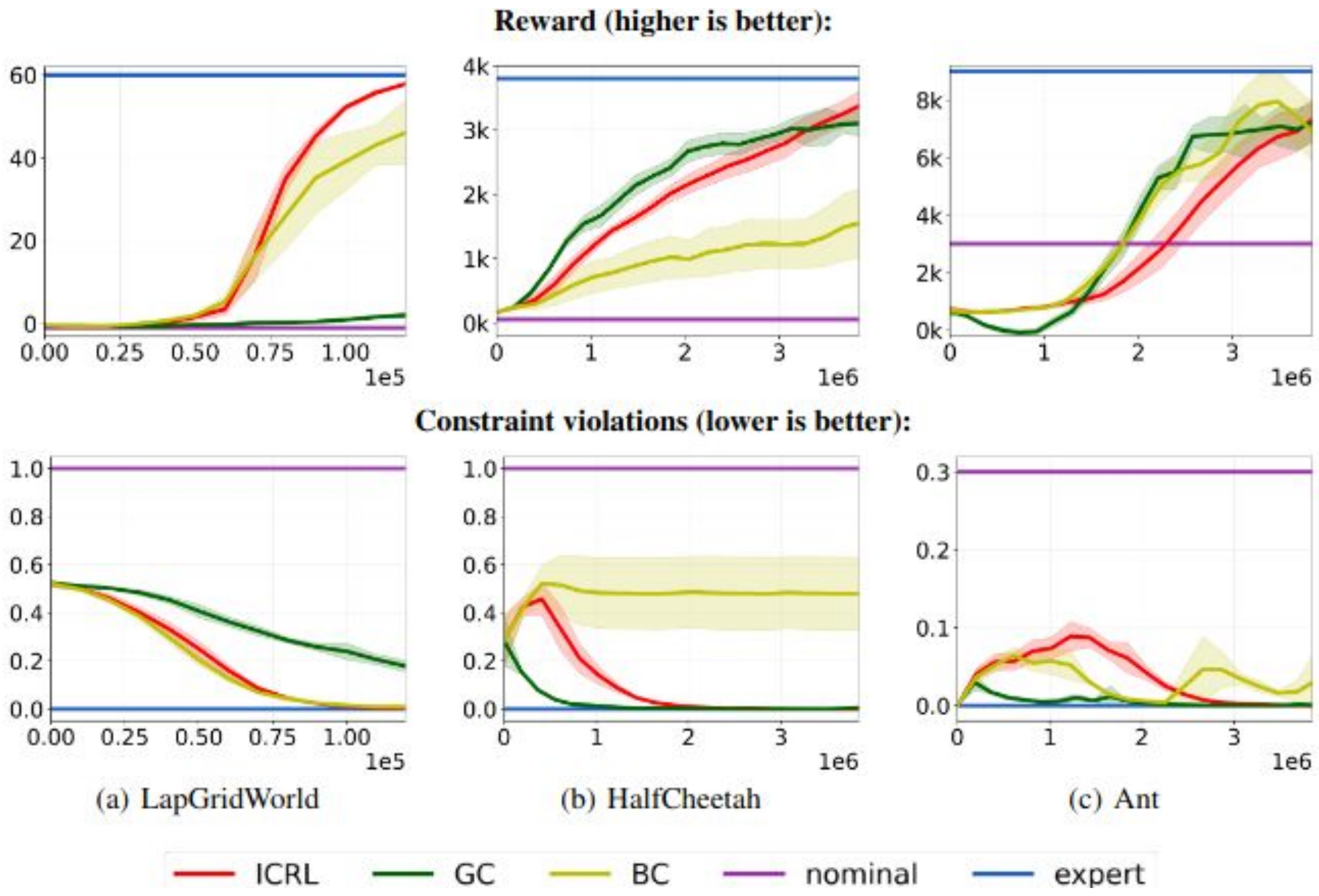
Current Neural Network

Sampling Neural Network

Use Importance Sampling to help reuse old samples and improve data efficiency.

Comparison With Baselines

We benchmark the algorithm against a *binary classifier baseline* and a *GAIL inspired baseline*.



Transfer Experiments

Question: Can constraint net obtained from demonstration of one agent transfer to other agents?

Answer: Yes. Even when new agents differ in morphology or have different dynamics.