Action Robust Reinforcement Learning and Applications in Continuous Control

Chen Tessler*, Yonathan Efroni* and Shie Mannor

*equal contribution

Poster #272
Robust MDPs

\[ \pi^* = \arg \max_{\pi} \min_{P \in \{P_1, \ldots, P_n\}} \mathbb{E}_P \left[ \sum_t \gamma^t r_t \right] \]

Important model, yet not feasible in practical applications.
Action Robustness in Robotics

Abrupt disturbances

Model uncertainty
Action Robust MDPs

\[ \pi^\text{mix}_\alpha (\pi, \pi') = \begin{cases} 
\pi & \text{w.p. } 1 - \alpha \\
\pi' & \text{otherwise}
\end{cases} \]

AR-MDPs are a special case of RMDPs, which consider uncertainty in the performed action.
Algorithm

Find optimal actor policy

Evaluate joint policy

Update adversary towards the 1-step greedy policy

Theorem 1. This procedure converges to the Nash equilibrium.
Results

Baseline

Ours (\(\alpha=1\))

Hopper

Walker2d
Conclusions

- Robustness enables coping with **uncertainty** and **transfer** to unseen domains
- A **gradient based** approach for robust reinforcement learning with convergence guarantees
- **Does not require** explicit definition of the **uncertainty set**
- Application to **Deep RL**
Come visit @ Poster #272