Jointly Inferring Human Irrationality and Intent

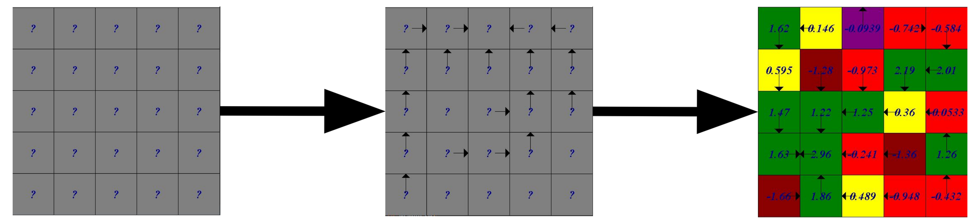
Emanuel Navarro-Oritz[1], Daniel S. Brown,[2] Anca Dragan[2]
[1]Cabrillo College
[2]Department of Electrical Engineering and Computer Sciences
University of California, Berkeley

2021 Transfer-to-Excellence Research Experiences for Undergraduates Program (TTE REU Program)

Abstract

Inverse reinforcement learning enables robots to learn new tasks from human demonstrations by learning a reward function that explains the human's intent.Our experimental results demonstrate that our novel joint Bayesian inference approach produces a better model of human intent, as it is intended to detect if the human demonstrator is systematically biased or irrational and compensate for the human's irrationality.

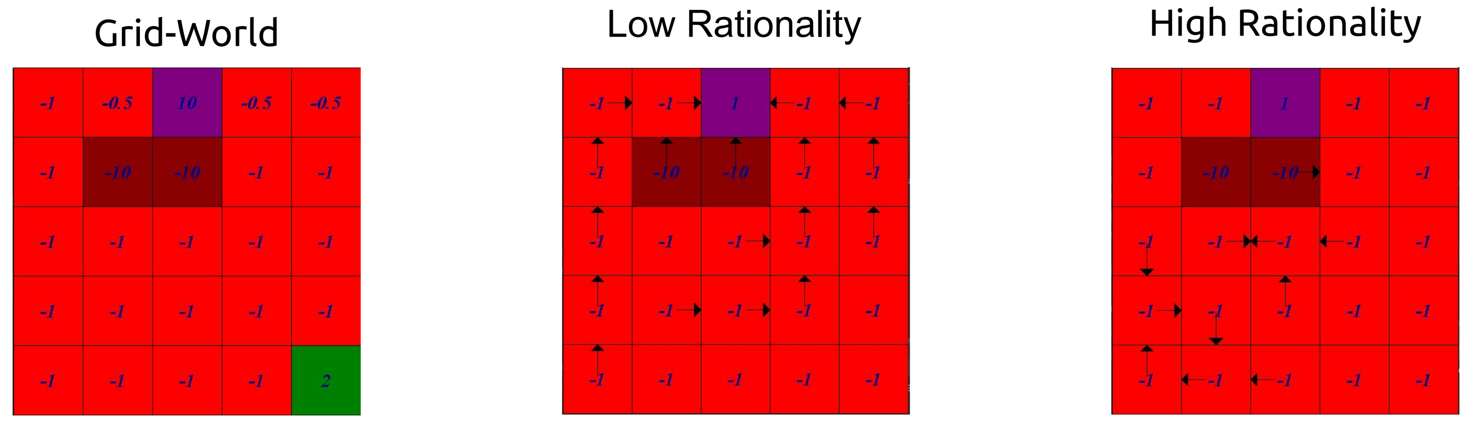
Introduction



Goal: We want to be able to train robots and computers to learn through human demonstrations.

- Inverse reinforcement learning algorithms provide a method for robots to learn a new task by inferring the demonstrator's reward function given an action.
- Given the standard inverse reinforcement [1] approach has led to successful models under the assumption that human demonstrators are rational.
- Our research approach has been developed to gauge the human demonstrator's rationality by not just inferring the reward function but by also parameterizing the rationality of the human's intent.

Background



Where we simulate a robot interface in a gridworld environment Each tile is state where a robot may traverse, which has a set of actions it may take; Up, down, left, and right. All actions have associated transition probability that tells us what happens to the robot under an action. A state has an associated value as a reward a robot may receive, also known as a reward function, which defines how favorable it is to transition to that state.
Gamma is a discounting factor that tells us how we encode and quantify how much each reward is worth in the future. Given the mdp we want to maximize the reward a robot receives under environment.
Examples of varying boltzmann rationality. Where low rationality exemplifies a random behavior, while higher rationality exhibits an intended behavior.

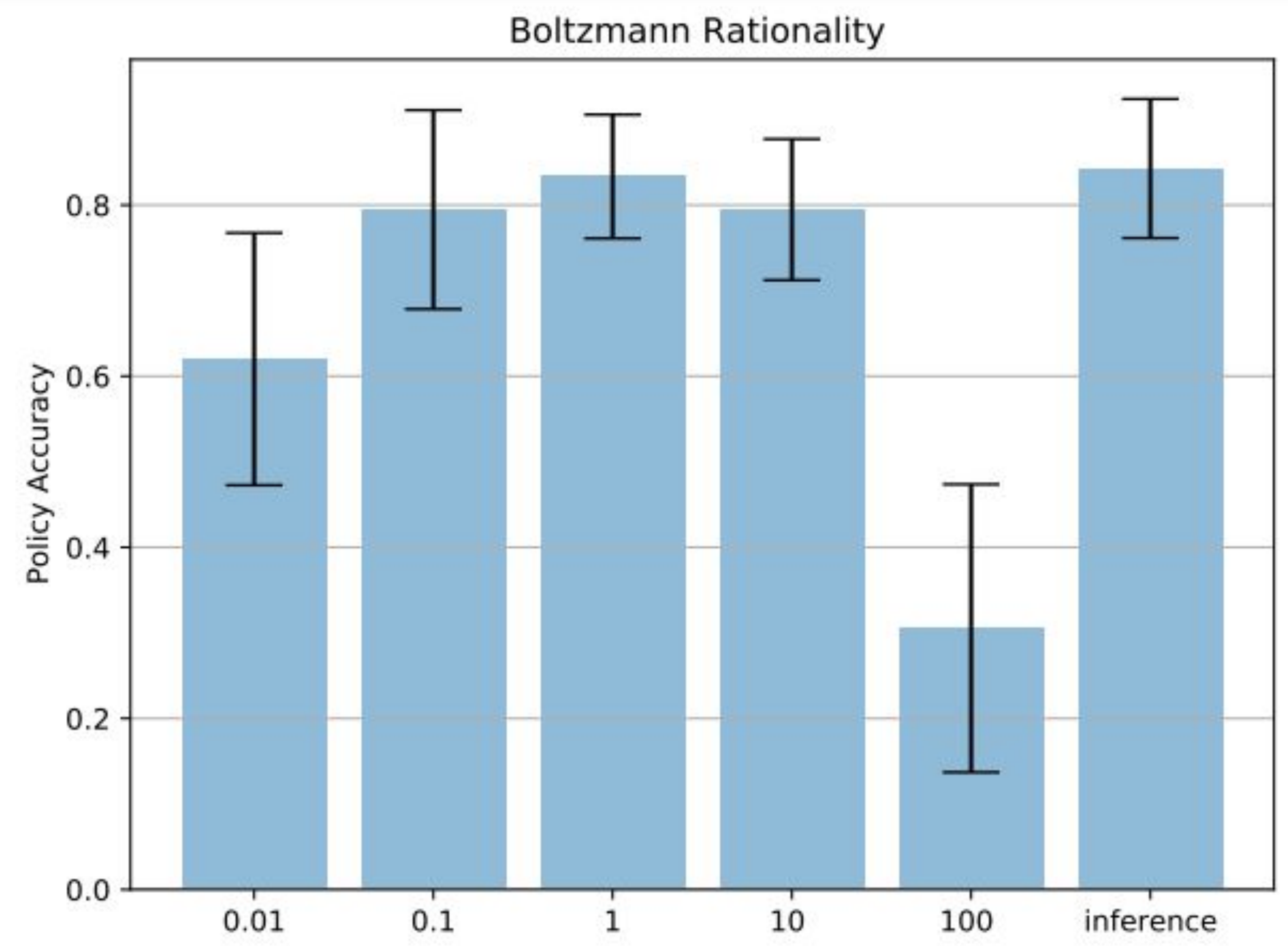
Experimental Method

Bayesian Inverse Reinforcement Research Approach

- Our likelihood function gives higher likelihood to reward function hypotheses that make the actions the demonstrator took look better than the alternative actions they could have chosen
- We take a Bayesian approach[2] to reward learning where want to have a distribution over reward functions that look likely given the demonstrations.

$$\prod_{(s,a)\in D} \frac{e^{\beta Q^*_R(s,a)}}{\sum_{b\in\mathcal{A}} e^{\beta Q^*_R(s,b)}}. \qquad P(R,\beta|D) \propto P(D|R,\beta)P(R,\beta)$$

Results



Conclusion

- Standard Inverse Reinforcement Learning Approach

When given wrongly assumed irrationality the data shows strong evidence that the standard approach fails to infer a reward function. In particular when we overestimate and underestimate the demonstrators competency the policy accuracy falls to a mean of 35 percent and 63 percent, respectively.

Research approach

Now if we look at the approach implemented where we inferred over reward function and the rationality, the model produces a better policy accuracy. This is exciting progress because this allows us to create deeper richer models with varying rationality. The result shows promise that yes we could infer a human's intent and under a bias.

Future Work

In future research we hope to learn from different forms of human irrationality such as myopia, non-determinism bias, and prospect bias.

Aknowledge

This work was supported by the Hopper Dean Foundation and National Science Foundation Award #1757690. The Transfer-to-Excellence Program is sponsored by the National Science Foundation and the Center for Energy Efficient Electronics.

References

[1]Abbeel, P. and Ng, A., 2000 et al *Apprenticeship Learning via Inverse Reinforcement Learning*
[2]Sutton and Barto 1998 et al *Reinforcement Learning: An Introduction*
[3]Ramachandran and Amir 2007 et *al Bayesian Inverse Reinforcement Learning*
[4]Brown, D. 2021 et al *Safe and Efficient Inverse Reinforcement Learning*

Contact Information
Email:emannavarro@gmail.com Cell: (831) 313-5229
linkedin: linkedin.com/in/emanuel-navarro-ortiz/