Recent Advancements in Inverse Reinforcement Learning

Alberto Maria Metelli
Politecnico di Milano
Piazza Leonardo da Vinci, 32
20133 Milano, Italy
albertomaria.metelli@polimi.it

Introduction

Inverse reinforcement learning (IRL) has seen significant advancements in recent years. This class of approaches aims to efficiently learn the underlying reward function that rationalizes the behavior exhibited by expert agents, often represented by humans. In contrast to mere behavioral cloning, the reconstruction of a reward function yields appealing implications, as it allows for more effective *interpretability* of the expert's decisions and provides a *transferable* specification of the expert's objectives for application in even different environments. Unlike the well-understood field of *reinforcement learning* (RL) from a theoretical perspective, IRL still grapples with limited understanding, significantly constraining its applicability. A fundamental challenge in IRL is the inherent *ambiguity* in selecting a reward function, given the existence of multiple candidate functions, all explaining the expert's behavior.

Content of the Talk

In this talk, I will survey three of my papers that have made notable contributions to the IRL field: "Provably Efficient Learning of Transferable Rewards" (Metelli et al. 2021), "Towards Theoretical Understanding of Inverse Reinforcement Learning" (Metelli, Lazzati, and Restelli 2023), and "Inverse Reinforcement Learning with Sub-optimal Experts" (Poiani et al. 2023).

The central innovation introduced by (Metelli et al. 2021) is a novel formulation of the IRL problem that overcomes the issue of ambiguity. IRL is reframed as the problem of learning the *feasible reward set*, which is the set of all rewards that can explain the expert's behavior. This approach postpones the selection of the reward function, thereby circumventing the ambiguity issues. Furthermore, the feasible reward set exhibits convenient geometric properties that enable the development of efficient algorithms for its computation.

Building on this novel formulation of IRL, (Metelli, Lazzati, and Restelli 2023) addresses the problem of efficiently learning the feasible reward set when the environment and the expert's policy are not known in advance. It introduces a novel way to assess the dissimilarity between feasible reward sets based on the Hausdorff distance and presents a new PAC (probabilistic approximately correct) framework. The most significant contribution of this paper is the introduction of the *first sample complexity lower bound*, which highlights the challenges inherent in the IRL problem. Deriving this lower bound necessitated the development of novel technical tools. The paper also demonstrates that when a generative model of the environment is available, a *uniform sampling strategy* achieves a sample complexity that matches the lower bound, up to logarithmic factors.

Finally, in (Poiani et al. 2023), the IRL problem in the presence of *sub-optimal experts* is investigated. Specifically, the paper assumes the availability of multiple sub-optimal experts, in addition to the expert agent, which provide additional demonstrations, associated with a known quantification of the maximum amount of sub-optimality. The paper shows that this richer information mitigates the ambiguity problem, significantly reducing the size of the feasible reward set while retaining its favorable geometric properties. Furthermore, the paper explores the associated statistical problem and derives novel lower bounds for sample complexity, along with almost matching algorithms.

These selected papers represent notable advancements in IRL, contributing to the establishment of a solid theoretical foundation for IRL and extending the framework to accommodate scenarios with sub-optimal experts.

Acknowledgments

This paper is supported by PNRR-PE-AI FAIR project funded by the NextGeneration EU program.

References

Metelli, A. M.; Lazzati, F.; and Restelli, M. 2023. Towards Theoretical Understanding of Inverse Reinforcement Learning. In *ICML*, volume 202, 24555–24591. PMLR.

Metelli, A. M.; Ramponi, G.; Concetti, A.; and Restelli, M. 2021. Provably Efficient Learning of Transferable Rewards. In *ICML*, volume 139, 7665–7676. PMLR.

Poiani, R.; Curti, G.; Metelli, A. M.; and Restelli, M. 2023. Inverse Reinforcement Learning with Sub-optimal Experts. Preprint.