Ph.D. in Information Technology:

Thesis Defense June 9th, 2021 online by Teams – at 17.00

Giorgia Ramponi – XXXIII cycle

Advisor: Prof. Marcello Restelli

PhD Thesis Title: **Challenges and Opportunities in Multi-Agent Reinforcement Learning**

Abstract:

Reinforcement Learning (RL) is a Machine Learning area that studies sequential decision- making problems, where a learning agent interacts with an unknown environment in order to maximize its rewards. In recent years, RL methods have made substantial progress in solving real-world problems. However, the most successful applications, such as beating the world champion player of Go, solving robotic control problems, managing the power consumption of households, and achieving promising results in autonomous driving, involve more than one agent and can be cast in the Multi-Agent Reinforcement Learning (MARL) setting. However, although the MARL setting is an important research area of practical interest, this framework is still poorly understood from a theoretical point of view. In general, the presence of many agents makes the learning problem more complex, and in many situations, single RL algorithms cannot be applied.

In this thesis, we take a step toward solving this problem, providing theoretically sound algorithms for this setting. We analyze the challenges and opportunities that a multi-agent environment creates in the RL framework, providing new approaches in three RL sub- problems while also showing how they are interconnected. The contributions of the thesis are theoretical, algorithmic, and experimental. We take inspiration from practical problems, we design new algorithms with desirable theoretical properties to solve them, and we show their performances in benchmarks domains and on real-world data.

The thesis is divided into four parts. In the first part, we provide the background and preliminaries necessary to follow the rest of the thesis. We start by introducing the RL problem and classical algorithms to solve it. Then, we introduce the Inverse RL problem, i.e., the problem of recovering reward functions from an expert's demonstrations. Finally, we provide the necessary background on game theory and MARL.

In the second part, we analyze how the presence of multiple agents affects the Inverse Reinforcement Learning problem and objective. We provide two novel algorithms: the first considers how to recover and cluster the intentions of (i.e., the rewards optimized by) a set of agents given demonstrations of near-optimal behavior; the second aims at inferring the reward function optimized by an agent while observing its actual learning process. The experimental evaluation is conducted on synthetic problems and two real-world problems.

We then show the importance of learning (or knowing) the other agent's intention to construct efficient algorithms. In particular, in the third part, we study online learning in the MARL scenario, showing how the presence of other agents can increase the hardness of the problem while proposing statistically efficient algorithms. We design two algorithms to solve the Configurable MDP problem, a setting where an external entity can partially control the transition model. Then, we analyze the statistical limits of general-sum stochastic games when we control only one agent, providing a new lower bound and a near-optimal algorithm.

Finally, in the fourth part, we study MARL from an optimization viewpoint while showing the difficulties that arise from multiple function optimization problems. Then, we present a new algorithm for this scenario, providing convergence results and extensively evaluating it against SotA baselines.