Guaranteeing Control Requirements via Reward Shaping in Reinforcement Learning

Francesco De Lellis, *Member, IEEE*, Marco Coraggio, *Member, IEEE*, Giovanni Russo, *Senior Member, IEEE*, Mirco Musolesi, *Member, IEEE*, and Mario di Bernardo, *Fellow, IEEE*

***Abstract*—In addressing control problems such as regulation and tracking through reinforcement learning (RL), it is often required to guarantee that the acquired policy meets essential performance and stability criteria such as a desired settling time and steady-state error before deployment. Motivated by this, we present a set of results and a systematic reward-shaping procedure that: 1) ensures the optimal policy generates trajectories that align with specified control requirements and 2) allows to assess whether any given policy satisfies them. We validate our approach through comprehensive numerical experiments conducted in two representative environments from OpenAI Gym: the Pendulum swing-up problem and the Lunar Lander. Utilizing both tabular and deep RL methods, our experiments consistently affirm the efficacy of our proposed framework, highlighting its effectiveness in ensuring policy adherence to the**

address complex control problems, relying solely on data and employing a reward maximization process. This approach finds diverse applications, spanning from attitude control [1] and wind farm management [2] to autonomous car-driving [3] and the regulation of plasma using high-fidelity simulators [4]. However, a significant challenge in this domain revolves around ensuring that the learned control policy demonstrates the desired closed-loop performance and steady-state error, posing a crucial open question in control system design.

It is often argued that accurate knowledge of system dynamics is necessary to provide analytical guarantees of stability and performance, which is crucial for industrial applications [5], [6]. In fact, in this article, we introduce a set