2024 IEEE International Conference on Robotics and Automation (ICRA)
May 13-17, 2024. Yokohama, Japan

Co-learning Planning and Control Policies Constrained by Differentiable Logic Specifications

Zikang Xiong, Daniel Lawson, Joe Eappen, Ahmed H. Qureshi, and Suresh Jagannathan

mation (ICRA) | 979-8-3503-8457-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICRA57147.2024.10610942

Abstract— **Synthesizing planning and control policies in robotics is a fundamental task, further complicated by factors such as complex logic specifications and high-dimensional robot dynamics. This paper presents a novel reinforcement learning approach to solving high-dimensional robot navigation tasks with complex logic specifications by co-learning planning and control policies. Notably, this approach significantly reduces the sample complexity in training, allowing us to train high-quality policies with much fewer samples compared to existing reinforcement learning algorithms. In addition, our methodology streamlines complex specification extraction from map images and enables the efficient generation of long-horizon robot motion paths across different map layouts. Moreover, our approach also demonstrates capabilities for high-dimensional control and avoiding suboptimal policies via policy alignment. The efficacy of our approach is demonstrated through experiments involving simulated high-dimensional quadruped robot dynamics and a real-world differential drive robot (TurtleBot3) under different types of task specifications.**

I. Introduction

Synthesizing planning and control policies is among the core tasks of robotics. The planning policy determines a path comprising a sequence of robot configurations between the given start and goal. In contrast, the control policy helps robots interact with the physical environment allowing them to follow the given plan to reach a goal. However, synthesizing planning and control policies is challenging since the high-dimensional robot dynamics introduce a vast

into RL presents its own challenges. For example, previous approaches [1]–[10] that have used temporal logic for reward shaping, involves the construction of complicated reward functions that require large amounts of samples [11]. Furthermore, handcrafting these temporal logic specifications can be challenging. For instance, situations where map layouts are represented as images or have irregular obstacle shapes present numerous challenges in defining precise, actionable logical rules. Although existing solvers [12] can find robot paths under a given logical specification, they lack scalability and can only handle linear and, to some extent, quadratic constraints, which is not ideal for practical scenarios, as highlighted in our experiments. Furthermore, these methods only account for path planning without considering errors from the underlying controller or vice versa, which often leads to failure in executions.

Therefore, we propose a novel approach, called Differentiable Specifications Constrained Reinforcement Learning (DSCRL), to address the above challenges using the following key features:

Lower Sample Complexity: We introduce a new methodology that integrates differentiable specifications into constrained RL, which significantly lowers the sample complexity of training.

Control and Planning Alignment: Learning both planning and control policies separately faces the challenges of sub-