

Safe Reinforcement Learning via Episodic Control

ZHUO LI, DERUI ZHU, JENS GROSSKLAGS

Abstract—Safe reinforcement learning (Safe RL) aims to learn policies capable of learning and adapting within complex environments while ensuring actions remain free from catastrophic consequences. This is a critical consideration in domains such as robotics, autonomous vehicles, and healthcare. Unlike traditional RL, which focuses mainly on maximizing episodic rewards, Safe RL integrates safety constraints to balance rewards and safety. Current Safe RL algorithms, while promising, lack sample efficiency as they require extensive environmental interactions to perform multi-objective optimization. This paper introduces an episodic-control-based method to enhance sample efficiency in safe policy optimization. In RL, methods based on episodic control involve the storing and replaying of past experiences or episodes, aiding in more efficient and precise policy optimization. Our proposed method involves clustering, measuring, and storing previous states based on a joint metric of returns and safety. Subsequently, we retrieve these state measurements and incorporate them into the policy optimization process through reward shaping. This approach effectively guides the policy towards high-return and safe decisions. We evaluate the performance of our method on established Safe RL benchmarks, including six safety-critical agent control tasks. The results demonstrate that our method can concurrently achieve higher episodic returns and fewer violations of safety constraints compared to the baseline methods, suggesting an effective balance between earning rewards and safety.

1. Contextual Challenges Introduction

2. Related Works

Summary:

3. Proposed Work

4. My Perspective on This Paper