Reinforcement Learning for Safe and Efficient Planning in Autonomous Driving

by

Anirban Santara

Under the supervision of

Prof. Pabitra Mitra (IIT Kharagpur)

and

Prof. Balaraman Ravindran (IIT Madras)

Abstract

The success of autonomous driving is contingent on the development of safe and efficient motion planning algorithms capable of working in multi-agent environments under stochasticity and partial observability. Reinforcement Learning (RL) provides a powerful framework for learning to behave in such environments. The focus of this thesis is to advance the state-of-the-art in reinforcement learning-based motion planning for autonomous driving. We present MADRaS, an open-source Multi-Agent DRiving Simulator that is capable of simulating challenging driving tasks of high variance. We demonstrate how MADRaS can be used as a curriculum learning platform for training RL agents to drive a wide range of cars in different road tracks, navigate through traffic in narrow roads and prevent congestion and deadlocks through multi-agent cooperation. In reinforcement learning, the agent's objective is specified in terms of a scalar reward function making its accurate description crucial for success in achieving the desired behavior. In order to bypass the need to hand-engineer reward functions, Imitation learning algorithms like Generative Adversarial Imitation Learning (GAIL) estimate an expert's reward function from its demonstrations and then maximize it using RL. We observe that although GAIL is effective in matching (and often, exceeding) the expert at mean performance, high-cost trajectories, corresponding to tail-end events of catastrophic failure, are more likely to be encountered by GAIL agents than the expert. To address this issue, we develop Risk-Averse Imitation Learning (RAIL) as an alternative to GAIL in risk-sensitive applications that achieves up to 89% reduction in tail-risk at benchmark continuous control tasks of OpenAI Gym. The sample efficiency and convergence time of an RL algorithm heavily depend on the exploration method used. While human beings use knowledge from prior experiences at related tasks while exploring a new task, most exploration algorithms for RL use the information only from the current task-environment. We develop ExTra, a framework for Transfer-guided Exploration in which we leverage a known optimal policy of a related task for efficient exploration in a new task. We demonstrate that ExTra is capable of exceeding and complementing the performance of traditional exploration algorithms.

Keywords: Autonomous Driving, Reinforcement Learning, Imitation Learning, Efficient Exploration