Project Proposal

Tianyu Shi
tianyu.shi3@mail.mcgill.ca

I. Problem Definition

In the coming years and decades, autonomous vehicles will become increasingly prevalent, providing new opportunities to improve mobility and transportation systems. Many practical applications of autonomous driving motion planning constrain agents to learn from a fixed batch of data which has already been gathered. A line of researches [1,2,3] for learning policies follow the end-to-end imitation learning that directly maps sensor inputs to vehicle control commands via supervised training on large amounts of human driving data. However, these systems cannot be generalized to unseen scenarios and their performances are severely limited by the coverage of human driving data. Furthermore, it is difficult to pose the learned policy as a purely supervised learning problem as the autonomous vehicle needs to heavily interact with the environment including other vehicles and roadways.

II. Motivation

Most imitation learning method may fail due to the distribution shift problem, which is introduced by the mismatch of state distribution between the collected dataset and the current environment [4]. Deep reinforcement Learning has great potential to learn such policies from exploration. However, the amount of exploration required will have high cost which prohibits its use in real applications.

III. Related Work

Imitation Learning: Imitation learning that learns policies via supervised training on human driving data has been applied to a variety of tasks, including modeling lane following [2], navigational behavior [5], off-road driving [6], etc. In [7], they proposed a joint learning algorithm that learns a shared cost function employed by behavior and trajectory components. It provides interpretable costing function for each driving constraints and make the end-to-end learning trainable.

Reinforcement Learning: Reinforcement learning allows agent to learns by exploring the environment, and does not require explicit human demonstration. Learning an optimal policy will be very time-consuming and easy to fall into local optimum after many episodes. It is thus desirable to find a feasible action space that can help avoid unnecessary exploration.

provide some references on using reinforcement learning in robotics

how is your proposed project solve this issue?

maybe you can mention that data-driven methods do not garantee stability

IV. Objective and contribution

In the previous research [8], we proposed an imitation learning approach that use optimization based method to generate the driving data set and neural networks to imitate the optimal lane-change trajectories. Based on the previous framework, we will investigate how to provable guarantee of stability of the learned policy as suggest in [9] by iteratively minimizing the Lyapunov risk. Secondly, we will investigate how to integrate Off-Policy Deep Reinforcement Learning [4] to learn from a batch of data and address the extrapolation error problem in dynamic changing environment.

great

References

[1] Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L.D., Monfort, M., Muller, U., Zhang, J., et al.: End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316 (2016)

[2] Xu, H., Gao, Y., Yu, F., Darrell, T.: End-to-end learning of driving models from large-scale video datasets. CVPR (2017)

[3] Kim,J.,Canny,J.:Interpretable learning for self-driving cars by visualizing causal attention. ICCV (2017)

[4] Fujimoto, S., Meger, D., Precup, D. (2019, May). Off-policy deep reinforcement learning without exploration. In International Conference on Machine Learning (pp. 2052-2062)

[5] Codevilla, F., Miiller, M., López, A., Koltun, V., Dosovitskiy, A. (2018, May). End-to-end driving via conditional imitation learning. In 2018 IEEE International Conference on Robotics and Automation (ICRA) (pp. 1-9). IEEE

[6] Silver,D.,Bagnell,J.A.,Stentz,A.:Learning from demonstration for autonomous navigation in complex unstructured terrain. The International Journal of Robotics Research 29(12) (2010) 1565–1592

[7] Sadat, A., Ren, M., Pokrovsky, A., Lin, Y. C., Yumer, E., Urtasun, R. (2019). Jointly learnable behavior and trajectory planning for self-driving vehicles. arXiv preprint arXiv:1910.04586

[8] Xi, C., Shi, T. (2019). Efficient Motion Planning for Automated Lane Change based on Imitation Learning and Mixed-Integer Optimization. arXiv preprint arXiv:1904.08784.

[9] Chang, Y. C., Roohi, N., Gao, S. (2019). Neural lyapunov control. In Advances in Neural Information Processing Systems(NIPS) (pp. 3245-3254)

In related work, you can describe off-policy deep reinforcement learning and extrapolation error