A Short Review on
"Deep Reinforcement Learning with Enhanced Safety for Autonomous Highway Driving"

Matin Moezzi
Department of Mathematics and Computer Science
Amirkabir University of Technology (Tehran Polytechnic)
matin.moezzi@gmail.com

I. Introduction

Recent successes in deep learning (DL) and reinforcement learning (RL) have led many researchers to employ deep RL in autonomous driving as a decision-making framework. Safety assurance in autonomous driving is a critical problem due to the uncertainty of the environment and unpredictable behavior of other drivers; therefore, many efforts have been made to ensure safety in autonomous driving. Professor Ali Baheri and his colleagues proposed a safe deep RL method consisting of two safety modules: the handcrafted rule-based and dynamically-learned module [1]. The dynamically-learned module comprises an offline trained RNN to predict future states and a handcrafted module that indicates whether future states lead to an accident. If so, that state is stored in the collision buffer with the negative reward; otherwise, it is stored in the safe buffer. This one-page summary discusses a challenge encountered by this method and proposes IRL as a possible solution.

II. Handcrafted versus Data-Driven Module

The handcrafted single-rule module cannot enumerate all possible driving rules and norms to deal with all the states. Therefore, the rule-based method fails to identify safe states when facing unexpected behavior from traffic cars such as sudden breaks, accelerations, and lane changes [2]. The following rule is used to distinguish a safe state from a collision state based on the relative distance between vehicles and a handcrafted threshold.

$$d_{TV} - T_{min} \times v_{TV} > d_{TV_{min}}$$

This rule-based method has two drawbacks: 1) It does not take The dynamics of the environment and traffic cars' behavior into account. 2) The rule classifies states into two categories (safe or collision), causing sparse reward.

Inverse RL (IRL) can approximate the unknown reward function using observed behaviors. Sharifzadeh et al. [3] have proposed an IRL approach to extract the reward function for large state-spaces and implemented their solution for autonomous driving. The IRL approach mitigates both of the mentioned issues. Unlike the rule-based method, IRL gives a real-valued reward function that evaluates the states' safety level continuously, preventing reward sparsity. In addition, IRL extracts this function based on previously observed behaviors; therefore, the environment's dynamic and traffic cars' behavior will directly affect the reward value. Furthermore, in the IRL method, the reward depends on the previous state and the action.

III. Deep RL and Safety Control Come Together

In the field of cognitive science, there are two major learning paradigms, empiricism, and speculation. Empiricism is a way of learning from historical experiences. Speculation is the way of logical thinking, which means taking measures by reasoning. The thinking process in human beings contains both empiricism and speculation, which are interactive during the process [4].

Deep RL would make unpredicted decisions in unfamiliar scenarios, which is also the shortcoming of the data-driven method. In addition, avoiding collision is the main goal when designing the control strategy for autonomous driving. The combination of deep RL and safety-based control for autonomous driving is proposed in [4]. Xiong et al. [4] combine artificial potential field method and path tracking, two safety control methods, as the speculation and logical thinking with the DDPG algorithm, which is very similar to the empiricism paradigm, to address the collision avoidance problem. First, the vehicle learns the driving policy via DDPG in a stable and familiar environment without any vehicles around. Then, safety-based control is used to design a collision avoidance mechanism with vehicles around.

References

[1] A. Baheri, S. Nageshrao, H. E. Tseng, I. Kolmanovsky, A. Girard, and D. Filev, “Deep reinforcement learning with enhanced safety for autonomous highway driving,” 2020.

[2] Z. Cao, E. Biyik, W. Z. Wang, A. Raventos, A. Gaidon, G. Rosman, and D. Sadigh, “Reinforcement learning based control of imitative policies for near-accident driving,” *CoRR*, vol. abs/2007.00178, 2020. [Online]. Available: https://arxiv.org/abs/2007.00178

[3] S. Sharifzadeh, I. Chiotellis, R. Triebel, and D. Cremers, “Learning to drive using inverse reinforcement learning and deep q-networks,” *CoRR*, vol. abs/1612.03653, 2016. [Online]. Available: http://arxiv.org/abs/1612.03653

[4] X. Xiong, J. Wang, F. Zhang, and K. Li, “Combining deep reinforcement learning and safety based control for autonomous driving,” *CoRR*, vol. abs/1612.00147, 2016. [Online]. Available: http://arxiv.org/abs/1612.00147

[5] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. K. Yogamani, and P. Pérez, “Deep reinforcement learning for autonomous driving: A survey,” *CoRR*, vol. abs/2002.00444, 2020. [Online]. Available: https://arxiv.org/abs/2002.00444

[6] C. Jianyu, W. Zining, and T. Masayoshi, “Deep hierarchical reinforcement learning for autonomous driving with distinct behaviors,” in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 1239–1244.

[7] V. Talpaert, I. Sobh, B. R. Kiran, P. Mannion, S. K. Yogamani, A. E. Sallab, and P. Pérez, “Exploring applications of deep reinforcement learning for real-world autonomous driving systems,” *CoRR*, vol. abs/1901.01536, 2019. [Online]. Available: http://arxiv.org/abs/1901.01536

[8] M. Hügle, G. Kalweit, B. Mirchevska, M. Werling, and J. Boedecker, “Dynamic input for deep reinforcement learning in autonomous driving,” *CoRR*, vol. abs/1907.10994, 2019. [Online]. Available: http://arxiv.org/abs/1907.10994

[9] B. Mirchevska, C. Pek, M. Werling, M. Althoff, and J. Boedecker, “High-level decision making for safe and reasonable autonomous lane changing using reinforcement learning,” in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 2156–2162.

[10] “Deep reinforcement learning for autonomous driving,” *CoRR*, vol. abs/1811.11329, 2018, withdrawn. [Online]. Available: http://arxiv.org/abs/1811.11329

[11] A. E. Sallab, M. Abdou, E. Perot, and S. Yogamani, “Deep reinforcement learning framework for autonomous driving,” *Electronic Imaging*, vol. 2017, no. 19, pp. 70–76, 2017. [Online]. Available: https://www.ingentaconnect.com/content/ist/ei/2017/00002017/00000019/art00012

[12] S. Kuutti, R. Bowden, and S. Fallah, “Weakly supervised reinforcement learning for autonomous highway driving via virtual safety cages,” *CoRR*, vol. abs/2103.09726, 2021. [Online]. Available: https://arxiv.org/abs/2103.09726

[13] S. Kuutti, R. Bowden, H. Joshi, R. Temple, and S. Fallah, *Safe Deep Neural Network-Driven Autonomous Vehicles Using Software Safety Cages*, 10 2019, pp. 150–160.