Social Compliant Automated Driving

I propose to incorporate Deep Reinforcement Learning (DRL), an adversarial learning mechanism together with social psychological factors (e.g., Social Value Orientation, SVO) with uncertainty into the model-predictive policy learning framework proposed in (Henaff et al., 2019) to tackle challenging driving manoeuvres involving both longitudinal and lateral control, e.g., on-ramp merging, driving (merging) during weaving segments, left (right) turn at non-priority/unprotected intersections. Adversarial learning and training can be implemented under the Naturalistic and Adversarial Driving Environment (NADE) (Feng et al., 2021) modelled with the Markov decision process (MDP) and probabilistic graphical models (PGM) and with additional information regarding social psychological factors if data are available. With the fine-constructed rewards function considering both safety, efficiency, and social compliance, the DRL model will be trained and upgraded on simulation platforms, e.g., SUMO, OpenCDA (Xu et al., 2021), Berkeley Flow (Wu et al., 2017), employing policy gradient methods, e.g., Proximal Policy Optimization (PPO) (Schulman et al., 2017), to ensure efficient, safe, and social compliant automated vehicle driving behaviour.

For social compliance, social psychology metrics that could reflect people's interactions will be considered. A promising candidate is Social Value Orientation, which indicates a person's preference about how to allocate resources between the self and another person and could represent how much weight a person attaches to the welfare of others in relation to their own. Social norms will be included, which is also very crucial for safety. Other suitable measures will also be positively explored. Other ways, suggestions, and advice on how to model social compliance and implement human-like social driving would be highly welcomed.

The interested driving scenes will be roundabouts, on-ramp merging, weaving sections, etc, which should involve both longitudinal and lateral control.

Another direction is to develop mechanisms to connect automated vehicles' sensing (with uncertainties and minor errors) with AVs' control. Since 100% of accuracy in sensing the environment is not currently (and in a short term) possible, the developed AVs' controlling algorithms should tolerate uncertainties and minor sensing errors. Therefore to bridge the not-perfect sensing with AVs' controlling would be a meaningful research direction.

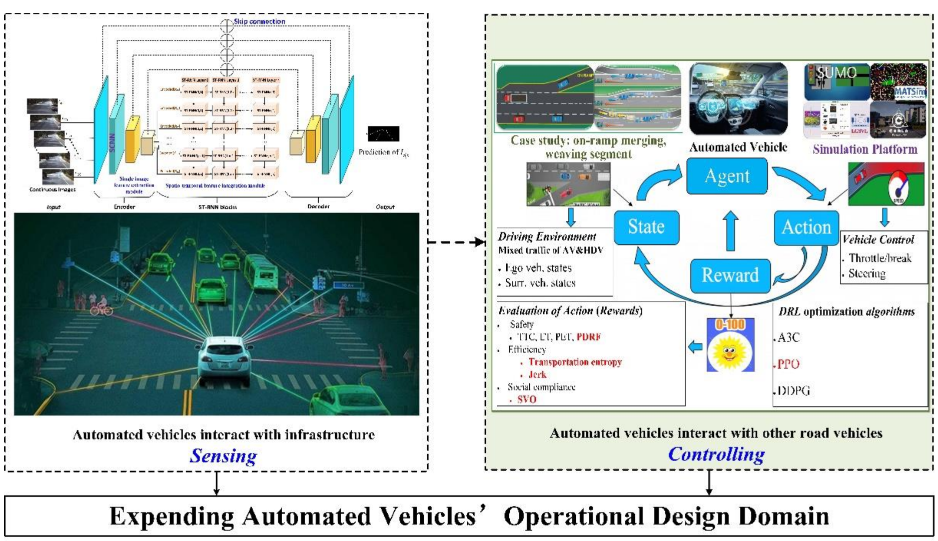

Figure 1. A framework illustration for connecting automated vehicles' sensing with controlling

Reference

Feng, S., Yan, X., Sun, H., Feng, Y., Liu, H.X., 2021. Intelligent driving intelligence test for autonomous vehicles with naturalistic and adversarial environment. Nat. Commun. 12, 1–14. https://doi.org/10.1038/s41467-021-21007-8

Henaff, M., LeCun, Y., Canziani, A., 2019. Model-predictive policy learning with uncertainty regularization for driving in dense traffic. 7th Int. Conf. Learn. Represent. ICLR 2019 1–20.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O., 2017. Proximal Policy Optimization Algorithms 1–12. *arXiv preprint arXiv:1707.06347*.

Wu, C., Kreidieh, A., Parvate, K., Vinitsky, E., Bayen, A.M., 2017. Flow: A Modular Learning Framework for Autonomy in Traffic 1–18. *arXiv preprint arXiv:1710.05465*.

Xu, R., Guo, Y., Han, X., Xia, X., Xiang, H., Ma, J., 2021. OpenCDA:An Open Cooperative Driving Automation Framework Integrated with Co-Simulation. 2021 IEEE Int. Intell. Transp. Syst. Conf. 1155–1162. https://doi.org/10.1109/ITSC48978.2021.9564825

[Note: This is my second contribution to TRAIL Congress 2022. This research is at an early stage. Therefore looking forward to feedback and suggestions.]