Lane-change decisions of connected autonomous vehicles using spatially-weighted information and deep reinforcement learning

Jiqian Dong[1,2], Sikai Chen[1,2,3,*], Paul (Young Joun) Ha[1,2], Runjia Du[1,2], Yujie Li[1,2], Samuel Labi[1,2]

[1] U.S. DOT Center for Connected and Automated Transportation (CCAT), Purdue University, West Lafayette, IN, USA
[2] Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA
[3] Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

Keywords: Deep reinforcement learning, lane changing decision, dynamic length input, connectivity range, autonomous vehicle.

1. Introduction

Due to the complexity of the CAV driving environment including the varying number of cars surrounding the CAV, the size of decision-making inputs to the CAV controller may vary considerably in each time step. Traditional control and planning techniques may fail because they cannot directly manipulate the variable size of the input. This study presents a method to transfer the original inputs into fixed-size inputs using Deep Reinforcement Learning (DRL). The paper's findings suggest that a DRL-based intelligent controller can perform efficiently. There is abundant research in the literature regarding the fixed input size, e.g., Mirchevska et al. (2017) assumes that the model can process only 20 features from surrounding vehicles; also, Saxena et al. (2019) use representations of a fixed number of occupancy grids. While these models are considered state-of-the-art, imposing a fixed size restricts the number of vehicles in the environment that the CAV can perceive. Further, to extract useful information, grid representation requires a Convolutional Neural Network (CNN) which has two shortcomings: (1) less precision due to its black-box property; (2) extreme computational expensiveness that limits integration in real-time decision-making.

To address the shortcomings of fixed-size representation as the CAV learns appropriate control policies, Huegle et al. (2019) combined Deep Sets (Zaheer et al., 2017) and Q-learning(Watkins and Dayan, 1992). However, due to Deep Sets's simple summation manipulation, the high-dimension features are condensed into a single fixed-size vector. Useful information such as downstream vehicles' speeds, locations, and lane positions are lost in the process. Additionally, the value of information among surrounding vehicles is not explicitly considered. As such, vehicles far away and vehicles nearby are treated equally through the same embedding network. Another research gap is that the Huegle DRL model is trained on datasets that only had successful and safe lane-changing transitions. Therefore, their model has limitations because collision-free decisions are never guaranteed in reality.

The main contributions of this paper are:

1) A DRL-based method that fuses local and system-wide information via Deep Sets modification.
2) An end-to-end framework that controls the CAV's collision-free lane-changing decision on the fused information.
3) An assessment of the relationship between connectivity range sufficiency and traffic density.

2. Methodology

Model overview

We use a classical Deep Q Network (DQN) to fuse all the input information, and output the CAV's high-level lane-changing decision (Figure 1). For inputs, we consider 3 blocks: downstream information (within connectivity range), locality information (within sensor range) and ego CAV information. For each component, we all use multilayer perceptron (MLP) with detailed structure as follows.

- Encoding network: $\mathrm{Dense}(64) + \mathrm{Dense}(32)$
- Q network: $3 \times \mathrm{Dense}(64) + \mathrm{Dense}(32) + \mathrm{Dense}(16) + \mathrm{Dense}(8)$
- Output layer: $\mathrm{Dense}(3)$

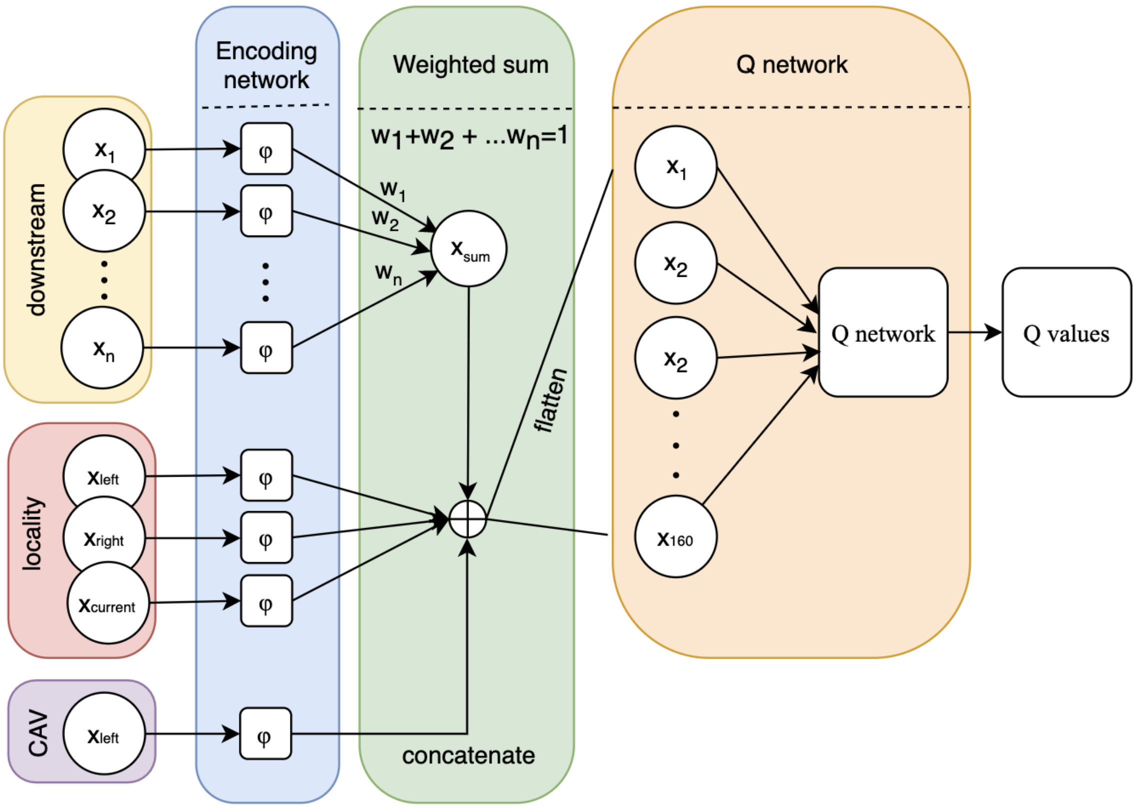

Figure 1. Proposed network architecture

3. Experiment settings

3.1 State space

We consider the following features in the input representation: relative distance, relative speed and relative lane.

$$\mathcal{S} = \begin{pmatrix} Downstream \\ Locality \\ CAV_{info} \end{pmatrix}$$

3.2 Weights

To explicitly weight the dynamic inputs, we define the weights as inverse proportional to the distance to the ego CAV, and all sum up to 1.

$$w_i = \frac{1/dl_i}{\sum_i^n 1/dl_1}$$

3.3 Action space

The action space is discrete for each time step indicating the possible actions the CAV can perform. $\mathcal{A} = \{change\ to\ left, keep\ lane, change\ to\ right\}$.

3.4 Reward function

The reward function consists of 2 rewards and 2 penalties: Speed reward R_v, destination reward R_D, collision penalty P_c and lane-changing penalty P_{LC}.
The overall reward function is defined as:

$$R_{total} = w_1 R_v + w_2 R_D - w_3 P_c - w_4 P_{LC}$$

Where: $w_1\ to\ w_4$ are the weights that can be tuned as a tradeoff between a vehicle's "aggression" and its "comfort".

4. Results

4.1 Comparative analysis

Our proposed model is compared with the 3 baseline models including: the unweighted classical Deep Set Q learning model, rule-based lane-changing model and no lane changing model. The mean and median performance are compared in Figure 2 and Table 1.

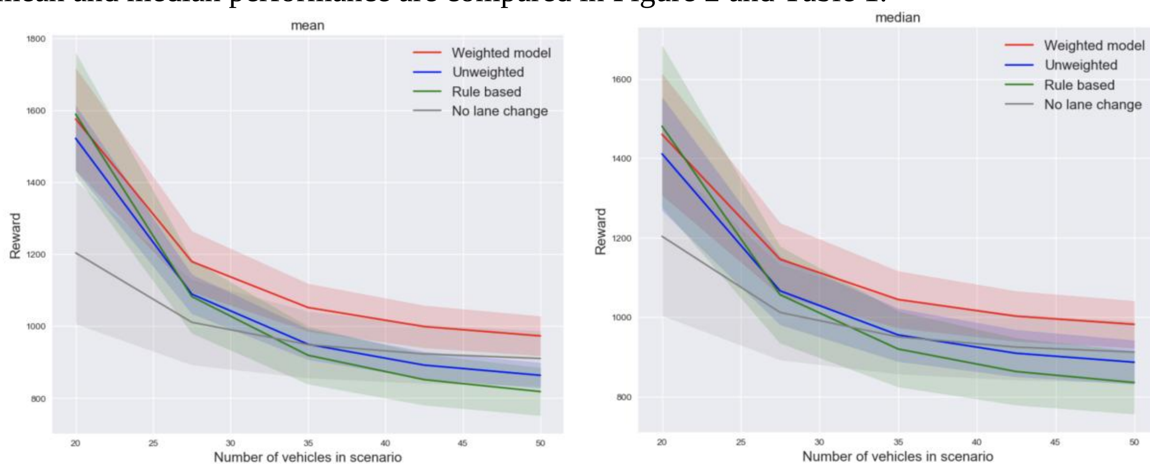

Figure 2. Mean and median performance for 10 episodes of test

Table 1 Performance comparison for different models in different scenarios.

number of vechicles	20			30			40			50		
models	mean	median	std	mean	median	std	mean	median	std	mean	median	std
No lane changing	1189.8	1191	19.5	1066.9	1062.4	21.8	828.1	825.5	51.9	942.2	951.5	28.1
Rule based (LC2013)	1570.4	1456.6	359.2	1103.6	1112.1	87.4	801	806.5	47.1	810.9	813.1	34.3
Unweighted input	1510.7	1393.3	287.2	1071.6	1085.5	38.4	914.7	935.7	79.2	822.4	820.2	26.3
Weighted input	1559.5	1442.2	382.4	1182.2	1166.8	42.3	1039.9	1039.8	30.5	902.7	906.6	25.4

It was observed that the proposed approach is most effective in the scenarios with 30 and 40 vehicles. Also, it was seen that the CAV efficiency is greatly enhanced, and due to the proactive nature of its decision-making, the proposed weighted model outperforms all the baseline models. It was also observed that in highly congested scenarios, all the models performed poorly due to a lack of available space for vehicle maneuvering; on the other hand, at extremely low density traffic, the rule-based lane-change model exhibited the best performance, as expected.

4.2 Optimal connectivity range analysis

Further, the proposed model can be used to identify the CAV's optimal connectivity range. As shown in Figure 3, as the connectivity range increases, the performance of the model increases dramatically initially and slows down with further increments of the range. The reason for this is initially, an increase in the connectivity range will introduce more downstream information to the CAV, which will be beneficial to the decision processor for taking proactive decisions. However, increasing the connectivity range will increase the variance due to the introduction of noise or unrelated information that in turn arises from the nature of human-driven vehicle (HDV) operations. From the results (Figure 3), the optimal connectivity range is approximately 300 meters; additional increments in the connectivity range yields little or no reward to the CAV.

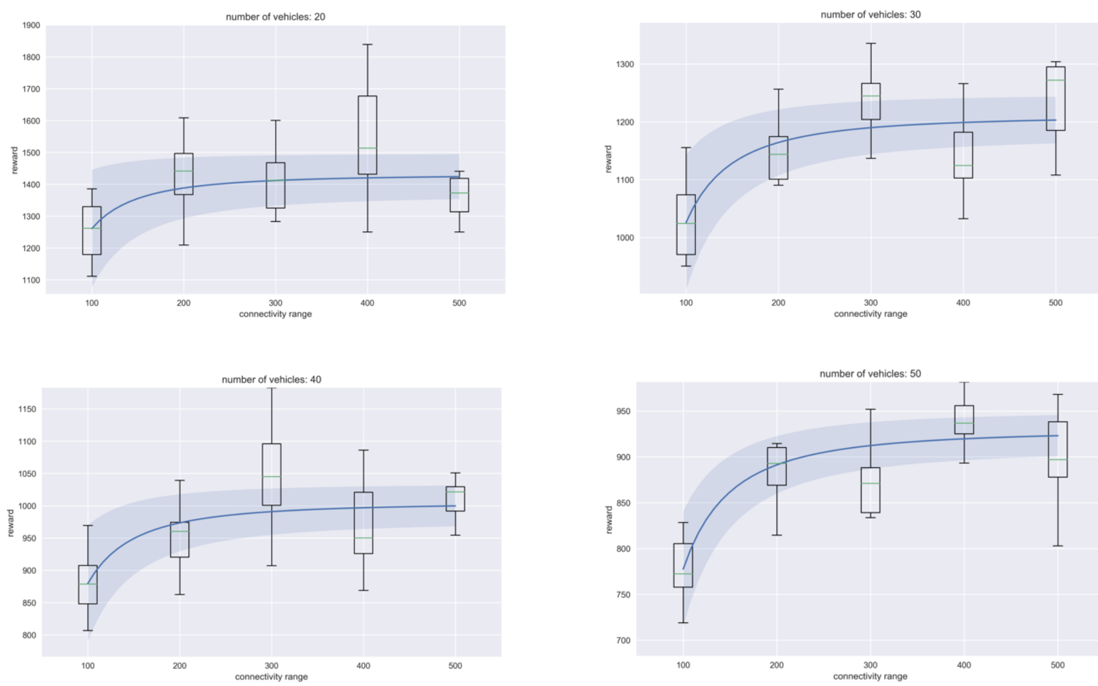

Figure 3. Reward vs connectivity range in different density scenario

5. Conclusion

In this paper, we present an end-to-end DRL based processor to make high level decisions to control CAV lane changes. We demonstrate that the model has the capability to increase the CAV efficiency in mixed traffic. Future research can consider using connectivity and a data storage system to process temporal information such as historical data on the vehicle position, speed, and acceleration over a longer time period, and to use this data in the CAV's decision process. Using such historical data, it will be easier to recognize the existence of barrier situations such as downstream workzones, accidents, or potholes that will require re-routing or preemptive evasive maneuvers. In addition, the DRL based method can be used in future research to make collaborative decisions that not only maximize the CAV's individual utility but also benefit all the agents in the entire network. Examples of such future work include studies of traffic string stability enhancement and cooperative crash avoidance in emergency situations.

References

Huegle, M., Kalweit, G., Mirchevska, B., Werling, M., Boedecker, J., 2019. Dynamic Input for Deep Reinforcement Learning in Autonomous Driving.

Mirchevska, B., Blum, M., Louis, L., Boedecker, J., Werling, M., 2017. Reinforcement Learning for Autonomous Maneuvering in Highway Scenarios Fitted Q-iteration with Extremely Randomized Trees 32–41.

Saxena, D.M., Bae, S., Nakhaei, A., Fujimura, K., Likhachev, M., 2019. Driving in Dense Traffic with Model-Free Reinforcement Learning.

Watkins, C.J.C.H., Dayan, P., 1992. Q-learning. Mach. Learn. https://doi.org/10.1007/bf00992698

Zaheer, M., Kottur, S., Ravanbhakhsh, S., Póczos, B., Salakhutdinov, R., Smola, A.J., 2017. Deep sets. Adv. Neural Inf. Process. Syst. 2017-Decem, 3392–3402.