Optimal Decision-Making Strategies for Self-Driving Car Inspired by Game Theory

Kyoungtae Ji
School of Mechanical Engineering, Kyungpook National University,
Daegu 41566, South Korea
Email: wlrudxo644@knu.ac.kr

Kyoungseok Han*
School of Mechanical Engineering, Kyungpook National University,
Daegu 41566, South Korea
Email: kyoungsh@knu.ac.kr

***Abstract*—This paper presents an optimal decision-making strategy for a self-driving car using a game-theoretic approach. To ensure the safety of the decision, Stackelberg game's maximin reward strategy, which considers concurrency, is applied. The receding horizon is included to increase the accuracy of the decision, but the computational burden is high. We assume that the follower takes only one prediction time, not the receding horizon, to relieve the computational burden. For an accurate prediction of interacting vehicles, the intention estimation model is suggested. We demonstrate the efficiency of our approach in a simulation environment and various traffic conditions.**

Keywords—self-driving car; decision-making; game theory

I. Introduction

The commercialization of self-driving cars (SCs) is currently underway. It will, however, take a long time for all vehicles to become self-driving. Interaction between autonomous vehicles and human drivers is unavoidable until all vehicles are driven automatically. If the interaction is not taken into account, the SC should drive cautiously to ensure safety. As a result, developing appropriate driving strategies in situations where SCs and human drivers coexist on the road is critical.

To resolve this problem, in this work, we propose to develop a game-theoretic decision-making approach of an SC when interacting with human drivers. The basic assumption of game theory is that "all game players are rational decision makers who want to maximize rewards." In this manner, SCs assume the interacting vehicle to be rational decision makers and take optimal action [1].

However, in real-world driving, each vehicle drives based on its strategy without following the game's rules. Human drivers interact smoothly using gestures and social norms, but it is difficult for SCs. To overcome this problem, an estimated driving strategy of the interacting vehicle is needed based on their behavior [2]. We assume that the human drivers have their intention, such as "politeness," and follow the game rule based on their intention. Politeness is defined as the intention of yield, represented by headway and acceleration [3]. During the interaction, the SC observes the opponent vehicle's behavior and estimates its politeness in real-time.

There are some representative approaches of game theory, such as Nash and Stackelberg. The Nash equilibrium is an approach to maximize reward based on an opponent's behavior [4]. However, in the Nash equilibrium, the worst case, such as a collision, is not considered significant. In contrast, the Stackelberg game approach is a robust game to maximize the worst-case reward for the opponent's aggressive behavior. We use the Stackelberg game approach to ensure driving safety.

We deal with various traffic scenarios where interaction with surrounding vehicles is unavoidable to validate our approach's efficacy. More specifically, lane merging in congested low-speed traffic and overtaking on highways when the vehicle ahead is not moving quickly enough are some of the scenarios.

This short paper is organized as follows: In Section II we propose the decision-making method of SCs using a game-theoretic approach. Simulation results of our approach for various traffic scenarios are presented in Section III, and the conclusion is given in Section IV.

II. Game Theory-based decision-making

A. Reward Function

In game theory, the goal of all game participants is the reward function. It should conduct a thorough assessment of their circumstances and the surrounding environment. In a driving situation, the overall goal is to avoid collisions, drive at the desired speed, and maintain a safe headway distance. Based on this, the basic form of the reward function is as follows [5]:

$$R^k = w_1\hat{v} + w_2\hat{s} + w_3\hat{h} \tag{1}$$

where $\hat{v}$ denotes velocity reword, $\hat{s}$ for safety rewards and $\hat{h}$ is the headway reward. $w_i, (i = 1,2,3)$ is a weight coefficient of each parameter that can be controlled for the importance of the parameters and intention of each driver. R^k is the reward for the current time step. Based on this equation, adjustable parameters can be added for specific objectives.

For more accurate decisions, the reward function of an SC accumulates in multiple time steps rather than one prediction time. The cumulative reward for receding horizon is given as follows:

$$R^t = \sum_{k=1}^{n} \lambda^k R^k \quad (2)$$

where n denotes the time step for k = 1, 2, ···, n. R^t is the total reward of n time step, and λ^k is the discount factor.

B. Stackelberg Game Approach

The Stackelberg game approach is a hierarchical, sequential game that the game participants are assumed to be leaders and followers. Leaders always act first. Followers observe the leader's behavior and then choose their behavior.

In this paper, we discuss traffic scenarios in which SCs drive to achieve their goals, such as merging and overtaking lanes, and how surrounding vehicles respond to the behavior of an SC. As a result, we assign SCs the role of a leader and human drivers the role of a follower.

The leader predicts the follower's reaction and takes optimal action based on the worst-predicted reaction of the follower. In this manner, the leader can ensure the minimum reward for the follower's aggressive action. The basic Stackelberg equilibrium is given as follows [6]:

$$\kappa_l^* \in \underset{\kappa_l \in \Gamma_l}{\operatorname{argmax}} \left(\min_{\kappa_f \in \Gamma_f^*(\kappa_l)} R_l(\kappa_l, \kappa_f) \right) \quad (3)$$

$$\Gamma_f^*(\kappa_l) \overset{\text{def}}{=} \{\kappa_f^* \in \Gamma_f \colon R_f(\kappa_l, \kappa_f^*) \geq, \forall \kappa_f \in \Gamma_f\} \quad (4)$$

where κ denotes possible action based on action space Γ, κ^* is the optimal action in optimal action space Γ^*, and R is the reward of the vehicles. The subscript l and f stand for leader and follower. The leader assumes that the follower takes optimal reaction κ_f^* according to the leader's action κ_l. The leader takes optimal action κ_l^* using the predicted $\Gamma_f^*(\kappa_l)$ to maximize the worst $R_l(\kappa_l, \kappa_f)$ (Maximin strategy).

The basic assumption of the Stackelberg game approach is the sequential action selection. However, in real-world driving, all vehicles drive based on current information, not sequential. To improve this unrealistic assumption, the follower's optimal action considering concurrency is given as [7]:

$$\kappa_f^* \in \underset{\kappa_f \in \Gamma_f}{\operatorname{argmax}} \left((\min_{\kappa_l \in \Gamma_l} R_f(\kappa_l, \kappa_f) \right) \quad (5)$$

The follower assumes that maximizing the minimum reward by the leader's possible behavior is the optimal behavior. Moreover, applying (2) can increase the accuracy of the decision, but it is not suitable for real-time due to the high computational burden. To relieve the computational burden, we include the reasonable assumption that human driver does not consider receding horizon, only consider the mid-term interaction.

The Stackelberg equilibrium reflects (2) and (5) can be expressed as follows:

$$\mathrm{K}_l^* \in \underset{\kappa_l \in \Gamma_l}{\operatorname{argmax}} \left(\min_{\kappa_f \in \Gamma_f^*} R_l^t(\kappa_l, \kappa_f) \right) \quad (6)$$

$$\Gamma_f^* \overset{\text{def}}{=} \left\{ \kappa_f^* \in \Gamma_f \colon \min_{\kappa_l \in \Gamma_l} \left(R_f^k(\kappa_l, \kappa_f^*) \right) \geq \min_{\kappa_l \in \Gamma_l} \left(R_f^k(\kappa_l, \kappa_f) \right), \forall \kappa_f \in \Gamma_f \right\} \quad (7)$$

where K_l^* is the optimal action combination of leader for n time step, i.e., $\mathrm{K}_l = [\kappa_l^1, \kappa_l^2, \cdots, \kappa_l^n]$. R_l^t is the cumulated reward of leader based on (2), R_f^k is the reward of follower at long-term one-time step.

The leader takes optimal action κ_l^* based on predicted optimal actions of the follower Γ_f^*. In this way, the leaders can prepare for possible actions of the follower, except highly unreasonable actions.

C. Real-Time Intention Estimation

In traffic scenarios that we deal with, the goal of the SC is similar to changing the lane. The interacting vehicle is located in the next lane and observes lane-change signals from the SCs. In that case, the interacting vehicle intends to ignore or yield, and it appears to be an acceleration change. We define the intention of yield as "Politeness" that $P \in [0,1]$. Based on this, the real-time intention estimation using velocity information that is observed by the SC is given as follows:

$$P(t+1) = \frac{P(t) + \alpha}{1 + \beta} \quad (8)$$

where $P(t) \in [0,1]$ is the estimated politeness of interacting vehicle at time step t, α and β are tunable parameters that $\alpha \in [0, \beta]$, $\beta \in (0, \infty]$. When a SC turns on the lane-change signal, if interacting vehicle decelerates, then the α is set to close to β for $P(t+1)$ increment. In contrast, if interacting vehicle accelerate or maintain the speed, then the α is set to close to 0 for $P(t+1)$ decrease.

III. Case Studies

In the following, we simulate two different traffic conditions to verify our approach: 1) merging lanes of SCs at the end of the lane under heavy traffic and 2) overtaking when front vehicles are slow. What the two situations have in common is that lane changes are made through interaction with surrounding vehicles. We modify only the reward function and apply the same decision-making strategy to an SC. The initial traffic condition is performed in 0 seconds of Fig. 1 and Fig. 2.

A. Test Environment Statement

We present simulation-based verification of our decision-making approach in case studies. Because vehicle dynamics are not the focus of our work, all game player's behavior is limited to finite discrete actions. Furthermore, the intelligent driver model, a reliable longitudinal car-following model [8], modified the surrounding vehicles to interact with an SC. Details of the simulation environment are omitted in this paper. The interested readers are referred to our previous paper [9]:

B. Simulation Results

The results of two different scenarios are shown in Fig. 1 and Fig. 2. The simulation starts at 0 seconds and displays snapshots every few seconds during the simulation. The red mark is an SC, and the blue ones are human vehicles (HVs).

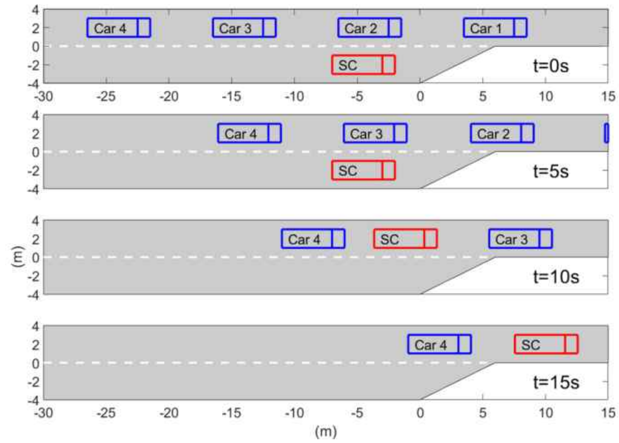

Fig. 1. Simulation of game theory-based lane merge of a self-driving car

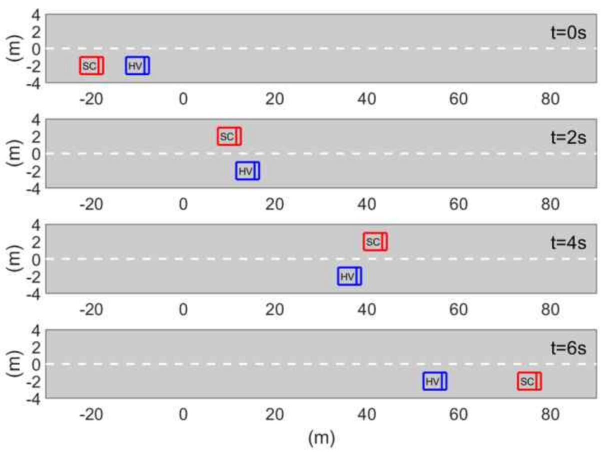

Fig. 2. Simulation of game theory-based overtaking of a self-driving car

Fig. 3 shows the estimated politeness of interacting vehicles for each scenario. The lane-change decision of an SC is based on game theory, and the estimated intention of interacting vehicles is shown in Fig. 3.

In Fig. 1, an SC presents the intention of lane change using lane-change signals for lane merge. Lane-merge interaction with Car 3 fails at 5 seconds, and Car 4 yields at 10 seconds to succeed in a lane merge. In Fig. 3(a), the estimated intention of Car 3 is "not yield," and the predicted action of Car 3 is to interrupt the lane change. After waiting for Car 3 to go, the estimated intention of Car 4 is "yield," and the predicted action of Car 4 is decelerating to create enough gap for lane change. Based on the Stackelberg game approach, an SC changes the lane at 10 seconds.

In Fig. 2, the human vehicle drives at a low speed, so an SC overtakes for a high reward. In this case, we included an overtaking lane penalty $w_4\hat{o}$ in (1). When an SC gives the intention of lane change, HVs decelerates to yield. As shown in Fig. 3(b), the estimated politeness of HVs is high and

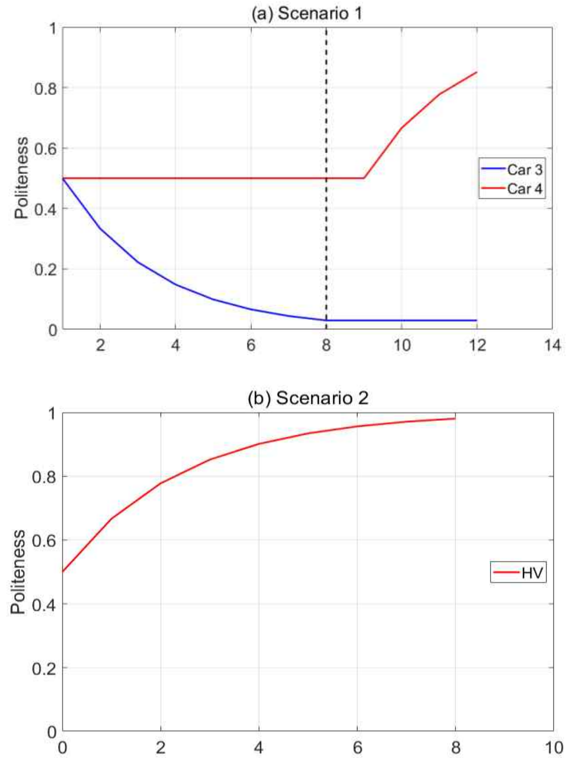

Fig. 3. (a) Estimated intention of Cars 3 and 4 for lane merge. (b) Estimated intention of a human vehicle for overtaking.

the predicted behavior of HVs based on the Stackelberg game does not prevent lane change, so the SC returns to the original lane properly.

IV. CONCLUSION

The interaction between an SC and a human driver is an area that needs more research. This paper presents the game theory-based optimal decision-making strategy of SCs when interacting with HVs. To overcome the unrealistic assumption of the basic Stackelberg game approach, we modified the Stackelberg game to take concurrency into account. To reduce the computational burden, we use the Stackelberg game approach to predict human driver behavior. Because predicting behavior requires the intention of interacting vehicles, a real-time intention estimation model based on speed variations is proposed. Our approach is validated using simulations for various traffic scenarios, and SCs successfully achieve their goals. In the future, we will generalize to be suitable for various scenarios.

ACKNOWLEDGMENT

This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2021R1A6A1A03043144); in part by the NRF grant, funded by the Korean government (MSIT) (NRF-2020K1A3A1A39112277).

REFERENCES

[1] R. B. Myerson, *Game Theory*. Harvard University Press, 2013.

[2] G. S. Sankar and K. Han, "Adaptive robust game-theoretic decision making strategy for autonomous vehicles in highway," *IEEE Transactions on Vehicular Technology,* vol. 69, no. 12, p. 14484, 2020.

[3] Q. Zhang, R. Langari, H. E. Tseng, D. Filev, S. Szwabowski, and S. Coskun, "A game theoretic model predictive controller with aggressiveness estimation for mandatory lane change," *IEEE Transactions on Intelligent Vehicles,* vol. 5, no. 1, p. 75, 2020.

[4] L. Sun, M. Cai, W. Zhan, and M. Tomizuka, "A game-theoretic strategy-aware interaction algorithm with validation on real traffic data," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, p. 11038, 2020.

[5] Li, Nan, et al. "Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems." IEEE Transactions on control systems technology 26.5 (2017): 1782-1797.

[6] T. Başar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1998.

[7] N. Li, Y. Yao, I. Kolmanovsky, E. Atkins, and A. R. Girard, "Game-theoretic modeling of multi-vehicle interactions at uncontrolled intersections," *IEEE Transactions on Intelligent Transportation Systems,* 2020.

[8] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical Review E, Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 62, no. 2, p. 1805, 2000.

[9] K. Ji, M. Orsag, and K. Han, "Lane-merging strategy for a self-driving car in dense traffic using the Stackelberg game approach," *Electronics,* vol. 10, no. 8, p. 894, 2021.