

Cooperative Driving of Connected Autonomous Vehicles in Heterogeneous Mixed Traffic: A Game Theoretic Approach

Shiyu Fang, Peng Hang, *Member, IEEE*, Chongfeng Wei, *Member, IEEE*,
Yang Xing, *Member, IEEE*, and Jian Sun

Abstract—High-density, unsignalized intersections have always been a bottleneck of efficiency and safety. The emergence of Connected Autonomous Vehicles (CAVs) results in a mixed traffic condition, further increasing the complexity of the transportation system. Against this background, this paper aims to study the intricate and heterogeneous interaction of vehicles and conflict resolution at the high-density, mixed, unsignalized intersection. Theoretical insights about the interaction between CAVs and Human-driven Vehicles (HVs) and the cooperation of CAVs are synthesized, based on which a novel cooperative decision-making framework in heterogeneous mixed traffic is proposed. Normalized Cooperative game is concatenated with Level-k game (NCL game) to generate a system optimal solution. Then Lattice planner generates the optimal and collision-free trajectories for CAVs. To reproduce HVs in mixed traffic, interactions from naturalistic human driving data are extracted as prior knowledge. Non-cooperative game and Inverse Reinforcement Learning (IRL) are integrated to mimic the decision-making of heterogeneous HVs. Finally, three cases are conducted to verify the performance of the proposed algorithm, including the comparative analysis with different methods, the case study under different Rates of Penetration (ROP) and the interaction analysis with heterogeneous HVs. It is found that the proposed cooperative decision-making framework is beneficial to driving conflict resolution and the traffic efficiency improvement of the mixed unsignalized intersection. Besides, due to the consideration of driving heterogeneity, better human-machine interaction and cooperation can be realized in this paper.

Index Terms—connected autonomous vehicles, heterogeneous mixed traffic, unsignalized intersection, level-k game, inverse reinforcement learning, cooperative driving

I. INTRODUCTION

THE past decade has witnessed the rapid development of Connected Autonomous Vehicles (CAVs) [1]. CAVs have

a wider perception range and a smaller perception error, giving them the potential to detect danger earlier. This further enables them to actively cooperate with other vehicles, resolving conflicts by properly arranging the Right of Way (ROW). However, challenges persist. On account of the complexity of real-world situations, CAVs may behave unreasonably when interacting with Human Vehicles (HVs) which may lead to collision or deadlock. Additionally, there is a further question about how to enhance system performance by coordinating the relationship between vehicles.

For decades, one of the ideas to ameliorate the above problem is to explore the latent rules and patterns from human interactions. Numerous studies have analyzed drivers' decision-making logic, cognitive structure, and inherent characteristics to design a human-like CAV. Game theory formulates the interaction between incentive structures and the relationship between players' strategies which has been widely used to replicate social decision-making [2, 3]. Except for the advantage of interpretability, game theory is also suitable for dissimilar scenarios such as roundabouts [4], on-ramp merging areas [5, 6], and unsignalized intersections [7, 8]. Instead of treating surrounding vehicles as moving obstacles, the game theoretic approach was adapted to mimic human behavior by dynamically interacting with surrounding vehicles for automatic lane changing [9]. Level-k game was combined with vehicle controller [10] and Monte Carlo tree search (MCTS) [11] respectively to model the time-extended, multi-step, and interactive decision-making of CAV. Meanwhile, identifying real-time intentions through human cognitive structure plays an important role in developing human-centric intelligent vehicles. Therefore, Chu et al. [12] explored the mechanisms behind distracted driving intentions based on Stimulus-Organism-Response (SOR) cognitive theory. In addition, Artificial Neural Network (ANN) was designed by mimicking the mechanism by which information is transmitted in the animal neural network and it was used to predict drivers' intentions in on-ramp merging [13], lane changing [14, 15], and left-turning [16]. Bi et al. [17] proposed a queuing network cognitive architecture to infer intentions with and without distraction tasks and the experiments carried out on a driving simulator have shown a good performance in inferring typical and rapid lane-changing behavior intentions. Similar to the diverse and fickle intentions, drivers themselves also possess different characteristics. Huang et al. [18] adopted Inverse Reinforcement Learning (IRL) to model the different driving

This work was supported in part by the National Key R&D Program of China (2022YFB2502901), the State Key Laboratory of Intelligent Green Vehicle and Mobility under Project No. KFZ2408, the Natural Science Foundation of China (52302502 and 52125208), the Shanghai Scientific Innovation Foundation (23DZ1203400), the Belt and Road Cooperation Program under the 2023 Shanghai Action Plan for Science, Technology and Innovation (23210750500), and the Fundamental Research Funds for the Central Universities.

S. Fang, P. Hang and J. Sun are with the Department of Traffic Engineering and Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China. (e-mail: {2111219, hangpeng, sunjian}@tongji.edu.cn)

C. Wei is with the School of Mechanical and Aerospace Engineering at Queen's University Belfast, BT7 1NN Belfast, UK. (email: c.wei@qub.ac.uk)

Y. Xing is with the School of Aerospace, Transport, and Manufacturing, Cranfield University, UK, MK43 0AL. (e-mail: yang.x@cranfield.ac.uk)

Corresponding author: P. Hang

behaviors from naturalistic human driving data. Schwarting et al. [19] addressed Social Value Orientation (SVO), which quantifies the degree of an agent's selfishness or altruism, and then modeled the interactions between agents as a best-response game. Despite a number of works that have shown that the interaction between CAVs and HVs can be solved by a best-response game, a rational collective decision might be the opposite of the socially optimum [20]. Therefore, a solution to tackle this phenomenon is to establish an institutional arrangement that can optimize the system performance via cooperative driving.

The cooperative driving of CAVs has been widely studied. Cooperation refers to the decision-making process of vehicles aimed at orchestrating vehicles' actions so as to achieve a goal that can not be achieved by each vehicle in isolation [21]. Furthermore, CAVs cooperative approaches can be classified into centralized, negotiation, agreement, and emergent according to CAVs' extent of autonomy. In the centralized control, a coordinator will act as an intersection manager to reserve certain space-time for each approaching vehicle [22] or to decide all vehicles' strategies for at least one global task [23]. However, the capability of the coordinator is a potential bottleneck. Although Xu et al. [24] combined MCTS with heuristic rules to find a nearly global-optimal passing order for CAVs and results showed that it can keep a good trade-off between system performance and computation flexibility. Under centralized control, CAVs only passively follow the instruction of the coordinator so their potentials are far from being tapped. Negotiation and agreement approaches use fixed and dynamics protocols respectively to allow CAVs to communicate with others. Carlino et al. [25] proposed a decentralized auction-based autonomous intersection management scheme to permit vehicle passage based on their value of time. Vu et al. [26] further migrated auctions to a cellular automaton model. However, although the above methods are distributed during cooperation, CAV still adhere to a meticulously designed negotiation protocol, which may give rise to robustness issues in real-world scenarios. Instead of following any predefined protocol, emergent approaches refer to each vehicle making its decision by estimating the rationality of others and deducing their actions based on the current state. Wang et al. [27] established a cooperation model for agents with different priorities through game theory. Chen et al. [28] combined depth-first spanning tree algorithm and communication topology to realized a conflict-free cooperation of vehicles. Results showed that real-time computation can be achieved at intersections involving more than 30 vehicles. To sum up, the above researches mainly focused on CAVs-only environment cooperation. Nevertheless, real-world traffic flow will be mixed with HVs for at least 40 years according to [29], so the effectiveness under a mixed traffic environment has yet to be proven.

Aiming at either the interactions between CAV and HV or the cooperation methods in a CAVs-only environment merely leads to a half-baked CAV, their performance under real-world interactive scenarios should be further evaluated. Comparing with above research field, there has been little discussion about the verification and validation of CAV in mixed traffic. Wang

et al. [30] demonstrated his controller by adding background traffic flow but only four vehicles at most were involved in their two-lane merging scenario. Besides, human-in-the-loop trial on the driving simulator is also a practical way to replicate mixed traffic. Sadigh et al. [31] pioneering excavated the ability for CAV to actively gather the information of surrounding HV, and experiments were carried out on a driving simulator. The fly in the ointment is that experiments only consider one-on-one interaction. Furthermore, multiple human-machine interactions seem unpractical on the driving simulator because of the equipment limitations. Therefore, existing studies about mixed traffic are mainly low-density which is an extraordinary simplification. While real-world heterogeneous mixed traffic is far more hazardous because the environment contains both heterogeneous HVs and CAVs, and the heterogeneous preferences of drivers will drive them to make different decisions, posing greater challenges to CAV decision-making.

To address the aforementioned challenges, the cooperative driving of CAVs in heterogeneous mixed traffic conditions is studied. The contributions of this paper are presented as follows:

1. This paper proposes a novel cooperative driving framework that enables CAVs to cooperate in high-density, mixed, unsignalized intersections and coordinate with heterogeneous human drivers, in favor of driving conflict resolution, human-machine interaction augmentation and traffic efficiency improvement.
2. The cooperative decision-making of CAVs is realized by concatenating the Normalized Cooperative game and Level-k game (NCL game), which outputs the k-allocation that corresponds to maximum system overall efficiency. Then Lattice planner is combined to generate an optimal and collision-free CAV trajectory. The cooperation performance of CAVs under different Rates of Penetration (ROP) is verified.
3. To simulate the human-machine interaction in authentic mixed traffic, heterogeneous HV decisions are modeled. By clustering, driver classification and composition are obtained and serve as prior knowledge. Non-cooperative game is used to reproduce driver's rational decision under an interactive environment and IRL is utilized to calibrate different drivers' decision-making preferences. Finally, heterogeneous HVs are introduced to the simulation experiment to examine the effectiveness and robustness of the cooperative driving framework.

The rest of the paper is organized as follows. Section II introduces the main problem of our research and the corresponding cooperative driving framework to solve it. Reproducing heterogeneous HV decision is described in Section III. Section IV discusses the process of CAV cooperative decision-making and trajectory planning. In Section V, we validate our framework through several simulation experiments. Finally, conclusions are made in Section VI.

II. PROBLEM STATEMENT AND COOPERATIVE DRIVING FRAMEWORK

CAVs are still much-maligned due to their intricate maneuvers and deficient control methods. Deadlock and collision

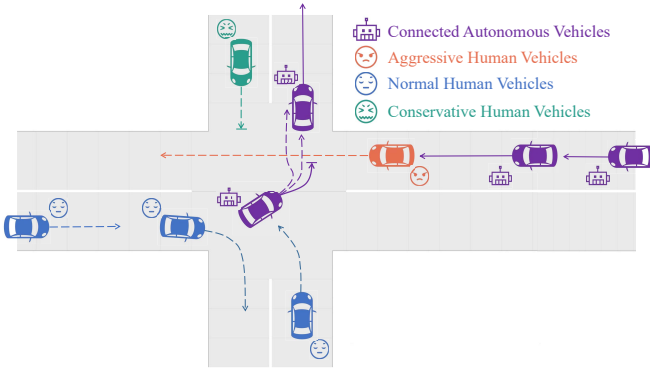


Fig. 1. Thumbnail of the high-density, unsignalized intersection with heterogeneous HVs involved.

may happen when CAVs interact with others in complex, interactive scenarios. Furthermore, human drivers are heterogeneous so they may adopt various maneuvers in the same circumstances based on their preferences. Therefore, CAVs always tend to passive defensive action such as excessive deceleration which leads to inefficiencies, or even worse, to the loss of trust in CAVs.

Among all urban traffic scenarios, unsignalized intersections are the most challenging owing to their numerous conflict points and fickle interaction opponents. Additionally, intersections in the United States witness over 2.8 million traffic accidents annually, accounting for 44% of all accidents, thereby rendering them the most perilous traffic scenario [32]. Unsignalized intersections pose a greater challenge to cooperative driving as the absence of signal control results in conflicts that cannot be separated by spatial and temporal domains.

Considering that investigating the interactions and cooperations of CAVs and HVs in sophisticated scenarios is of great significance to the validation and application of a cooperative driving algorithm for CAVs. Therefore, the main problem of this paper is to enable CAVs to properly respond to or even actively coordinate with heterogeneous HVs in a high-density, unsignalized intersection. Fig. 1 shows a thumbnail of the intersection.

To achieve the above purpose, an elaborate framework is proposed, shown in Fig. 2. Firstly, naturalistic human driving data are collected and analyzed to investigate the driver decision type and its distribution from a real-world intersection. Then a series of feasible trajectories are generated through the non-cooperative game and compared with expert demonstrations from naturalistic human driving data. Through maximum entropy IRL, different decision-making preferences θ^T are calibrated. We then reproduce the decision of heterogeneous HVs through Nash Equilibrium and generate the corresponding next state through vehicle dynamics.

To overcome the challenges of cooperative driving at mixed unsignalized intersections, a three-layer hierarchy is introduced for the cooperative decision-making and trajectory planning of CAVs. At the top of the hierarchy, the level-k game is introduced to imitate the human reasoning depths and also serves as the next layer's independent variable. Then, based on the cooperative game, k-allocation which leads to system

optimum can be achieved by enumerating. At the bottom layer, the Lattice planner will generate an optimal and collision-free trajectory that conforms to the vehicle's dynamic constraints.

The simulation will be updated every $\Delta t = 0.1s$ through vehicles dynamically competing and coordinating with each other. Finally, we compare the reproduced HV trajectories with actual decisions and trajectories at first. Then, we validate our framework with three case studies, including the comparative analysis with different methods, the case study under different Rates of Penetration (ROP), and the interaction analysis with heterogeneous HVs.

III. REPRODUCING HETEROGENEOUS HV DECISIONS IN MIXED TRAFFIC

Despite the uncertainty due to the absence of traffic lights, unpredictable and uncontrollable HV latent intention may lead to even greater strait for CAVs. While most studies adopt the Intelligent Driver Model (IDM) as a human decision model, it fails to account for the reciprocal effects of decision-making, thereby depriving HVs of their ability to actively interact [33, 34].

In this section, sophisticated real-world drivers' demonstrations are extracted for reproducing heterogeneous HV decisions. The main process consists of four steps. Firstly, interaction data from a real-world intersection are collected. Secondly, drivers' decisions are classified into three groups by the results of the K-means cluster and the elbow method. Then, a non-cooperative game is introduced to generate HVs' decisions. Finally, the decision-making preference of each group is calibrated through IRL.

A. Data Collection

As mentioned before, investigating the interactions of HVs is of great use in establishing human-like CAVs. In our paper, data acquisition was carried out at an intersection in Shanghai, China: XianXia Rd-JianHe Rd (XXJH) based on previous research of Ni et al. [35]. The data used for training only included interaction events where the Post Encroachment Time (PET) was less than 3 seconds to ensure the presence of an influence relationship between vehicles.

After screening, the trajectories of the HVs were extracted through a semi-automated video image processing tool and then served as prior knowledge. Furthermore, their speed and acceleration are calculated for clustering.

B. Driver Classification by K-means Cluster

In the complex real-world environment, drivers often make different decisions when faced with identical situations due to their diverse driving styles. Thus, it is essential to analyze the drivers' characteristics during interaction to identify typical driving patterns. Instead of directly training and calibrating all data through IRL, targeted training is used to fine-tune the preferences of heterogeneous drivers. Therefore, the priority is to determine how many decision clusters the drivers at the intersection of XXJH can be divided into.

To determine the optimal number of clusters for a dataset, the elbow method can be used. The elbow method is based

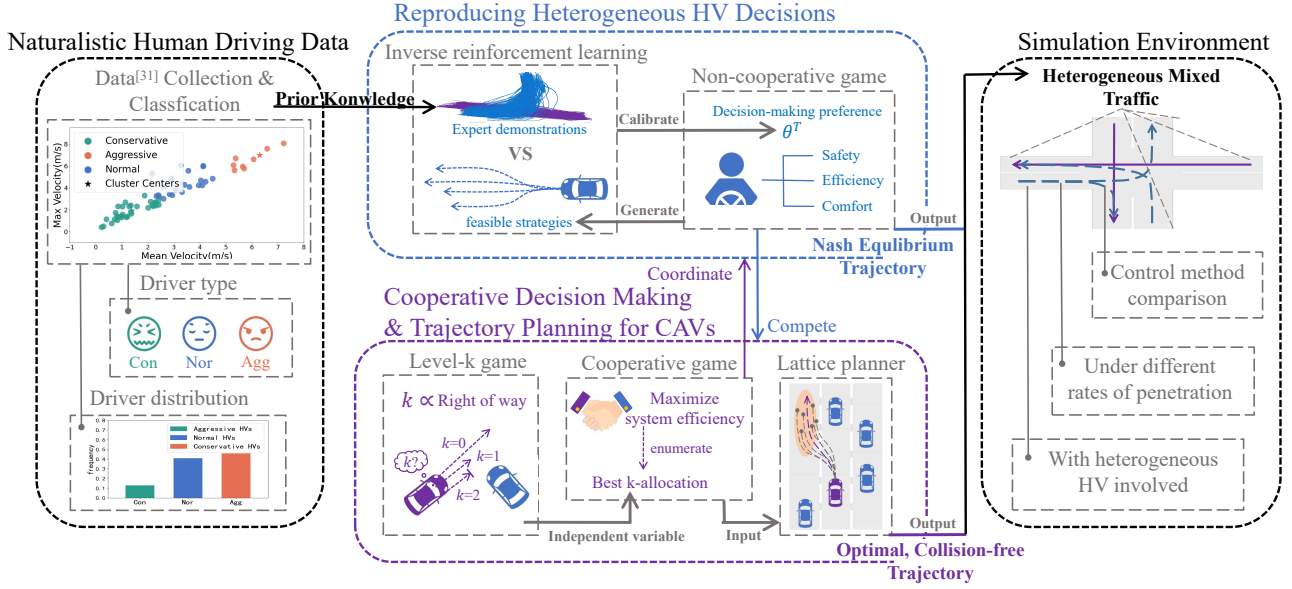


Fig. 2. Cooperative driving framework in mixed traffic.

on the principle that as the number of clusters increases, each cluster tends to show a higher degree of aggregation. This leads to a natural decrease in the Sum of Squared Errors (SSE). However, beyond a certain point, the rate at which SSE decreases slows down and eventually plateaus as the number of clusters increases further. This results in a graph that resembles an elbow shape, with the corresponding value at the elbow point representing the actual cluster number in the data. In our case, the result of the elbow method indicates that three clusters would be optimal for our dataset.

Considering that speed and acceleration are the most intuitive manifestation of decision-making during driving, the average, maximum, minimum, and standard deviation of speed and acceleration of vehicles are taken as the clustering parameters refer to Chen et al. [36]. K-means clustering is then used to classify naturalistic human driving data samples in XXJH.

TABLE I
CLUSTERING CENTERS OF XXJH DATA

Params	$V(m/s)$			
	mean	max	min	std
Cluster-1	3.31	4.42	2.36	0.66
Cluster-2	1.34	1.60	0.99	0.27
Cluster-3	6.29	6.98	5.40	0.50
Params	$A(m/s^2)$			
	mean	max	min	std
Cluster-1	-0.07	0.49	-0.72	0.37
Cluster-2	0.08	0.55	-0.34	0.28
Cluster-3	-0.77	-0.30	-1.14	0.28

After the clustering algorithm converges, the center of each cluster can be obtained, as shown in Table. I. Furthermore, their trajectories are used as expert demonstrations and then compared with HVs' feasible trajectories to calibrate different decision-making preferences. In this paper, a non-cooperative game is chosen to estimate possible decisions human may adopt when confronted with other drivers because it achieves more realistic human behavior when performing conflicting

maneuvers at intersections [37, 38]. Together with vehicle dynamics, feasible trajectories of each frame can be concluded.

C. HVs' Feasible Trajectories Generation through Non-cooperative Game

As mentioned before, game theory describes human as a rational decision-maker who takes action dependencies into account. Hence, there is a growing application in the field of modeling human decisions. In this paper, we regard human drivers as rational decision-makers and are aware of the consequence that their actions will influence other drivers who have conflict with them in temporal and spatial dimensions. The non-cooperative game is therefore established to mimic how drivers conjecture and compete with each other while driving in a shared space.

We consider a sampling interval, $\Delta t = 0.1s$, and an action set including six strategies that represent common driving maneuvers in urban traffic based on Tian et al. [39], listed in Table. II.

TABLE II
ACTION SET U

Action u	$A(m/s^2)$	$\omega(rad/s)$
Maintain u_1	0	0
Accelerate u_2	2	0
Decelerate u_3	-2	0
Brake u_4	-4	0
Turn left u_5	0	$\pi/4$
Turn right u_6	0	$-\pi/4$

where U is the action set, $u_t^i = [a_t^i, \omega_t^i]^T$ is the action of vehicle i at t frame, a_t^i is acceleration, and ω_t^i is heading angle rate. Therefore, six feasible trajectories can be generated in each frame by constituting action u_t^i with vehicle dynamics in Eq. 1.

$$f(s_t^i, u_t^i) = \begin{bmatrix} x_t^i + v_t^i \cos(\gamma_t^i) \Delta t \\ y_t^i + v_t^i \sin(\gamma_t^i) \Delta t \\ v_t^i + a_t^i \Delta t \\ \gamma_t^i + \omega_t^i \Delta t \end{bmatrix} \quad (1)$$

where f denotes the uni-cycle model, $s_t^i = [x_t^i, y_t^i, v_t^i, \gamma_t^i]^T$ denotes the state of a vehicle, (x_t^i, y_t^i) , v_t^i and γ_t^i are the position, speed, and yaw angle of vehicle i at t frame, respectively.

Furthermore, in a non-cooperative game, a discreet trade-off between the consequence of each action and the rival's possible corresponding response will be made in every $\Delta t = 0.1s$. Finally, after taking every action in Table II into account, the best response to cope with the rival's each action can be concluded. Once both players have found their best responses, Nash equilibrium is reached. In Nash equilibrium, none of the players can acquire a lower cost by unilaterally adjusting their actions. Therefore, the mathematical definition of Nash equilibrium is defined by

$$R(u_t^{i*}, u_t^{-i*}) \geq R(u_t^i, u_t^{-i*}) \quad (2)$$

where R is the reward function, u_t^{i*} and u_t^{-i*} are the action performed by player i and players other than player i when game achieves Nash equilibrium at t frame, u_t^i is arbitrary action of player i in action set U .

It's worth noting that Nash equilibrium isn't always the best solution in some cases [20]. However, we do not regard this as a drawback, but as a proxy for the limitation of human reasoning and the consequence of being absolutely rational. As in the real world, not all games have a happy ending.

In general, the decision-making process is often influenced by three key factors: efficiency, comfort, and safety. In this paper, efficiency is represented by distance of the destination, comfort is represented by jerk, and safety is represented by deviation from the expected path and Time to Collision (TTC). Destination refers to the last conflict point along the expected path. Then, the reward of each possible action can be calculated by s_t^i , u_t^i , and u_t^{-i} in Eq. 3.

$$R(s_t^i, u_t^i, u_t^{-i}) = [-d_t^i - j_t^i - o_t^i + \frac{1}{TTC_t^i}] \theta^T \quad (3)$$

where d_t^i is the distance between the contemporary position of vehicle i and its destination at t frame, j_t^i is the jerk which is inversely proportional to comfort, o_t^i is the offset to the expected path, TTC_t^i is the time to collision, and θ^T is preference weight which describes the driver's preference in efficiency, comfort, and safety. By incorporating different values of θ^T , we can effectively model drivers with heterogeneous preferences.

By combining Eq. 1 with Table II, six feasible trajectories can be generated in each frame. Max entropy IRL is then introduced to quantify the difference between feasible trajectories and expert demonstrations to calibrate HVs' different decision-making preferences.

D. Decision-making Preference Calibration by IRL

After dividing drivers into groups and generating feasible trajectories through a non-cooperative game, IRL is introduced to uncover inherent characteristics that influence expert decision-making. The concept of IRL emerged after behavior cloning, despite their sharing many similarities. Differing from simply imitating expert maneuvers, IRL tries to infer the reason why experts make their decisions and then optimize the strategy. In other words, except directly learning the state-action mapping, IRL infers the form of reward weight and optimizes maneuvers through it.

Because of the aforementioned advantages, IRL has been widespread in many fields. Among many IRL algorithms, the maximum entropy method stands out due to its ability to address ambiguity. As a human driver follows the stochastic policies that may lead to a different distribution of candidate decisions [18, 40], the principle of maximum entropy is introduced to choose the distribution that does not exhibit any additional preference. Therefore, ambiguity is resolved. Generally, the probability of a trajectory is proportional to the exponential of the reward in Eq. 4.

$$P(\zeta|\theta) = \frac{1}{Z(\theta)} e^{R(\zeta)} = \frac{1}{Z(\theta)} e^{\theta^T f_\zeta} \quad (4)$$

where $P(\zeta|\theta)$ is the probability of trajectory ζ , $R(\zeta)$ is the corresponding reward function of expert trajectory ζ which is the multiplication of reward weight vector θ^T and trajectory feature vector f_ζ , and $Z(\theta)$ is the partition function.

When confronted with a continuous or high-dimensional space problem, the partition function may fail to converge. However, in a finite horizon problem, the reward weights in maximizing entropy are certain to be convergent [41]. Therefore, limited feasible trajectories are generated and Eq. 4 could be rewritten as follows.

$$P(\zeta|\theta) = \frac{e^{\theta^T f_\zeta}}{\sum_{i=1}^M e^{\theta^T f_{\tilde{\zeta}_i}}} \quad (5)$$

where $f_{\tilde{\zeta}}$ is the feature of feasible trajectory $\tilde{\zeta}$, and M is the finite feasible trajectory. By this approximation, probability is much easier to calculate.

The kernel of IRL is to adjust the weights of the reward function to yield a maneuver that matches expert demonstrations. To fulfill that purpose, the likelihood of expert demonstrations should be maximized by adjusting θ^T .

$$\max_{\theta} L(\theta) = \max_{\theta} \sum_{\zeta \in D} \log P(\zeta|\theta) \quad (6)$$

where $L(\theta)$ is the likelihood function and also the objective function, $D = \{\zeta_i\}_{i=1}^N$ is the trajectory set containing N expert demonstrations which are the real-world human trajectories we screened and clustered before.

Substituting $P(\zeta|\theta)$ in Eq. 6 with Eq. 5, $L(\theta)$ can be rewritten as

$$L(\theta) = \sum_{\zeta \in D} [\theta^T f_\zeta + \log \sum_{i=1}^M e^{\theta^T f_{\tilde{\zeta}_i}}] \quad (7)$$

This function is convex and gradient-based optimization is used for solving optimal reward weight. The gradient can be written as the difference between the expert demonstrations f_ζ and the feasible trajectories $f_{\tilde{\zeta}}$ in Eq. 8.

$$\nabla_{\theta} L(\theta) = \sum_{\zeta \in D} [f_\zeta - \sum_{i=1}^M \frac{e^{\theta^T f_\zeta}}{\sum_{i=1}^M e^{\theta^T f_{\tilde{\zeta}}}}] \quad (8)$$

L2 regularization is introduced to the likelihood function to prevent overfitting and $\lambda > 0$ is the regularization parameter. Thus, Eq. 7 and Eq. 8 can be rewritten as

$$L(\theta) = \sum_{\zeta \in D} [\theta^T f_\zeta + \log \sum_{i=1}^M e^{\theta^T f_{\tilde{\zeta}}}] - \lambda \theta^2 \quad (9)$$

$$\nabla_{\theta} L(\theta) = \sum_{\zeta \in D} [f_\zeta - \sum_{i=1}^M \frac{e^{\theta^T f_\zeta}}{\sum_{i=1}^M e^{\theta^T f_{\tilde{\zeta}}}}] - 2\lambda \theta \quad (10)$$

Then, the pseudocode of maximum entropy IRL is summarized in Algorithm. 1. Based on the clustering results, three groups of expert demonstrations are respectively trained by IRL to calibrate the preferences of different drivers. Feasible trajectories and corresponding features can be obtained by utilizing a non-cooperative game and Eq. 3, respectively.

Algorithm 1: Maximum entropy inverse reinforcement learning.

Data: Expert demonstrations dataset D , learning rate α , regularization parameter λ , maximum training epoch E

Result: Optimized reward weights θ^* , including efficiency, comfort, and safety

```

1 Initialize  $\theta \leftarrow \mathcal{N}(0, 0.05)$ ;
2 Initialize buffer  $B \leftarrow []$ ;
3 Compute human features  $f \leftarrow \sum_{i=1}^M f_{\tilde{\zeta}}$ ;
4 while  $\zeta_i \in D$  do
5   Generate feasible trajectories  $f_{\tilde{\zeta}}$ ;
6   Calculate features of  $\tilde{\zeta}$ ;
7   Add features to buffer  $B \leftarrow B \cup \tilde{\zeta}$ ;
8    $i \leftarrow i + 1$ ;
9 end
10 while  $epoch < E$  do
11   Calculate the likelihood function of each sample in buffer  $B$ ;
12    $L(\theta) = \sum_{\zeta \in D} [\theta^T f_\zeta + \log \sum_{i=1}^M e^{\theta^T f_{\tilde{\zeta}}}]$ ;
13   Calculate the gradient  $\nabla_{\theta} L(\theta)$ ;
14   Update reward weights  $\theta \leftarrow \theta + \alpha \nabla_{\theta} L(\theta)$ ;
15 end
16 Get optimized reward weights  $\theta^* \leftarrow \theta$  after epoch reached  $E$ ;
```

After training, the reward weights of different driver groups can be calibrated through maximizing entropy and iterating, shown in Table III.

TABLE III
IRL CALIBRATION RESULTS

Type	Reward weights		
	efficiency	comfort	safety
Cluster3(Aggressive HV)	8.33	1.56	3.69
Cluster1(Normal HV)	8.2	1.72	5.7
Cluster2(Conservative HV)	7.79	2.1	8.44

According to the results, drivers in Cluster 3 exhibit the highest weight for efficiency and the lowest weight for safety and comfort, indicating a preference for high-speed overtaking rather than stopping and waiting. Hence, these drivers are classified as aggressive type. On the other hand, Cluster 1 drivers exhibit intermediate values across these three features without any distinct preferences which are classified as normal type. Lastly, Cluster 2 shows a predominant inclination towards safety, while displaying a relatively lower preference towards efficiency compared to the other two clusters. As a result, these drivers tend to be more cautious when making decisions and are classified as conservative type.

In conclusion, the above results are in line with real-world expectations. After calibrating different drivers' preferences, the action corresponding to Nash equilibrium can be inferred and the next state of the vehicle can be calculated by substituting Nash equilibrium action into Eq. 1.

So far, We have modeled heterogeneous HVs to establish a mixed environment. Then, a cooperative driving framework that enables CAVs to interact and coordinate with other vehicles is designed, thereby advancing traffic efficiency and safety.

IV. COOPERATIVE DECISION-MAKING AND TRAJECTORY PLANNING FOR CAVS

Unlike the unpredictable and uncontrollable HVs' actions, CAVs are capable of cooperating and even coordinating with others to nudge the system to greater efficiency and safety. As shown in Fig. 2, the cooperative driving algorithm for CAVs is introduced in three subsections. Level-k game and cooperative game are introduced in subsection A and subsection B to search for an optimal k-allocation solution that refers to the highest system efficiency. In subsection C, the Lattice planner is used for planning an optimal and collision-free trajectory based on k-allocation results.

A. Resolving Multiple Traffic Streams Conflict by Level-k Game

Level-k game assumes that players generate strategies based on their depths of reasoning. In addition, k can also be defined as ROW [42], driving style [39, 43]. Here, considering the absence of traffic lights will cause multiple traffic streams to intersect in time and space, we denote k as ROW to resolve conflicts.

In the case of trajectory planning, level-0 drivers regard other vehicles as static obstacles and plan their trajectory on this basic assumption. Then level-1 drivers will infer the planning trajectory of level-0 drivers and then plan their trajectories while avoiding collision with level-0 trajectory. By

that analogy, all drivers may obtain the expected paths based on their depth of reasoning. Therefore, the expected path of vehicle i , denoted as ζ_i , can be calculated from

$$\zeta_i(k) = \max R(k, s_u^i, \zeta_j(k-1)) \quad (11)$$

where $\zeta_i(k)$ is the trajectory of vehicle i with k level, and the trajectory is determined by maximizing the reward function R .

In this paper, our CAV is free to generate its depths of reasoning to maximize the system's overall efficiency. When multiple traffic streams intersect, a unique k will be assigned to each stream to prevent collision or deadlock. Researchers also pointed out that normally reasoning depths of human are less than or equal to 2 according to [43, 44]. Therefore, to prevent CAVs from behaving unreasonably, traffic streams with k greater than 2 are corrected to k equal to 2 based on the limited human reasoning depths.

Level- k game provides a theoretical basis for the orderly passage of vehicles. Though unsignalized intersections fail to resolve conflict from the temporal dimension, countable streams lead to finite k -allocation solutions. By optimizing the k -allocation of streams, system optimum can be achieved and the conflict can be resolved. Therefore, a cooperative game is utilized to quantify the system performance of each k -allocation solution.

B. Searching Best k -allocation through Cooperative Game

Reservation-based control methods such as First Come First Serve (FCFS) and the Batch-strategy have a long history and are widely used due to their succinct form [22]. FCFS control simply assumes that ROW is proportionate to the entry order and firstly serves as a reservation mechanism for autonomous vehicles ascend to Dresner and Stone [45]. The Batch-strategy believes that intersections will be more efficient from a capacity of view if served in batches. Therefore, vehicles in each stream are grouped into batches and served at FCFS principle. While these methods do not consider system efficiency and rely solely on the order of entry into the ROW, they result in underutilization of the cooperation capabilities of CAVs.

To give full play to the cooperative capabilities of CAVs, the cooperative game is combined with the aforementioned level- k game to optimize system efficiency in this paper. Differing from the non-cooperative game that regards other participants as rivals and emphasizes maximizing individual profit. The cooperative game refers to participants uniting as several coalitions and trying to maximize system goals through collaborating with other coalitions. However, the cooperative game must meet the requirements of superadditivity, individual rationality, and group rationality [46].

Superadditivity refers to the inequality of the reward of coalitions and the total reward of separate individual vehicles. It can be expressed by

$$R(V_i) + R(V_j) \leq R(V_i \cup V_j) \quad (12)$$

where $R(V_i)$ is the reward function of vehicle i , and $V_i \cup V_j$ stands for the coalitions consists of vehicle i and vehicle j .

In addition, individual rationality should satisfy the following inequality.

$$R(V_i) \leq R_i(V_i \cup V_j \dots \cup V_n) \quad (13)$$

where $R_i(V_i \cup V_j \dots \cup V_n)$ represents the individual reward of vehicle i under cooperative game that n vehicles are united as a coalition. This inequality stipulates that each vehicle could obtain a higher reward when controlled with a cooperative game than before. Finally, group rationality means that the system goals are the summation of individual rewards, i.e.,

$$O = \sum_{i \in N} R(V_i) \quad (14)$$

where O is the system goal of a cooperative game and the sum of individual reward at the same time. System goals are often in the form of minimum overall delay or maximum efficiency such as moving faster or further.

Through experiments, we finally choose the inverse distance to the destination to represent system efficiency as it has shown a better performance than other surrogate variables. Therefore, the objective function O is as follows.

$$O(k) = \max \sum_{i \in \text{stream}} \sum_{j \in \text{group}} \frac{1}{\bar{d}_{i,j}} \quad (15)$$

where $O(k)$ is the value of the objective function under a certain k -allocation solution, $\bar{d}_{i,j}$ is the average distance to the destination of the expected path belonging to vehicles in i stream and j group. The destination refers to the last conflict point along the expected path. Once a vehicle surpasses this point, it is considered unaffected by other vehicles. Each k -allocation solution will lead to a different $\bar{d}_{i,j}$.

It should be noted that this distance is not the current distance to the destination but the average distance to the destination of the planning trajectory under a certain k -level. The value of $\bar{d}_{i,j}$ can be achieved by giving a specific k to the Lattice planner, which will be discussed in the next subsection.

Virtual lane rotation is a commonly used method for transferring vehicles from different traffic streams into the same dimension [47, 48]. In this paper, when calculating $\bar{d}_{i,j}$, all vehicles are also converted to the virtual lane to improve the calculation efficiency. However, previous studies often overlook the distinctions between turning vehicles and going straight vehicles, and assume that their impacts on the system are identical. Simultaneously, we are aware that vehicles going straight and those making turns may have different impacts on the system. For example, in the relevant traffic rules, it is stipulated that the left-turning vehicle should yield to oncoming straight traffic at the intersection. Moreover, in real-world scenarios, left-turning drivers may occasionally not yield to oncoming straight traffic when they are closer to the point of potential collision or under some other circumstances. Henceforth, we aim to investigate the initial conditions under which a left-turning vehicle exhibits a greater inclination towards yielding versus choosing to proceed first. Furthermore, this difference will consequently influence their right of way which corresponds to the k value in our study.

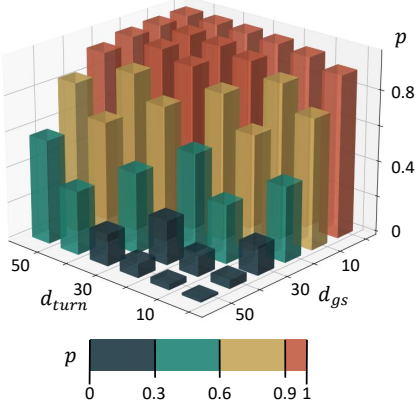


Fig. 3. The probability of going straight vehicle pass first under different initial state.

We introduce a relative weight α to represent the different effects of direct driving and turning vehicles on the system. Therefore, coefficients of distance to the destination should be various and Eq. 15 can be rewritten as

$$\max(\sum_{j \in \text{group}} \frac{1}{\bar{d}_{gs,j}} + \sum_{j' \in \text{group}'} \frac{\alpha}{\bar{d}_{turn,j'}}) \quad (16)$$

where α represents the weight of turning vehicles relative to going straight vehicles, d_{gs} and d_{turn} are the distance to destination of going straight vehicle and turning vehicle, respectively. Then, an investigation is carried out to unify the dimension of $\bar{d}_{gs}$ and $\bar{d}_{turn}$.

To investigate the real-world relationship between going straight vehicles and turning vehicles, the interactions between going straight vehicles and turning vehicles are further extracted based on the XXJH data. Next, a logistic model is introduced to explore the relation between the initial distance to the destination and the actual order of passage. After calibrating by interaction data, results are presented as follows.

$$\ln \frac{p}{1-p} = 0.0925d_{turn} - 0.1332d_{gs} + 2.35 \quad (17)$$

where p is the probability of going straight vehicle pass first. The relationship between p and d_{turn} , d_{gs} can be shown in Fig. 3.

Substituting $p = 0.5$ into Eq. 17, there is

$$d_{turn} = 1.44d_{gs} - 25.4 \quad (18)$$

According to the coefficients of d_{turn} and d_{gs} , a turning vehicle will cost more to maintain equality with a going straight vehicle. This conclusion is consistent with relevant traffic rules which stipulate that turning vehicles should give precedence to going straight vehicles when in conflict at an intersection. In addition, by substituting Eq. 18 into Eq. 16, the objective function is derived as follows.

$$\max(\sum_{j \in \text{group}} \frac{1}{\bar{d}_{gs,j}} + \sum_{j' \in \text{group}'} \frac{1}{1.44\bar{d}_{gs,j'} - 25.4}) \quad (19)$$

Through the logistic model, the impacts of going straight vehicles and turning vehicles to the system are normalized.

Instead of directly stipulating that turning vehicles must give priority to straight traffic, social norms that going straight vehicles possess higher priority than turning vehicles at the intersection are taken into consideration subtly during this normalization. Here, we named it a normalized cooperative game.

As indicated above, in a cooperative game, vehicles will unite as several coalitions. Maximizing the system's overall efficiency is the same goal that all coalitions share. Since the number of traffic stream at an intersection is limited, there are also only a few possible k-allocation solutions. By combining the normalized cooperative game with the level-k game, the k-allocation solution that corresponds to optimal system efficiency $O(k^*)$ can be easily obtained by enumerating finite k-allocation solutions.

$$k^* = \arg \max O(\mathbf{K}) \quad (20)$$

where $\mathbf{K}$ is all k-allocation possibilities vector which equals to A_{stream}^2 . Finally, the Lattice planner is introduced to generate the optimal collision-free trajectories based on the concept of the level-k game to evaluate each k-allocation solution's performance.

C. Planning Optimal Trajectory with Lattice Planner

Lattice planner is a well-accepted sample-based planner to generate optimal and collision-free trajectories according to expected path and obstacles. It mainly consists of coordinates transforming, sampling, and curvature polynomials fitting.

During the coordinates transforming step, the corresponding state in Frenet coordinates is transformed according to the initial state in Cartesian coordinates. Then, planning horizons and maximum steering angles are sampled. For improving computation speed, the initial sampling space is relatively small. When it comes to no available solution, sampling space will be enlarged. Simulation will be shut down and print out no solution if there are still no available solutions. When it comes to multiple solutions, the optimal trajectory is determined by reward. Finally, after sampling, curvature polynomials are used to generate the entire trajectory between the start point and end point.

After obtaining the entire trajectory, the average distance to the destination is calculated. Furthermore, $\bar{d}_{gs,j}$ and $\bar{d}_{turn,j'}$ in Eq. 16 is obtained. Therefore, the objective function value can be computed for each k-allocation solution, and subsequently, the optimal k-allocation is determined by selecting the solution with the highest value.

During each calculation, CAV assumes that HV will adhere to the optimal system efficiency k-allocation, referred to as the estimation of HV's k value. This estimation will further generate a prediction of HV's action. However, considering the uncontrolled nature of HVs, HV actual action may deviate from the predicted action. Different from accommodating the behavior of HV regardless, designed CAV prioritizes the implementation of optimal system schemes, aiming to guide HV towards behaviors that correspond to optimal system efficiency k-allocation during continuous interaction. Only in the rare event that HVs are completely uncooperative and

a collision is imminent, the CAV will remove k^* from the feasible k-allocation $\mathbf{K}$, while searching for a new k-allocation until collision avoidance is achieved. Algorithm. 2 describes the procedure of CAVs' cooperative decision-making and trajectory planning.

Algorithm 2: Cooperative decision-making and trajectory planning for CAVs.

Input: Vehicle's state s_t^i , stream number *stream*, vehicle number in each stream *group*

Output: Vehicle decision-making depths k and next state s_{t+1}^i according to planning trajectory

```

1 Calculate possible k-allocation  $\mathbf{K} \leftarrow A_{stream}^2$ ;
2 Initialize objective function buffer  $O(\mathbf{K}) \leftarrow []$ ;
3 foreach  $k \in \mathbf{K}$  do
4   Evaluate each  $k$  performance  $O(k)$ ;
5   foreach  $i \in stream$  do
6     foreach  $j \in group$  do
7       Calculate average distance to destination  $\bar{d}_j$ 
         through  $Lattice(k)$ ;
8       Normalize  $\bar{d}_{gs,j} \leftarrow \bar{d}_j$ ;
9       Calculate objective function  $O(k)$ ;
10    end
11  end
12  Add  $O(k)$  to buffer  $O(\mathbf{K}) \leftarrow O(k)$ 
13 end
14 Find the best k-allocation solution
     $k^* = \arg \max O(\mathbf{K})$ ;
15 if collision then
16   Remove  $O(k^*)$  from  $O(\mathbf{K})$ ;
17   Return to line 14;
18 end
19 Get each vehicle  $k$  and plan its next state  $s_{t+1}^i$  through
     $Lattice(k^*)$  and k-allocation solution;
```

Steps in Algorithm 2 show the whole process of CAV cooperative decision-making and trajectory planning. At each step, we use the level-k model to simplify the possible passing scenarios and search for the optimal k-allocation solution. Further, in combination with the lattice planner, the collision-free trajectory under this optimal scheme is generated. It will repeat every $\Delta t = 0.1s$ until the simulation duration is reached. After finishing the modeling of HVs and CAVs, several experiments are conducted to validate our cooperative driving framework.

V. SIMULATION VALIDATION

A. Simulation Design

To evaluate the performance of the proposed algorithm framework, an isolated four-approach unsignalized intersection in Fig. 4 is used as the simulation environment. All approaches are 40 meters long and each approach contains one or two traffic streams.

Although we simplify the number of traffic streams, nearly ten conflicts still exist according to the topological relationship. Except for crossing conflicts, the simulation environment also includes converging conflicts and diverging

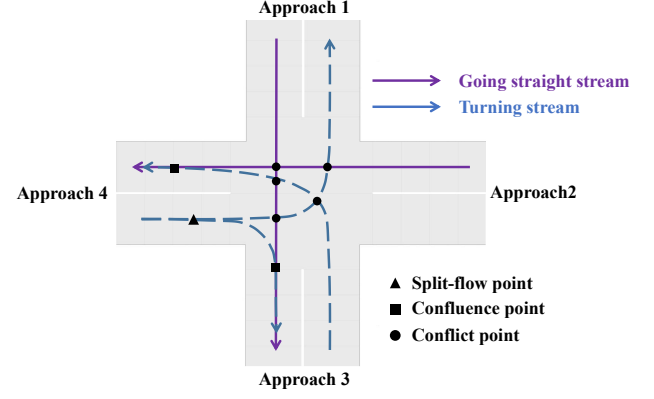


Fig. 4. Simulation environment.

conflicts which further examine the proposed algorithm framework under various circumstances [49]. Together with high-density traffic flow and heterogeneous HVs, we posit that this unsignalized intersection presents sufficient complexity to evaluate the proposed algorithm framework.

Furthermore, driver preferences not only influence their decision-making but also impact their initial and target speeds to produce a more realistic environment. Therefore, based on our prior knowledge derived from XXJH data, the initial speed of HV is set to match the average speed, while the target speed is aligned with the maximum speed. For instance, as shown in Table I, an aggressive driver would enter the simulation environment at a velocity of 6.3m/s and aim to maintain a speed of 7m/s if feasible. To evaluate the efficacy and robustness of our framework, four simulations were conducted in this study.

We started by testing the integration of the non-cooperative game and IRL, and found that it effectively reproduced the trajectory of heterogeneous HVs through the replay test method. Then, we compared NCL game and CL game (NCL game without normalization) with reservation-based control methods under different traffic densities to evaluate the method we proposed. FCFS and the Batch-strategy were used as the representative of the reservation-based method based on [22]. Next, in order to verify the effectiveness of the proposed cooperative driving algorithm, a simulation was conducted under different CAV rates of penetration. Finally, we further introduced drivers with different personal preferences to our simulation environment based on IRL calibration results in Table III.

B. Comparison between reproduced HV decision and actual HV decision

To verify whether the combination of the non-cooperative game and IRL can effectively reproduce the behavior of the human driver, we adopt the replay test method. Specifically, the subject vehicle utilizes a game theory framework consisting of six discrete strategies in Table. II for decision-making, while the interactive vehicles replays extracted trajectories accordingly.

Firstly, we conducted a comparison to assess the consistency between the reproduced sequence of passages and the actual

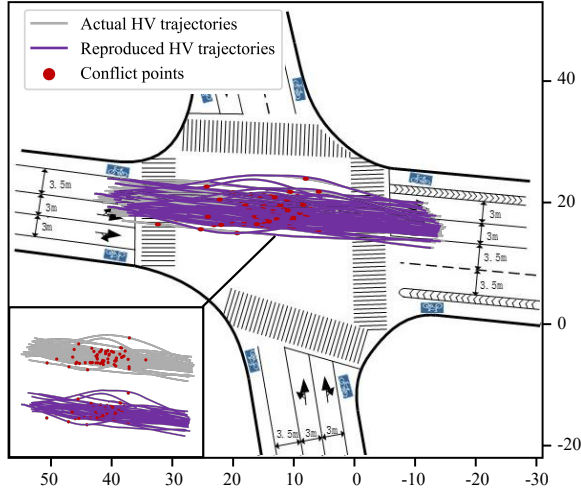


Fig. 5. Comparison between reproduced HV trajectories and actual HV trajectories.

sequence. The results reveal that 89% of cases reproduced the same sequence of passages as observed in reality. In contrast to employing IDM as the HV decision model, it exhibits 45% improvement. The primary reason for IDM's poor performance is that IDM is commonly employed as a car-following model, depicting the relationship between the following vehicle and the leading vehicle. Therefore, it is difficult to describe the interdependence of actions during the interaction process. The above results demonstrate that the proposed method outperforms IDM in reproducing HV decisions, thereby facilitating the establishment of a more realistic simulation environment for examining our CAV algorithm.

Secondly, we visualize the reproduced HV trajectories and the actual trajectories, as shown in Fig. 5. The result shows that the reproduced trajectories highly coincide with the actual trajectories. To further compare the similarity of the trajectories between reproduced HV trajectories and the actual trajectories, we employ Dynamic Time Warping (DTW) which minimizes the effects of shifting and distortion in time through a flexible transformation/mapping of time [50, 51]. However, given that this paper primarily focuses on the modeling and verification of the proposed CAV algorithm, a comprehensive analysis of trajectory similarity is available at: <https://fangshiyuu.github.io/NCLGame-IRL/>

Next, three case studies were carried out to evaluate the efficacy and robustness of the proposed cooperative driving framework. All simulations last 2 minutes, which is 1200 frames. After the simulation, the trajectories of all vehicles were recorded for further evaluation.

C. Case1: Comparison of Different Control Methods

For the purpose of evaluating the results of different control methods. Simulations were conducted in a full CAV environment at first. Average travel speed is applied to evaluate system efficiency.

Fig. 6 shows the average travel speed under different lane volumes. Clearly, it presents a negative relationship with lane volume, which is consistent with basic cognizance about the relationship between speed and volume according to

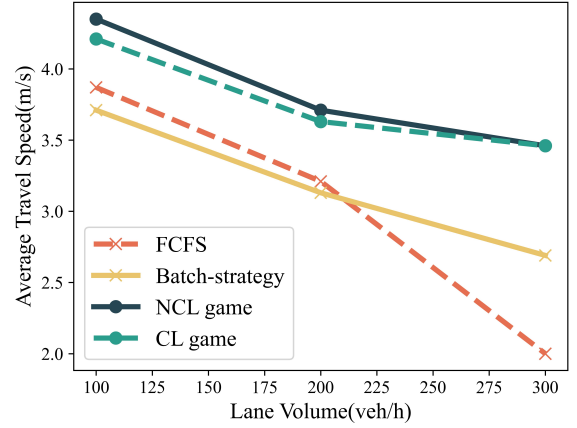


Fig. 6. Average travel speed comparison of control method under different lane volume.

Greenshields et al. [52]. In addition, travel speed decreases dramatically under FCFS control when lane volume increases, which is caused by the overly rigid rules of FCFS. Besides, when lane volume is relatively low (under 200veh/h), FCFS control even outperforms the Batch-strategy because grouping vehicles into batches in low density may add unnecessary steps that lead to low efficiency.

As seen in Fig. 6, the CL game and NCL game exhibit higher travel speeds compared to the reservation-based control method across all lane volumes, thereby demonstrating the effectiveness of combining cooperative games and level-k games from a system perspective. Furthermore, as lane volume increases, the disparity between the CL game and NCL game diminishes. This can be attributed to the number of vehicles in conflicting streams becoming more significant than the types of streams during simulation (turning stream or going straight stream). Consequently, with an increase in lane volume, the normalization effect becomes less pronounced.

Furthermore, Fig. 7 illustrates a scenario in which eight vehicles coexist in the simulation environment simultaneously under the Batch-strategy control and NCL game. Fig. 7 (a) depicts their initial position and the numbers beside vehicles stand for their entry sequence. Fig. 7 (b)-(c) demonstrates the evolution of vehicle positions over time. Owing to the Batch-strategy serving vehicles as batches, V_7 (stands for vehicle 7 in the figure) obtained the same ROW as V_3 and passed in a queue successfully. However, the Batch-strategy still follows the principle of FCFS which led to V_1 passing first even if there is only one vehicle in its stream. This phenomenon causes other streams conflicting with V_1 to yield regardless of the number of vehicles within those groups, leading to inefficiency.

A pattern is noticeable when the NCL game is put into action. Vehicles in Approach 2 have the highest right-of-way (ROW), followed by vehicles in Approach 3. The evaluation of all possible k-allocation solutions and selection of the most efficient one leads to this result. Fig. 7 (d) shows the average travel speed of the whole system at different simulation times. The proposed NCL game exhibits superior efficiency compared to the Batch-strategy, with a nearly 50% reduction in total intersection delay.

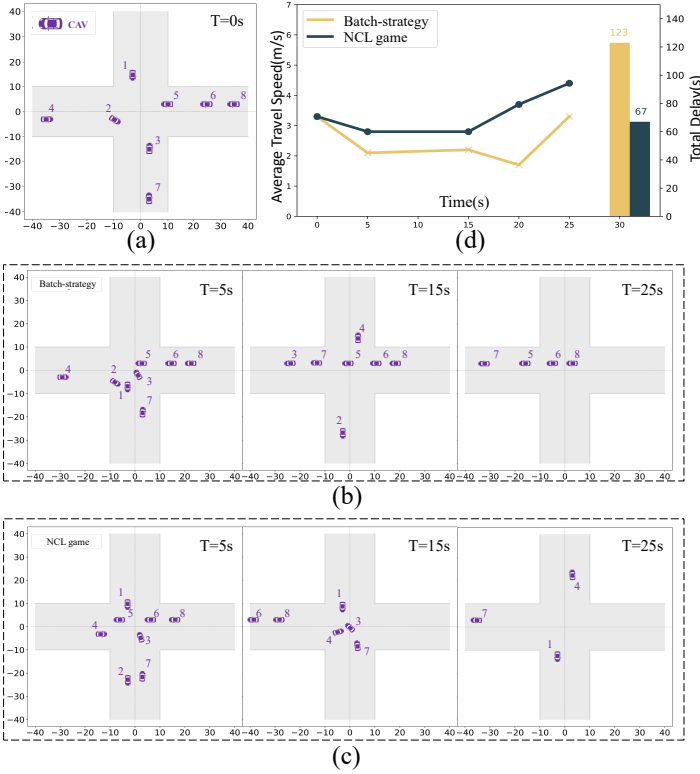


Fig. 7. Interaction case: comparison of NCL game and the Batch-strategy: (a) shows the initial position of the vehicles; (b)-(c) show three subsequent steps of CAVs control with the Batch-strategy and NCL game, respectively. (d) summarizes the average travel speed and total delay of this interaction case; numbers next to the vehicles refer to the order in which they enter the simulation environment.

Therefore, considering that continuous traffic and human drivers pose further challenges to CAV, we maintain the lane volume at 300 veh/h in the subsequent experiments to assess the applicability of our work in continuously mixed scenarios.

D. Case2: Experiment under Different ROP

Given the long-term expectation for CAVs to operate alongside HVs in traffic, it is imperative to thoroughly investigate their interactions under various circumstances. Consequently, we further assessed intersection performance using the NCL game and the Batch-strategy under different ROP. Each control method was studied under three ROP levels: 100%, 60%, and 20%.

The average travel speed and total delay are calculated and standardized in Fig. 8. A clear trend is observed as the ROP increases, with the average travel speed increasing and the total delay decreasing, indicating an improvement in system efficiency. Furthermore, when CAVs are controlled with NCL game approach, the average travel speed across is higher under all ROP. Besides, the minimum average travel speed when using NCL game control is even higher than the maximum travel speed using Batch-strategy control. Similar conclusions can be drawn for total delay, providing that the proposed NCL game method outperforms traditional reservation-based methods and is suitable for mixed traffic flows containing both CAVs and HVs.

The above analysis aims to evaluate the efficiency of intersections. Safety should also be noted due to its reliance on

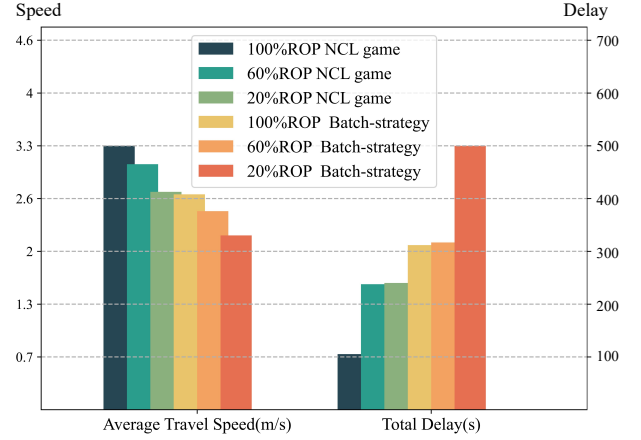


Fig. 8. Average travel speed and delay under different ROP.

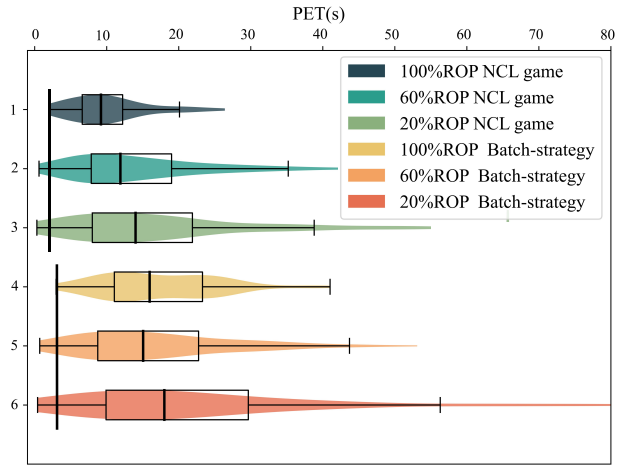


Fig. 9. Distribution of PET under different ROP.

public trust in CAVs, which currently poses a significant obstacle to the popularization of CAVs. Therefore, we introduce Post Encroachment Time (PET) to assess model performance from a safety perspective. Initially, PET was introduced by Allen et al. [53] as the time between the moment when the front vehicle leaves the conflict area or invasion line and the rear vehicle reaches the conflict area or invasion line. The distribution of PET under different ROP and control methods is illustrated in Fig. 9.

The box plots above show that the distribution of PET becomes more discrete when ROP decreases, suggesting that interactions between vehicles are more chaotic and inefficient. Simultaneously, the decline in ROP also results in higher maximum PET and average PET, indicating a longer interval between vehicle passages based on the concept of PET. In other words, as the ROP decreases, interactions become more conservative, making deadlocks more likely to occur.

Another line of evidence came from minimum PET, which indicates the most dangerous interaction case during the whole simulation duration. Fig. 9 shows a positive correlation between ROP and minimum PET, proving that increasing the proportion of our CAV helps to improve system safety. The rationale behind this lies in the fact that CAV decision-making not only aligns with the optimal solution for system efficiency

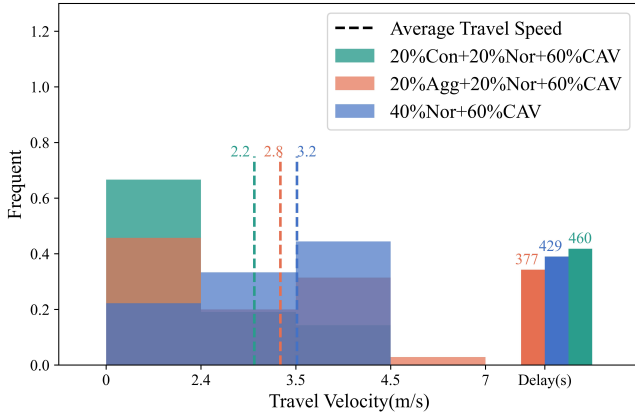


Fig. 10. Average travel speed and total delay under different vehicle compositions.

but also dynamically estimates and adjusts the k value of HV during decision-making, thereby improving system safety.

Through the analysis of average travel speed, total delay, and PET, we have proven that the CAV we designed has the ability to nudge the intersection system toward greater efficiency and safety. However, in addition to high-density and mixed traffic, human drivers possess different decision-making preferences will lead to heterogeneous traffic. To further evaluate the proposed cooperative driving framework, heterogeneous HVs are introduced to reproduce a more realistic driving environment in the next subsection.

E. Case3: Heterogeneous Human Driver Involved

We conducted three experiments at 60% ROP at first. Instead of generating all HVs make normal decisions while interacting with other vehicles like in Case 2. Half of HVs were replaced by aggressive type or conservative type, which may produce more unpredictable behavior. Total delay and the distribution of average travel speed are shown in Fig. 10.

Fig. 10 shows that when there is only normal type HV mixed with CAVs in a simulation environment, the average travel speed is significantly higher than aggressive HV or conservative HV involved. In addition, when 20% of normal HV was replaced by aggressive HV, though aggressive HV has a higher initial speed and target speed as we appointed in subsection A, the proportion of low-speed (velocity $< 2.4m/s$) vehicles increases on the contrary. Not to mention that the proportion of low-speed vehicles will increase sharply from 22% to 67% if half of the normal vehicles are replaced by conservative vehicles. Though little difference is found in terms of total delay between the three experiments, the above experiments emphasize the difficulty when interacting with heterogeneous HVs whose decisions are fickle.

Therefore, further experiments should be carried out to examine if the proposed NCL game can coordinate with heterogeneous HVs. The real-world environment is established according to the HV composition in XXJH data. Finally, the experimental group consists of 13% aggressive HV, 41% normal HV, and 46% conservative HV. In the control group, normal HVs are replaced by CAVs. To control variables, the composition of rewards and reward weights for CAV is designed to be consistent with normal HV. The difference is

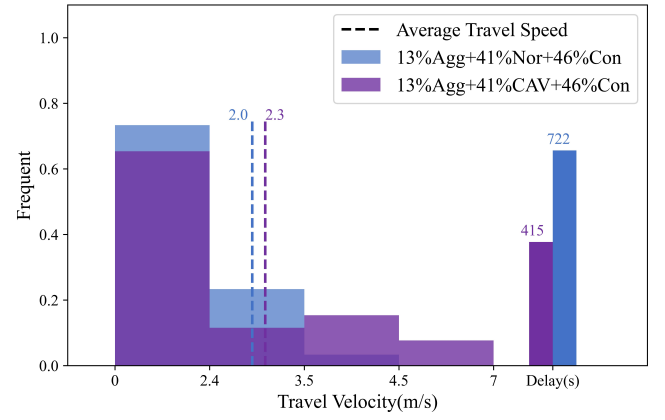


Fig. 11. Average travel speed and total delay based on XXJH vehicle composition.

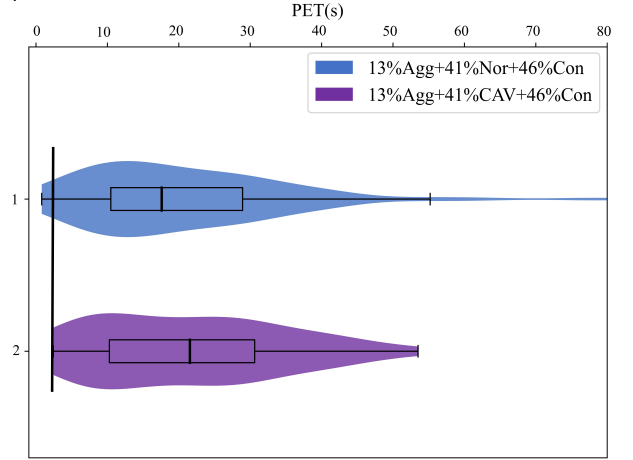


Fig. 12. Distribution of PET based on XXJH vehicle composition.

that HVs made decisions by searching Nash equilibrium in a non-cooperative game and then generated trajectory through vehicle dynamics while CAVs replaced these steps with the NCL game we proposed and the Lattice planner respectively.

According to Fig. 11, after replacing normal HV with our CAV, the performance of the intersection improves in all aspects. Specifically, the average travel speed of the intersection system increases by 0.3m/s, and the proportion of high-travel-speed vehicles increases. The more striking change is a 42.5% reduction in total delay which confirmed the robustness and the superiority of our CAV algorithm framework under a heterogeneous mixed environment.

Similar to Case 2, we calculate the PET of the interactions during the simulation duration. As depicted in Fig. 12, after replacing normal HV with CAV, the minimum PET increases, indicating that our CAV possesses the capability to enhance driving safety while improving system efficiency. Furthermore, we conducted a simulation case to visually illustrate the interaction and cooperation between heterogeneous HVs and CAVs.

By comparing Fig. 7 (d) and Fig. 13 (d), it is evident that the inclusion of heterogeneous HVs results in a decrease in average travel speed and increase in total delay, posing greater challenges to our cooperative driving algorithm framework. The detailed simulation process is illustrated in Fig. 13 (b)-

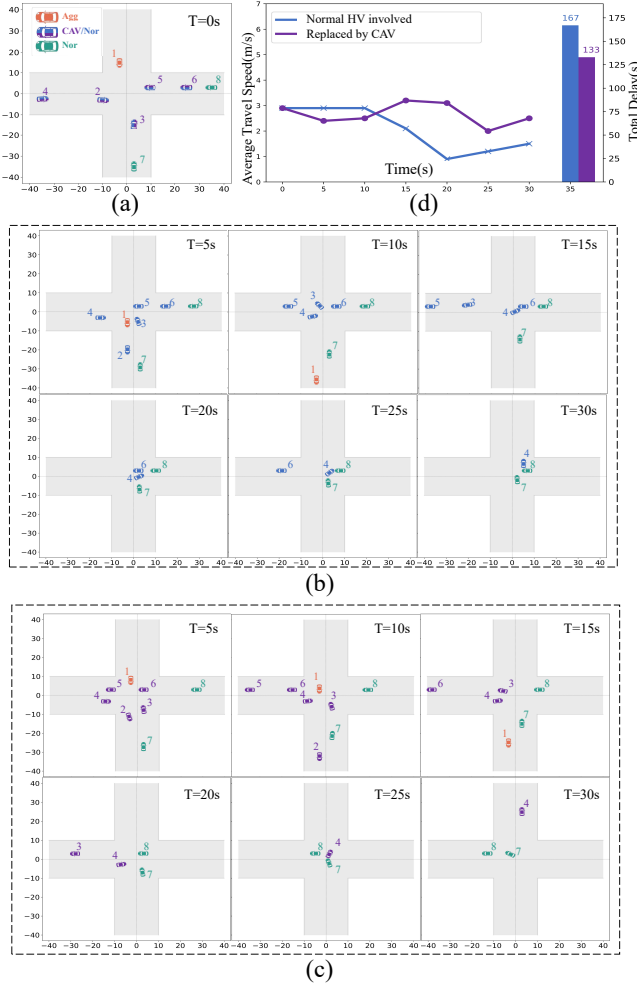


Fig. 13. Interaction case: comparison of normal HV and CAV involved. (a) shows the initial position of the vehicles and the colors of vehicles depict their type; (b) shows the six subsequent steps of the interaction that only HVs participants; (c) shows the corresponding steps after replacing normal HVs with CAVs; (d) summarizes the average travel speed and total delay of this interaction case; numbers next to the vehicles refer to the order in which they enter the simulation environment.

(c). In an environment with only HVs, aggressive vehicle V_1 prioritizes its own interests and strives to pass through the intersection as quickly as possible. Initially, this behavior leads to a higher average travel speed for the system; however, once the aggressive vehicle exits the intersection due to its selfish maneuvering that may endanger other vehicles' interests over long distances, the average travel speed plummets.

After replacing normal HVs with CAVs, the CAVs actively induce the aggressive vehicle to decelerate by demonstrating a willingness to pass first (by comparing the status of Fig. 13 (b)-(c) at $T = 5s$ and $T = 10s$). This can be attributed to our cooperative method, where their cooperative behaviors emerge rather than being predetermined by a predefined protocol, thus offering potential applicability in more complex scenarios. Compared with the reckless yield to aggressive vehicles, the proposed CAV algorithm combines the optimal k-allocation solution and the dynamic correction of HV estimation, which effectively improves the efficiency of the system. Although the average travel speed of the system is slightly lower at first based on Fig. 13 (d), the subsequent simulation dura-

tion exhibits significantly improved speeds and reduced total delay compared to all HV environments. Thus far, we have demonstrated the effectiveness and robustness of our CAV cooperative driving framework across various environments.

In conclusion, through replay tests, we first demonstrate that the actions and trajectories of real-world heterogeneous HVs can be well reproduced by combining non-cooperative games with IRL. Then, three simulation cases are conducted for the verification and validation of our cooperative driving framework. The simulation results show that with the designed framework, CAVs are capable of coordinating with heterogeneous vehicles in high-density, mixed intersections. Besides, as the proportion of CAVs increases, the efficiency and safety of the intersection system improve significantly.

VI. CONCLUSION

With the combination of the normalized cooperative game and level-k game, a cooperative driving framework is proposed for CAVs to address the driving conflict in high-density, mixed intersections.

Differing from reservation-based control methods, the proposed NCL game theoretic approach is capable of cooperating and even actively coordinating with other vehicles and shows a superior performance across different ROP. In addition to the cooperation between CAVs, CAVs can actively collaborate with heterogeneous HVs. Three simulation cases were conducted for verification and validation, including the comparative analysis with different methods, the case study under different ROP, and the interaction analysis with heterogeneous HVs. The performance of the intersection system is analyzed through average travel speed, total delay, and PET. Experiment results show that the proposed cooperative driving framework is capable of confronting complex, mixed traffic scenarios and therefore nudging the intersection system to greater efficiency and safety.

REFERENCES

- [1] Y. Xing, Z. Hu, P. Hang, and C. Lv, "Learning from the dark side: A parallel time series modelling framework for forecasting and fault detection on intelligent vehicles," *IEEE Transactions on Intelligent Vehicles*, 2023.
- [2] C. F. Camerer and E. Fehr, "When does 'economic man' dominate social behavior?" *science*, vol. 311, no. 5757, pp. 47–52, 2006.
- [3] J. Liu, D. Zhou, P. Hang, Y. Ni, and J. Sun, "Towards socially responsive autonomous vehicles: A reinforcement learning framework with driving priors and coordination awareness," *IEEE Transactions on Intelligent Vehicles*, 2023.
- [4] P. Hang, Y. Zhang, N. de Boer, and C. Lv, "Conflict resolution for connected automated vehicles at unsignalized roundabouts considering personalized driving behaviours," *Green Energy and Intelligent Transportation*, vol. 1, no. 1, p. 100003, 2022.
- [5] C. Xu, W. Zhao, L. Li, Q. Chen, D. Kuang, and J. Zhou, "A nash q-learning based motion decision algorithm with considering interaction to traffic participants," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 12 621–12 634, 2020.
- [6] M. Garzón and A. Spalanzani, "Game theoretic decision making for autonomous vehicles' merge manoeuvre in high traffic scenarios," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 3448–3453.
- [7] D. Li, A. Liu, H. Pan, and W. Chen, "Safe, efficient and socially-compatible decision of automated vehicles: a case study of unsignalized intersection driving," *Automotive Innovation*, pp. 1–16, 2023.
- [8] Y. Yan, L. Peng, T. Shen, J. Wang, D. Pi, D. Cao, and G. Yin, "A multi-vehicle game-theoretic framework for decision making and planning of

- autonomous vehicles in mixed traffic,” *IEEE Transactions on Intelligent Vehicles*, 2023.
- [9] A. Aksjonov and V. Kyrki, “A safety-critical decision-making and control framework combining machine-learning-based and rule-based algorithms,” *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, vol. 7, no. 10-07-03-0018, 2023.
 - [10] N. Li, I. Kolmanovsky, A. Girard, and Y. Yildiz, “Game theoretic modeling of vehicle interactions at unsignalized intersections and application to autonomous vehicle control,” in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 3215–3220.
 - [11] S. Karimi and A. Vahidi, “Receding horizon motion planning for automated lane change and merge using monte carlo tree search and level-k game theory,” in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1223–1228.
 - [12] P. Chu, Y. Yu, J. Yang, and C. Huang, “Understanding the mechanism behind young drivers’ distracted driving behaviour based on sor theory,” *Journal of Transportation Safety & Security*, vol. 14, no. 10, pp. 1655–1673, 2022.
 - [13] R. Yuan, M. Abdel-Aty, Q. Xiang, Z. Wang, and X. Gu, “A temporal multi-gate mixture-of-experts approach for vehicle trajectory and driving intention prediction,” *IEEE Transactions on Intelligent Vehicles*, 2023.
 - [14] J. Do, K. Han, and S. B. Choi, “Lane change-intention inference and trajectory prediction of surrounding vehicles on highways,” *IEEE Transactions on Intelligent Vehicles*, 2023.
 - [15] K. Yu, L. Lin, M. Alazab, L. Tan, and B. Gu, “Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5g-enabled intelligent transportation system,” *IEEE transactions on intelligent transportation systems*, vol. 22, no. 7, pp. 4337–4347, 2020.
 - [16] R. Yao, W. Zeng, Y. Chen, and Z. He, “A deep learning framework for modelling left-turning vehicle behaviour considering diagonal-crossing motorcycle conflicts at mixed-flow intersections,” *Transportation research part C: emerging technologies*, vol. 132, p. 103415, 2021.
 - [17] L. Bi, G. Gan, J. Shang, and Y. Liu, “Queueing network modeling of driver lateral control with or without a cognitive distraction task,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1810–1820, 2012.
 - [18] Z. Huang, J. Wu, and C. Lv, “Driving behavior modeling using naturalistic human driving data with inverse reinforcement learning,” *IEEE Transactions on Intelligent Transportation Systems*, 2021.
 - [19] W. Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, and D. Rus, “Social behavior for autonomous vehicles,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 50, pp. 24972–24978, 2019.
 - [20] I. V. Chremos, L. E. Beaver, and A. A. Malikopoulos, “A game-theoretic analysis of the social impact of connected and automated vehicles,” in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020, pp. 1–6.
 - [21] S. Mariani, G. Cabri, and F. Zambonelli, “Coordination of autonomous vehicles: taxonomy and survey,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 1, pp. 1–33, 2021.
 - [22] C. Yu, W. Sun, H. X. Liu, and X. Yang, “Managing connected and automated vehicles at isolated intersections: From reservation-to optimization-based methods,” *Transportation research part B: methodological*, vol. 122, pp. 416–435, 2019.
 - [23] W. Zhong, K. Li, J. Shi, J. Yu, and Y. Luo, “Reservation-prioritization-based mixed-traffic cooperative control at unsignalized intersections,” *IEEE Transactions on Intelligent Vehicles*, 2024.
 - [24] H. Xu, Y. Zhang, L. Li, and W. Li, “Cooperative driving at unsignalized intersections using tree search,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4563–4571, 2019.
 - [25] D. Carlino, S. D. Boyles, and P. Stone, “Auction-based autonomous intersection management,” in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. IEEE, 2013, pp. 529–534.
 - [26] H. Vu, S. Aknine, and S. D. Ramchurn, “A decentralised approach to intersection traffic management,” in *IJCAI*, 2018, pp. 527–533.
 - [27] H. Wang, Y. Li, and H. V. Zhao, “Performance analysis of road intersections based on game theory and dynamic level-k model,” in *2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*. IEEE, 2020, pp. 1112–1119.
 - [28] C. Chen, Q. Xu, M. Cai, J. Wang, J. Wang, and K. Li, “Conflict-free cooperation method for connected and automated vehicles at unsignalized intersections: Graph-based modeling and optimality analysis,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 21897–21914, 2022.
 - [29] T. Litman, *Autonomous vehicle implementation predictions*. Victoria Transport Policy Institute Victoria, BC, Canada, 2017.
 - [30] M. Wang, S. P. Hoogendoorn, W. Daamen, B. van Arem, and R. Happee, “Game theoretic approach for predictive lane-changing and car-following control,” *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 73–92, 2015.
 - [31] D. Sadigh, N. Landolfi, S. S. Sastry, S. A. Seshia, and A. D. Dragan, “Planning for cars that coordinate with people: leveraging effects on human actions for planning and active information gathering over human internal state,” *Autonomous Robots*, vol. 42, no. 7, pp. 1405–1426, 2018.
 - [32] R. Azimi, G. Bhatia, R. R. Rajkumar, and P. Mudalige, “Stip: Spatio-temporal intersection protocols for autonomous vehicles,” in *2014 ACM/IEEE international conference on cyber-physical systems (ICCPS)*. IEEE, 2014, pp. 1–12.
 - [33] M. Sharath and N. R. Velaga, “Enhanced intelligent driver model for two-dimensional motion planning in mixed traffic,” *Transportation Research Part C: Emerging Technologies*, vol. 120, p. 102780, 2020.
 - [34] X. Zhao, Y. Tian, and J. Sun, “Yield or rush? social-preference-aware driving interaction modeling using game-theoretic framework,” in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021, pp. 453–459.
 - [35] Y. Ni, M. Wang, J. Sun, and K. Li, “Evaluation of pedestrian safety at intersections: A theoretical framework based on pedestrian-vehicle interaction patterns,” *Accident Analysis & Prevention*, vol. 96, pp. 118–129, 2016.
 - [36] K.-T. Chen, H.-Y. W. Chen, and H.-Y. W. Chen, “Driving style clustering using naturalistic driving data,” *Transportation research record*, vol. 2673, no. 6, pp. 176–188, 2019.
 - [37] Y. Rahmati, M. K. Hosseini, and A. Talebpour, “Helping automated vehicles with left-turn maneuvers: a game theory-based decision framework for conflicting maneuvers at intersections,” *IEEE Transactions on Intelligent Transportation Systems*, 2021.
 - [38] Y. Zhang, W. Wu, and W. Zhang, “Noncooperative game-based cooperative maneuvering of intelligent surface vehicles via accelerated learning-based neural predictors,” *IEEE Transactions on Intelligent Vehicles*, 2022.
 - [39] R. Tian, S. Li, N. Li, I. Kolmanovsky, A. Girard, and Y. Yildiz, “Adaptive game-theoretic decision making for autonomous vehicle control at roundabouts,” in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 321–326.
 - [40] J. Zhang, J. Pu, J. Xue, M. Yang, X. Xu, X. Wang, and F.-Y. Wang, “Hiveqpt: Human-machine-augmented intelligent vehicles with generative pre-trained transformer,” *IEEE Transactions on Intelligent Vehicles*, 2023.
 - [41] B. D. Ziebart, A. L. Maas, J. A. Bagnell, A. K. Dey *et al.*, “Maximum entropy inverse reinforcement learning,” in *Aaai*, vol. 8. Chicago, IL, USA, 2008, pp. 1433–1438.
 - [42] S. Karimi, A. Karimi, and A. Vahidi, “Level- k reasoning, deep reinforcement learning, and monte carlo decision process for fast and safe automated lane change and speed management,” *IEEE Transactions on Intelligent Vehicles*, 2023.
 - [43] R. Tian, N. Li, I. Kolmanovsky, Y. Yildiz, and A. R. Girard, “Game-theoretic modeling of traffic in unsignalized intersection network for autonomous vehicle control verification and validation,” *IEEE Transactions on Intelligent Transportation Systems*, 2020.
 - [44] X. Wang, S. Zhang, and H. Peng, “Comprehensive safety evaluation of highly automated vehicles at the roundabout scenario,” *IEEE Transactions on Intelligent Transportation Systems*, 2022.
 - [45] K. Dresner and P. Stone, “A multiagent approach to autonomous intersection management,” *Journal of artificial intelligence research*, vol. 31, pp. 591–656, 2008.
 - [46] R. Xing, Z. Su, Q. Xu, N. Zhang, and T. H. Luan, “Secure content delivery for connected and autonomous trucks: A coalition formation game approach,” *IEEE Transactions on Intelligent Transportation Systems*, 2022.
 - [47] C. Chen, Q. Xu, M. Cai, J. Wang, J. Wang, and K. Li, “Conflict-free cooperation method for connected and automated vehicles at unsignalized intersections: Graph-based modeling and optimality analysis,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 21897–21914, 2022.
 - [48] B. Xu, S. E. Li, Y. Bian, S. Li, X. J. Ban, J. Wang, and K. Li, “Distributed conflict-free cooperation for multiple connected vehicles at unsignalized intersections,” *Transportation Research Part C: Emerging Technologies*, vol. 93, pp. 322–334, 2018.
 - [49] J. Guo, S. Cheng, and Y. Liu, “Merging and diverging impact on mixed traffic of regular and autonomous vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1639–1649, 2020.

- [50] L. Hu, T. Lu, G. Li, X. Zhang, and H. Cai, "Automatic generation of intelligent vehicle testing scenarios at intersections based on natural driving datasets," *IEEE Transactions on Intelligent Vehicles*, 2023.
- [51] J. Taylor, X. Zhou, N. M. Roupail, and R. J. Porter, "Method for investigating intradriver heterogeneity using vehicle trajectory data: A dynamic time warping approach," *Transportation Research Part B: Methodological*, vol. 73, pp. 59–80, 2015.
- [52] B. Greenshields, J. Bibbins, W. Channing, and H. Miller, "A study of traffic capacity," in *Highway research board proceedings*, vol. 1935. National Research Council (USA), Highway Research Board, 1935.
- [53] B. L. Allen, B. T. Shin, and P. J. Cooper, "Analysis of traffic conflicts and collisions," Tech. Rep., 1978.



Shiyu Fang received the B.S. degree in transportation engineering from Jilin University, Changchun, China. He is currently pursuing the Ph.D. degree with the Department of Traffic Engineering, Tongji University, Shanghai, China. His main research interests include decision making and motion planning for autonomous vehicles.



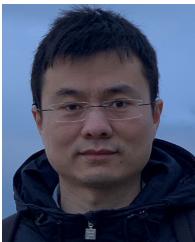
Yang Xing (Member, IEEE), received his Ph.D. degree from Cranfield University, UK. He is currently a Lecturer in Applied Artificial Intelligence at Cranfield University. He was a Research Associate at the University of Oxford and a Research Fellow at Nanyang Technological University. His research interests include human behaviour modelling, intelligent multi-agent collaboration, and intelligent vehicles. He is serving as an Associate Editor of IEEE TRANSACTIONS ON INTELLIGENT VEHICLES.



Jian Sun received the Ph.D. degree from Tongji University, Shanghai, China, in 2006. In 2011, he was with Tongji University, as a Lecturer, and then promoted to the position as a Professor, where he is currently a Professor with the College of Transportation Engineering and the Dean of the Department of Traffic Engineering. His main research interests include traffic flow theory, traffic simulation, connected vehicle-infrastructure system, and intelligent transportation systems.



Peng Hang (Member, IEEE) received the Ph.D. degree with the School of Automotive Studies, Tongji University, Shanghai, China, in 2019. He is currently a Research Professor with the Department of Traffic Engineering, Tongji University, Shanghai, China. In 2018, he was a Visiting Researcher with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. From 2020 to 2022, he was a Research Fellow with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. His research interests include vehicle dynamics and control, decision making, motion planning, and motion control for autonomous vehicles. He is an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, Journal of Field Robotics, IET Smart Cities, and SAE International Journal of Vehicle Dynamics, Stability, and NVH.



Chongfeng Wei (Member, IEEE) received the Ph.D. degree in mechanical engineering from the University of Birmingham, U.K., in 2015. He is currently an Assistant Professor with Queen's University Belfast, U.K. His current research interests include decision making and control of intelligent vehicles, human-centric autonomous driving, cooperative automation, and dynamics and control of mechanical systems. He is serving as an Associate Editor for IEEE OPEN JOURNAL OF INTELLIGENT TRANSPORTATION SYSTEMS, IEEE TRANSACTIONS ON INTELLIGENT VEHICLES, and IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS.



DSPACE

<https://dspace.org/>

Cooperative driving of connected autonomous vehicles in heterogeneous mixed traffic: a game theoretic approach

Fang, Shiyu; Hang, Peng; Wei, Chongfeng; Xing, Yang; Sun, Jian

2024-05-13

Institute of Electrical and Electronics Engineers (IEEE)

<https://doi.org/10.1109/TIV.2024.3399694>

Fang S, Hang P, Wei C, et al., (2024) Cooperative driving of connected autonomous vehicles in heterogeneous mixed traffic: a game theoretic approach. IEEE Transactions on Intelligent Vehicles. Available online 13 May 2024