2021 IEEE Intelligent Transportation Systems Conference (ITSC)
Indianapolis, USA. September 19-21, 2021

2021 IEEE International Intelligent Transportation Systems Conference (ITSC) | 978-1-7281-9142-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/ITSC48978.2021.9564702

Yield or Rush? Social-Preference-Aware Driving Interaction Modeling Using Game-Theoretic Framework

Xiaocong Zhao, Ye Tian, Jian Sun*

Abstract— **The newly formed hybrid traffic flow where Human-driven Vehicles (HV) share roads with Autonomous Vehicles (AV) is a foreseeable trend in the modern transportation system, making HV-AV interactions inevitable. This paper proposes a Parallel-Game-based Interaction Model (PGIM) to describe social-preference-aware driving interactions. Firstly, quantitively defined social preference is integrated into a game-theoretic model to capture the social patterns in interactions between heterogeneous drivers. Secondly, an active semantic decision-making method is developed within PGIM via a counterfactual reasoning process. This method enables agents to gradually understand the social preference of an unacquainted interacting agent during the interaction and make semantic decisions social-compliantly. We verify the interactive driving performance of PGIM in simulations by showcasing in the unprotected left-turn scenario. The results show that the proposed PGIM could trigger distinct interaction evolvement patterns by varying social preferences of interacting agents. Therefore, the PGIM brings a potential way from the perspective of the social property to explicitly reveal the mechanism in human-involved driving interactions.**

I. Introduction

With the speeding deployment of AVs both in localized road tests and as regular private vehicles[1], the interaction between HV and AV in hybrid traffic is becoming a new normal. During such a transition, collision-free is no longer the only requirement for a production-ready AV. To drive in a human-involved environment, AV needs to conduct navigations that are not only collision-free but also social-compliant. As reported in many works, interactive driving scenarios, where involved road users affect each other in the behavior selection [2], are still challenging for the current AVs [3], [4]. For example, the unprotected left-turn scenario where a left-turn vehicle interacts with a front-coming vehicle in an unprotected (two-phase) intersection is commonly observed in daily driving (see in Fig. 1). In this scenario, it is legally correct to guarantee the right-of-way of the front-coming streams, while it is socially preferable in the human driving environment to balance the delay of both left-turn and front-coming streams[5]. AV usually suffers from the dilemma between *legally correct* and *socially preferable*, and tends to yield excessive conservation for so-called safe operation in interactive driving scenarios like left-turns [6], [7]. In dense traffic, front-coming vehicles may never end, which means

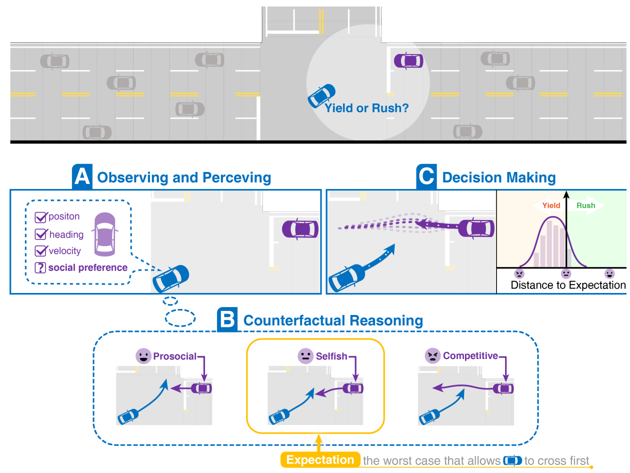

Figure 1. Diagram of social-preference-aware decision-making in the unprotected left-turn interaction. *(A) Unknowing the social preference of the front-coming car (purple), the left-turn car (blue) get into the dilemma of whether to rush or yield; (B) By constructing virtual agents with distinct social preferences and conducting virtual interactions, the blue car could find cases where it could cross first and set the social preference of the virtual front-coming car in the worst case as the expectation to the front-coming car in the real interaction; (C) By comparing the observed trajectory and those in virtual interactions, the blue car found the purple car's trajectory negatively deviates from the expectation and therefore decided to yield.*

that without active driving maneuvers, the left-turn vehicle could fall into a long-time standstill. As such, potential mismatches between human drivers' expectations and AVs' conductions could lead to the decay of traffic efficiency, road rage of human drivers, and sometimes even accidents [8].

Two main issues have impeded the behavioral congruency of AVs in interactive driving scenarios. The first one is the lack of adaptability to a specific interacting agent. AV is usually designed to behave like an average human driver[9], [10] or switch behaviors among several modes trained from featured human driving data [11]. While in interactive driving scenarios, interaction evolvement highly depends on interacting agents' individual behavioral preferences, which could inevitably deviate from the average level of a certain driver group [12]. Consequently, the decision-making process suffers from the predilection to use such vague averages rather than individualized information.

The second issue is the lack of counterfactual reasoning ability. A common solution to navigating in interactive driving scenarios is generating a belief on and further predicting the driving situation based on the observable traffic dynamics and then acting accordingly. However, solutions of this kind usually neglect the influence of self-actions on the evolvement of the traffic situation. It is the dependence between behavior

This research was sponsored by the Natural Science Foundation of China (U1764261) and the National Key Research and Development Program of China (2018YFB1600505).

Xiaocong Zhao, Ye Tian, and Jian Sun are with the Department of Traffic Engineering and Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China.
*Corresponding author: Jian Sun (E-mail: sunjian@tongji.edu.cn)

selections of the subjective vehicle and other road users that develops the interactive driving event. Neglecting such dependence would inevitably bring the cost of considerable uncertainty and consequently leads to passive driving maneuvers. Hence, to actively plan and navigate in interactive scenarios, AV needs to be aware of the dependence between behavior selection and situation evolvement, and bases the planning process on (1) the fact that other road users will respond to self-actions and (2) the knowledge of how other road users might react to self-actions.

Motivated by the above two issues, this paper proposes a Parallel-Game-based Interaction Model (PGIM) to serve active and social-compliant driving interactions. As shown in Fig. 1 (A) through (C), the heterogeneity in human social preference is taken into account to perform individualized navigation in interactive driving scenarios. The main contribution of this work is three-folded:

- Proposing a game-theoretic modeling framework to describe the social-preference-aware driving interaction between heterogeneous drivers.
- Designing a parallel-game-based method to enable active decision-making in interactive driving scenarios.
- Presenting insights on analyzing the quality of driving interactions from the perspective of social preference.

The rest of this paper is arranged as follows: Firstly, previous works related to interaction modeling are reviewed in §2. Then, we formulate the problem in §3 to mathematically describe the general two-player interactive driving scenario. In §4, we present the PGIM to depict driving interactions and solve the semantic decision-making problem therein. Simulation results are presented and discussed in §5. Finally, we conclude our work in §6.

II. Related Works

A simple-yet-fundamental criterion for evaluating autonomous driving performance is whether the target scenario is congruent to the actual driving environment. The actual driving environment is interactive due to the limited road resources and is stochastic due to the heterogeneous road users. This section will roll out previous works on driving interaction modeling according to whether the *interactivity* and *heterogeneity* of the driving environment are considered.

Interactivity. Theoretically, to verify the interactive driving performance of a specific method, all the background vehicles should possess the interactive driving ability. However, calling for interactive driving ability to verify interactive driving performance could then become a "chicken and egg" problem [13]. Therefore, assumptions, to varying degrees, on background vehicles were always made to give a reasonably interactive environment.

Based on the "like-me" assumption, the subjective vehicle could make decisions assuming interacting vehicles share the same strategy. A typical method for modeling the "like-me" interaction is the game theory. Tian [14] developed interactive background vehicles using the level-k game-theoretic framework. As such, the background vehicles are online identified to be level-1 or level-2 agents to serve case-orientated AV control. Assuming the cost function of background vehicles is known to the subjective vehicle, Wang [15] proposed a differential-game-based framework that allows every single agent to plan for other surrounding agents. Similarly, Bahram's work assumes the surrounding vehicles are capable of re-planning themselves based on the maneuvers of the cognitive (subjective) vehicle, thus forming an interactive lane-changing scenario [16]. Besides, efforts on game-theoretic robot interaction modeling have also been widely rolling out in recent years, serving robot racing [17], drone racing [18], etc.

Heterogeneity of environmental agents has been taken into account in previous works from roughly two perspectives. In the first type, environmental agents follow a weighted reward/cost function, and the difference between individuals is characterized by varied weight metrics. A typical dividing criterion for background vehicles of this type is aggressiveness, defining the conservative and the aggressive driver as who cares relatively more about safety and travel delay, respectively. Cases of the second type differentiate background vehicles via how they allocate the self-reward and the reward of others during the planning process. A psychological term named Social Value Orientation (SVO), determining how an agent allocates the reward of his/herself and that of others[19], [20], is employed in [21] to quantify the social preference of individual drivers. Aside from directly embedding social preference into the reward function, Herman et al. [22] proposed a bounded continuous variable named social acceptability to indicate how humans perceive a driving style that follows a specific reward function.

POMDP-based methods allow heterogeneity-involved interaction handling by taking the intention of case-orientated surrounding vehicles as hidden variables, however, at the cost of expensive computation. Efforts have been made to solve POMDP online for vehicle operation via distinct approaches, including the simplification on state-space by predefining paths [23], discretization of interacting vehicles intentions[24], limitation on the region (of spatial interacting range and brief state) of interest [25] and nonparametric reinforcement learning [26], etc.

The state-of-the-art learning-based methods also provide promising solutions to implicit interaction modeling (e.g. [27]). However, with the main interest of abstracting explicit patterns in humanlike interaction, learning-based works will not be covered in this work. A systematic survey in [28] is recommended for a well-rounded knowledge of learning-based interaction modeling methods.

III. Problem Description

Consider the scenario in which two agents traverse across an intersection with conflict along their future paths (see in Fig. 2 as an example). Assume that two agents cannot communicate with each other through any over-the-air connection and only temporal position $\boldsymbol{p}_i \in \mathbb{R}^2$ and heading angle $\psi_i \in \mathbb{R}$ could be observed.

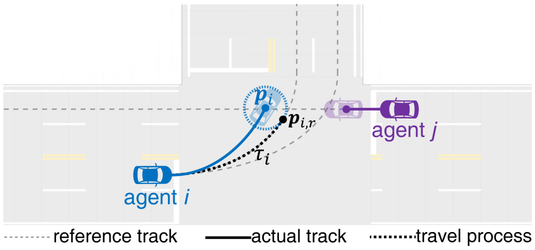

Figure 2. Definition of driving progress in the unprotected left-turn scenario

As the problem is mathematically symmetric for both two agents, we would roll out the formulation from the perspective of agent i for the sake of briefness. The general problem could be stated as: agent i tries to plan an *N*-segments collision-free *trajectory* $\boldsymbol{p}_i^{0\to N} \in \mathbb{R}^{2\times(N+1)}$ to cross the intersection as soon as possible. Here, *trajectory* refers to a track coupled with a speed profile that can be followed by a vehicle. The position $\boldsymbol{p}_i^n$ is the end of the n-th segment ($n \in [1, N]$) and $\boldsymbol{p}_i^0$ is the starting position. Agent i travels each segment within a constant time interval δt, constructing a speed profile naturally. Besides, considering a common rule in real-world driving, agent i would try to follow the reference path indicated by the road structure (see the dashed central line in Fig. 2). To depict the points above, we write the reward function as

$$R_i(\boldsymbol{p}_i^{0\to N}) = \tau_i(\boldsymbol{p}_i^N) - \alpha \sum_{n=1}^{N} \left\|\boldsymbol{p}_{i,r}^n - \boldsymbol{p}_i^n\right\| \tag{1}$$

where $\tau_i(\boldsymbol{p}_i^N)$ denotes the length from the starting position $\boldsymbol{p}_{i,r}^0$ to the end position $\boldsymbol{p}_{i,r}^N$ along the reference path. The subscript r indicates that the point is on the reference path, for example, $\boldsymbol{p}_{i,r}^n(\boldsymbol{p}_i^n)$ represents the on-reference point closest to $\boldsymbol{p}_i^n$ (see illustration in Fig. 2). $\alpha \geq 0$ is a free parameter. In Eq. (1), the first term shows the desire for driving progress, and the second term indicates the preference to avoid overly deviating from the central line of the road.

IV. Methodology

§3 presents a decoupled driving interaction problem from the perspective of a single agent, while interaction cannot be solved unilaterally. In this section, we develop the Parallel-Game-based Interaction Model (PGIM) to further concrete the bilateral driving interaction problem. Firstly, the individual social preference is quantitively defined to represent the personalized driving strategy. Then, we employ the game-theoretic formulation to capture between-agents driving patterns in an interactive form. Finally, a parallel-game-based method is proposed to serve active semantic decision-making in driving interactions.

A. Social Preference

To capture the social patterns in real-world driving, we adopt Social Value Orientation (SVO) [19]–[21], a time-variable psychological term that indicates how an individual allocates the rewards of him/herself and that of others to quantify individuals' social preference. As such, the utility of each interacting agent comprises the rewards of both two agents, that is

$$U_i = \cos(\varphi_i) \cdot R_i + \sin(\varphi_i) \cdot R_j \tag{2}$$

where $\varphi_i \in [-\pi, \pi]$ is the SVO of agent i. The value of φ affects the interacting strategy of an agent. For example, a prosocial agent ($\varphi = \pi/4$) cares equally about itself and others, while $\varphi = 0$ will leads to an egoistic agent which is totally self-interested. According to Eq. (1), the reward of agent j, namely R_j, only depends on j's own actions. This means during the trajectory optimization for agent i, R_j cannot be directly influenced by agent i. However, agent i affects the feasible solution set of agent j, and in turn affects agent j's optimal reward. Hence, we consider the optimal reward of j R_j^* instead in the following formulations.

Thus, considering the personalized SVO, the planning problem for agent i is written as

$$\boldsymbol{p}_i^* = \arg\max_{\boldsymbol{p}_i^{0\to N} \in \mathbb{P}_i} U_i(\boldsymbol{p}_i^{0\to N}) \tag{3a}$$

$$s.t.\ U_i(\boldsymbol{p}_i^{0\to N}) = \cos(\varphi_i) \cdot R_i(\boldsymbol{p}_i^{0\to N}) + \sin(\varphi_i) \cdot R_j^*(\boldsymbol{p}_i^{0\to N}) \tag{3b}$$

$$\boldsymbol{p}_i^n = \mathcal{D}(\boldsymbol{p}_i^{n-1}, \boldsymbol{u}_i^n) \tag{3c}$$

$$\|\boldsymbol{u}_i^n\| \leq \overline{u_i} \in \mathbb{R}^+ \tag{3d}$$

$$\left\|\boldsymbol{p}_i^n - \boldsymbol{p}_j^n\right\|^2 \geq 2 l_{veh}^2 \tag{3e}$$

$$\left\|\boldsymbol{p}_{i,r}^n(\boldsymbol{p}_i^n) - \boldsymbol{p}_i^n\right\| \leq w_{lane} - w_{veh} \tag{3f}$$

where

- U_i is the utility function of agent i;
- $R_j^* = R_j|_{\boldsymbol{p}_j^{0\to N} = \boldsymbol{p}_j^*}$ is the optimal reward of agent j;
- $\mathbb{P}_i$ is the feasible solution set for agent i;
- $\mathcal{D}(\cdot)$ denotes the vehicle dynamics which is a bicycle model in this work; $\boldsymbol{u}_i^n$ is a constant control input during time interval $t \in [(n-1)\delta t,\ n\delta t]$;
- $\overline{u_i}$ represents the control limit;
- with the length of the vehicle l_{veh}, the inequality constraints (3e) is applied to avoid collisions between interacting agents;
- with the width of the lane w_{lane} and the width of the vehicle w_{veh}, the inequality constraints (3f) prevent the vehicle from leaving the lane.

B. Interactive Motion Planning

We could solve the problem described in Eq. (3a-f) unilaterally from the perspective of agent i to this end. However, considering two agents share a similar strategy which is known to each other, agent j goes through the same planning process as denoted by Eq. (3a-f). Therefore, agent i needs to update $\boldsymbol{p}_j^*$ for agent j by taking i's own current plan into account. Thus, agent i plans for itself and its interacting agent in turn, and the iterative process will theoretically stop when agent i could not further refine any one of two agents' plans. For the k-th iteration of this iterative process, we could rewrite the Eq. (3a-f) in an iterative form:

$$\boldsymbol{p}_i^{(k),*} = \arg \max_{\boldsymbol{p}_i^{(k)} \in \mathbb{P}_i\left(\boldsymbol{p}_j^{(k-1),*}\right)} U_i\left(\boldsymbol{p}_i^{(k)}\right) \tag{4a}$$

$$s.t.\ U_i\left(\boldsymbol{p}_i^{(k)}\right) = \cos(\varphi_i) \cdot R_i\left(\boldsymbol{p}_i^{(k)}\right) + \sin(\varphi_i) \cdot R_j^*\left(\boldsymbol{p}_i^{(k)}\right) \tag{4b}$$

$$\boldsymbol{p}_i^n = \mathcal{D}(\boldsymbol{p}_i^{n-1}, \boldsymbol{u}_i^n) \tag{4c}$$

$$\|\boldsymbol{u}_i^n\| \leq \overline{u_i} \in \mathbb{R}^+ \tag{4d}$$

$$\left\|\boldsymbol{p}_i^n - \boldsymbol{p}_j^n\right\|^2 \geq 2 l_{veh}^2 \tag{4e}$$

$$\left\|\boldsymbol{p}_{i,r}^n(\boldsymbol{p}_i^n) - \boldsymbol{p}_i^n\right\| \leq w_{lane} - w_{veh} \tag{4f}$$

It is worth noting that the inequality constraints (4e) set limits according to the position of agent j, which means the feasible solution set of agent i depends on the solution of agent j, i.e. $\mathbb{P}_i = \mathbb{P}_i\left(\boldsymbol{p}_j^{(k-1),*}\right)$. Hence, to trigger the iteration process, an initial guess of $\boldsymbol{p}_j^{(0),*}$ is made by maximizing j's rewards to produce $\mathbb{P}_i\left(\boldsymbol{p}_j^{(0),*}\right)$. Besides, the utility function (4b) of agent i comprises an agent-j-related term $R_j^*(\boldsymbol{p}_i^{0\to N})$. However, without inter-vehicle communication, agent i is not supposed to obtain the plan of its interacting counterpart. Hence, it is difficult to denote the effect of $\boldsymbol{p}_i^{(k)}$ on the optimal reward of agent j, i.e., $R_j^*\left(\boldsymbol{p}_i^{(k)}\right)$, in a closed-form expression. To address this, we applied a numerical solution based on sensitivity analysis presented in [18] to give a local approximation of $R_j^*\left(\boldsymbol{p}_i^{(k)}\right)$.

The whole Iterative Motion Planning (IMP) process for agent i could be denoted as $(\boldsymbol{p}_i^*, \widehat{\boldsymbol{p}}_j^*)$=$IMP(\boldsymbol{p}_i^0, \boldsymbol{p}_j^0, \varphi_i, \varphi_j)$ and is detailed by **Algorithm 1** (see in *Appendix A*). When the stationarity condition $\left\|\boldsymbol{p}_i^{(k),*} - \boldsymbol{p}_i^{(k-1),*}\right\| < \epsilon$ is satisfied, a solution tuple $(\boldsymbol{p}_i^*, \widehat{\boldsymbol{p}}_j^*)$ that satisfies the necessary condition for a Nash equilibrium is reached by the definition of *the best response of the best response*.

C. Active Decision-Making based on Parallel Game

As is needed in the IMP process, the interacting agent's SVO plays a significant role in the evolving trend of the interaction event, affecting the result of i's semantic decision-making, i.e., whether to yield or rush. From agent i's perspective, the more precisely i knows about j's SVO, the more targeted i's plan will be. Theoretically speaking, the semantic decision and the corresponding trajectory that maximizes i's utility could be produced by executing **Algorithm 1** given the exact SVO value of agent j. However, in a real-world situation, we usually have no prior knowledge of interacting counterpart's social preferences. The only information that may allow us to deduce the social preferences is the trajectory history. To address this, we proposed a parallel-game-based method for semantic decision-making by estimating interacting agents' SVO from the observed trajectory. The estimated SVO value is associated with measurable uncertainty.

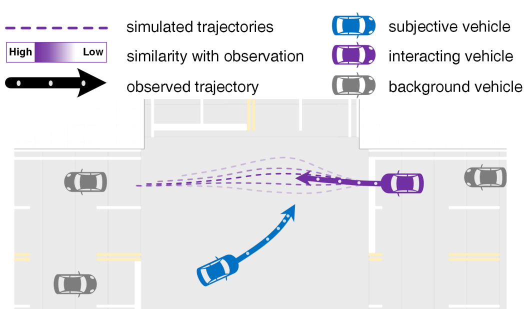

Figure 3. Estimating interacting agent's SVO by trajectory comparison from the perspective of blue car (agent i). *(Simulated trajectories are generated under varying settings of interacting agent's SVO. The lower transparency of simulated trajectories represents a higher similarity between observed and simulated trajectories.)*

As mentioned above, we could give a fixed answer on whether agent i could cross first given agent j's SVO at the beginning of the interaction by executing **Algorithm 1**. This allows us to conduct counterfactual reasoning to check whether a front-coming agent with a specific SVO would yield to agent i given the initial state of the scenario. As we currently have no prior knowledge about the SVO distribution in the human driver group, we first sample SVO values from a uniform distribution. Then, virtual agents with varying social preferences are constructed from these sampled SVO values. Afterward, parallel virtual games between agent i and each virtual agent are simulated simultaneously. The initial state of all virtual interactions is settled as the current state of the real interaction. By finding the most selfish agent that allows i to cross first, we obtain the loosest SVO expectation $\tilde{\varphi}_j$ to the agent j. Note that agent i expects agent j to be as altruistic as possible, but as long as the agent j is more altruistic than the loosest expectation, the left-turn maneuver is theoretically allowed. In this worst case, we could pre-calculate a trajectory that allows i to cross first. The process of obtaining the expectation to agent j's SVO, i.e. $\tilde{\varphi}_j$, and the pre-calculated trajectory $\boldsymbol{p}_i^*$ are detailed in **Algorithm 2** (see in *Appendix B*).

During the real-time interaction, agent i needs to verify j's SVO through the observation on j's actual track $\overline{\boldsymbol{p}}_j^{0\to n}$ at each step n. As shown in Fig. 3, the observation-based SVO is calculated under a straightforward assumption that the more precise the estimation of j's social preference, the closer the trajectory produced in the simulation to the actual. This process allows each agent to optimize actions through continuous information gathering and is detailed in **Algorithm 3** (see in *Appendix C*) using a particle filter. Thus, by comparing the expected SVO $\tilde{\varphi}_j$ and the observation-based SVO $\bar{\varphi}_j^n$, i could tell whether the pre-calculated trajectory $\boldsymbol{p}_i^*$ is feasible. We define the gap between the expectation and estimation of SVO as *Expectation Surplus* (*ES*), which could be mathematically denoted as $\bar{\varphi}_j^n - \tilde{\varphi}_j$ with a standard deviation of $\bar{\sigma}_{\varphi_j}^n$. A positive *ES* means a social cue indicating that the behaviors of the real interacting agent are advantageous for the rushing maneuver. Therefore, the left-turn agent i could theoretically conduct $\boldsymbol{p}_i^*$ to cross the intersection without yielding to the current interacting agent j given a positive *ES* confirmed at a certain confidence level. Otherwise, $\boldsymbol{p}_i^*$ will be updated in the next step.

V. Simulation and Discussion

Compared with the conventional decision-making process, we further take into account the social preference of interacting agents. This allows agents controlled by PGIM to be gradually conscious of *how the interacting agent may act* during the interaction progress. Therefore, even if the scenarios are identical (mainly in the manner of kinematic state and position of vehicles), PGIM-based agents may make different semantic decisions or conduct different trajectories when interacting with agents that have varying social preferences. In this section, simulations are conducted in the left-turn scenario to reveal the social-preference-aware driving performance of PGIM.

A. *Simulation Settings*

Simulations in group 1 are conducted to visualize the detailed motions in PGIM-based driving interactions. In this group, a prosocial (SVO=$\pi/4$) LT (short for Left-Turn here and after) agent interacts with respectively egoistic (SVO=0), prosocial (SVO=$\pi/4$), and competitive (SVO=$-\pi/4$) FC (short for Front-Coming here and after) agents. The initial states of the three cases are the same, and two agents in each case are designed to have potential conflicts, which means they reach the conflict area at the same time if no additional control is made.

Simulations in group 2 are conducted to reveal the effect of social preference on semantic driving decision-making from a statistical manner. We simulate 500 left-turn interaction cases in this group. For each case, the LT agent keeps the prosocial preference (SVO=$\pi/4$), while the SVO of the FC agent is sampled from a uniform distribution ranging from $-\pi/2$ to $\pi/2$. This value range covers the social preference from sadistic, competitive to egoistic, and to prosocial, altruistic. The setting of the initial state is roughly the same as that in group 1, but we add a Gaussian noise in the initial position of both two agents, still ensuring the potential conflicts.

Note that, in both two simulation groups, agents are not aware of the SVO of its interacting counterpart.

B. *Social-Preference-Aware Driving Performance*

As shown in Fig. 4, the driving maneuvers of the LT agent differ in three cases from the manner of both real-time plans and actual conductions. In the prosocial case only, the LT agent crosses the intersection without yielding to the FC agent. As depicted in Fig. 4(a), the LT agent plans to yield initially. However, after the FC agent showing obvious prosocial preference by slowing down, plans for rushing are generated and continuously modified by the LT agent. The path of the egoistic FC agent is marginally affected when the LT agent is nearby (see in Fig. 4(b)). A few plans for rushing are generated by the LT agent during the interaction but soon discarded after further confirmation on the FC agent's egoistic social preference. In Fig. 4(c), a similar pattern could be observed in the competitive case where the path of the FC agent is also slightly influenced. While in the competitive case, we notice the upward curve in the path generated by the FC agent (see in the lower right corner of Fig. 4(c) for details). Obviously, the

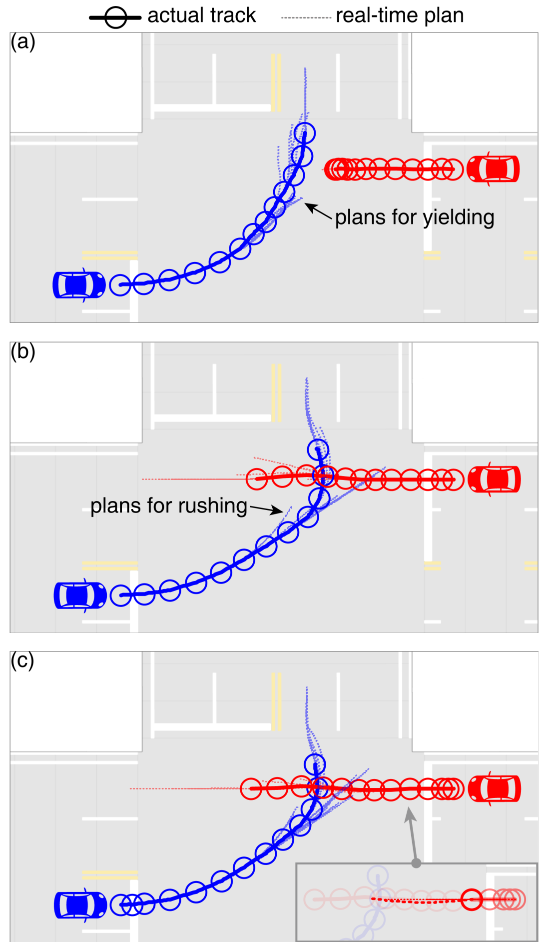

Figure 4. Left-turn interactions with distinct FC agents: (a) prosocial; (b) egoistic; (c) competitive.

upward curve brings no benefits to the FC agent's own driving progress but sets further limitations to the planning of the LT agent by getting in the potential ways of the LT agent, making LT agent's rushing plan even more unachievable. Note that the initial settings of the three cases are identical. Therefore, the distinguished interaction evolvement is triggered only by different social preferences of FC agents. This conveys that it may be difficult to predict the evolvement of driving interactions without the consideration of drivers' social preferences. The social value orientation may provide a semantically understandable way to further abstract information from gathered frame-by-frame observations to serve more predictable driving interactions.

Fig. 5 shows the number of cases that the LT yields or rushes with respect to the FC agent's SVO. We could find that the LT agent is more likely to cross first when interacting with an FC agent that values more about others' interests (indicated by a higher value of SVO). This finding is congruent to the real-world driving experience and indicates that PGIM-based agents could properly adjust driving behaviors according to the agent interacting with.

C. *Analysis on Interaction Quality*

Fig. 6 depicts the driving progress of the interacting agents in group 1. Interestingly, we find that the travel of both two

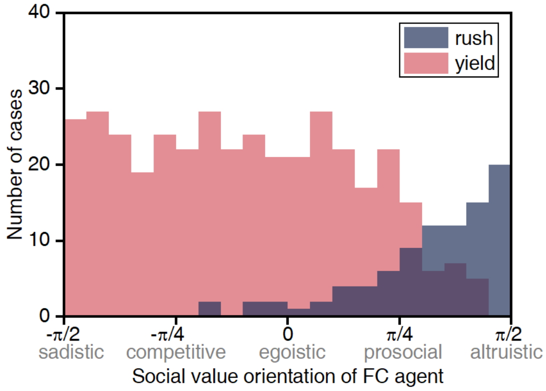

Figure 5. Semantic decisions of the LT agent

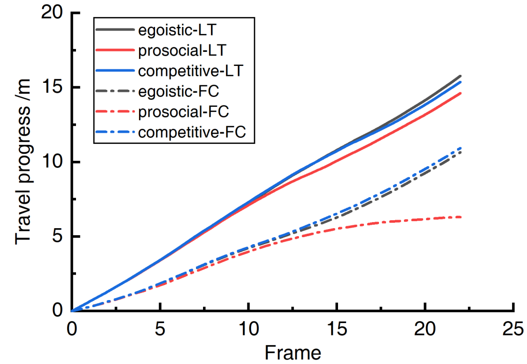

Figure 6. Driving progress in left-turn interactions

agents delays the most in the prosocial case where the FC agent does nothing aggressive, being even worse than that in the competitive case where the FC agent deliberately impedes the LT vehicle. At first glance, this finding might seem to convey that the prosocial FC agent acts non-cooperatively. However, the feeling of *the prosocial agent is not cooperative* does not transfer in our daily driving.

To dig into the findings above, we firstly give a narrow definition of the FC agent's cooperativeness from the perspective of the LT agent, that is, the willingness to let the LT agent cross the intersection as soon as possible. By this definition, the competitive FC agent is by no means cooperative because the nature of competitiveness will trigger behaviors that are detrimental to the LT agent's planning. However, we could also interpret the interaction between the LT agent and the competitive FC agent from another point of view, that is, the FC agent show aggressiveness as a part of the strategy to make it clear to the LT agent that the FC agent would cross first so that the LT agent could modify its plan (for yielding) more advance. As such, both the competitive FC agent's identifiable aggressiveness and egoistic FC agent's temporal-consistent selfishness make the interaction *crisper* to the LT agent and consequently benefit the LT agent's planning. Conversely, the prosocial FC agent shows a yielding signal to the LT agent midway through the interaction, which affects the LT agent's behavior selection and impedes the temporal consistency of its motion plan. This also explains why vagueness exists the most in interactions between two prosocial agents (see in Fig. 5). After all, the room for interpreting interactions is usually larger than that for happening, and the social preference of drivers may provide a new point of view for interpreting interaction quality.

VI. Conclusion

In this paper, we proposed a driving interaction model, the PGIM, that integrates the perception of social preference and the capability of counterfactual reasoning. Our main findings are as follows:

- By characterizing drivers with the social value orientation, PGIM could mirror the distinct interaction patterns resulted from the individualized social property of drivers.
- Through the counterfactual reasoning process, PGIM-based agents could identify the social preference of the interacting agents and adjust behaviors compliantly.
- The social preferences of interacting agents have an influence on the temporal consistency of involved agents' behaviors, which in turn affects the interaction quality (e.g., driving progress).

Ongoing work includes a well-rounded analysis of the interactions between agents of different pairs of social preferences. We would further investigate the effects of social preference on interaction quality and explore the possibility of utilizing such patterns to serve congruent and efficient interactive autonomous driving.

Appendix

A. *Algorithm 1-Iterative Motion Planning*

Input: $\boldsymbol{p}_i^0, \boldsymbol{p}_j^0, \varphi_i, \varphi_j$
Output: $(\boldsymbol{p}_i^*, \hat{\boldsymbol{p}}_j^*)$
Initialize:
$k \leftarrow 0$,
$\boldsymbol{p}_i^{(0),*} \leftarrow \boldsymbol{0}^{2\times(N+1)}$,
$\hat{\boldsymbol{p}}_j^{(0),*} \leftarrow \arg\max_{\boldsymbol{p}_j^{(0)}\in\mathbb{P}_i(\boldsymbol{p}_j^0)} R_j(\boldsymbol{p}_j^{(0)})$ Eq. (2)
while $k = 0$ or $\|\boldsymbol{p}_i^{(k),*} - \boldsymbol{p}_i^{(k-1),*}\| \geq \epsilon$ **do**
set $k \leftarrow k + 1$
update i's solution
$\boldsymbol{p}_i^{(k),*} \leftarrow \arg\max_{\boldsymbol{p}_i^{(k)}\in\mathbb{P}_i(\boldsymbol{p}_j^{(k-1),*})} U_i(\boldsymbol{p}_i^{(k)})$ Eq. (4a)
update the estimation of j's solution
$\hat{\boldsymbol{p}}_j^{(k),*} \leftarrow \arg\max_{\boldsymbol{p}_j^{(k)}\in\mathbb{P}_j(\boldsymbol{p}_i^{(k),*})} U_j(\boldsymbol{p}_j^{(k)})\big|_{\boldsymbol{p}_i^{(k),*}}$
set $\boldsymbol{p}_i^* \leftarrow \boldsymbol{p}_i^{(k),*}$, $\hat{\boldsymbol{p}}_j^* \leftarrow \hat{\boldsymbol{p}}_j^{(k),*}$
return $(\boldsymbol{p}_i^*, \hat{\boldsymbol{p}}_j^*)$

B. *Algorithm 2-Calculation of SVO Expectation*

Input: $\boldsymbol{p}_i^0, \boldsymbol{p}_j^0, \varphi_i$
Output: expectation of j's SVO $\tilde{\varphi}_j$, pre-calculated path $\boldsymbol{p}_i^*$
for n **from** $N/2$ **to** $-N/2$ **do**
set $\varphi_j \leftarrow n \cdot \pi/N$
run Algorithm 1
$(\boldsymbol{p}_i^*, \hat{\boldsymbol{p}}_j^*) \leftarrow IMP(\boldsymbol{p}_i^0, \boldsymbol{p}_j^0, \varphi_i, \varphi_j)$
if j yields to i according to $(\boldsymbol{p}_i^*, \hat{\boldsymbol{p}}_j^*)$ **do**
set $\tilde{\varphi}_j \leftarrow \varphi_j$
return $\tilde{\varphi}_j, \boldsymbol{p}_i^*$

C. ***Algorithm 3**-Observation-Based SVO Estimation*

1	**Input**: observed path $\overline{\boldsymbol{p}}_j^{0\to n}$, φ_i, $\tilde{\varphi}_j$
2	**Output**: observation-based SVO at step n (mean value $\bar{\varphi}_j^n$, standard deviation $\bar{\sigma}_{\varphi_j}^n$)
3	**Initialize**: M particles $\bar{\varphi}_j^{[m]}$, corresponding weights $\omega^{[m]}$
4	**while** $\frac{1}{\sum_{m=1}^{M}(\omega^{[m]})^2} < 0.5M$ **do**
5	**for** m **from** 1 **to** M **do**
6	sample $\bar{\varphi}_j^{[m]}$ from $\mathcal{N}\left(\tilde{\varphi}_j, \sigma_\varphi^2\right)$
7	run **Algorithm 1** $\left(\boldsymbol{p}_i^{*,[m]}, \widehat{\boldsymbol{p}}_j^{*,[m]}\right) \leftarrow IMP\left(\boldsymbol{p}_i^0, \boldsymbol{p}_j^0, \varphi_i, \bar{\varphi}_j^{[m]}\right)$
8	update $\omega^{[m]} \leftarrow \omega^{[m]} \times p\left(\overline{\boldsymbol{p}}_j^{0\to n} \middle| \widehat{\boldsymbol{p}}_j^{*,[m]}\right)$
9	normalize $\omega \leftarrow \frac{\omega}{\sum_{m=1}^{M}\omega^{[m]}}$
10	$\bar{\varphi}_j^n \leftarrow \sum_{m=1}^{M}(\omega^{[m]}\,\bar{\varphi}_j^{[m]})$
11	$\mu^n \leftarrow \sum_{m=1}^{M}\bar{\varphi}_j^{[m]}$
12	$\bar{\sigma}_{\varphi_j}^k \leftarrow \sqrt{\sum_{m=1}^{M}\left[\omega^{[m]}\left(\bar{\varphi}_j^{[m]} - \mu^k\right)^2\right]}$

REFERENCES

[1] B. Brown and E. Laurier, ‘The Trouble with Autopilots: Assisted and Autonomous Driving on the Social Road’, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver Colorado USA, May 2017, pp. 416–429. doi: 10.1145/3025453.3025462.

[2] G. Markkula *et al.*, ‘Defining interactions: a conceptual framework for understanding interactive behaviour in human and automated road traffic’, *Theor. Issues Ergon. Sci.*, vol. 21, no. 6, pp. 728–752, Nov. 2020, doi: 10.1080/1463922X.2020.1736686.

[3] A. Schieben, M. Wilbrink, C. Kettwich, R. Madigan, T. Louw, and N. Merat, ‘Designing the interaction of automated vehicles with other traffic participants: design considerations based on human needs and expectations’, *Cogn. Technol. Work*, vol. 21, no. 1, pp. 69–85, Feb. 2019, doi: 10.1007/s10111-018-0521-z.

[4] A. Rasouli and J. K. Tsotsos, ‘Autonomous Vehicles That Interact With Pedestrians: A Survey of Theory and Practice’, *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 900–918, Mar. 2020, doi: 10.1109/TITS.2019.2901817.

[5] R. J. Troutbeck and S. Kako, ‘Limited priority merge at unsignalized intersections’, *Transp. Res. Part Policy Pract.*, vol. 33, no. 3, pp. 291–304, Apr. 1999, doi: 10.1016/S0965-8564(98)00046-9.

[6] W. Zhan, C. Liu, C. Chan, and M. Tomizuka, ‘A non-conservatively defensive strategy for urban autonomous driving’, in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Nov. 2016, pp. 459–464. doi: 10.1109/ITSC.2016.7795595.

[7] S. Noh, ‘Decision-Making Framework for Autonomous Driving at Road Intersections: Safeguarding Against Collision, Overly Conservative Behavior, and Violation Vehicles’, *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3275–3286, Apr. 2019, doi: 10.1109/TIE.2018.2840530.

[8] H. Alambeigi, A. D. McDonald, and S. R. Tankasala, ‘Crash Themes in Automated Vehicles: A Topic Modeling Analysis of the California Department of Motor Vehicles Automated Vehicle Crash Database’, *ArXiv200111087 Stat*, Jan. 2020, Accessed: Mar. 15, 2021. [Online]. Available: http://arxiv.org/abs/2001.11087

[9] X. Zhao, R. He, and J. Wang, ‘How do drivers respond to driving risk during car-following? Risk-response driver model and its application in human-like longitudinal control’, *Accid. Anal. Prev.*, vol. 148, p. 105783, Dec. 2020, doi: 10.1016/j.aap.2020.105783.

[10] R. Fu, Z. Li, Q. Sun, and C. Wang, ‘Human-like car-following model for autonomous vehicles considering the cut-in behavior of other vehicles in mixed traffic’, *Accid. Anal. Prev.*, vol. 132, p. 105260, Nov. 2019, doi: 10.1016/j.aap.2019.105260.

[11] B. Zhu, Y. Jiang, J. Zhao, R. He, N. Bian, and W. Deng, ‘Typical-driving-style-oriented personalized adaptive cruise control design based on human driving data’, *Transp. Res. Part C Emerg. Technol.*, vol. 100, pp. 274–288, Mar. 2019, doi: 10.1016/j.trc.2019.01.025.

[12] H. Yu, H. E. Tseng, and R. Langari, ‘A human-like game theory-based controller for automatic lane changing’, *Transp. Res. Part C Emerg. Technol.*, vol. 88, pp. 140–158, Mar. 2018, doi: 10.1016/j.trc.2018.01.016.

[13] D. Isele, ‘Interactive Decision Making for Autonomous Vehicles in Dense Traffic’, in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Oct. 2019, pp. 3981–3986. doi: 10.1109/ITSC.2019.8916982.

[14] 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)*, Dec. 2018, pp. 321–326. doi: 10.1109/CDC.2018.8619275.

[15] M. Wang, S. P. Hoogendoorn, W. Daamen, B. van Arem, and R. Happee, ‘Game theoretic approach for predictive lane-changing and car-following control’, *Transp. Res. Part C Emerg. Technol.*, vol. 58, pp. 73–92, Sep. 2015, doi: 10.1016/j.trc.2015.07.009.

[16] M. Bahram, A. Lawitzky, J. Friedrichs, M. Aeberhard, and D. Wollherr, ‘A Game-Theoretic Approach to Replanning-Aware Interactive Scene Prediction and Planning’, *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 3981–3992, Jun. 2016, doi: 10.1109/TVT.2015.2508009.

[17] Z. Wang, R. Spica, and M. Schwager, ‘Game Theoretic Motion Planning for Multi-robot Racing’, in *Distributed Autonomous Robotic Systems*, vol. 9, N. Correll, M. Schwager, and M. Otte, Eds. Cham: Springer International Publishing, 2019, pp. 225–238. doi: 10.1007/978-3-030-05816-6_16.

[18] R. Spica, D. Falanga, E. Cristofalo, E. Montijano, D. Scaramuzza, and M. Schwager, ‘A Real-Time Game Theoretic Planner for Autonomous Two-Player Drone Racing’, *ArXiv180102302 Cs*, Jan. 2018, Accessed: Apr. 29, 2020. [Online]. Available: http://arxiv.org/abs/1801.02302

[19] C. K. W. de Dreu and P. A. M. van Lange, ‘The Impact of Social Value Orientations on Negotiator Cognition and Behavior’, *Pers. Soc. Psychol. Bull.*, vol. 21, no. 11, pp. 1178–1188, Nov. 1995, doi: 10.1177/01461672952111006.

[20] R. O. Murphy and K. A. Ackermann, ‘Social Value Orientation: Theoretical and Measurement Issues in the Study of Social Preferences’, *Personal. Soc. Psychol. Rev.*, vol. 18, no. 1, pp. 13–41, Feb. 2014, doi: 10.1177/1088868313501745.

[21] W. Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, and D. Rus, ‘Social behavior for autonomous vehicles’, *Proc. Natl. Acad. Sci.*, p. 201820676, Nov. 2019, doi: 10.1073/pnas.1820676116.

[22] M. Herman, V. Fischer, T. Gindele, and W. Burgard, ‘Inverse reinforcement learning of behavioral models for online-adapting navigation strategies’, in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 3215–3222. doi: 10.1109/ICRA.2015.7139642.

[23] C. Hubmann, M. Becker, D. Althoff, D. Lenz, and C. Stiller, ‘Decision making for autonomous driving considering interaction and uncertain prediction of surrounding vehicles’, in *2017 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2017, pp. 1671–1678. doi: 10.1109/IVS.2017.7995949.

[24] W. Liu, S. Kim, S. Pendleton, and M. H. Ang, ‘Situation-aware decision making for autonomous driving on urban road using online POMDP’, in *2015 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2015, pp. 1126–1133. doi: 10.1109/IVS.2015.7225835.

[25] S. Ulbrich and M. Maurer, ‘Towards Tactical Lane Change Behavior Planning for Automated Vehicles’, in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Gran Canaria, Spain, Sep. 2015, pp. 989–995. doi: 10.1109/ITSC.2015.165.

[26] S. Shalev-Shwartz, S. Shammah, and A. Shashua, ‘Safe, Multi-Agent, Reinforcement Learning for Autonomous Driving’, *ArXiv161003295 Cs Stat*, Oct. 2016, Accessed: Nov. 13, 2020. [Online]. Available: http://arxiv.org/abs/1610.03295

[27] J. Li, F. Yang, M. Tomizuka, and C. Choi, ‘EvolveGraph: Multi-Agent Trajectory Prediction with Dynamic Relational Reasoning’, *ArXiv200313924 Cs*, Oct. 2020, Accessed: Nov. 03, 2020. [Online]. Available: http://arxiv.org/abs/2003.13924

[28] W. Schwarting, J. Alonso-Mora, and D. Rus, ‘Planning and Decision-Making for Autonomous Vehicles’, *Annu. Rev. Control Robot. Auton. Syst.*, vol. 1, no. 1, pp. 187–210, 2018, doi: 10.1146/annurev-control-060117-105157.