

# Planning Persuasive Trajectories Based on a Leader-Follower Game Model

Chaozhe R. He<sup>1</sup>, Yichen Dong<sup>2</sup>, Nan Li<sup>2</sup>

**Abstract:** We propose a framework that enables autonomous vehicles (AVs) to proactively shape the intentions and behaviors of interacting human drivers. The framework employs a leader-follower game model with an adaptive role mechanism to predict human interaction intentions and behaviors. It then utilizes a branch model predictive control (MPC) algorithm to plan the AV trajectory, persuading the human to adopt the desired intention. The proposed framework is demonstrated in an intersection scenario. Simulation results illustrate the effectiveness of the framework for generating persuasive AV trajectories despite uncertainties.

*Keywords:* Human-Machine and Human-Robot Systems; Path Planning and Motion Control; Robotics

## 1. INTRODUCTION

Autonomous vehicles (AVs) have the potential to transform the landscape of ground transportation by enhancing safety, increasing efficiency, and expanding accessibility. However, successfully integrating AVs into existing road systems requires them to handle complex interactions with other road users, particularly human-driven vehicles (HVs) Schwarting et al. (2019); Wang et al. (2022). A key source of complexity stems from the variability in human driver intentions and corresponding behaviors (e.g., proceeding or yielding). Various factors, including personal preferences, emotional states, and the driving context, influence driver intentions Xing et al. (2019). In traffic scenarios that involve extensive interactions (e.g., negotiation, competition, and cooperation), such as at unsignalized intersections, highway on/off-ramps, or in congested traffic, AVs may not only respond to human driver behaviors but also proactively shape their intentions to promote cooperation. This paper aims to develop an autonomous driving algorithm that can safely respond and further effectively shape the intention and behavior of human drivers, thereby improving AVs' ability to navigate complex interaction scenarios.

*Related Work* Game theory provides a mathematical framework for studying the interactions of intelligent agents. Various researchers have utilized it in developing algorithms for modeling interaction traffic scenarios and predicting vehicle behaviors in these scenarios, e.g., in Bahram et al. (2015); He and Li (2024); Li et al. (2017, 2020); Liu et al. (2022); Tian et al. (2020); Yu et al. (2018); Zhang et al. (2019).

To account for variable human driver intentions, Zhang et al. (2019) incorporates the driver's aggressiveness as a model parameter and estimates it online based on observed vehicle behavior; Li et al. (2017) and Tian et al. (2020) model human drivers of different types based on cognitive levels and different driver types can have different intentions under the same/similar traffic condition; Lefkopoulous et al. (2020) uses multiple models to predict vehicle behaviors under different driver intentions and uses a hybrid Kalman filtering scheme to identify the most likely model.

In Li et al. (2020), we presented a leader-follower game model (LFG, which is extended from the classical Stackelberg game model) for modeling the interactions of multiple vehicles. In this model, different driver intentions in an interaction traffic scenario, particularly the intention to proceed ahead or yield to other vehicles, are represented by the asymmetric leader and follower roles of the game players. In Liu et al. (2022) and He and Li (2024), we utilized this leader-follower game model to develop algorithms for AVs to navigate traffic scenarios involving extensive interactions with HVs safely. LFG Provides lightweight conflict resolving mechanism when interacting with Humans.

However, the above approaches can, at most, estimate the driver's intentions based on observed past behaviors of the interacting vehicles and then optimally respond to the estimated intentions – none of these approaches can proactively shape the intentions of human drivers. In light of the benefits of promoting human-robot cooperation, methods for enabling robots to proactively shape human intentions and behaviors were proposed in Parekh et al. (2022) and Pandya et al. (2024). However, human-robot interactions are not represented by a game model in the methods of Parekh et al. (2022) and Pandya et al. (2024), and the proposed methods may not be readily applicable to autonomous driving applications.

*Main Contributions* In this work, we propose a novel framework for autonomous agents to proactively influence the intentions and behaviors of interacting humans using game theory. The framework is particularly well-suited for autonomous driving applications, promoting AV-HV cooperation and enhancing both safety and efficiency.

- We extend the leader-follower game model of He and Li (2024); Li et al. (2020); Liu et al. (2022) to one with a role-adapting mechanism and utilize it for predicting human interactive intention and behavior.
- We design a branch MPC algorithm that can predict HV role changes in response to AV behaviors and utilize it for planning persuasive AV trajectories.
- We demonstrate, through a simulated intersection scenario, that the proposed planning framework can generate proactive AV behavior that influences human driver intention without compromising safety.

The remainder of this paper is organized as follows. First, we formulate the problem in Section 2. We present the leader-follower game model with an adaptive role in Section 2 and the interactive planning algorithm in Section 4. A case study on autonomous driving at the intersection is given in Section 5 to showcase their effectiveness. We conclude the paper 6 and discuss future directions.

<sup>1</sup> C. R. He is with the Department of Mechanical and Aerospace Engineering, University at Buffalo, Buffalo, NY, US. [chaozheh@buffalo.edu](mailto:chaozheh@buffalo.edu)

<sup>2</sup> Y. Dong and N. Li are with the School of Automotive Studies, Tongji University, Shanghai, China. [dych.happy@163.com](mailto:dych.happy@163.com), [li\\_nan@tongji.edu.cn](mailto:li_nan@tongji.edu.cn)

## 2. PROBLEM STATEMENT

In this section, we present the AV-HV interaction problem of interest as a general human-robot interaction problem. In particular, we consider the following 2-player human-robot interaction game:

$$x_{t+1} = f(x_t, u_t^H, u_t^R) \quad (1)$$

where  $x = (x^H, x^R)$ , the “H” stands for “human” and the “R” stands for “robot.” We assume the robot has a reward function that encodes its goal:

$$R(x, u^R, u^H) \quad (2)$$

which depends on the states and actions of both players. With this reward function, the robot adopts an MPC-based control strategy: At each time  $t$ , the robot computes an optimal action sequence  $\mathbf{u}_t^{R,*} = \{u_t^{R,*}, u_{t+1}^{R,*}, \dots, u_{t+N-1}^{R,*}\}$  that maximizes the cumulative reward over the planning horizon, i.e.,

$$\begin{aligned} \mathbf{u}_t^{R,*} \in \arg \max_{\mathbf{u}_t^R} \sum_{\tau=0}^{N-1} \lambda^\tau R(x_{t+\tau+1}, u_{t+\tau}^R, \hat{u}_{t+\tau}^H) \\ \text{s.t. } x_{t+\tau+1} = f(x_{t+\tau}, \hat{u}_{t+\tau}^H, u_{t+\tau}^R) \\ x_{t+\tau} \in \mathcal{X}_{\text{safe}}, \quad \forall \tau = 1, \dots, N \end{aligned} \quad (3)$$

where  $\lambda \in (0, 1]$  is a discount factor, and  $\mathcal{X}_{\text{safe}}$  represents a safe state set used to represent strict safety specifications.

A decision process as above requires prediction of human actions, i.e., the  $\hat{u}_{t+\tau}^H$  in (3), where “ $\wedge$ ” is used to distinguish predicted actions from the human player’s actual actions. Traditional methods typically utilize a “predict then plan” strategy, where  $\hat{u}_{t+\tau}^H$  are first determined and remain fixed as the robot plans the actions  $u_{t+\tau}^R$ . Such a strategy is lightweight in computation but misses the action-reaction interactive nature of the game during planning. To develop a computationally tractable algorithm that enables interactive planning, in particular, proactively shaping predicted human actions  $\hat{u}_{t+\tau}^H$  during planning  $u_{t+\tau}^R$ , we leverage a leader-follower game model with a role-adapting mechanism, introduced in the next section.

## 3. ADAPTIVE LEADER-FOLLOWER GAME

In this section, we first briefly review the leader-follower game (LFG) model of He and Li (2024); Li et al. (2020); Liu et al. (2022) and then present the proposed role-adapting mechanism. The extended LFG model with the role-adapting mechanism will be utilized to predict human interactive intentions and behaviors.

### 3.1 Leader-Follower Game

Consider a pair of players, one is a “leader” and the other is a “follower”, and rewrite the cumulative reward in (3) as

$$\bar{R}_\sigma(x_t, \mathbf{u}_l, \mathbf{u}_f, t) = \sum_{\tau=0}^{N-1} \lambda^\tau R_\sigma(x_{t+\tau+1}, u_{l,t+\tau}, u_{f,t+\tau}) \quad (4)$$

where the subscript  $\sigma \in \{\text{leader(l)}, \text{follower(f)}\}$  indicates the role of the player in the game;  $\mathbf{u}_{l,t} = \{u_{l,t+\tau}\}_{\tau=0}^{N-1} \in \mathbf{U}_l$  and  $\mathbf{u}_{f,t} = \{u_{f,t+\tau}\}_{\tau=0}^{N-1} \in \mathbf{U}_f$  denote the action sequences of the leader and the follower;  $\mathbf{U}_l$  and  $\mathbf{U}_f$  are the admissible sets of action sequences of both roles; and  $R_\sigma(\cdot, \cdot, \cdot)$  is a single-step reward similar to that defined in (2), where the subscript  $\sigma$  implies that the reward functions for the two different roles can be different.

The leader and the follower both aim to maximize their cumulative rewards but follow different strategies: The follower maximizes the worst-case reward due to uncertain

leader’s actions, i.e., it takes the following “max-min” strategy:

$$\mathbf{u}_{f,t}^* \in \arg \max_{\mathbf{u}_f \in \mathbf{U}_f} Q_f(x_t, \mathbf{u}_f) \quad (5a)$$

where

$$Q_f(x_t, \mathbf{u}_f) = \min_{\mathbf{u}_l \in \mathbf{U}_l} \bar{R}_f(x_t, \mathbf{u}_l, \mathbf{u}_f). \quad (5b)$$

This strategy represents an intention to yield to the leader Li et al. (2020). In contrast, the leader assumes that the other player is a follower who uses the above “max-min” strategy. Hence, the leader can predict the follower’s actions and optimize its actions according to:

$$\mathbf{u}_{l,t}^* \in \arg \max_{\mathbf{u}_l \in \mathbf{U}_l} Q_l(x_t, \mathbf{u}_l) \quad (6a)$$

where

$$Q_l(x_t, \mathbf{u}_l) = \min_{\mathbf{u}_f \in \mathbf{U}_f^*(x_t)} \bar{R}_l(x_t, \mathbf{u}_l, \mathbf{u}_f) \quad (6b)$$

$$\mathbf{U}_f^*(x_t) = \{\mathbf{u}_f \in \mathbf{U}_f : Q_f(x_t, \mathbf{u}_f) \geq Q_f(x_t, \mathbf{u}_f'), \forall \mathbf{u}_f' \in \mathbf{U}_f\}. \quad (6c)$$

This strategy represents a player that assumes the other player will yield and hence decides to proceed ahead Li et al. (2020). It is shown in He and Li (2024); Li et al. (2020); Liu et al. (2022) that the above LFG model can accurately predict vehicle interactive behaviors in various traffic scenarios, including intersections Li et al. (2020) and highways He and Li (2024); Liu et al. (2022). We note that in the considered human-robot interaction game, either player,  $\bullet \in \{H, R\}$ , can be a leader or follower, and they may not hold complementary roles (i.e., they may both be leaders or followers). For more comprehensive discussions on the above LFG model, please refer to Li et al. (2020) or Liu et al. (2022).

### 3.2 Leader-Follower Game with Adaptive Roles

In the LFG model, the role  $\sigma \in \{l, f\}$  represents the player’s intention and is a key parameter determining the player’s behavior. It is assumed to be constant over the interaction period in Li et al. (2020) and Liu et al. (2022). In this work, we relax this assumption and allow  $\sigma$  to change over time, motivated by the observation that a human driver can dynamically adjust their intention in response to changing traffic conditions Tian et al. (2021).

Let  $\bullet$  denote the ego player and  $\circ$  denote the other player. We assume that the ego player holds a belief on the constant role of the other player,  $\sigma^\circ$ , and updates the belief based on Bayesian inference as follows Hwang et al. (2006):

$$\mathbb{P}(\sigma^\circ = \varepsilon | \xi_t) \propto \mathbb{P}(x_t | \sigma^\circ = \varepsilon) \mathbb{P}(\sigma^\circ = \varepsilon | \xi_{t-1}) \quad (7)$$

where  $\propto$  indicates “proportional to”;  $\xi_t$  denotes the collection of observed information up to time  $t$ , i.e.,

$$\xi_t = \{x_0, \dots, x_{t-1}, x_t, u_0^\bullet, \dots, u_{t-1}^\bullet\} \quad (8)$$

where  $x_0, \dots, x_t$  are observed game states and  $u_0^\bullet, \dots, u_{t-1}^\bullet$  are the ego player’s own actions taken at previous times; and the “likelihood function”  $\mathbb{P}(x_t | \sigma^\circ = \varepsilon)$  is given as

$$\mathbb{P}(x_t | \sigma^\circ = \varepsilon) = \mathcal{N}(r_t | \sigma^\circ = \varepsilon, 0, \mathcal{W}). \quad (9)$$

In (9),  $\mathcal{N}(\cdot, 0, \mathcal{W})$  denotes the probability density function of the multivariate normal distribution with zero mean and covariance  $\mathcal{W}$  evaluated at  $(\cdot)$ ;  $r_t | \sigma^\circ = \varepsilon$  is the residual between observed state and predicted state assuming the other player’s role is  $\varepsilon$ , i.e.,

$$r_t | \sigma^\circ = \varepsilon = x_t - \hat{x}_t | \sigma^\circ = \varepsilon. \quad (10)$$

The predicted state  $\hat{x}_t | \sigma^\circ = \varepsilon$  is given by

$$\hat{x}_t | \sigma^\circ = \varepsilon = f(x_{t-1}, u_{t-1}^\bullet, \hat{u}_{t-1}^\circ | \sigma^\circ = \varepsilon) \quad (11)$$

where  $f(\cdot, \cdot, \cdot)$  is the dynamics equation in (1);  $x_{t-1} = (x^\bullet, x^\circ)$  is the observed game state at the previous time;

$u_{t-1}^\bullet$  is the ego player's previous action; and  $\hat{u}_{t-1}^\circ | \sigma^\circ = \varepsilon$  is the predicted previous action of the other player assuming its role to be  $\sigma^\circ = \varepsilon$ , which is determined by the LFG model in Section 3.1.

With a belief on the other player's role, we assume that the ego player adjusts its role probabilistically. In particular, we define a transition matrix to describe the probabilistic transition of the ego player's role:

$$\begin{bmatrix} \pi_{ll,t}^\bullet & 1 - \pi_{ff,t}^\bullet \\ 1 - \pi_{ll,t}^\bullet & \pi_{ff,t}^\bullet \end{bmatrix} = \Pi_t^\bullet = \Pi_t^\bullet(x_t, \mathbb{P}^\mathbb{P}(\sigma_t^\bullet)), \quad (12)$$

where  $\pi_{ll,t}^\bullet$  (resp.  $\pi_{ff,t}^\bullet$ ) represents the probability of transitioning to role l (resp. f) when the ego player's previous role is l (resp. f);  $\mathbb{P}^\mathbb{P}(\sigma_t^\bullet)$  represents a "plausible" distribution of the ego player's role, determined by the ego player's belief on the other player's role, i.e.,

$$\mathbb{P}^\mathbb{P}(\sigma_t^\bullet) = \Phi(\mathbb{P}(\sigma^\circ | \xi_t)). \quad (13)$$

The relation  $\Phi$  between the belief on the other player's role  $\mathbb{P}(\sigma^\circ | \xi_t)$  and the "plausible" distribution of the ego player's role  $\mathbb{P}^\mathbb{P}(\sigma_t^\bullet)$  can be designed. A reasonable design is

$$\begin{bmatrix} \mathbb{P}^\mathbb{P}(\sigma_t^\bullet = l) \\ \mathbb{P}^\mathbb{P}(\sigma_t^\bullet = f) \end{bmatrix} = \begin{bmatrix} \mathbb{P}(\sigma^\circ = f | \xi_t) \\ \mathbb{P}(\sigma^\circ = l | \xi_t) \end{bmatrix} \quad (14)$$

where ego tends to take a complementary role to the other player. We remark that alternative designs of  $\Phi$  accounting for other considerations are also possible. Meanwhile, in (12),  $\Pi_t^\bullet(x_t, \mathbb{P}^\mathbb{P}(\sigma_t^\bullet))$  indicates that the transition probabilities can in general depend on the game state  $x_t$  and the "plausible" distribution  $\mathbb{P}^\mathbb{P}(\sigma_t^\bullet)$ . For instance, when traffic state  $x_t$  implies no conflict between two vehicles, the ego driver may ignore the "plausible" distribution  $\mathbb{P}^\mathbb{P}(\sigma_t^\bullet)$  when determining his role.

#### 4. INTERACTIVE PREDICTION AND PLANNING BASED ON LEADER-FOLLOWER GAME WITH ADAPTIVE ROLES

In this section, we present the interactive planning algorithm for the robot, assuming the human is making decisions based on an LFG model with adaptive roles. We first describe the method for getting interactive predictions, then present the planning algorithm.

##### 4.1 Interactive Prediction

The robot assumes that the human behaves as an LFG model with role adaptation. Because the role of the human is not directly observable to the robot, the robot holds a belief on the role of the human and updates it based on Bayesian inference similarly to (7).

Denote the belief at time  $t$  as  $\mathbb{P}(\sigma_t^H)$ , then a corresponding open-loop prediction of human's behavior can be given by

$$\hat{x}_{t+\tau}^H = x_{\sigma_t^H, t+\tau}^H, \quad \tau = 0, \dots, T_1, \quad \sigma_t^H \sim \mathbb{P}(\sigma_t^H) \quad (15)$$

where it is assumed that the human will not change his role for  $[t, t + T_1]$ . At  $t + T_1$ , the human updates his belief on the robot's role  $\mathbb{P}(\sigma_{t+T_1}^R)$ , based on the game states resulting from the actions planned by the robot between  $[t, t + T_1]$ , determines a plausible distribution of his role  $\mathbb{P}^\mathbb{P}(\sigma_{t+T_1}^H)$ , and adjusts his role probabilistically, all using the methods in Section 3.2. Specifically, an adjusted role of the human is given as

$$\sigma_{t+T_1}^H \sim \mathbb{P}(\sigma_{t+T_1}^H) = \hat{\Pi}(\hat{x}_{t+T_1}, \mathbb{P}^\mathbb{P}(\sigma_{t+T_1}^H)) \mathbb{P}(\sigma_t^H). \quad (16)$$

Note that the robot may not know the actual transition matrix of the human; thus, the " $\wedge$ " in (16) indicates that the transition matrix used by the robot to predict the human role adjustment may be different from the human's actual transition matrix (12).

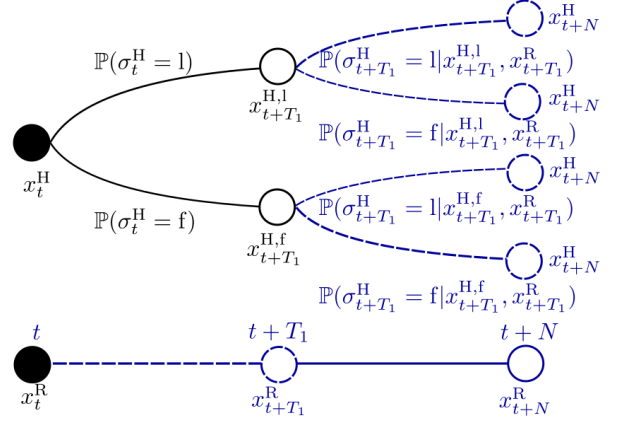


Fig. 1. Closed-loop MPC-based strategy. The top branch corresponds to the prediction of human's behavior, while the lower branch corresponds to the robot's planned motion. The black solid branches in the human prediction are affected by the state at  $t$ . The blue dashed branches are affected by the robot's behavior between  $[t, t + T_1]$ , which is also marked as a blue dashed branch, through the change of role  $\sigma_{t+T_1}^H$ .

With the adjusted role of the human,  $\sigma_{t+T_1}^H$ , a corresponding prediction of human's behavior over the remaining horizon  $[t + T_1, t + N]$  can be given as

$$\hat{x}_{t+\tau}^H = x_{\sigma_{t+T_1}^H, t+\tau}^H, \quad \tau = T_1 + 1, \dots, N, \quad \sigma_{t+T_1}^H \sim \mathbb{P}(\sigma_{t+T_1}^H). \quad (17)$$

The predicted human actions over the horizon are given as

$$\hat{u}_{t+\tau}^H = \begin{cases} \hat{u}_{\sigma_t^H, t+\tau}^H, & \tau = 0, \dots, T_1 - 1, \\ \hat{u}_{\sigma_{t+T_1}^H, t+\tau}^H, & \tau = T_1, \dots, N - 1. \end{cases} \quad (18)$$

We note that due to the dependence on the human's adjusted role and hence his future behavior on the planned actions by the robot as shown in (16), the predictions are action-reaction interactive and closed-loop. That is, the predictions change as the robot's planned trajectory changes; see the interactive prediction scheme in Fig. 1.

##### 4.2 Interactive and Closed-Loop Planning

With interactive predictions of the human behavior, the robot's MPC policy (3) is updated to have two stages with four branches, as shown in Fig. 1:

$$\begin{aligned} \mathbf{u}_t^{R,*} &\in \arg \max \\ &\left\{ \mathbb{E}_{|\Upsilon|} \left[ \sum_{\tau=0}^{T_1} \lambda^\tau R(x_{t+\tau+1}, u_{t+\tau}^R, \hat{u}_{\sigma_t^H, t+\tau}^H) \middle| \xi_t \right] + \right. \\ &\quad \left. \mathbb{E}_{|\Upsilon|^2} \left[ \sum_{\tau=T_1+1}^N \lambda^\tau R(x_{t+\tau+1}, u_{t+\tau}^R, \hat{u}_{\sigma_{t+T_1}^H, t+\tau}^H) \middle| \xi_t \right] \right\} \\ \text{s.t.} \quad &x_{t+\tau+1} = f(x_{t+\tau}, u_{t+\tau}^R, u_{t+\tau}^H), \\ &\mathbb{P}(x_{t+\tau} \in \mathcal{X}_{\text{safe}}, \tau = 1, \dots, N | \xi_t) \geq 1 - \epsilon. \end{aligned} \quad (19)$$

The expectations in (19) are with respect to the probabilities  $\mathbb{P}(\sigma_t^H)$  and  $\mathbb{P}(\sigma_{t+T_1}^H)$ . The  $|\Upsilon|$  denotes the number of roles. For each MPC decision made by the robot on the trajectories, there will be two possible futures, and the chance constraints will be evaluated uniformly for all states along the trajectory.

Note that in the robot's prediction step, we assume that the initial role estimation by the human on the robot at  $t$  is  $[0.5, 0.5]$ . In practice, the human could also use a Bayesian filter. This means that, in contrast to the robot generating predictions on the human's role continuously, the robot's

perception of the human's estimation on the robot's role is not carried over between executions of (19).

## 5. CASE STUDY: AUTONOMOUS DRIVING IN INTERACTIVE TRAFFIC SCENARIOS

In this section, we apply the proposed closed-loop interactive planning to autonomous vehicles in the intersection where they need to interact with human-driven vehicles<sup>3</sup>.

### 5.1 Scenario Setup

*Simulation Setup and Goals* The top view of the simulation scenario is shown in Fig. 3, where the autonomous vehicle (AV, red) is traveling eastbound, while the human-driven vehicle (HV, blue) is traveling northbound in a 2-way intersection. The merging point is at (0, 0) [m], and the white solid crossing line 6.5 [m] from the merging point indicates the entrance to the intersection. Whoever reaches this crossing line first is considered to arrive first and will have the right of way at the intersection. Both the AV and HV aim to travel safely through the intersection in the fastest possible way. To achieve this, the AV could pass through the intersection faster than the HV by persuading the human driver to give it the right of way.

We use the following discrete-time model to represent vehicle kinematics in a 2-way 1-lane intersection; c.f., (1).

$$\begin{bmatrix} s_{t+1}^{\bullet,k} \\ v_{t+1}^{\bullet,k} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_t^{\bullet,k} \\ v_t^{\bullet,k} \end{bmatrix} + \begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \end{bmatrix} a_t^{\bullet,k}, \quad (20)$$

where  $s, v, a$  denotes position, velocity, and acceleration, respectively. The subscript  $t$  represents the discrete time, the first superscript  $\bullet \in \{H, R\}$  distinguishes the two players in the game, while the second superscript  $k \in \{x, y\}$  denotes the  $x$  or  $y$ -direction,  $\Delta t$  is the sampling period. Since it is 1-lane, no lane change action is needed.

*Safety Constraints* To avoid collisions in planning, the safety requirements require staying within the safe set

$$\mathcal{X}_{\text{safe}} := \{(s^1, s^2) \mid \|s^1 - s^2\|_2 \geq 1.5l_{\text{veh}}\}, \quad (21)$$

where  $s^\bullet = [s^{\bullet,x}, s^{\bullet,y}]^\top$ ,  $\|\cdot\|_2$  represents Euclidean norm, and  $l_{\text{veh}} = 5$  [m] is the vehicle length. Given the uncertain nature of human prediction, (21) is enforced as the chance constraint over the planning horizon in (19), with  $\epsilon = 0.02$

$$\mathbb{P} \left\{ (s_{\tau|t}^1, s_{\tau|t}^2) \in \mathcal{X}_{\text{safe}}, \forall \tau = 1, \dots, N \mid \xi_t \right\} \geq 1 - \epsilon. \quad (22)$$

*Reward* The reward function is set to the following linear combination of terms, corresponding to liveliness (goal achievement), control effort, and collision He and Li (2024). When the collision is inevitable, to avoid that vehicle establishing dangerous behavior by traveling at maximum speed, the collision penalty terms have speed-dependent coefficients Li et al. (2020).

$$\begin{aligned} c_t &= (1 + v_t)r_{\text{collision}}, \\ R_t &= s_t^{\bullet,k} - s_0^{\bullet,k} - w_1 c_t - w_2 |u_{t-1}|, \end{aligned} \quad (23)$$

where  $(\bullet, k) = (R, x)$  for the AV and  $(\bullet, k) = (H, y)$  for the HV. Horizon  $N$  is set to 5 seconds to allow reaching and passing the intersection within the planning horizon from the beginning. Note that the indicator  $r_{\text{collision}} = 1$  when a collision happens at  $t$ , and it is defined as the safety constraints (21) being violated, i.e.,  $s^\bullet \notin \mathcal{X}_{\text{safe}}$ . Note that (19) also uses reward (23) to deal with infeasible cases.

<sup>3</sup> Source codes and parameters: [https://github.com/CHELabUB/game\\_theoretic\\_model/tree/master/2025\\_mecc](https://github.com/CHELabUB/game_theoretic_model/tree/master/2025_mecc)

*Action Space* We carefully choose a set of trajectories as the action space by setting a list of target speeds and using an optimal velocity model to generate smooth and human-like trajectories that achieve different target speeds at designated spots (e.g., stop at the crossing line); see He and Li (2024) for details. Given the acceleration range  $a_{\text{max}}$  and  $a_{\text{min}}$ , the target speed over the horizon is selected to be  $n_{\text{mesh}}$  value between the maximum (capped at speed limit  $v_{\text{max}}$ ) and minimum reachable speed (capped at stop). Note that the trajectory associated with the minimum reachable speed as the target speed may be subject to changes for efficiency considerations: if the minimum reachable speed is 0, and it is possible to stop at the intersection, then the trajectory that leads to a stop at the intersection is used. One additional trajectory that keeps the current speed is also added to  $\mathbf{U}_\bullet$ , making the total action trajectory of  $|\mathbf{U}_\bullet| = n_{\text{mesh}} + 1$ . In this case study, we set  $n_{\text{mesh}} = 10$ . In both the LFG model (5, 6) and the MPC-based model (19), the trajectories are selected from this set of action trajectories, that is  $\mathbf{u} \in \mathbf{U}_\bullet$ .

*Role Transition Probability Matrix* We assume the HV uses an LFG model with the plausible role distribution  $\mathbb{P}(\sigma_t^H)$  given by maximum likelihood estimation

$$\sigma_t^H = \begin{cases} 1 & \text{if } \mathbb{P}(\sigma_t^R = 1) < \mathbb{P}(\sigma_t^R = f), \\ f & \text{if } \mathbb{P}(\sigma_t^R = 1) > \mathbb{P}(\sigma_t^R = f), \\ \sigma_t & \text{if } \mathbb{P}(\sigma_t^R = 1) = \mathbb{P}(\sigma_t^R = f). \end{cases} \quad (24)$$

Note that the last condition is a tie-breaker: the agent will keep the current role unchanged when the two probability values are identical. Based on  $\mathbb{P}(\sigma_t^H)$ , the role transition probability matrix is given by

$$\Pi = \mathbb{P}(\sigma_t^H = 1) \begin{bmatrix} 1 & p_a \\ 0 & 1 - p_a \end{bmatrix} + \mathbb{P}(\sigma_t^H = f) \begin{bmatrix} 1 - p_a & 0 \\ p_a & 1 \end{bmatrix} \quad (25)$$

c.f., (12).  $p_a \in [0, 1]$  captures the likelihood of the human to alter its role to the plausible role if different, which also capture how easy the human may be influenced. Note that with (25), human will keep the role if it is the same as the plausible role. We remark that the main parameter  $p_a$  can be different from those used in (19), which is denoted as  $\hat{p}_a$ . This corresponds to the fact that  $\hat{\Pi}$  in (16) can be different from  $\Pi$  in (12), i.e., HV's  $\Pi$  is unknown to AV.

### 5.2 Results

*2-LFG Case* We first simulate the case when both AV and HV use the LFG model. The initial roles of HV and AV are complementary, and their estimation of the other vehicle starts from neutral. For both LFG models, we set  $p_a = 1$ . The initial speed and position for the AV are fixed at  $-20$  [m] and  $4$  [m/s], while those for the human-driven vehicle are centered around  $-20$  [m] and  $4$  [m/s] with uniform distribution of  $\pm 1$  [m/s] and  $\pm 5$  [m].

Initial role by HV	AV [%]	HV [%]
Leader	0.0	100.0
Follower	97.9	2.1

Table 1. Percentage of arrival first at the crossing line based on 1000 runs with the random initial condition by the HV holding leader role vs follower role.

We ran 1000 runs, and the results are summarized in Table 1. No collision or safety constraint violation occurs. If the HV starts with a follower role, it will finish after the AV most of the time. By contrast, if it starts with a leader role, it will finish before the AV all the time. The bias arises from the fact that the initial condition is randomized for HV, whereas AV always starts from the same place at the same speed. One example of such runs is shown in the left column of Fig. 2 for the time profile and Fig. 3 for the top view lapse. The role estimation by both players

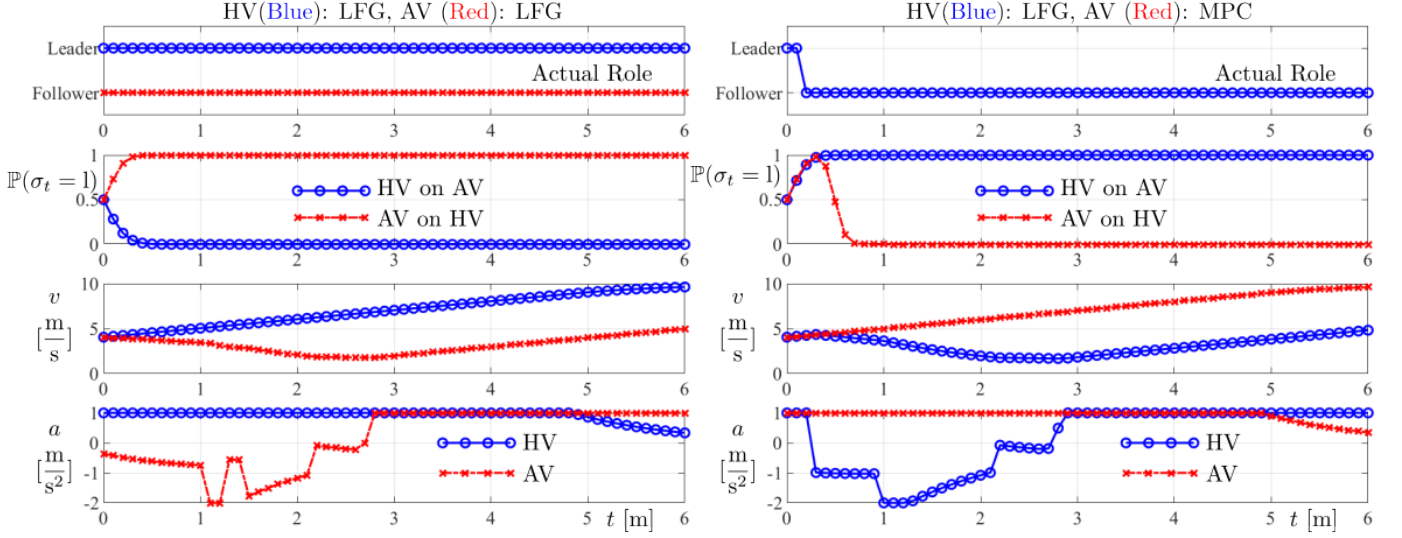


Fig. 2. Time profile comparison plots. The left column corresponds to the run when both human and AV are using the leader-follower game (LFG) to make decisions, while the right column corresponds to the run when human and AV are using LFG and MPC model (19) to make a decision. In all panels, the blue profile corresponds to a blue human-driven vehicle that is traveling northbound, while the red profile corresponds to the red autonomous vehicle (AV) that is traveling eastbound. The first row shows the actual role of those who use the LFG model. The second row shows the estimated probability of holding a leader role that one agent holds on the other agent. The third row shows the speed profile, and the fourth row shows the acceleration profiles.

on the other converges to the correct role quickly, and neither agent changes their role. In these cases, the initial conditions and role (effectively fixed role) determine the outcome. These results show that while LFG produces causal and meaningful behaviors, the AV is unable to influence the HV to make persuasive planning decisions.

Initial Role by HV, $p_a/\hat{p}_a$	AV [%]	HV[%]
Leader, 1.00/1.00	87.8	12.2
Follower, 1.00/1.00	96.4	3.6
Leader, 0.70/1.00	86.3	13.7
Leader, 0.50/1.00	82.9	17.1
Leader, 0.30/1.00	78.1	21.9
Follower, 0.50/1.00	96.5	3.5
Follower, 0.70/1.00	96.5	3.5
Follower, 0.30/1.00	96.6	3.4
Leader, 1.00/0.98	78.9	21.1
Leader, 1.00/0.95	0.0	100.0
Leader, 1.00/0.70	0.0	100.0
Follower, 1.00/0.98	96.4	3.6
Follower, 1.00/0.95	38.9	61.1
Follower, 1.00/0.70	1.2	98.8

Table 2. Percentage of arrival first at the crossing line based on 1000 runs with the random initial condition by the human-driven vehicle.  $p_a = \hat{p}_a$  when AV knows HV’s willingness to change the role;  $p_a > \hat{p}_a$  ( $p_a < \hat{p}_a$ ) when AV underestimates (overestimates) HV’s willingness to change the roles.

**LFG + MPC Cases** Next, we simulated the cases with the same setup except that the HV uses an LFG model while the AV uses the MPC algorithm (19). We started the case when the parameter  $p_a$  is known by the AV, that is  $p_a = \hat{p}_a = 1$ , the prediction by the AV on the HV is accurate. One example is shown in the right column of Fig. 2 for the time profile and Fig. 3 for the top view lapse. Despite the HV starting with a leader role, and the AV holding a neutral estimation of the HV’s role, the AV can persuade the HV to change roles, slow down, and allow the AV to cross the intersection first. The statistics for this case are summarized in the first two rows of Table 2. Compared to the statistics in Table 1, for the majority

of the cases, the AV can successfully persuade humans to give it priority.

To account for uncertainty, we then simulated cases where the parameter  $p_a$  varies at different levels unknown to AV, and report the results in Table 2. No collision or safety constraint violation occurs in any of the cases, demonstrating the safety of the design. When the AV overestimates the HV’s likelihood to adapt role ( $\hat{p}_a > p_a$ ), it can still persuade the HV to change its role and give it the priorities, more likely when the HV starts with a follower role than a leader role, and can tolerate large gap between  $\hat{p}_a$  and  $p_a$ ; See the second group in Table 2. When the AV underestimates the HV’s likelihood to adapt role ( $\hat{p}_a < p_a$ ), the AV with MPC algorithm would choose not to persuade HV and instead give HV more priority as the gap between  $\hat{p}_a$  and  $p_a$  increases; See the third group in Table 2. We acknowledge that the big swing from AV finish first mostly to HV finish first mostly between  $\hat{p}_a = 0.98$  and 0.95 is associated with the chance constraint level set  $\epsilon = 0.02$ ; cf. (22).

In summary, bearing uncertainties in the key parameters, the proposed MPC-based interactive planning algorithm (19) enables the AV to be persuasive, influence human behavior, and gain an advantage.

## 6. CONCLUSION

In this study, we presented a framework that enables autonomous vehicles to proactively shape the intentions and behaviors of interacting human drivers. The framework employed a leader-follower game model with an adaptive role mechanism to predict human interaction intention and behavior, and utilized a branch model predictive control algorithm to plan the AV trajectory that persuades the human to adopt a desired intention. We demonstrated the proposed framework in an intersection scenario, and simulation results illustrated its effectiveness in generating persuasive AV trajectories under uncertainties. Future directions include implementing and validating the proposed framework in more complex driving scenarios, with human-in-the-loop, and baseline against other non-persuasive motion planning algorithms.



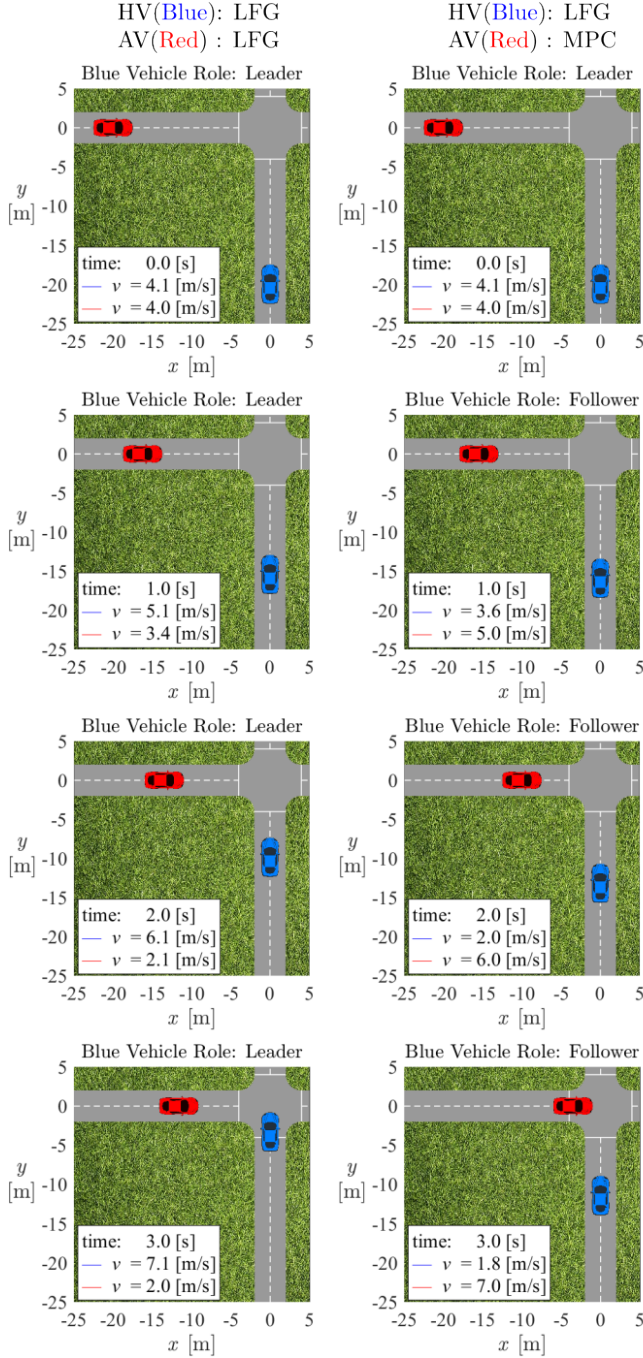


Fig. 3. Topview lapse comparison. The left column corresponds to the run when both human and AV are using the leader-follower game (LFG) model to make decisions, while the right column corresponds to the run when human and AV are using LFG and MPC model (19) to make a decision. In all frames, the blue vehicle and legend correspond to the HV, while the red vehicle corresponds to the AV, and their time profiles are shown as blue and red curves in Fig. 2.

## REFERENCES

Bahram, M., Lawitzky, A., Friedrichs, J., Aeberhard, M., and Wollherr, D. (2015). A game-theoretic approach to replanning-aware interactive scene prediction and planning. *IEEE Transactions on Vehicular Technology*, 65(6), 3981–3992.

He, C.R. and Li, N. (2024). Eco-driving accounting for interactive cut-in vehicles. In *International Conference on Mobility, Operations, Services and Technologies*

(MOST), 90–100. IEEE.

Hwang, I., Balakrishnan, H., and Tomlin, C. (2006). State estimation for hybrid systems: applications to aircraft tracking. *IEEE Proceedings-Control Theory and Applications*, 153(5), 556–566.

Lefkopoulou, V., Menner, M., Domahidi, A., and Zeilinger, M.N. (2020). Interaction-aware motion prediction for autonomous driving: A multiple model Kalman filtering scheme. *IEEE Robotics and Automation Letters*, 6(1), 80–87.

Li, N., Oyler, D.W., Zhang, M., Yildiz, Y., Kolmanovsky, I., and Girard, A.R. (2017). Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems. *IEEE Transactions on Control Systems Technology*, 26(5), 1782–1797.

Li, N., Yao, Y., Kolmanovsky, I., Atkins, E., and Girard, A.R. (2020). Game-theoretic modeling of multi-vehicle interactions at uncontrolled intersections. *IEEE Transactions on Intelligent Transportation Systems*, 23(2), 1428–1442.

Liu, K., Li, N., Tseng, H.E., Kolmanovsky, I., and Girard, A. (2022). Interaction-aware trajectory prediction and planning for autonomous vehicles in forced merge scenarios. *IEEE Transactions on Intelligent Transportation Systems*, 24(1), 474–488.

Pandya, R., Wang, Z., Nakahira, Y., and Liu, C. (2024). Towards proactive safe human-robot collaborations via data-efficient conditional behavior prediction. In *International Conference on Robotics and Automation (ICRA)*, 12956–12963. IEEE.

Parekh, S., Habibian, S., and Losey, D.P. (2022). RILI: Robustly influencing latent intent. In *International Conference on Intelligent Robots and Systems (IROS)*, 1–8. IEEE.

Schwartz, W., Pierson, A., Alonso-Mora, J., Karaman, S., and Rus, D. (2019). Social behavior for autonomous vehicles. *Proceedings of the National Academy of Sciences*, 116(50), 24972–24978.

Tian, R., Li, N., Kolmanovsky, I., Yildiz, Y., and Girard, A.R. (2020). Game-theoretic modeling of traffic in unsignalized intersection network for autonomous vehicle control verification and validation. *IEEE Transactions on Intelligent Transportation Systems*, 23(3), 2211–2226.

Tian, R., Tomizuka, M., and Sun, L. (2021). Learning human rewards by inferring their latent intelligence levels in multi-agent games: A theory-of-mind approach with application to driving data. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4560–4567. IEEE.

Wang, W., Wang, L., Zhang, C., Liu, C., Sun, L., et al. (2022). Social interactions for autonomous driving: A review and perspectives. *Foundations and Trends in Robotics*, 10(3-4), 198–376.

Xing, Y., Lv, C., Wang, H., Ai, Y., Cao, D., Velenis, E., and Wang, F.Y. (2019). Driver lane change intention inference for intelligent vehicles: Framework, survey, and challenges. *IEEE Transactions on Vehicular Technology*, 68(5), 4377–4390.

Yu, H., Tseng, H.E., and Langari, R. (2018). A human-like game theory-based controller for automatic lane changing. *Transportation Research Part C: Emerging Technologies*, 88, 140–158.

Zhang, Q., Langari, R., Tseng, H.E., Filev, D., Szwabowski, S., and Coskun, S. (2019). A game theoretic model predictive controller with aggressiveness estimation for mandatory lane change. *IEEE Transactions on Intelligent Vehicles*, 5(1), 75–89.