

# Game Theoretic Modeling of Vehicle Interactions at Unsignalized Intersections and Application to Autonomous Vehicle Control

Nan Li, Ilya Kolmanovsky, Anouck Girard, and Yildiray Yildiz

**Abstract**—In this paper, we discuss a game theoretic approach to model the time-extended, multi-step, and interactive decision making of vehicles at unsignalized intersections. The vehicle interaction model is then used to define an autonomous vehicle controller. Simulation results for a common intersection scenario are reported.

## I. INTRODUCTION

Some of the most complicated scenarios faced by a driver or an autonomous vehicle in urban traffic involve intersections, where vehicles interact with each other as well as with other traffic participants such as pedestrians. At an intersection, vehicles make consecutive maneuvers, to achieve their individual objectives, e.g., reaching their desired lanes via crossing the intersection or via making a turn, while avoiding accidents. In the decision making regarding which maneuver to take next, each vehicle should take into account the interactions with the others.

For unsignalized intersections, it is particularly important to account for such interactions. Due to the lack of guidance from traffic lights (that may be viewed as centralized traffic controllers), the drivers/vehicles need to decide on their own whether, when, and how they should enter and cross the intersection. A failure to understand the interactions may lead to deadlocks – no one decides to enter the intersection or everyone gets stuck in the middle of the intersection. In more extreme cases, collisions may occur.

The vehicle-to-vehicle or vehicle-to-pedestrian interactions may be analyzed and accounted for using reachability based approaches, e.g., in [1], [2]. However, such approaches usually address worst-case scenarios and may lead to conservative results. To reduce conservatism, in [2] the desired path of each traffic participant is assumed to be known, and uncertain interactions are described by probabilistic deviations from the desired path. However, such approaches do not address the high-level decision making problem to create the desired path. Furthermore, assuming accurate knowledge of the desired paths of other traffic participants may not be realistic – one may only be able to predict the desired paths of the others via his/her own reasoning, while the reasoning and the prediction may be imperfect.

Alternatively, game theory may be used to study strategic reasoning in multi-agent systems, such as the interactive



Fig. 1: An unsignalized intersection in Kansas (image provided by Google street view). Unsignalized intersections can usually be found in residential areas in the US [3].

decision making of multiple traffic participants in a traffic scenario. Game theoretic approaches take into account the rationality of intelligent agents in modeling their interactions, which may lead to less conservative results compared to reachability-based analyses that account for worst-case and unlikely-to-realize scenarios. In [4], the vehicle-to-vehicle interaction at an intersection is modeled as an one-shot normal-form game – each vehicle selects an action between “Stop” and “Go” based on a payoff matrix, and without taking into account vehicle dynamics. Game theoretic modeling of vehicle-to-vehicle interactions in highway driving scenarios are considered in [5], [6] and [7]. In [7], a simple intersection scenario for two one-lane roads and two vehicles whose objectives are to cross the intersection by going straight is considered, which is different from the unsignalized intersection scenario considered in this paper, where there are multiple lanes and the vehicles can also turn.

Driving a vehicle crossing an unsignalized intersection is a rather complicated process – the driver may need to adjust his/her maneuver decisions rapidly according to the latest situation he/she is facing. Hence, crossing an intersection may be formulated as a time-extended, multi-step decision making process wherein driver-to-driver/vehicle-to-vehicle interactions occur continually over the entire process.

The contributions of this paper are: 1) we present a game theoretic approach to model the time-extended, multi-step, and interactive decision making processes of human-driven vehicles at unsignalized intersections; 2) we define a controller to operate an autonomous vehicle at such intersections based on the vehicle interaction models and a multi-model strategy; 3) we test our algorithm performance in various simulated traffic scenarios and show that it has a good capability of resolving conflicts.

We note that in this paper we do not consider autonomous intersection management [8], [9], where a centralized controller schedules and manages the motion of all cars at

\*This research has been supported by the National Science Foundation Award Number CNS 1544844.

Nan Li, Ilya Kolmanovsky, and Anouck Girard are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109, USA {nanli, ilya, anouck}@umich.edu. Yildiray Yildiz is with the Department of Mechanical Engineering, Bilkent University, Ankara, Turkey yyildiz@bilkent.edu.tr.

the intersection. Instead, we are focusing on developing an autonomous vehicle controller that can handle existing unmanaged intersections, and we leave the consideration of how our developments can also benefit the realization of intersection automation to future work.

## II. VEHICLE MODEL

### A. Dynamic model

We use the following discrete-time model to represent the vehicle dynamics at intersections,

$$\begin{aligned} x(t+1) &= x(t) + v(t) \cos(\theta(t)) \Delta t, \\ y(t+1) &= y(t) + v(t) \sin(\theta(t)) \Delta t, \\ v(t+1) &= v(t) + a(t) \Delta t, \\ \theta(t+1) &= \theta(t) + \omega(t) \Delta t, \end{aligned} \quad (1)$$

where  $t$  denotes the discrete time instant,  $(x, y)$  [m] represent the vehicle's global position,  $v$  [m/s] denotes the vehicle's speed,  $\theta$  [rad] denotes the vehicle's yaw angle (the angle between the vehicle's heading direction and the global  $x$ -direction),  $a$  [m/s<sup>2</sup>] denotes the vehicle's acceleration,  $\omega$  [rad/s] denotes the vehicle's yaw angle rate, and  $\Delta t$  [s] denotes the time step. The acceleration  $a(t)$  and yaw angle rate  $\omega(t)$  are two inputs that are determined by the action the vehicle selects, as described in the next subsection.

Since in this paper we focus on the high-level decision making, the model (1) is sufficient – more detailed vehicle dynamics and controls can be engineered at lower levels, as is done in [10].

### B. Action set

We let  $\Gamma$  designate a finite set of maneuver actions a vehicle may apply at an intersection, based on the “action-point” human driver models [11], [12]. Each action  $\gamma \in \Gamma$  is an  $(a, \omega)$  pair that determines the inputs to the model (1). In this paper, we consider an action set  $\Gamma$  containing the following actions:

- 1) “Maintain”:  $(a, \omega) = (0, 0)$ ;
- 2) “Turn left”:  $(a, \omega) = (0, \frac{\pi}{4})$ ;
- 3) “Turn right”:  $(a, \omega) = (0, -\frac{\pi}{4})$ ;
- 4) “Accelerate”:  $(a, \omega) = (2.5, 0)$ ;
- 5) “Decelerate”:  $(a, \omega) = (-2.5, 0)$ ;
- 6) “Brake”:  $(a, \omega) = (-5, 0)$ .

### C. Action selection

The approach to selecting actions is based on receding horizon optimization. At each time step  $t$ , the vehicle selects a sequence of actions,  $\gamma := \{\gamma_0, \gamma_1, \dots, \gamma_{n-1}\}$ ,  $\gamma_i \in \Gamma$ ,  $i = 0, \dots, n-1$ , to maximize a cumulative reward,  $\mathcal{R}$ , over a horizon of length  $n$ , assuming that the actions are applied sequentially at each step within the horizon. After the action sequence that achieves the highest cumulative reward,  $\hat{\gamma}$ , has been found, the vehicle applies the first action  $\hat{\gamma}_0$  of  $\hat{\gamma}$  at the current step. After the motion proceeds with  $\hat{\gamma}_0$  applied over one time step, the process is repeated at the next time step  $t+1$ . The cumulative reward is given by

$$\mathcal{R}(\gamma) = \sum_{i=0}^{n-1} \lambda^i R_i(\gamma), \quad (2)$$

where the stage reward  $R_i(\gamma) = R(\gamma_i | s_i)$  is a function of the action  $\gamma_i$  given the predicted traffic state at step  $i$ ,  $s_i$ , and  $\lambda \in [0, 1]$  is a discount factor. The procedure to predict the traffic state  $s_i$  is described in Section III. Note that  $s_i$  includes the states of all vehicles at the intersection. The procedure to compute  $R$  is given in the next subsection.

### D. Reward function

The stage reward function is defined as follows,

$$R = w_1 \hat{c} + w_2 \hat{s} + w_3 \hat{o} + w_4 \hat{l} + w_5 \hat{d}, \quad (3)$$

where  $w_i > 0$ ,  $i = 1, \dots, 5$ , are weighting factors, and the metrics  $\hat{c}$ ,  $\hat{s}$ ,  $\hat{o}$ ,  $\hat{l}$ , and  $\hat{d}$  are explained below:

$\hat{c}$  (collision avoidance): We define a collision avoidance-zone (c-zone) for each vehicle as a rectangular area that bounds the geometric contour of the vehicle. The term  $\hat{c}$  is assigned  $-1$  if the ego vehicle's c-zone has an overlap with the c-zone of any other vehicle, which indicates a danger of collision, and is assigned 0 otherwise.

$\hat{s}$  (safe distance): We define a safe-zone (s-zone) for each vehicle as a larger rectangular area that contains its c-zone with a safety margin. The term  $\hat{s}$  is assigned  $-1$  if the ego vehicle's s-zone has an overlap with the s-zone of any other vehicle, and assigned 0 otherwise. This term is to encourage the ego vehicle to keep a reasonable distance from other vehicles.

$\hat{o}$  (off road): The ego vehicle should drive on the road rather than off the road. If the ego vehicle's c-zone gets outside the road boundaries, that is, it has an overlap with the grey shadowed areas in Fig. 2(a), the term  $\hat{o}$  is set to  $-1$ , and it is set to 0 otherwise.

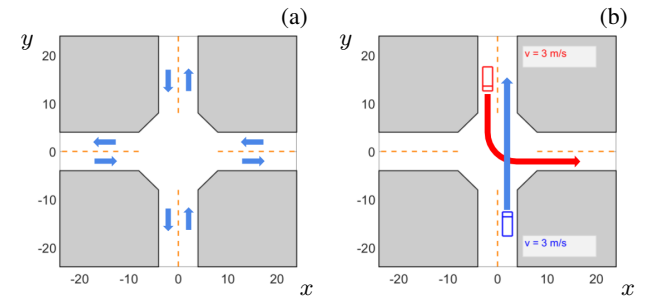


Fig. 2: Traffic scenarios at an intersection. (a) White designates roads, grey designates off-road areas, orange dashed lines are lane markings, and blue arrows indicate the traffic direction in each lane. (b) The blue (red) rectangle represents car 1(2), and the ends with double lines are the front of the cars. The blue (red) arrow indicates the objective lane of car 1(2). The speed of car 1(2) is shown in the blue (red) box.

$\hat{l}$  (lane for opposite direction traffic): The ego vehicle should not 1) drive across the lane markings that separate lanes of traffic moving in opposite directions (double yellow lines in the US); 2) drive in the lanes for traffic of opposite direction to its orientation. If either 1) or 2) are violated, the term  $\hat{l}$  is set to  $-1$ , and it is set to 0 otherwise.

$\hat{d}$  (distance to objective): The ego vehicle has an objective lane to go to. The distance to the objective lane is penalized

so that the ego vehicle is encouraged to take actions to approach/reach it as quickly as it can. In particular, we define a reference point in the objective lane, denoted by the pair  $(x_r, y_r)$  [m], and the term  $\hat{d}$  is assigned value based on

$$-\hat{d} = |x - x_r| + |y - y_r|. \quad (4)$$

### III. LEVEL- $k$ GAME THEORETIC DECISION MAKING

To evaluate the cumulative reward (2), the traffic states  $s_i$ ,  $i = 1, \dots, n-1$ , need to be predicted over the prediction horizon. We use a game theoretic model for this prediction.

Numerous experimental results from psychology, cognitive science, and economics have suggested a hierarchical structure in human reasoning in games, see [13], [14], [15], [16]. The study of this reasoning hierarchy and its applications in game theoretic settings are addressed by the “level- $k$  game theory.” In [5], [6], [17], the level- $k$  game theory is exploited to model vehicle interactions in highway traffic. The model has been compared to human traffic data in [6]. In this paper, we also exploit level- $k$  game theory to model vehicle interactions, in particular, at intersections. Recently, level- $k$  modeling of human agents was also considered in aerospace and energy applications [18], [19], [20], where human-to-human and human-to-automation interactions play a central role.

The model is premised on the idea that strategic agents (drivers/vehicles at an intersection, in our setting) have different reasoning levels. In particular, the level  $k$  indicates an agent’s reasoning depth. The reasoning hierarchy starts from level-0. A level-0 agent makes instinctive decisions to pursue its goal without considering the interactions between itself and the others. On the contrary, a level-1 agent takes into account such interactions in its decision making process, in particular, by assuming that all the other agents in the game are level-0. Specifically, a level-1 agent assumes that all the other agents are level-0 so they make instinctive decisions; the level-1 agent predicts their actions as well as the evolutions of the game resulting from their actions based on this assumption; the level-1 agent then makes its own decision as the best response to such evolutions to pursue its own goal. Similarly, a level- $k$  agent assumes that all the other agents are level- $(k-1)$ , makes predictions based on this assumption, and responds accordingly.

In this paper, a level-0 driver/vehicle treats the other vehicles at the intersection as stationary obstacles, and selects its action sequence,  $\hat{\gamma}_0$ , accordingly. In this setting, a level-0 driver/vehicle may represent an aggressive driver in real traffic, who usually assumes that other drivers will yield the right of way.

After the level-0 is defined, the action selection procedure for level- $k$ ,  $k \geq 1$ , in the case of 2-agent interactions has the following form,

$$\mathcal{R}_k(\gamma) = \mathcal{R}(\gamma | \hat{\gamma}_{k-1}^{\text{other}}) = \sum_{i=0}^{n-1} \lambda^i R_i(\gamma | \hat{\gamma}_{k-1}^{\text{other}}), \quad (5)$$

$$\hat{\gamma}_k^{\text{ego}} \in \arg \max_{\gamma_i \in \Gamma} \mathcal{R}_k(\gamma), \quad (6)$$

where  $\hat{\gamma}_{k-1}^{\text{other}} = \{\hat{\gamma}_{(k-1,0)}^{\text{other}}, \dots, \hat{\gamma}_{(k-1,n-1)}^{\text{other}}\}$  denotes the selected action sequence of the other agent that is assumed to be level- $(k-1)$ , and

$$R_i(\gamma | \hat{\gamma}_{k-1}^{\text{other}}) = R(\gamma_i | s_0, \gamma_0, \dots, \gamma_{i-1}, \hat{\gamma}_{(k-1,0)}^{\text{other}}, \dots, \hat{\gamma}_{(k-1,i)}^{\text{other}}),$$

that is, the obtained reward value for action  $\gamma_i$  at each step  $i$  over the horizon depends on the current traffic state  $s_0$ , the ego agent’s actions  $\{\gamma_0, \dots, \gamma_{i-1}\}$ , and the other agent’s actions  $\{\hat{\gamma}_{(k-1,0)}^{\text{other}}, \dots, \hat{\gamma}_{(k-1,i)}^{\text{other}}\}$ . This dependence reflects the interactions between agents.

To make this hierarchical reasoning decision making procedure more clear, we explain it through an example: Suppose the ego agent is level-2. At first, the ego agent computes an action sequence  $\hat{\gamma}_0^{\text{ego}}$  that is its optimal action sequence if the other agent is treated as a stationary obstacle. Then, the ego agent computes the action sequence  $\hat{\gamma}_1^{\text{other}}$  that is the optimal action sequence for the other agent if the ego agent applies  $\hat{\gamma}_0^{\text{ego}}$  over the horizon and results in traffic states evolving accordingly. Finally, the ego agent computes the action sequence  $\hat{\gamma}_2^{\text{ego}}$  that is its optimal action sequence if the other agent applies  $\hat{\gamma}_1^{\text{other}}$ .

We remark that here we make a simplification: When a level- $k$  ego agent computes  $\hat{\gamma}_k^{\text{ego}}$ , it assumes that the other agent applies  $\hat{\gamma}_{k-1}^{\text{other}}$ , which is computed by assuming that the ego agent was level- $(k-2)$ . Thus,  $\hat{\gamma}_{k-1}^{\text{other}}$  is independent of  $\hat{\gamma}_k^{\text{ego}}$  and can be determined first. With this simplification, we avoid the problem of nested back-and-forth reasoning and make the computation tractable.

While the above formulation accounts for 2 agents, it can be straightforwardly generalized to the case of more agents, and is also computationally scalable. Details will appear elsewhere.

### IV. GAME THEORETIC AUTONOMOUS VEHICLE CONTROLLER

Based on the developed level- $k$  game theoretic driver models, we can define an autonomous vehicle controller for unsignalized intersections. The controller is based on a multi-model strategy, and selects optimal actions according to the model of each of the drivers it is interacting with. Specifically, the ego agent (the autonomous vehicle using this controller) holds a belief about the model of each of the other agents (the drivers), which is a probability distribution over all models. It modifies this belief at every update step by analyzing the actual action applied by each of the other agents.

The principle of this controller is based on the observation that although humans usually do not have perfect knowledge about each other’s characteristics initially, they quickly gain better understandings of each other via interactions, and, as a result, become capable to resolve conflicts.

In this paper, we restrict our treatment to 2-agent interactions. The ego agent’s belief is represented by  $P(k)$ , which is the probability that the other agent can be modeled as level- $k$ . It is shown in [16] that humans are usually level-0, 1 and 2 reasoners in their interactions and level-3 reasoners

are rarely encountered, thus in this paper we consider  $k \leq 2$ , and  $P(0) + P(1) + P(2) = 1$ .

The expected cumulative reward of an action sequence is given by

$$\mathcal{R}_P(\gamma) = \sum_{k=0}^2 P(k) \mathcal{R}(\gamma | \hat{\gamma}_k^{\text{other}}), \quad (7)$$

where  $\mathcal{R}(\gamma | \hat{\gamma}_k^{\text{other}})$  is defined by (5).

The ego agent selects the action sequence that achieves the highest cumulative reward expectation,

$$\hat{\gamma}_P^{\text{ego}} \in \arg \max_{\gamma_i \in \Gamma} \mathcal{R}_P(\gamma). \quad (8)$$

After the action sequence  $\hat{\gamma}_P^{\text{ego}}$  has been found, the ego agent applies the first action in this sequence over one time step.

Moreover, after the other agent applies its action,  $\gamma^{\text{other}}$ , the autonomous vehicle controller compares  $\gamma^{\text{other}}$  to the first action  $\hat{\gamma}_{(k,0)}^{\text{other}}$  of the predicted action sequence  $\hat{\gamma}_k^{\text{other}}$  associated with each level- $k$  model. It improves its belief in the model  $k$  that matches the actual action best, by

$$\begin{aligned} k^* &\in \arg \min_{k=0,1,2} \|\gamma^{\text{other}} - \hat{\gamma}_{(k,0)}^{\text{other}}\|, \\ P(k^*) &\leftarrow P(k^*) + \Delta P, \\ P(k) &\leftarrow P(k) / (\sum_{k=0}^2 P(k)), \end{aligned} \quad (9)$$

where we use  $\|\cdot\|$  to represent a measure of the difference between the action  $\gamma^{\text{other}}$  and the action  $\hat{\gamma}_{(k,0)}^{\text{other}}$  for  $k = 0, 1, 2$ . In our simulation, we use  $\|\gamma^1 - \gamma^2\| = |a^1 - a^2| + |\omega^1 - \omega^2|$ . Since we restrict the action choices of every vehicle to the finite action set  $\Gamma$  defined in Section II,  $P(k^*)$  gets increased when  $\gamma^{\text{other}} = \hat{\gamma}_{(k^*,0)}^{\text{other}}$ .

We remark that (9) is triggered only when there is at least one  $\hat{\gamma}_{(k,0)}^{\text{other}}$  different from others for some  $k$ . Otherwise, the controller has no information to improve its belief because the action selections of all models are the same.

We note that [21] considers a multi-policy strategy for autonomous driving, in which policies are selected from a finite policy set for the ego car to use; while in our multi-model approach, the ego car uses multiple models to predict the other cars' decisions, and then computes its own decisions based on both the prediction associated with each model and its belief in each model, where each model is created based on a game theoretic scheme.

## V. RESULTS

### A. Simulation results of level- $k$ interactions

In this subsection, we present simulation results of level- $k$  driver interactions. The traffic scenario we consider is demonstrated in Fig. 2(b). We consider this particular scenario as it is frequently encountered in urban traffic, and may be particularly challenging for autonomous vehicles.

The simulations are configured as follows: the c-zone and s-zone of each car are, respectively, a 5[m]×2[m] rectangle and a 8[m]×2.4[m] rectangle centered at the geometric center of the car. The width of a lane is 4[m]. The central

area of the intersection is a regular octagon shown as in Fig. 2(b).

We use  $\Delta t = 0.25$ [s],  $n = 8$ ,  $\lambda = 0.9$ , and we select the weights  $w_i$ ,  $i = 1, \dots, 5$ , in (3) as

$$w_1 = 200, w_2 = 20, w_3 = 100, w_4 = 10, w_5 = 1. \quad (10)$$

Different weight selections may reflect different driving styles. Reasonable driving behavior is observed with the weight selection (10).

In particular, we initialize the scenario as: the two cars, car 1 (blue) and car 2 (red), are initialized, respectively, heading straight upward and heading straight downward, with the same distance to the entry of the intersection ( $|y_1(0)| = |y_2(0)| = 16$ [m]), and with the same speed (4[m/s]). This scenario is challenging because it may cause a deadlock. In the sequel, we refer to this initialization as “*Scenario 1*.” We test the interactive behaviors of the two cars when they are of different combinations of reasoning levels. Snapshots of some of the simulations are shown in Fig. 3.

From the simulations, we can observe that for the case of level- $k$  versus level- $(k-1)$ , that is, when one vehicle is modeled as level- $k$  and the other vehicle is modeled as level- $(k-1)$ , the two vehicles can resolve the conflict and successfully get to their objective lanes. This is reasonable since when at least one of the two agents makes a correct assumption about the other as well as correctly predicts the other's actions, it can resolve the conflict by correctly responding to the traffic state evolutions. When both vehicles make wrong assumptions about each other, for example, for the case of level- $k$  versus level- $k$ , the vehicles may not always be able to resolve conflicts. Specifically, for level-1 versus level-1, the two vehicles eventually get to their objective lanes although at the beginning the blue car makes unnecessary decelerations (to yield the right of way to the red car). However, for level-0 versus level-0 and level-2 versus level-2, the two vehicles eventually collide.

Next, we consider another scenario in which the two cars, car 1 (blue) and car 2 (red), are initialized, respectively, heading straight upward and heading straight downward, with their  $y$ -positions randomly selected from [12, 20][m] based on the uniform distribution, and with their speeds randomly selected from [3, 5][m/s] based on the uniform distribution. We refer to this initialization as “*Scenario 2*.” We run simulations multiple times for each combination of levels. The percentage of simulation runs where both cars can resolve their conflicts and successfully reach their objective lanes within 10[s] and without colliding with each other, without getting off the road or getting on the lanes for opposite direction traffic are shown in Table I.

Interaction type	Percentage of conflict resolution
L1 vs. L0	99%
L2 vs. L1	95%
L0 vs. L0	41%
L1 vs. L1	84%
L2 vs. L2	57%
L2 vs. L0	41%

TABLE I: Conflict resolution percentage of level- $k$  interactions.

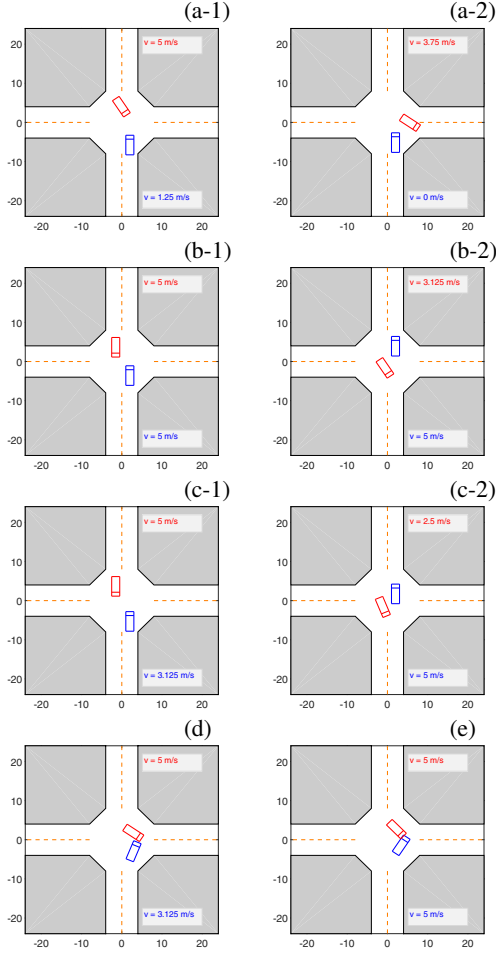


Fig. 3: Level- $k$  interaction simulation results. (a-1,2) L1 (blue) vs. L0 (red) at 2.5[s], 4[s]; (b-1,2) L2 (blue) vs. L1 (red) at 2.5[s], 4[s]; (c-1,2) L1 (blue) vs. L1 (red) at 2.5[s], 4[s]; (d) L0 (blue) vs. L0 (red) at 3[s]; (e) L2 (blue) vs. L2 (red) at 3[s].

From these results, we can observe that for level- $k$  versus level- $(k-1)$ , the conflicts are very likely to be resolved; while, for level- $k$  versus level- $k$ , the two vehicles may fail in resolving the conflicts. Also, we observe that the percentages of failures are higher than observed in everyday transportation. The reasons are: 1) In our simulation, the level is fixed during the entire simulation; while, a human driver may be able to adjust his/her level based on his/her interactions with the other driver. Indeed, we have exploited this fact to develop the autonomous vehicle controller, whose performance is shown in the next subsection. 2) Human drivers can communicate with each other through other ways, such as eye contact or making gestures, to help them resolve the conflict; such communications are not considered in our simulations.

### B. Simulation results of autonomous vehicle controller

In this subsection, we present simulation results of the proposed multi-model based autonomous vehicle controller. Based on an experimental study of human behavior conducted in [16], we initialize the autonomous vehicle's belief

distribution about the other vehicle's model as  $P(0) = 0.1$ ,  $P(1) = 0.6$ , and  $P(2) = 0.3$ . At each update step, the probability of the best model(s) is improved by  $\Delta P = 0.5$ , and the distribution is then re-normalized, see (9).

We note that if the controller can identify the driver model efficiently and accurately, its design enables it to make correct responses. Thus, we choose to present the time histories of  $P(k)$ ,  $k = 0, 1, 2$ , with the controller interacting with different driver types, to demonstrate its performance. We first initialize the simulation as *Scenario 1*, and let car 1 (blue) be controlled by the autonomous vehicle controller and car 2 (red) be modeled using different level- $k$  models. The results are shown in Fig. 4.

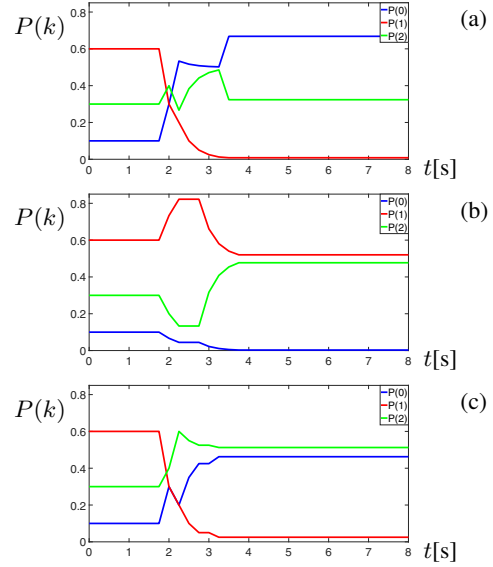


Fig. 4: Model identification history. (a) Autonomous vehicle controller vs. level-0 driver; (b) Autonomous vehicle controller vs. level-1 driver; (c) Autonomous vehicle controller vs. level-2 driver.

From these results, we can observe that the controller efficiently identifies the true model  $k^*$  of the driver it is interacting with (by increasing  $P(k^*)$ ). As a result, the controller gives higher weight to the cumulative reward associated with the level- $k^*$  model in (7), which motivates the controller's preference to the optimal actions corresponding to model  $k^*$ . In this experiment, both the autonomous vehicle and the human-driven vehicle reach their objective lanes safely and in a timely manner, so the simulation snapshots are omitted.

It is not surprising that the controller performs well in the above experiment, because its design is based on the level- $k$  models. To show the robustness of this controller, two additional experiments are performed, evaluating: 1) the controller's performance when facing a driver model it is not familiar with; and 2) the controller's performance when facing a vehicle controlled based on the same control logic.

For the first case, we define a driver model (D) as:

$$\mathcal{R}_D(\gamma) = \frac{1}{2}\mathcal{R}(\gamma|\hat{\gamma}_0^{\text{other}}) + \frac{1}{2}\mathcal{R}(\gamma|\hat{\gamma}_1^{\text{other}}),$$

$$\hat{\gamma}_D^{\text{ego}} \in \arg \max_{\gamma_i \in \Gamma} \mathcal{R}_D(\gamma), \quad (11)$$



that is, the driver assumes that half of the other drivers are level-0 and the other half are level-1, and responds accordingly. This driver model is different from any level- $k$  models,  $k = 0, 1, 2$ , and thus is not included in the model set of the multi-model based autonomous vehicle controller.

For the second case, we run simulations where both cars are controlled by the developed multi-model controller.

We first test both cases by initializing the simulation as *Scenario 1*, and let car 1 (blue) be controlled by the autonomous vehicle controller and car 2 (red) be controlled by, respectively, the new model (11) and the autonomous vehicle controller. The simulation snapshots are shown in Fig. 5.

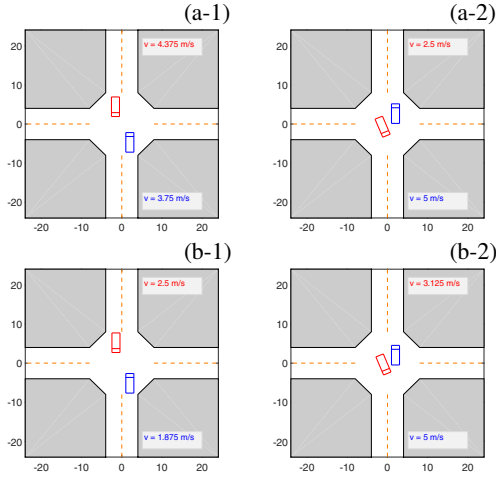


Fig. 5: Autonomous vehicle controller simulation results at 2.5[s], 4[s]. (a-1,2) Autonomous car (blue) vs. unfamiliar driver model (red); (b-1,2) Autonomous car (blue) vs. autonomous car (red).

We then test the controller versus level- $k$  models,  $k = 0, 1, 2$ , versus the driver model (11), and versus the same controller by initializing the simulations as *Scenario 2*, and run multiple times. The percentage of simulation runs where both car 1 and car 2 can successfully reach their objective lanes within 10[s] and without colliding with each other, without getting off the road or getting on the lanes for opposite direction traffic are shown in Table II.

Interaction type	Percentage of conflict resolution
Auto vs. L0	94%
Auto vs. L1	95%
Auto vs. L2	94%
Auto vs. D	90%
Auto vs. Auto	93%

TABLE II: Conflict resolution percentage of the AV controller.

## VI. SUMMARY

In this paper, we exploited a level- $k$  game theoretic approach to model the time-extended, multi-step vehicle-to-vehicle interactions at unsignalized intersections. The conditions under which the conflicts between vehicles at the intersection can be resolved were studied. Moreover, an autonomous vehicle controller based on the driver interaction models and online model estimation was proposed. It was

shown by simulations to have a good capability of resolving conflicts versus different driver types at the intersection.

## REFERENCES

- [1] Y. Chen, H. Peng, and J. Grizzle, "Fast trajectory planning and robust trajectory tracking for pedestrian avoidance," *IEEE Access*, 2017.
- [2] M. Althoff, O. Stursberg, and M. Buss, "Model-based probabilistic collision detection in autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 2, pp. 299–310, 2009.
- [3] Institute of Transportation Engineers, "Unsignalized intersection improvement guide," Tech. Rep., <http://www.ite.org/uiig/types.asp>.
- [4] R. Mandiau, A. Champion, J.-M. Auberlet, S. Espié, and C. Kolski, "Behaviour based on decision matrices for a coordination between agents in a urban traffic simulation," *Applied Intelligence*, vol. 28, no. 2, pp. 121–138, 2008.
- [5] N. Li, D. Oyler, M. Zhang, Y. Yildiz, A. Girard, and I. Kolmanovsky, "Hierarchical reasoning game theory based approach for evaluation and testing of autonomous vehicle control systems," in *Decision and Control, 55th Conference on*. IEEE, 2016, pp. 727–733.
- [6] N. Li, D. W. Oyler, M. Zhang, Y. Yildiz, I. Kolmanovsky, and A. R. Girard, "Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems," *IEEE Transactions on Control Systems Technology*, vol. PP, no. 99, pp. 1–16, 2017.
- [7] D. Sadigh, S. Sastry, S. A. Seshia, and A. D. Dragan, "Planning for autonomous cars that leverage effects on human actions," in *Robotics: Science and Systems*, 2016.
- [8] K. Dresner and P. Stone, "A multiagent approach to autonomous intersection management," *Journal of Artificial Intelligence Research*, vol. 31, pp. 591–656, 2008.
- [9] J. Wu, A. Abbas-Turki, and A. El Moudni, "Cooperative driving: an ant colony system for autonomous intersection management," *Applied Intelligence*, vol. 37, no. 2, pp. 207–222, 2012.
- [10] N. Li, H. Chen, I. Kolmanovsky, and A. Girard, "An explicit decision tree approach for automated driving," in *Dynamic Systems and Control Conference*. ASME, 2016.
- [11] R. Michaels, "Perceptual factors in car following," in *Proceedings of the 2nd International Symposium on the Theory of Road Traffic Flow (London, England)*, OECD, 1963.
- [12] D. Hoefs, "Entwicklung einer messmethode uber den bewegungsablauf des kolonnenverkehrs," *Universitat (TH) Karlsruhe, Germany*, 1972.
- [13] D. O. Stahl and P. W. Wilson, "On players models of other players: Theory and experimental evidence," *Games and Economic Behavior*, vol. 10, no. 1, pp. 218–254, 1995.
- [14] T. Hedden and J. Zhang, "What do you think I think you think?: Strategic reasoning in matrix games," *Cognition*, vol. 85, no. 1, pp. 1–36, 2002.
- [15] M. A. Costa-Gomes and V. P. Crawford, "Cognition and behavior in two-person guessing games: An experimental study," *The American economic review*, vol. 96, no. 5, pp. 1737–1768, 2006.
- [16] M. A. Costa-Gomes, V. P. Crawford, and N. Iriberri, "Comparing models of strategic thinking in van huyck, battalio, and beil's coordination games," *Journal of the European Economic Association*, vol. 7, no. 2-3, pp. 365–376, 2009.
- [17] D. W. Oyler, Y. Yildiz, A. R. Girard, N. I. Li, and I. V. Kolmanovsky, "A game theoretical model of traffic with multiple interacting drivers for use in autonomous vehicle development," in *American Control Conference*. IEEE, 2016, pp. 1705–1710.
- [18] S. Backhaus, R. Bent, J. Bono, R. Lee, B. Tracey, D. Wolpert, D. Xie, and Y. Yildiz, "Cyber-physical security: A game theory model of humans interacting over control systems," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2320–2327, 2013.
- [19] Y. Yildiz, A. Agogino, and G. Brat, "Predicting pilot behavior in medium-scale scenarios using game theory and reinforcement learning," *Journal of Guidance, Control, and Dynamics*, 2014.
- [20] N. Musavi, D. Onural, K. Gunes, and Y. Yildiz, "Unmanned aircraft systems airspace integration: A game theoretical framework for concept evaluations," *Journal of Guidance, Control, and Dynamics*, 2016.
- [21] A. G. Cunningham, E. Galceran, R. M. Eustice, and E. Olson, "MPDM: Multipolicy decision-making in dynamic, uncertain environments for autonomous driving," in *Robotics and Automation (ICRA), International Conference on*. IEEE, 2015, pp. 1670–1677.