

# Courteous Autonomous Cars

Liting Sun<sup>1</sup>, Wei Zhan<sup>1</sup>, Masayoshi Tomizuka<sup>1</sup>, and Anca D. Dragan<sup>2</sup>

**Abstract**—Typically, autonomous cars optimize for a combination of safety, efficiency, and driving quality. But as we get better at this optimization, we start seeing behavior go from too conservative to too aggressive. The car’s behavior exposes the incentives we provide in its cost function. In this work, we argue for cars that are not optimizing a purely selfish cost, but also try to be *courteous* to other interactive drivers. We formalize courtesy as a term in the objective that measures the increase in another driver’s cost induced by the autonomous car’s behavior. Such a courtesy term enables the robot car to be aware of possible irrationality of the human behavior, and plan accordingly. We analyze the effect of courtesy in a variety of scenarios. We find, for example, that courteous robot cars leave more space when merging in front of a human driver. Moreover, we find that such a courtesy term can help explain real human driver behavior on the NGSIM dataset.

## I. INTRODUCTION

Autonomous cars are getting better at generating their motion not only in isolation, but also around people. We now have many strategies for dealing with interactions with people on the road, each modeling people in substantially different ways.

Most techniques first anticipate what people plan on doing, and generate the car’s motion to be efficient, but also to safely stay out of their way. This prediction can be as simple as assuming the person will maintain their current velocity within the planning horizon [1]–[3], or as complicated as learning a human driver policy or cost function [4]–[7].

Other techniques account for the interactive nature of coordinating on the road, and model people as changing their plans depending on what the car does. Some do it via coupled planning, assuming that the person and the robot are on the same team, optimizing the same joint cost function [8]–[10], while others capture interaction as a game in which the human and robot have different utilities, but they influence each other’s actions [11]–[13].

All of these works focus on *how* to optimize the robot’s cost when the robot needs to interact with people. In this paper, we focus on *what* the robot should optimize in such situations, particularly if we consider the fact that humans are not perfectly rational.

<sup>1</sup>Liting Sun, Wei Zhan and Masayoshi Tomizuka are with the Department of Mechanical Engineering, University of California, Berkeley, CA, USA, 94720. {litingsun, wzhan, tomizuka}@berkeley.edu

<sup>2</sup>Anca D. Dragan is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA, 94720. anca@berkeley.edu

Typically, when designing the robot’s cost function, we focus on safety and driving quality of the ego vehicle. Arguably, that is rather *selfish*.

Selfishness has not been a problem with approaches that predict human plans and react to them, because that led to conservative robots that always try to stay out of the way and let people do what they want. But, as we are switching to more recent approaches that draw on the game-theoretic aspects of interaction, our cars are starting to become more aggressive. They cut people off, or inch forward at intersections to go first [11] [14]. While this behavior is good sometimes, we would not want to see it all the time.

Our observation is that as we get better at solving the optimization problem for driving by better models of the world and of the people in it, there is an increased burden on the cost function we optimize to capture what we want. We propose that purely selfish robots that care about their safety and driving quality are not good enough. They should also be *courteous* to other drivers. This is of crucial importance since humans are not perfectly rational, and their behavior will be influenced by the aggressiveness of the robot cars.

*We advocate that a robot should balance minimizing the inconvenience it brings to another driver, and that we can formalize inconvenience as the increase in the other driver’s cost due to the robot’s behavior to capture one aspect of human behavior irrationality.*

We make the following contributions:

**A formalism for courtesy incorporating irrational human behavior.** We formalize courteous planning as trading off between the robot’s selfish objective and a courtesy term, and introduce a mathematical definition for this term for irrational human behavior – we measure the increase of the vehicle’s best cost under the robot’s planned behavior, compared to the vehicle’s best cost under an alternative “best case scenario”, and define the cost increase as the courtesy term.

**An analysis of the effects of courteous planning.** We show the difference between courteous and selfish robots under different traffic scenarios. The courteous robot leaves the person more space when it merges, and might even block another agent (not a person) to ensure that the human can safely proceed.

**Showing that courtesy helps explain human driving.** We do an Inverse Reinforcement Learning (IRL)-based analysis [7], [16]–[18] to study whether our courtesy term helps in better predicting how humans drive. On the NGSIM dataset [19] of real human driver trajec-

ries, we find that courtesy produces trajectories that are significantly closer to the ground truth.

We think that the autonomous car of the future should be safe, efficient, and courteous to others, perhaps even more so than represented in our current human-only driving society. Our paper enables autonomous car designers to decide to make that happen.

## II. PROBLEM STATEMENT

In this paper, we consider an interactive robot-human system with two agents: an autonomous car  $\mathcal{R}$  and a human driver  $\mathcal{H}$ <sup>1</sup>. Our task is to enable a courteous robot car which cares about the potential inconvenience it brings to the human driver's utilities, and generates trajectories that are socially predictable and acceptable.

Throughout the paper, we denote all robot-related terms by subscript  $(\cdot)_{\mathcal{R}}$  and all human-related terms by  $(\cdot)_{\mathcal{H}}$ .

Let  $x_{\mathcal{R}}$  and  $u_{\mathcal{R}}$  denote, respectively, the robot's state and control input, and  $x_{\mathcal{H}}$  and  $u_{\mathcal{H}}$  for the human's.  $x = (x_{\mathcal{R}}^T, x_{\mathcal{H}}^T)^T$  represents the states of the interaction system. For each agent, we have

$$x_{\mathcal{R}}^{t+1} = f_{\mathcal{R}}(x_{\mathcal{R}}^t, u_{\mathcal{R}}^t), \quad (1)$$

$$x_{\mathcal{H}}^{t+1} = f_{\mathcal{H}}(x_{\mathcal{H}}^t, u_{\mathcal{H}}^t), \quad (2)$$

and the overall system dynamics are

$$x^{t+1} = f(x^t, u_{\mathcal{R}}^t, u_{\mathcal{H}}^t). \quad (3)$$

We assume that both the human driver and the autonomous car are optimal planners, and they use Model Predictive Control (MPC) with a horizon of length  $N$ . Let  $C_{\mathcal{R}}$  and  $C_{\mathcal{H}}$  be, respectively, the cost functions of the robot car and the human driver over the horizon:

$$C_i(x^t, u_{\mathcal{R}}, u_{\mathcal{H}}; \theta_i) = \sum_{k=0}^{N-1} c_i(x^{t+k}, u_{\mathcal{R}}^k, u_{\mathcal{H}}^k; \theta_i), i \in \{\mathcal{R}, \mathcal{H}\} \quad (4)$$

where  $\mathbf{u}_i = (u_i^0, u_i^1, \dots, u_i^{N-1})^T$  are sequences of control actions of the robot car ( $i=\mathcal{R}$ ) and the human driver ( $i=\mathcal{H}$ ), and  $x^{t,k}$  with  $k=0, 1, \dots, N-1$  are the corresponding sequence of system states.  $\theta_i$  represent, respectively, the preferences of the robot car ( $i=\mathcal{R}$ ) and the human driver ( $i=\mathcal{H}$ ). At every time step  $t$ , the robot car and the human driver generate their optimal sequences of actions  $\mathbf{u}_{\mathcal{R}}^*$  and  $\mathbf{u}_{\mathcal{H}}^*$  by minimizing  $C_{\mathcal{R}}$  and  $C_{\mathcal{H}}$ , respectively, execute the first steps  $u_{\mathcal{R}}^{*0}$  and  $u_{\mathcal{H}}^{*0}$  (i.e., set  $u_i^t = u_i^{*0}$  in (3)), and replan for step  $t+1$ .

Such an optimization-based state feedback strategy formulates the closed-loop dynamics of the robot-human interaction system as a game. To simplify the game, we assume that the robot car has access to  $C_{\mathcal{H}}$ , and that the human only computes a best response to the robot's actions rather than trying to influence them,

<sup>1</sup>If there are multiple robot cars that we control, we treat them all as a single  $\mathcal{R}$ . If there are multiple human drivers, we reason about how each of them affects the robot's utility separately.

as in [11]. This means that the robot car can compute, for any control sequence it considers, how the human would respond and what cost the human will incur:

$$\mathbf{u}_{\mathcal{H}}^* = \arg \min_{\mathbf{u}_{\mathcal{H}}} C_{\mathcal{H}}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}) \triangleq g(x^t, \mathbf{u}_{\mathcal{R}}; \theta_{\mathcal{H}}) \quad (5)$$

$$C_{\mathcal{H}}^*(\mathbf{u}_{\mathcal{R}}) = C_{\mathcal{H}}(x^t, \mathbf{u}_{\mathcal{R}}, g(x^t, \mathbf{u}_{\mathcal{R}}; \theta_{\mathcal{H}}); \theta_{\mathcal{H}}). \quad (6)$$

Here  $g(x^t, \mathbf{u}_{\mathcal{R}}; \theta_{\mathcal{H}})$  represents the response curve of the human driver towards the autonomous car.

Armed with this model, the robot can now compute what it should do, such that when the human responds, the combination is good for the robot's cost:

$$\mathbf{u}_{\mathcal{R}}^* = \arg \min_{\mathbf{u}_{\mathcal{R}}} C_{\mathcal{R}}(x^t, \mathbf{u}_{\mathcal{R}}, g(x^t, \mathbf{u}_{\mathcal{R}}; \theta_{\mathcal{H}}); \theta_{\mathcal{R}}). \quad (7)$$

Our goal is to generate courteous robot behavior to the human, i.e. that takes into consideration the inconvenience it brings to the human driver. We will do so by changing the cost function of the robot to reflect this inconvenience.

## III. COURTEOUS PLANNING

We propose a *courteous planning* strategy based on one key observation: human is not perfectly rational, and one of the irrationality is that they weight losses higher than gains when evaluating their actions [15]. Hence, a courteous robot car should balance the minimization of its own cost function and the inconvenience (loss) it brings to the human driver.

Therefore, we construct  $C_{\mathcal{R}}$  in (7) as

$$C_{\mathcal{R}}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{R}}, \theta_{\mathcal{H}}, \lambda_c) = C_{\mathcal{R}}^{\text{self}}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{R}}) + \lambda_c C_{\mathcal{R}}^{\text{court}}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}), \quad (8)$$

where  $C_{\mathcal{R}}^{\text{self}}$  is the cost function for a regular (selfish) robot car which cares about only its own utilities (safety, efficiency, etc), and  $C_{\mathcal{R}}^{\text{court}}$  models the courtesy term of the robot car to the human driver. It is a function of the robot car's behavior, the human's behavior, the human's cost parameters ( $\theta_{\mathcal{H}}$ ) and some alternative costs (see Section III.A).  $\lambda_c \in [0, \infty)$  captures the trade-off. If we want the robot car to be just as courteous as a human driver, we could learn  $\lambda_c$  from human driver demonstration, as we do in Section V. As robot designers, we might set this parameter higher than regular human driving to enable more courteous autonomous cars, particularly when they do not have passengers on board.

### A. Alternative Costs

With any robot plan  $\mathbf{u}_{\mathcal{R}}$ , the robot car changes the human driver's environment and therefore induces a best cost for the human,  $C_{\mathcal{H}}^*(\mathbf{u}_{\mathcal{R}})$ . Our courtesy term compares this cost with the *alternative*,  $C_{\mathcal{H}}^{\text{alt},*}$  – the best case scenario for the person. It is not immediately clear how to define this best case scenario since it may vary depending different on driving scenarios. We explore three alternatives.

**What the human could have done, had the robot car not been there.** We first consider a world in which the robot car wouldn't even exist to interfere the person. In such a world, the person gets to optimize their cost without the robot car:

$$C_{\mathcal{H}}^{alt,*}(x^t, \theta_{\mathcal{H}}) = \min_{\mathbf{u}_{\mathcal{H}}} C_{\mathcal{H}}(x^t, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}) \quad (9)$$

This induces a very generous definition of courtesy: the alternative is for the robot car to not have been on the road at all. In reality though, the robot car is there, which leads to our second alternative.

**What the human could have done, had the robot car only been there to help the human.** Our second alternative is to assume that the robot car already on the road could be completely altruistic. The robot car could actually optimize the human driver's cost, being a perfect collaborator:

$$C_{\mathcal{H}}^{alt,*}(x^t, \theta_{\mathcal{H}}) = \min_{\mathbf{u}_{\mathcal{H}}, \mathbf{u}_{\mathcal{R}}} C_{\mathcal{H}}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}) \quad (10)$$

For this alternative, the robot car and the human would perform a joint optimization for the human's cost. For example, the robot car can brake to make sure that the human could change lanes in front of it, or even block another traffic participant to make sure the human has space.

**What the human could have done, had the robot car just kept doing what it was previously doing.** A fully collaborative robot car is still perhaps not the fairest one to compute inconvenience against. After all, the autonomous car does have a passenger sometimes, and it is fair to take their needs into account too. Our third alternative computes how well the human driver could have done, had the robot car kept acting the same way as it was previously doing:

$$C_{\mathcal{H}}^{alt,*}(x^t, \theta_{\mathcal{H}}) = \min_{\mathbf{u}_{\mathcal{H}}} C_{\mathcal{H}}(x^t, \mathbf{u}_{\mathcal{R}}^{t-1}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}) \quad (11)$$

This means that the person is now responding to a constant robot trajectory  $\mathbf{u}_{\mathcal{R}}^{t-1} = (u_{\mathcal{R}}^{t-1}, \dots, u_{\mathcal{R}}^{t-1})$ , for instance, maintaining its current velocity.

Our experiments below explore these three different alternative options for the courtesy term.

### B. Courtesy Term

We define the courtesy term based on the difference between what cost the human has, and what cost they would have had in the alternative:

**Definition 1** (Courtesy of the Robot Car)

$$C_{\mathcal{R}}^{court}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}) = \max\{0, C_{\mathcal{H}}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{H}}) - C_{\mathcal{H}}^{alt,*}(x^t; \theta_{\mathcal{H}})\} \quad (12)$$

Note that we could have also sent the courtesy term to simply be the human cost, and have the robot trade off between its cost and the human's. However, that would have penalized the robot for any cost the human

incurs, even if the robot does not bring any inconvenience to the human. That might cause too conservative behavior. In fact, if we treat the alternative cost as the reference point in Prospect Theory – a human irrationality model [15], then the theory suggests that human weigh losses more than gains. This means that our courteous robot car should care more about avoiding additional inconvenience, rather than providing more convenience, i.e., helping to reduce the human cost lower than the alternative one. Mathematically, this concept is formulated via Definition 1: the robot does not get any bonus for bringing the human cost lower than  $C_{\mathcal{H}}^{alt,*}$  (possible with some definitions of  $C_{\mathcal{H}}^{alt,*}$ ), it only gets a penalty for making it higher.

### C. Solution

Thus far, we have constructed a compound cost function  $C_{\mathcal{R}}(x^t, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta_{\mathcal{R}}, \theta_{\mathcal{H}}, \lambda_c)$  to enable a courteous robot car, considering three alternative costs. At every step, the robot needs to solve the optimization problem in (7) to find the best actions to take. We approximate the solution by alternatively fixing one of  $\mathbf{u}_{\mathcal{R}}$  or  $\mathbf{u}_{\mathcal{H}}$ , and solving for the other.

## IV. ANALYSIS OF COURTEOUS PLANNING

In this section, we analyze the effect of courteous planning on the robot's behavior in different simulated driving scenarios. In Section V, we study how courteous planning can help better explain real human driving data, enabling robots to be more human-like and predictable, as well as better able at anticipating human driver actions on the road.

*Simulation Environment:* We implement the simulation environment using Julia [20] on a 2.5 GHz Intel Core i7 processor with 16 GB RAM. We set the horizon length to  $N=10$ , and the sampling time to 0.1s. Our simulated environment is 1/10 scale of the real world: 1/10 road width, car sizes, maximum acceleration ( $0.5m/s^2$ ) and deceleration ( $-1.0m/s^2$ ), and low speed limit (1.0m/s).

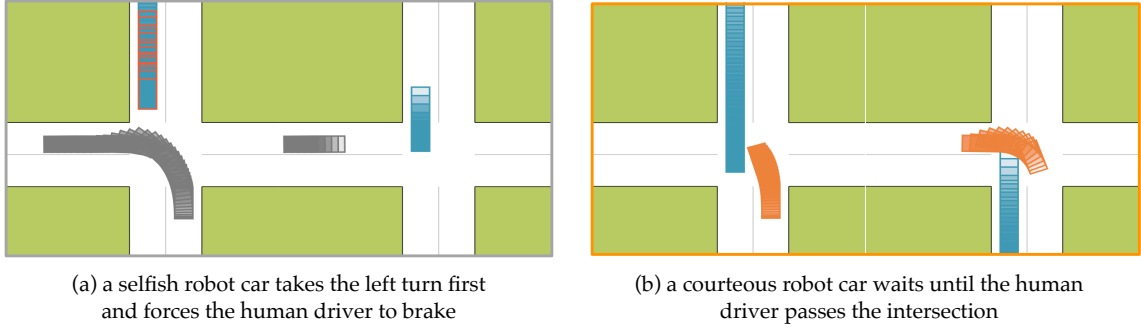
Regarding the cost functions  $C_{\mathcal{H}}$  and  $C_{\mathcal{R}}$  in (6)-(8), except for the courtesy term formulated above, we penalize safety, car speed, comfort level and goal distances in both  $C_{\mathcal{H}}$  and  $C_{\mathcal{R}}^{self}$ . Details about this can be found later in Section V.

For all results, we denote a selfish (baseline) autonomous car with gray rectangle, a courteous one as orange, and the human driver as dark blue.

### A. The Effect of Courtesy

1) *Lane Changing:* We first consider a lane changing driving scenario, as shown in Fig. 1. The autonomous car wants to merge into the human driver's lane from an adjacent lane. We assume that the goal of the human driver is to maintain speed. Then all three different alternatives lead to the same alternative optimal behavior and cost of the human: the human would go in their lane undisturbed by the robot. Hence, with constant





**Fig. 4:** Interaction between a straight-driving human and a left-turning autonomous car: (a) a selfish (baseline) robot car takes a left turn immediately and forces the human driver to brake (red frames); (b) a courteous robot car waits in the middle of the intersection and takes the left turn after the human passes so that the human can maintain speed.

native world we are considering to evaluate how much inconvenience the autonomous car is causing. Here, we consider a scenario in which that is no longer the case to highlight the differences generated by the alternative formulations of courtesy in the robot car’s behavior.

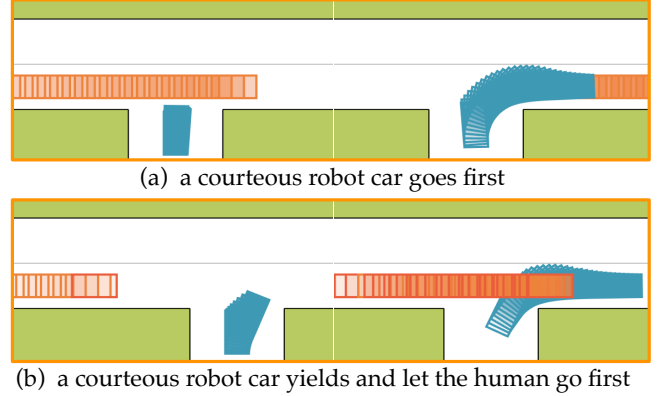
We consider a scenario where the human is turning right, with a straight-driving robot car coming from their left. In this scenario, the three alternative costs are different, which leads to different courtesy terms:

- Alternative I—Robot car not being there: the optimal human behavior would be to take a right turning directly;
- Alternative II—Robot car being collaborative: the robot would take the necessary yielding maneuver to let the human driver take the right turn first, leading to the same alternative optimal human behavior of performing the right turn directly;
- Alternative III—Robot car maintaining behavior: the robot car would maintain its speed, and the optimal human behavior would be to slow down.

Figure 5 summarizes the results of using these different courtesy terms. In Alternative III, a courteous robot car goes first, as shown in Fig. 5(a). Intuitively, this is because  $C_H^{alt,*}$  is initially high, and by maintaining its speed (or even accelerating depending on  $C_R^{self}$ ), no further inconvenience is brought to the human by the robot car, i.e.,  $C_R^{court}$  remains zero. Hence, the robot car goes first (Had the robot try to brake, it only increases  $C_R^{self}$  without changing  $C_R^{court}=0$ , and therefore  $C_R$  increases). The other two alternatives (I and II) are much more generous to the human. Results in Fig. 5(b) show that a courteous robot car finds it too expensive to force the human to go second, and slows down to let the human go first. The red frames in Fig. 5(b) indicate the time instants when the autonomous car brakes.

### C. Extension to environments with multiple agents

We study a scenario on a two-way road. The robot car and the human are driving towards opposite directions, but the robot car is blocked and it has to temporarily merge into the human driver’s lane to get through, as in Fig. 6. We use the collaborative robot as our alternative formulation of the courtesy term in this scenario.

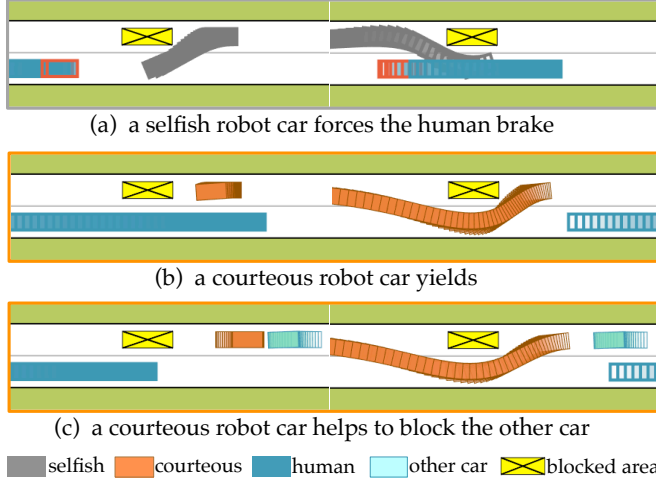


**Fig. 5:** Interaction between a right-turning human driver and a courteous autonomous car with different courtesy terms: (a) the robot car goes first when it evaluates the courtesy term using going forward as an alternative world; (b) the robot car yields and let the human go first when it evaluates the courtesy term based on a collaborative or not-being-there alternative world.

When there are only two agents in the environment, i.e., the autonomous car and the human driver, the results for a selfish and a courteous autonomous car are shown in Fig. 6(a)-(b): A selfish autonomous car directly merges into the human’s lane and forces the human driver to brake; while a courteous autonomous car decides to wait until the human driver passes by since the courtesy term becomes too expensive to go first.

Such courtesy-aware planning becomes much more interesting when there is a third agent in the environment, as shown in Fig. 6(c). We assume that the third agent is a responsive agent to the autonomous car and the autonomous car is courteous only to the human driver (and not to both). In this case, for  $C_H^{alt,*}$ , the human would ideally want to pass undisturbed by either the robot or the other agent: the courtesy term captures the difference in cost to the human between the robot’s behavior and the alternative of a collaborative robot, and this cost to the human depends on how much progress the human is able to make and how fast. As a result, a very courteous robot has an incentive to produce behavior that is as close as possible to making that happen.

Then an interesting behavior emerges: the autonomous car first backs up to block the third agent (the following car) from interrupting the human driver until the human driver safely passes them, and then the robot car finishes its task. This displays truly collaborative behavior, and only happens with high enough weight on the courtesy term. This may not be practical for real on-road driving, but it enables the design of highly courteous robots in some particular scenarios where human have higher priority over all other autonomous agents.



**Fig. 6:** A blocking-area overtaking scenario: (a) with a selfish cost function, the robot car overtakes first and forces the human driver to brake; (b)(c) a courtesy-aware robot car yields to the human driver and even helps to block other cars depending on its formulation of the human driver's alternative world

## V. COURTESY HELPS EXPLAIN HUMAN DRIVING

Thus far, we have shown that courtesy is useful for enabling cars to generate actions that do not cause inconvenience to other drivers. We have also seen that the larger the weight we put on the courtesy term, the more the car behavior becomes social. A natural next question is – are humans courteous?

Our hypothesis is that our courtesy term can help explain human driving behavior. If that is the case, this has two important implications: it means that it can enable robots to better predict human actions by giving them a more accurate model of how people drive, and it also means that robot can use courtesy to produce more human-like driving.

We put our hypothesis to the test by learning a cost function from human driver data, with and without a courtesy feature. We find that using the courtesy feature leads to a more accurate cost function that is better at reproducing human driver data, lending support to our hypothesis.

### A. Learning Cost Functions from Human Demonstrations

1) *Human Data Collection:* The human data is collected from the Next Generation SIMulation (NGSIM)

dataset [19], which captures the highway driving behaviors/trajectories by digital video cameras mounted on top of surrounding buildings. We selected 153 left-lane-changing driving trajectories on Interstate 80 (near Emeryville, California), and separated them into two sets: a *training* set of size 100 (denoted by  $\mathcal{U}_D$ , i.e., the human demonstrations), and the other 53 trajectories as the *test* set.

2) *Learning Algorithm:* We use Inverse Reinforcement Learning (IRL) [7], [16]–[18] to learn an appropriate cost function from human data.

We assume that cost function is parameterized as a linear combination of features:

$$c(x^t, u_{\mathcal{R}}^t, u_{\mathcal{H}}^t; \theta) = \theta^T \phi(x^t, u_{\mathcal{R}}^t, u_{\mathcal{H}}^t). \quad (13)$$

Then over the trajectory length  $L$ , the cumulative cost function becomes

$$\begin{aligned} C(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta) &= \theta^T \sum_{t=0}^{L-1} \phi(x^t, u_{\mathcal{R}}^t, u_{\mathcal{H}}^t) \\ &= \theta^T \Phi(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}) \end{aligned} \quad (14)$$

where  $\mathbf{u}_{\mathcal{R}}$  and  $\mathbf{u}_{\mathcal{H}}$  are, respectively, the actions of the robot car and the human over the trajectory. Our goal is to find the weights  $\theta$  which maximizes the likelihood of the demonstrations:

$$\theta^* = \arg \max_{\theta} P(\mathcal{U}_D | \theta) \quad (15)$$

Building on the principle of maximum entropy, we assume that trajectories are exponentially more likely when they have lower cost:

$$P(\mathbf{u}_{\mathcal{H}}, \theta) \propto \exp \left( -C(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H}}; \theta) \right). \quad (16)$$

Thus the probability (likelihood) of the demonstration set becomes

$$P(\mathcal{U}_D | \theta) = \prod_{i=1}^n \frac{P(\mathbf{u}_{\mathcal{H},i}^D, \theta)}{P(\theta)} = \prod_{i=1}^n \frac{P(\mathbf{u}_{\mathcal{H},i}^D, \theta)}{\int P(\tilde{\mathbf{u}}_{\mathcal{H}}, \theta) d\tilde{\mathbf{u}}_{\mathcal{H}}} \quad (17)$$

where  $n$  is the number of trajectories in  $\mathcal{U}_D$ .

To tackle the partition term  $\int P(\tilde{\mathbf{u}}_{\mathcal{H}}, \theta) d\tilde{\mathbf{u}}_{\mathcal{H}}$  in (17), we approximate  $C(x^0, \mathbf{u}_{\mathcal{R}}, \tilde{\mathbf{u}}_{\mathcal{H}}; \theta)$  with its Laplace approximation as proposed in [7]:

$$\begin{aligned} C(x^0, \mathbf{u}_{\mathcal{R}}, \tilde{\mathbf{u}}_{\mathcal{H}}; \theta) &\approx C(x^0, \mathbf{u}_{\mathcal{R}}, \mathbf{u}_{\mathcal{H},i}^D; \theta) + \left( \tilde{\mathbf{u}}_{\mathcal{H}} - \mathbf{u}_{\mathcal{H},i}^D \right)^T \frac{\partial C}{\partial \mathbf{u}_{\mathcal{H}}} \\ &\quad + \frac{1}{2} \left( \tilde{\mathbf{u}}_{\mathcal{H}} - \mathbf{u}_{\mathcal{H},i}^D \right)^T \frac{\partial^2 C}{\partial \mathbf{u}_{\mathcal{H}}^2} \left( \tilde{\mathbf{u}}_{\mathcal{H}} - \mathbf{u}_{\mathcal{H},i}^D \right). \end{aligned} \quad (18)$$

With the assumption of locally optimal demonstrations, we have  $\frac{\partial C}{\partial \mathbf{u}_{\mathcal{H}}} \big|_{\mathbf{u}_{\mathcal{H},i}^D} = 0$  in (18). This simplifies the partition term  $\int P(\tilde{\mathbf{u}}_{\mathcal{H}}, \theta) d\tilde{\mathbf{u}}_{\mathcal{H}}$  as a Gaussian Integral where a closed-form solution exists (see [7] for details). Substituting (17) and (18) into (15) yields the optimal parameter  $\theta^*$  as the maximizer.



## B. Experiment Design

**Hypothesis.** Within human interactions, human drivers show courtesy to others, i.e., they optimize a compound cost function in the form of  $C = C^{self} + \lambda_c C^{court}$  as (8) instead of a selfish one as  $C^{self}$ .

**Independent Variable.** To test our hypothesis, we run two sets of IRL on the same set of human data, but with one different feature. For the selfish cost function  $C^{self}$ , four features are selected as follows:

- speed feature  $f_d$ : deviation of autonomous car's speed compared to the speed limit:

$$f_d = (v - v_d)^2 \quad (19)$$

- comfort features  $f_{acc}$  and  $f_{steer}$ : jerk and steering rate of the autonomous car;
- goal feature  $f_g$ : distance to the target lane:

$$f_g = e^{\frac{d_g}{w_l}}, \quad (20)$$

where  $d_g$  is the Euclidean distance and  $w_l$  is the lane width.

- safety feature  $f_s$ : relative positions with respect to surrounding cars;

$$f_s = \sum_{i=1}^{n_s} e^{-d_i}, \quad (21)$$

where  $n_s$  is the number of surrounding cars and  $d_i, i=0, 1, \dots, n_s$  is the distance to each of them.

For the courtesy-aware cost function  $C = C^{self} + \lambda_c C^{court}$ , we use the same four features as above, plus one additional feature that equals to the courtesy term.

**Dependent Measures.** We measured the similarity between trajectories planned with the learned cost functions and human driving trajectories on the *test* set (another 53 left-lane changing scenarios that are different from the training set from the NGSIM dataset).

## C. Analysis

**Training performance.** The training results are shown in Fig. 7 and Table I. One can see that with the additional courtesy term, better learning performance (in terms of training loss) has been achieved. This is a sanity check: having access to one extra DOF can lead to better training loss regardless, but if it did not that would invalidate our hypothesis.

	$\theta_g$	$\theta_d$	$\theta_{acc}$	$\theta_{steer}$	$\theta_s$	$\lambda_c$
$C^{self}$	1.0	2.08e+04	5.80e+02	3.91e+02	4.37	–
C	1.0	1.96e+02	6.7e+04	2.36e+02	6.53	9.89e+04

TABLE I: The parameters in C learned via IRL

**Trajectory similarity.** Figure 8 shows one demonstration example of the trajectories for a selfish car (grey) and a courteous car (orange), with four surrounding vehicles. The dark blue rectangle is the human driver in our two-agent robot-human interaction system and all other vehicles (cyan) are treated as moving obstacles.

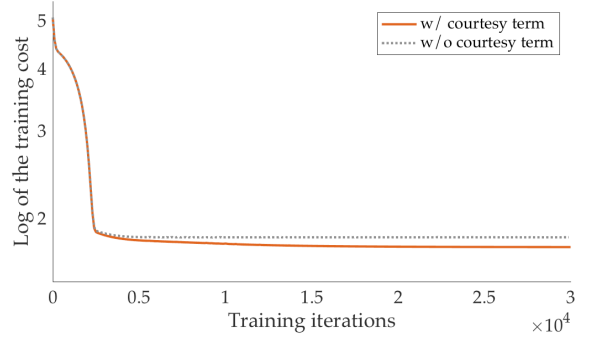


Fig. 7: Training curves for cost functions with and without the courtesy term

It shows that a simulated car with C that includes courtesy manages to reduce its influence on the human driver by choosing a much smoother and less aggressive merging curve, while a car driven by  $C^{self}$  merges in much aggressively.

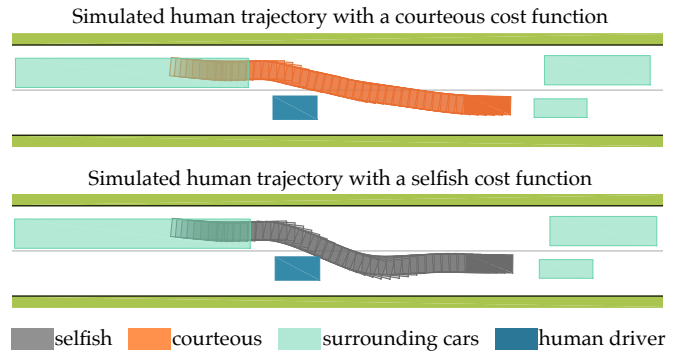


Fig. 8: An example pair of simulated trajectories with courteous (top) and selfish (bottom) cost functions

Results for all 53 left-lane changing test trajectories are given in Fig. 9 (left). To describe the similarities among trajectories, we adopted the Mean Euclidean Distance (MED) [21]. As shown in Fig. 9 (right), the courtesy-aware trajectories are much similar to the ground truth trajectories, i.e., a courteous robot car behaves more human-like. We have also calculated the space headways of the following human driver on the robot car's target lane for all 53 test scenarios, and the statistical results are given in Fig. 9 (middle). Compared to a selfish robot car, a courteous robot car can achieve safer left-lane changing behaviours in terms of following gaps for the human driver behind.

## VI. CONCLUSION

**Summary.** We introduced courteous planning based on the fact that human irrationally care more about additional inconvenience they are brought to by others. Courteous planning enables an autonomous car to take into consideration such inconvenience when evaluating its possible plans. We saw that not only this leads to more courteous robot behavior, but it also helps explain real human driving data, because humans too are likely trying to be courteous.

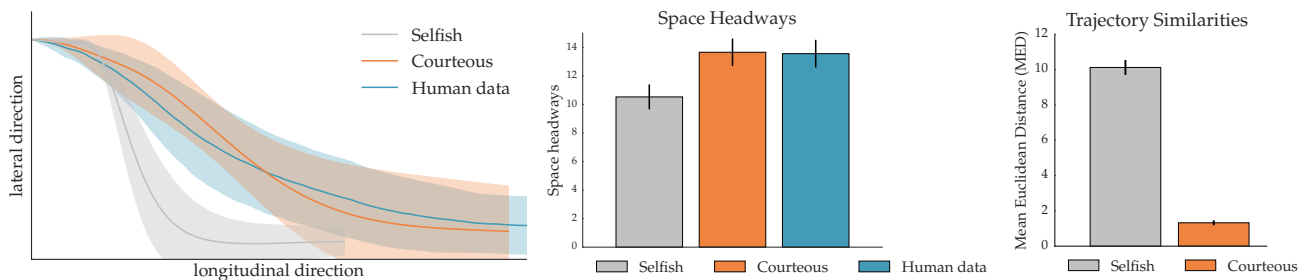


Fig. 9: The courtesy term helps fit test set human driver data significantly better: we can see this from the actual trajectories (left), the following gaps (middle), and the mean euclidean distances from the ground truth human data (right).

**Limitations and Future Work.** Despite the fact that courtesy is not absolute, but relative to how well off the human driver could be, the trade-off between courtesy and selfishness remains a meta-parameter that is difficult to set. In general, defining the right trade-off parameters in the objective function for autonomous cars and robots more broadly remains a challenge. With autonomous cars, this is made worse by the fact that it is not necessarily a good idea to rely on Inverse Reinforcement Learning—this might give us models of human drivers, as it did in our last experiment, but that might not be what we want the car to optimize for.

Further, we studied courtesy with a single human driver to be courteous toward (we had other agents, but the robot did not attempt courtesy toward them). In real life, there will be many people on the road, and it becomes difficult to be courteous to all. To some extent, this is alleviated by our definition of courtesy: it is not maximizing everyone’s utility, but it is minimizing the inconvenience we cause. But further work needs to push courtesy to the limits of interacting with multiple people in cases where it is difficult to be courteous to all.

#### ACKNOWLEDGEMENT

This work was partially supported by Mines Paris-Tech Foundation, “Automated Vehicles—Drive for All” Chair, and NSF CAREER.

#### REFERENCES

- [1] Y. Kuwata, J. Teo, G. Fiore, S. Karaman, E. Frazzoli, and J. P. How, “Real-time motion planning with applications to autonomous urban driving,” *IEEE Transactions on Control Systems Technology*, vol. 17, no. 5, pp. 1105–1118, 2009.
- [2] Z. Liang, G. Zheng, and J. Li, “Automatic parking path optimization based on bezier curve fitting,” in *Automation and Logistics (ICAL)*, 2012 IEEE International Conference on. IEEE, 2012, pp. 583–587.
- [3] W. Zhan, J. Chen, C. Y. Chan, C. Liu, and M. Tomizuka, “Spatially-partitioned environmental representation and planning architecture for on-road autonomous driving,” in *2017 IEEE Intelligent Vehicles Symposium (IV)*, June 2017, pp. 632–639.
- [4] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, “Social LSTM: Human trajectory prediction in crowded spaces,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 961–971.
- [5] W. Zhan, C. Liu, C. Y. 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.
- [6] M. Shimosaka, K. Nishi, J. Sato, and H. Kataoka, “Predicting driving behavior using inverse reinforcement learning with multiple reward functions towards environmental diversity,” in *Intelligent Vehicles Symposium (IV)*, 2015 IEEE. IEEE, 2015, pp. 567–572.
- [7] S. Levine and V. Koltun, “continuous inverse optimal control with locally optimal examples,” in *the 29th International Conference on Machine Learning (ICML-12)*, 2012.
- [8] G. R. de Campos, P. Falcone, and J. Sjöberg, “Autonomous co-operative driving: a velocity-based negotiation approach for intersection crossing,” in *Intelligent Transportation Systems (ITSC)*, 2013 16th International IEEE Conference on. IEEE, 2013, pp. 1456–1461.
- [9] M. Hafner, D. Cunningham, L. Caminiti, and D. Del Vecchio, “Automated vehicle-to-vehicle collision avoidance at intersections,” in *Proceedings of world congress on intelligent transport systems*, 2011.
- [10] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard, “Socially compliant mobile robot navigation via inverse reinforcement learning,” *The International Journal of Robotics Research*, vol. 35, no. 11, pp. 1289–1307, 2016.
- [11] 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.
- [12] M. Bahram, A. Lawitzky, J. Friedrichs, M. Aeberhard, and D. Wollherr, “A Game-Theoretic Approach to Replanning-Aware Interactive Scene Prediction and Planning,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 3981–3992, June 2016.
- [13] 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.
- [14] S. S. A. D. Dorsa Sadigh, Shankar S. Sastry, “Information gathering actions over human internal state,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, October 2016, pp. 66–73.
- [15] A. Tversky and D. Kahneman, “Advances in prospect theory: Cumulative representation of uncertainty,” *Journal of Risk and uncertainty*, vol. 5, no. 4, pp. 297–323, 1992.
- [16] P. Abbeel and A. Y. Ng, “Apprenticeship learning via inverse reinforcement learning,” in *Proceedings of the twenty-first international conference on Machine learning*. ACM, 2004, p. 1.
- [17] B. D. Ziebart, A. L. Maas, J. A. Bagnell, and A. K. Dey, “Maximum entropy inverse reinforcement learning,” in *AAAI*, vol. 8. Chicago, IL, USA, 2008, pp. 1433–1438.
- [18] P. Abbeel and A. Y. Ng, “Inverse reinforcement learning,” in *Encyclopedia of machine learning*. Springer, 2011, pp. 554–558.
- [19] V. Alexiadis, J. Colyar, J. Halkias, R. Hranac, and G. McHale, “The Next Generation Simulation Program,” *Institute of Transportation Engineers. ITE Journal*, Washington, vol. 74, no. 8, pp. 22–26, Aug. 2004.
- [20] “https://julialang.org/”
- [21] J. Quehl, H. Hu, O. S. Tas, E. Rehder, and M. Lauer, “How good is my prediction? finding a similarity measure for trajectory prediction evaluation,” in *2017 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 120–125.