Encoding Defensive Driving with Noncooperative Differential Games

David Fridovich-Keil and Chih-Yuan Chiu and Claire J. Tomlin
University of California, Berkeley
Email: {dfk, chihyuan_chiu, tomlin}@berkeley.edu

***Abstract*—This paper presents a novel formulation of robustness modeled on defensive driving. Like adversarial safety formulations, our method is also based in differential game theory; however, our approach is general-sum and applies to N-player scenarios. We present an overview of our method and preliminary results for two interactive driving examples in which the user may pre-specify the severity of avoidance maneuvers.**

I. Introduction

In designing autonomous systems, practitioners employ different notions of safety and robustness. Broadly, there are two prevailing definitions of "safety:" *adversarial* and *probabilistic* constraint satisfaction. In this work, we introduce a third, distinct notion of safety. Like adversarial formulations, our work is based upon noncooperative differential game theory; however, unlike such methods, our approach is not equivalent to a single two-player zero-sum differential game, and naturally extends to an arbitrary number of agents.

Our work is heavily based upon the literature in differential game theory and adversarial reachability. For a more complete reference on differential games, please see Isaacs [6, 7] and Başar and Olsder [1], and for early formulations of Nash equilibria in these games, please see [11, 10]. Adversarial reachability methods [3, 8, 4, 2] pose a given optimal control problem as a zero-sum differential game whose Nash equilibrium satisfies a Hamilton-Jacobi-Isaacs partial differential equation. These approaches seek to identify when a system's state can be driven, despite worst-case bounded disturbance, toward one set and away from another set.

Like such methods, we study safety and robustness in the context of differential games; however, in our case the players are the ego agent and an arbitrary number of other agents. We divide the time horizon into two parts: an *adversarial* part followed by a *cooperative* part. During the adversarial portion, the ego agent presumes that other agents wish to harm it and encodes such behavior in the cost structure of the game, and during the cooperative portion of the time horizon it presumes that other agents will try to help it (e.g., to avoid collision).

II. Problem Formulation

Suppose one "ego" agent with input u_1 interacts with other agents (input $u_i \in \mathbb{R}^{m_i}, i \in \{2 \dots N\}$) over a time horizon of length T, each agent has a distinct objective J_i which each depends upon the strategies $\{\gamma_i\}_{i=1}^N$ of all agents, and the state follows known dynamics f. This is precisely a differential game. We presume that each player's cost is expressed as $J_i(\gamma_{1:N}) = \int_0^T g(t, x, u_{1:N}) dt$, where the strategy of each player is taken to be a state feedback strategy, i.e., $u_i(t) = \gamma_i(t, x)$.

Our novel formulation of robustness is best understood through the lens of defensive driving, in which one ought to imagine all other cars as momentarily distracted. To encode this "imagined scenario," the ego agent divides the time horizon into two sections such that $T = T_{\text{adv}} + T_{\text{coop}}$ and reimagines the cost for non-ego agents (indices $2 \dots N$) as:

$$\tilde{J}_i := \int_0^{T_{\text{adv}}} g_{\text{adv},i}(x, u_{1:N}) dt + \int_{T_{\text{adv}}}^T g_{\text{coop},i}(x, u_{1:N}) dt\,. \quad (1)$$

Such problems may be solved approximately to local feedback Nash equilibria in real time using the recent iterative linear-quadratic game algorithm of [5]. An adversarial running cost, $\{g_{\text{adv},i}\}_{i=2}^N$, encodes adversarial behavior for the non-ego agents for the first part of the time horizon, while $\{g_{\text{coop},i}\}_{i=2}^N$ encodes cooperative behavior for the remainder of the time horizon. In effect, by constructing the game this way, the ego agent imagines other agents to be "momentarily distracted" during T_{adv} and will act "defensively." Note that we have dropped the dependence of these costs on t for clarity.

The custom iterative linear-quadratic game solver from [5] also accounts for inequality constraints on $x(t)$ and $u_i(t)$ via barrier methods. Such constraints force the ego agent to bear responsibility for satisfying joint state constraints (e.g., non-collision). All agents also must satisfy individual constraints (e.g. staying within a range of speeds). Pang and Scutari [9] provide a thorough treatment of constraints in games.

III. Methods

To test this construction, we simulate two traffic encounters that involve significant interaction (see Sec. IV), in which a responsible human driver would likely drive defensibly. Our method attempts to capture the spectrum of this "defensive" behavior. In each setting every player (here, car) has augmented bicycle dynamics, i.e.:

$$\begin{aligned} \dot{p}_{x,i} &= v_i \sin\theta_i, & \dot{v}_i &= a_i, \\ \dot{p}_{y,i} &= v_i \cos\theta_i, & \dot{\phi}_i &= \omega_i, \\ \dot{\theta}_i &= (v_i/L_i)\tan\phi_i, & \dot{a}_i &= j_i, \end{aligned} \quad (2)$$

where $x = (p_{x,i}, p_{y,i}, \theta_i, v_i, \phi_i, a_i)_{i=1}^N$ represents each vehicle's position, heading, front wheel angle, speed, and acceleration, and $u_i = (\omega_i, j_i)$ represents each vehicle's front wheel rate and tangent jerk. L_i is each player's inter-axle distance.

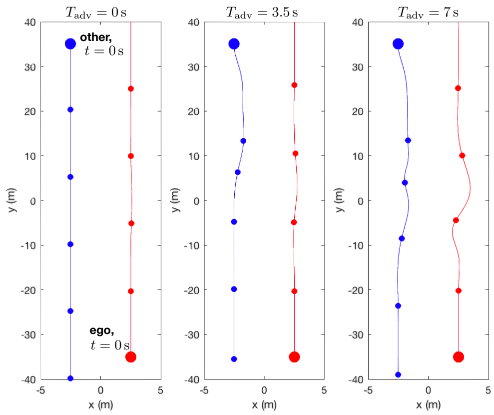

Fig. 1. Oncoming example. Ego (red) and oncoming (blue) cars perform more extreme maneuvers as T_{adv} increases. Dots are placed at 3 s intervals.

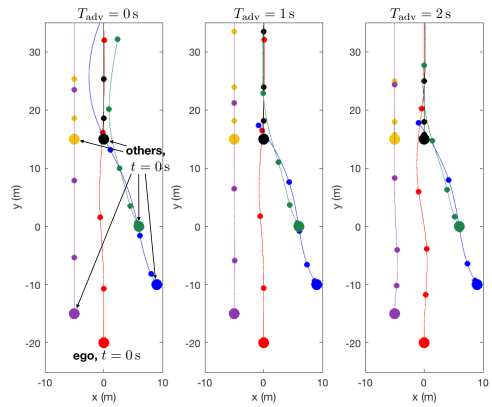

Fig. 2. Merging example, with a similar pattern of more extreme maneuvers. Again, the ego player is colored red and dots are placed at 3 s intervals.

We define $g_{\text{adv},i}$ and $g_{\text{coop},i}$ as a weighted combination of the following functions, each of which encourages the specified behavior. We use the shorthand $p_i = (p_{x,i}, p_{y,i})$ for the position of each car, $d_{\ell_i}(p_i) = \min_{p_\ell \in \ell_i} \|p_\ell - p_i\|$ for the distance between a car and the corresponding lane centerline ℓ_i, and d_{prox} for a constant desired minimum proximity between cars:

$$\text{input: } u_i^T R_{ii} u_i \tag{3}$$

$$\text{lane center: } d_{\ell_i}(p_i)^2 \tag{4}$$

$$\text{ideal speed: } (v_i - v_{\text{ref},i})^2 \tag{5}$$

$$\text{cooperative: } \mathbf{1}\{\|p_i - p_j\| < d_{\text{prox}}\}(d_{\text{prox}} - \|p_i - p_j\|)^2 \tag{6}$$

$$\text{adversarial: } \|p_i - p_j\|^2 \,. \tag{7}$$

Recall that, for non-ego agents, the "adversarial" cost is only present during T_{adv} and the "cooperative" cost is present thereafter. We also enforce inequality constraints (where d_{lane} is the lane half-width, and $\underline{v}_i$ and $\overline{v}_i$ are speed limits):

$$\text{proximity: } \|p_i - p_j\| > d_{\text{prox}} \tag{8}$$

$$\text{lane: } |d_{\ell_i}(p_i)| < d_{\text{lane}} \tag{9}$$

$$\text{speed range: } \underline{v}_i < v_i < \overline{v}_i \,, \tag{10}$$

where the "proximity" constraint is enforced for only the ego agent but the other constraints are enforced for all agents, and all are enforced over the entire time horizon.

For all tests we use a time horizon $T = 15\,\text{s}$ and discretize time (following [5] and [1]) at $0.1\,\text{s}$ intervals. We intend that the ego agent follow its own Nash strategy for a short time and then re-solve the game in a receding horizon; we shall study the implications on robustness in future work.

IV. Results

We present preliminary results for two different traffic scenarios in which a responsible human driver would likely drive defensively. First, we consider a simple situation involving oncoming vehicles on a straight road, as a proof of concept. Later, we shall consider a more complicated highway ramp merging example. Each instantiation is solved in well under $0.75\,\text{s}$ in single-threaded operation on a standard laptop, via an open-source, purely C++ algorithm cubic in the number of players [5]. This performance indicates real-time capabilities which will be explored in future work on hardware.

A. Oncoming example

In this example, the ego car is traveling North on a straight road when it encounters another car traveling South. Since the road has a lane in each direction, "ideally" the ego vehicle would not deviate too far from its lane or speed. However, to drive more defensively, the ego vehicle should plan as though the oncoming Southbound car were to act noncooperatively. Our method encodes precisely this type of defensive planning. Fig. 1 shows the planned trajectories that emerge for increasing T_{adv}. As shown, the ego vehicle (red) imagines more aggressive maneuvers for itself and the oncoming car as T_{adv} increases. Note, however, that these are merely *imagined* trajectories and that (a) the ego vehicle will only begin executing this trajectory before finding another in a receding time horizon, and (b) the oncoming vehicle will make its own decisions and will *not* generally follow this "partially adversarial" trajectory. With increasing T_{adv}, the ego car imagines an increasingly adversarial encounter and acts more and more defensively as a result. In practice, the user or system designer would select a suitable T_{adv} before operation, e.g., by choosing the largest T_{adv} such that the solution deviates from a nominal solution with $T_{\text{adv}} = 0$ sufficiently little.

B. Merging example

Here, we present a six-player example in which the ego car must merge from an on-ramp into highway traffic. Fig. 2 shows the resulting approximate local Nash equilibria of the defensive driving game for increasing T_{adv}. As before, the ego vehicle imagines increasingly extreme pursuit-evasion behavior as T_{adv} grows.

V. Conclusions

We have presented a novel formulation of robustness modeled on defensive driving. Our method draws upon earlier work in differential game theory that forms the basis for adversarial safety methodologies common in Hamilton-Jacobi reachability, and uses new solution methods to solve these "defensive" problems in real-time. We are eager to implement this method in hardware and test its performance in a receding time horizon with other (human) drivers.

REFERENCES

[1] Tamer Başar and Geert Jan Olsder. *Dynamic noncooperative game theory*, volume 23. SIAM, 2nd edition, 1999.

[2] Somil Bansal, Mo Chen, Sylvia Herbert, and Claire J Tomlin. Hamilton-jacobi reachability: A brief overview and recent advances. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 2242–2253. IEEE, 2017.

[3] L. C. Evans and P. E. Souganidis. Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations. *Indiana University mathematics journal*, 33(5):773–797, 1984.

[4] Jaime F Fisac and S Shankar Sastry. The pursuit-evasion-defense differential game in dynamic constrained environments. In *54th Conference on Decision and Control (CDC)*, pages 4549–4556. IEEE, 2015.

[5] David Fridovich-Keil, Ellis Ratner, Anca D Dragan, and Claire J Tomlin. Efficient iterative linear-quadratic approximations for nonlinear multi-player general-sum differential games. *arXiv preprint arXiv:1909.04694*, 2019.

[6] Rufus Isaacs. Games of pursuit. Technical report, Rand Corporation, 1951.

[7] Rufus Isaacs. *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*. Courier Corporation, 1999.

[8] Ian M Mitchell, Alexandre M Bayen, and Claire J Tomlin. A time-dependent hamilton-jacobi formulation of reachable sets for continuous dynamic games. *Transactions on Automatic Control*, 50(7):947–957, 2005.

[9] Jong-Shi Pang and Gesualdo Scutari. Nonconvex games with side constraints. *SIAM J. on Optimization*, 21 (4):1491–1522, December 2011. ISSN 1052-6234. doi: 10.1137/100811787. URL https://doi.org/10.1137/100811787.

[10] Alan W Starr and Yu-Chi Ho. Further properties of nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3(4):207–219, 1969.

[11] Alan W Starr and Yu-Chi Ho. Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3(3):184–206, 1969.