

Game-Theoretic Planning for Autonomous Driving among Risk-Aware Human Drivers

Rohan Chandra^{*1}, Mingyu Wang^{*2}, Mac Schwager³ and Dinesh Manocha^{1,4}

Abstract—We present a novel approach for risk-aware planning with human agents in multi-agent traffic scenarios. Our approach takes into account the wide range of human driver behaviors on the road, from aggressive maneuvers like speeding and overtaking, to conservative traits like driving slowly and conforming to the right-most lane. In our approach, we learn a mapping from a data-driven human driver behavior model called the CMetric to a driver’s entropic risk preference. We then use the derived risk preference within a game-theoretic risk-sensitive planner to model risk-aware interactions among human drivers and an autonomous vehicle in various traffic scenarios. We demonstrate our method in a merging scenario, where our results show that the final trajectories obtained from the risk-aware planner generate desirable emergent behaviors. Particularly, our planner recognizes aggressive human drivers and yields to them while maintaining a greater distance from them. In a user study, participants were able to distinguish between aggressive and conservative simulated drivers based on trajectories generated from our risk-sensitive planner. We also observe that aggressive human driving results in more frequent lane-changing in the planner. Finally, we compare the performance of our modified risk-aware planner with existing methods and show that modeling human driver behavior leads to safer navigation.

I. INTRODUCTION

Risk-aware planning involves sequential decision-making in dynamic and uncertain environments, where agents must consider the risks associated with their actions and corresponding costs and rewards [6]. *Risk-seeking* agents are willing to take lower expected reward in exchange for a higher reward variance (more risk), while *risk-averse* agents are willing to take a lower expected reward in exchange for lower reward variance (less risk). Agents that are risk-averse or risk-seeking are collectively referred to as risk-aware. Human drivers are risk-aware by nature [24], [31], [22]. For example, aggressive drivers frequently speed, overtake, and perform sharp cut-ins, whereas conservative drivers drive more cautiously. To navigate successfully among human drivers, autonomous vehicles (AVs) must identify the risk

* denotes equal contribution to the work.

Wang and Schwager are supported in part by ONR grant N00014-18-1-2830. Chandra and Manocha are supported by ARO Grants W911NF1910069, W911NF2110026, U.S. Army Grant No. W911NF2120076 and, Semiconductor Research Corporation (SRC) and Intel. Toyota Research Institute provided funds to support this work.

¹Department of Computer Science, University of Maryland, USA {rchandr1, dmanocha}@umd.edu

²Department of Mechanical Engineering, Stanford University, Stanford, USA mingyuw@stanford.edu

³Department of Aeronautics and Astronautics, Stanford University, Stanford, USA schwager@stanford.edu

⁴Department of Electrical and Computer Engineering, University of Maryland, USA schwager@stanford.edu

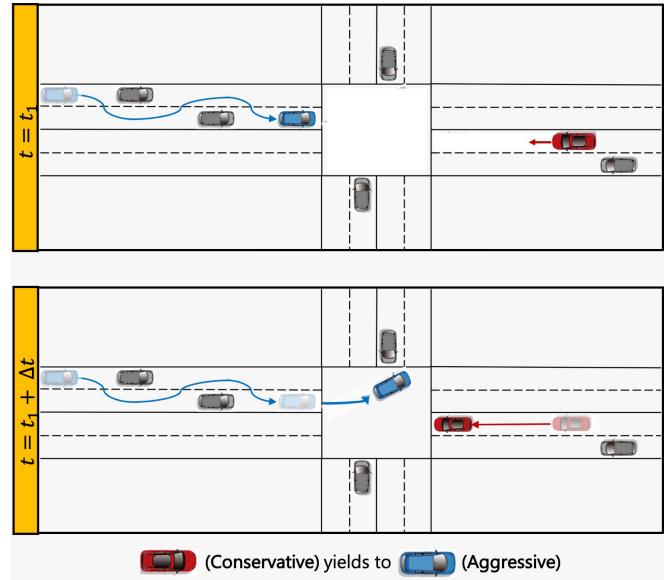


Fig. 1: **Risk-aware planning with human agents:** We present a novel risk-aware planner that takes into account human driver behavior. In the first step (*top*), we characterize the human agent as aggressive (blue agent) with the ego-vehicle as a conservative agent (red). In the second step (*bottom*), we derive the corresponding risk sensitivity of the human agent and generate a game-theoretically optimal and safe risk-aware trajectory for the red agent that suggests the red agent yields to the blue human driver allowing it to cross first).

preferences of human drivers online, and predict and plan future motion with the risk preferences of all agents in mind, including the AV’s own risk preferences.

The most common risk measures utilized in risk-sensitive planning are entropic risk [15] and conditional value at risk (CVaR) [32]. A popular approach for risk-aware planning in multi-agent traffic scenarios is to model risk-aware agent interactions via dynamic games [45] wherein agents act while considering their impact on other agents as well as the intentions of the other agents. In [45], the authors compute the Nash equilibrium solution of the game by iteratively solving a set of LEQ equations [48], [14]. The main benefits of this approach compared to prior risk-aware planning methods include improved time-to-goal and, more importantly, generation of emergent behaviors. For instance, risk-averse agents learn to maintain a greater distance from risk-seeking agents and generally yield more frequently to risk-seeking agents at intersections, at roundabouts, and during merging.

Despite its performance and benefits, the main drawback of the approach proposed by [45] is that it does not model the risk tolerance of human drivers, as it assumes the AV knows

the synthetically chosen risk tolerances for all other driving agents). Extending game-theoretic risk-aware planning to human drivers will allow AVs to act more confidently around human drivers and reduce time-to-goal via more efficient and safer navigation. Estimating the risk tolerances of human drivers, however, requires computationally tractable human driver behavior models that can characterize drivers.

Some of the state-of-the-art approaches for modeling human driver behavior [30], [20], [37], [34] are data-driven and require a large volume of clean training data. These methods classify behaviors as aggressive and conservative [30], [20] or selfish and altruistic [37]. In contrast, deterministic models [7] do not require data and assign a real-valued score to each agent to quantify its behavior. These approaches can be integrated with risk-aware planning frameworks to incorporate planning for human agents.

Main Contributions: We propose a novel approach for risk-aware planning in multi-agent traffic scenarios that takes into account human driver behaviors. We extend an existing risk-aware planner [45] by incorporating interactions with human drivers using a data-driven human driver behavior model [7]. We derive a linear mapping between the driver behavior and risk tolerance (Equation 2), which serves as the key component of our proposed approach.

To evaluate our approach, we validate the mapping between driver behavior and risk tolerance by measuring the number of lane changes, and test the accuracy of this model via K-Means clustering. Our results show that aggressive human driving results in more frequent lane changing. We confirm that the final trajectories obtained from the risk-aware planner generate emergent behaviors. We measure the yield % and minimum distance between human drivers at intersections, at roundabouts, and during merging where we observe that conservative drivers generally yield to aggressive drivers while maintaining a greater distance from them. We also conduct a user study in which we show that users are able to distinguish between aggressive and conservative trajectories generated by the planner.

Finally, we compare our modified risk-aware planner with existing planners that do not model human drivers and show that modeling human drivers results in safer navigation. Specifically, [45] (and similar planners) assign a fixed neutral risk tolerance to human drivers and the ego-vehicle generates to the human driver accordingly. However, when the human driver is, in fact, either aggressive or conservative, then we show that the error (absolute value of minimum relative distance between the agents) increases by 10%. We explain this in detail in Section V-D.

II. RELATED WORK

A. Risk-Aware Planning

Risk sensitivity-based planning [27], [10], [11], [36], [8] considers the risk associated with the actions of agents to avoid unsafe situations. The most common risk measures utilized in risk-sensitive planning are entropic risk [15] and conditional value at risk (CVaR) [32]. Entropic risk has been widely used in optimal control due to its simplicity

and tractability [47], while CVaR has recently been incorporated in trajectory optimization due to its interpretability [11]. Risk-aware planning has been used extensively in autonomous underwater vehicles [28], ground vehicles [38], [42], and unmanned aerial vehicles (UAVs) [9]. While the CVaR risk model has been used in the latter two cases, [28] used the entropic measure of risk. In addition to CVaR and the entropic models, several other models are also used in various applications such as the dynamic risk density function for collision avoidance [39] and semantic maps for simultaneous localization and mapping (SLAM) [40], [41]

B. Data-Driven Methods for Driver Behavior Prediction

Data-driven methods broadly follow two approaches. In the first approach, various machine learning algorithms such as clustering, regression, and classification predict or classify driver behavior as either aggressive or conservative. These methods have been studied in traffic psychology and the social sciences [44], [17], [16], [12], [19], [13], [18], [4], [1], [21], [3], [5], [26], [26], [25], [29], [43], [46]. So far, there has been relatively little work to improve the robustness and ability to generalize to different traffic scenarios, steps that require ideas from computer vision and robotics.

The second approach uses trajectories to learn reward functions for human behavior using inverse reinforcement learning (IRL) [30], [20], [37], [34]. IRL-based methods, however, have certain limitations. IRL requires large amounts of training data, and the learned reward functions are unrealistically tailored towards scenarios only observed in the training data [30], [34]. For instance, [30] requires 32 million data samples for optimum performance. Additionally, IRL-based methods are sensitive to noise in the trajectory data [37], [34]. Consequently, current IRL-based methods are restricted to simple and sparse traffic conditions.

III. PRELIMINARIES AND PROBLEM FORMULATION

In this section, we briefly summarize the CMetric algorithm [7] for behavior modeling. We will also define the problem statement of risk-aware planning.

A. CMetric: Modeling Human Driver Behavior

We briefly summarize the CMetric algorithm [7] because it provides an objective measure of aggressiveness based on driving maneuvers such as speeding, overtaking, and so on. To determine if an agent is aggressive or conservative, the algorithm begins by reading the trajectories of the agent and surrounding vehicles via cameras or lidars during a time period ΔT . The trajectory of an agent i is represented by

$$\Xi^i = \{x_t^i \mid t = t_0, t_1, \dots, t_0 + T\}.$$

In CMetric, these trajectories are represented via weighted undirected graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ in which the vertices denote the positions for each agent and the edges correspond to the distances between agents. The algorithm proceeds by using these graphs to model the likelihood and intensity of driving behavior indicators like speeding, overtaking, sudden zigzagging, and lane-changes via the closeness and degree

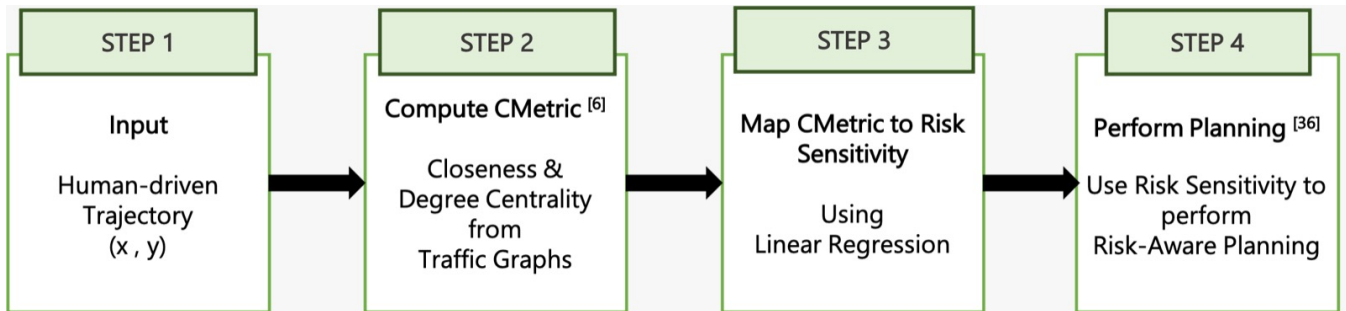


Fig. 2: **Overview:** We outline the steps of our approach for risk-sensitive planning with human agents. (Block 1) In the first step, humans drive a vehicle in an OpenAI simulator using a keyboard. (Block 2) The next step is to compute the CMetric value for each human user that encodes the aggressive (or conservative) nature of the human driver via certain indicators such as speeding, overtaking, and zigzagging through the simulated traffic. (Block 3) We map the human drivers' behaviors to their risk sensitivity using a linear transformation. (Block 4) Finally, based on the risk sensitivity, we perform game-theoretic risk-aware planning using the planner developed by Wang et al. [45].

centrality functions [33] represented by $\Phi : \mathcal{G} \rightarrow \mathbb{R}$. These behavior indicators determine whether an agent is aggressive or conservative [35]. The behavior profile for agent i is denoted by ζ_i and is computed as,

$$\zeta_i(\Xi_{\Delta t}) = \Phi(\mathcal{G}), \quad (1)$$

where $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is constructed using $\Xi_{\Delta t}$.

B. Problem Formulation

We consider N agents consisting of a mixture of human drivers and AVs with the system dynamics defined by Equation 3, and we define the cost function for each agent by Equation 5. A human driver is simulated using a user-controlled keyboard with the following features: acceleration, braking, and lane changing. For simplicity, we test with one human driver, but our approach can work with more than one human driver. We further assume that agents are non-ideal in that agents are not provided the risk tolerance of other agents. The input to our approach consists of the state and control signals of every agent at time t . Then, our goal is to compute the Nash equilibrium trajectories for all agents. The trajectories for the human agents are predictions, while the trajectories for the AVs can be executed in a receding horizon planning loop. Finally, none of the agents are assumed to follow constant velocity models.

IV. ALGORITHM

We describe our algorithm (Figure 2). The first step is to read the trajectories for every agent over a finite horizon T , denoted by Ξ_T . These trajectories correspond to human agents. The second step is to compute the CMetric value, ζ , for each agent during T ; the CMetric value encodes the aggressive (or conservative) nature of the driver via certain indicators such as speeding, overtaking, and zigzagging. The third step consists of mapping an agent's CMetric to their risk sensitivity. This is performed using a linear transformation obtained by simple linear regression. This is discussed in detail in Section IV-A. Finally, based on the risk sensitivity, we perform game-theoretic risk-aware planning using the planner developed by Wang et al. [45].

A. CMetric to Risk Sensitivity

We denote the risk sensitivity parameter by θ . We first compute a linear mapping $\mathcal{M} : Z \rightarrow \Theta$. Since both $\zeta \in Z$ and $\theta \in \Theta$ are scalars, we can use simple one dimensional linear regression to estimate \mathcal{M} . We create a training dataset by first generating trajectories corresponding to a fixed array of risk sensitivity values ranging from -5.0 (risk-seeking) to $+5.0$ (risk-averse). We denote these risk sensitivity values as $\hat{\theta}$ to indicate they are training values. We then evaluate the CMetric values, also represented using $\hat{\zeta}$, corresponding to each of these trajectories using the algorithm described in the previous section. The risk sensitivity and CMetric pair constitute the training dataset on which we apply linear regression to estimate linear coefficients β_0 and β_1 . \mathcal{M} is then defined as follows,

$$\mathcal{M}(\zeta) = \beta_1 \zeta + \beta_0, \quad (2)$$

where ζ is the CMetric value of a human agent at test time.

B. Risk-Aware Planning

The system dynamics for each agent are given by,

$$x_{t+1} = A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2 + w_t. \quad (3)$$

To simplify notation, we describe a two-player system, although our approach can easily generalize to n agents. $x_t = [x_t^1, x_t^2] \in X$ represents the system state and $x_t^i = [p_x^i, p_y^i, v_x^i, v_y^i]$ denotes the position (in meters) and velocity (in meters/second) of an agent. $u_t^1 = a_1, u_t^2 = a_2 \in U$ are the control inputs for both agents denoting the acceleration of both agents, $w_t \sim \mathcal{N}(0, W_t)$ is the system noise, and A_t, B_t^1, B_t^2 are fixed matrices of appropriate dimensions. An agent incurs the following cost during a finite horizon T :

$$\Psi^i = \sum_{t=0}^{T-1} \left[\frac{1}{2} x_t^T Q_t^i x_t + l_t^i x_t + \frac{1}{2} \sum_j u_t^{jT} R_t^{ij} u_t^j \right] + \frac{1}{2} x_T^T Q_T^i x_T + l_T^i x_T, \quad (4)$$

where $Q_t \succeq 0$ and $R_t \succ 0$. To model risk, we use the exponential risk cost function used in [45],

$$J(\Psi) = \frac{1}{\mathcal{M}(\zeta)} \log \mathbb{E} \left[e^{\mathcal{M}(\zeta)\Psi} \right] = R_{\mathcal{M}(\zeta)}(\Psi), \quad (5)$$

where $\mathcal{M}(\zeta)$ is the risk tolerance of a human driver.

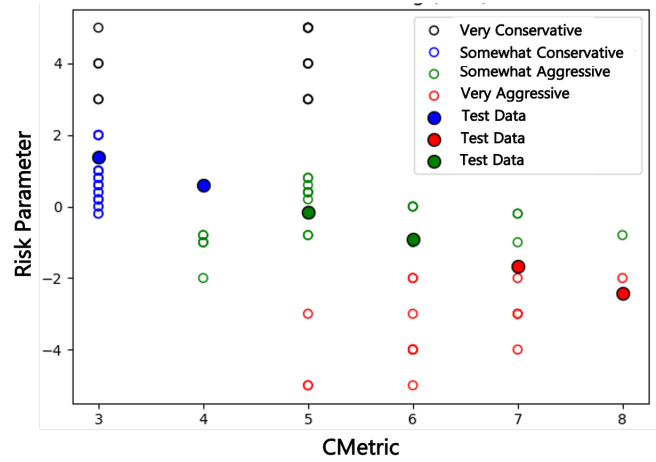
Remark 1: The difference between Equation 5 and the risk cost function described in [45] is that the risk parameter in the latter work includes a fixed value for every agent, whereas in this work, we automatically generate the risk parameter for human agents in a data-driven fashion.

The optimal strategies for each player can be obtained by minimizing $J(\Psi^i)$ for each agent i and obtaining the Nash equilibrium using Riccati recursion [2, Chap. 6]. However, Equation 5 is constrained by the fact that $\mathcal{M}(\zeta)$ is bounded. If the $\mathcal{M}(\zeta)$ is too low or too high, then the cost function value approaches ∞ , also known as “neurotic breakdown” [45]. Due to the data-driven nature of Equation 5, in order to ensure optimality, certain traffic parameters such as traffic density is assumed, since they affect the CMetric value [7], and by Equation 2, the risk sensitivity of the human agent used in Equation 5.

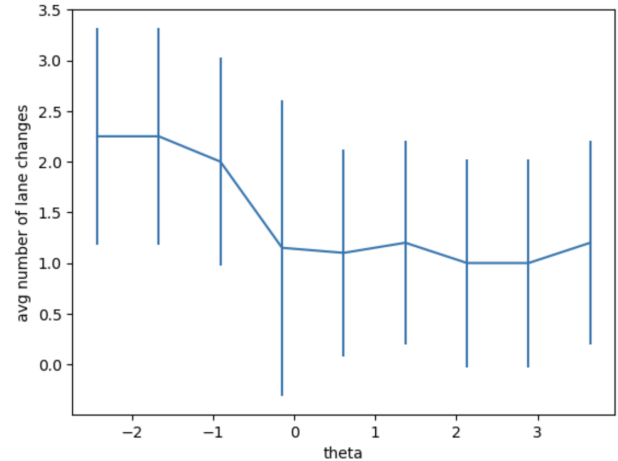
V. EXPERIMENTS AND RESULTS

In this section, we present the results of extensive experiments testing the accuracy of the linear mapping between human driver aggressiveness and risk tolerance. We then evaluate the emergent behaviors associated with the final trajectories generated by the iterative risk sensitive game theoretic solver, compare with [45], which is chosen as the baseline (in which human driver behavior is ignored), and finally, discuss using alternative human driver behavior models. All experiments are performed using a 12-core 2.60GHz Intel i7 processor. We conduct open-loop tests that follow the pipeline outlined in Figure 2. We use the OpenAI traffic simulator [23] to compute the CMetric values representing human driver behavior and the python-based controller provided by Wang et al. [45] to generate the final trajectories based on the risk tolerances obtained from the corresponding CMetric values. The configurations of both the simulator and the controller (which include the dynamics of the vehicles, traffic density, number of lanes etc.) are kept identical so that all vehicles generated using the controller are tracked in the simulator.

We compute the CMetric values of the human driver in a highway scenario since we require a fixed duration of time (5s) during which we must observe the vehicle’s trajectory and its interaction with other vehicles. For the risk-aware trajectory controller, we consider a merging scenario where a human agent must merge onto a highway with another human agent in the target merging lane. Here, the human agent is equivalent to an agent whose risk sensitivity value is obtained from the CMetric value. We assume vehicles follow the center line in their current driving lanes and only consider the vehicle’s speed to finish the merging maneuver. In other words, we assume a steering controller will be executed separately for each car to remain in its lane.



(a) **KMeans clustering:** As human driver behavior becomes more aggressive (higher CMetric value), the risk parameter θ tends to decrease. This trend is consistent with the definition of risk sensitivity in [45]. Furthermore, our approach can effectively categorize the risk sensitivity of new human drivers (solid colored points) regardless of traffic density, number of lanes, etc.

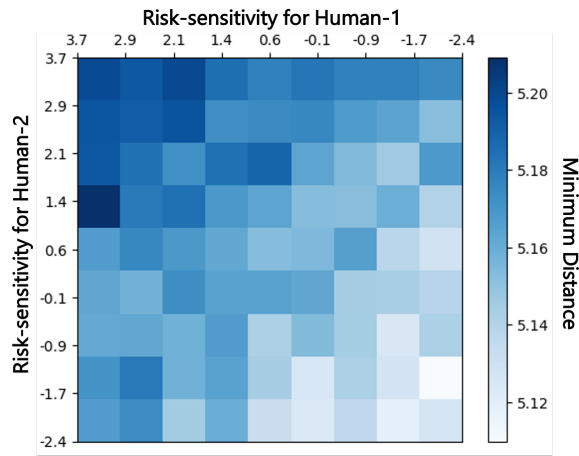


(b) **Number of lane changes:** Aggressive drivers ($\theta < 0$) yield a greater number of lane changes than conservative drivers ($\theta > 0$).

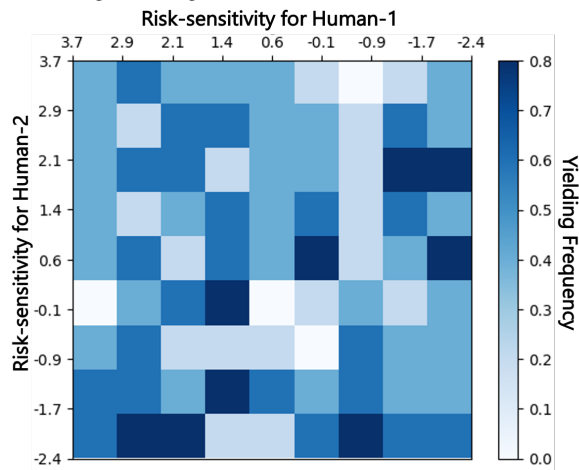
Fig. 3: We highlight the relationship between the CMetric value and the risk parameter θ . Refer to Section V-A for further details.

A. Verifying the accuracy of \mathcal{M}

In Figure 3a, we plot the risk parameter θ (y-axis) obtained from a given CMetric value (x-axis) via the linear mapping. When computing the risk parameters corresponding to each CMetric value, we vary the simulation configuration (traffic density, number of lanes etc.) to include a range of environments. This results in θ belonging to a range (as opposed to a fixed value). This is desirable since, in practice, the traffic will vary according to place and time. The risk parameters are clustered into four categories: “very conservative”, “conservative”, “aggressive”, and “very aggressive”. Each cluster is identified by a color. The empty circles are training data. The goal of this experiment is to cluster a test set of CMetric values (solid-colored points) based on their risk sensitivity. The test data are generated by a human driving the OpenAI simulator [23] in a randomly selected environment consisting of eleven vehicles and four lanes. The



(a) **Minimum distance:** Darker colors (indicating larger minimum distance) corresponding to two risk-averse human agents and lighter colors (indicating smaller minimum distance) corresponding to two risk-seeking human agents.

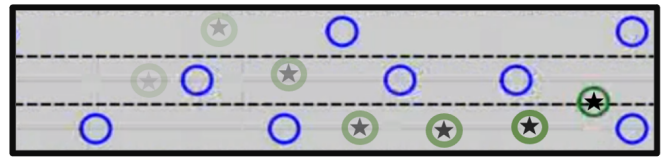


(b) **Yielding behaviors:** Darker colors (indicating higher likelihood of yielding) corresponding to interactions between a risk-seeking agent and risk-averse agent. As the risk tolerances for both the human drivers are data-driven, and therefore noisy, both agents are adapting to the other. As a result, when either agent is risk-averse, we see higher yielding likelihood (darker colors).

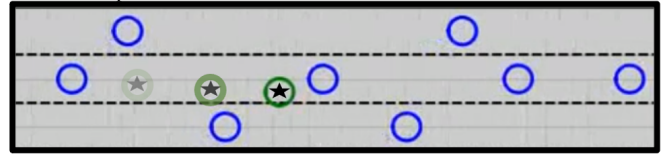
Fig. 4: We demonstrate a range of emergent behaviors. The relative risk sensitivity between the two human agents determines the minimum distance between the two agents and the yielding outcome.

results (Figure 3a) demonstrate that given the CMetric value, the linear regression mapping can accurately identify the risk sensitivity among a wide range of traffic configurations.

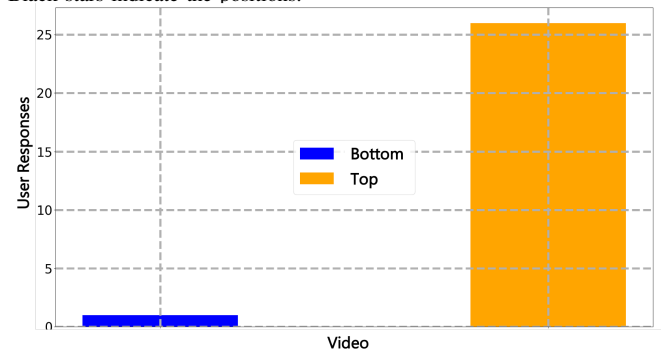
Another metric we use to validate the correlation between the CMetric and corresponding risk sensitivity is the average number of lane changes. Based on the final trajectories generated from the risk sensitivity parameter (obtained from corresponding CMetric values) and using the controller provided by [45], we measure the average number of lane changes made by the ego-vehicle. The reason for using average number of lane changes as a metric is that aggressive drivers change lanes more frequently than non-aggressive and conservative drivers. The aim of the experiment, therefore, is to check if an aggressive human-driven vehicle (modeled using the keyboard of the OpenAI



(a) (top) Risk-aware planner generates an aggressive trajectory for the green agent consisting of zigzagging, overtaking, and overspeeding. Black stars indicate the positions.



(b) (bottom) Risk-aware planner generates a conservative trajectory where the green agent prefers to stay in the center lane instead of switching to the top lane, which has sufficient space for overtaking the blue agent in front. Black stars indicate the positions.



(c) **Q. Indicate which agent (top vs. bottom) is more aggressive.**

Fig. 5: **Qualitative analysis.** We conduct a user study to qualitatively assess the emergent nature of the final trajectories. 26 out of the 27 participants correctly identified the aggressive driver.

simulator) results in more lane changes by the final simulated ego-vehicle (simulated using the python controller) and conversely, if a conservative human driver results in fewer simulated lane changes. In Figure 3b, we confirm this is indeed the case; aggressive drivers ($\theta < 0$) yield a greater number of lane changes than conservative drivers ($\theta > 0$).

B. Emergent behaviors

We evaluate the final trajectories generated using the learned risk sensitivity of human drivers in a merging scenario where a human agent attempts to merge onto the highway. Different risk sensitivities yield a range of emergent behaviors. For example, in [45], Wang et al. showed that two risk-averse agents maintain a larger minimum distance between them, while risk-seeking agents may allow a smaller gap. Further, in an interaction between a risk-averse and a risk-seeking agent, there is a higher likelihood of the risk-averse agent yielding to the risk-seeking agent.

The experiments conducted by Wang et al. modeled synthetic agents for which the risk sensitivity must be manually chosen. Here, we run the same set of experiments for human agents. In Figure 4a, we can observe darker colors (indicating larger minimum distance) corresponding to two risk-averse human agents and lighter colors (indicating smaller minimum distance) corresponding to two risk-seeking human agents.

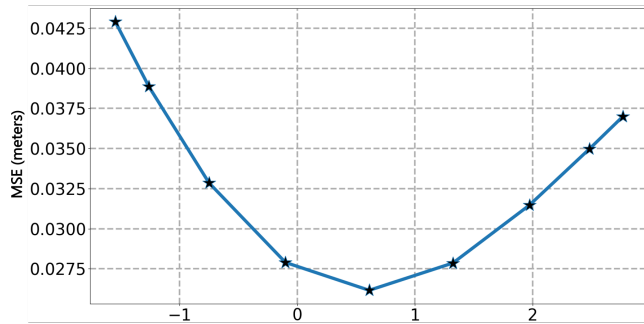


Fig. 6: **Comparison with [45]:** The approach by Wang et al. assumes a neutral risk sensitivity for human agents. However, when an ego-agent interacts with a human driver who may be aggressive or conservative (indicated by the negative and positive values on the x axis, respectively), then the assumption of neutral risk tolerance results in an error in terms of absolute value of minimum relative distance between the two agents. See Section V-D for more details.

In Figure 4b, we can observe darker colors (indicating a higher likelihood of yielding for the risk-averse agent) corresponding to interactions between a risk-seeking agent and a risk-averse agent.

C. User studies

We recruited 27 participants to respond to a user study consisting of two questions. The first question involved showing two video clips of final trajectories. The first clip (top) consisted of a risk seeking trajectory ($\theta = -2.429$), while the second (bottom) consisted of a risk averse trajectory ($\theta = 3.651$). Participants were not told the risk preferences that generated the trajectories, and were asked to identify which trajectory corresponded to an aggressive driver. The goal of this question is to qualitatively assess the emergent nature of the final trajectories. That is, based on simply observing the nature of the trajectory, can a human distinguish between the generated trajectories? We answer in the affirmative; 26 out of the 27 participants correctly identified the risk seeking driver as the aggressive driver.

D. Comparing with the baseline

We compare our modified risk-aware planner with existing planners that do not model human drivers and show that modeling human drivers results in safer navigation. Specifically, [45] (and similar planners) assign a fixed neutral risk tolerance to human drivers and the ego-vehicle generates to the human driver accordingly. There are two outcomes:

- 1) Suppose the human driver is, in fact, aggressive. Then, by modeling the driver with a neutral risk tolerance, the ego-vehicle may stray close to the aggressive driver as opposed to keeping a safe distance from them.
- 2) Conversely, suppose the human driver is conservative. Then, by modeling the driver with a neutral risk tolerance, the ego-vehicle may enter a brief deadlock during which both agents wait to see who moves first.

We aim to capture these inefficiencies via a single error metric, which is the absolute value of the minimum relative distance between the two agents. This metric is ideal since in both cases, it measures the discrepancy between the expected

distance and the actual observed distance. For example, we show that in the first case, the expected minimum relative distance between both agents is more than the observed distance while in the second case, we show the observed minimum distance is more than the actual distance. In both cases, the error is positive by virtue of the absolute value. Empirically, the maximum RMSE observed is 0.0425m or 10% as shown in Figure 6.

E. Using alternative human driver behavior models

Thus far, we have successfully demonstrated that CMetric can be effectively integrated with risk-aware planning to generate game-theoretic behavior-rich trajectories. Alternative models for human driver behavior such as the SVO can theoretically be used. However, there are practical issues when it comes to integrating these alternative models in risk-aware planning. Here, we discuss some of these challenges. SVO is an offline technique that requires a large volume of training data to learn a data-driven reward function via inverse reinforcement learning. Our technique is meant to be deployed in realtime and, as such, we test in an open-loop simulation and use *active* metrics such as yield %, frequency of lane changes, and minimum distance between agents. The SVO approach, on the other hand, uses RMSE to measure the deviation of the prediction trajectories from the ground-truth trajectories. We do not assume the availability of ground-truth data. In future, we will conduct experiments comparing CMetric with SVO once the source code for SVO is public.

VI. CONCLUSION, LIMITATIONS, AND FUTURE WORK

We presented an approach for risk-aware planning in multi-agent traffic with human agents. The basic intuition of our approach is that aggressiveness of a driver is linearly correlated with risk preference. That is, aggressive drivers are more risk-seeking while conservative drivers are more risk-averse. Accordingly, we integrate a human driver behavior model [7] with the risk-aware dynamic game solver in [45] via simple linear regression to derive a mapping between driver behavior and risk tolerance. Our results show that aggressive human driving results in more frequent lane-changing. We show that conservative drivers generally yield to aggressive drivers while maintaining a greater distance from them. Finally, we confirm that the final trajectories obtained from the risk-aware planner generate emergent behaviors though a comprehensive user study in which participants were able to distinguish between aggressive and conservative drivers.

There are some limitations to our method. Currently, we have tested our approach in an open-loop simulation where we use two different simulators for the human behavior model and the trajectory planner. To use both simulators in open-loop simulation effectively, the environment configuration must be kept identical, which is cumbersome and a hindrance. In the future, we will explore a closed-loop simulator that combines the human behavior model and the risk-aware trajectory planner.

REFERENCES

- [1] Ahmad Aljaafreh, Nabeel Alshabatat, and Munaf S. Najim Al-Din. Driving style recognition using fuzzy logic. *2012 IEEE International Conference on Vehicular Electronics and Safety (ICVES 2012)*, pages 460–463, 2012.
- [2] Tamer Başar and Geert Jan Olsder. *Dynamic noncooperative game theory*. SIAM, 1998.
- [3] Kenneth H. Beck, Bina Ali, and Stacey B Daughters. Distress tolerance as a predictor of risky and aggressive driving. *Traffic injury prevention*, 15 4:349–54, 2014.
- [4] Aniket Bera, Tanmay Randhavane, and Dinesh Manocha. Aggressive, tense or shy? identifying personality traits from crowd videos. In *IJCAI*, 2017.
- [5] J Christopher Brill, Mustapha Mouloua, Edwin Shirkey, and Pascal Alberti. Predictive validity of the aggressive driver behavior questionnaire (adbq) in a simulated environment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 53, pages 1334–1337. SAGE Publications Sage CA: Los Angeles, CA, 2009.
- [6] Kuanqi Cai, Chaoqun Wang, Shuang Song, Haoyao Chen, and Max Q-H Meng. Risk-aware path planning under uncertainty in dynamic environments. *Journal of Intelligent & Robotic Systems*, 101(3):1–15, 2021.
- [7] Rohan Chandra, Uttaran Bhattacharya, Trisha Mittal, Aniket Bera, and Dinesh Manocha. Cmetric: A driving behavior measure using centrality functions. *arXiv preprint arXiv:2003.04424*, 2020.
- [8] Margaret P Chapman, Jonathan Lacotte, Aviv Tamar, Donggun Lee, Kevin M Smith, Victoria Cheng, Jaime F Fisac, Susmit Jha, Marco Pavone, and Claire J Tomlin. A risk-sensitive finite-time reachability approach for safety of stochastic dynamic systems. In *2019 American Control Conference*, pages 2958–2963. IEEE, 2019.
- [9] Arnav Choudhry, Brady Moon, Jay Patrikar, Constantine Samaras, and Sebastian Scherer. Cvar-based flight energy risk assessment for multirotor uavs using a deep energy model. *arXiv preprint arXiv:2105.15189*, 2021.
- [10] Yin-Lam Chow and Marco Pavone. A framework for time-consistent, risk-averse model predictive control: Theory and algorithms. In *2014 American Control Conference*, pages 4204–4211. IEEE, 2014.
- [11] Yinlam Chow, Aviv Tamar, Shie Mannor, and Marco Pavone. Risk-sensitive and robust decision-making: a cvar optimization approach. In *Advances in Neural Information Processing Systems 28*, pages 1522–1530. Curran Associates, Inc., 2015.
- [12] Jerry L Deffenbacher, Eugene R Oetting, and Rebekah S Lynch. Development of a driving anger scale. *Psychological reports*, 1994.
- [13] Zhong-Xiang Feng, Jing Liu, Yang-Yang Li, and Wei-Hua Zhang. Selected model and susceptibility analysis of aggressive driving behavior. *Zhongguo Gonglu Xuebao(China Journal of Highway and Transport)*, 25(2):106–112, 2012.
- [14] Wendell H Fleming and William M McEneaney. Risk sensitive optimal control and differential games. In *Stochastic theory and adaptive control*, pages 185–197. Springer, 1992.
- [15] Hans Föllmer and Thomas Knispel. Entropic risk measures: Coherence vs. convexity, model ambiguity and robust large deviations. *Stochastics and Dynamics*, 11(02n03):333–351, 2011.
- [16] Davina J French, Robert J West, James Elander, and John Martin WILDING. Decision-making style, driving style, and self-reported involvement in road traffic accidents. *Ergonomics*, 36(6):627–644, 1993.
- [17] E Gulian, G Matthews, Aleck Ian Glendon, DR Davies, and LM Debnay. Dimensions of driver stress. *Ergonomics*, 1989.
- [18] Stephen J. Guy, Sujeong Kim, M. Chiao Lin, and Dinesh Manocha. Simulating heterogeneous crowd behaviors using personality trait theory. In *Symposium on Computer Animation*, 2011.
- [19] Motonori Ishibashi, Masayuki Okuwa, Shun’ichi Doi, and Motoyuki Akamatsu. Indices for characterizing driving style and their relevance to car following behavior. In *SICE Annual Conference 2007*, pages 1132–1137. IEEE, 2007.
- [20] Julian Jara-Ettinger. Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences*, 29:105–110, 2019.
- [21] Barbara Krahé and Ilka Fenske. Predicting aggressive driving behavior: The role of macho personality, age, and power of car. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 28(1):21–29, 2002.
- [22] Minae Kwon, Erdem Biyik, Aditi Talati, Karan Bhasin, Dylan P. Losey, and Dorsa Sadigh. When humans aren’t optimal: Robots that collaborate with risk-aware humans. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, page 43–52. Association for Computing Machinery, 2020.
- [23] Edouard Leurent and Jean Mercat. Social attention for autonomous decision-making in dense traffic. *arXiv preprint arXiv:1911.12250*, 2019.
- [24] Anirudha Majumdar, Sumeet Singh, Ajay Mandlekar, and Marco Pavone. Risk-sensitive inverse reinforcement learning via coherent risk models. In *Proceedings of Robotics: Science and Systems*, Cambridge, Massachusetts, July 2017.
- [25] Ishak Mohamad, Mohd. Alauddin Mohd. Ali, and Mahamod Ismail. Abnormal driving detection using real time global positioning system data. *Proceeding of the 2011 IEEE International Conference on Space Science and Communication (IconSpace)*, pages 1–6, 2011.
- [26] Yi Lu Murphey, Richard Milton, and Leonidas Kiliaris. Driver’s style classification using jerk analysis. *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*, pages 23–28, 2009.
- [27] Takayuki Osogami. Robustness and risk-sensitivity in markov decision processes. In *Advances in Neural Information Processing Systems 25*, pages 233–241. Curran Associates, Inc., 2012.
- [28] Arvind A Pereira, Jonathan Binney, Geoffrey A Hollinger, and Gaurav S Sukhatme. Risk-aware path planning for autonomous underwater vehicles using predictive ocean models. *Journal of Field Robotics*, 30(5):741–762, 2013.
- [29] Geqi Qi, Yiman Du, Jianping Wu, and Ming Xu. Leveraging longitudinal driving behaviour data with data mining techniques for driving style analysis. *IET intelligent transport systems*, 9(8):792–801, 2015.
- [30] Neil C Rabinowitz, Frank Perbet, H Francis Song, Chiyuan Zhang, SM Eslami, and Matthew Botvinick. Machine theory of mind. *arXiv preprint arXiv:1802.07740*, 2018.
- [31] Lillian J Ratliff and Eric Mazumdar. Inverse risk-sensitive reinforcement learning. *IEEE Transactions on Automatic Control*, 65(3):1256 – 1263, 2019.
- [32] R Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.
- [33] Francisco Aparecido Rodrigues. Network centrality: An introduction. *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems*, page 177, 2019.
- [34] Dorsa Sadigh, Shankar Sastry, Sanjit A Seshia, and Anca D Dragan. Planning for autonomous cars that leverage effects on human actions. In *Robotics: Science and Systems*, 2016.
- [35] Fridulv Sagberg, Selpi, Giulio Francesco Bianchi Piccinini, and Johan Engström. A review of research on driving styles and road safety. *Human factors*, 57(7):1248–1275, 2015.
- [36] Samantha Samuelson and Insoon Yang. Safety-aware optimal control of stochastic systems using conditional value-at-risk. In *2018 American Control Conference*, pages 6285–6290. IEEE, 2018.
- [37] Wilko Schwarting, Alyssa Pierson, Javier Alonso-Mora, Sertac Karaman, and Daniela Rus. Social behavior for autonomous vehicles. *Proceedings of the National Academy of Sciences*, 116(50):24972–24978, 2019.
- [38] Vishnu D Sharma, Maymoonah Toubeh, Lifeng Zhou, and Pratap Tokekar. Risk-aware planning and assignment for ground vehicles using uncertain perception from aerial vehicles. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11763–11769. IEEE, 2020.
- [39] Vishnu D Sharma, Maymoonah Toubeh, Lifeng Zhou, and Pratap Tokekar. Risk-aware planning and assignment for ground vehicles using uncertain perception from aerial vehicles. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11763–11769. IEEE, 2020.
- [40] Vishnu D Sharma, Maymoonah Toubeh, Lifeng Zhou, and Pratap Tokekar. Risk-aware planning and assignment for ground vehicles using uncertain perception from aerial vehicles. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11763–11769. IEEE, 2020.
- [41] Vishnu D Sharma, Maymoonah Toubeh, Lifeng Zhou, and Pratap Tokekar. Risk-aware planning and assignment for ground vehicles using uncertain perception from aerial vehicles. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11763–11769. IEEE, 2020.
- [42] Vishnu Dutt Sharma and Pratap Tokekar. Risk-aware path planning

- for ground vehicles using occluded aerial images. *arXiv preprint arXiv:2104.11709*, 2021.
- [43] Bin Shi, Li Xu, Jie Hu, Yun Tang, Hong Jiang, Wuqiang Meng, and Hui Liu. Evaluating driving styles by normalizing driving behavior based on personalized driver modeling. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45:1502–1508, 2015.
- [44] Orit Taubman-Ben-Ari, Mario Mikulincer, and Omri Gillath. The multidimensional driving style inventory—scale construct and validation. *Accident Analysis & Prevention*, 36(3):323–332, 2004.
- [45] Mingyu Wang, Negar Mehr, Adrien Gaidon, and Mac Schwager. Game-theoretic planning for risk-aware interactive agents. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6998–7005. IEEE, 2020.
- [46] Wenshuo Wang, Junqiang Xi, Alexandre Chong, and Lin Li. Driving style classification using a semisupervised support vector machine. *IEEE Transactions on Human-Machine Systems*, 47:650–660, 2017.
- [47] Peter Whittle. Risk-sensitive linear/quadratic/gaussian control. *Advances in Applied Probability*, 13(4):764–777, 1981.
- [48] Peter Whittle and Peter R Whittle. *Risk-sensitive optimal control*, volume 20. Wiley New York, 1990.