Influencing Interactions in Autonomous Driving

Dorsa Sadigh

Computer Science Department, Stanford University

1. Introduction

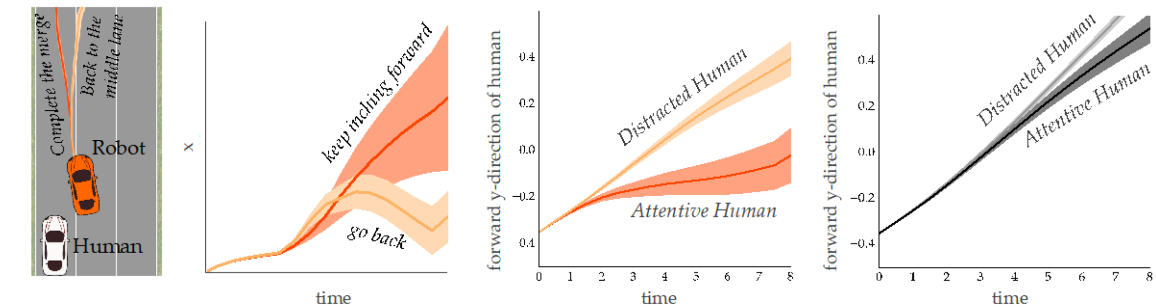
Today's society is rapidly advancing towards autonomous systems that interact and collaborate with humans, e.g., semiautonomous vehicles interacting with drivers and pedestrians, medical robots used in collaboration with doctors, or service robots interacting with their users in smart homes. One of the key aspects of safe and seamless interaction between autonomous systems and humans is studying how robots such as autonomous cars can *influence* humans' actions in one-on-one or group settings. This is usually overlooked by autonomous driving industry assuming humans act as external disturbances just like moving obstacles, or assuming that automation can always help societies without actually considering how humans can be impacted. Humans are not simply a disturbance that needs to be avoided; and they also do not easily adapt to any automation being inserted in their lives; humans are intelligent agents with approximately rational strategies who can be influenced and act in novel ways when interacting with other autonomous and intelligent agents.

In this paper, we will discuss a unifying framework for *influencing interactions* in autonomous driving, i.e., actions of autonomous vehicles that can positively influence human-driven vehicles at vehicle-to-vehicle level or large-scale interactions. We believe influencing interactions is a significant component for enabling safe and reliable integration of autonomous vehicles in our society.

2. Influencing Interactions at Vehicle Level

In our work, we have designed a novel framework for understanding the interaction between autonomous vehicles and human-driven vehicles. We model this interaction as a dynamical system, where the state of the environment evolves based on the actions of the human-driven car and the autonomous car at each time step: $x^{t+1} = f(x^t, u_A^t, u_H^t)$. Here, x^t denotes the state of the environment computed based on the sensor values at each time step including the coordinates, velocity and heading of each vehicle present in this interaction, and the road and lane boundaries. The set of actions of each vehicle u_A^t for the autonomous car and u_H^t for the human-driven car include the steering angle and acceleration of each vehicle.

Our key insight is that autonomous cars can take actions that influence the behavior of the human-driven cars on the same road. We observe this in our own driving behavior, when a car tries to change lanes, it starts nudging into the destination lane, which further affects the cars in that lane to slow down. Similarly, the actions of an autonomous car can result in the human changing lanes, slowing down, or speeding up. Our approach for planning for influencing interactions for autonomous vehicles has a few fundamental components. We first develop imitation learning techniques to build predictive models of human driving behavior, and have designed interaction-aware controllers that model the interaction between a human and a robot as a Stackelberg game. Leveraging optimization-based and game theoretic techniques, our work produces robot policies that influence human behavior towards safer and more interactive outcomes in vehicle-to-vehicle interaction with autonomous cars [1, 2, 4].



(a) Robot trajectories based on ! ‟ (#) in a lane change example.

(b) Robot trajectories based on ! ‟ (#).

(c) Human trajectories with active info gathering.

(d) Human trajectories with passive info gathering.

Fig 1. Planning for an autonomous car that influences a human driven car and actively gathers information about its driving style.

2.1. **Human Driver Models**

Imitation learning address the problem of learning computational models of humans in many robotics settings. Here, we leverage similar techniques, where we model each human driver as an agent who approximately optimizes a reward function: $u_H^* = \arg\max_{u_H} R_H(x, u_H, u_R)$. We assume $R_H(x, u_H, u_R) = w \cdot \varphi(x, u_H, u_R)$, is this underlying reward function and is represented as a linear combination of a set of features $\varphi(x, u_H, u_R)$. These features in the setting of driving can include distances to the road boundaries, lane boundaries, or other cars, velocity and heading of the vehicles. In our work, we collect training driving data in a driving simulator, and use this expert data in the form of demonstrations or preferences to learn the parameters w of the reward function using techniques such as maximum entropy inverse reinforcement learning or active preference based learning of reward functions [1-5, 9, 11].

2.2. Planning for interaction-aware controllers.

Once we have a predictive human driving model, we can plan for autonomous cars that better interact with humans by being mindful of how their actions influences humans. We consider a setting where the autonomous car optimizes for its own reward function: $u_R^* = \arg\max_{u_R} R_R(x, u_R, u_H^*)$. Here, the robot's reward function directly depends on and influences u_H^*, the learned human policy.

This interaction modeling results in a two-player game between a human-driven car and an autonomous car. The actions of the autonomous car influences the actions of the human driven car, while the actions of the human driven car influences the actions of the autonomous car. In order to efficiently solve this interaction game and plan for autonomous vehicles, we approximately solve the game as a Stackelberg (leader-follower) game. Our work results in influencing actions from the autonomous vehicle that are more assertive, more efficient, and in many settings safer. Some of these trajectories are shown in Fig 1. Our user studies suggest that autonomous cars that are mindful of their interactions with humans can achieve tasks such as lane changing or coordinating at intersections safely and efficiently [1, 2, 4].

3. Influencing Interactions at the Global Level

So far, we have discussed influencing interactions at the vehicle level, where the actions of an autonomous car influences the decision making of a human-driven car in its vicinity. This can be observed in many driving settings such as changing lanes, merging, or exiting from a highway and has substantial effects on the larger traffic system [6-8, 10, 12]. Our insight is that autonomous vehicles have the potential to impact human decision making. For instance, the presence of a large number of autonomous vehicles on roads can influence the state of traffic such as congestion, delay, or flow on each road and hence influence human's routing choices. We now would like to discuss the challenges arising in mixed-autonomy traffic settings where a large number of autonomous vehicles and human-driven vehicles interact.

3.1. **Equilibria in Mixed-Autonomy Traffic**

Traffic congestion has large economic and social costs. The introduction of autonomous vehicles can potentially reduce this congestion, both by increasing network throughput and by enabling a social planner to incentivize users of autonomous vehicles to take longer routes that can alleviate congestion on more direct roads.

We formalize the effects of *altruistic autonomy* on roads shared between human drivers and autonomous vehicles. We develop a formal model of road congestion on shared roads based on the fundamental diagram of traffic. We consider a network of parallel roads and provide algorithms that compute optimal equilibria that are robust to additional unforeseen demand. Our results show that even with arbitrarily small altruism, total latency can be unboundedly better than without altruism, and that the best selfish equilibrium can be unboundedly better than the worst selfish equilibrium. We validate our theoretical results through microscopic traffic simulations and show average latency decrease of a factor of 4 from worst-case selfish equilibrium to the optimal equilibrium when autonomous vehicles are altruistic [6].

3.2. Humans' Routing Choice Models

When users of a road network choose their routes selfishly, the resulting traffic configuration may become very inefficient. Because of this, we consider how to influence human routing decisions so as to decrease congestion on these roads.
Similar to previous section, we consider a network of parallel roads with two modes of transportation: (i) human drivers who will choose the quickest route available to them, and (ii) a ride hailing service which provides an array of autonomous vehicle ride options, each with different prices, to users.
We design a pricing scheme for the autonomous vehicles so that when autonomous service users choose from their options and human drivers selfishly choose their routes, road usage is maximized and transit delay is minimized. To do so, we formalize a model of how autonomous service users make choices between routes with different price v.s. delay values. Developing a preference-based algorithm similar to our work in learning reward functions in Section 2.1 to learn the preferences of the users, and using a vehicle flow model related to the Fundamental Diagram of Traffic, we formulate a planning optimization to maximize a social objective and demonstrate the benefit of the proposed routing and learning scheme [12].

3.3. Dynamic Routing in Mixed-Autonomy Traffic

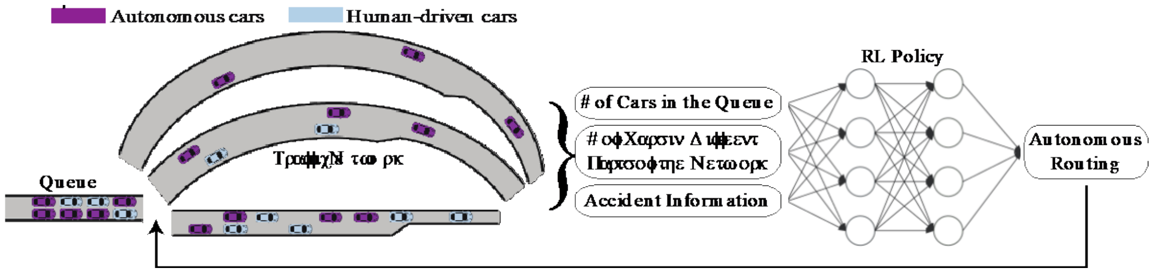


Fig 2. Using deep reinforcement learning to dynamically route autonomous cars.

We now consider a social planner that can control autonomous cars, which are a fraction of all present cars in mixed-autonomy traffic networks. We study a dynamic routing game, in which the route choices of autonomous cars can be controlled and the human drivers react selfishly and dynamically to autonomous cars' actions. As the problem is prohibitively large, we use deep reinforcement learning to learn a policy for controlling the autonomous vehicles. This policy influences human drivers to route themselves in such a way that minimizes congestion on the network.

To gauge the effectiveness of our learned policies, we establish theoretical results characterizing equilibria on a network of parallel roads and empirically compare the learned policy results with best possible equilibria. Moreover, we show that in the absence of these policies, high demands and network perturbations would result in large congestion, whereas using the policy greatly decreases the travel times by minimizing the congestion.

4. Summary

In this paper, we summarized our work in planning for influencing interactions in autonomous driving at two levels: i) vehicle-to-vehicle interaction, where an autonomous car plans to influence human-driven cars for safer and more efficient driving behavior, and ii) global-level interaction, where a larger number of autonomous vehicles and human-driven vehicles interact with each other on the same traffic network. We design routing decisions for autonomous vehicles that influence humans routing choices in order to decrease the total delay of the traffic network for a more desirable societal objective.

It is only now that autonomous systems are finally weaving their way into our lives. Robots are moving into our homes. Smart cities and intelligent vehicles are becoming a reality. Our long-term goal is to develop a theory for modeling and designing the effects of automation and robotics on humans' decision making, and this work is a first step towards developing efficient robotics algorithms that lead to safe and transparent autonomous systems cognizant of the interactions and influences they can have on humans and our society.

5. References

[1] Planning for Autonomous Cars that Leverage Effects on Human Actions. Dorsa Sadigh, S. Shankar Sastry, Sanjit A. Seshia, Anca D. Dragan. *Proceedings of Robotics: Science and Systems (RSS), June 2016*

[2] Information Gathering Actions over Human Internal State. Dorsa Sadigh, S. Shankar Sastry, Sanjit A. Seshia, Anca Dragan. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 2016*

[3] Active Preference-Based Learning of Reward Functions. Dorsa Sadigh, Anca D. Dragan, S. Shankar Sastry, Sanjit A. Seshia. *Proceedings of Robotics: Science and Systems (RSS), July 2017*

[4] Planning for Cars that Coordinate with People: Leveraging Effects on Human Actions for Planning and Active Information Gathering over Human Internal State. Dorsa Sadigh, Nick Landolfi, S. Shankar Sastry, Sanjit A. Seshia, Anca D. Dragan. *Autonomous Robots (AURO), October 2018*

[5] Batch Active Preference-Based Learning of Reward Functions. Erdem Bıyık, Dorsa Sadigh. *Proceedings of the 2nd Conference on Robot Learning (CoRL), October 2018*

[6] Altruistic Autonomy: Beating Congestion on Shared Roads. Erdem Bıyık, Daniel A. Lazar, Ramtin Pedarsani, Dorsa Sadigh. *Proceedings of the 13th International Workshop on Algorithmic Foundations of Robotics (WAFR), December 2018*

[7] Maximizing Road Capacity Using Cars that Influence People. Daniel A. Lazar, Kabir Chandrasekher, Ramtin Pedarsani, Dorsa Sadigh. *Proceedings of the 57th IEEE Conference on Decision and Control (CDC), December 2018*

[8] Hierarchical Game-Theoretic Planning for Autonomous Vehicles. Jaime F. Fisac, Eli Bronstein, Elis Stefansson, Dorsa Sadigh, S. Shankar Sastry, Anca D. Dragan. *International Conference on Robotics and Automation (ICRA), May 2019*

[9] Learning Reward Functions by Integrating Human Demonstrations and Preferences. Malayandi Palan, Nicholas C. Landolfi, Gleb Shevchuk, Dorsa Sadigh. *Proceedings of Robotics: Science and Systems (RSS), June 2019*

[10] Human-Robot Interaction for Truck Platooning Using Hierarchical Dynamic Games. Elis Stefansson, Jaime Fisac, Dorsa Sadigh, Shankar Sastry, Karl H. Johansson. *European Control Conference (ECC), June 2019*

[11] Active Learning of Reward Dynamics from Hierarchical Queries. Chandrayee Basu, Erdem Bıyık, Zhixun He, Mukesh Singhal, Dorsa Sadigh. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), November 2019*

[12] The Green Choice: Learning and Influencing Human Decisions on Shared Roads. Erdem Bıyık, Daniel A. Lazar, Dorsa Sadigh, Ramtin Pedarsani. *Proceedings of the 58th IEEE Conference on Decision and Control (CDC), December 2019*