Cooperative Planing for Autonomous Lane Merging

Shray Bansal[1], Akansel Cosgun[2], Alireza Nakhaei[2] and Kikuo Fujimura[2]

***Abstract*— Autonomously performed lane merging as collaborative planning in a scenario with human drivers. Showed that planning to co-operate with the human significantly outperforms maximizing selfish reward in this domain.**

I. INTRODUCTION

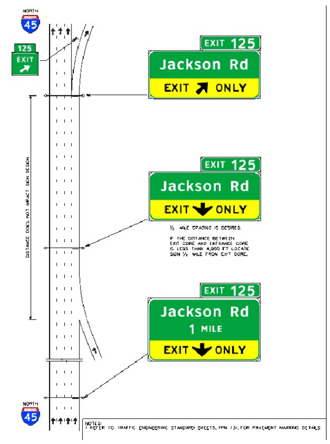
Merging into another lane is easy in isolation, however, traffic density and road design make the situation complex. For example, in Figure 1a there is only half a mile between the entry and the exit. This creates a scenario where cars exiting and entering have to negotiate with each other to safely navigate this lane merge in the time available. Humans do so by inferring intentions, driving styles, etc. of those around them and conditioning their behavior on that. This is in contrast to scenarios like lane-following or intersection navigation where one could get away with treating other cars like obstacles to avoid instead of agents to interact with.

Prior work in autonomous driving disregards the complexity of these agents, assuming them to be constant velocity obstacles, or controlled by simple rules. We take inspiration from recent work by Sadigh et al. [1] which grants agency to the other (human) driver and take into account their goals and actions. However, their self-interested AV only maximizes it's own reward, they assume also that the AV is able to choose its actions first. This works well in scenarios where the aims AV to influence the human's actions like making them slow down by moving in front of them but can be uncomfortable or unsafe in real-world situations. On the other hand, it is not uncommon for human drivers to slow down and let others into their lane. Behavioral economics also supports the argument that people are influenced by notions of fairness and reciprocity when making decisions, even at the expense of their own self-interest.

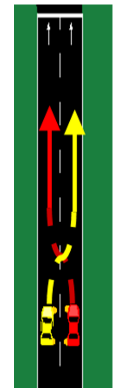
Our goal is to successfully navigate the Double Lane merging scenario autonomously in the presence of human drivers. Here, two cars start in adjacent lanes and must merge into each others' lanes in a limited road length (Figure 1b). Our hypothesis is that *an AV which considers the goals of other agents in addition to it's own selfish reward should benefit the human driver without adversely affecting itself.*

We simulated Double Lane merging and measured the performance of the AV with different levels of cooperation in the presence of people. Our analysis shows that having a more balanced reward, when planning for the AV, leads to better human performance. We also found that it positively affected the AV's performance as well.

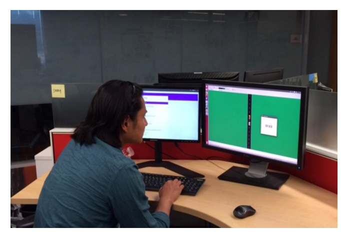
[1] Georgia Institute of Technology. sbansal34@gatech.edu
[2] Honda Research Institute (HRI), Mountain View, CA, USA.



(a) Challenging Highway



(b) Problem Setup



(c) User controls the human car.

Fig. 1: (a) Highway in San Diego where the entry and exit are only 0.5 mile apart. (b) Double Lane Merge Scenario: where two cars start side-by-side with and have to merge into each other's lanes on a road of limited length.

II. PROBLEM

The goal is for the autonomous agent to safely move to it's target lane as quickly as possible in the presence of a human driver. We use the scenario of Figure 1b where either car changing lanes independent of the other can lead to a collision. So the challenge here is to take actions based upon the predicted behavior of the other car.

Our world is a fully-observed dynamical system where the state x contains the position and speed of both cars. The state at the next time-step of the system is determined by applying the control for both human (u_H) and robot (u_R) according to the deterministic function T, $x_{t+1} = T(x_t, u_R(t), u_H(t))$. Each agent has five available actions that get applied instantly. (1) *accelerate* action increases the speed, while (2) *decelerate* decreases it by a constant factor, (3) *stay* maintains it, (4) *turn-left*, and (5) *turn-right* introduce a positive, and negative lateral component to the velocity while maintaining the speed.

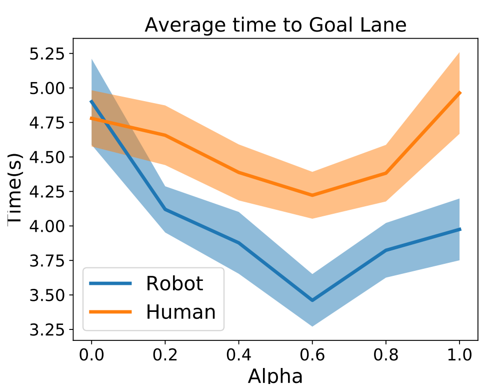
III. APPROACH

The goal of our robot car is to maximize a given reward $R_R(x)$ over the fixed length of time N. So the optimal set of actions u_R^* from the current state x_0 will be given by,

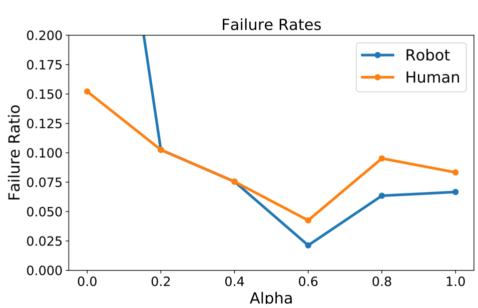
$$u_R^*(x_0) = argmax_{u_R^{(0)},\ldots,u_R^{(N-1)}} \sum_{t=0}^{N-1} R_R(T(x_t, u_H^{(t)}, u_R^{(t)})). \tag{1}$$

However, $u_H^{(t)}$ are the *future* actions of the human and so are not known. So the optimal robot actions can not be computed without predicting the human's actions. To do this, we assume that the human acts rationally to attain their goal. Due to the non zero-sum nature of this multi-agent game, as well as the dependence of each agent's reward on the other agent's actions, we will approach this as a collaboration. In it, to determine the optimal set of actions for the robot u_R^* we will maximize the joint reward $R_J = \alpha R_R + (1-\alpha) R_H$. with respect to the joint action space (u_R^*, u_H^*). α here is the co-operation factor which determines the relative importance of each reward in the collaboration. For example, $\alpha = 1$ is an aggressive robot with no consideration for the human and $\alpha = 0$ is the reverse, while $\alpha = 0.5$ equally considers the goal of both agents. We do not assume apriori knowledge of α and plan to study its effect with the user study.

We implement this optimization as a limited depth (6 seconds) search using a simple heuristic to speed it up. We execute the first action for the robot from the optimal plan and repeat the search after every human action at every time-step. The robot rewards are positive for reaching the goal and negative for collisions.



(a)



(b)

Fig. 2: Results from the user study.

IV. EXPERIMENTS AND ANALYSIS

We used the SUMO simulator [2] and the in-built visualizer to render the cars, the simulation time-step is 0.2 seconds. Our user-study involved 20 participants each performing 18 trials. They were instructed to safely drive the car to the goal lane with an AV that communicated its goal with blinkers. The cooperation factor (α) and road lengths were randomly sampled from $\{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$, and$\{100, 200\}m$ respectively. We studied the effect of α, having around 50 trials for each, in Figure 2.

Figure 2a shows the average time it takes for a car to merge into it's goal lane for each α. Lower is better here and the minima is near the middle and is high at either end, implying the higher effectiveness of the *Fair* AV (FAV) with $\alpha = 0.6$. We compared the FAV to the Selfish AV (SAV) with $\alpha = 1.0$ using an unpaired one-sided t-test and found that the merge-time was significantly lower for both the human ($p < 0.05$) and the AV ($p < 0.05$) in the case of the fairer vehicle. The SAV is equivalent to the agent proposed by Sadigh et al. [1] for our scenario. Since the FAV is more considerate towards the human's goal it improves the humans' ability to reach the goal. However, the FAV fairs better even for the robot, we attribute this to its possession of a more accurate model of the user's behavior. During planning, the SAV assumes that the human actively plans for the AV's goal and not her own. However, this is false and probably leads to it modifying its plan often.In Figure 2a, we also observe that the human is slower at lane merging. This can be explained by the fact that the robot's reward favors the fastest lane merge plan but the human has only wishes to merge before the road ends and is not incentivized to do it as soon as possible.

When calculating the average times above we only considered the successful trials, i.e. trials where the car was able to successfully merge into it's goal lane before the end of the road was reached. In Figure 2b we plot the Failure Rate, which is the ratio of unsuccessful trials to the total trials for each alpha. Again, we found that the FAV outperformed all others, including the SAV, which means that both human and robot were able to reach their goal lanes more often when the AV had a fair reward function.

V. CONCLUSION

We formulated a planning framework with different co-operation levels and used it to control an autonomous car interacting with a human in simulation. We observed that in this mixed-autonomy set-up collaboration outperformed selfishness. In the future we would like to study this effect in other human-robot interaction domains.

REFERENCES

[1] Sadigh, Dorsa, et al. "Planning for Autonomous Cars that Leverage Effects on Human Actions." Robotics: Science and Systems. 2016.

[2] Krajzewicz, Daniel, et al. "Recent development and applications of SUMO-Simulation of Urban MObility." IJOASM 2012.