Altruistic Maneuver Planning for Cooperative Autonomous Vehicles Using Multi-agent Advantage Actor-Critic

Behrad Toghi[1,2], Rodolfo Valiente[1], Dorsa Sadigh[2], Ramtin Pedarsani[3], Yaser P. Fallah[1]

[1]*University of Central Florida, Orlando, FL* [2]*Stanford University, Stanford, CA* [3]*University of California, Santa Barbara, CA*

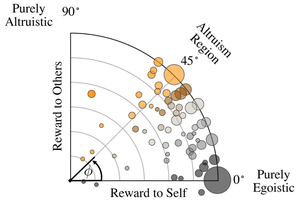
CVPR VIRTUAL JUNE 19-25

1. Overview

- Autonomous and human-driven vehicles must learn to co-exist by sharing the same road infrastructure.
- To attain socially-desirable behaviors, autonomous vehicles must be instructed to consider the utility of other vehicles around them in their decision-making process. This is a challenging problem due to the ambiguity of a human driver's willingness to cooperate with an autonomous vehicle.
- We take an end-to-end approach and let the autonomous agents to implicitly learn the decision-making process of human drivers only from experience. We introduce a multi-agent variant of the synchronous Advantage Actor-Critic (A2C) algorithm and train agents that coordinate with each other.

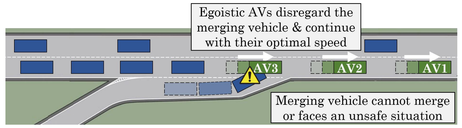
2. Social Preferences

- The Social Value Orientation (SVO) ring demonstrates different behaviors based on a human/robot's preference to account for others.

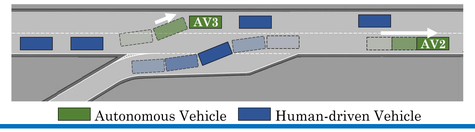


Agent Type	Cares about itself	Cares about its allies	Cares about humans
Egoistic			
Cooperative			
Cooperative Sympathetic			

- Egoistic AVs solely optimize for their own utility.



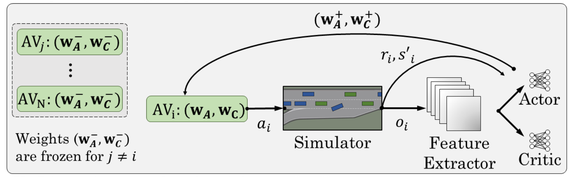
- Altruistic AVs compromise on their individual utility.



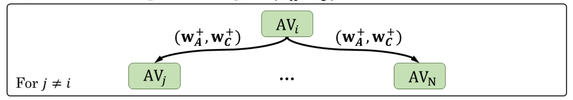
3. Altruistic Cooperative Driving

- Multi-agent Advantage Actor-Critic framework and policy dissemination process enables concurrent training of multiple agents.

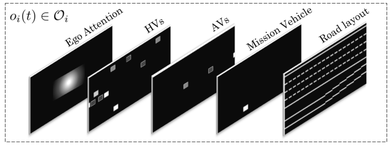
I) Repeat the weight update k times for agent i:



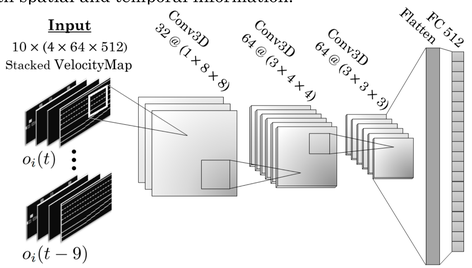
II) Disseminate updated weights $(\mathbf{w}_A^+, \mathbf{w}_C^+)$:



- Stacked multi-channel **VelocityMap** state representation embeds the speed and position of vehicles as well as the road layout.

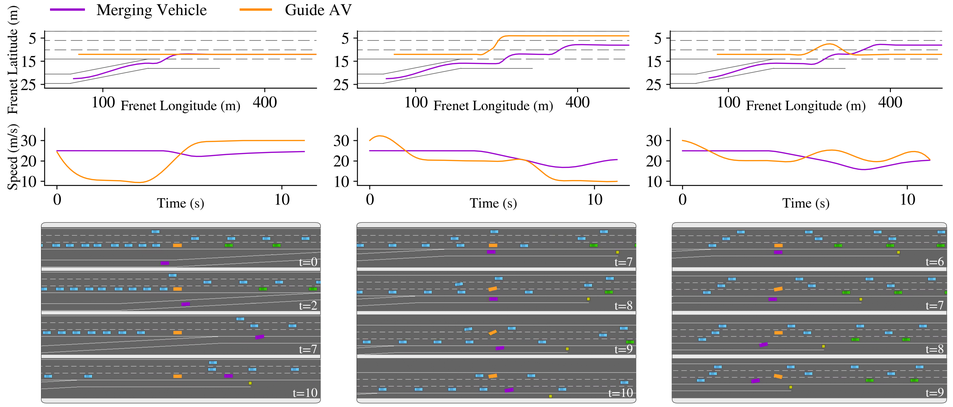


- Architecture of our 3D convolutional Feature Extractor Network (FEN) that captures both spatial and temporal information.

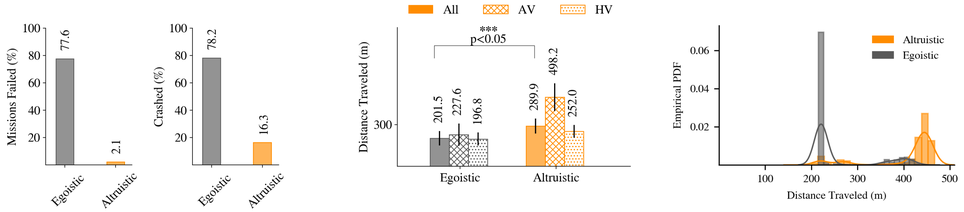


4. Experiments & Results

- Sample behaviors recorded during the training of autonomous agents.



- Sympathetic cooperative AVs improve traffic-level metrics.



5. Summary

- Using our proposed decentralized multi-agent learning scheme, we are able to induce altruism into the decision-making process of autonomous vehicles and adjust their SVO.
- Our altruistic agents not only learn to drive on the highway environment from scratch but also are able to coordinate with each other and affect the behavior of humans around them to realize socially-desirable outcomes that eventually improve traffic safety and efficiency.

Questions and/or Comments: toghi@knights.ucf.edu, https://symcodrive.toghi.net