iPLAN: Intent-Aware Planning in Heterogeneous Traffic via Distributed MARL

Xiyang Wu[1], *Rohan Chandra*[2], *Tianrui Guan*[1], *Amrit Singh Bedi*[1], *Dinesh Manocha*[1]

[1] University of Maryland, College Park [2] The University of Texas at Austin

Main Contribution

- A fully decentralized MARL algorithm embedded with a classical trajectory forecasting architecture for intent-aware planning in dense and heterogeneous environments.
- An explicit representation of agents' private incentives, including
  - **Behavioral Incentive:** High-level decision-making strategy that sets planning sub-goals.
  - **Instant Incentive:** Low-level motion planning that executes sub-goals.
- An incentive inference mechanism that allows agents to infer incentives from their opponents solely with their local observations.

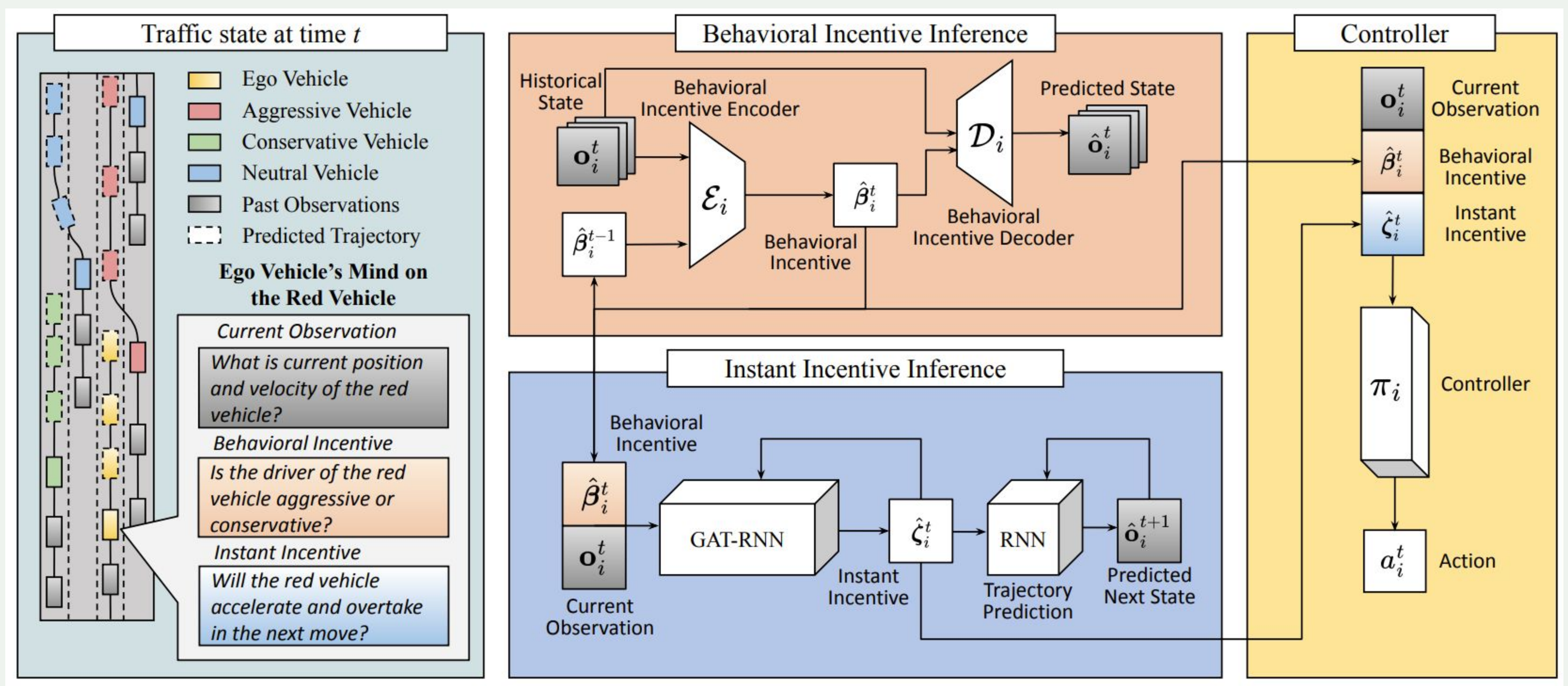
Problem Statement

Each vehicle (yellow vehicle) performs decision -making in heterogeneous traffic with its mind over another vehicle (red vehicle) in three aspects:

- **Current Observation:** *What is the current position and velocity of the red vehicle?*
- **Behavioral Incentive:** *What's the most likely action of this driver to take next?*
- **Instant Incentive:** *How should I execute this sub-goal/high-level action/plan using my controller so that I'm safe and still on track towards my goal?*

The behavior incentive biases the motion forecasting by setting planning sub-goals with intention awareness, whereas the instant incentive executes the sub-goals to perform motion planning.

Methodology



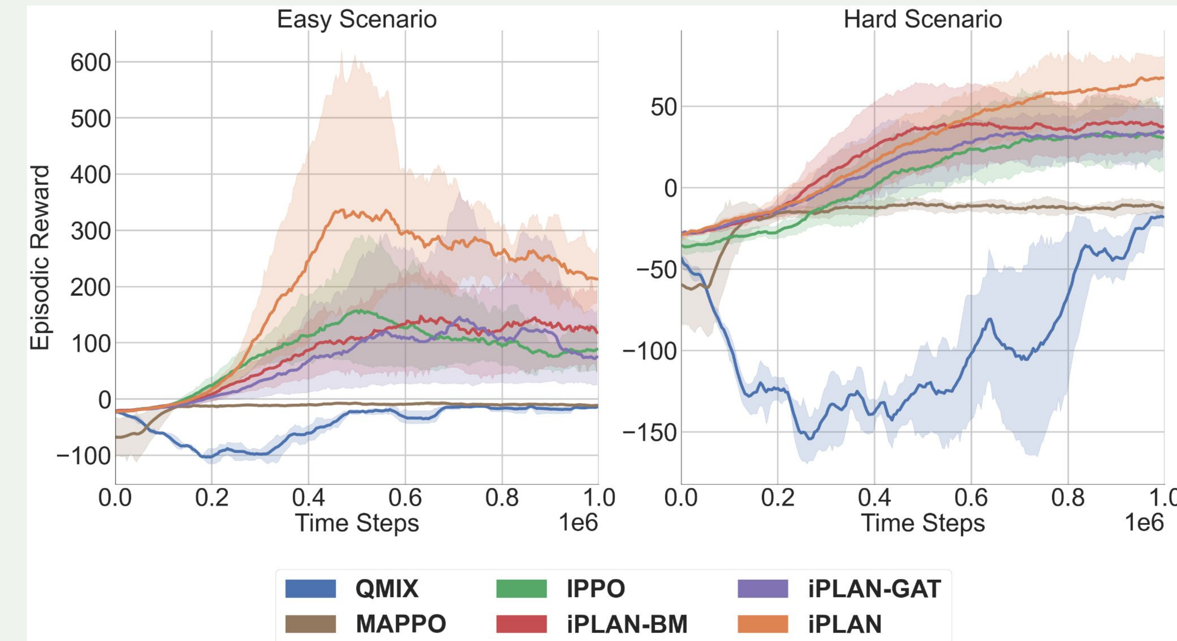
- **Behavior Incentive Inference**: Use historical observation sequences to predict observation sequences in the next few time steps.
- **Instant Incentive Inference:** Use current observation and behavior incentive inference to generate instant incentives for trajectory prediction.
- **Controller:** Combine two incentive inferences with current observations for decision making.

Navigation Metrics

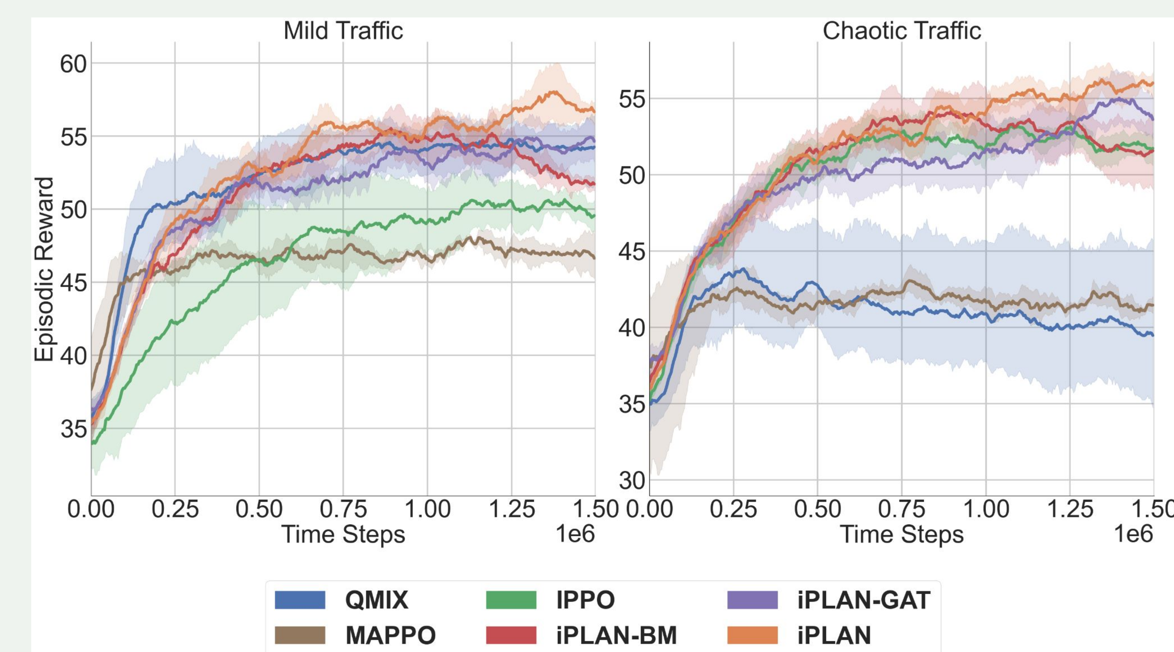
	Approach	Avg. Speed (m/s)	Avg. Survival Time (# Time Steps) ↑	Success Rate (%) ↑
Mild	QMIX [24]	21.24 ± 0.09	**75.98 ± 3.67**	67.50 ± 6.34
	MAPPO [36]	**27.85 ± 0.40**	48.94 ± 3.11	32.81 ± 5.22
	IPPO [11]	22.63 ± 0.17	66.13 ± 4.13	49.06 ± 7.35
	iPLAN-GAT	22.05 ± 0.11	75.54 ± 3.61	**68.44 ± 6.64**
	iPLAN-BM	22.61 ± 0.16	64.11 ± 4.28	45.63 ± 6.33
	iPLAN	22.91 ± 0.15	70.56 ± 3.81	**68.44 ± 5.86**
Chaotic	QMIX [24]	27.06 ± 0.47	39.38 ± 2.64	19.69 ± 3.72
	MAPPO [36]	**29.46 ± 0.05**	42.31 ± 2.43	16.25 ± 3.76
	IPPO [11]	22.28 ± 0.13	67.01 ± 3.64	42.50 ± 7.12
	iPLAN-GAT	20.91 ± 0.13	71.24 ± 3.83	61.88 ± 6.41
	iPLAN-BM	21.65 ± 0.28	63.20 ± 3.51	35.31 ± 5.66
	iPLAN	21.61 ± 0.16	**76.20 ± 3.33**	**67.81 ± 5.91**

iPLAN has a higher success rate and longer survival time than all other approaches, while it tends to be conservative in its average speed.

Empirical Results



Non-Cooperative Navigation: iPLAN (**Orange**, DTDE) has higher episodic reward than QMIX (**Blue**, CTDE), MAPPO (**Brown**, CTDE), and IPPO (**Green**, DTDE), in both easy and hard scenarios.



Heterogeneous Highway: iPLAN (**Orange**, DTDE) has higher episodic reward than QMIX (**Blue**, CTDE), MAPPO (**Brown**, CTDE), and IPPO (**Green**, DTDE), in mild and chaotic traffic.

For More Details