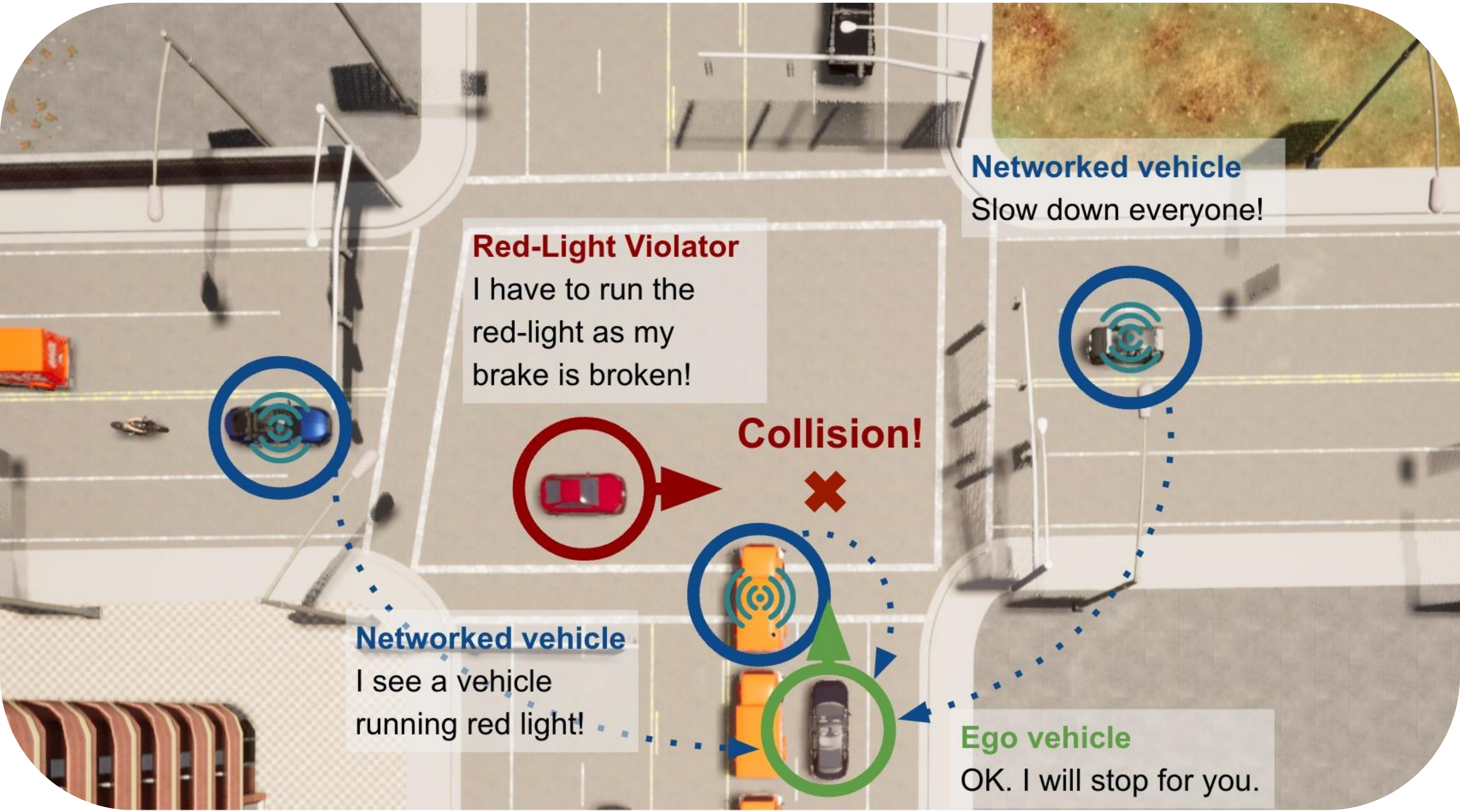
Talking Vehicles: Cooperative Driving via Natural Language

Jiaxun Cui, Chen Tang, Jarrett Holtz, Janice Nguyen, Alessandro G. Allievi, Hang Qiu, Peter Stone

Presenter: Jiaxun Cui

TEXAS
The University of Texas at Austin
BOSCH
UC RIVERSIDE
Sony AI


Networked vehicle
Slow down everyone!
Red-Light Violator
I have to run the red-light as my brake is broken!
Collision!
Networked vehicle
I see a vehicle running red light!
Ego vehicle
OK. I will stop for you.

Motivation

1. Driving involves multi-agent coordination
2. Prior works has done generating latent representations but they are impossible to be understand by humans
3. Coordinating human drivers is a potential in a mixed-autonomy situation
4. Enabling negotiation among vehicles to achieve individual tasks

Problem Formulation

Partially **O**bservable **S**tochastic **G**ame $\langle \mathcal{I}, \mathcal{S}, \{\mathcal{O}_i\}, \{\mathcal{A}_i\}, \mathcal{P}, \{\mathcal{R}_i\} \rangle$

- $\mathcal{I}$ is the set of all N agents,
- $\mathcal{S}$ is the joint state space of the environment,
- $\mathcal{A}_i$ is the action space of agent i, and $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_N$ is the joint action space of all agents,
- $\mathcal{P}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ is the state transition functions or the environment dynamics,
- $\mathcal{R}_i$: $\mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the reward function for agent i,
- $\mathcal{O}_i$: $\mathcal{S} \times \mathcal{A} \to \mathbb{O}$ is the observation function for agent i,

Problem Formulation

Partially **O**bservable **S**tochastic **G**ame $\langle \mathcal{I}, \mathcal{S}, \{\mathcal{O}_i\}, \{\mathcal{A}_i\}, \mathcal{P}, \{\mathcal{R}_i\} \rangle$

Focal Group A group of agents of interest that optimizes for social welfare given a set of background agents

$$\max_{\{\pi_i\}_{i\in\mathcal{F}}} \mathbb{E}\left[\sum_{i\in\mathcal{F}} \sum_{t=0}^{t=\infty} R_i(s_t, \mathbf{a}_t) \middle| \{\pi_j\}_{j\notin\mathcal{F}, j\in\mathcal{I}} \right]$$

1. How can learned agents understand the situation and **generate** meaningful messages to help others perceive the environment or potentially negotiate about motion plans in natural languages;
2. How can learned agents **comprehend** the received natural language messages and **incorporate** them into high-level cooperative driving decisions?

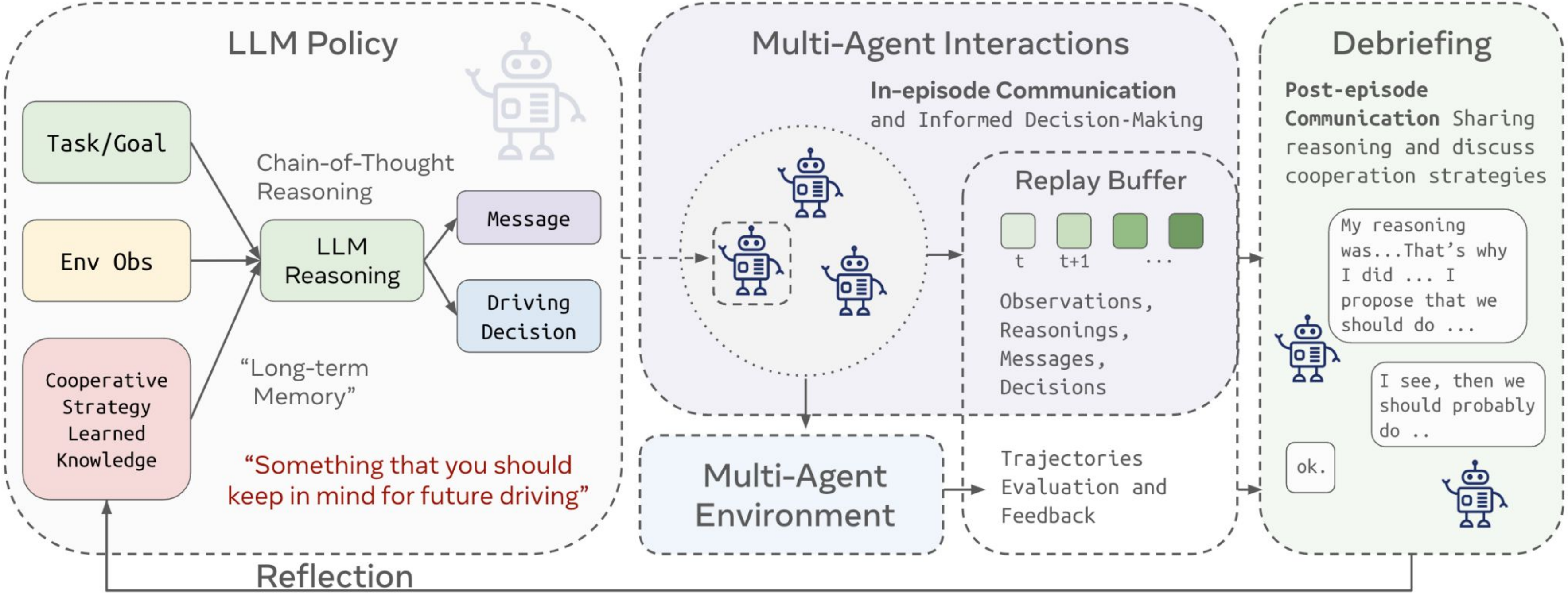
Scenarios

Individual tasks and rewards that may not fully aligned
Text scenario description with partial observability awareness
Action space: natural language message and atomic controls

Method: Agent Framework $P_i(\{t_i^m\}; \{t_i^c\} | \{t_i^o\}; \{\{t_j^m\}\}_{j \in \mathcal{F}})$

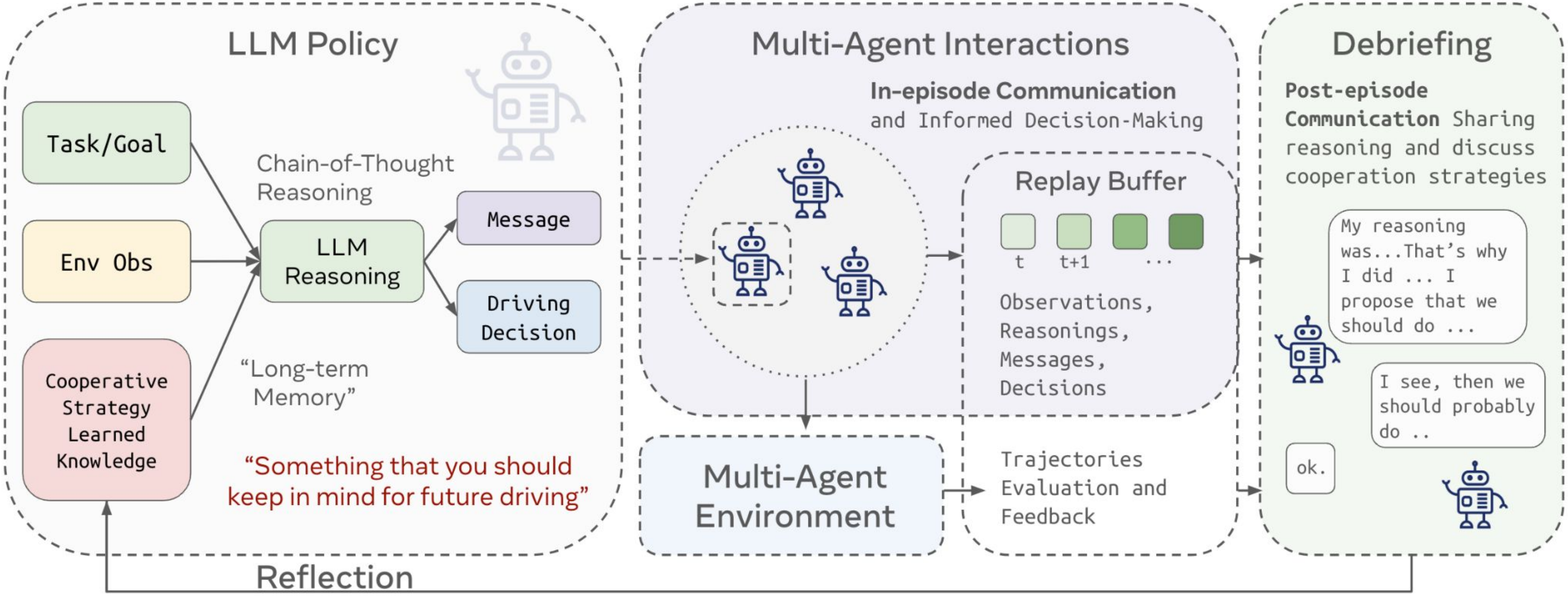
In-Context Knowledge $P_i(\{t_i^m\}; \{t_i^c\} | \{\boldsymbol{t_i^k}\}; \{\boldsymbol{t_i^s}\}; \{t_i^o\}; \{\{t_j^m\}\}_{j \in \mathcal{F}})$

Chain-of-Thought Reasoning $P_i(\{t_i^m\}; \{t_i^c\} | \{t_i^k\}; \{t_i^s\}; \{t_i^o\}; \{\{t_j^m\}\}_{j \in \mathcal{F}}; \{\boldsymbol{t_i^r}\})$



Method: LLM+Debrief

Replay Buffer + Batch Context Sampling
Debrief: Propose cooperation strategy and Reflect Individually



Hypothesis

(1) LLM agents can establish collaboration without prior interactions;

(2) Natural language communication enhances the performance and coordination of LLM agents compared to those without communication;

(3) Decentralized reflection enables LLM agents to improve their collaborative ability over interactions;

(4) Centralized discussion among LLM agents provides additional improvements in collaboration and communication compared to decentralized reflection.

Metrics and Baselines

60 episodes Round Robin among [risky, safe] scenarios

Metrics

Average Total Reward normalized by group size
Average Number of Success normalized by group size
Average Number of Collisions normalized by group size

Baselines

1. Zero-shot
2. Decentralized Reflection
3. Centralized Debrief
4. Correction + RAG

Results

Method \ Scenario			Overtake (Negotiation)			Highway Merge			Highway Exit		
Name	LLM	Comm	R ↑	CR ↓	SR ↑	R ↑	CR ↓	SR ↑	R ↑	CR ↓	SR ↑
Zero-shot (Silent)	Yes	No	-0.13	55.0	41.7	-0.87	93.3	6.7	-0.53	63.3	36.5
+Reflection (Silent)	Yes	No	0.80	3.0	83.3	-0.37	68.3	31.7	0.20	40.0	60.0
+Correction+RAG (Silent)	Yes	No	0.00	50.0	50.0	0.03	48.3	51.7	-0.16	58.3	41.7
Zero-shot (Comm)	Yes	Yes	0.53	23.3	76.5	-1.00	100.0	0.0	-0.60	65.0	35.0
+Reflection (Comm)	Yes	Yes	0.73	11.7	85.0	0.53	23.3	76.7	0.32	33.3	65.0
+Correction+RAG (Comm)	Yes	Yes	0.83	6.7	90.0	-0.07	53.3	46.7	-0.16	58.3	41.7
+Debrief (Comm)	Yes	Yes	**1.00**	**0.0**	**100.0**	**1.00**	**0.0**	**100.0**	**0.63**	**10.0**	**73.3**

Method \ Scenario			Overtake (Perception)			Red Light			Left Turn		
Name	LLM	Comm	R ↑	CR ↓	SR ↑	R ↑	CR ↓	SR ↑	R ↑	CR ↓	SR ↑
Zero-shot (Silent)	Yes	No	-0.87	93.3	6.7	-0.87	93.3	6.7	-0.93	96.7	3.3
+Reflection (Silent)	Yes	No	-0.26	36.7	10.0	-0.87	93.3	6.7	-0.27	63.3	36.7
+Correction+RAG (Silent)	Yes	No	0.07	33.3	40.0	-0.73	86.7	13.3	0.20	40.0	60.0
Zero-shot (Comm)	Yes	Yes	-0.46	73.3	26.7	-0.33	66.7	33.3	-1.00	100.0	0.0
+Reflection (Comm)	Yes	Yes	0.40	30.0	70.0	0.07	10.0	26.7	**0.60**	**20.0**	**80.0**
+Correction+RAG (Comm)	Yes	Yes	**0.70**	**6.7**	76.7	0.73	13.3	87.7	-0.60	76.7	16.7
+Debrief (Comm)	Yes	Yes	0.63	16.7	**80.0**	**1.00**	**0.0**	**100.0**	**0.60**	**20.0**	**80.0**
Coopernaut (Comm)	No	Yes	1.00	0.0	100.0	0.97	0.0	96.7	1.00	0.0	100.0

Results

- Communication does not guarantee collaboration in zero-shot interactions
- Decentralized reflection and correction with RAG memory reduce collisions.
- Centralized debriefing enhances coordination more than decentralized reflection.
- The potential of natural language communication for multi-agent

Example Generated Coordination Strategy

Car1 (Merging Vehicle)

Knowledge To effectively merge onto the highway when in a hurry, I should initiate communication by clearly stating my intention to merge, including my current speed and position. As I approach the merge point, I will gradually accelerate to match the highway speed. The vehicle in the rightmost lane should proactively create a gap by slightly reducing its speed earlier, allowing me to merge smoothly. Continuous communication is essential to coordinate actions and avoid collisions. I must maintain a safe merging distance and speed, avoiding abrupt maneuvers, and ensure my actions do not cause stagnation or disrupt traffic flow. I should also be prepared to adjust my speed dynamically based on real-time feedback from the highway vehicle.,

Cooperative Strategy In our cooperative strategy for merging onto the highway, the merging vehicle (myself) will initiate the process by clearly communicating its intention to merge, including its current speed and position. As the merging vehicle approaches, it will gradually accelerate to match the highway speed, ensuring a smooth transition. The vehicle in the rightmost lane of the highway will proactively create a gap by slightly reducing its speed earlier, allowing the merging vehicle to enter the lane without needing to match speeds precisely. Throughout the process, both vehicles will maintain clear communication, with the merging vehicle leading in signaling its actions and the highway vehicle responding by adjusting speed to facilitate a safe merge. This approach prevents collisions and avoids traffic stagnation by ensuring a coordinated and efficient merging process.

Qualitative Videos

https://utexas.box.com/s/dhnpd03gq0088ctiofqsoe4q6cmn46uj

Discussion

1. Inference time reduction
2. Ad hoc teamwork instead of pre-coordination
3. More agents
4. Involving Visual/Multimodal Language Models