Toward a Multi Agent Approach for LLM-Based Dynamic Vehicle Control and Communication in Accidental Condition

Rafi Md Ashifujjman
Graduate School of Integrated Science and Technology
Shizuoka University
Hamamatsu, Japan
ashifujjman@gmail.com

Naoki Fukuta
College of Informatics,Academic Institute
Shizuoka University
Hamamatsu, Japan
fukuta@inf.shizuoka.ac.jp

***Abstract*—**

I. Introduction

~~This paper explores the potential of integrating Large Language Models (LLMs) [1] into a multi agent framework for enhancing decision making in Autonomous Vehicles (AVs), particularly in unknown and unsafe domains. Our main concern would be~~ ~~**To explore the use of Large Language Models(LLMs) for autonomous driving systems(ADS) as the main decision-making agent within a multi-agent framework for evaluating its reasoning abilities in handling long-tail, False positive, and False negative scenarios.**~~ ~~We will investigate whether LLMs can effectively improve upon current mechanisms in handling complex traffic scenarios, with a focus on replicating human-like reasoning and adaptability. To this aim, we have prepared some unknown unsafe scenarios that self driving cars will not properly handle well today and these will serve as tests for the LLM powered decision making system. We will see how well they perform compared to current systems. We will investigate whether they can reason like a human and adapt to unexpected situations. The proposed approach aims to address challenges in autonomous driving systems, especially in scenarios that require common sense reasoning and experience based decision making.~~

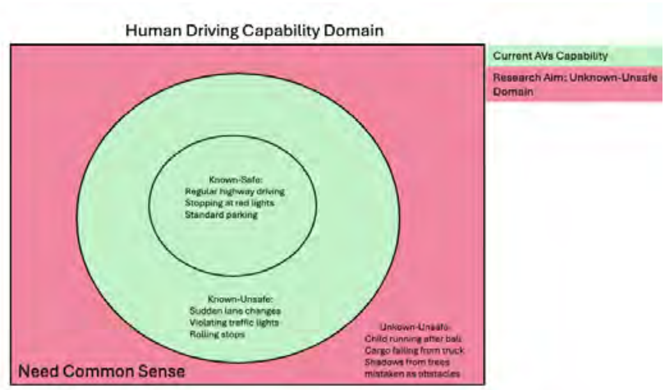
~~II. Backgruound~~

~~III. Methodology~~

IV. Conclusion

~~This paper explores the potential of integrating Large Language Models (LLMs) within a multi-agent framework for enhancing decision making in autonomous vehicles (AVs), particularly in challenging corner cases where traditional AV systems struggle. We focus on using LLMs as the main "decision-makers" in AV systems to see if they can think and react more like humans.~~

~~As autonomous vehicles become more common, they face many challenges, particularly in unpredictable driving environments. One of the biggest challenges is dealing with unknown unsafe driving situations which are complicated because they can happen without warning, require immediate decision-making, often combine multiple challenges at once, and may not have been experienced before by the driver.~~



~~Fig. 1: The domain is divided into Known Safe, Known Unsafe, and Unknown Unsafe regions.~~

~~Imagine that you are driving near a playground when a ball suddenly rolls onto the road. As a human, you can predict the possibility that a child might chase the ball. You can recognize this situation using your reasoning ability and previous experience. You instinctively identify it as a potential danger and take immediate action, such as slowing down or stopping to prevent an accident. The driving domain is divided into~~

- ~~Known Safe: These are situations where you know it is safe. For example, driving straight on an open highway with no obstacles or stopping at a red light and waiting for it to turn green.~~
- ~~Known Unsafe: These are situations where the danger is obvious. For example, running a red light, speeding in a school zone, or driving on the wrong side of the road.~~
- ~~Unknown Unsafe: These are unexpected and unplanned situations, for example, a ball rolling on the road or~~

~~a pedestrian suddenly crossing outside of a crosswalk. These are unpredictable situations that require quick judgment. Humans can use common sense and past experiences to react quickly and avoid danger.~~

~~Traditional autonomous vehicles have successfully handled known safe and known unsafe situations because these scenarios have clear rules that can be pre-programmed into their systems. However, unknown unsafe situations are much more challenging. They occur without warning and require immediate decision making while considering other vehicles and pedestrians. Human drivers use intuition and past experiences to navigate these unpredictable scenarios. In contrast, AVs rely on pre-defined models and programmed rules, making it difficult for them to recognize the danger or respond appropriately. This limitation poses a serious safety risk in real-world driving, where unexpected situations can frequently occur. In this research, we focus on using LLMs as the main decision makers within AV systems to assess whether they can think and react more like humans. Current AVs face difficulties in complex scenarios, particularly in uncertain situations where traditional pre-programmed responses can lead to issues such as false positives, incorrectly identifying safe situations as dangerous, or false negatives, failing to recognize actual dangers. We propose a system that combines two key potential components:~~

- ~~Showing the possibility of using LLMs as the core decision-making agents to provide human-like reasoning and contextual understanding, and~~
- ~~Showing the possible implementations of Multi Agent System that enables vehicle to vehicle communication and coordination with LLMs.~~

~~Our goal is to contribute to the development of, safer and more reliable Level 5 [2] autonomous vehicles. We aim to contribute for level 5 (complete automation [3]). We are seeing many examples where traditional AVs can take decisions in situations where the outcome is known, using pre trained models. However, they may struggle with unpredictable situations where their pre-programmed knowledge falls short. Human drivers can draw on common sense and past experiences to make sense of unusual situations, such as recognizing that traffic cones on a truck are not a threat. However, current AV systems, even those that use rule-based approaches or reinforcement learning (RL [4]), will struggle in these scenarios. This limitation is especially evident in what we call "unknown unsafe" region situations where the correct action is not immediately clear. [5] For AVs, this can lead to mistakes such as:~~

- ~~False Positives: The system incorrectly identifies something as dangerous when it is not, such as mistaking a shadow for an obstacle.~~
- ~~False Negatives: The system fails to identify something that is actually dangerous, such as not recognizing a small animal on the road.~~

~~As we mentioned, AVs may fail in these cases. Furthermore, even now autonomous vehicles (AV)and human drivers are actually sharing the road, creating mixed traffic scenarios. This co-existence presents significant challenges for traffic flow.~~

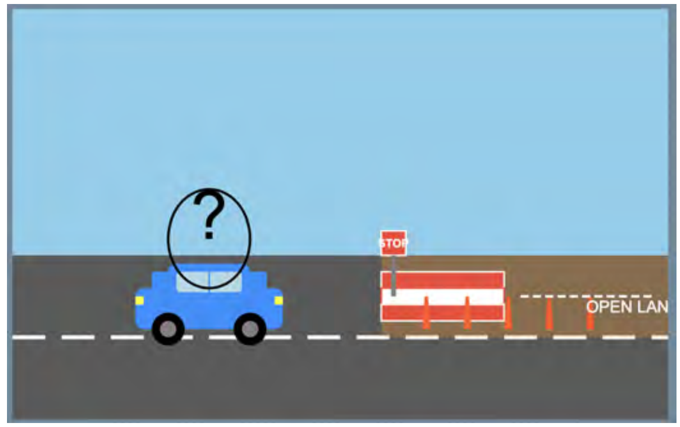


Fig. 2: What current AV might do in this situation?

Imagine a situation where a road is partially closed due to construction, with a "STOP" sign and traffic cones marking off part of the road. A typical autonomous vehicle (AV) might see the "STOP" sign and interpret it as a complete road closure, coming to a complete stop because that is what its pre-programmed data tell it to do. However, a human driver in the same situation might use common sense, realize that the road is only partially closed, and safely continue through the open section.

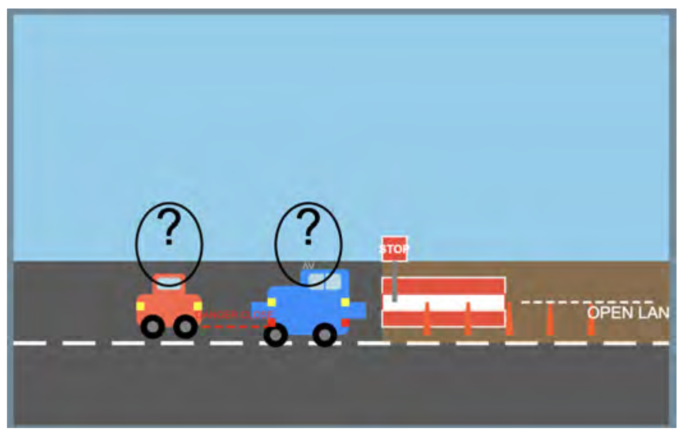


Fig. 3: What current AV might do in this situation?

Now, let us make it more complicated: an autonomous truck comes to a sudden stop when it sees the "STOP" sign. There is another car, driven by a human or another AV, right behind it. This sudden stop could easily cause a crash. Now imagine a human driver seeing a partially blocked road and choosing to carefully go through the open part. As they proceed, they notice an autonomous vehicle (AV) approaching in the opposite direction. Using common sense, the human driver would slow down to avoid any risk of collision. However, the AV does not recognize the approaching car as a possible obstacle. It might not slow down and could keep going at its normal speed, which could lead to a crash.

However, if the AV had better reasoning abilities, it might recognize that the road is only partially closed and proceed safely through the open part. And if the AV could communicate with the vehicle behind it or if the lane was suddenly closed, it could alert the other car about its actions, preventing a possible accident.

To address these challenges, we propose a possible approach on integrating Large Language Models (LLMs) into a multi-agent framework for autonomous driving control system. LLMs have the potential to provide commonsense reasoning that traditional systems lack, helping AVs better understand the context of complex, real-world scenarios. Incorporating Vehicle-to-Vehicle communication using multi-agent, which allows vehicles to exchange information with other vehicles. ~~This capability is also essential for creating realistic and complex driving scenarios.~~

~~**Can LLMs effectively enhance decision-making for AVs in Unknown-Unsafe domains within a multi-agent framework compared to current mechanisms?**~~

~~A fundamental challenge in autonomous vehicles is to prepare for truly unknown situations [6]. By definition, any scenario we design is "known" to us, even when we try to create situations that are unknown to the autonomous agents. This creates a paradox in testing genuine unknown-unsafe scenarios. Furthermore, creating realistic reproductions of complex real-world scenarios in simulation environments presents a significant problem. For example, unpredictable human behaviors and intricate environmental interactions can be difficult to replicate accurately. Verifying the effectiveness of LLM-based decision-making in truly unknown situations is also challenging. It is nearly impossible to test every possible edge case or unexpected scenario that could occur in real life. Human drivers rely on their common sense to navigate these situations, so we need to investigate whether LLMs can effectively mimic human-like reasoning. In our research, we will create specific scenarios and use them as inputs for LLM to see if they can demonstrate human-like reasoning in unknown unsafe situations. By examining how well LLMs can handle these scenarios, our objective is to determine whether they can improve the current struggles faced by AVs.~~

We are investigating the effectiveness of various Large Language Models (LLMs) for use as decision-making agents in autonomous driving systems. Our primary objective is to answer the research question, **Can LLMs effectively enhance decision making for AVs in Unknown-Unsafe domains within a multi-agent framework compared to current mechanisms?** During this investigation with these various LLM, we proceed under two assumptions: 1) Certain AIs can transform real-world situations into text-based explanations, which are subsequently used as inputs for the LLMs, and 2) The outputs generated by the LLMs can be translated into actual decisions made by car parking allocation systems by interpreting the responses from the LLMs.

Hugging Face is a platform where various open LLM models are available. LM Studio is a platform to test and integrate LLMs available at Hugging Face into a locally available system on ordinary computing devices even without powerful GPUs.In our research, we used google/flan-t5-large from

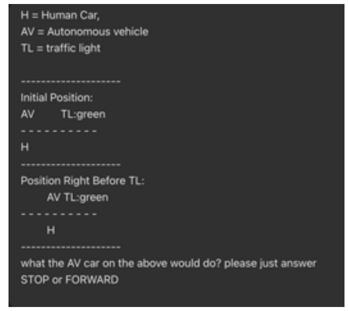


Fig. 4: Text-base scenario

Hugging Face and checked what responses it gave. We also tested cloud-based LLMs such as the GPT-4-turbo variant, DeepSeek-R1, and Claude 3.7 Sonnet to compare their results. For running models on our own computer using LM Studio, we picked We presented a text-base scenario showing the autonomous vehicle (AV) and human-driven car (H) approaching a green traffic light similar to the one in Fig. 4, to the language models. We used the following prompt: 'What would the AV do in this situation? Please, just answer STOP or FORWARD.' This prompt was given to all selected LLMs. The output is shown in Fig. 5

References

[1] R. Zhang, X. Guo, W. Zheng, C. Zhang, K. Keutzer, and L. Chen, "Instruct large language models to drive like humans," *arXiv preprint arXiv:2406.07296*, 2024.

[2] A. Balasubramaniam and S. Pasricha, "Object detection in autonomous vehicles: Status and open challenges," *arXiv preprint arXiv:2201.07706*, 2022.

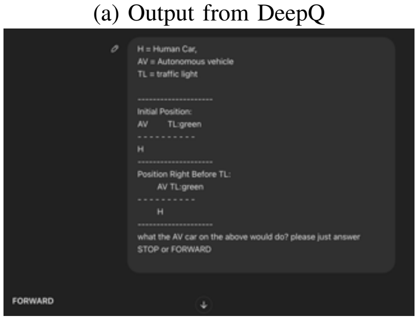
[3] M. Raza, "Autonomous vehicles: levels, technologies, impacts and concerns," *International Journal of Applied Engineering Research*, vol. 13, no. 16, pp. 12710–12714, 2018.

[4] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[5] D. Fu, X. Li, L. Wen, M. Dou, P. Cai, B. Shi, and Y. Qiao, "Drive like a human: Rethinking autonomous driving with large language models," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 910–919, 2024.

[6] T. Singh, E. van Hassel, A. Sheorey, and M. Alirezaei, "A systematic approach for creation of sotif's unknown unsafe scenarios: An optimization based method," tech. rep., SAE Technical Paper, 2024.

(a) Output from DeepQ



(b) Output from ChatGpt

(c) Output from google/flan-t5-larg

Fig. 5: Response from LM Studio suggested models