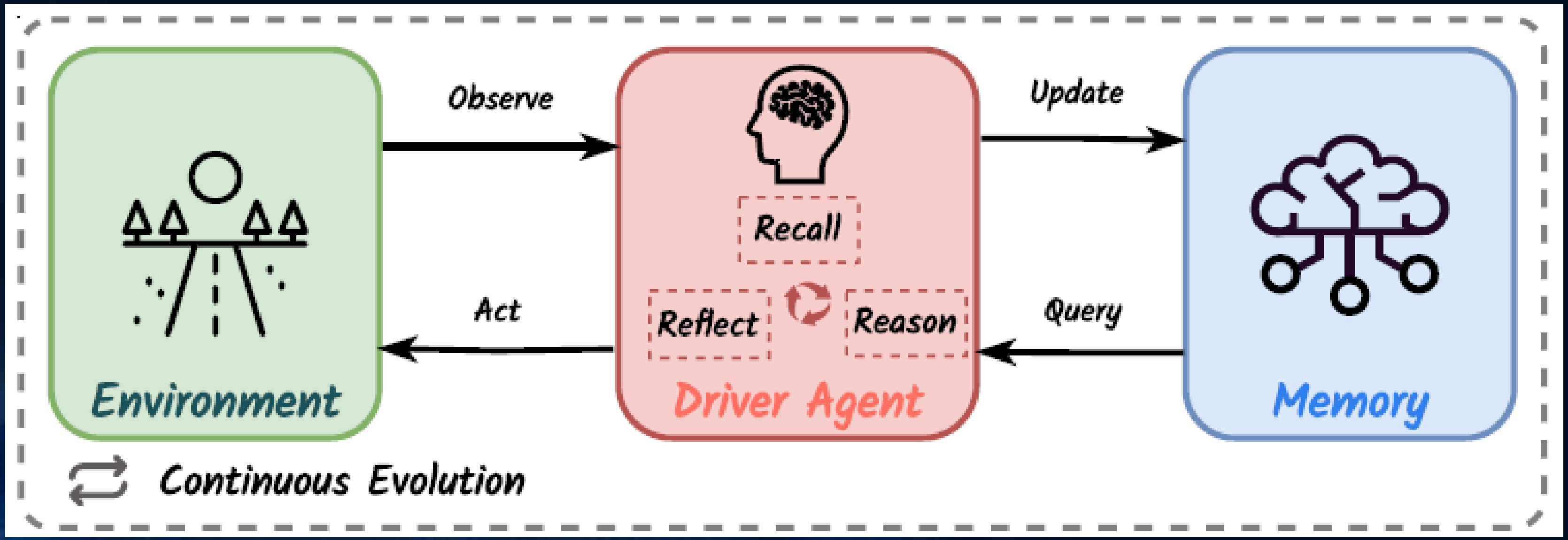
A Knowledge-driven approach to autonomous driving with large language models

BY SURESH RAVURI

Introduction to DILU

- The DILU (Decision-making Integration via Language Understanding) approach addresses the need for autonomous vehicles (AVs) to adapt to complex, dynamic environments using a hybrid of traditional data-driven methods and advanced knowledge-driven strategies. This integration helps AVs make decisions similar to human reasoning, ensuring safer and more efficient navigation.

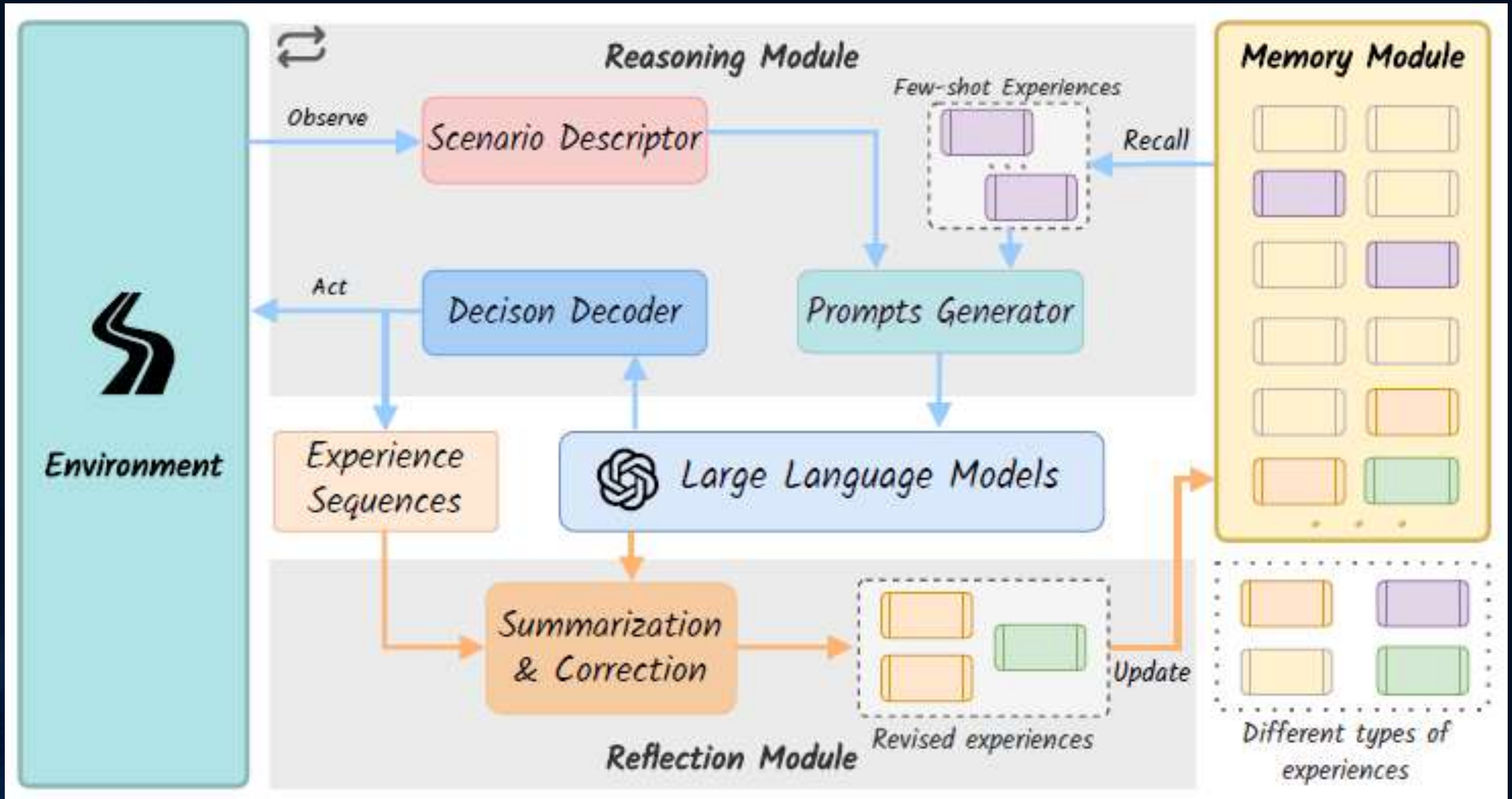


DILU Framework overview

The DILU framework consists of three main components:

- **Reasoning Module:** Utilizes LLMs to interpret sensor data and environmental context to formulate potential actions.
- **Reflection Module:** Evaluates the proposed actions based on past experiences and simulated future scenarios, ensuring decisions are safe and optimal.
- **Memory Module:** Stores outcomes of past decisions to refine future reasoning and reflection processes.

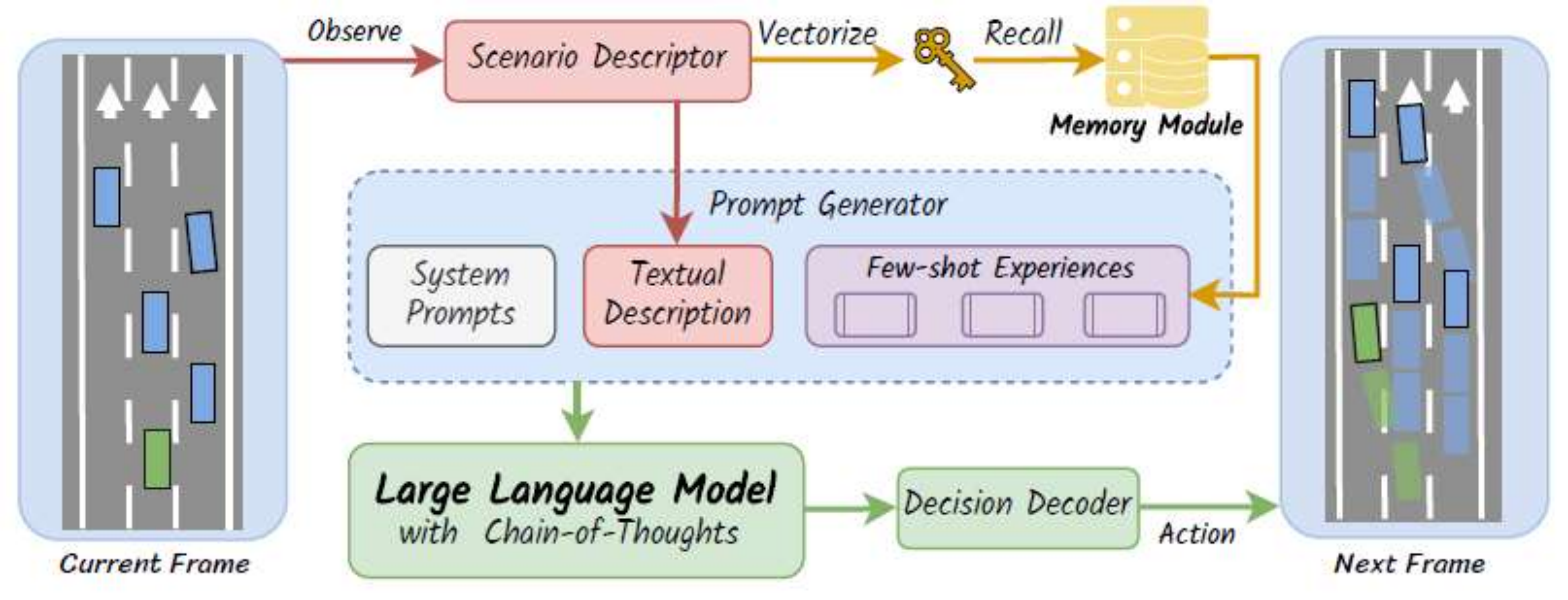
The framework of DiLu: It consists of four modules: Environment, Reasoning, Reflection, and Memory. In DiLu, the Reasoning module can observe the environment, generate prompts by combining scenario descriptions and experiences in the Memory module, and decode responses from the LLM to finish decision-making. Concurrently, the Reflection module evaluates these decisions, identifies the unsafe decision to the experiences, and finally updates the revised experiences into the Memory module.



Deep dive into the reasoning module

- **Scenario Encoding:** It begins by encoding the current traffic scenario into a descriptive text using a standard sentence structure, allowing for a comprehensive understanding of both static and dynamic elements within the driving scenario.
- **Memory Recall:** The module then recalls relevant past experiences from the Memory Module using the current scenario's description. This recall process involves a vectorized representation of the scenario to find similar past situations.
- **Prompt Generation:** With the recalled experiences, a tailored prompt consisting of the system prompts, scenario description, and few-shot experiences is generated. This prompt prepares the LLM for the reasoning task.
- **LLM Processing and Decision Decoding:** The LLM receives the prompt and uses Chain-of-Thought reasoning to articulate a step-by-step logic that culminates in a decision. This decision is then translated into actionable commands for the vehicle .

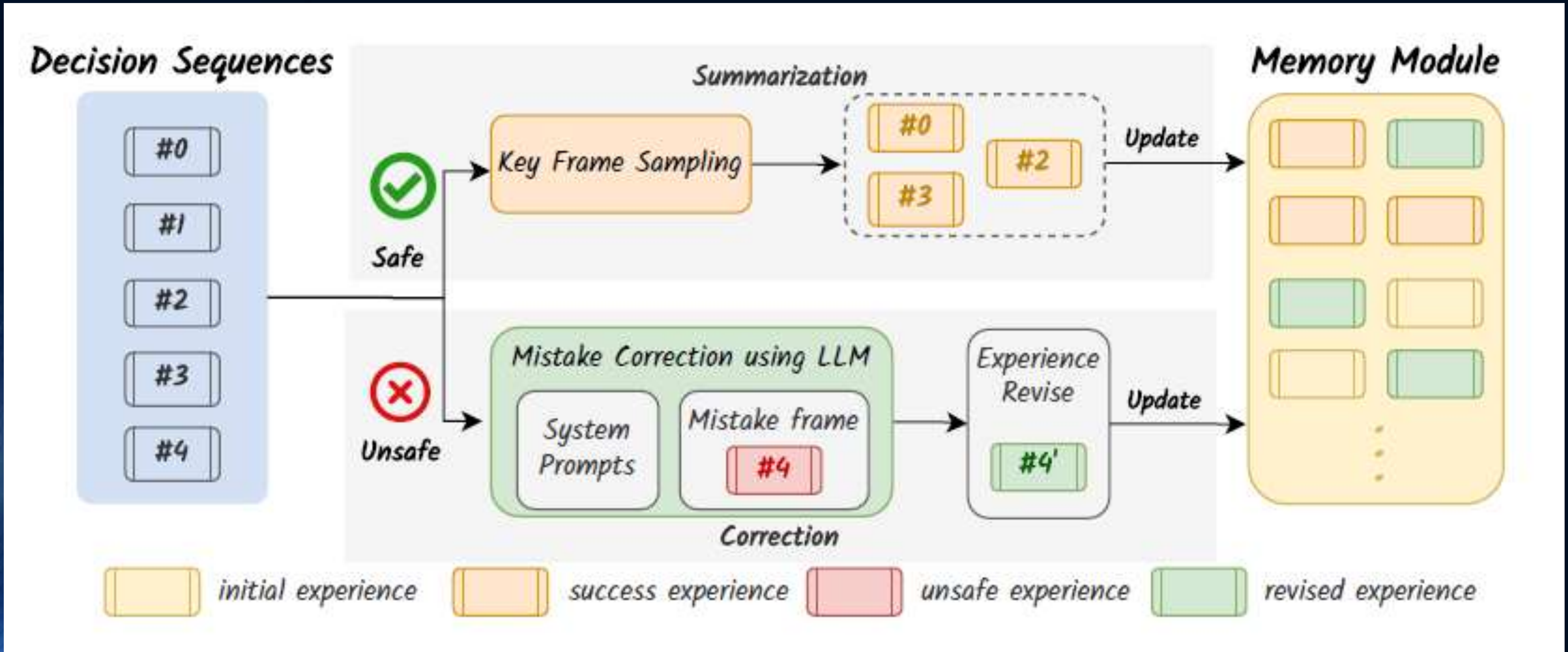
Reasoning module. We leverage the LLM's common-sense knowledge and query the
experiences from Memory module to make decisions based on the scenario observation.



Reflection Module Explained

- **Assesses Decisions:** The Reflection Module evaluates the actions and decisions made by the Reasoning Module during driving simulations, categorizing them as either safe or unsafe.
- **Error Analysis and Correction:** For unsafe decisions, such as those leading to collisions, the Reflection Module uses LLM to understand and correct the faulty reasoning that led to the decision.
- **Updates Memory Module:** The correct reasoning and decisions, after reflection, are stored in the Memory Module, enriching it with new, safe driving experiences and strategies for avoiding past mistakes.
- **Continuous Learning:** By processing the outcomes of past decisions, the Reflection Module allows DiLu to learn from experience, akin to a novice driver becoming more skilled over time .

Reflection module: The Reflection module takes recorded decisions from closed-loop driving tasks as input, it utilizes a summarization and correction module to identify safe and unsafe
decisions, subsequently revising them into correct decisions through the human knowledge embedded
in LLM. Finally, these safe or revised decisions are updated into Memory module.



Role of the Memory Module

- **Stores Experiences:** Archives past driving scenarios and decisions, facilitating a growing knowledge base.
- **Initial Setup:** Begins with pre-loaded basic scenarios to establish foundational decision-making processes.
- **Recalls Past Scenarios:** During operation, it helps recall similar past experiences to aid current decision-making.
- **Informs Decisions:** Uses stored experiences to provide context and support for new decisions.
- **Updates Continuously:** Adds new experiences and outcomes to enhance and update the knowledge base.

Experimental Results

- Detailed discussion of experimental setups using the Highway-env simulation environment, showcasing how DILU outperforms traditional methods in terms of decision accuracy, response time, and adaptability under various driving conditions.

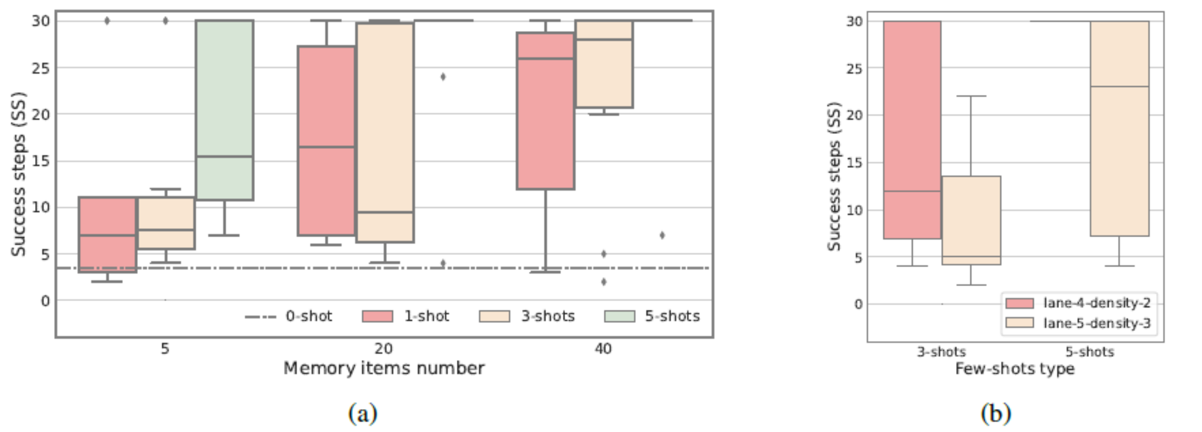


Figure 6: (a) Quantitative experiments with different experiences in Memory module and different few-shot numbers. Notably, the 5-shots setting achieve a maximum(30) simulation steps with both 20 and 40 memory items. (b) Generalizability experiment on environment with different traffic density using 20 memory items.

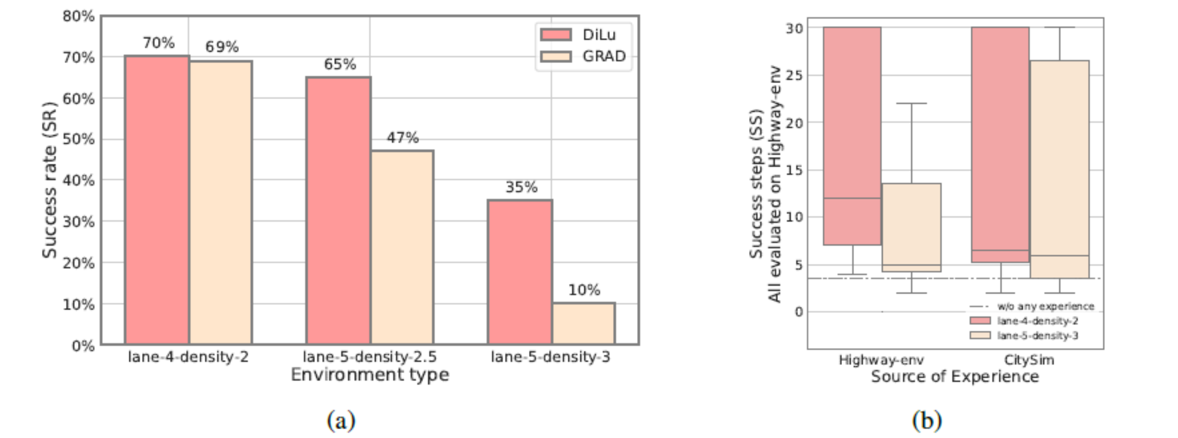


Figure 7: (a) Performance comparison with GRAD in different types of motorway environments. Both two methods are **only** optimized on `lane-4-density-2` settings and evaluate on `lane-4-density-2`, `lane-5-density-2.5` and `lane-5-density-3` respectively. (b) Experiments with experience from different domains.

Conclusion and Future Directions

- DiLu integrates human-level knowledge into autonomous systems, significantly enhancing decision-making and outperforming traditional methods in generalization.
- It currently experiences decision-making latencies of 5-10 seconds and occasional hallucinations from LLMs.
- Future improvements will focus on LLM compression and optimization to reduce latency and enhance both efficiency and effectiveness.

Acknowledgments and References

- The research was supported by Shanghai Artificial Intelligence Laboratory, the National Key R&D Program of China (Grant No. 2022ZD0160104) and the Science and Technology Commission of Shanghai Municipality (Grant Nos. 22DZ1100102 and 23YF1462900).

- Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. arXiv preprint arXiv:2303.08774, 2023.
- Daniel Bogdoll, Jasmin Breitenstein, Florian Heidecker, Maarten Bieshaar, Bernhard Sick, Tim Fingscheidt, and Marius Z¨ollner. Description of corner cases in automated driving: Goals and challenges. In Proceedings of the IEEE/CVF International Conference on Computer VisionWorkshop, pp. 1023–1028, 2021.
- Jan-Aike Bolte, Andreas Bar, Daniel Lipinski, and Tim Fingscheidt. Towards corner case detection for autonomous driving. In 2019 IEEE Intelligent vehicles symposium (IV), pp. 438–445. IEEE, 2019.

Thank You!