References

[1] J Mao, J Ye, Y Qian, M Pavone, Y Wang. "A Language Agent for Autonomous Driving," arXiv:2311.10813.
Tries to break down reasoning further than other methods to emulate human thinking processes better. Has the LLM choose to call or not call certain functions that give it more context based on the information given to it.

[2] K Jiang, X Cai, Z Cui, A Li, Y Ren, H Yu, H Yang, D Fu, L Wen, P Cai. "KoMA: Knowledge-driven Multi-agent Framework for Autonomous Driving with Large Language Models," In *IEEE Transactions on Intelligent Vehicles*, 2024. arXiv:2407.14239.
Created a framework that has multiple different agents controlled by LLMs, there is a shared memory module between all of the agents that allows all agents to learn faster as well as learn from each others mistakes. Uses Chain of Thought (CoT) reasoning to have the LLMs first analyze their goal based on the current environment, then make a plan for what to do, and finally take a singular action based on that plan. The memory module includes a ranking for each action taken that ranks them based on safety and efficiency.

[3] X Tian, J Gu, B Li, Y Liu, Y Wang, Z Zhao, K Zhan, P Jia, X Lang, H Zhao. "DriveVLM: The Convergence of Autonomous Driving and Large Vision-Language Models," arXiv:2402.12289.
Uses Visual Language Models (VLMs) that have chain of thought (CoT) reasoning that uses scene description, scene analysis, and hierarchical planning. VLMs by themselves aren't perfect so they combine VLMs with traditional 3d perception and planning modules to get the best possible results. The model is trained to output certain conditions that include weather, time, road, and lane.

[4] L Wen, D Fu, X Li, X Cai, T Ma, P Cai, M Dou, B Shi, L He, Y Qiao. "DiLu: A Knowledge-Driven Approach to Autonomous Driving with Large Language Models," arXiv:2309.16292.
Try to make driving knowledge driven not data driven, meaning to understand why an action should be taken in a scenario even if it hasn't been seen yet by generalizing from past experiences. Uses three core modules: reasoning, reflection, and memory. Memory stores experiences and is used by the reasoning module to make decisions on what actions to take. The reasoning module decides what to do based on the current state of the environment and past experiences. The reflection module categorizes whether or not an action taken was safe or unsafe. Unsafe decisions are revised and updated in the memory module. To identify correct unsafe actions they feed incorrect actions and the given reasoning at the time to the LLM and ask it to explain why the chosen action was wrong. This serves as a future explanation of why not to do what was done.

[5] C Cui, Y Ma, X Cao, W Ye, Z Wang. "Drive as You Speak: Enabling Human-Like Interaction with Large Language Models in Autonomous Vehicles," In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024. arXiv:2309.10228.
Used LLMs to plan maneuvers on command or in certain situations to keep driver safe, or assist them when needed.

[6] Y Jin, R Yang, Z Yi, X Shen, H Peng, X Liu, J Qin, J Li, J Xie, P Gao, G Zhou, J Gong. "SurrealDriver: Designing LLM-powered Generative Driver Agent Framework based on Human Drivers' Driving-thinking Data," In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024. arXiv:2309.13193.
Split driving memory into short term and long term, where short term keeps track of last few actions to ensure smooth sequences and long term memory is used to create longer plans. Some additional mandatory and recommended rules are given to the LLM to ensure that it follows safe driving practices.