

# AgentMatrix: Advancing towards Nested and Distributed Multi-Agent LLM Infrastructures

Qingxu Fu

**Abstract**—Recent advances in Large Language Models (LLMs) have propelled substantial developments in artificial intelligence (AI). Many researchers have leveraged LLMs as the foundation to build AI agent systems for two primary purposes: a) developing robust, efficient and user-friendly applications to overcome LLMs’ performance and input-domain limitations. b) refining datasets or reward models from existing datasets and LLMs, thereby augmenting the capability of LLMs in a “bootstrapping” way. Nevertheless, the existing LLM-agent frameworks have many limitations, such as the dependence on specific LLMs from certain companies, the lack of adaptability for novel scenarios, the insufficient scalability for extensive agent interactions, and the inability to respond in real time to user inputs like audio. To address these problems, we propose AgentMatrix, a multi-agent framework for LLM-based agents. AgentMatrix features a nested and distributed design: (1) The agents can be joined together by Directed Acyclic Graphs (DAGs), Directed Cyclic Graphs (DCGs), or autonomous free chat groups, thus composing into higher-level agents with stronger capability. Furthermore, the composed agents can join with agents (of any level) once again, forming a sophisticated and well-organized nested structure. (2) Different agents can be distributed into different physical devices, allowing the user terminal devices to integrate agents with high resource consumptions and prompting the scalability and flexibility of the system. AgentMatrix is accessible at <https://github.com/binary-husky/agent-matrix/>. [1][1]

**Index Terms**—Multi-agent systems, nested agents, large language models, supervised fine-tuning, reinforcement learning.

## REFERENCES

- [1] D. H. Wolpert and W. G. Macready, “No free lunch theorems for optimization,” *IEEE transactions on evolutionary computation*, vol. 1, no. 1, pp. 67–82, 1997.

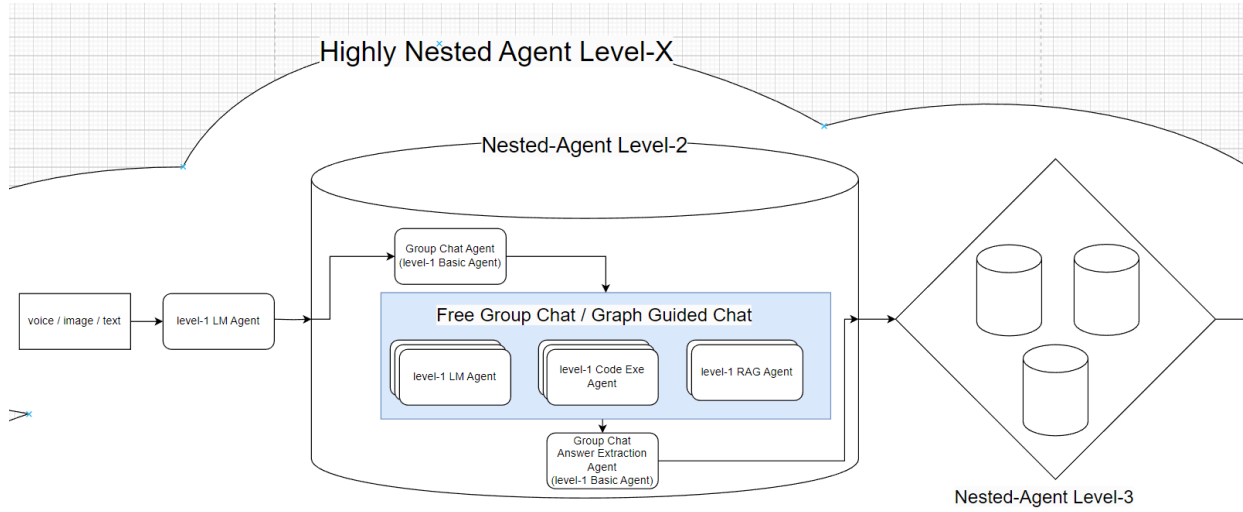


Fig. 1: Agents are allowed to establish nested cooperation structure in AgentMatrix.

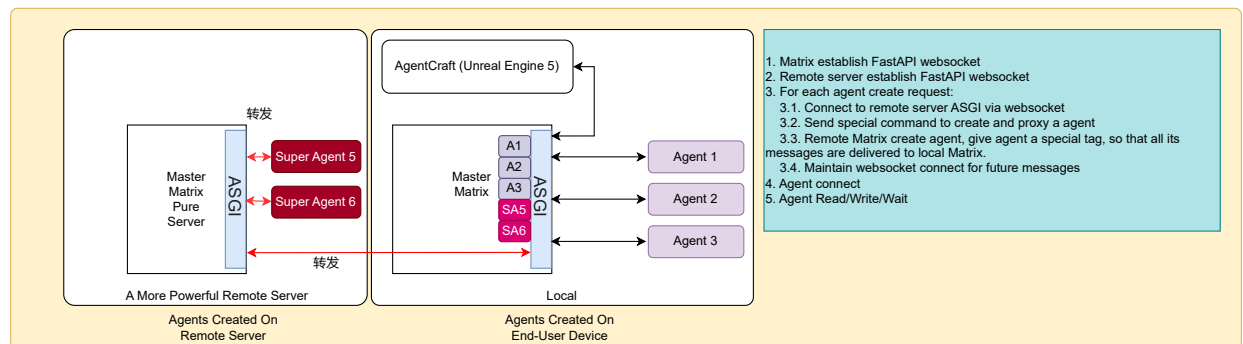


Fig. 2: The communication and distribution of agents in AgentMatrix framework.

## 智能体的任务与生命周期

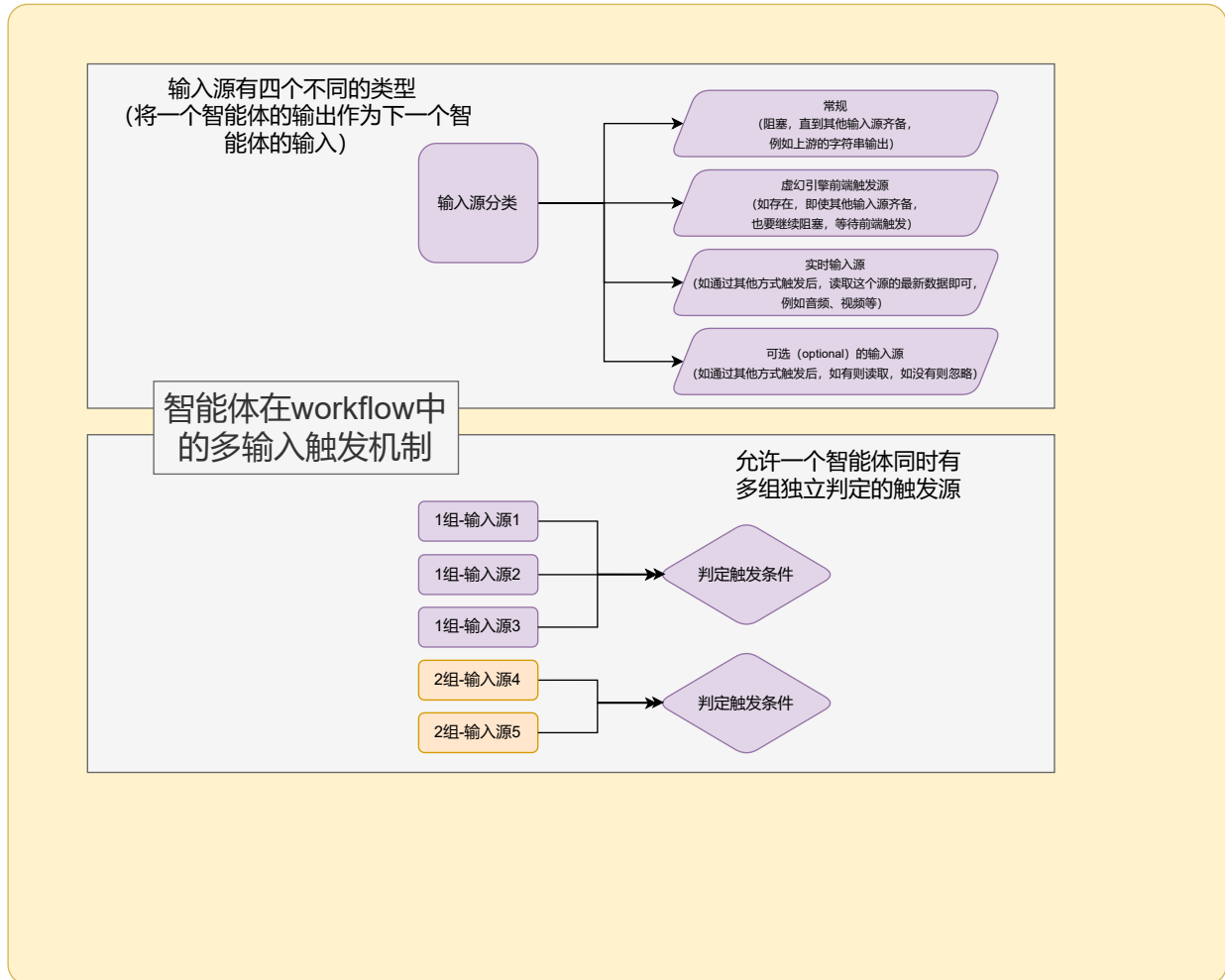


Fig. 3: Agent begin its task when receiving triggers. In most cases, agents should begin working when it receives message from its upstream agents. However, an agent can connect to multiple upstream agents simultaneously, and in some applications an agent also has to deal with real-time signal. Thereby, a multi-group triggering mechanism is designed to resolve these challenges.