MegaAgent

A Large-Scale Autonomous LLM-based Multi-Agent System Without Predefined SOPs

Qian Wang*, Tianyu Wang*, Zhenheng Tang, Qinbin Li, Nuo Chen, Jingsheng Liang, Bingsheng He

ACL 2025 VIENNA

ZHIYUAN COLLEGE

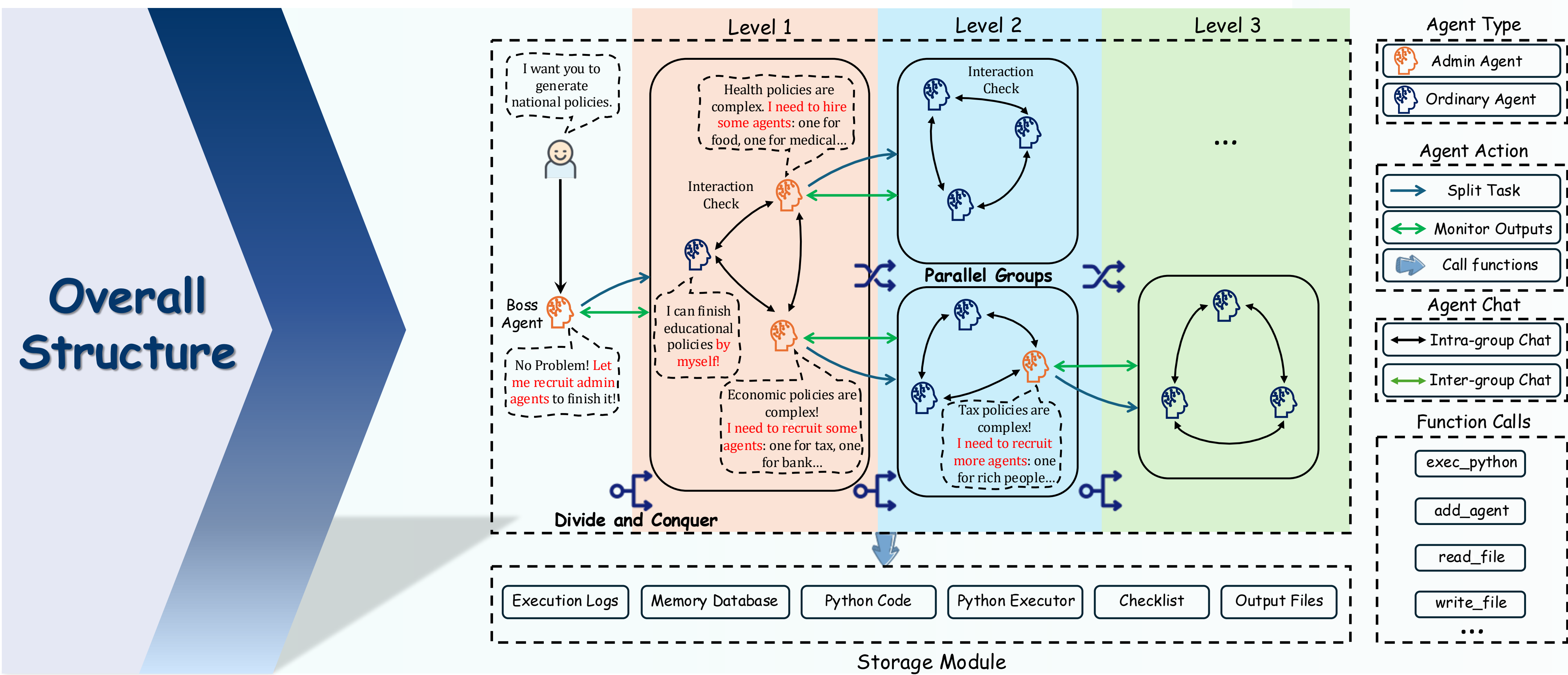NUS National University of Singapore

TL; DR

A multi-agent **SYSTEM** which pioneers **Large-Scale** autonomous multi-agent coordination

Employs concepts from OS to operate **without predefined SOPs**.

Outperforms existing multi-agent systems like MetaGPT **in scalability and efficiency**.

Get & Star Our Code!

Key Features

- Autonomous Coordination
- Dynamic Task Decomposition
- Scalability Enabled by **Psudo-Tree** (Up to 590+ Agents)

- **Parallel Execution**
- Hierarchical Monitoring
- **No Predefined SOPs**
- OS for multi-agent! (mutex, message queue, producer-consumer, ...)

Overall Structure

Experiments and Key Findings

RQ1: Can MegaAgent **perform well on standard benchmarks**?

YES! MegaAgent can perform well!

Model	MBPP	HumanEval	MATH	GSM-8k
MetaGPT	81.7	82.3	N/A	N/A
Camel	78.1	57.9	22.3	45.6
AgentVerse	82.4	89.0	54.5	81.2
AutoGen	85.3	85.9	**69.5**	87.8
MegaAgent	**92.2**	**93.3**	69.0	**93.0**

RQ2: Can MegaAgent complete a task requiring extensive coordination **without a predefined SOP**?

YES! MegaAgent can perform well while other baselines failed!

FUN FACT: Most Tokens Are Consumed in Task-Solving Stage. Notably, # of Input Tokens >> # of Output Tokens

Model	Error-Free Execution	User Move	AI Move	Game Termination	# of Agents	Time(s)	Time/Agent (s)
AutoGen	✓	✓	✗	✗	2	180	90
MetaGPT	✓	✓	✗	✗	6	480	80
CAMEL	✗	✗	✗	✗	2	1,830	915
AgentVerse	✗	✗	✗	✗	4	1,980	495
MegaAgent	✓	✓	✓	✓	7	800	114

Stage	# Input Tokens	# Output Tokens	# Total Tokens	Time (s)
Planning	42,947	12,347	55,294	0–60
Task-Solving	1,098,573	55,022	1,153,595	30–840
Merging	22,099	1,493	23,592	840–870
Total	1,163,619	68,862	1,232,481	870

RQ3: Can MegaAgent **be effectively scaled to more complex tasks** involving significant large number of agents?

YES! MegaAgent produces complete policies with 590 agents in 2,991 seconds!

Model	Outputs	# Agents	Time (s)	Time/Agent (s)
AutoGen	Outline	1	40	40
MetaGPT	Python Program	6	580	97
CAMEL	Plans	2	1,380	690
AgentVerse	None	4	510	128
MegaAgent	Complete Policies	590	2,991	5

Model	No Predefined SOP	Multi-file Support	Parallelism	Scalability
AutoGen	✗	✗	✗	✗
MetaGPT	✗	✓	✗	✗
CAMEL	✗	✗	✗	✗
AgentVerse	✓	✗	✗	✗
MegaAgent	✓	✓	✓	✓