← Go to **ICLR 2025 Workshop FM-Wild** homepage (/group?id=ICLR.cc/2025/Workshop/FM-Wild)

All It Takes Is One Prompt: An Autonomous LLM-MA System

PDF (/pdf?id=mSKj2bCWRr)

Qian Wang (/profile?id=~Qian_Wang25),
Tianyu Wang (/profile?id=~Tianyu_Wang13),
Zhenheng Tang (/profile?id=~Zhenheng_Tang2),
Qinbin Li (/profile?id=~Qinbin_Li1), Nuo Chen (/profile?id=~Nuo_Chen4),
Jingsheng Liang (/profile?id=~Jingsheng_Liang1),
Bingsheng He (/profile?id=~Bingsheng_He1)

Published: 06 Mar 2025, Last Modified: 06 Mar 2025 ICLR 2025 FM-Wild Workshop
Everyone Revisions (/revisions?id=mSKj2bCWRr) BibTeX

Keywords: Hierarchical Agent System

Abstract:
LLM-based multi-agent (LLM-MA) systems have shown promise in tackling complex tasks. However, existing solutions often suffer from limited agent coordination and heavy reliance on predefined Standard Operating Procedures (SOPs), which demand extensive human input. To address these limitations, we propose \textit{MegaAgent}, a framework designed for autonomous coordination in LLM-MA systems. \textit{MegaAgent} generates agents based on task complexity and enables dynamic task decomposition, parallel execution, efficient communication, and comprehensive system monitoring of agents. In evaluations, \textit{MegaAgent} demonstrates exceptional performance, successfully developing a Gobang game within 800 seconds and scaling up to 590 agents in a national policy simulation to generate multi-domain policies. It significantly outperforms existing systems, such as MetaGPT, in both task completion efficiency and scalability. By eliminating the need for predefined SOPs, \textit{MegaAgent} demonstrates exceptional scalability and autonomy, setting a foundation for advancing true autonomy in LLM-MA systems. Code is available at \url{https://anonymous.4open.science/r/MegaAgent-dev-DEF0} (https://anonymous.4open.science/r/MegaAgent-dev-DEF0%7D)

Submission Number: 62

Filter by reply type... Filter by author... Search keywords... Sort: Newest First
Everyone Program Chairs Submission62 Reviewers Submission62 Authors
Submission62... Submission62... Submission62...
3 / 3 replies shown

Paper Decision

Decision by Program Chairs 06 Mar 2025, 03:50 (modified: 06 Mar 2025, 10:20)
Program Chairs, Reviewers, Authors Revisions (/revisions?id=BkkFEnz4eJ)

Decision: Accept

The paper introduces the MegaAgent framework, which enhances multi-agent system coordination and efficiency by autonomously generating agents and managing tasks hierarchically. However, its effectiveness can be compromised by initial task decomposition inaccuracies and the head agent's validation errors, potentially leading to inefficiencies and outputs that don't meet user needs in

(https://openreview.net/forum?id=mSKj2bCWRr&noteId=QrUuKZ6pzT)

Official Review by Reviewer mwsK 28 Feb 2025, 19:14 (modified: 06 Mar 2025, 11:25)

Program Chairs, Reviewers Submitted, Reviewer mwsK, Authors Revisions (/revisions?id=QrUuKZ6pzT)

Review:

This paper introduces the MegaAgent framework, designed to enhance the coordination and efficiency of large language model-based multi-agent systems.

Pros: (1) It autonomously generates agents and dynamically decomposes tasks based on complexity, facilitating efficient task execution without predefined Standard Operating Procedures. (2) The framework uses a hierarchical structure for task management and communication. (3) It incorporates sophisticated monitoring and error-handling mechanisms to ensure accurate task completion.

Potential Cons: (1) While MegaAgent's autonomous capabilities significantly enhance operational efficiency, the system's effectiveness is highly contingent upon the initial task decomposition executed by the managing agent. If the manager prematurely halts task decomposition—assuming a simpler task structure when complexity actually warrants further breakdown—this initial oversight can lead to inefficiencies or incomplete task execution across the system. (2) The framework's reliance on autonomous agent groups to effectively communicate and converge on consistent results could be undermined if these groups do not effectively share insights or synchronize their understanding and responses to evolving task demands. This can result in fragmented outputs or inconsistencies that detract from the system's overall performance and reliability in dynamic or complex operational environments. (3) The effectiveness of MegaAgent relies heavily on the head agent’s ability to validate outputs for user intentions. If the head agent misjudges task nuances or the relevance of the results, it may approve technically correct but contextually inappropriate outputs, leading to solutions that fail to meet the actual needs of the user.

Rating: 8: Top 50% of accepted papers, clear accept
Confidence: 4: The reviewer is confident but not absolutely certain that the evaluation is correct

MegaAgent, an autonomous LLM-based multi-agent system that permits scalable, parallel, and adaptive task coordination, is

presented in the paper. With successful 590-agent simulations, it significantly outperforms existing frameworks like MetaGPT and AutoGen in terms of efficiency and scalability. While highly promising, improvements in cost efficiency, failure analysis, and generalization across domains would enhance its impact.

Official Review by Reviewer ToVj (Vivek Kumar Mishra (/profile?id=~Vivek_Kumar_Mishra1))

25 Feb 2025, 08:46 (modified: 06 Mar 2025, 11:25)

Program Chairs, Reviewers Submitted, Reviewer ToVj, Authors Revisions (/revisions?id=xxPwCFbd3w)

Review:

Quality: A well-structured and technically sound approach to LLM-based multi-agent systems (LLM-MA) is presented in this study, addressing key framework limitations. Experiments in software development and large-scale simulations are used to rigorously test the MegaAgent model, which demonstrates superior scalability and efficiency.

Strengths : Well-Designed Evaluation: A comprehensive performance analysis is ensured by the paper's clear comparisons to AutoGen, MetaGPT, CAMEL, and AgentVerse. Scalability: The ability to orchestrate 590 agents in a simulation is a significant improvement over previous methods, demonstrating its efficacy. Technical Depth: Provides a robust architecture for agent-based LLM collaboration and covers hierarchical task decomposition, parallel execution, and real-time monitoring.

Limitations: Even though there are monitoring mechanisms, the paper does not go into detail about failure scenarios like possible deadlocks in multi-agent interactions. Comparative Fairness: It is not explicitly mentioned that some improvements could be the result of implementation optimizations rather than fundamental architectural superiority.

Clarity: The paper is well-organized and breaks down the framework step by step. However, some sections appear excessively technical, making them difficult for non-experts to access.

Strengths: Logical Flow: The hierarchical task delegation and monitoring mechanisms are clearly articulated, making it easy to understand how MegaAgent operates.
Illustrative Tables and Figures: The inclusion of novelty comparisons and agent hierarchy diagrams improves the paper's clarity. Mathematical Justifications: Execution time improvements, such as the transition from O(n) to O(log n) complexity, are explained with technical precision.

Areas for Improvement
Dense Terminology: The descriptions of agent recruitment strategies and inter-agent communication could be simplified for broader comprehension.
Lack of Real-World Analogies: While the OS-inspired framework is intuitive, more real-life examples would improve engagement.

Originality The paper introduces an autonomous coordination strategy for LLM-based multi-agent systems, eliminating manual SOPs—a key bottleneck in previous approaches.

Strengths: New Approach to Task Coordination MegaAgent recruits and assigns tasks dynamically, in contrast to MetaGPT and AutoGen, which rely on predefined agent roles. Inspiration from OS Design: A novel and scalable strategy, the Boss-Agent/Admin-Agent hierarchy mirrors process-thread execution. Self-Correction Mechanism:

Implements memory persistence via vector databases, preventing task repetition and hallucination a common LLM issue.

Areas of improvement: Although novel, previous multi-agent LLM research has investigated agent communication and task decomposition. The paper focuses solely on hierarchical management, ignoring graph-based or decentralized agent systems. Limited Exploration of Alternative Coordination Models

Significance
The study offers scalable solutions for software development, policymaking, and large-scale simulations with real-world implications for multi-agent collaboration in AI-driven automation.

Strengths
Scalability Breakthrough: Successfully scales to 590 agents, far surpassing previous models.
Potential for Industry Adoption: Could revolutionize automated coding, strategic planning, and AI-driven research collaboration.
Open-Source Contribution: The availability of code promotes reproducibility and encourages further innovation.

Challenges
Computational Overhead: Requires significantly more tokens than competitors (8.8M tokens in policy simulation), raising concerns about practical implementation costs.
Domain Generalization Isn't Tried: There aren't any experiments done on different problem areas like legal technology, finance, or medical AI, which limits its applicability.
Rating: 8: Top 50% of accepted papers, clear accept
Confidence: 4: The reviewer is confident but not absolutely certain that the evaluation is correct