Journal of Computer Technology and Software
ISSN:2998-2383
Vol. 5, No. 9, 2026

An LLM-Agent Framework for Adaptive Task Decomposition and Continual Strategy Updating in Non-Stationary Environments

Yi Hu[1*]
[1]University of Southern CaliforniaLos Angeles, USA
*Corresponding author: Yi Hu; xqfqhu@gmail.com

Abstract: The study addresses the challenges faced by agents in dynamic and uncertain environments, where decision-making is easily disrupted, task structures are difficult to maintain, and strategies lack continuous adaptability. It proposes an adaptive task decomposition and strategy updating method grounded in large model reasoning. The approach first introduces state modeling and semantic context encoding mechanisms that capture environmental non-stationarity, allowing the agent to acquire and integrate temporal information throughout long-term interactions. Building on this foundation, an adaptive task decomposition module dynamically generates hierarchical task structures through semantic reasoning, enabling the agent to preserve coherent execution even when task goals change, disturbances intensify, or feedback becomes incomplete. In parallel, the strategy updating mechanism adjusts decision distributions based on real-time feedback, allowing rapid recovery and stable behavior when action deviations, sudden scene changes, or state noise occur. These components are integrated into a unified closed-loop reasoning framework that equips the agent with structural understanding, behavioral adjustment, and robust execution capabilities in complex scenarios. Systematic evaluation across multiple key metrics demonstrates that the method improves task completion, planning consistency, error recovery, and decision stability, highlighting its potential value for complex tasks in diverse application domains.

Keywords: Adaptive task decomposition; strategy update; large language model agent; dynamic uncertain environment

1. Introduction

Against the backdrop of continuously evolving complex systems, agent technology is moving from static and rule-based task settings toward more dynamic, variable, and uncertain real-world environments. As environmental states shift, decision chains lengthen, and external disturbances accumulate, traditional agents that rely solely on predefined strategies or fixed planning patterns struggle to maintain stable and effective task execution[1]. At the same time, the rise of large models has endowed agents with new cognitive, reasoning, and abstraction capabilities. These capabilities offer a promising path to overcome long-standing limitations in task understanding, multi-step reasoning, contextual association, and strategy generalization. However, when large model-driven agents operate in dynamic and uncertain scenarios, they still face fundamental challenges related to understanding evolving task structures and making effective decisions in the absence of complete prior knowledge. These challenges remain central issues for both the research community and industry[2].

As task complexity increases, single-level planning approaches reveal clear limitations. Task decomposition has therefore become an important means for agents to handle uncertainty and complex dependency structures. In real applications, tasks often span long time periods, involve multiple objectives, and require adaptation to changing constraints. Agents must infer the overall task structure and adjust the granularity and order of decomposition based on environmental changes. The semantic reasoning capability of large models provides new opportunities for task decomposition. It allows agents to extract structured information from multimodal inputs and high-dimensional contexts and to form a high-level understanding of task hierarchies. Yet current large model agents rely heavily on static prompts or manually crafted decomposition templates. These designs limit their ability to remain consistent and reliable when external conditions shift, information is incomplete, or task goals evolve. This creates an urgent need for adaptive and dynamic decomposition mechanisms that improve agent robustness in open environments[3].

Uncertain environments require flexibility not only in task understanding but also in strategy execution. Agents must continually update their decision patterns when confronted with random environmental transitions, unpredictable state changes, or delayed feedback. Fixed policies often fail to maintain long-term performance. They lead to policy rigidity, loss of contextual information, and drift from the original objective. Although large models have strong generative and reasoning abilities, they lack the inherent capability to iteratively refine strategies based on environmental feedback. Building a decision framework that can absorb new information, detect task changes, and update strategies in real time is therefore essential. This is a key step in advancing large model agents from being functional to becoming reliable. The challenge becomes even more pronounced in multi-task environments, cross-domain transfer scenarios, and tasks that require long-range reasoning. These constraints form a major

bottleneck that limits the development of higher agent autonomy[4].

With the growth of deep learning and the rise of semantic capabilities in large models, agents are moving toward multimodal interaction, open-ended task handling, and multi-stage reasoning. This evolution introduces new difficulties. The semantic space of tasks expands. Environmental states become non-stationary. Feedback signals contain noise or bias. These factors require stronger structural understanding and more flexible adaptation mechanisms throughout the execution process. Adaptive task decomposition helps agents reorganize task structures. Dynamic strategy updating improves decision quality over time. Together, they provide a foundation for building agents that can operate effectively over long horizons, generalize across environments, and maintain stable performance under high uncertainty. These abilities are not only technical improvements but also central to enhancing agent autonomy. They support the transition of large model agents from reactive tools to autonomous intelligent systems.

It is therefore of significant importance to investigate adaptive task decomposition and strategy updating for large model agents in dynamic and uncertain environments. Such methods provide critical support for applying large models to complex real-world tasks. They help agents overcome the limitations of static configurations and single-step reasoning and move toward continuous learning and continuous decision making. They also strengthen the stability, reliability, and interpretability of agent reasoning and decision-making in practical scenarios such as production, intelligent manufacturing, risk management, scheduling, autonomous driving, and intelligent operations. Moreover, adaptive mechanisms establish a foundation for future artificial intelligence systems. These systems must maintain task execution capability when rules are uncertain, information is incomplete, or objectives shift. They also need long-term autonomy, transferability, and sustainable cooperation. This research direction is therefore essential for building the next generation of intelligent agent systems.

2. Related work

In recent years, large model-driven agents have become an important direction in intelligent decision-making and task execution. With strong capabilities in semantic understanding, logical reasoning, and contextual modeling, these agents can handle more complex task pipelines in open environments and demonstrate stronger cross-domain and multimodal generalization than traditional methods. Existing studies explore instruction following, chain reasoning, tool use, multi-step planning, and autonomous execution. These advances allow agents to parse task intentions from natural language and generate high-level strategies[5]. However, many of these methods rely on fixed prompt structures or predefined interaction templates. They struggle when task objectives evolve, when environmental feedback is unstable, or when contextual information is incomplete. In addition, many frameworks emphasize decision-making while overlooking adaptability to environmental changes. As a result, they fail to adjust strategies or restructure tasks under varying conditions, which limits their reliability in real deployments.

Task decomposition is a key component of structured decision processes. It has been widely studied in complex planning, multi-stage reasoning, and cross-domain task execution. Traditional approaches depend on manually designed hierarchical planning systems or fixed templates to divide tasks into executable subtasks. With the development of deep models, task decomposition has shifted toward semantic understanding, contextual reasoning, and large model-based generation. Agents can now identify task logic, constraint relations, and execution order more effectively. Yet these methods often assume static task structures. When task goals change, external disturbances increase, or environmental constraints are redefined, decomposition strategies lack flexibility. Moreover, most existing approaches focus on one-time decomposition. They rarely monitor or revise decomposition results during execution. This limitation makes it difficult for agents to maintain robustness in long-horizon tasks[6].

Research on strategy updating and adaptive decision making focuses on how agents adjust their actions based on environmental feedback. Early studies rely on reinforcement learning and improve strategies through trial and error. However, for complex, high-dimensional, and dynamic real-world environments, pure reinforcement learning often suffers from slow convergence, high cost, and weak knowledge transfer. With the progress of large model reasoning capabilities, some methods integrate language reasoning with policy correction. Agents can modify their decision processes by interpreting feedback, describing environmental changes, or summarizing past behaviors. Yet these methods remain passive. Agents adjust only after failures or significant deviations. They cannot actively monitor environmental changes, detect potential risks, and adjust strategies in advance. Even when certain correction mechanisms exist, most agents cannot link task decomposition with strategy updating. This disconnect prevents the formation of a closed-loop adaptive framework[7].

Existing studies show that large models enhance language understanding and reasoning for agents. However, there are clear gaps in achieving continuous task decomposition and adaptive strategy updating in dynamic and uncertain environments. Decomposition mechanisms lack sensitivity to environmental variation. Strategy updating lacks awareness of task structure. Coordination between the two is also missing. Current research mainly focuses on improving reasoning ability while paying less attention to long-term stability, robustness, generalization, and self-recovery in open environments. Therefore, building a unified framework that integrates task decomposition and strategy updating for large model agents in dynamic and uncertain environments is essential. It is a key step toward practical deployment and an important direction for improving long-term autonomy and environmental adaptability. Research in this area is likely to bring structural progress to large model agents. It can enable them to understand environmental changes continuously,

adjust task structures proactively, and maintain stable and controllable decision evolution.

3. Proposed Framework

In dynamic and uncertain environments, the first step is to construct an environmental model capable of representing state evolution in real time, enabling agents to capture key trends in the decision-making chain. To this end, the environment is treated as a non-stationary process, and a state transition kernel is used to characterize its changes over time, allowing the agent to make inferences even with incomplete information. The evolution of the environmental state s_t can be described as follows:

$$s_{t+1} = f(s_t, a_t, \xi_1)$$

Here, a_t represents the current action, ξ_1 represents unpredictable disturbances, and $f(\cdot)$ characterizes the time-varying properties of the dynamic environment. Based on this, a contextual understanding function based on a large model is introduced to compress historical information into an abstract semantic representation to support task structure inference.

$$h_t = \Phi(s_1, s_2, \ldots, s_t)$$

Here, h_t is used to support subsequent task decomposition and policy generation. By modeling environmental uncertainties, the agent can continuously capture environmental drift during the execution process, providing a foundation for adaptive reasoning. Its model architecture is shown in Figure 1.

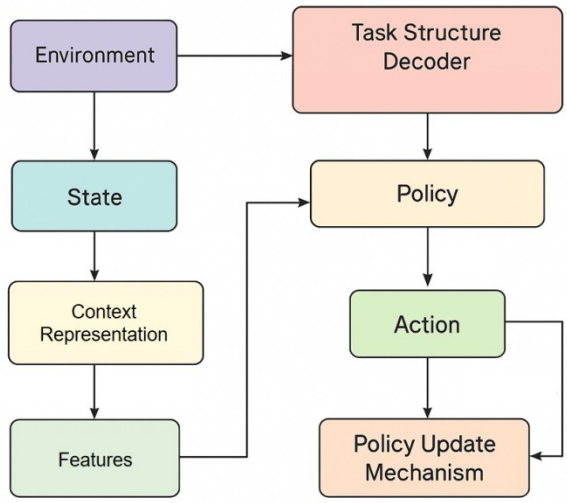


Figure 1. Overall Model Architecture

To enable agents to understand complex tasks at the semantic level, task decomposition is viewed as a process of generating a sequence of sub-objectives g from a high-level objective $\{g_1, \ldots, g_n\}$. Based on the reasoning capabilities of the large model, a task structure decoder is defined to generate an executable sub-task graph structure:

$$\{g_1, \ldots, g_n\} = \psi(g, h_t)$$

Here, ψ is a generator function that infers the hierarchical relationships of tasks using contextual semantics and environmental state. To ensure task executability, consistency constraints are introduced to guarantee that the dependencies between subtasks satisfy the topological order:

$$C(g_i, g_j) = 1$$

Here, C represents the dependency on consistency checks. This deconstruction approach allows the agent to reorganize the task structure according to the current environment, rather than relying on a static template, thus naturally adapting to external changes.

During the policy generation and update phase, the agent needs to continuously improve its decision-making behavior based on the decomposed sub-tasks and environmental feedback. The agent's policy can be defined as:

$$\pi(a_t | s_t, g_i, h_t)$$

This is achieved by mapping the current subtask, state, and semantic context to the optimal action. To enable dynamic updates, a policy adjustment term is introduced, allowing the agent to self-correct using feedback information.

$$\pi_{t+1} = \Omega(r_t, s_{t+1}, h_{t+1})$$

Where r_t is the feedback signal, and Ω represents the function that extracts policy correction information from the feedback. The final policy is updated iteratively.

$$\pi_{t+1} = \pi_t + \Delta\pi_t$$

This enables dynamic adaptation to long-term task chains, making the decision-making process no longer a static rule matching, but a continuously optimized evolutionary system.

By unifying and integrating environment modeling, task decomposition, and policy updates, a closed-loop, large-scale agent adaptive inference framework can be formed. In this system, the environment model provides real-time state change information, the task decomposition module reconstructs the sub-task structure based on the new context, and the policy module continuously optimizes the action plan based on feedback. The cyclical relationship between these three components can be abstracted as follows:

$$(s_t, g) \rightarrow \{g_1, \ldots, g_n\} \rightarrow a_t \rightarrow s_{t+1}$$

This unified framework enables agents to maintain reasoning consistency, goal stability, and behavioral robustness in dynamic environments. It not only allows agents to maintain continuous task execution under external disturbances but also enables them to automatically adjust their deconstruction and strategies when goals change or constraints are redefined, thereby achieving a highly adaptable and generalizable autonomous decision-making system.

4. Experimental Analysis

4.1 Dataset

This study adopts the ALFRED dataset as the primary evaluation benchmark. The dataset is built on an interactive indoor simulation environment and provides complex interaction sequences across multiple scenes, tasks, and action

chains. It covers navigation, object manipulation, task planning, and other types of instruction execution. The layout, object states, and interaction outcomes in the environment are highly variable. Agents must handle significant uncertainty and long-term dependencies during execution. These characteristics align well with the research objectives of adaptive task decomposition and strategy updating in dynamic environments.

Tasks in ALFRED consist of high-level goals and multi-step subtasks. Agents must understand the semantics of the instructions and perform long-horizon reasoning, state tracking, and behavioral adjustment in a dynamic environment. The dataset includes rich scene states, visual observations, action sequences, and natural language descriptions. These elements provide strong supervision for building large model-driven context understanding modules. In addition, the interactability and variability of objects, together with the diversity of task chains, make ALFRED an ideal platform for evaluating structured decision making.

The dataset is fully open source and offers a unified simulation interface and a reproducible environment. This design enables fair comparison under the same task configurations and uncertain conditions. Its dynamic properties, cross-task transfer settings, and high-level semantic structures required for task decomposition match the focus of this study. Research based on this dataset can verify the reasoning consistency of agents in complex environments. It can also evaluate their sensitivity to environmental changes and their ability to evolve strategies in long-horizon tasks.

4.2 Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table 1. Comparative experimental results

Method	Task Success Rate	Path Efficiency	Error Recovery Rate	Action Prediction Entropy
V2xpnp[8]	41.3	0.54	18.7	2.91
Muma-tom[9]	47.8	0.59	22.4	2.66
SciAgents[10]	52.1	0.62	27.9	2.53
Mapcoder[11]	55.6	0.65	31.2	2.47
Ours	63.4	0.71	39.5	2.11

From the table, it can be observed that all methods show a clear stratified performance pattern under dynamic and uncertain environments. Traditional methods face evident limitations when handling changing environments, long task chains, and complex interactions. Their task success rate and path efficiency remain low due to insufficient sensitivity to environmental variations, limited global semantic understanding, and the inability to reconstruct task structures in real time. As the methods progress from V2xpnp and Muma tom to SciAgents and Mapcoder, overall performance improves. However, they still struggle to maintain stable task completion when environmental disturbances are frequent, and task structures change dynamically. This trend reflects the weaknesses of traditional architectures in long-horizon reasoning and structured planning.

In terms of path execution efficiency, Path Efficiency increases as model complexity and semantic understanding improve. Yet all methods still show delayed responses to environmental changes and limited understanding of task chain structures. V2xpnp and Muma tom often fail to adjust their trajectories when the environment shifts, which leads to low path efficiency. Even SciAgents and Mapcoder exhibit trajectory drift, redundant actions, and inconsistent strategies in tasks with long-term dependencies. These observations indicate that strategies relying only on static planning or local adjustments cannot maintain globally consistent behavior in complex scenarios.

The Error Recovery Rate provides further insight into differences in adaptive strategy capability. Most baseline methods do not include an explicit strategy updating mechanism. When deviations occur during execution or when object states or environmental conditions change, the agents often fail to recover the task sequence effectively. As a result, their recovery rates remain low. In contrast, models with adaptive strategy updating can rapidly reconstruct task understanding and adjust action sequences when facing errors or environmental shifts. Their recovery rates are significantly higher than those of traditional methods. This pattern highlights the importance of strategy updating modules for improving long-term robustness.

The comparison of Action Prediction Entropy further illustrates differences in decision stability among methods. Baseline methods often show dispersed action selection distributions, indicating high uncertainty in their strategies under dynamic environments. As environmental complexity increases, the randomness and inconsistency of their action outputs become more pronounced. The entropy score of the proposed method is the lowest, which shows that its strategy maintains a more stable action distribution under environmental disturbances. It relies less on random search and more on structured reasoning and feedback-driven policy adaptation. Overall, the results demonstrate that adaptive task decomposition and dynamic strategy updating can significantly enhance task completion, decision stability, and recovery capability in uncertain environments. These improvements align closely with the core objectives of the proposed model.

In the adaptive task decomposition and strategy update framework, the learning rate is a crucial hyperparameter that directly affects the convergence speed of the strategy and the stability of task execution. When the agent continuously iterates its strategy in a dynamic and uncertain environment, an excessively large or small learning rate will affect its ability to respond to environmental changes and maintain the task chain. Therefore, it is necessary to conduct a systematic sensitivity analysis of different learning rate settings to clarify their impact on task success rate and guide a more reasonable hyperparameter selection. The experimental results are shown in Figure 2.

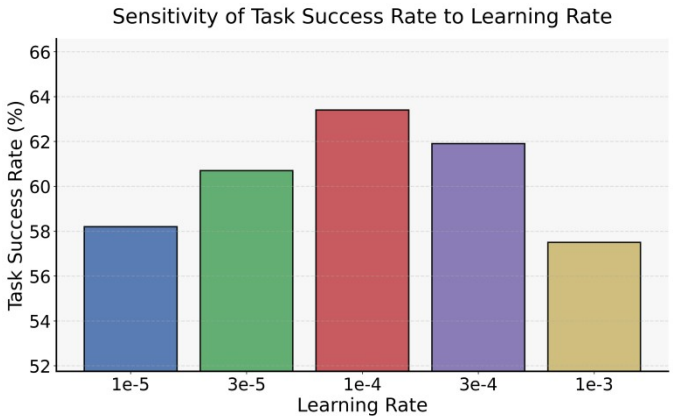


Figure 2. Hyperparameter sensitivity experiment of Task Success Rate under different learning rate settings

The figure illustrates that the choice of learning rate has a clear impact on the stability of task execution in dynamic and uncertain environments. A very small learning rate slows the update process. The agent cannot adapt to ongoing changes in the environment. A very large learning rate causes unstable updates and fluctuating behavior, which prevents the formation of a reliable policy. A moderate learning rate achieves a balance between update speed and stability. This balance allows the agent to remain responsive to environmental variation while maintaining consistent behavior, which results in higher task success rates.

The changes in path efficiency further reveal how learning rate influences long sequence execution. In scenarios with long task chains and frequent environmental shifts, insufficient policy updates cause redundant actions and increasing trajectory deviation. Excessively rapid updates can repeatedly overwrite the execution path and weaken global planning consistency. When the learning rate falls within a suitable range, the agent preserves stable behavior while correcting local errors. This leads to improved path efficiency and reflects the positive effect of adaptive task decomposition and strategy updating on execution quality.

The results for error recovery and action prediction entropy show that learning rate directly affects recoverability and decision stability. Under disturbances or task drift, a low learning rate prevents timely reconstruction of the task structure. A high learning rate makes unnecessary resets more likely, which disrupts the recovery process. With an appropriate learning rate, the model adjusts its policy based on feedback in a more structured and controlled manner. Recovery performance improves. At the same time, lower action entropy indicates more stable and consistent decision-making. These characteristics help the agent maintain reliable performance in complex and changing environments.

The number of task decomposition layers directly determines the granularity of the agent's structure when handling complex tasks, and this granularity affects the accuracy of path planning and the redundancy of action sequences. When the number of decomposition layers is too small, the agent struggles to capture fine-grained operational logic, while too many layers can increase the computational burden and affect path execution efficiency. Therefore, it is necessary to systematically examine the impact of different task decomposition layer numbers on path efficiency to determine the optimal decomposition depth for dynamic and uncertain environments. The experimental results are shown in Figure 3.

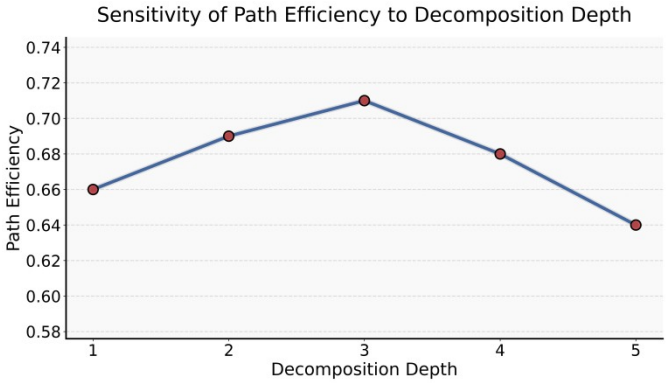


Figure 3. Experiments on the hyperparameter sensitivity of Path Efficiency under different levels of task decomposition.

The curve in the figure shows that the impact of task decomposition depth on path efficiency follows a clear structural pattern. As the decomposition granularity shifts from coarse to more refined levels, path efficiency first increases and then decreases. This indicates that deeper decomposition is not always better, and shallow decomposition is not always preferable. There exists an optimal range that matches the dynamics of the environment and the complexity of the task. When the decomposition depth is low, the structural guidance available to the agent is limited. As a result, the agent struggles to maintain a stable decision sequence under complex environmental changes. This leads to redundant actions and frequent turning in path execution.

Increasing the decomposition depth provides the agent with more complete hierarchical cues. This allows the agent to inherit contextual information more accurately during reasoning and to reduce ineffective actions. Path efficiency reaches its peak within this range. However, the improvement does not continue indefinitely. When the decomposition depth exceeds a reasonable threshold, the task structure becomes overly fragmented. The amount of structural information that must be processed at each step increases. This slows the agent's response to dynamic and uncertain conditions and adds unnecessary reasoning overhead during execution.

At deeper decomposition levels, excessive structural load may reduce the agent's ability to adjust its strategy when abrupt environmental changes occur. For example, when object positions shift or interference events appear, too many structural nodes may delay the triggering of policy updates. This delay causes the planning process to lag behind the environment, which decreases path efficiency. This observation suggests that task decomposition and strategy updating must operate together to maintain efficient behavior in dynamic scenarios.

Overall, the results show that an appropriate decomposition depth is essential for building large model agents with

adaptive capabilities. A moderate number of decomposition layers achieves a balance between structural guidance and reasoning burden. It allows the agent to capture fine-grained task logic while avoiding the reaction slowdown caused by structural overload. These findings further support the importance of the adaptive task decomposition framework proposed in this work. Dynamically adjusting structural granularity based on environmental variation and task demands helps maintain stable and efficient path execution under uncertainty.

This paper also presents the impact of different levels of perceptual noise on the experimental results, which are shown in Figure 4.

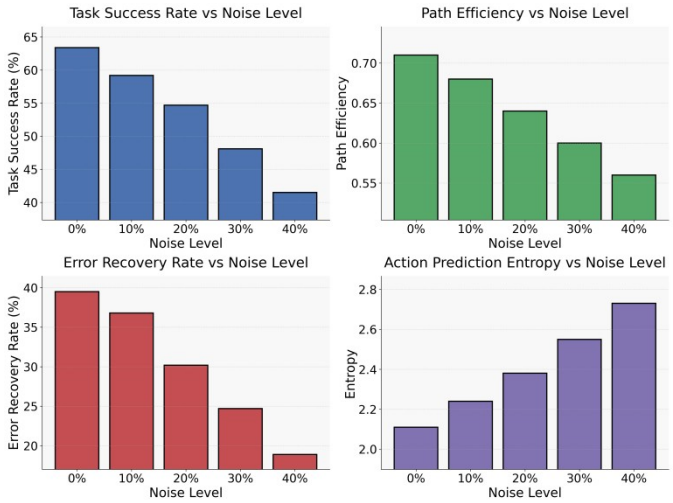


Figure 4. The impact of different levels of perceptual noise on experimental results.

The four subplots together show that increasing perceptual noise significantly weakens task execution quality in dynamic and uncertain environments. As noise levels rise, state recognition, scene understanding, and action feedback all become disturbed. These disruptions lead to biased environmental representations. Since the perception module provides the input foundation for task decomposition and policy generation, corrupted input signals affect both structural inference and subsequent decision updates. As a result, task success rate and path efficiency decline steadily.

The changes in path efficiency further reveal how noise damages structured planning. When environmental features received by the agent are inaccurate or distorted, the agent is more likely to deviate from planned routes and produce redundant or repeated actions. Even with a certain level of reasoning capability, the agent cannot maintain consistent planning under misleading perceptual inputs. Therefore, path efficiency decreases as noise increases, which reflects the difficulty of sustaining stable execution strategies under uncertain observations.

The variation in error recovery ability highlights the fragility of strategy updating mechanisms in high noise settings. With higher noise injection, the agent deviates from the correct action path more easily. Its ability to adjust strategies through environmental feedback also weakens. This indicates that under heavy noise, the agent becomes slower at identifying errors during reasoning. Strategy updates fail to trigger in time, and the recovery rate shows a persistent downward trend. This finding underscores the importance of strengthening policy updating modules to enhance robustness.

The rising trend of action prediction entropy reflects increasing decision uncertainty. Higher noise creates greater ambiguity in environmental interpretation. The action distribution becomes more dispersed and unstable. The agent exhibits stronger randomness when selecting actions and struggles to maintain a stable reasoning chain about the task structure. Higher entropy values indicate weaker operational preferences and greater sensitivity to momentary disturbances. Overall, these results suggest that perceptual quality is critical for maintaining accurate task decomposition and effective strategy updating. Improving perceptual robustness is therefore essential for building highly stable large model agents.

5. Conclusion

This work investigates large model-based agents in dynamic and uncertain environments and proposes an adaptive task decomposition and strategy updating method for real and complex scenarios. By integrating environmental modeling, semantic reasoning, task structure generation, and feedback-driven policy evolution, the study builds an agent framework capable of continuously interpreting external changes, adjusting task structures in real time, and stabilizing decision patterns. The proposed method overcomes the limitations of traditional agents that rely on static planning or fixed prompting schemes. It allows the system to maintain high behavioral consistency and task completion performance even when goals shift, disturbances increase, or feedback becomes incomplete.

At the methodological level, this study presents an effective integration of large model reasoning and structured task planning. The agent can build a robust mapping between high-dimensional semantic spaces and low-dimensional interaction spaces. The adaptive task decomposition mechanism allows the agent to adjust its reasoning granularity based on real-time environmental dynamics. The strategy updating module absorbs feedback throughout the execution process and refines behavior patterns. Together, these components form a closed loop and self-evolving decision system. The experimental trends further show that this structured and adaptive design enhances stability and generalization in long-horizon tasks and provides clear advantages when dealing with complex environmental variations.

In terms of application, the proposed method has strong potential in intelligent control, intelligent manufacturing, human agent collaboration, robotics, autonomous driving, and multi-agent decision making. As environments become more complex, conventional systems struggle to complete real-time, high-risk, and highly dynamic tasks with fixed strategies. Agents equipped with adaptive decomposition and evolving strategies can achieve more reliable planning and more

flexible execution in these uncertain settings. The proposed framework can serve as a foundation for future application systems. It improves long-term deployment capability, environmental transferability, and autonomous cooperation. It also provides essential support for advancing agent technologies toward real-world deployment.

Looking forward, the autonomy and adaptability of agents still have broad room for development. Large model reasoning is expected to expand across more modalities and larger open-world datasets. This progress will provide deeper semantic support for complex task structure understanding. As environmental dynamics move from simulated settings to real physical scenarios, future agents will require stronger perceptual robustness, more efficient long-term policy evolution, and better cross-task transfer ability. Collaborative task decomposition, shared knowledge structures, and collective strategy updating across multiple agents will also become important directions. Continued exploration of large model-driven autonomous agent systems will enable agents to operate in more complex, dynamic, and realistic environments and bring new advances to a wide range of intelligent applications.

References

[1] Woo H, Yoo G, Yoo M. Structure learning-based task decomposition for reinforcement learning in non-stationary environments[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022, 36(8): 8657-8665.

[2] Yoo M, Cho S, Woo H. Skills regularized task decomposition for multi-task offline reinforcement learning[J]. Advances in Neural Information Processing Systems, 2022, 35: 37432-37444.

[3] Prasad A, Koller A, Hartmann M, et al. Adapt: As-needed decomposition and planning with language models[C]//Findings of the Association for Computational Linguistics: NAACL 2024. 2024: 4226-4252.

[4] Zhang, Q., Wang, Y., Hua, C., Huang, Y. and Lyu, N., "Knowledge-Augmented Large Language Model Agents for Explainable Financial Decision-Making, " arXiv preprint arXiv:2512.09440, 2025.

[5] Ruiz-Gonzalez U, Andres A, Del Ser J. Large Language Models for Structured Task Decomposition in Reinforcement Learning Problems with Sparse Rewards[J]. Machine Learning and Knowledge Extraction, 2025, 7(4): 126.

[6] Wang, F., Ma, Y., Guan, T., Wang, Y. and Chen, J., "Autonomous Learning Through Self-Driven Exploration and Knowledge Structuring for Open-World Intelligent Agents, " 2026.

[7] Bidochko A, Vyklyuk Y. Thought Management System for long-horizon, goal-driven LLM agents[J]. Journal of Computational Science, 2025: 102740.

[8] Zhou Z, Xiang H, Zheng Z, et al. V2xpnp: Vehicle-to-everything spatio-temporal fusion for multi-agent perception and prediction[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2025: 25399-25409.

[9] Shi H, Ye S, Fang X, et al. Muma-tom: Multi-modal multi-agent theory of mind[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2025, 39(2): 1510-1519.

[10] Ghafarollahi A, Buehler M J. SciAgents: automating scientific discovery through bioinspired multi-agent intelligent graph reasoning[J]. Advanced Materials, 2025, 37(22): 2413523.

[11] Islam M A, Ali M E, Parvez M R. Mapcoder: Multi-agent code generation for competitive problem solving[J]. arXiv preprint arXiv:2405.11403, 2024.