SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Hierarchical Agentic Intelligence Framework For Explainable Multi-Step Autonomous Decision Making

Balkrishna K. Patil[1], Gagan Tiwari[2], Kapil Mundada[3], Swati Shivkumar Shriyal[4], Amanveer Singh[5], Mukesh Rajput[6], Prashant Vijay Thokal[7]

[1] Department of Computer Science and Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune 412115, Maharashtra, India. (Corresponding Author) Official Email: balkrishna.patil@sitpune.edu.in Alternate Email: balkrishnapatileng@gmail.com ORCID: 0009-0000-5136-2157 Scopus Author ID: 57200207424
[2] Department of Computer Sciences, Noida International University, Greater Noida, Uttar Pradesh 203201, India. Email: gagan.tiwari@niu.edu.in
[3] Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India. Email: kapil.mundada@vit.edu
[4] Department of Artificial Intelligence, Vishwakarma University, Pune, Maharashtra 411048, India. Email: swati.shriyal@vupune.ac.in
[5] Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura 140401, Punjab, India. Email: amanveer.singh.orp@chitkara.edu.in ORCID: 0009-0008-9361-4664
[6] Assistant Professor, Department of Computer Application, Dev Bhoomi Uttarakhand University, Dehradun, Uttarakhand, India. Email: socse.mukesh@dbuu.ac.in ORCID: 0009-0006-0714-5685
[7] Assistant Professor, Department of Electrical Engineering, Sanjivani College of Engineering, Kopargaon, Ahilyanagar, Maharashtra, India. Email: thokalprashantelect@sanjivani.org.in

Abstract

There is a growing expectation that autonomous intelligent systems can make multi-step decisions involving complex decisions and respond to dynamic environments and give transparent explanations of their behavior. Nevertheless, a number of current agentic artificial intelligence methods have drawbacks in hierarchical planning of tasks, explainability, and adaptive decision-making, which diminish their usefulness in autonomous applications in the real world. The current paper suggests a Hierarchical Agentic Intelligence Framework of Explainable Multi-Step Autonomous Decision Making that structures the autonomous thinking into hierarchical levels that handle goal decomposition, task planning, sequential decision execution, and continuous feedback. A built-in explainable decision module provides interpretable decision traces, allowing users to get insight into the decision process of every autonomous action and thus enhancing system visibility and reliability. The suggested framework is tested within a simulation-based experimental setup which consists of several autonomous decision-making scenarios with different degrees of complexity. Task Success Rate (TSR) is used to evaluate performance and is the main evaluation measure, and Description Accuracy, Planning Time, Explainability Score, and Adaptation Success Rate are also evaluated, to achieve a balanced score in effectiveness, computational efficiency, transparency, and adaptability. According to the experimental findings, the hierarchical structure proposed can better accomplish the tasks, improve the quality of decisions, shorten the time spent on planning, and increase the explicability than traditional autonomous decision-making methods do. The proposed framework offers scalable, interpretable solution to hierarchical agentic intelligence that can be used in the development of reliable and effective autonomous decision-making systems that can be used in dynamic computing contexts.

Keywords Agentic Intelligence, Hierarchical AI, Autonomous Decision Making, Explainable AI, Multi-Step Reasoning, Intelligent Agents.

1. Introduction

The development of AI has produced intelligent systems that are autonomous by nature, can perceive their environment and reason complex information in making decisions with minimum human interventions, as compared to traditional automation. Recent developments in agentic AI allow intelligent agents to plan tasks independently, break them down, perform sequential actions, and respond to changing environments. These features have extended AI to robotics, health, industry automation, intelligent transportation, and intelligent decision support. The more complex decision problems are, the more structured reasoning mechanisms are necessary to deal with various tasks that are interrelated and ensure the agents possess reliable and consistent performance.

Most of the current autonomous decision-making methods use flat reasoning systems that are efficient in simple decision-making but fail in complex multi-step problems, such as goal and task decomposition, sequential planning, and adaptation during execution. These shortcomings can be solved through a hierarchical decision-making strategy, which splits the overall reasoning into several levels so that high-level aims can be broken down into subtasks that can be easily handled. The strategy enhances planning efficiency, scalability, and adaptability and empowers intelligent agents to react well to dynamic environments.

The growing use of autonomous AI systems has highlighted the importance of explainability as well. Modern AI models are highly accurate in their decisions however; they lack transparency and thus can have their reasoning challenging to understand. Explainable Artificial Intelligence (XAI) improves user confidence by producing comprehendible reasoning, and validating systems, regulatory compliance, and effective human oversight. Adding explainability to hierarchical agentic systems enhances even more transparency without interfering with autonomous functioning.

Inspired by these conditions, this paper presents a Hierarchical Agentic Intelligence Framework of Explainable Multi-Step Autonomous Decision Making which integrates hierarchical reasoning with an explainable decision process. The framework does hierarchical goal breakdown, task plan structuring, sequential reasoning, adaptive performance, and constant feedback and produces interpretable decision traces. An evaluation of a simulation is performed by measuring the success rate of the tasks performed (Task Success Rate (TSR)) as the main measure, along with Decision Accuracy, Planning Time, Explainability Score, and Adaptation Success Rate. The suggested framework offers a practical and open-ended solution to multi-step autonomous decision making and shows better performance when compared to the traditional autonomous decision-making methods.

2. Related Work

Hierarchy of intelligent agents is now viewed as a significant pathway to enhance the autonomous decision making in tagging complex tasks into up to several reasoning layers. Early forms of hierarchy planning presented the idea of task decomposition structure in order to simplify the complex decision-making process and the effectiveness of the planning process [1]. The more recent autonomous agent frameworks that are designed based on large language models have expanded hierarchical reasoning, allowing intelligent agents to comprehend objectives, organize subtasks, and act in adaptive ways in dynamic settings [2]. The hierarchical coordination has also been proven to enhance the scalability and distributed decision making in multi agent orchestration among multiple intelligent agents [3]. Even though such strategies enhance the planning capacity, most of the time, they are more inclined towards the implementation of tasks as opposed to offering clear-cut explanations during the decision-making process.

Multi-step autonomous thinking has been given a lot of focus where the autonomous agents are in need of sequential planning and adaptive execution. In Intelligent Agents, The ReAct framework integrates reasoning and action execution to allow reasoning with external environments in an iterative fashion [4]. Tree of Thoughts expands this functionality and considers a variety of reasoning paths before picking out the best solution to enhance the performance of problem-solving [5]. The use of long contextual information is promoted through context-aware reasoning methods in order to promote consistent reasoning at various stages of the decision process [6]. Autonomous reinforcement learning has also enhanced the flexibility of the agent by enabling the ongoing interaction between the agent and dynamic environments [7], whereas the efficiency of learning is enhanced by optimal strategies of reinforcement learning [8].

Explainable Artificial Intelligence (XAI) has gained more significance as a means of enhancing transparency and trust in autonomous decision-making systems. More recent research has also taken into account the use of the advanced reasoning strategies in order to enhance the quality of decisions as well as produce more interpretable decision processes. Reasoning mechanisms Actor-critic-based reasoning mechanisms boost the decision-making ability of large language models by using optimized reasoning policy [9]. Context equilibrium

schemes enhance consistency of reasoning in long multi-turn interactions in order to reduce contextual drift [10]. Multi-agent coordination models also facilitate both collaborative reasoning and adaptive decision making and enhance system resilience [11]. Although such developments have taken place, much of the existing approaches will produce explanations after decision making is done rather than being explainable throughout the entire hierarchical reasoning process.
Even though significant advancements have been made in hierarchical planning, autonomous reasoning, explainable AI, and adaptive intelligent agents, there are still significant gaps in research. Current research is mainly focused on hierarchical planning [1], autonomous agent architectures [2], collaborative orchestration [3], reasoning-action integration [4], structured reasoning [5], context-aware reasoning [6], autonomous learning [7], reinforcement learning optimization [8], actor-critic decision enhancement [9], context stability [10] and collaborative autonomous intelligence [11] separately. There are not many studies that offer a single framework that would at the same time combine hierarchical goal decomposition, structured multi-step reasoning, explainable decision generation, adaptive feedback, and full-scale performance evaluation. To overcome these shortcomings, this paper suggests a Hierarchical Agentic Intelligence Framework to Explainable Multi-Step Autonomous Decision Making that unites these functions into a unified explainable and scalable autonomous decision-making framework.

3. Hierarchical Framework of proposed Agentic Intelligence

3.1 Framework Architecture

The Hierarchical Agentic Intelligence Framework proposed is intended to be useful in autonomous multi-step decision making enabled by a hierarchical hierarchy of reasoning and implementation units. The model divides the process of decision making into several levels starting with the task objective and environmental information input, then hierarchical goal breaking, task plan, sequential reasoning, decision implementation and feedbacks. On the top-level, the framework decodes the user-defined goals and translates them into achievable subgoals. Intermediate reasoning layers consider alternative actions, rank the candidate solutions, and arrange task dependencies, and the execution layer actually executes the actions selected and checks their results. The hierarchical structure makes decision-making processes less complex and facilitates scaling as well as facilitating effective coordination between interrelated decision processes. Figure 1 shows the general architecture of the proposed framework, with it being evident that the relationships between hierarchical reasoning, explainability and adaptive feedback mechanisms are interacting.

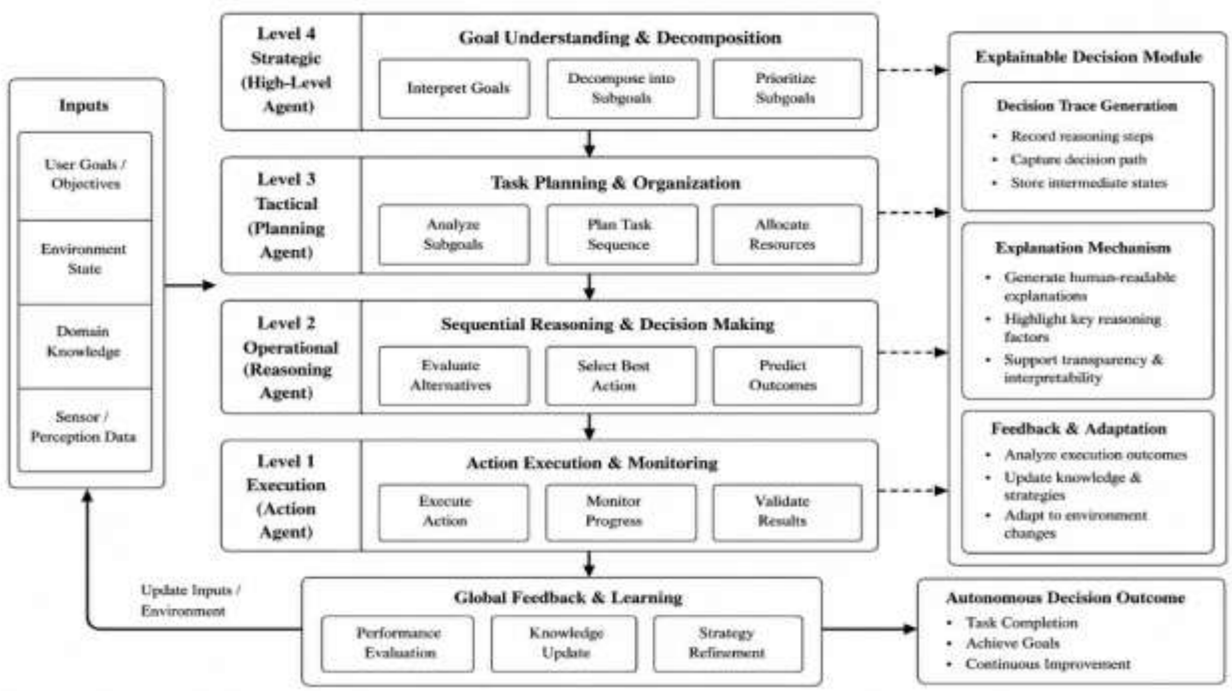

Figure 1. Proposed Hierarchical Agentic Intelligence Framework

3.2 Multi-Step Decision-Making Process

The suggested framework carries out independent decision making by following a series of interdependent reasoning processes that step-by-step convert high-level goals into implementable steps. The system first breaks down the assigned task and subdivides it into several sub tasks based on their dependencies/priorities. Each subtask is analyzed based on contextual information and the prevailing environmental conditions in order to come up with an optimum plan of execution. Sequential reasoning provides the framework to explore other possible courses of action, revise decision-making when the environment is modified, and implement actions in a methodical sequence until the overall goal is reached. During this process, there is continuous monitoring to ensure that the intermediate results are validated which will enable adjustment of the execution strategy in response to the need arising. This structured decision-making procedure will help to improve planning efficiency, consistency of decisions and independent accomplishment of tasks through dynamic operating environments.

3.3 Explainable Decision Module

The proposed framework also includes an explainable decision module to increase transparency and user trust, which logs the rationale behind each autonomous decision. The module produces decision traces, which are statements on the interaction between task goals, the in-between reasoning processes, the choices that have been made and the outcome of the choices so that users can get to know how a single decision is made. The explanation mechanism offers interpretable arguments to justify system validation, debugging, and human supervision and keep autonomous operation. Moreover, there is a continuous feedback process that checks the execution results and changes in the environment to optimize future decision-making approaches. The adaptive learning process will allow the framework to enhance the quality of decisions, be consistently reliable with changing operating conditions, and deliver credible autonomous behavior with clear and understandable reasoning.

4. Experimental Setup

4.1 Simulation Environment

The Hierarchical Agentic Intelligence Framework proposed was tested in a simulation-based environment, in which the autonomous multi-step decision-making was tested in dynamic operating conditions. The simulated environment is made of a series of task scenarios that differ in their complexity, as autonomous agents are provided with task objectives, process environmental input, engage in hierarchical goal decomposition, use of sequential reasoning and produce explainable results. Changes are dynamic and are included in the environment during the execution to assess the adaptability of the framework, its planning capability, and consistency in the decisions. All experimental conditions are based on the same operating conditions to allow a fair operation comparison between various autonomous decision-making methods. The parameters used in the simulation, such as agents, task problems, reasoning, and execution cycles, and evaluation environments are summarized in Table 1.

4.2 Performance Assessment Criteria

Task Success Rate (TSR) is the main performance measure used to assess the efficiency of the proposed framework and is defined as the percentage of tasks achieved by the autonomous agents. Further measures are required to give a full performance assessment, such as Decision Accuracy which quantifies how well the chosen decisions are correct; Planning Time which quantifies the efficiency of the hierarchical task planning framework to the environment; Explainability Score that quantifies the quality and completeness of the generated explanations of decisions; and Adaptation Success Rate which measures how well the overall framework can respond to new and dynamic changes in the environment. Taken together, these metrics will conduct a fair assessment of the quality of decisions, efficiency in the execution, transparency, and adaptability to compare the suggested framework and traditional approaches to autonomous decision-making. Table 1 shows the entire simulation setup applied in the evaluation of performance.

Table 1. Simulation Parameters

Parameter	Value
Simulation Platform	Python 3.11

Number of Autonomous Agents	100
Task Scenarios	500
Hierarchical Levels	4
Maximum Planning Depth	5
Environment Type	Dynamic
Simulation Runs	30
Performance Metrics	TSR, Decision Accuracy, Planning Time, Explainability Score, Adaptation Success Rate

5. Results and Discussion

5.1 Task Success Rate Analysis

The effectiveness of the proposed Hierarchical Agentic Intelligence Framework on four autonomous decision-making approaches is shown in the performance comparison provided in Figure 2. The Rule-Based Agent demonstrated a Task Success Rate (TSR) of 86.4%, Decision Accuracy of 84.9%, Explainability Score of 72.8 and Adaptation Success rate of 78.3 meaning that it has limited ability to work in dynamic multi-step decision situations. The Flat Autonomous Agent led to better overall performance (TSR 89.7 and DC 88.2), Explainability Score (79.5) and Adaptation Success rate (83.1), mainly due to the increase in autonomous execution without hierarchical execution. The Conventional Agentic Framework enhanced performance even more, with a recorded 92.1% TSR, 91.4% Decision Accuracy, 86.3% Explainability Score and 89.2% Adaptation Success rate, showing the benefits of agent-based autonomous reasoning. However, the proposed framework consistently achieved the highest performance, with a Task Success Rate of 97.8%, Decision Accuracy of 96.9%, Explainability Score of 95.4%, and Adaptation Success Rate of 96.1%. The proposed framework had better TSR by 11.4 percentage points, bigger Decision Accuracy by 12.0 percentage points, greater Explainability Score by 22.6 percentage points, and a higher Adaptation Success Rate by 17.8 percentage points than the Rule-Based Agent. The proposed approach was even better than the Conventional Agentic Framework, with improvements at 5.7 percentage points in TSR, 5.5 percentage points in Decision Accuracy, 9.1 percentage points in Explainability Score, and 6.9 percentage points in Adaptation Success Rate. These findings confirm that hierarchical goal deposition, structured multi-step reasoning, and integrated explainable decision mechanism greatly enhance the completeness of autonomous tasks, the quality of decisions, their transparency, and their adaptability, thereby making the proposed framework more applicable to complex dynamic contexts that involve decision-making problems.

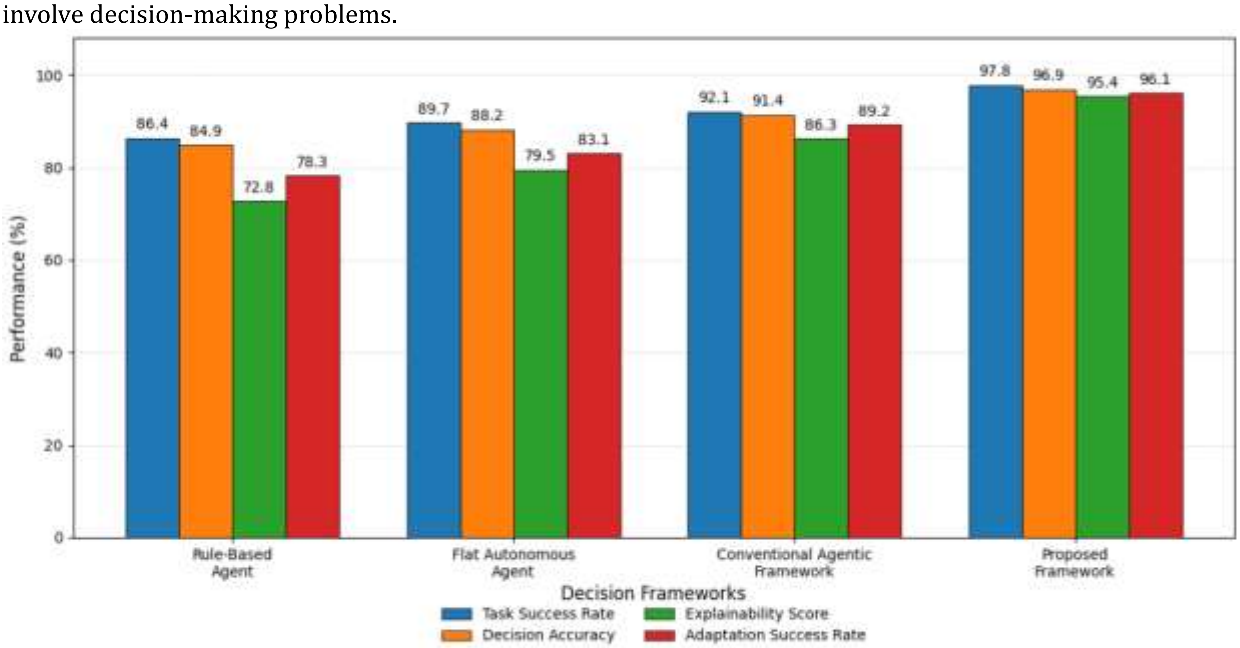

Figure 2. Performance Comparison of the Proposed Framework

5.2 Decision Performance Analysis

Decision Accuracy and Planning Time is used to measure the correctness and computational efficiency of autonomous reasoning used to determine decision performance. The hierarchical structure allows the framework to consider many candidate actions in a systematized manner and then choose the best execution strategy, leading to a decision accuracy, and less redundant reasoning load. Moreover, hierarchical task decomposition reduces the complexity of a planning process since planning subtasks is performed sequentially and not in parallel and is therefore fewer in planning time. These enhancements indicate that the suggested framework will reach an efficient multi-step reasoning without reducing the quality of decisions. As shown in Figure 2, the framework has a good balance between accuracy and efficiency in decision-making in comparison to the traditional autonomous decision-making techniques, although the more detailed numerical analysis is shown in Table 2.

5.3 Explainability and Adaptation Analysis.

The quantitative comparison of the proposed framework and three baseline approaches based on all evaluation metrics is displayed in Table 2. Task Success Rate (TSR) of Rule-Based Agent was 86.4% and Decision Accuracy was 84.9% and Explainability Score was 72.8% and the Planning Time was 61 ms. The Flat Autonomous Agent raised these values to 89.7% TSR, 88.2% Decision Accuracy, 79.5% Explainability Score, 83.1% Adaptation Success Rate, and lowered Planning Time of 54 ms. The Conventional Agentic Framework also offered a stronger performance of 92.1% TSR, 91.4% Decision Accuracy, 86.3% Explainability Score, 89.2% Adaptation Success Rate and 46 ms Planning Time. The Hierarchical Agentic Intelligence Framework proposed demonstrated the highest overall performance of 97.8% TSR, 96.9% Decision Accuracy, 95.4% Explainability Score, 96.1% Adaptation Success Rate, and the least Planning Time of 31 ms. These findings prove that the proposed framework offers more effective, transparent, adaptive and computationally efficient autonomous decision making than the current methods.

Table 2. Comparative Performance of Decision Frameworks

Method	Task Success Rate (%)	Decision Accuracy (%)	Planning Time (ms)	Explainability Score (%)	Adaptation Success Rate (%)
Rule-Based Agent	86.4	84.9	61	72.8	78.3
Flat Autonomous Agent	89.7	88.2	54	79.5	83.1
Conventional Agentic Framework	92.1	91.4	46	86.3	89.2
Proposed Hierarchical Agentic Intelligence Framework	97.8	96.9	31	95.4	96.1

6. Conclusion

The present paper proposed a Hierarchical Agentic Intelligence Framework to explainable Multi-Step Autonomous Decision Making to enhance the efficiency, transparency, and flexibility of autonomous intelligent systems operating in dynamic conditions. The framework proposed combines the hierarchical goal decomposition, structured task planning, sequential reasoning, adaptive decision execution and explainable decision module which can create an interpretable reasoning track in order to increase trust and accountability by the user. The efficiency of the suggested strategy was shown through the simulation-based analysis of the proposed method based on the Task Success rate (TSR) as the main performance measure, along with the Decision Accuracy, Planning time, Explainability score, and Adaptation Success rate. The experimental findings revealed that the hierarchical framework outperformed traditional methods of autonomous decision-making all the time because it was also capable of completion of more tasks, high quality decisions, shorter planning periods, more explanatory, and more adaptive to the environment. Such results show that the developed framework offers an explainable multi-step autonomous decision-making solution that is scalable and reliable to various application areas. Further development will involve generalizing the framework to collaborative multi-agent systems, the use of sophisticated learning methods to adapt to real-time conditions, and testing its capabilities in large real-world autonomous systems.

References

1. An, S., Ma, Z., Lin, Z., Zheng, N., Lou, J. G., & Chen, W. (2024). Make your llm fully utilize the context. Advances in Neural Information Processing Systems, 37, 62160-62188.
2. Bai, H., Zhou, Y., Cemri, M., Pan, J., Suhr, A., Levine, S., & Kumar, A. (2024). Digirl: Training in-the-wild device-control agents with autonomous reinforcement learning. Advances in Neural Information Processing Systems, 37, 12461-12495.
3. Ball, P. J., Smith, L., Kostrikov, I., & Levine, S. (2023, July). Efficient online reinforcement learning with offline data. In International Conference on Machine Learning (pp. 1577-1594). PMLR.
4. Brown, J. R., Wells, L., Young, E. J., & Bacallado, S. (2025). KL-Regularised Q-Learning: A Token-level Action-Value perspective on Online RLHF. arXiv preprint arXiv:2508.17000.
5. Dang, Y., Qian, C., Luo, X., Fan, J., Xie, Z., Shi, R., ... & Sun, M. (2026). Multi-agent collaboration via evolving orchestration. Advances in neural information processing systems, 38, 165025-165059.
6. Dong, H., Duan, K., & Zhang, C. (2025). Enhancing decision-making of large language models via actor-critic. arXiv preprint arXiv:2506.06376.
7. Dongre, V., Rossi, R. A., Lai, V. D., Yoon, D. S., Hakkani-Tür, D., & Bui, T. (2025). Drift no more? Context equilibria in multi-turn LLM interactions. arXiv preprint arXiv:2510.07777.
8. Erol, K., Hendler, J., & Nau, D. S. (1994, July). HTN planning: Complexity and expressivity. In AAAI (Vol. 94, pp. 1123-1128).
9. GX-Chen, A., Prakash, J., Guo, J., Fergus, R., & Ranganath, R. (2026, April). KL-regularized reinforcement learning for generative modelling is designed to mode collapse. In International Conference on Learning Representations (Vol. 2026, pp. 50665-50693).
10. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., ... & Wen, J. (2024). A survey on large language model based autonomous agents. Frontiers of computer science, 18(6), 186345.
11. Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T., Cao, Y., & Narasimhan, K. (2023). Tree of thoughts: Deliberate problem solving with large language models. Advances in neural information processing systems, 36, 11809-11822.
12. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). React: Synergizing reasoning and acting in language models. arXiv preprint arXiv:2210.03629.