Journal of Computer Technology and Software

ISSN: 2998-2383

Vol. 4, No. 12, 2025

Conceptualizing LLM-Based Cognitive Chains for Intelligent Command-and-Control Applications

Robert Kosowski
University of North Texas, Denton, USA
rak0456@unt.edu

Abstract: This study responds to the practical need for applying large language models (LLMs) to military command-and-control (C2) tasks and proposes a new reference model for constructing and integrating cognitive chains. The aim is to provide methodological guidance and developmental pathways for next-generation intelligent C2 information systems. First, based on the current development status and application practices of LLM technologies, the paper analyzes the major challenges that arise in military scenarios and the corresponding approaches for addressing them. Then, it introduces a technical roadmap for enabling intelligent command entities through LLMs and derives a scenario-driven mechanism that supports multi-agent organizational operations and capability formation. Finally, it examines the concept and technical foundation of cognitive chains in depth and presents an application vision for a "four-chain-fusion" LLM-empowered C2 system.

Keywords: Large Language Models (LLMs); Command-and-Control Systems; Cognitive Chains; Intelligent Command Entities

1. Introduction to the Problem

In November 2022, the release of the generative pre-trained transformer model ChatGPT triggered a rapid global surge of interest within just a few months. This phenomenon ignited an intense "large-model race", during which foundation-model products experienced explosive growth, with model types continuously expanding and usage methods evolving at an accelerating pace. The overall developmental stages of large models are illustrated in Figure 1. In the civilian domain, large-model applications have already penetrated multiple sectors, including healthcare, financial services, and intelligent transportation.

For example, an intelligent monitoring and adaptive analysis system for assessing the psychological states of specific populations adopts a technical architecture that integrates "large models + wearable devices + multimodal intelligent fusion", enabling the detection of anomalous individual or group behaviors and the early prediction of potential risks [1]. In the military domain, Palantir has released concept videos of its artificial intelligence platform (AIP), demonstrating an OODA process in which conversational AI collaborates with human operators to perform observation, judgment, decision-making, and action tasks during land-combat scenarios [2]. These demonstrations reveal a new paradigm for human-AI cooperation, a deployment model for integrating large models with simulation, data governance, and trusted-service mechanisms, and a set of emerging application patterns that have drawn significant attention from domestic and international defense-technology communities. The technical architecture of Palantir AIP is shown in Figure 2

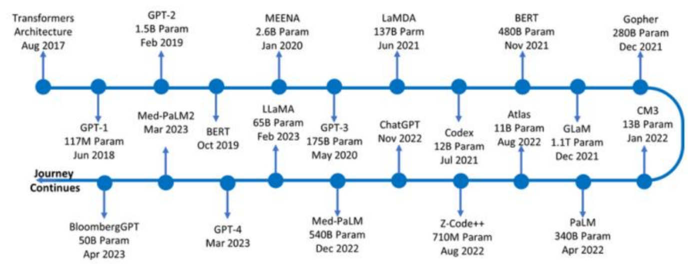

Figure 1. Schematic Diagram of Large-Model Development Stages

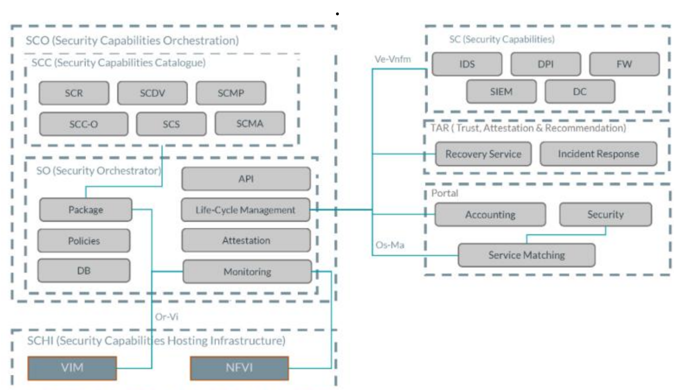

Figure 2. Schematic Diagram of the Palantir AIP Technical Architecture

Drawing on the author's and the team's practical experience over the past two years in applying large language models (LLMs) within various domains, it has become evident that, compared with traditional discriminative AI models, generative AI and foundation models demonstrate remarkable few-shot and zero-shot learning capabilities. These models exhibit broad applicability across diverse contexts and scenarios, including semantic understanding and intent recognition in human-machine interaction, multi-turn dialogue strategy generation, intelligence retrieval and event-forecasting in information-processing domains, document drafting, task-factor decomposition in mission planning, and process-level situation monitoring in integrated data-management systems. Combined with advancements such as chain-of-thought reasoning and human-level reinforcement learning, the rapid progress of large-model technologies not only significantly improves operational efficiency but also offers new cognitive frameworks for higher-level decision-making and problem-solving. A comparison between discriminative AI and generative AI is illustrated conceptually in Figure 3.

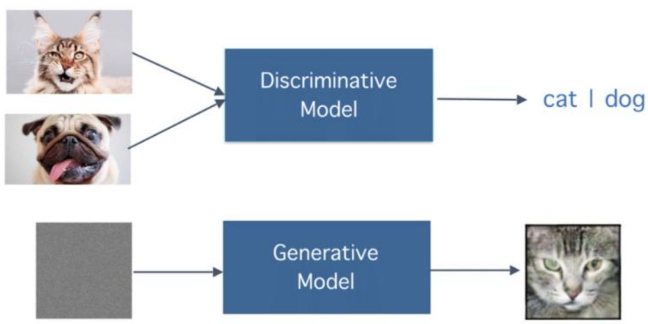

Figure 3. Schematic Comparison of Discriminative AI and Generative AI Characteristics

Large models possess certain general artificial intelligence characteristics, such as extensive and deeply connected knowledge, strong reasoning abilities, and enhanced learning and decision-making competence. However, their performance in specialized domains still requires improvement. Since large models rely on massive datasets and high-dimensional neural network architectures, their essence can be viewed as an engineered transformation of AI from "qualitative change to quantitative change". Although theoretically scalable, practical engineering implementation remains challenging due to stringent requirements on computation and data. Computational resources form the foundation for building and deploying large models, yet most organizations cannot afford such costs. Furthermore, high-quality, large-scale, and diverse datasets are essential, often requiring collection and preprocessing at a societal scale.

Therefore, a key question is how to maintain clarity when navigating this wave of new technologies and how to more effectively align large-model development with real-world needs. Understanding how to guide large models toward meaningful and practical integration within specific domains and scenario-oriented applications is an issue that merits thoughtful consideration.

2. Challenges in Applying Large Models to Intelligent Command-and-Control

2.1 Related Work

The integration of large language models (LLMs) into intelligent systems has drawn considerable attention in recent research. Foundational works have explored the self-organization and tool-use capabilities of LLMs [3], as well as bootstrapped reasoning processes that serve as precursors to cognitive-chain operations [4]. These advances lay essential groundwork for enabling interpretable reasoning and task decomposition in command applications.

Recent developments in generative multi-agent environments have demonstrated the potential for interactive, role-based collaboration among intelligent agents [5][6]. These systems simulate human-like decision-making patterns and enable scalable coordination across complex scenarios, providing inspiration for constructing C2-oriented agent ecosystems. Reinforcement learning frameworks for resource orchestration [7] and multi-agent control under uncertainty [8] further support adaptive autonomy and role-specialized intelligence within large-scale deployments.

Ensuring the robustness of LLM outputs under uncertainty has become a focal point for enhancing trustworthiness. Methods for uncertainty quantification and risk-aware summarization [9], privacy-preserving instruction adaptation [10], and federated fine-tuning with semantic alignment [11] contribute significantly to dependable LLM deployment in sensitive military contexts. These works reflect growing attention to the operational reliability and policy compliance of language models in real-world environments.

Supporting tools such as LLM-centric retrieval-augmented generation [12], sparse retrieval mechanisms for fact verification [13], and transformer-based modeling for clinical or financial domains [14][15] have further expanded the frontier of LLM applicability. These approaches highlight the capacity of LLMs to integrate structured and unstructured data for decision support and situational analysis, core functions in command and control systems.

Innovations in intelligent agent collaboration within microservice systems [16] and simulated agent societies [17] reveal technical architectures conducive to multi-role coordination and knowledge sharing. Vision-based editing using diffusion models [18], language grounding in robotics [19], and adaptive human–computer interaction through reinforcement learning [20] illustrate the multi-modal and interaction-rich pathways that support the evolution of large-model-based command agents.

From a strategic perspective, transformer models have also shown value in domain-specific applications such as risk monitoring in finance [21] and dynamic portfolio optimization under multi-agent reinforcement learning [22], underscoring the versatility of LLMs in operational decision-making under constraints.

Lastly, the reasoning capabilities of LLMs in zero-shot scenarios [23] further emphasize their potential in dynamic and unpredictable military environments, supporting the design of flexible, reasoning-capable agents that can perform with limited supervision.

2.2 Key Challenges and Limitations

As described earlier, large models can significantly enhance efficiency in tasks such as human-machine interaction, information retrieval, text generation, and workflow automation, and they often exhibit characteristics of partial artificial general intelligence in specific scenarios. However, large models are not omnipotent and still show limitations, including hallucination problems, weak interpretability of generated content, and deficiencies in scientific computing and logical reasoning. Among these, hallucination is a major bottleneck hindering the widespread deployment of large models at this stage.

The causes of hallucination in large models mainly arise from three factors:

Limitations in AI model design: Large models are fundamentally based on statistical patterns, making them unable to reliably distinguish facts from fabrications, nor strictly adhere to logical and rule-based reasoning.

Knowledge limitations in training data: Inaccurate corpora, missing edge-case information, and hidden biases may all lead to erroneous model outputs.

User-induced bias in prompts: If a prompt contains misleading or fabricated information, the model may shift its output toward this false context due to probabilistic tendencies. An example illustrating ChatGPT's knowledge limitations and prompt-induced bias is shown in Figure 4.

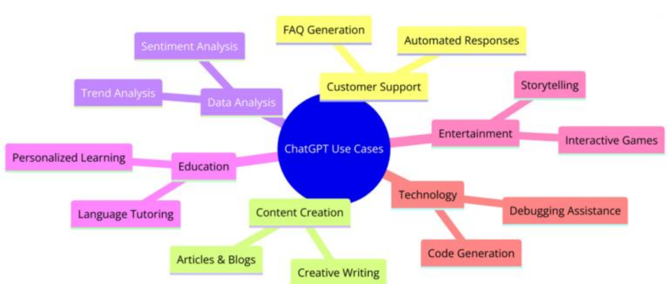

Figure 4. Example of ChatGPT's Knowledge Limitation and User Prompting

Currently, academic and industrial communities have proposed several technical solutions to address these weaknesses, such as enhanced information retrieval, domain-specific knowledge augmentation, and numerical reasoning enhancement. These approaches enable large models to produce more accurate and reliable outputs in constrained scenarios. A conceptual illustration of controlled output mechanisms in scenario-specific applications is provided in Figure 5. Nevertheless, the application of large models in military command-and-control still involves many issues requiring careful consideration.

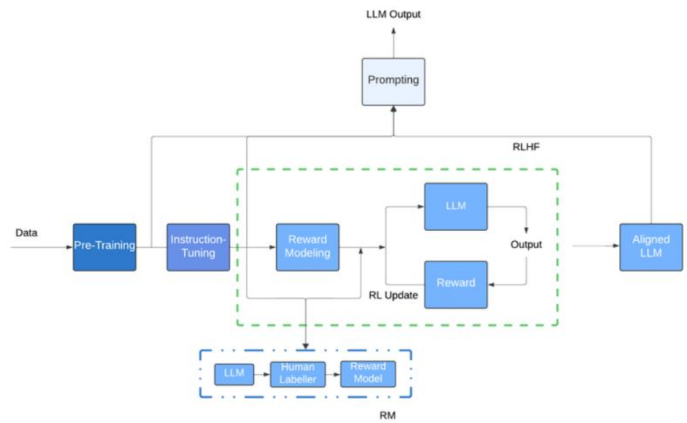

Figure 5. Schematic Diagram of Controllable Output Mechanisms for Large-Model Scenario Applications

Professionalism challenges: Modern warfare demands high precision and dynamic adaptation. Military intelligence must address complex, real-time battlefield tasks and rapidly changing operational needs. Most military scenarios require strong task-specific specialization rather than generalization or simplification.

Confrontation challenges: Military operations involve constant information confrontation and countermeasures. Data obfuscation, camouflage, deception, and adversarial interference are widespread. Large variations in data quality, low signal-to-noise ratios, and difficulties in distinguishing authenticity pose far greater challenges to military large-model applications compared with civilian domains.

Complexity challenges: Combat operations involve multi-domain integration, diverse mission elements, and highly variable battlefield environments. Military data often feature non-structural characteristics and complex interdependencies. Single-mode or single-state large models cannot meet such needs, requiring multi-state and multi-modal combined applications.

Security challenges: Military operations concern life-or-death decision-making and require highly reliable large-model systems. At the same time, models deployed in warfare face adversary reconnaissance, deception, and interference. It is necessary to prevent hostile forces from exploiting vulnerabilities in academic research or system implementation to mislead or manipulate model outputs.

3. Intelligence Development Requirements for Command-and-Control Systems

3.1 "+Intelligence" and "Intelligence+"

Observing the evolution of command-and-control (C2) information systems over the past decade, the pathway toward intelligent C2 capabilities has generally followed a "two-step" development pattern. The first step is the "+Intelligence" phase, which focuses on advancing informatization, service-oriented architectures, and network foundations while integrating AI technologies and embedded intelligent computing models. This phase provides new methods, tools, and mechanisms for traditional reconnaissance, early warning, C2 operations, information confrontation, and battlefield data applications. The second step is the "Intelligence+" phase, in which artificial intelligence takes the leading role, enabling proactive guidance, dynamic coordination, and organizational-level decision augmentation. Ultimately, this leads to fully integrated, streamlined, and system-level intelligent C2 networks.

The transition from "+Intelligence" to "Intelligence+" manifests in two major dimensions.

From intelligent components to holistic intelligent systems. Under the "+Intelligence" model, the C2 system embeds intelligent modules or small models, enabling users to obtain decision support by selecting specific software functions. This is a "point-based" approach for enhancing local C2 intelligence. In contrast, "Intelligence+" places AI as the dominant driver, autonomously guiding, coordinating, and orchestrating functional modules and auxiliary tools across the entire C2 system, thereby achieving full-process and system-wide intelligent transformation.

From typical tasks to all-domain missions. Small embedded models under the "+Intelligence" paradigm tend to have limited generalization, making it difficult to adapt to diverse operational demands. Continuous model iteration is required to fit new application contexts. In contrast, "Intelligence+" leverages unified information architectures and highly generalized intelligent model clusters, enabling adaptive support across full-spectrum battlefield operations.

Advances in large-model technologies have created feasible technological pathways for transitioning from "+Intelligence" to "Intelligence+", positioning large models as a driving force behind intelligent C2 development. By leveraging large models' powerful human-machine interaction, intent inference, and cognitive generalization capabilities, C2 information systems can achieve significant improvements in flexibility, usability, and robustness. Presently, next-generation C2 information systems are under active exploration and research, with an expected transition from typical "+Intelligence" configurations to architectures where large models serve as the core orchestrating element. This enables multi-level intelligent model clusters to guide operational applications across diverse scenarios and mission types.

Small models tailored for specific conditions can autonomously execute well-structured tasks, such as software-supported planning, rapid matching of operational templates derived from previous campaign cases, or intelligent troop formation planning that automatically recommends organizational schemes based on current requirements. Such systems can also integrate expert feedback to refine planning results and generate rapid planning outputs.

3.2 Large Models and Small Models

Small models and traditional AI algorithms exhibit strong applicability and specialization in fields such as image recognition, target detection, path planning, and intelligent control. These models are technically mature within specific domains and task scenarios, and they hold clear advantages over large models in terms of domain expertise and operational stability. At the same time, due to their small size and low computational demand, small models are better suited for resource-constrained battlefield command-information systems where computing power is limited and scalability is difficult. Therefore, small-model-based intelligent decision support will continue to serve as a core component in enhancing intelligent C2 capabilities for a considerable period.

Consequently, multi-state and multi-modal AI model integration provides a feasible pathway for addressing the challenges of deploying large models in C2 applications. This integration enables scenario-driven collaboration between large models and small models. Large-model products such as AutoGPT [11] released in 2023 and multi-agent collaboration frameworks like HuggingGPT demonstrate large models' ability to autonomously coordinate external tools. Inspired by the HuggingGPT paradigm, a high-level planning architecture can be constructed as shown in Figure 6. Large models can decompose complex problems into smaller tasks and determine which tools or small models are best suited to solve each subproblem. In this way, large models become orchestrators capable of coordinating small-model resources, leveraging small-model strengths to compensate for large-model weaknesses, and forming a complementary and efficient intelligent application mechanism.

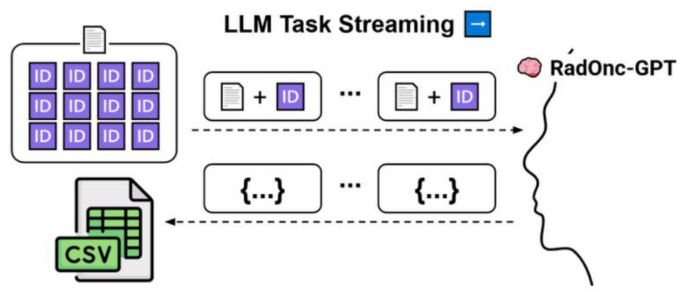

Figure 6. Reference Architecture for Large-Model Planning and Coordination

4. Large-Model-Enabled Intelligent Command Agents

4.1 Intelligent Command Agents

An intelligent agent refers to an entity capable of perceiving its environment and taking actions to achieve specific objectives. An agent may take the form of software, hardware, or an integrated system and is characterized by autonomy, reactivity, proactivity, and interactive capability. By sensing environmental changes, such as through sensor inputs or data streams, an agent uses acquired knowledge and algorithms to make judgments and decisions, then conducts actions that influence the environment or help achieve designated goals. Intelligent agents are widely applied across artificial intelligence domains, including automation systems, robotics, virtual assistants, and game characters. Their core lies in the ability to maintain autonomous operation and continuous evolution, enabling them to better accomplish tasks and adapt to complex environments.

Based on this definition, and considering the application contexts and capability requirements of military command-and-control systems, this paper approaches command agents (C^2-AI Agents) from two perspectives: a broad interpretation and a narrow interpretation.

From a broad perspective, a command agent represents a new paradigm for the design, development, integration, delivery, and iterative evolution of future military C2 information systems. During the design and development phase, unlike traditional software analysis and requirement-design methods, system engineers will not only complete scenario design, workflow design, and input-output modeling but will also place greater emphasis on using natural language to refine user needs with precision. Fine-grained annotation of operational scenarios will accumulate high-quality labeled datasets essential for model training and prompt-engineering refinement.

During the development and integration phase, system input-output modules (I/O, including network communication), data management components, user management modules, and model-service interfaces will become more generalized and standardized. This will enhance reusability and interoperability, enabling tighter integration among different service branches. C2 information systems will further unify interfaces such as machine interaction protocols, communication primitives, and even user authentication. Cross-system interoperability will become smoother, and cloud-based system architectures will be further advanced.

During delivery and iterative evolution, increased openness of system-development interfaces will reduce user-side maintenance complexity. With natural-language-driven system description and large-model-assisted code generation or template-based injection, users will be able to rapidly construct scenario-specific system upgrades. This will improve system adaptability and provide industrial developers with clearer insights into post-fielding maintenance tasks.

From a narrow perspective, a command agent encompasses a series of perceivable, interpretable, and executable military-intelligence models. The conceptual model of command agents is illustrated in Figure 7. This model adopts a large-model-plus-small-model architecture, where model scheduling, algorithmic integration, meta-learning, and transfer learning techniques are combined to dynamically integrate data, knowledge, tools, and models for military C2 tasks. This enables continuous learning, optimization, and quantitative evaluation, driving progressive evolution of C2 systems.

Command agents may include intelligence-processing agents, situational-perception agents, decision-support agents, and action-control agents. These agents are expected to evolve from manual intelligence assessment toward automated inference transformation, from situational understanding toward predictive situational transition, and from decision analysis toward adversarial reasoning and strategy generation. For example, situational-awareness agents may evolve into models capable of predicting transitions based on probabilistic inference, while action-control agents may evolve into semi-autonomous or fully autonomous models capable of executing mission-level command transitions. Ultimately, such advancements will support the transition toward "Intelligence+" command-system capabilities.

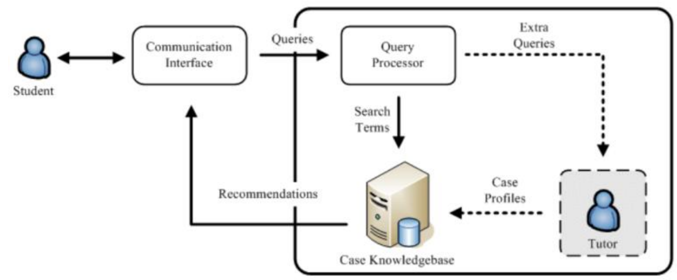

Figure 7. Conceptual Framework of the Intelligent Command Agent

4.2 Vision for Large-Model-Enabled Intelligent Command Agents

According to the four fundamental characteristics of intelligent agents - autonomy, reactivity, pro-activeness, and social ability - the feasibility of enabling command agents with large models can be analyzed as follows.

Autonomy. Autonomy refers to an agent's ability to operate without external intervention and maintain a certain degree of internal and behavioral control. This implies that an agent should not only follow explicit human instructions to accomplish tasks but also demonstrate the ability to independently initiate and execute actions. Large models possess strong conversational comprehension and tool-use capabilities, which can serve as new forms of human-system interaction media. This significantly enhances human-machine collaboration efficiency and reduces human workload.

Large models' abstract knowledge-learning and reasoning capabilities will allow them to move beyond the limitations of purely textual or rule-based knowledge. They will increasingly exhibit adaptive abilities through real-time environmental input, enabling dynamic adjustment and integrated decision-making. Simultaneously, their multimodal perception and predictive abilities, including image, audio, and signal processing, will allow them to extract information from massive operational datasets, such as those generated in training or adversarial simulations. This will trigger the emergence of new intelligence patterns, many of which do not require manual or rule-based encoding before being integrated into command systems.

Reactivity. Reactivity refers to an agent's ability to respond rapidly to environmental changes and external stimuli. This means that an agent must perceive shifts in the surrounding environment and take timely and appropriate actions.

Traditionally, the perceptual abilities of language models were limited to text input, while action models were constrained to text outputs. With multimodal integration, however, the perceptual space of language models expands to include visual and auditory information from the environment. This significantly enhances a large-model-driven agent's ability to interact effectively with the physical world and execute missions.

One of the key challenges for large-model-driven agents is the intermediate step required when converting textual reasoning into non-textual actions. The agent must first produce a textual plan before transforming it into executable operations, which may increase latency. However, this is consistent with human reasoning processes, as human behaviors are also guided by prior deliberation and planning.

Pro-activeness. Pro-activeness means that an agent not only reacts to the environment but also actively takes actions aimed at achieving goals. This characteristic emphasizes an agent's ability to reason, plan, and execute proactive measures to accomplish goals or adapt to changes. Intuitively, future iterations of large models will increasingly embed predictive and anticipatory capabilities that implicitly generate such proactive behaviors. Large models excel in conceptual reasoning and high-level planning and can be prompted with instructions such as "let us think step by step" to activate their reasoning ability, including logical and mathematical inference.

Social ability. Social ability refers to an agent's capability to interact with other agents or humans through various forms of communication, including natural language. Large models demonstrate strong natural-language comprehension and generation, and when combined with structured or protocol-based communication channels, they can interact with other models or humans in an interpretable manner. This forms the basis of agent-level social capability.

Building on this, multi-agent collaboration frameworks such as MetaGPT extend social ability by assigning different roles to intelligent agents and enabling them to collaborate on multi-role tasks. The design philosophy of MetaGPT resembles a "project team" composed of intelligent agents, including product managers, architects, software engineers, and QA testers. Each agent autonomously generates instructions based on its assigned role and executes specific tasks. Through inter-agent communication and data exchange, they cooperate to complete complex missions and ultimately deliver results to the user.

5. New Conceptualization of Deep Cognitive-Chain Applications for Large Models

5.1 Understanding and Interpreting Cognitive Chains

Section 4.1 provided both broad and narrow perspectives on the concept of command agents. From a narrow perspective, a command agent can be viewed as a set of transferable, interpretable, and executable composite military-intelligence agents, essentially a collection of individual agents with specialized capabilities. How to activate various types of agents, construct connections among them, and achieve scenario-driven multi-agent organization and operation requires the integration of the "cognitive chain" concept and associated technologies.

Cognitive chains (CoT) refer to a series of logically connected and sequentially linked reasoning steps or thought processes that form a complete chain of reasoning. This serves as a method for guiding individuals to think about, analyze, and solve problems. Cognitive-chain technology enables complex reasoning tasks to be decomposed into multiple interpretable steps, providing clearer logical pathways and improving overall interpretability.

Based on the reasoning capability inherent in cognitive chains, intelligent agents can autonomously learn and perform relevant inference when prompted, without requiring complex iterative training. When given simultaneous prompts, agents can generate appropriate reasoning steps and correct conclusions, demonstrating an approximation of human-level structured thinking and logical processing.

In the command domain, cognitive chains have long served as a guiding framework for commanders and staff to develop decision-making thought processes during command activities. They have also functioned as a method for supporting C2 task execution and organizational application. However, cognitive-chain methods have historically been expressed only in theoretical form and have not been fully implemented as operational tools through human-computer interaction.

5.2 Conceptualization of Cognitive-Chain Applications

In typical operational scenarios, cognitive-chain applications extend beyond traditional reasoning chains and can incorporate judgment chains, computational chains, and system-level cognitive chains to address complex military tasks. This forms a "four-chain fusion" approach that enhances the intelligent capability of command-information systems.

1) Judgment Cognitive-Chain Applications

Judgment cognitive chains combine one or multiple known judgments to infer new conclusions based on causal relationships, correlation patterns, or other logical links. The key characteristic of this cognitive-chain type is "accuracy". For example, in situational awareness, large vision models can fuse multi-source intelligence data, integrate historical datasets, and analyze indicators of adversarial intent to produce precise early-warning assessments. During operational planning and formulation, judgment chains allow commanders to analyze operational requirements and incorporate relevant data from adversarial analysis, historical cases, doctrinal knowledge, and battlefield examples. These steps enhance the accuracy and robustness of operational-plan generation by providing systematic and logically traceable pathways. The resulting multi-branch or distributed operational plans support flexible maneuver strategies, each validated through stepwise logical reasoning, thereby improving interpretability and confidence.

2) Computational Cognitive-Chain Applications

Computational cognitive chains focus on decomposing difficult or complex problems into a series of solvable steps

[20], typically applying decomposition-based methods to control complex missions. Under uncertainty, planning, learning, and adjustment can be achieved through "precision-oriented" reasoning. In mission-planning scenarios, language-model-based cognitive chains can be used to analyze military contexts, interpret situational elements, and transform complex environments into modular information-processing tasks. The agent can then match each task to corresponding tools and models, enabling organized analysis and solution generation.

For example, when a commander designates a strike mission, a large model can automatically generate a reasoning chain to plan the F2T2EA (find, fix, track, target, engage, assess) process. Through logical sequencing, the model can evaluate reconnaissance inputs, adjust target classification, perform strike-asset allocation, and integrate multiple intelligence modules. This produces a coherent and executable kill chain that supports rapid and high-confidence mission execution.

3) Parallel Cognitive-Chain Applications

Parallel cognitive chains enable intelligent agents to process multiple tasks concurrently, improving task execution speed. Their key characteristic is "speed". In intelligence contexts, multimodal large models can simultaneously process data from synthetic aperture radar (SAR), electro-optical sensors, and video streams to support high-efficiency target detection and identification. This facilitates rapid integration of diverse intelligence sources to identify new threats based on zero-shot and few-shot capabilities. Consequently, identification accuracy increases while source-to-action latency is significantly reduced.

By combining diverse inputs, parallel cognitive chains enable operational forces to detect risks at the earliest stage, assess threats in real time, and autonomously generate multiple alternative courses of action to ensure mission safety.

4) System Cognitive-Chain Applications

System cognitive chains adopt a system-engineering perspective and construct reasoning patterns based on the "1 + 1 > 2" principle. Their defining characteristic is "fusion". In intelligent operations, large models leverage powerful human-machine interaction and rapid contextual understanding to interpret commander intent and provide precise recommendations. These recommendations can then be automatically converted into machine-interpretable instructions, forming unified command expressions that support interoperability among sensors, platforms, and equipment.

System cognitive chains enable seamless coordination across domains and ensure that various operational components function as an integrated whole. For example, multimodal large models can fuse real-time sensor data with historical knowledge bases to generate adaptive situational insights. Language-model-based reasoning interfaces can further integrate human-domain insights with machine-recognized patterns, aligning intelligent systems for precise control and cross-domain synchronization. This enables coordinated decision support between humans and unmanned systems, achieving high-efficiency collaborative operations.

6. Conclusion

New technologies must be integrated with new operational concepts in order to fully realize their potential. Even with a strong foundation in artificial intelligence, failure to adopt appropriate concepts and coordinated methods of employment will still place forces at a disadvantage on the battlefield. Current research on operational demands, capability requirements, and functional needs for intelligentized warfare remains insufficient, often leading to partial or overly generalized conclusions.

Therefore, before intelligent technologies reach full maturity, it is essential to advance theoretical research and develop requirements for intelligent C2 systems, with a strategic focus on the potential of large-model applications. Future intelligentized warfare will require the integration of large models into C2 frameworks, force composition, and decision-support mechanisms. This includes modes ranging from human-in-the-loop to machine-in-the-loop C2, as well as workflow integration, operational-content alignment, and information-exchange requirements. Clarifying how commanders should leverage intelligent information services will be critical to improving decision-making capability.

At the same time, the development of AI models is highly scenario-dependent. Even large models with strong generalization still require extensive scenario-oriented training to form reliable domain-specific capabilities. In the military field, the scarcity of high-quality datasets is a major challenge. Data often reside in fragmented, isolated repositories across units, lacking standardization and unified formats. This limits the coverage, diversity, and realism needed for large-model training. As a result, large-model descriptions of scenarios are often insufficiently aligned with actual operational data.

Addressing these issues requires task-driven scenario refinement and the deliberate application of large-model technologies in representative tasks. Improving adaptability across full-spectrum missions will require the creation of diversified task datasets, iterative updates to scenario representations, and enhanced multimodal sample quality. Ultimately, breakthroughs must be made in technologies that integrate scenario requirements with large-model learning capabilities, forming cohesive solutions that respond to urgent operational needs.

References

[1] S. Hu, C. Zheng, Z. Zhou, C. Chen and G. Sukthankar, "LAMP: Leveraging Language Prompts for Multi - person Pose Estimation," arXiv preprint arXiv:2307.11934, 2023.

[2] A. Younesi, M. Ansari, M. Fazli, A. Ejlali, M. Shafique and J. Henkel, "A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends," IEEE Access, vol. 12, pp. 41180-41218, 2024.

[3] T. Schick, J. Dwivedi-Yu, R. Dessì, R. Raileanu, M. Lomeli, E. Hambro et al., "Toolformer: Language models can teach

themselves to use tools", Advances in Neural Information Processing Systems, vol. 36, pp. 68539-68551, 2023.

[4] E. Zelikman, Y. Wu, J. Mu and N. Goodman, "Star: Bootstrapping reasoning with reasoning", Advances in Neural Information Processing Systems, vol. 35, pp. 15476-15488, 2022.

[5] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang and M. S. Bernstein, "Generative agents: Interactive simulacra of human behavior", Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, pp. 1-22, 2023.

[6] G. Li, H. Hammoud, H. Itani, D. Khizbullin and B. Ghanem, "Camel: Communicative agents for 'mind' exploration of large language model society", Advances in Neural Information Processing Systems, vol. 36, pp. 51991-52008, 2023.

[7] G. Yao, H. Liu and L. Dai, "Multi-agent reinforcement learning for adaptive resource orchestration in cloud-native clusters", Proceedings of the 2nd International Conference on Intelligent Computing and Data Analysis, pp. 680-687, 2025.

[8] R. Ying, J. Lyu, J. Li, C. Nie and C. Chiang, "Dynamic Portfolio Optimization with Data-Aware Multi-Agent Reinforcement Learning and Adaptive Risk Control", 2025. [Unpublished]

[9] S. Pan and D. Wu, "Trustworthy summarization via uncertainty quantification and risk awareness in large language models", arXiv preprint arXiv:2510.01231, 2025.

[10] Y. Huang, Y. Luan, J. Guo, X. Song and Y. Liu, "Parameter-Efficient Fine-Tuning with Differential Privacy for Robust Instruction Adaptation in Large Language Models", arXiv preprint arXiv:2512.06711, 2025.

[11] S. Wang, S. Han, Z. Cheng, M. Wang and Y. Li, "Federated fine-tuning of large language models with privacy preservation and cross-domain semantic alignment", 2025. [Unpublished]

[12] X. Guo, Y. Luan, Y. Kang, X. Song and J. Guo, "LLM-Centric RAG with Multi-Granular Indexing and Confidence Constraints", arXiv preprint arXiv:2510.27054, 2025.

[13] P. Xue and Y. Yi, "Sparse Retrieval and Deep Language Modeling for Robust Fact Verification in Financial Texts", Transactions on Computational and Scientific Methods, vol. 5, no. 10, 2025.

[14] A. Xie and W. C. Chang, "Deep Learning Approach for Clinical Risk Identification Using Transformer Modeling of Heterogeneous EHR Data", arXiv preprint arXiv:2511.04158, 2025.

[15] Y. Wu, Y. Qin, X. Su and Y. Lin, "Transformer-based risk monitoring for anti-money laundering with transaction graph integration", Proceedings of the 2025 2nd International Conference on Digital Economy, Blockchain and Artificial Intelligence, pp. 388-393, 2025.

[16] Y. Li, S. Han, S. Wang, M. Wang and R. Meng, "Collaborative evolution of intelligent agents in large-scale microservice systems", arXiv preprint arXiv:2508.20508, 2025.

[17] R. Liu, Y. Zhuang and R. Zhang, "Adaptive Human-Computer Interaction Strategies Through Reinforcement Learning in Complex", arXiv preprint arXiv:2510.27058, 2025.

[18] R. Mokady, A. Hertz, K. Aberman, Y. Pritch and D. Cohen-Or, "Null-text inversion for editing real images using guided diffusion models", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6038-6047, 2023.

[19] M. Ahn, A. Brohan, N. Brown, Y. Chebotar, O. Cortes, B. David et al., "Do as I can, not as I say: Grounding language in robotic affordances", arXiv preprint arXiv:2204.01691, 2022.

[20] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo and Y. Iwasawa, "Large language models are zero-shot reasoners", Advances in Neural Information Processing Systems, vol. 35, pp. 22199-22213, 2022.

[21] Brown T., Lee A., Gomez F. When AI hallucinates: mechanisms and mitigation strategies [J]. World Science Review, 2024, 3: 22-29.

[22] Johnson M., Davis K., Clark P., et al. Enhanced entity recognition under uncertain observation conditions: a unified detection framework [J]. Information Science Review, 2023, 37(7): 65-78.