Creativity and Innovation
2025, VOL. 9, NO. 3, 138-142
DOI: 10.47297/wspciWSP2516-252729.20250903

Research on the Application of Large Language Models in Game Design for Strategic Games

Ziyu Wu

Qingdao University, Qingdao, Shandong, 266071, China

ABSTRACT
This study focuses on the application of large language models (LLMs) in the field of game design for strategic games. By analyzing the developmental needs of game design, it explores the application of LLMs in strategy generation and optimization, character behavior modeling, and natural language interaction. The research finds that LLMs can enhance gaming experiences and improve game intelligence, but they also face challenges such as high training costs, the need to strengthen model reliability, and improving controllability. Further research is required to promote their better application in game design.

KEYWORDS
Large language models; strategic game design; natural language interaction; strategy generation

1 Introduction

Strategic games, as a key intersection of artificial intelligence (AI) and game design, have long been a focal point of research. From simple board games to complex strategy games, strategic games continuously drive advancements in AI technology. Large language models, as a major breakthrough in natural language processing (NLP), provide powerful tools for strategic game design with their robust language understanding and generation capabilities. However, there are both opportunities and challenges in leveraging LLMs to enhance the intelligence of game design.

LLMs have demonstrated potential in creating richer and more expansive narrative spaces and excel in NLP tasks. By leveraging their generative and comprehension abilities, LLMs enable natural interactions with players. In strategic game design, this application can significantly enhance the gaming experience, making games more vivid and engaging. Additionally, LLMs possess strong knowledge representation and reasoning capabilities, offering intelligent decision-making support for game characters in areas such as plot development, dialogue, and mission descriptions, thereby increasing the strategic depth and challenge of games.

Despite the widespread exploration and utilization of LLMs, their ability to exhibit rational behavior in strategic environments, as represented by game theory, remains an unsolved mystery [1]. Therefore, researching the application of LLMs in strategic game design holds significant theoretical and practical importance.

2 The Development and Research of Strategic Game Design and Large Language Models

2.1 The Evolution of Strategic Game Design

With the advancement of artificial intelligence (AI) technology, a multitude of concepts have emerged, marking a paradigm shift from "computational intelligence" to "perceptual intelligence" and further to "cognitive intelligence." This evolution has introduced specialized domains such as "decision intelligence" and "game intelligence" [2].

Early strategic games like chess and Go were primarily designed around fixed rules and algorithms, with limited decision-making capabilities for in-game characters. For example:

·Chess is a two-player strategy game played on a 64-square board. Each player commands an army of 16 pieces: one king, one queen, two rooks, two bishops, two knights, and eight pawns—each with unique movement rules.

·The objective is to deliver an inescapable attack (checkmate) on the opponent's king. A "check" occurs when the king is threatened, and "checkmate" is declared if no legal move can prevent its capture, ending the game.

However, as AI technologies—particularly deep learning and reinforcement learning—have matured, game characters have progressively acquired more sophisticated decision-making abilities. These advancements enable dynamic strategy adaptation, realistic behavioral modeling, and human-like interactions, transforming traditional game design paradigms.

*** Corresponding Author:** 1256083698@qdu.edu.cn

2.2 Current Research on Large Language Models

In recent years, significant progress has been made in LLMs, making them a focal point in artificial intelligence research. Models such as OpenAI's GPT series, Google's Gemini, and Anthropic's Claude, pre-trained on massive text corpora, have demonstrated increasingly sophisticated capabilities not only in natural language understanding, generation, and reasoning, but also in a wide range of applications including text classification and question-answering systems.

Furthermore, the performance of these models has been enhanced through scaling laws, with larger parameter counts (e.g., GPT-4's estimated trillion-level parameters) and broader training datasets enabling more accurate and context-aware outputs. A key trend is the shift from unimodal (text-only) to multimodal (text, image, audio) systems, as exemplified by the Flamingo model's achievement in joint vision-language understanding [3], which enables richer interactions. Additionally, techniques such as reinforcement learning from human feedback (RLHF) and constitutional AI have improved alignment with human values and safety. Researchers are also exploring specialized applications, such as fine-tuning LLMs for domains like legal, medical, or game design.

In strategic games, preliminary explorations have been conducted with LLMs, showing potential in dynamic NPC dialogues and strategy optimization. Some studies have leveraged LLMs to generate dialogue content for game characters, enhancing immersion, while others have attempted to use LLMs for game strategy optimization and decision support. However, their reasoning capabilities in complex game-theoretic scenarios remain limited. Future research directions include efficiency optimization (e. g., mixture of experts), real-time learning, and better integration with symbolic reasoning systems.

2.3 Current Applications of LLMs in Strategic Game Design

Currently, numerous research teams and game development companies worldwide are actively exploring innovative applications of LLMs in game design. For instance, in the internationally authoritative LMSYS Chatbot Arena Leaderboard evaluations, China-developed LLMs such as Yi-Large-preview and Chat GLM have achieved top-tier global rankings. This not only demonstrates China's technical prowess in fundamental LLM research but also highlights its leading advantages in gaming industry applications.

From a practical application perspective, LLMs can significantly enhance the interactivity of game characters. Through emotion recognition and analysis of players, they can generate dynamic dialogues for continuous in-game interactions. With continuous advancements in intelligent technology, LLMs also deliver more realistic and immersive gaming experiences.

Particularly in strategic game design, LLMs have garnered increasing attention due to their unique developmental value. For example, in social deduction games like *Werewolf*, LLMs can simulate complex social behaviors among players, including forming trust-based alliances, confronting rival teams, executing strategic deception, and leading group decision-making—enabling diverse strategic experiences across different roles. By generating more natural and fluid dialogue systems and simulating varied character behaviors, LLMs make player interactions more authentic and engaging. This not only deepens immersion but also allows players to experience the nuances of different strategies.

However, current research still has several limitations. For instance, the application of large language models in complex game-theoretic scenarios remains insufficiently explored, and the integration between LLMs and strategic game design methodologies still requires further development.

3 Analysis of LLM Applications in Strategic Game Design

3.1 Strategy Optimization

3.1.1 Rule-Driven: Explicit Encoding and Logical Reasoning

In 1956, the Dartmouth Conference marked the formal birth of AI as a concept and academic discipline. The development of rule-driven approaches traces back to the early stages of artificial intelligence research, which initially focused on the "symbolic" paradigm - simulating human reasoning through explicitly encoded rules (IF-THEN logic). For instance, expert systems addressed specific problems (such as medical diagnosis and chemical analysis) using rules manually crafted by domain experts. By the 1970s, researchers began experimenting with knowledge organization through ontological semantic networks, though these efforts remained limited due to the high costs of manual knowledge base construction. Nevertheless, Bullock, Gelman and Baillargeon (1982) along with Schultz and Kestenbaum (1985) provided excellent overviews of the general constraint types utilized during the learning of specific causal rules [4].

The period from the 1950s through the late 1980s represented the dominant era of rule-driven technological paradigms. In traditional game design for strategic games, strategy generation primarily relied on rule-driven methods -

reasoning and decision-making based on predefined rules and logic, where the formulation of rules constrained player behavior to promote positive strategic development.

3.1.2 Learning-Driven: Efficient Conversion from Knowledge to Ability

Learning-driven approaches serve as the core impetus for motivating individuals to actively learn and continuously grow. This paradigm transcends passive knowledge acquisition by emphasizing intrinsic motivation, goal orientation, self-efficacy, proactive exploration, and continuous feedback—key characteristics that collectively transform learners from "passive recipients" to "active creators," enabling effective conversion of knowledge into competence . A prime example is GPT-3, an exemplary LLMs that achieves remarkable language generation capabilities through training on massive datasets (45TB of diverse texts, including web pages, books, and academic papers). This enables the model to precisely capture linguistic patterns and produce natural, coherent text, demonstrating strong potential in dialogue systems, content creation, and beyond.

With advancements in LLMs, strategy generation is transitioning toward learning-driven paradigms. Breakthroughs in deep neural networks and end-to-end learning facilitate dynamic adaptation and intelligent strategy optimization. For instance:

·DeepMind's AlphaGo[5] (as shown in the Figure 1) mastered Go strategies through self-play learning

·Modern LLMs can generate zero-shot strategies[6] by combining natural language game rules with learned patterns

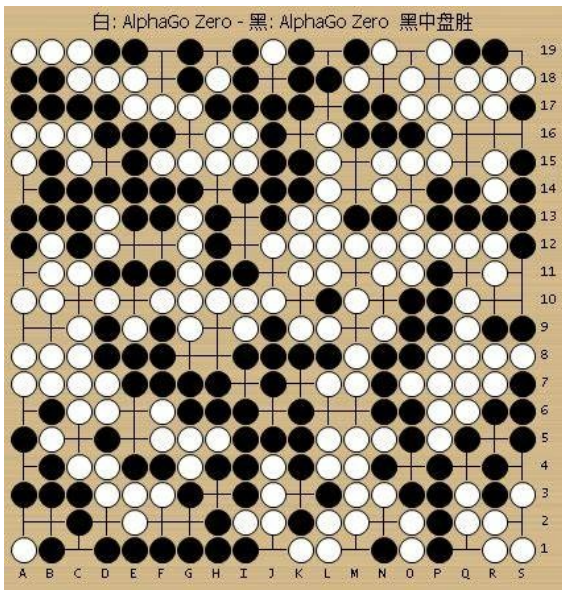

Figure 1 Go strategies learned by AlphaGo through self-play

(*"Image obtained from: Author Name (if available), 'AlphaGo Zero: A Game Changer in AI,' CSDN Blog, [Online]. Available: [URL]. Accessed: [Date]."*)

3.1.3 The Evolution from Rule-Driven to Learning-Driven

The rule-driven phase centered on manually defined explicit rules, implementing functionality through "if-then" statements - as seen in early grammar checkers that relied on syntactic rules for text parsing. While offering advantages like logical transparency and ease of debugging, this approach suffered from poor scalability, weak generalization, and the necessity to manually cover all scenarios, resulting in high maintenance costs. Furthermore, exponential growth in rule complexity led to the "knowledge acquisition bottleneck."

The statistics-driven stage introduced machine learning to reduce manual rule dependency through systematic statistical methods like decision trees and Bayesian networks. For instance, spam filtering systems learned feature correlations from email data to achieve automated classification. While overcoming the rule explosion problem, this phase faced interpretability challenges due to immature neural network technology and model opacity.

The learning-driven deep learning phase established neural networks as its core, with their revival directly catalyzing the birth of Deep Learning and triggering comprehensive transformation in AI. Its strengths lie in end-to-end optimization and contextual understanding capabilities. However, neural networks demand substantial training resources, and enhancing their safety, controllability and intelligence remains a critical focus for future development.

In summary, we observe that the evolution from rule-driven to learning-driven paradigms fundamentally represents a shift from explicit rules to data-driven intelligent emergence. This progression can be concretely illustrated through the following timeline:(as shown in the Table 1)

Table 1 Timeline of Evolution from Rule-driven to Learning-driven Paradigms

Year	Event	Impact
1956	Dartmouth Conference, birth of AI concept	Symbolic AI dominated early research
1980	Commercialization of expert systems (e.g., MYCIN)	Rule-driven technologies implemented in healthcare, finance, etc.
1986	Backpropagation algorithm proposed	Neural network renaissance, pioneering learning-driven approaches
2006	Hinton introduces Deep Belief Networks	Deep learning enters academic mainstream
2012	AlexNet wins ImageNet competition	Deep learning breakthrough in computer vision
2016	AlphaGo defeats Lee Sedol	Integration of RL and DL, AI surpasses human performance
2020-Present	Release of GPT-3, DALL-E, etc.	Large-scale pretrained models become research focus

3.2 Character Behavior Modeling: From Scripted to Human-Like

Character modeling represents a core technology for creating virtual NPCs (Non-Player Characters) in games, where NPCs serve to advance plotlines, provide player hints, and deliver fragmented narrative information - enabling players to reconstruct complete storylines while enhancing immersion. Through pre-training on massive textual datasets, LLMs excel at capturing nuanced linguistic patterns, empowering multidimensional NPC design across personality traits, behavioral logic, and speech styles. Their advanced natural language understanding and generation capabilities allow for richer, more authentic character behaviors by simulating human decision-making processes and behavioral patterns. This results in NPCs that align with player expectations so convincingly that distinguishing between AI and human behavior becomes challenging.

For instance, in simulation/strategy games, LLMs can emulate player business decisions (investments, production, sales, etc.), with NPCs dynamically responding to market fluctuations and competitor actions. This capability not only increases gameplay challenge but also significantly enhances entertainment value through:

·Context-aware strategic adaptation

·Behavior consistency with game world rules

·Emergent interpersonal dynamics between characters

3.3 Natural Language Interaction

In strategic games, natural language interaction enhances player immersion and engagement. LLMs can serve as intelligent dialogue systems within games, enabling real-time conversations with players. For example, in role-playing strategy games, players can engage in natural language exchanges with in-game characters, creating dynamic and interactive experiences. LLMs excel at interpreting player intent and generating contextually appropriate responses by analyzing player decisions, thereby enriching gameplay and heightening realism.(as shown in the Figure 2)

Beyond player-NPC interactions, natural language systems also facilitate strategic collaboration among players. By processing voice or text inputs from teammates, LLMs leverage their trained understanding of human language to: Semantically analyze tactical discussions and shared information, Assess battle conditions and evolving game states,

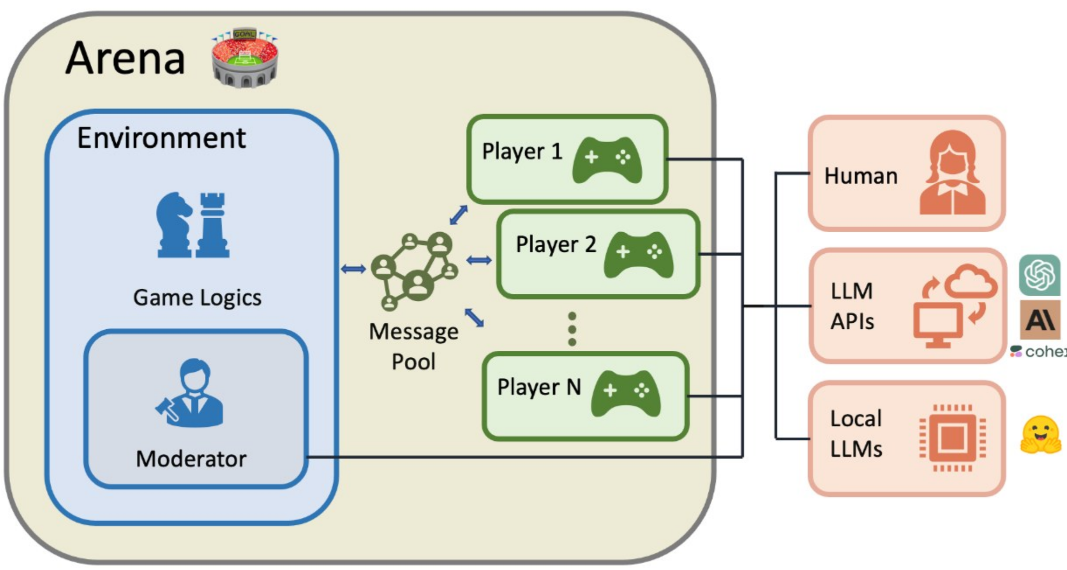

Figure 2 Applications of Large Language Models in Role-Playing Strategic Games

(Res Obscura. (2023, October 15). *LLM-based educational games will [subtitle if any]*. Res Obscura Substack. https://resobscura.substack.com/p/llm-based-educational-games-will)

Generate actionable recommendations to optimize team strategy.

This dual application—both for immersive storytelling and competitive strategy refinement—demonstrates how LLMs transform traditional gameplay into adaptive, linguistically rich experiences.

4 Conclusion

LLMs hold significant promise for strategic game design, with applications ranging from strategy optimization to character behavior standardization. These capabilities enhance player immersion, improve interactive experiences, and expand creative possibilities within games. While LLMs deliver richer, more engaging gameplay and drive innovation across the industry, practical implementation faces challenges including high training costs, reliability concerns, and controllability limitations.

For instance, Phelps et al. [7] demonstrated GPT-3.5's performance in iterated prisoner's dilemma scenarios, revealing that while LLMs can exhibit both altruistic and self-interested behaviors, they still struggle with dynamic interactions and adaptive conditional reciprocity strategies [8].

·Future advancements may address these limitations through:

·Multimodal integration (combining text, visual, and auditory inputs)

·Multi-task learning frameworks

·Interactive generation techniques

·Hybrid architectures merging causal reasoning with reinforcement learning

Such developments could solidify LLMs' reliability in gaming contexts while propelling the industry toward new frontiers of intelligent gameplay design.

References

[1] Hua, W., Liu, O., Li, L., Amayuelas, A., Chen, J., Jiang, L., ... & Zhang, Y. (2023). *Game-theoretic LLM: Agent Workflow for Negotiation Games*. [Unpublished manuscript]. Rutgers University.

[2] Luo, J., Zhang, W., Su, J., et al. (2024). Research progress of multi-agent learning in games. *Systems Engineering and Electronics*, *46*(5), 1628-1655.

[3] Alayrac, J.-B., Donahue, J., Luc, P., Miech, A., Barr, I., Hasson, Y., ... & Simonyan, K. (2022). Flamingo: A visual language model for few-shot learning. *Advances in Neural Information Processing Systems*, *35*, 23716-23736.

[4] Schultz, T. R., & Kestenbaum, N. R. (1993). Learning causal patterns: Making a transition from data-driven to theory-driven learning. *Machine Learning*, *11*(2-3), 173-194.https://doi.org/10.1023/A:1022607123329

[5] Silver, D., Hubert, T., Schrittwieser, J., et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play.*Science*,*362*(6419), 1140-1144.https://doi.org/10.1126/science.aar6404

[6] Huang, M., Tang, J., et al. (2023). SPaR: Self-play with tree-search refinement to improve instruction-following in large language models. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 10289-10305.

[7] Phelps, S., Russell, Y., et al. (2023). Investigating emergent goal-like behaviour in large language models using experimental economics. *arXiv preprint*. https://arxiv.org/abs/2305.07970

[8] Hua, W., et al. (2023). *Supplementary materials for: Game-theoretic LLM: Agent Workflow for Negotiation Games*. [Dataset]. Rutgers University.