Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD27362

The Atlas of Technological Evolution in Chess AI: Paradigm Shifts from Classical Models, Algorithmic Revolutions to Boundary Expansion and Challenges

Zhile Fang[1], Xinyang Li[2*], Jiawei Su[3]

[1]*St Bees School, Cumbria County, The United Kingdom*
[2]*College of Computer and Control Engineering, Qiqihar University, Qiqihar, China*
[3]*Artificial Intelligence (Sino-Foreign Cooperative Education), Chang'an University, Xi'an, China*
**Corresponding Author. Email: 2023026027@qqhru.edu.cn*

Abstract. The technological evolution of artificial intelligence in game playing, particularly in chess, exhibits distinct generational shifts. This paper systematically categorizes existing research into three technical paradigms: the Rule-driven Paradigm, the Data-driven Paradigm, and the General Algorithm Paradigm, analyzing their intrinsic evolutionary logic—including optimization of search mechanisms, migration of knowledge representations, and paradigm shifts in training objectives. Through comparative analysis, it reveals a synergistic evolution between enhanced algorithmic autonomy and optimized computational efficiency. Addressing current research bottlenecks, the study explores potential pathways toward general game intelligence, such as incorporating meta-learning architectures and learning universal game representations. Findings demonstrate that chess AI's progression not only expands the boundaries of game intelligence but also offers methodological insights for designing general decision-making agents. This advancement fundamentally hinges on progressively reducing reliance on human priors while achieving adaptability to complex gaming environments through algorithmic innovations—particularly diverse strategy generation. This trajectory underscores the continuous pursuit of autonomy and efficiency in AI evolution.

Keywords: Artificial Intelligence Chess, Algorithm Paradigm Shift, Reinforcement Learning

1. Introduction

Against the backdrop of rapid advancements in artificial intelligence, strategic games—particularly chess—have emerged as a critical testing ground for evaluating AI capabilities. As the quintessential example of perfect-information games, chess not only features an enormous state space and profound strategic complexity but has also witnessed pivotal milestones in AI evolution. From DeepBlue's historic victory over Garry Kasparov to AlphaZero's revolutionary self-play learning and DeepNash's breakthroughs in imperfect-information scenarios, AI applications in chess have consistently redefined the boundaries of machine intelligence. A thorough examination of this evolutionary trajectory offers invaluable insights into the underlying logic of AI strategy

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD27362

optimization while providing actionable frameworks for implementing general AI in decision-making contexts. Consequently, systematically analyzing the developmental paradigms of chess AI carries significant theoretical and practical implications for advancing game intelligence toward higher-order, universally adaptable decision-making agents.

Current research on board-game AI reveals three distinct evolutionary phases: The first generation, exemplified by DeepBlue, established the rule-driven paradigm through human expert knowledge and brute-force search. The second generation, represented by AlphaZero, discarded manual rules to pioneer the data-driven paradigm via deep reinforcement learning and self-play, achieving end-to-end strategy optimization. The third generation, including systems like DeepNash, extends into imperfect-information games through innovations such as Regularized Nash Dynamics (R-NaD), marking the exploration of a general algorithmic paradigm [1]. These systems demonstrate clear generational shifts in technical architecture, training mechanisms, computational dependencies, and knowledge representation. Nevertheless, limitations persist: early systems suffered from over-reliance on expert knowledge and poor generalizability; mid-generation systems, while self-learning, demanded excessive hardware resources and lacked interpretability; modern systems, despite adapting to imperfect information, remain constrained by heavy dependence on large-scale simulations, with room for improvement in real-time performance and scalability. These challenges necessitate a unified, paradigm-level deconstruction of AI's evolution in game theory to uncover the intrinsic relationships and complementary potential across generations—a crucial step toward methodological guidance for future AI system design.

To this end, this study conducts an in-depth analysis of paradigm shifts in game AI development, focusing on three landmark chess systems: Deep Blue, AlphaZero, and DeepNash. By dissecting their core technical modules—including search mechanisms, knowledge representation, and training logic—the paper systematically maps the characteristics, evolutionary pathways, and synergistic relationships among these paradigms. The investigation extends beyond the interplay between algorithmic mechanisms and hardware resources to scrutinize fundamental differences in data acquisition methods, intrinsic strategy-generation mechanisms, and decision-optimization objectives across generations. Through this longitudinal comparative framework, the research elucidates the logical progression from rule-driven to data-driven and ultimately general algorithmic paradigms while exploring the technical drivers and boundary challenges in the evolution of game AI toward higher-order intelligent agents.

2. Methods: paradigm shifts in chess AI evolution

The technological evolution of AI in games exhibits clear generational transitions, driven by progressive reductions in reliance on human priors and innovations in algorithmic architectures. Early systems like DeepBlue established the rule-driven paradigm, combining expert-designed evaluation functions with brute-force search. Mid-stage systems like AlphaZero pioneered the data-driven paradigm, using self-play and neural networks for strategy generation. Recent systems like DeepNash advanced the general algorithmic paradigm, solving equilibrium-computation challenges in imperfect-information games. By deconstructing the core technical components—search mechanisms, knowledge representation, and training logic—this study reveals the iterative motivations and complementary relationships across these paradigms.

DOI: 10.54254/2755-2721/2025.LD27362

2.1. Foundation and limitations of the rule-driven paradigm

As a representative of the rule-driven paradigm, Deep Blue achieved its core breakthrough through the synergy of hardware-accelerated brute-force search and expert knowledge systems. The system employed the α-β pruning algorithm to optimize minimax search, reducing node evaluations by 70% by anticipating opponents' optimal responses [2]. It incorporated over 8,000 handcrafted features (e.g., pawn structure weights) into its evaluation function. The expert system designed heuristic rules through game record analysis and human knowledge (e.g., piece values, positional advantages) to assist search pruning, relying on manually tuned evaluation functions—such as opening books and endgame tablebases—to enumerate possible moves within limited depths. At the hardware level, Deep Blue utilized custom-designed VLSI chess chips and dedicated search accelerators, constructing an FPGA-based parallel interconnection architecture with 480 chess chips. Leveraging the parallel computing power of the RS/6000 supercomputer, it achieved an evaluation speed of 200 million nodes per second, with an average search depth of 6-8 moves (extending beyond 20 moves on critical paths). Technologies like global hash table sharing reduced redundant computations by over 40% [3].

However, while this rule-driven paradigm achieved groundbreaking success in chess, it also revealed significant limitations. Its strategy generation relied entirely on pre-programmed expert knowledge bases and fixed evaluation functions—a rigid architecture that lacked self-evolution and adaptation capabilities. When confronted with games featuring more complex state spaces or scenarios requiring dynamic strategy adjustments, DeepBlue's constraints became particularly pronounced. This heavy dependence on handcrafted rules, coupled with strict requirements for perfect information environments, greatly restricted its applicability. Nevertheless, the fundamental framework combining brute-force search with expert systems, established by DeepBlue, provided a crucial technical reference and evolutionary foundation for subsequent advancements in game AI.

2.2. Collective intelligence revolution in the data-driven transition paradigm

Stockfish exemplifies the data-driven transition paradigm, primarily addressing dynamic optimization of evaluation functions and enhancement of search efficiency. By incorporating techniques such as Null-Move Pruning and Lazy Move Generation, it increased the average search depth to over 20 plies. Its evaluation function dynamically adjusts piece weights (e.g., pawn structure, king safety) through machine learning, integrating more than 100 dynamic features. Leveraging collective intelligence from open-source communities and CPU multithreading optimization, it transcended traditional rule-based limitations. However, this system still relies on knowledge inputs from human experts and lacks self-play learning capabilities, preventing autonomous strategy evolution—a limitation later resolved by the next-generation system AlphaZero.

2.3. Revolutionary breakthrough in the data-driven paradigm

The emergence of AlphaZero marks the critical transition toward a data-driven paradigm. As a mid-term breakthrough in AI board game algorithms, AlphaZero completely discarded human knowledge by adopting a deep reinforcement learning architecture: its core innovation lies in the synergy between Monte Carlo Tree Search (MCTS) and residual neural networks (ResNet), training from scratch through self-play to dynamically generate strategies and value assessments. Specifically, the policy network generates candidate moves, the value network evaluates position win probabilities,

DOI: 10.54254/2755-2721/2025.LD27362

and MCTS dynamically expands high-value paths; through self-play, it generated 4.4 million training games within 4 hours, driving a 592-point Elo rating improvement. This breakthrough enabled AlphaZero to surpass human performance across Go, chess, and other domains—defeating traditional engines like Stockfish after merely 4 hours of training—demonstrating the immense potential of data-driven approaches over handcrafted rules. Subsequent research further validated the framework's scalability: for chess variants (e.g., Xiangqi), incorporating supervised pretraining into the AlphaZero framework combined with MCTSenabled strategy optimization under limited computational resources [4], significantly reducing randomness during the cold-start phase. Furthermore, to explore the boundary expansion potential of the AlphaZero framework, follow-up research such as AZdb achieved multi-agent collaboration by introducing a latent-conditioned architecture. This method employs a behavioral diversity reward mechanism to encourage distinct agents to generate differentiated strategies, integrating optimal decisions through sub-additive planning. Experiments demonstrate that AZdb outperformed standalone AlphaZero by 50 Elo in puzzle-solving tasks and resolved Penrose chess positions unsolvable by traditional engines [5].

This paradigm first achieved cross-game universality (Go/Shogi/Chess) with a single neural network architecture capable of handling multiple board games. However, it also revealed new limitations: dependence on intensive computational resources from high-performance TPU clusters (thousands of chips), and underperformance in imperfect-information games like poker. These constraints delineate key improvement priorities for subsequent research.

2.4. General algorithmic paradigm: breaking the imperfect-information barrier

DeepNash achieved dual breakthroughs in imperfect-information games (e.g., Stratego): Algorithmically, it proposed R-NaD (Regularized Nash Dynamics), compressing Stratego's state space (10^535) into a 256-dimensional latent space (VAE-encoded), improving Nash equilibrium efficiency by 100x [6]. Training-wise, it employed elastic weight consolidation to adaptively update adversarial module parameters hourly.

Unlike perfect-information games (e.g., Go), DeepNash handled information asymmetry and hidden states, extending AI's reach to complex real-world scenarios. It achieved human-expert performance (84% win rate) in Stratego using only CPUs, though still reliant on massive parallel simulations (tens of millions of games). This paradigm proved that equilibrium strategies can be discovered purely through data-driven methods without perfect information, laying new groundwork for general game intelligence.

Further advancing this frontier, the Student of Games (SoG) algorithm unified perfect- and imperfect-information equilibrium-solving under a single framework, combining guided search (GT-CFR), self-play learning, and game-theoretic reasoning [7]. This marked a paradigm shift toward broader generality.

2.5. Iterative logic and future challenges

The evolution follows a clear progression: DeepBlue's rigid rules spurred AlphaZero's self-learning, while AlphaZero's imperfect-information limitations drove DeepNash's equilibrium innovations. This reflects a redefinition of computational goals—from rule execution (DeepBlue) to win-rate optimization (AlphaZero) and finally to Nash equilibrium discovery (DeepNash). These generational shifts are systematically contrasted in Table 1 across three critical dimensions: algorithmic architectures, hardware dependencies, and data acquisition methods.

DOI: 10.54254/2755-2721/2025.LD27362

Table 1. Contrasts the three paradigms across algorithms, hardware, and data sources

Era	Algorithms	Hardware	Data Sources
Early	α-β pruning +expertsystems	Supercomputers(CPU)	Human games+manual rules
Mid	MCTS+neuralnetworks(RL)	TPU/GPU clusters	Self-play generated data
Modern	R-NaD(Nash Dynamics)	CPU/GPU hybrid	Imperfect-information self-play

Current bottlenecks include real-time strategy latency and multi-agent equilibrium convergence. While SoG unified perfect/imperfect-information solving, its GT-CFR-based search suffers from high computational complexity in real-time scenarios [7]. Emerging directions like quantum Monte Carlo search and meta-game agents aim to push toward general decision architectures.

3. Experiments

3.1. AlphaZero experimental results

AlphaZero (2017) demonstrated its general game-playing prowess in Science: Trained via self-play (440K games in 4 hours), it achieved a 592-point Elo gain, surpassing Stockfish 8 (then the top chess engine). In Go, it defeated AlphaGo Lee 100-0 with 1/8th the computation. Hardware: 5,000 TPU v1 chips; 800 MCTS simulations per move; ResNet updates every second. Key conclusion: Pure reinforcement learning, devoid of human priors, can achieve superhuman performance across multiple games [4-6].

3.2. Qualitative analysis

AlphaZero's revolutionary contributions:

1.Cross-game generality: A unified neural architecture mastered chess, Go, and shogi without game-specific modifications, demonstrating unprecedented transferability in deep reinforcement learning.

2.Training efficiency: Achieved superhuman performance within 4 hours of self-play, surpassing engines like Stockfish that required decades of human expertise refinement.

3.Strategic paradigm shift: Emergent behaviors (e.g., material sacrifice for long-term initiative) fundamentally challenged centuries of established human heuristics.

This work established the tabula rasa learning paradigm, creating a foundational framework for subsequent breakthroughs in imperfect-information games (e.g., DeepNash's R-NaD algorithm) [6,8].

4. Challenges and prospects

Overall, the application of artificial intelligence in chess has undergone three evolutionary stages: from rule-based systems to data-driven approaches, and subsequently to the paradigm of general-purpose algorithms. This progression has significantly advanced game-playing intelligent systems in areas such as strategic generation, autonomous learning, and imperfect information handling. Nevertheless, despite the considerable technological achievements to date, the field still faces several fundamental challenges that demand in-depth investigation and systematic breakthroughs. These challenges can be broadly categorized into the following three aspects:

DOI: 10.54254/2755-2721/2025.LD27362

4.1. Insufficient interpretability of decision-making

Deep learning systems, exemplified by AlphaZero and DeepNash, have far surpassed traditional engines in gameplay performance. However, their strategic generation processes exhibit a high degree of black-box characteristics, lacking clear and explicit explanations for key decision-making nodes. This limitation, to some extent, hinders human experts' ability to trust and understand the system's behavior. Current research attempts to enhance model transparency through methods such as visual feature mapping and strategic path tracing, but these efforts largely remain at a relatively superficial level of interpretability.

Future directions can be explored in greater detail from multiple perspectives. First, the Chain-of-Thought (CoT) approach can be integrated into chess AI, enabling the model to produce interpretable reasoning processes that enhance human comprehension of its strategic behavior. Second, natural language generation techniques can be employed to allow the system to articulate its move-selection logic in real time, using language that is readily understandable to humans. Furthermore, symbolic logic modules can be combined with neural networks, whereby the symbolic knowledge graph performs causal structure analysis on top of deep model outputs, achieving "causal tracing" of strategic decision-making paths. By integrating these approaches, it is possible to construct a new game-playing AI framework in which "interpretation is strategy" [9].

4.2. High dependence on hardware resources

Systems such as AlphaZero rely on TPU clusters to conduct large-scale self-play training, which raises the entry threshold for research on high-performance game-playing AI to an excessively high level, hindering the inclusive development of the technology and limiting participation by smaller teams. Some existing studies have proposed lightweight network architectures and model distillation mechanisms, aiming to reduce computational costs at the expense of some performance. However, their effectiveness in highly complex game environments remains constrained.

In the future, reliance on high-performance hardware can be further mitigated from both model design and training mechanisms. In terms of model architecture, lightweight network structures such as efficient Transformers and TinyNet can be adopted, while techniques such as model pruning, quantization, and knowledge distillation can be applied to compress model size and adapt to mid- and low-end computing resources. In terms of training mechanisms, imitation learning and offline reinforcement learning frameworks can be introduced to pre-train models using historical game records, significantly reducing the demand for self-play-generated data [10]. At the same time, incorporating mechanisms such as federated learning to distribute the training process across multiple low-power devices may enable the development of low-threshold, widely accessible game-playing agents.

4.3. Weak transferability

At present, most existing AI systems focus primarily on specific board game environments and lack the practical ability to transfer strategies across different rules or game structures. Although AlphaZero demonstrates a certain degree of cross-game adaptability, it still requires network retraining in practice. Some related studies have attempted to enhance strategy generalization capabilities through multi-task learning and neural architecture search (NAS), yet they lack a unified theoretical framework to provide foundational support.

DOI: 10.54254/2755-2721/2025.LD27362

To improve the cross-environment generalization capabilities of game-playing AI, future work could explore a deeply integrated pathway combining unified representation frameworks with multimodal learning. First, a cross-game shared semantic embedding of game states could be established, enabling models to identify structural commonalities in strategies. Second, NAS can be combined with transfer learning to automatically adapt network architectures to different game environments. Furthermore, large language models (LLMs), such as GPT, could be leveraged for their embedded game knowledge and abstract reasoning capabilities, transforming them into "knowledge engines" for strategy generation. These engines could work collaboratively with deep strategy networks to produce interpretable transfer strategies [11]. This direction holds the potential to drive the shift from specialized models toward general-purpose game-playing agents.

5. Conclusion

This paper systematically reviews the technological evolution of artificial intelligence (AI) in chess. Addressing the paradigm shifts from classical models to cutting-edge algorithms, we establish an analytical framework centered on three generations of technical paradigms: The early rule-driven paradigm, exemplified by DeepBlue, relied on expert knowledge bases and hardware-accelerated brute-force search. The mid-term data-driven paradigm, revolutionized by AlphaZero, achieved breakthroughs through deep reinforcement learning integrated with Monte Carlo Tree Search (MCTS). By generating data via self-play, it entirely discarded human priors. Recently, the generalized algorithmic paradigm, marked by DeepNash, resolved equilibrium-solving challenges in imperfect-information games through R-NaD. It achieved human-expert performance in Stratego (84% win rate), proving that purely data-driven equilibrium strategy discovery requires no full information observability. Concurrently, research based on the data-driven paradigm (e.g., AZdb) has advanced strategy diversity and multi-agent collaboration.

In examining the evolutionary logic, this paper conducts a focused comparative analysis of the significant differences and advancements across three generations of paradigms in key dimensions: search mechanisms (from α-β pruning to MCTS and then to R-NaD), knowledge representation (from handcrafted features to neural networks and then to latent space encoding), and training objectives (from rule enforcement/win-rate maximization to Nash equilibrium solving). This longitudinal comparison clearly reveals the core trajectory of technological development: computational objectives continuously escalate, hardware reliance transitions from supercomputers to TPU clusters and then to heterogeneous CPU/GPU architectures, and system autonomy progressively strengthens (from manual rule input to self-play data generation and then to model-free policy iteration). Concurrently, this paper objectively identifies the primary challenges in current research: the insufficient interpretability of policy generation limits human understanding and trust; the heavy dependence on high-performance computing resources (e.g., TPU clusters) hinders accessibility; and the generalization and transfer capabilities of policies across different board games or rule sets remain weak.

In summary, by deconstructing the paradigm shift trajectory of chess AI, this paper not only delineates the intrinsic evolutionary logic and complementary relationships spanning rule-driven, data-driven, and general algorithm paradigms, but also systematically synthesizes existing core bottlenecks. Building upon this analysis, the paper further explores potential future breakthroughs: integrating symbolic reasoning and causal modeling to enhance interpretability; reducing computational costs via techniques such as meta-learning, model distillation, and supervised pretraining—for instance, recent research demonstrates that methods incorporating supervised pretraining can achieve policy convergence within 4 days on GPUs, offering an empirical solution

DOI: 10.54254/2755-2721/2025.LD27362

for resource-constrained scenarios; and exploring neural-semantic networks incorporating large language models to strengthen policy generalization and transfer capabilities. Concurrently, collaborative frameworks based on diverse policies prove that reducing computational redundancy through subsumbtion planning (e.g., requiring only 625 simulations to maximize diversity gains) can provide lightweight solutions for resource-limited settings. This study contends that a profound understanding of this evolutionary map provides crucial methodological insights for advancing game intelligence toward higher-order, more universal general decision agents.

Authors contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Klein, D. (2022). Neural networks for chess. arXiv preprint arXiv: 2209.01506.
[2] Hsu, F.-H. (2002). Behind Deep Blue: Building the computer that defeated the world chess champion. Princeton University Press.
[3] Hsu, F.-H. (1999). IBM's Deep Blue chess grandmaster chips. IEEE Micro, 19(2), 70–81.
[4] Zahavy, T., Veeriah, V., Hou, S., Waugh, K., Lai, M., Leurent, E., Tomašev, N., Schut, L., Hassabis, D., & Singh, S. (2024). Diversifying AI: Towards Creative Chess with AlphaZero. arXiv preprint arXiv: 2308.09175v3.
[5] Zahavy, T., Veeriah, V., Hou, S., Waugh, K., Lai, M., Leurent, E., Tomašev, N., Schut, L., Hassabis, D., & Singh, S. (2024). Diversifying AI: Towards Creative Chess with AlphaZero. arXiv preprint arXiv: 2308.09175v3.
[6] Perolat, J., De Vylder, B., Hennes, D., Tarassov, E., Strub, F., de Boer, V., Muller, P., Connor, J. T., Burch, N., Anthony, T., McAleer, S., Elie, R., Cen, S. H., Wang, Z., Gruslys, A., Malysheva, A., Khan, M., Ozair, S., Timbers, F., Pohlen, T., Eccles, T., Rowland, M., Lanctot, M., Lespiau, J.-B., Piot, B., Omidshafiei, S., Lockhart, E., Sifre, L., Beauguerlange, N., Munos, R., Silver, D., Singh, S., Hassabis, D., & Tuyls, K. (2022). Mastering the game of Stratego with model-free multiagent reinforcement learning. Science, 378(6622), 990-996.
[7] Schmid, M., Moravčík, M., Burch, N., Kadlec, R., Davidson, J., Waugh, K., Bard, N., Timbers, F., Lanctot, M., Holland, G. Z., Davoodi, E., Christianson, A., & Bowling, M. (2023). Student of Games: A unified learning algorithm for both perfect and imperfect information games. Science Advances, 9(44), eadg3256.
[8] Sadler, M., & Regan, N. (2019). Game Changer: AlphaZero's groundbreaking chess strategies and the promise of Al. New In Chess.
[9] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems, 35, 24824-24837.
[10] Alajlan, N. N., & Ibrahim, D. M. (2022). TinyML: Enabling of inference deep learning models on ultra-low-power IoT edge devices for AI applications. Micromachines, 13(6), 851.
[11] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., ... & Lample, G. (2023). Llama: Open and efficient foundation language models. arXiv preprint arXiv: 2302.13971.