

Review

Not peer-reviewed version

From Minimax to Self-Play: A Taxonomy of Decision Quality, Optimality, and Robustness in Game-Playing Search

[Vinuka Silva](#)* and Nipuna Senanayake

Posted Date: 6 May 2026

doi: 10.20944/preprints202605.0057.v1

Keywords: game playing artificial intelligence; neural guided search; algorithm taxonomy; performance evaluation; metaheuristic optimization; board games; adversarial game tree search




Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

From Minimax to Self-Play: A Taxonomy of Decision Quality, Optimality, and Robustness in Game-Playing Search

Vinuka Silva ^{1,2,*}  and Nipuna Senanayake ¹

¹ School of Computing, Informatics Institute of Technology, 10 Trelawney Pl, Colombo 00400, Sri Lanka

² School of Computing, Engineering and Technology, Robert Gordon University, Aberdeen, Scotland

* Correspondence: vinukassilva@gmail.com

Abstract

Search algorithms have long underpinned game-playing artificial intelligence, however as game domains expanded from small deterministic board games to large scale, partially observable, real-time environments, no survey has systematically organized this evolving literature into a unified taxonomy or evaluated algorithms through a consistent design space. This taxonomy based survey addresses that gap through domain-scoped literature clustering, analyzing 61 works from 1946 to 2025 across six thematic clusters: spatial pathfinding and navigation, adversarial game tree search, Monte Carlo tree search and bandit based planning, metaheuristic optimization, learning augmented search, and search under uncertainty and partial observability. A four dimensional design space — covering interaction topology, information structure, computational regime, and source of search guidance enables consistent cross-cluster comparison of hybrid approaches. Analysis reveals a paradigm shift from analytic correctness and proof driven evaluation toward empirical benchmarking, sampling based planning, and neural guided search. Cross-cluster synthesis identifies fundamental tensions among decision quality, formal guarantees, and resilience under uncertainty, and documents an evolution in evaluation methodology from deterministic metrics to distributional robustness testing. Open challenges are identified, pointing toward principled frameworks for managing trade offs among quality, optimality, and robustness. This survey provides artificial intelligence researchers and game developers with a structured reference for selecting and evaluating gaming search algorithms across diverse environments.

Keywords: game playing artificial intelligence; neural guided search; algorithm taxonomy; performance evaluation; metaheuristic optimization; board games; adversarial game tree search

1. Introduction

Search algorithms have been central to the development of artificial intelligence for games since the earliest days of computer science. From the first attempts to program computers to play Chess in the mid twentieth century to contemporary systems capable of mastering Go, Chess, Shogi, Atari games, and large scale real-time strategy (RTS) environments [1,2], games have served as both a testbed and a driver for advances in search methodology. Across this long history, three interrelated but conceptually distinct dimensions have consistently shaped algorithm design and evaluation in gaming contexts: **decision quality**, **optimality**, and **robustness**.

Decision quality in games concerns the ability of an algorithm to consistently select strong actions and strategies that lead to effective game play outcomes. Unlike classical optimization problems, games demand decisions that are not only locally good but also strategically coherent over long horizons, sensitive to opponent behavior, and adaptive to changing game phases. Early game playing programs emphasized handcrafted evaluation functions and heuristic guidance to approximate high quality play in tractable domains such as Chess [3]. As games grew in complexity and scale, decision

quality increasingly became an empirical concept, measured through win rates, scores, or relative performance against strong baselines rather than through explicit guarantees [4,5].

Optimality, by contrast, is rooted in theoretical notions from game theory and search theory, including minimax optimality, Nash equilibrium, and optimal control. In small or well structured games, search algorithms can, in principle, compute optimal play through exhaustive or pruned exploration of the game tree [3,6]. However, most modern games exhibit state spaces, branching factors, and temporal constraints that render exact optimality infeasible. As a result, gaming research has progressively shifted from exact optimality toward bounded, asymptotic, or approximate notions of optimal play, as exemplified by alpha–beta pruning [6], best-first minimax variants [7], and later by Monte Carlo Tree Search (MCTS) methods with asymptotic convergence guarantees [8,9]. The tension between theoretical optimality and practical feasibility is particularly acute in real-time and multi agent games, where strict decision deadlines preclude deep deliberation [10,11].

Robustness addresses a complementary challenge: maintaining acceptable performance under uncertainty, variability, and adversarial pressure. Games routinely expose search algorithms to stochastic outcomes, imperfect information, hidden state, opponent adaptation, and fluctuating computational budgets. Classical search theory explicitly addressed uncertainty through probabilistic models of detection and pursuit–evasion [12–14], while later work extended robustness considerations to imperfect information games and partial observability [15,16]. In contemporary gaming AI, robustness has emerged as a critical requirement, particularly in RTS games with fog of war [17], in real-time environments with variable time constraints [18,19], and in systems that rely on learned models or policies that may generalize imperfectly outside their training distribution [20,21].

Although decision quality, optimality, and robustness are deeply intertwined, they are often in tension. Algorithms that aggressively pursue optimality may be brittle under uncertainty or time pressure, while highly robust methods may sacrifice peak decision quality or theoretical guarantees. Over time, gaming research has repeatedly rebalanced these dimensions in response to new game domains, computational resources, and methodological innovations. Notable paradigm shifts include the transition from deterministic minimax search to sampling based planning with MCTS [8,9], and the subsequent integration of learned policies and value functions to guide search in large, complex games [5,20,22]. Each shift reflects a changing understanding of how best to trade off decision quality, optimality, and robustness in practice.

This survey examines the **evolution of decision quality, optimality, and robustness in search algorithms for gaming** through a unified, gaming specific lens. Rather than organizing the literature purely by algorithmic families or chronological order, the survey adopts a taxonomy grounded in the unique characteristics of games, including interaction topology, information structure, computational regime, and sources of search guidance. Within this framework, the survey traces how different classes of search algorithms ranging from spatial pathfinding and adversarial game tree search to Monte Carlo planning, optimization based methods, learning augmented search, and search games under uncertainty have addressed the three core dimensions over time. Figure 1 provides a structural overview of the survey.

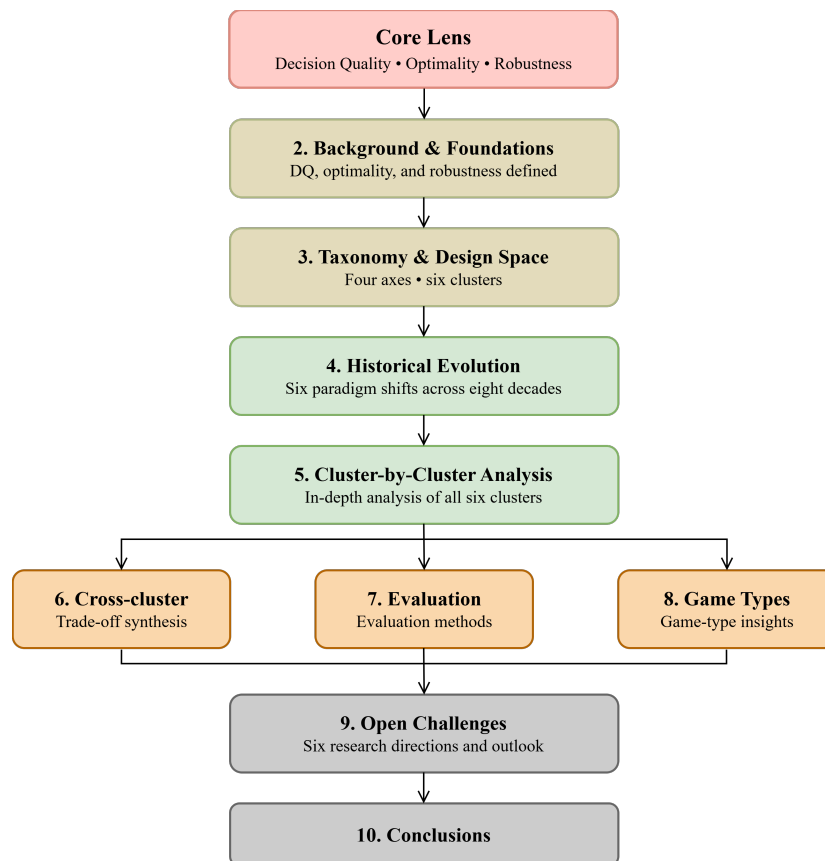


Figure 1. Structural overview of the survey, showing the flow from historical analysis through cluster-by-cluster examination to cross-cluster synthesis.

2. Background and Conceptual Foundations

This section establishes the conceptual foundations required to analyze search algorithms in gaming through the lenses of **decision quality**, **optimality**, and **robustness**. While these notions originate in classical search theory, game theory, and control, their meaning and practical realization in games differ fundamentally from traditional optimization or planning problems. Games introduce strategic interaction, adversarial dynamics, uncertainty, and real-time constraints that reshape how quality, optimality, and robustness are defined, measured, and achieved.

2.1. Decision Quality in Gaming Contexts

Decision quality in games refers to the degree to which an algorithm's chosen actions lead to strong, effective, and strategically coherent gameplay outcomes. Unlike single shot optimization, gaming decisions are embedded in sequential, interactive processes in which local action quality must be evaluated relative to long term consequences and opponent responses. Early work on computer game playing framed decision quality through handcrafted evaluation functions that approximated the desirability of game states, as exemplified by early Chess programs that relied on material balance and positional heuristics to guide search [3]. In these settings, decision quality was closely tied to the fidelity of the evaluation function and the depth of search achievable within computational limits.

As games increased in complexity, explicit state evaluation became increasingly approximate, and decision quality shifted toward empirical performance measures. In large game trees, such as those encountered in Chess and Go, search algorithms began to emphasize relative performance such as winning against strong opponents over provable state value accuracy [6,7]. The emergence of MCTS further reframed decision quality as an outcome of statistical sampling and exploration–exploitation balance, where high quality decisions arise from sufficient simulation coverage and effective selection policies rather than exact evaluation [8,9].

In modern gaming systems, particularly those incorporating learning, decision quality is often assessed through win rates, scores, or tournament-style evaluations rather than through intrinsic measures of move optimality [5,20,22]. In real-time and multi-agent games, decision quality also encompasses responsiveness and stability across different game phases, as delays or inconsistent behavior can degrade perceived gameplay quality even if individual decisions are strong in isolation [11,18,19]. Consequently, decision quality in gaming is best understood as a composite notion that integrates action effectiveness, strategic coherence, and temporal suitability.

2.2. Optimality and Its Variants in Games

Optimality in games traditionally derives from game theoretic concepts such as minimax optimality in zero sum games and equilibrium notions in more general settings. In early board game research, optimality was treated as an explicit objective: given sufficient search depth and accurate evaluation, minimax algorithms could, in principle, converge to optimal play [3]. Alpha-beta pruning and related techniques were introduced to make this objective computationally feasible by reducing the effective branching factor without sacrificing correctness [6].

However, strict optimality quickly became impractical as game complexity grew. Even in deterministic, perfect information games, exponential growth in the game tree limited achievable search depth, motivating best first and fixed depth approximations that traded optimality for efficiency [7]. In this context, optimality was often reinterpreted as *depth limited optimality*-optimal play relative to a truncated horizon and an approximate evaluation function.

Sampling based approaches introduced a further shift. MCTS methods provide asymptotic guarantees: with infinite simulations, action selection converges to the optimal choice under certain assumptions. In practice, however, finite time performance dominates, and optimality becomes a probabilistic and resource dependent concept. Algorithms such as Information Set MCTS extended these ideas to imperfect information games, where exact optimality is often ill defined or computationally unattainable [15].

Learning augmented approaches further blur the boundary between theoretical and practical optimality. Systems that combine search with learned policies, value functions, or dynamics models can achieve levels of play previously associated with optimal or near optimal strategies, despite lacking explicit guarantees of convergence to equilibrium solutions [5,20,22]. In these cases, optimality is operationalized through empirical dominance rather than formal proof, raising questions about the conditions under which such performance reflects genuine strategic optimality versus exploitation of specific training distributions.

2.3. Robustness in Gaming Search Algorithms

Robustness in gaming refers to an algorithm's ability to maintain acceptable performance under uncertainty, variability, and adversarial adaptation. Games present multiple sources of uncertainty, including stochastic state transitions, hidden information, unpredictable opponents, and fluctuating computational resources. Classical search theory explicitly addressed uncertainty through probabilistic formulations of detection and pursuit evasion, where robustness was measured by expected performance over uncertain target locations or movement patterns [12–14]. These formulations emphasized worst-case and expected-case guarantees, laying an early foundation for robust decision-making.

In adversarial and imperfect information games, robustness also encompasses resistance to exploitation by opponents. Algorithms must perform well not only against a fixed strategy but across a diverse set of opponent behaviors. Techniques such as minimax search inherently incorporate a form of robustness by assuming worst case opponent play, but this assumption can be overly conservative in practice [6]. Sampling based and learning based methods often adopt a more empirical notion of robustness, evaluating performance across varied opponents, scenarios, or game variants [5,8,15].

Real-time constraints introduce an additional robustness dimension. Algorithms that perform well when given ample computation may degrade sharply under tight time budgets. Real-time heuristic search and rolling horizon methods explicitly address this challenge by prioritizing anytime behavior

and graceful degradation under reduced computation [18,19]. In complex, partially observable environments such as RTS games, robustness further requires adaptability to incomplete information and dynamic game evolution, motivating hybrid approaches that integrate learning, search, and domain knowledge [17,21].

2.4. Evaluation Metrics as Conceptual Anchors

The notions of decision quality, optimality, and robustness are inseparable from how they are evaluated. Early work relied on analytic measures such as path optimality, detection probability, or proof based correctness [12,23]. As gaming research progressed, empirical metrics became dominant, including win rates, scores, and ratings for competitive games [4,5], as well as computational metrics such as runtime, node expansions, and simulation throughput [8,24]. Robustness oriented evaluations introduced scenario diversity, cross map testing, and performance under uncertainty as key considerations [15–17].

These evolving evaluation practices reflect deeper conceptual shifts: from exact optimality to bounded rationality, from static correctness to dynamic performance, and from single instance success to distributional robustness. Understanding these shifts is essential for interpreting claims about decision quality, optimality, and robustness across different eras and algorithmic families.

2.5. Evaluation Metrics as Conceptual Anchors

While analytically separable, decision quality, optimality, and robustness are tightly coupled in gaming search algorithms. Improvements in decision quality often arise from sacrificing strict optimality in favor of heuristics or learning that perform well under realistic constraints. Conversely, robustness mechanisms such as conservative assumptions or stochastic exploration may reduce peak decision quality in favorable conditions while improving average case performance across diverse scenarios. The history of gaming search algorithms can thus be viewed as a sequence of rebalancing acts among these dimensions, driven by changes in game complexity, computational resources, and methodological innovation.

The remainder of this survey builds on these conceptual foundations. Using a gaming specific taxonomy and design space, it traces how different classes of search algorithms have navigated the trade offs between decision quality, optimality, and robustness, and how these trade offs have evolved from early board games to modern, large scale, and partially observable gaming environments.

3. Gaming Specific Taxonomy and Design Space

A central challenge in surveying search algorithms for games lies in organizing a heterogeneous body of work that spans radically different game genres, interaction models, information structures, and computational regimes. Traditional classifications based solely on algorithmic technique (e.g., minimax, MCTS, evolutionary methods) or chronological development obscure critical distinctions that directly affect **decision quality**, **optimality**, and **robustness** in gaming contexts. This survey therefore adopts a **gaming specific taxonomy and design space** that reflects how search algorithms are actually deployed, evaluated, and evolved in games.

The taxonomy presented in this section is not an abstract or generic framework; it is derived directly from the patterns observed across the papers in the provided literature corpus. It is intentionally multi dimensional, allowing individual methods to be positioned along several axes simultaneously and enabling **multi label classification** when algorithms span multiple conceptual categories. This structure provides the foundation for the historical and cluster based analyses that follow. Figures 2 and 3 show two dimensional projections of the design space — the first along the uncertainty structure and computational regime axes, and the second along the interaction topology and source of search guidance axes.

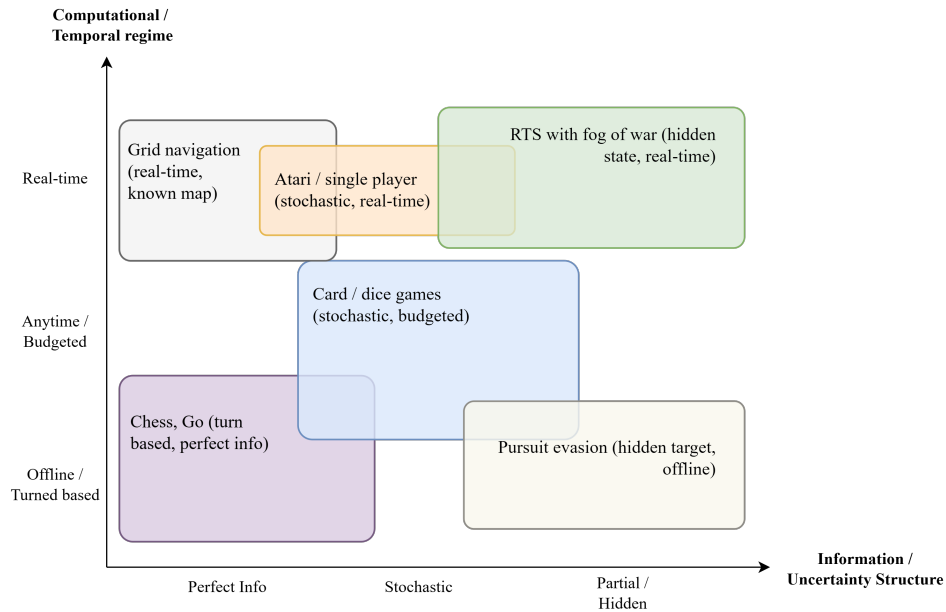


Figure 2. First two dimensional projection of the gaming specific design space, showing **Uncertainty Structure** (horizontal) versus **Computational Regime** (vertical).

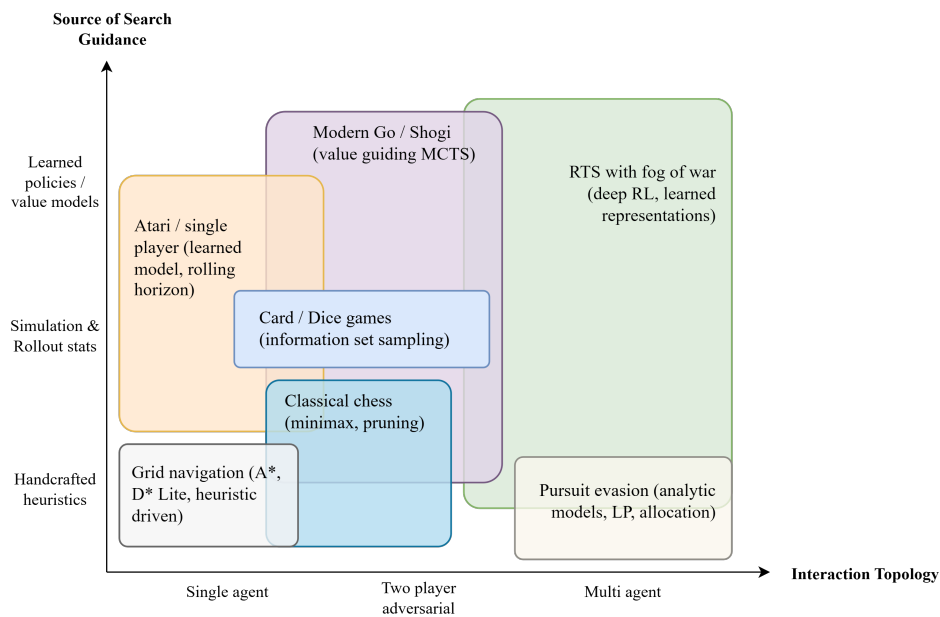


Figure 3. First two dimensional projection of the gaming specific design space, showing **Interaction Topology** (horizontal) versus **Source of Search Guidance** (vertical).

3.1. Design Space Axes for Gaming Search Algorithms

The proposed design space consists of four orthogonal axes that capture the dominant structural factors shaping search behavior and performance in games. Each axis is tightly linked to the challenges of decision quality, optimality, and robustness.

3.1.1. Interaction Topology

The interaction topology axis characterizes **who or what the algorithm is planning against**, fundamentally shaping the search objective and evaluation criteria.

- **Single-agent planning/search:** The algorithm optimizes decisions against a static environment or implicit dynamics, as in pathfinding and navigation tasks [23,25–27]. Decision quality is measured

through solution quality and efficiency, while optimality is often defined in terms of shortest paths or minimal cost.

- **Two-player adversarial interaction:** The algorithm explicitly models an opponent, typically under worst case assumptions, as in Chess, Go, and Shogi [3,6,7]. Optimality is framed through minimax values, and robustness is tied to resistance against strong adversarial play.
- **Multi-agent (cooperative or competitive) interaction:** The algorithm operates in environments with multiple interacting agents, such as RTS games, where coordination, partial observability, and real-time constraints dominate [17,21,28]. Here, decision quality emphasizes collective effectiveness and adaptability, while optimality is usually approximate or implicit.

This axis explains why methods that perform well in board games may fail in RTS settings despite sharing similar search primitives.

3.1.2. Information and Uncertainty Structure

Games differ substantially in how much of the state is observable and how outcomes are determined, making information structure a critical axis.

- **Perfect information deterministic games:** All state variables are observable and transitions are deterministic, as in classical board games and grid-based navigation [3,24,25]. Optimality is well defined, and decision quality can be directly tied to search depth and evaluation accuracy.
- **Stochastic games:** Randomness affects transitions or outcomes, requiring expectation based reasoning and robustness to variance [8,9].
- **Imperfect information and partial observability:** Some state variables are hidden or only indirectly observable, as in pursuit evasion problems and RTS games with fog of war [14,15,17]. Robustness becomes a dominant concern, and exact optimality may be ill-defined or computationally infeasible.

This axis highlights the shift from deterministic guarantees toward probabilistic and empirical performance measures as uncertainty increases.

3.2. Computational and Temporal Regime

Unlike many classical search problems, games impose strict constraints on **when** decisions must be made.

- **Offline or turn based deliberation:** The algorithm can spend substantial time per decision, enabling deeper search and closer approximations to optimal play [6,7].
- **Real-time or strict compute budget regimes:** Decisions must be made within tight time limits, often repeatedly and continuously, as in FPS and RTS games [11,18,19]. Decision quality must be balanced against responsiveness, and robustness includes graceful degradation under reduced computation.
- **Training time plus play time regimes:** Learning augmented systems invest significant computation offline or online during training, then combine fast inference with limited search during gameplay [5,20,22]. Optimality is evaluated empirically rather than analytically, and robustness depends on generalization beyond training conditions.

This axis clarifies why asymptotic guarantees may be less relevant than anytime behavior in many gaming scenarios.

3.3. Source of Search Guidance

The final axis describes **what guides the search** toward promising actions or states.

- **Handcrafted heuristics or analytic models:** Early and classical approaches rely on domain knowledge encoded in evaluation functions or cost heuristics [3,25,29]. Decision quality depends heavily on expert design.

- **Simulation and rollout statistics:** MCTS based methods guide decisions through empirical estimates obtained via repeated simulation, balancing exploration and exploitation [8,9,15].
- **Learned policies, value functions, or models:** Modern systems incorporate learning to guide or replace components of search, achieving high decision quality in large games without explicit guarantees of optimality [5,20–22].

This axis captures a major evolutionary trend in gaming search: the gradual replacement or augmentation of handcrafted knowledge with data driven guidance.

3.4. Practical Clusters of Gaming Search Algorithms

Building on the design space, the literature naturally groups into six practical clusters. These clusters are not mutually exclusive; many influential algorithms span multiple clusters, reflecting hybrid design choices.

3.4.1. Spatial Pathfinding and Navigation Search

This cluster includes algorithms for navigating graphs and spatial environments, such as shortest path search, incremental replanning, and hierarchical abstraction [23–27]. Decision quality is measured by path optimality and responsiveness, optimality often has clear formal definitions, and robustness concerns dynamic environments and real-time constraints.

3.4.2. Adversarial Game Tree Search

This cluster encompasses minimax based approaches and their refinements for two player adversarial games [3,4,6,7]. Decision quality is tied to evaluation accuracy and search depth, optimality is framed through minimax values, and robustness arises from worst case opponent modeling.

3.4.3. Monte Carlo Tree Search and Bandit Planning

MCTS based methods form a distinct cluster characterized by sampling, statistical estimation, and asymptotic convergence properties [8,9]. Extensions to imperfect information games, such as Information Set MCTS, further integrate robustness to hidden state [15]. Decision quality improves with simulation budget, while optimality is asymptotic rather than guaranteed in finite time.

3.4.4. Optimization and Metaheuristic Search (Including Rolling Horizon)

This cluster includes evolutionary algorithms, rolling horizon optimization, and other metaheuristic approaches used as search mechanisms in games [19]. Decision quality is often evaluated empirically, optimality is approximate, and robustness stems from population diversity and stochastic exploration.

3.4.5. Learning Augmented Search and Neural Planning

This cluster captures methods that integrate learning with search, including deep policy and value networks, learned dynamics models, and self play training [5,20–22]. These approaches achieve high decision quality in large games but replace formal optimality guarantees with empirical dominance and introduce new robustness challenges related to generalization.

3.4.6. Uncertainty, Partial Observability, and Search Games

Rooted in classical search theory and extended to modern settings, this cluster focuses on decision making under uncertainty, including pursuit evasion, moving target search, and partial observability [12–14,16,30]. Robustness is central, while optimality is typically probabilistic or expectation based.

Figure 4 illustrates the full gaming specific taxonomy of search algorithms across all six clusters and Table 1 summarizes how each cluster typically operationalizes decision quality, optimality, and robustness in gaming contexts, and serves as a reference throughout the cluster-by-cluster analysis that follows.

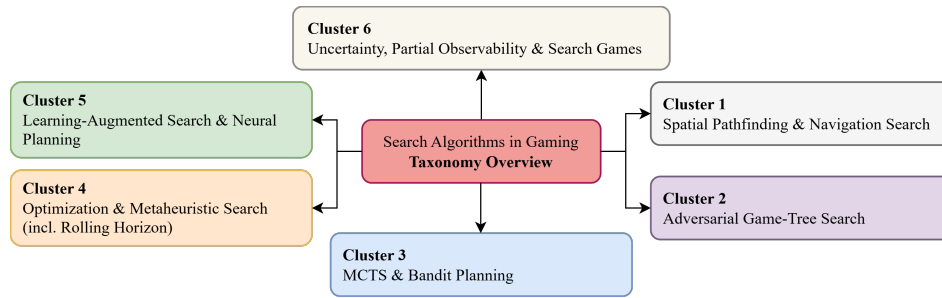


Figure 4. Gaming specific taxonomy of search algorithms used in this survey (six clusters).

Table 1. Survey taxonomy summary: how each cluster typically operationalizes decision quality, optimality, and robustness in gaming contexts

Cluster	Core Idea	Decision Quality	Optimality	Robustness	Typical Game Types
C1: Pathfinding / Navigation	Heuristic search + replanning for spatial movement	Path cost/length; smoothness; responsiveness under time constraints	Exact (admissible heuristics) or bounded/approx via abstraction and replanning	Stability under dynamic maps + compute variability; anytime behavior	Navigation, grid maps, game world traversal
C2: Adversarial Game-Tree	Minimax search, pruning, heuristic evaluation	Move/strategy strength vs opponents; depth limited state evaluation	Minimax correct under pruning; practical depth limited "optimality"	Worst case opponent modeling; conservative safety against exploitation	Board games (Chess-like), perfect info adversarial
C3: MCTS / Bandits	Sampling based planning with exploration exploitation control	Empirical win rate/score; improvement with simulation budget	Asymptotic convergence under assumptions; finite budget approximation	Robustness to stochasticity + variable compute budgets (anytime)	Large branching games; stochastic games; some RTS via abstractions
C4: Metaheuristics / Rolling Horizon	Budgeted optimization over action sequences (e.g., evolutionary RH)	Objective driven performance under tight budgets; diversity of candidate plans	Approximate best so far within budget; horizon limited	Graceful degradation; diversity driven resilience; noise tolerance	Real-time single player control; puzzle/optimization like game settings
C5: Learning Augmented Search	Learned priors (policy/value/model) guide or replace parts of search	High empirical strength at scale; learned strategic structure	Benchmark defined "near-optimality"; guarantees typically replaced by empirical dominance	Generalization and distributional robustness become central (shift, scale, fog of war)	Go/ Chess/ Shogi/ Atari-like; complex RTS and multi agent domains
C6: Uncertainty / Search Games	Belief and model based policies for hidden state and moving targets	Belief conditioned effectiveness; detection/capture probability; expected performance	Model relative optimality (LP/DP/control); equilibrium under incomplete info	Primary objective: robust performance under uncertainty, partial observability, adversarial hiding	Pursuit evasion; moving target search; fog of war-like uncertainty settings

3.5. Multi Label Classification and Hybridization

A defining feature of the gaming literature is the prevalence of hybrid approaches that draw from multiple clusters. For example, modern Go-playing systems combine adversarial game tree search, MCTS, and learning augmented guidance [5,22], while RTS systems integrate real-time constraints, partial observability, and learned representations [17,21,28]. The taxonomy therefore permits multi label classification, enabling a single paper to be analyzed through multiple conceptual lenses without forcing artificial boundaries.

3.6. Role of the Taxonomy in the Survey

This gaming specific taxonomy and design space serve three purposes in the remainder of the survey. First, they provide a systematic organizational framework that supports comprehensive coverage without fragmentation. Second, they enable a structured analysis of how decision quality, optimality, and robustness are addressed differently across game types and algorithm families. Third, they make explicit the evolutionary pathways through which ideas migrate between clusters, leading to increasingly integrated and hybrid approaches.

The subsequent sections leverage this taxonomy to trace the historical evolution of gaming search algorithms and to conduct a detailed, cluster by cluster analysis of state of the art methods and their trade offs.

4. Historical Evolution of Decision Quality, Optimality, and Robustness

The evolution of search algorithms in gaming is inseparable from the evolving interpretation of **decision quality**, **optimality**, and **robustness**. Across successive historical phases, these concepts have been repeatedly redefined in response to new game domains, computational resources, and modeling assumptions. Rather than progressing linearly, the field has undergone a series of paradigm shifts in which one dimension is emphasized at the expense of others, only to be later rebalanced through new algorithmic ideas. This section traces that evolution chronologically, grounding each phase in the gaming literature surveyed and highlighting how the three dimensions were understood and operationalized at each stage.

4.1. Analytic Foundations and Early Search Theory (1940s–1960s)

The earliest roots of gaming related search lie not in entertainment games but in analytic models of search, detection, and pursuit evasion. In this foundational period, **decision quality** was defined explicitly and quantitatively, often as probability of detection or expected time to capture, derived from probabilistic models of uncertainty [12,13]. Decisions were evaluated relative to mathematically specified objectives, leaving little ambiguity about what constituted a “good” action.

Optimality during this era was central and formal. Optimal strategies were derived analytically under clearly stated assumptions about sensing, target motion, and resource constraints [13]. These results established that optimality could be rigorously defined and achieved provided the environment and uncertainty models were sufficiently simple and well specified. Importantly, optimality was model-relative: it held only insofar as the assumed probabilistic structure accurately reflected reality.

Robustness was implicitly embedded in these formulations through expectation and worst case reasoning over uncertainty distributions. Rather than treating uncertainty as noise, early search theory treated it as the defining feature of the problem, anticipating later robustness centric perspectives in gaming AI [12]. This period thus established a conceptual template in which robustness and optimality coexist through probabilistic modeling, albeit under strong assumptions.

4.2. The Emergence of Heuristic and Adversarial Search (1970s–1990s)

The rise of computer game playing, particularly in Chess, marked a decisive shift in both domain and methodology. In adversarial, perfect information games, **decision quality** became relative rather

than absolute: a decision was good if it led to favorable outcomes against a rational opponent. Minimax search formalized this notion by embedding opponent modeling directly into decision-making [3].

Optimality in this period was sharply defined through minimax correctness. Theoretical advances such as alpha-beta pruning demonstrated that large portions of the game tree could be ignored without compromising optimality, provided the pruning rules preserved minimax equivalence [6]. Subsequent enhancements, including history based move ordering and best first variants, improved efficiency while maintaining correctness guarantees [7,29]. Optimality thus remained the organizing principle of algorithm design, with efficiency treated as a secondary but increasingly important concern.

Robustness was conceptualized primarily as resistance to adversarial exploitation. By assuming worst case opponent behavior, minimax search offered a strong form of robustness against strategic manipulation [3]. However, this robustness was narrow: it applied only under perfect information and deterministic dynamics. The limitations of this framing would become increasingly apparent as games grew more complex and less structured.

In parallel, heuristic search for navigation and planning matured during this era, with algorithms such as A* establishing **decision quality** as path optimality under admissible heuristics [25]. Here again, optimality was achievable and provable, reinforcing the prevailing belief that strong theoretical guarantees should be the primary goal of search algorithm design.

4.3. *Scaling, Engineering, and Practical Dominance (1990s–early 2000s)*

As computational resources increased, research attention shifted toward scaling search methods to larger problem instances. In board games, this culminated in highly engineered systems that combined deep adversarial search, extensive domain knowledge, and specialized hardware. The Deep Blue system exemplifies this phase, achieving world class decision quality through massive search depth and carefully tuned evaluation, without fundamentally altering the minimax paradigm [4].

During this phase, **decision quality** was increasingly assessed empirically, through competitive performance against strong opponents. While minimax optimality remained the theoretical foundation, practical success depended on engineering choices that went beyond formal analysis. This marked an early departure from purely proof driven evaluation.

In navigation and planning domains, **robustness** to environmental change and computational constraints gained prominence. Real-time heuristic search explicitly prioritized responsiveness, ensuring that acceptable decisions could be produced under strict time limits [18]. Incremental and replanning algorithms addressed dynamic environments by trading strict optimality for adaptability [26,31]. These developments signaled a growing recognition that robustness and real-time feasibility could outweigh exact optimality in practical gaming contexts.

4.4. *Sampling Based Planning and Statistical Decision Quality (mid-2000s–2010s)*

The introduction of Monte Carlo Tree Search represented a fundamental rethinking of how decision quality and optimality should be pursued in games with large branching factors. Rather than attempting exhaustive or pruned enumeration, MCTS framed decision making as a problem of **statistical estimation through sampling** [8,9,32].

In this paradigm, **decision quality** is inherently empirical and budget dependent: actions are selected based on observed outcomes of simulations, and quality improves gradually as more samples are collected. This replaced the deterministic link between search depth and decision quality with a probabilistic relationship between simulation budget and expected performance.

Optimality was reinterpreted as an asymptotic property. Algorithms such as UCT are provably convergent to minimax optimal actions in the limit of infinite simulations, but offer no finite time guarantees [9]. This reframing made optimality a long term aspiration rather than an immediate design constraint, enabling practical success in games previously considered intractable.

Robustness expanded in scope during this period. Sampling based methods naturally tolerate stochastic dynamics and variable computation budgets, degrading gracefully when resources are limited. Extensions to imperfect information games further emphasized robustness under hidden state,

even as theoretical guarantees weakened [15]. The application of these ideas to multiplayer imperfect information settings was most dramatically demonstrated by Pluribus, the first superhuman AI for six player Texas Hold'em Poker [33]. This era marked a decisive shift toward robustness as a practical, empirical property rather than a formally guaranteed one.

4.5. *Learning Augmented Search and Empirical Optimality (2010s–Present)*

The integration of machine learning with search represents the most transformative phase in the historical trajectory. Neural networks trained through supervised learning and self play began to serve as powerful priors for guiding search, dramatically improving **decision quality** in complex games such as Go [5]. Subsequent systems eliminated reliance on human data altogether, demonstrating that self play could generate strategic knowledge sufficient for superhuman performance [22]. A single, unified algorithm (AlphaZero) further demonstrated that these principles generalize simultaneously to Chess, Shogi, and Go without game specific tuning [1], establishing self-play and neural-guided MCTS as broadly applicable mechanisms for decision quality.

In this phase, **optimality** is largely operationalized through empirical dominance rather than formal proof. Systems achieve near optimal play by outperforming strong baselines across extensive evaluations, even though their relationship to theoretical equilibria remains implicit [5,22]. Learned model planning extends this approach by replacing explicit simulators with learned dynamics, further decoupling performance from classical notions of correctness [1,20].

Robustness becomes both more critical and more fragile. Learned priors improve robustness to combinatorial explosion and opponent diversity, but introduce new vulnerabilities related to distribution shift, partial observability, and scale variation. Most recently, the assumption that search is necessary for high decision quality has been directly challenged: a large Transformer trained purely on game records achieves grandmaster level chess without any search procedure [34]. These challenges are especially visible in real-time strategy games, where fog of war and multi agent coordination stress both learning and search components [17,21,35].

4.6. *Persistent Uncertainty Centric Traditions*

Running parallel to these mainstream developments is a persistent tradition focused on uncertainty, partial observability, and search games. From moving target search and constrained pursuit evasion problems to modern partial observability game formulations, this line of work consistently prioritizes robustness under hidden state [14,16,30,36–40]. Decision quality is evaluated relative to beliefs, and optimality is defined with respect to uncertainty models or equilibrium concepts under incomplete information.

This tradition provides an important counterpoint to performance driven gaming AI. It emphasizes that robustness cannot be an afterthought and that formal reasoning about uncertainty remains essential, particularly as modern games increasingly resemble partially observable, stochastic environments.

4.7. *Synthesis*

Across these historical phases, the evolution of gaming search algorithms reflects a progressive broadening of priorities. Early work emphasized exact optimality and analytic decision quality under strong assumptions. Subsequent phases progressively relaxed these assumptions in favor of scalability, empirical performance, and robustness under real world constraints. The modern landscape is characterized by hybrid systems that blend search, sampling, optimization, and learning, each contributing different strengths to the enduring challenge of balancing decision quality, optimality, and robustness.

Understanding this historical evolution is essential for interpreting current approaches and for identifying future directions. The tensions and trade offs observed today are not anomalies but the latest expressions of a long standing dialogue between theory and practice in gaming search. Figure 5 illustrates the high level evolution of dominant paradigms across the surveyed corpus over time.

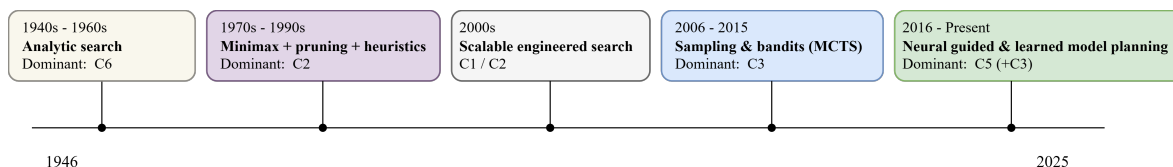


Figure 5. High level evolution of gaming search paradigms in the surveyed corpus, emphasizing shifts in dominant clusters over time.

5. Cluster by Cluster Deep Analysis

This section constitutes the analytical core of the survey. Each cluster is examined in depth with respect to how it addresses **decision quality**, **optimality**, and **robustness** in gaming contexts, tracing internal evolution, highlighting state of the art practices, and critically evaluating trade offs. Consistent with the taxonomy established in Section 3, clusters are analyzed independently while acknowledging cross cluster interactions where relevant.

5.1. Cluster 1: Spatial Pathfinding and Navigation Search

Spatial pathfinding and navigation search form one of the earliest and most enduring clusters in gaming AI. This cluster encompasses algorithms designed to compute action sequences that move an agent through a spatial environment efficiently, typically represented as graphs or grids. Although often framed as single agent problems, pathfinding algorithms have had a profound influence on broader gaming search paradigms by shaping notions of decision quality, optimality, and real-time robustness.

5.1.1. Foundations and Decision Quality

The conceptual foundations of spatial pathfinding are rooted in graph theory and heuristic search. Early work on shortest paths established the baseline notion of decision quality as **cost optimality**, where a high quality decision corresponds to selecting an action that lies on a minimum cost path to a goal [23]. The introduction of heuristic search formalized this notion further by allowing informed estimates of remaining cost to guide exploration, thereby improving decision quality under limited computation [25].

In gaming contexts, decision quality in pathfinding is typically evaluated using metrics such as path length, path cost, and responsiveness. Unlike adversarial games, where decision quality is relative to an opponent, pathfinding quality is intrinsic to the environment and task specification. However, gaming environments introduce additional constraints, including dynamic obstacles, frequent replanning, and strict per frame decision deadlines. As a result, decision quality extends beyond optimal path cost to include smoothness of motion, avoidance of erratic behavior, and timely reaction to environmental changes [10,11,41].

5.1.2. Optimality: From Exact Guarantees to Approximate Solutions

Optimality has historically been a defining feature of this cluster. Algorithms such as A* provide formal guarantees of optimality under admissible heuristics, making them attractive for games that demand predictable and explainable behavior [25]. These guarantees were further refined through analyses of memory and time trade offs, such as depth first iterative deepening approaches that preserve optimality while reducing space requirements [10].

As game environments grew larger and more complex, exact optimality became increasingly expensive. Incremental and dynamic replanning algorithms, such as D* Lite, relaxed strict optimality in favor of efficiency, enabling rapid updates to paths when the environment changes while maintaining near optimal solutions [26]. Hierarchical approaches, notably HPA*, introduced abstraction to reduce problem size, explicitly trading solution optimality for scalability [27]. In these methods, optimality is preserved only within the abstracted representation, and decision quality is evaluated in terms of acceptable approximation error rather than exact minimality.

Empirical studies comparing pathfinding algorithms in gaming environments further reinforced this shift, demonstrating that slightly suboptimal paths can yield superior gameplay performance when computation time and responsiveness are considered [41,42]. Thus, within this cluster, optimality evolved from a strict theoretical objective to a flexible, context dependent criterion.

5.1.3. Robustness Under Dynamic and Real-Time Constraints

Robustness in spatial pathfinding arises primarily from environmental dynamics and computational constraints rather than from adversarial uncertainty. Games frequently require agents to navigate partially known or changing environments, necessitating algorithms that can adapt without catastrophic performance degradation. Early formulations addressed robustness implicitly through replanning strategies that respond to newly discovered obstacles or changes in terrain [31].

Real-time heuristic search marked a significant conceptual shift by explicitly prioritizing robustness to time constraints [18]. These methods emphasize anytime behavior, ensuring that a valid, if suboptimal, action can always be selected within a fixed computation budget. In commercial game AI, such as the navigation systems described in the F.E.A.R. architecture, robustness is further enhanced through tight integration with game engines, allowing navigation decisions to be synchronized with animation, physics, and perception systems [11].

Pruning and acceleration techniques, including online graph pruning and Jump Point Search, contribute to robustness by reducing sensitivity to map size and branching factor, enabling consistent performance across diverse levels and scenarios [24,43]. Collectively, these developments reflect a shift from robustness as an afterthought to robustness as a first class design objective.

5.1.4. Evolution Within the Cluster

The internal evolution of spatial pathfinding mirrors broader trends in gaming search. Early work prioritized **formal optimality and correctness**, grounded in static problem formulations [23,25]. The 1990s introduced **adaptation and replanning** as environments became more dynamic [18,31]. The early 2000s emphasized **scalability and abstraction**, responding to the growing size of game worlds [26,27]. More recent work has focused on **engineering robustness and empirical efficiency**, often through algorithmic hybrids and systematic evaluation [24,41–45].

The most recent evolution within this cluster involves the integration of large language models with classical pathfinding algorithms. LLM-A* combines LLM global reasoning which provides semantically informed waypoints with the precise local search of A*, reducing both time and memory complexity in path planning tasks [45]. This hybridization represents an emerging bridge between the **Cluster 1** and **Cluster 5** paradigms, introducing data driven guidance into a traditionally heuristic dominated domain.

This evolution demonstrates a gradual rebalancing of decision quality, optimality, and robustness. While optimality remains conceptually important, practical gaming systems increasingly favor robust, responsive behavior that maintains acceptable decision quality under real world constraints.

5.1.5. Comparative Insights and Trade offs

Compared to other clusters, spatial pathfinding occupies a unique position. It offers some of the clearest formal guarantees of optimality, yet it also provides early examples of how such guarantees must be relaxed in practice. Unlike adversarial game tree search, robustness here is not primarily about opponent modeling but about environmental uncertainty and computational variability. Compared to learning augmented approaches, pathfinding relies far less on data and generalization, resulting in more predictable but less adaptive behavior.

The principal trade off in this cluster lies between **solution optimality and real-time robustness**. Hierarchical and pruning based methods improve scalability and responsiveness at the cost of path optimality, while exact methods risk infeasibility in large or dynamic environments. These trade offs foreshadow similar tensions that later emerge in MCTS based and learning based clusters, where guarantees are sacrificed to achieve performance at scale.

5.2. Cluster 2: Adversarial Game Tree Search

Adversarial game tree search represents the classical core of game playing artificial intelligence and provides the most explicit historical grounding for the concepts of decision quality, optimality, and robustness in gaming. This cluster encompasses minimax based search algorithms and their refinements for two player, zero sum, perfect information games, most notably Chess and related board games. Although later clusters introduce sampling, learning, and uncertainty handling, adversarial game tree search established the formal language and evaluation criteria that continue to influence gaming AI.

5.2.1. Foundations and Decision Quality in Adversarial Settings

The defining characteristic of adversarial game tree search is the explicit modeling of an intelligent opponent. Decision quality in this cluster is fundamentally relative: a decision is high quality if it leads to favorable outcomes under optimal or near optimal opponent play. Early work on programming computers to play Chess formalized this idea by framing gameplay as a recursive minimax optimization problem over a game tree [3]. In this formulation, decision quality is inseparable from the accuracy of the evaluation function and the depth to which the tree can be searched.

Handcrafted evaluation functions played a central role in shaping decision quality. These functions encoded expert knowledge about material balance, positional features, and strategic considerations, serving as proxies for long term outcomes beyond the search horizon [3]. As search depth increased, decision quality improved monotonically in theory, motivating intense focus on techniques that could push the depth boundary further within fixed computational budgets. This tight coupling between decision quality and search depth distinguishes adversarial game tree search from later sampling based approaches, where quality emerges statistically rather than deterministically.

5.2.2. Optimality and the Minimax Paradigm

Optimality is most explicitly defined and theoretically grounded in adversarial game tree search. In zero sum, perfect information games, minimax search computes strategies that are optimal in the game theoretic sense, assuming both players act optimally. Alpha–beta pruning was a critical breakthrough in making this objective computationally feasible, demonstrating that large portions of the game tree could be pruned without affecting the correctness of the minimax result [6]. This result established that optimality could be preserved even under aggressive pruning, provided that pruning rules respected the minimax structure.

Subsequent refinements focused on improving the practical efficiency of optimal search. Best first and fixed depth minimax variants explored different traversal orders and selection strategies to concentrate computational effort on the most promising parts of the tree [7]. The history heuristic further enhanced alpha–beta performance by learning which moves were historically effective at causing cutoffs, thereby increasing effective search depth without compromising optimality [29]. Collectively, these techniques illustrate a central theme of this cluster: **optimality is preserved by design, while efficiency is optimized through search control.**

However, the reliance on fixed depth search and heuristic evaluation functions introduces a subtle limitation. While the minimax value is optimal relative to the truncated horizon and evaluation function, it may deviate significantly from true optimal play in deep or complex positions. Thus, even within this cluster, practical optimality is conditional rather than absolute, constrained by available computation and heuristic fidelity.

5.2.3. Robustness Through Worst Case Reasoning

Robustness in adversarial game tree search is primarily achieved through **worst case opponent modeling**. By assuming that the opponent will always choose actions that minimize the algorithm's payoff, minimax search inherently guards against exploitation by adversarial strategies [3]. This pessimistic assumption provides a strong form of robustness: if an algorithm performs well under worst case play, it will perform at least as well against weaker or suboptimal opponents.

Alpha–beta pruning and its enhancements preserve this robustness, as pruning decisions do not alter the minimax outcome [6]. However, robustness in this sense is limited to adversarial optimality and does not extend to other forms of uncertainty. Classical game tree search typically assumes perfect information and deterministic transitions, making it brittle in the presence of stochasticity, hidden information, or unexpected rule variations. Moreover, worst case reasoning can lead to overly conservative play, sacrificing decision quality against non optimal opponents in favor of safety.

5.2.4. Scaling Limits and the Transition to Engineering Solutions

The scaling limits of adversarial game tree search became increasingly apparent as game complexity grew. Chess, while complex, has a relatively structured state space that allowed extensive engineering and hardware acceleration to push minimax based methods to unprecedented levels. The Deep Blue system exemplifies this trajectory, combining massive parallel computation, domain specific heuristics, and deep alpha–beta search to achieve world class performance [4]. Importantly, Deep Blue did not fundamentally alter the minimax paradigm; instead, it demonstrated how far decision quality and practical optimality could be pushed through engineering within this cluster.

AlphaZero subsequently rendered purely engineering driven minimax approaches obsolete in board game domains by mastering Chess, Shogi, and Go simultaneously through self play combined with neural guided MCTS [1], signalling the transition away from handcrafted adversarial search as the primary driver of decision quality.

Despite such successes, the limitations of adversarial game tree search became evident in games with larger branching factors, deeper horizons, or less structured evaluation functions. In these domains, the cost of maintaining strict minimax optimality grew prohibitive, motivating the exploration of alternative paradigms that relaxed guarantees in exchange for scalability and robustness.

5.2.5. Evolution Within the Cluster

The internal evolution of adversarial game tree search reflects a gradual shift from **theoretical purity to practical dominance**. Early work emphasized formal correctness and optimality proofs [3,6]. Subsequent research focused on search control heuristics and efficiency improvements that preserved optimality while increasing depth [7,29]. The culmination of this trajectory is exemplified by large scale engineered systems that achieve exceptional empirical decision quality through computational power and domain expertise [4].

Over time, however, the cluster’s influence began to wane as a standalone solution. The difficulty of extending minimax based methods to games with uncertainty, real-time constraints, or massive state spaces exposed fundamental robustness limitations. These pressures set the stage for the emergence of Monte Carlo Tree Search and learning augmented approaches, which reinterpret optimality and robustness in probabilistic and empirical terms [1].

The most striking recent challenge to this cluster’s foundational assumptions comes from systems that achieve grandmaster level performance without any search at all, relying entirely on learned pattern recognition from large scale game data [34]. This development raises fundamental questions about whether game tree search is a necessary component of high quality play or merely one path toward it.

5.2.6. Comparative Insights and Trade offs

Compared to spatial pathfinding, adversarial game tree search offers richer theoretical foundations for optimality but weaker robustness to uncertainty and time constraints. In contrast to MCTS based methods, it provides strong worst case guarantees but lacks graceful degradation under limited computation. The key trade off within this cluster lies between **theoretical optimality and practical applicability**: preserving minimax correctness constrains scalability and adaptability, while relaxing these constraints undermines the very guarantees that define the cluster.

These trade offs have profound implications for decision quality. Against strong adversaries in well structured domains, adversarial game tree search can deliver exceptionally high quality decisions.

In more diverse or uncertain settings, however, its conservative assumptions and computational demands limit both robustness and empirical performance.

5.3. Cluster 3: Monte Carlo Tree Search and Bandit Planning

Monte Carlo Tree Search (MCTS) and bandit based planning mark a fundamental paradigm shift in gaming search algorithms, redefining how **decision quality**, **optimality**, and **robustness** are conceptualized and balanced. Unlike adversarial game tree search, which relies on deterministic evaluation and worst case reasoning, this cluster is characterized by **sampling**, **statistical estimation**, and **anytime behavior**. Its emergence was driven by the need to scale decision making to games with large branching factors, long horizons, and limited domain knowledge, while maintaining acceptable decision quality under strict computational constraints.

5.3.1. Decision Quality Through Sampling and Statistical Estimation

In MCTS based methods, decision quality arises from empirical evidence accumulated through simulation rather than from exhaustive tree expansion or handcrafted evaluation functions. Bandit based Monte Carlo planning formalized this idea by treating action selection as a multi armed bandit problem, explicitly balancing exploration of uncertain actions with exploitation of actions that have performed well in prior simulations [9,32]. The Upper Confidence bounds applied to Trees (UCT) algorithm operationalized this balance, enabling effective decision making even when only a fraction of the game tree can be explored [32].

This statistical framing introduces a fundamentally different notion of decision quality. Instead of guaranteeing correctness relative to a heuristic evaluation, MCTS seeks to minimize empirical regret over simulations. High quality decisions are those supported by sufficient sampling evidence, making decision quality inherently budget dependent. As the simulation budget increases, action value estimates become more reliable, and decision quality improves monotonically in expectation [8,9]. This property makes MCTS particularly attractive for games where deep, accurate evaluation is difficult or infeasible.

In gaming practice, decision quality under MCTS is often measured through win rates or scores rather than intrinsic action optimality, reflecting the empirical nature of the approach [5]. Importantly, MCTS exhibits strong anytime behavior: even with very limited computation, it can provide a reasonable action choice, and additional computation directly translates into improved decision quality.

5.3.2. Optimality: Asymptotic Guarantees and Practical Approximation

Optimality in this cluster is defined asymptotically rather than deterministically. Under standard assumptions, UCT converges to the minimax optimal action as the number of simulations approaches infinity [8,9]. This guarantee contrasts sharply with fixed depth minimax search, where optimality is exact within a truncated horizon but does not necessarily improve beyond it.

In practical gaming settings, however, infinite simulation is unattainable, and optimality becomes a **probabilistic and approximate concept**. The quality of approximation depends on factors such as rollout policy quality, exploration constants, and tree expansion strategies. Research on selective backups and alternative backup operators demonstrated that convergence speed and finite time performance can vary significantly depending on these design choices [8,46].

Extensions to imperfect information games further complicate the notion of optimality. Information Set MCTS replaces state based nodes with information sets, enabling planning under hidden information but sacrificing clear convergence guarantees to equilibrium strategies in general settings [15,33]. In such games, optimality is often ill defined or replaced by empirical dominance against benchmark opponents.

Thus, within this cluster, optimality is best understood as a limit behavior rather than a practical objective. The focus shifts from proving exact optimal play to ensuring steady improvement with increased computation and reasonable performance under realistic budgets.

5.3.3. Robustness to Uncertainty, Opponents, and Computation

Robustness is one of the defining strengths of MCTS based methods. Sampling inherently provides resilience to stochastic transitions and noisy outcomes, as value estimates are based on averages over multiple simulations rather than single deterministic evaluations [9,47]. This property makes MCTS naturally suited to stochastic games and environments with randomized dynamics.

Robustness to opponent variability is achieved through repeated simulation against diverse action sequences. Unlike minimax search, which assumes a perfectly adversarial opponent, MCTS implicitly adapts to the observed distribution of opponent behavior during simulations. While this can reduce worst case guarantees, it often improves empirical robustness against a wide range of opponents in practice.

Computational robustness is equally important. MCTS degrades gracefully under reduced time budgets, maintaining usable decision quality even with few simulations. This anytime property enables deployment in real-time gaming contexts and under variable computational resources. Recent work on parallel and GPU accelerated MCTS further enhances robustness by increasing simulation throughput and reducing sensitivity to hardware limitations [48,49].

In imperfect information settings, robustness becomes more nuanced. Information Set MCTS provides a practical mechanism for handling hidden information, but it can suffer from strategy fusion and other pathologies that undermine robustness in certain games [15]. These limitations highlight that robustness in this cluster is empirical rather than guaranteed, contingent on careful algorithm design and evaluation.

5.3.4. Evolution Within the Cluster

The evolution of Monte Carlo Tree Search reflects a steady broadening of scope and ambition. Early work focused on establishing the bandit based framework and demonstrating its effectiveness in large, deterministic games [8,9]. A comprehensive survey in 2012 systematically cataloged enhancements, extensions, and applications across game types [32], and a more recent review covers subsequent modifications and hybrid approaches [49]. Subsequent research refined selection, expansion, and backup strategies to improve finite time performance and decision quality, including Boltzmann based exploration mechanisms that address limitations of standard UCT [46] and novelty based selection that integrates online generalization of uncertainty during search [47].

The extension to imperfect information games marked a significant expansion of applicability, bringing robustness to hidden state at the cost of theoretical clarity [15]. Later developments emphasized practical integration into complex game genres, such as RTS games, where MCTS was guided by domain scripts to manage enormous branching factors and real-time constraints [28].

Most recently, this cluster has increasingly intersected with learning based methods. Neural guided MCTS systems leverage learned policies and value estimates to bias sampling toward promising regions of the search space, dramatically improving decision quality under fixed budgets [5]. MCTS has also been applied as a non expert demonstration source within distributed deep RL training, providing search derived action guidance that accelerates policy learning [50]. Comparative analysis in multi agent environments has examined when MCTS outperforms or under performs rolling horizon evolutionary approaches, revealing behavioral differences that shape robustness under competitive pressure [51]. Parallelization and hardware acceleration represent another evolutionary trajectory [48].

5.3.5. Comparative Insights and Trade offs

Compared to adversarial game tree search, MCTS based methods sacrifice strict worst case optimality guarantees in exchange for scalability, flexibility, and empirical robustness. Decision quality improves smoothly with additional computation rather than discretely with increased depth. Compared to optimization and metaheuristic approaches, MCTS provides clearer convergence behavior and stronger theoretical grounding, albeit still asymptotic.

The primary trade offs within this cluster involve exploration versus exploitation, finite time performance versus asymptotic optimality, and robustness versus worst case guarantees. Aggressive exploration improves robustness to uncertainty but can degrade short term decision quality, while overly exploitative strategies risk premature convergence to suboptimal actions. These trade offs become especially salient in real-time and multi agent games, where simulation budgets are severely constrained.

5.4. Cluster 4: Optimization and Metaheuristic Search (Including Rolling Horizon)

Optimization and metaheuristic search methods occupy a distinctive position in gaming AI: they typically forgo strong formal guarantees in favor of **flexible objective formulations**, **stochastic exploration**, and **robust empirical performance under tight computational budgets**. Within the established taxonomy, this cluster includes rolling horizon evolutionary approaches used directly for gameplay decision making, as well as optimization centric formulations that shape how “good decisions” are defined and measured in game like search settings. This cluster is particularly relevant to the survey theme because it operationalizes **decision quality** as an optimization objective, treats **optimality** as approximate or bounded by practical constraints, and often pursues **robustness** through diversity, stochasticity, and resilience to local minima.

5.4.1. Decision Quality as Objective Driven Search

A central defining trait of metaheuristic approaches is that decision quality is explicitly tied to a chosen objective function, often optimized over a finite horizon. Rather than proving that an algorithm returns the optimal decision under a formal model, these methods treat decision making as a black box optimization problem: candidate action sequences are generated, evaluated via simulation or heuristic scoring, and iteratively improved.

The rolling horizon evolutionary paradigm illustrates this most clearly. In real-time, single player game navigation settings, rolling horizon evolutionary search was positioned as a direct competitor to tree search, demonstrating that competitive decision quality can be achieved by optimizing short horizon action sequences under strict time constraints [19,52]. Subsequent work extended this paradigm to a broad range of general video game environments, demonstrating state of the art performance across twenty diverse game types with parameter optimization tuned per game category [52]. Here, decision quality is measured empirically by in-game performance, and algorithmic success depends on the ability to rapidly identify high performing action sequences, often under non smooth, delayed reward structures. This emphasis aligns strongly with game development practice: the “best” decision is the one that produces strong gameplay outcomes under real runtime constraints, even when its relation to theoretical optimality is unclear.

Decision quality in this cluster also extends beyond immediate performance to include **strategic diversity** and **behavioral variety**, which can be crucial in games that reward unpredictability or multi-modal solutions. Although such metrics are not uniformly standardized across the corpus, the design patterns in this cluster commonly exploit population based search and stochastic operators precisely because they can produce a portfolio of plausible strategies rather than converging narrowly on a single line of play.

5.4.2. Optimality: Approximation, Practicality, and the Limits of Guarantees

Optimality in optimization and metaheuristic search is typically empirical: the objective is to find a “good” solution quickly rather than to guarantee global optimality. This does not imply that optimality is irrelevant; rather, it is reframed as a practical, resource bounded target. Rolling horizon approaches, for example, optimize over a finite horizon and execute only the first action before replanning, implicitly trading long horizon optimality for responsiveness and tractability [19,52]. In such settings, global optimality is neither achievable nor necessarily desirable, because the environment may evolve, the game state may change unpredictably, and computational budgets are fixed.

This practical stance toward optimality also appears in optimization style formulations of search problems outside classical gameplay, such as constrained moving target search, allocation games, and optimization based adversarial formulations. While some of these appear more directly in the uncertainty/search games cluster, their presence in the overall corpus emphasizes a recurring methodological theme: the use of optimization to define and pursue high quality decisions when exact optimality is difficult to characterize [12,13]. Within the metaheuristic cluster specifically, optimality tends to be treated as “best so far within budget,” often evaluated through comparative baselines rather than bound satisfaction.

A key conceptual consequence is that this cluster naturally motivates anytime evaluation culture: decision quality is not a single number but a function of time. Even when not explicitly plotted, rolling horizon methods inherently produce quality trajectories over time budgets, making them directly comparable to real-time tree search and MCTS variants under similar constraints [19].

5.4.3. Robustness Through Stochasticity, Diversity, and Graceful Degradation

Robustness in metaheuristic search is commonly achieved through three mechanisms:

1. **Stochastic exploration:** Random mutation, recombination, and probabilistic acceptance mechanisms reduce sensitivity to local minima and noisy evaluations, which are prevalent in simulation driven gaming settings.
2. **Population diversity:** Maintaining multiple candidate solutions enhances robustness to deceptive reward landscapes and supports contingency planning when the environment changes or when evaluation is uncertain.
3. **Anytime behavior and re planning:** Rolling horizon execution provides robustness to state changes by continuously re optimizing as the game evolves [19,51].

The comparison between rolling horizon evolutionary search and tree search in real-time single player games demonstrates that robustness is not merely about handling stochastic environments; it also encompasses resilience to compute variability and the ability to sustain acceptable performance under tight decision time regimes [19]. This aligns strongly with real world game AI requirements, where decision routines must be reliable under fluctuating frame budgets and varying hardware.

Robustness also appears in this cluster in the form of measurement robustness—the recognition that evaluation metrics must capture variability rather than average case performance. Later work on puzzle difficulty measurement through generalized entropy and solution information, while not a gameplay search algorithm per se, is directly relevant to robustness thinking: it formalizes how problem difficulty (and thus decision quality assessment) can vary across instances and structural configurations [53]. Such work supports the broader survey narrative that robustness increasingly demands distributional evaluation rather than single instance reporting.

5.4.4. Evolution Within the Cluster

The evolutionary trajectory of optimization and metaheuristic methods in gaming search reflects a gradual shift toward **real-time applicability** and **comparative evaluation against mainstream search methods**. While early optimization centric thinking in the corpus appears in foundational search theory that formalizes optimal allocation of effort under uncertainty [12,13], the modern gaming oriented expression of this cluster is exemplified by rolling horizon evolutionary decision making positioned explicitly against tree search in real-time game contexts [19], extended to general video game playing with systematic parameter optimization [52], and critically compared with MCTS in multi agent settings [51].

More recent work extends the cluster’s relevance beyond action selection into evaluation and problem characterization, such as quantifying puzzle difficulty using information theoretic measures [53]. This expansion reflects an important trend: optimization and metaheuristic approaches are not only decision making tools but also instruments for defining and measuring decision quality itself, especially in game domains where “quality” is multi dimensional and context sensitive.

5.4.5. Comparative Insights and Trade offs

Relative to adversarial game tree search and MCTS, metaheuristic search methods offer a distinct trade off profile:

- **Decision quality vs. interpretability:** Metaheuristics can find strong decisions without explicit modeling assumptions, but the resulting strategies may be harder to interpret or justify compared to heuristic minimax or A* style methods.
- **Optimality vs. flexibility:** These methods rarely offer formal optimality guarantees; however, they adapt easily to novel objective functions, constraints, and game mechanics—an important advantage in diverse game genres.
- **Robustness vs. sample efficiency:** Population based and stochastic exploration can be robust to noise and deceptive landscapes, but may require many evaluations to reach high quality decisions, making compute efficiency a critical limitation under tight real-time budgets.

In real-time settings, rolling horizon evolutionary search exposes a particularly important trade off: short horizon optimization can yield strong immediate decisions yet remain vulnerable to long term strategic traps if the evaluation function or simulation horizon fails to capture delayed consequences [19,51,52]. This is a core tension between decision quality and robustness: optimizing what is measurable in the short term can degrade robustness in long term strategic contexts.

Compared to learning augmented approaches, metaheuristics are less dependent on training data and generalization assumptions, which can improve robustness in domains with sparse or shifting distributions. However, they may struggle to match the peak decision quality of neural -guided systems in highly structured, data rich game domains.

5.5. Cluster 5: Learning Augmented Search and Neural Planning

Learning augmented search and neural planning constitute the most consequential modern development in gaming search algorithms, fundamentally reshaping how **decision quality**, **optimality**, and **robustness** are pursued and evaluated. Within the established taxonomy, this cluster includes approaches that use learned **policies**, **value functions**, **representations**, and/or **environment models** to guide, accelerate, or partially replace classical search. The defining methodological shift is that search performance is no longer driven primarily by handcrafted heuristics or purely statistical rollouts, but by **data driven inductive bias** learned from self play, interaction, or supervised signals.

This cluster is also where the trade off frontier among the three survey dimensions becomes most explicit: learning can dramatically elevate decision quality at scale, yet it typically replaces formal optimality guarantees with empirical dominance and introduces new robustness concerns related to generalization, distribution shift, and model error.

5.5.1. Decision Quality: From Handcrafted Evaluation to Learned Priors

A central contribution of learning augmented approaches is the transformation of decision quality from “search depth + heuristic accuracy” into “search guided by learned priors.” The canonical pattern is the integration of neural policy and value estimation with search, where learned components focus computation on promising regions of the decision space. The Go system that couples deep neural networks with tree search exemplifies this shift: decision quality emerges from a synergy between (i) a learned policy that narrows the effective branching factor and (ii) a learned value estimate that reduces reliance on long rollouts [1,5]. In this setting, high quality decisions are not primarily a function of brute force depth but of where computation is allocated.

A second, equally important shift is the replacement of human knowledge with self generated experience. The system that mastered Go without human knowledge demonstrates that decision quality can be obtained through self play training that iteratively improves policy and value estimates, effectively internalizing strategic structure that would be difficult to encode manually [22,54]. This development changes the meaning of “decision quality” in games: quality becomes tightly coupled to training processes and the diversity and intensity of self play rather than to a static evaluation function.

A third step in this evolution is the move from learned evaluation to learned dynamics models used directly for planning. Planning with a learned model enables decision quality improvements even in environments where rollouts are expensive or where accurate simulation is not readily available, by allowing search to operate within a learned latent dynamics space [20,55]. This expands the applicability of learning augmented search beyond classical board games to broader classes of environments, including those characterized by complex observations and reward structures.

In modern complex game types, particularly RTS environments, decision quality requires learned representations of spatial and strategic structure. Work on map feature learning and pathfinding in partially observable RTS settings illustrates how representation learning is increasingly central to decision quality, supporting both navigation and strategic decision making under incomplete information [17]. Similarly, work targeting scale flexibility in RTS shows that decision quality depends not only on raw policy strength but also on the ability of learned systems to maintain performance across different map sizes, unit counts, or scenario scales [21]. Survey level analyses of deep reinforcement learning in RTS further reinforce that decision quality in this domain is inseparable from representation choices, multi agent coordination, and environment complexity [35].

5.5.2. Optimality: Empirical Dominance Versus Formal Guarantees

Learning augmented gaming systems challenge classical notions of optimality. In adversarial, perfect information domains, optimality is traditionally framed via minimax values and correctness preserving pruning [3,6]. In contrast, neural guided search methods generally do not provide formal proofs of optimal play. Instead, optimality is operationalized through empirical evidence of dominance e.g., consistently defeating strong baselines or achieving superior performance on standardized tasks.

Neural guided Go systems demonstrate that search guided by learned priors can achieve performance indistinguishable from “near optimal” play in practice, yet the method’s guarantees remain largely empirical [5,56]. Self play systems further complicate optimality by producing strategies that are extremely strong but whose relationship to theoretical equilibria is typically not formally certified [22]. Learned model planning continues this trend: planning in a learned dynamics space can deliver remarkable performance, but model error introduces an additional layer of approximation that obscures any direct connection to exact optimality [20].

Optimality in this cluster is therefore best understood using conditional and approximate interpretations:

- **Conditional optimality relative to learned components:** Given a learned policy/value/model, the search procedure may be optimal with respect to the induced surrogate objective, but not necessarily optimal in the underlying game.
- **Asymptotic or limit optimality replaced by scaling laws:** Improvements are often tied to compute scaling (more training, larger models, deeper search), rather than to convergence proofs.
- **Empirical optimality as a benchmark construct:** “Optimality” is effectively defined by outperforming contemporary baselines under agreed evaluation regimes [1,2,5,20,22,34].

This does not imply that optimality is irrelevant; rather, the cluster reflects a historical redefinition of what the community accepts as “optimal enough” in games too large for exact solution.

5.5.3. Robustness: Generalization, Distribution Shift, and Uncertainty

Robustness is both a major motivation and a major vulnerability of learning augmented search. Learned priors can improve robustness to combinatorial explosion by guiding computation toward strategically relevant regions, reducing dependence on brittle handcrafted heuristics [5]. Self play can also enhance robustness against diverse opponents by exposing the learner to an evolving distribution of strategies, reducing over specialization to a fixed opponent class [22].

Learning introduces new failure modes that classical search often avoids:

- **Distribution shift and overfitting:** A learned policy or model can be highly effective in distribution yet degrade under scenario changes, rule variations, or different opponent styles.

- **Model bias and compounding error:** Learned model planning can suffer when model inaccuracies systematically distort search, potentially producing confident but incorrect decisions [20].
- **Partial observability sensitivity:** In RTS environments with fog of war, robustness requires maintaining performance under hidden information and incomplete state estimation. Map feature learning in partially observable RTS settings highlights this challenge directly, indicating that robust decision making depends on representations that remain informative under occlusion and uncertainty [17].
- **Scale brittleness:** Robustness in modern RTS settings includes scaling across map sizes and scenario configurations. Approaches explicitly focused on scale flexibility underscore that robustness is not merely about noise tolerance but about maintaining competence under structural changes in game instances [21]. Survey analyses of RTS deep reinforcement learning emphasize that evaluation practices vary widely and that robustness claims often depend heavily on experimental protocol and environment configuration [35].

Robustness in this cluster increasingly demands distributional evaluation (across maps, scenarios, opponents, and scales) rather than point estimates of performance. However, the corpus indicates that robustness is not uniformly standardized across game genres, especially in complex RTS environments where experimental comparability is difficult [21,35].

5.5.4. Evolution Within the Cluster

The internal evolution of learning-augmented gaming search can be characterized as a sequence of integrations that progressively expand what search can handle:

1. **Learned evaluation and action priors integrated with search:** Neural policy/value guidance combined with tree search elevated decision quality in Go by targeting search effort effectively [5].
2. **Self play as a general mechanism for producing strategic priors:** Removal of human data dependencies established self play as a scalable pathway to strong decision quality and broad opponent robustness within a domain [22].
3. **Learned dynamics models enabling planning beyond explicit simulators:** Planning with learned models extended the search paradigm to settings where dynamics are not directly encoded or where simulation is expensive, changing the role of search from tree expansion over known transitions to planning over learned latent dynamics [20].
4. **Representation and scaling challenges in multi agent, partially observable RTS:** Modern work emphasizes representations (map features) and robustness under partial observability and scaling, reflecting the domain shift from board games to complex, real-time multi agent environments [17,21,35].
5. **Mastery of complex imperfect information games through model free MARL (Stratego/DeepNash)** [57].
6. **The emergence of search free grandmaster level play via large scale language modelling:** A development that may herald a new phase in which the role of explicit search is fundamentally reconsidered [34].

5.5.5. Comparative Insights and Trade offs

Learning augmented search sits at a crossroads of the taxonomy, inheriting strengths from classical search while introducing new trade offs.

- **Decision quality vs. compute and data dependence:** The highest decision quality in this cluster often relies on substantial training compute and data generation through self play or large scale interaction [20,22]. This contrasts with classical search methods whose performance depends more directly on runtime search budgets than on training pipelines.

- **Empirical performance vs. provable optimality:** These methods tend to deliver exceptional empirical decision quality without formal optimality guarantees [5,20]. This represents a fundamental trade off relative to classical adversarial search.
- **Robustness to strategic diversity vs. brittleness to distribution shift:** Self play and learned guidance can enhance robustness to opponent diversity, yet systems may remain brittle under shifts in environment structure, partial observability, or scale changes unless explicitly designed and evaluated for these regimes [17,21].
- **Guidance strength vs. failure severity:** When learned priors are accurate, they dramatically improve search efficiency; when incorrect, they can misdirect search systematically, potentially reducing robustness below that of less informed but more conservative methods [20].

Relative to MCTS only methods, learning augmentation tends to improve decision quality under fixed budgets by guiding exploration and improving value estimation [5]. Relative to metaheuristics, learning based systems can achieve higher peak performance but may require careful robustness engineering to avoid overfitting and brittleness, especially in complex RTS domains.

5.6. Cluster 6: Uncertainty, Partial Observability, and Search Games

Uncertainty, partial observability, and search games form a foundational and conceptually distinct cluster in the gaming search literature. Whereas adversarial game tree search formalizes competition under perfect information and MCTS emphasizes sampling based decision quality, this cluster centers on decision making when critical state variables are unknown, hidden, or dynamically changing. It includes classical search theory and pursuit evasion formulations, moving target search, incomplete information allocation games, and modern partial observability game formulations. Across these works, the dominant theme is that robustness is not an auxiliary property but the primary objective, with decision quality and optimality defined relative to probabilistic beliefs, worst case uncertainty, or strategic information asymmetry.

This cluster is essential to the survey's title because it shows how the field's notions of decision quality, optimality, and robustness evolved long before modern machine learning initially through analytic models of detection and search allocation, later through game theoretic formulations on networks, and most recently through explicitly partially observable game constructs.

5.6.1. Decision Quality Under Uncertainty: Belief Conditioned Effectiveness

In uncertain or partially observable settings, decision quality cannot be defined purely in terms of path cost, minimax value, or rollout averages over known states. Instead, decisions must be evaluated relative to beliefs about hidden variables (e.g., target location, opponent type, unseen units) and the expected or worst case consequences of actions under those beliefs.

Classical work on search and screening established decision quality in terms of probability of detection and the allocation of search effort under uncertainty [12]. In this framing, a "good" decision is one that maximizes the likelihood of locating a hidden target (or minimizes expected time to detection) given limited resources and uncertain target distributions. The theoretical structure here is notable: decision quality is intrinsically probabilistic and inherently distributional, anticipating later robustness evaluation norms in gaming AI.

Koopman style search theory formalized this further by deriving the optimum distribution of searching effort over space and time, explicitly linking decision quality to probabilistic detection models and constrained resource allocation [13]. Importantly, these formulations treat the environment as uncertain and evolving, meaning decision quality is inseparable from assumptions about target motion, sensing, and prior distributions. This establishes an early template for belief aware decision making that remains relevant for modern games with fog of war and hidden state.

Moving target search extends these ideas by introducing dynamic hidden variables. The FAB algorithm for searching a moving target represents a step from static detection probability optimization toward sequential decision making under target dynamics [14,39]. Decision quality here is no longer

about choosing a single best allocation but about selecting search policies that remain effective as uncertainty evolves over time.

5.6.2. Optimality: From Exact Solutions to Model Relative Optimal Policies

Optimality in this cluster is more subtle than in perfect information adversarial search. In many problems, the “optimal” policy is defined relative to a probabilistic model of uncertainty rather than to a fully observed deterministic game state.

Early formulations in search theory often aim for exact optimality with respect to a well defined objective—e.g., maximizing detection probability given a search budget under explicit assumptions about sensing and target behavior [12,13]. In these settings, optimality can be formal and analytic, but it is conditional: it holds only insofar as the uncertainty model is correct.

Networked and game theoretic formulations extend this conditional optimality. The linear programming approach to a network search game with a mobile hider illustrates how optimal policies can be derived in structured environments using optimization methods, with optimality defined relative to the network model and the assumed motion capabilities of the hider [37]. Similarly, constrained moving target search using branch and bound reflects an attempt to preserve optimality (or near optimality) under additional constraints that arise in practical pursuit evasion like settings [36,40]. These methods demonstrate an important evolutionary feature of the cluster: as models become more realistic and constrained, optimality often requires increasingly heavy computational machinery and becomes harder to maintain in real time.

Later work introduces explicit information asymmetry and private information, where optimality becomes intertwined with strategic disclosure and inference. The search allocation game with private information about the initial target position highlights that “optimal” allocation depends not only on physical uncertainty but also on information structure and incentives [58]. In such settings, optimality may correspond to equilibrium behavior under incomplete information rather than to a single agent optimum.

Modern partial observability game formulations push this evolution further. Partially Observable Off Switch Games (PO OSG) explicitly embed partial observability into a strategic game framework, where optimality is defined over policies that must act under hidden state and evolving beliefs [30,39]. Complementing this, heuristic search value iteration for partially observable stochastic shortest path games provides depth limited precision bounds that formally extend classical shortest path optimality to stochastic PO settings [39], bridging classical search guarantees and modern partial observability formalisms.

5.6.3. Robustness as a First Class Objective

Robustness is the defining concern of this cluster. In uncertain games, robustness is not merely “performance under noise”; it is the ability to sustain performance when key aspects of the environment or opponent are unknown, evolving, or strategically manipulated.

Search theory papers directly optimize performance over uncertain target distributions and dynamics [12–14]. Robustness here is often expectation based, aiming to maintain high detection probability across plausible target locations and motions. Network search games and allocation games incorporate adversarial or strategic uncertainty, where robustness includes resistance to worst case hiding strategies or equilibrium responses [37,58]. This differs from minimax robustness because the uncertainty includes both hidden state and strategic incentives.

Modern partial observability formulations highlight that robustness depends on the quality of belief updating and on the policy’s ability to act effectively despite incomplete information [30]. Reconnaissance Blind Chess—a chess variant with imperfect information has been used to study how supervised and reinforcement learning from observations can confer robustness under hidden state without traditional belief tracking search [38]. The mastery of Stratego through model free multi agent RL [57] and the Player of Games framework that unifies perfect and imperfect information game play [2] further expand the empirical robustness frontier of this cluster. Real time worst case

robustness under partial observability is formalized in R2PS, which extends dynamic programming pursuit strategies to partial observability settings with belief preservation and worst case guarantees on capture time [40].

A key insight from this cluster is that robustness often requires **distributional evaluation** rather than average case reporting. Because uncertainty induces variability across instances, robust performance must be assessed across diverse scenarios, target behaviors, or information realizations. This perspective is consolidated and systematized in the survey of search games, which emphasizes the breadth of models and evaluation approaches used to quantify performance under uncertainty [16].

5.6.4. Evolution Within the Cluster

The evolutionary trajectory of this cluster can be summarized as a movement from **analytic probability models toward structured optimization and game formulations** and finally toward **explicit partial observability games**:

- **Foundations (1940s–1960s):** Probabilistic detection and search allocation models formalize decision quality and robustness under uncertainty [12,13]. Differential game perspectives provide continuous time adversarial dynamics that anticipate pursuit evasion in game-like settings [59].
- **Algorithmic and constrained formulations (1980s–1990s):** Moving target and constrained search problems introduce dynamics and constraints that push the cluster toward algorithmic solutions (e.g., FAB, branch and bound, LP) [14,33,36,37,39,40].
- **Synthesis and information structure (2010s):** Survey work consolidates models and highlights the diversity of assumptions and evaluation practices in search games [16], while private information allocation games emphasize the strategic role of information asymmetry [58].
- **Modern partial observability games (2020s):** Explicit partial observability game constructs such as PO-OSG extend the cluster into broader AI/game theoretic contexts where uncertainty, incentives, and belief based decision making are central [30].

This evolution demonstrates that uncertainty driven gaming search did not simply “arrive” with modern RL; it has a deep lineage with increasingly sophisticated models of hidden state and strategic interaction.

5.6.5. Comparative Insights and Trade offs

Compared with other clusters, uncertainty and search games invert the typical priority ordering: robustness is primary, while decision quality and optimality are defined relative to uncertain beliefs or strategic information.

- **Versus adversarial game tree search:** Minimax provides robustness to a worst case opponent under perfect information, but search games address robustness to hidden state and uncertainty a different and often harder condition. The trade off is that formal minimax optimality is replaced by belief or model relative optimal policies [3,6,37].
- **Versus MCTS:** Sampling methods can approximate decision quality in stochastic settings, but search games often require explicit modeling of uncertainty and information structure; robustness may depend more on model correctness than on simulation budget [9,15,16].
- **Versus learning augmented methods:** learning can produce powerful priors for partially observable environments, but this cluster emphasizes that robustness requires principled handling of hidden information and evaluation across uncertainty distributions [2,38,57]. Modern partial observability formulations underscore that strong empirical performance alone does not resolve the conceptual difficulty of defining optimal behavior under partial observability [30].

The central trade off throughout this cluster is model specificity versus robustness under mismatch. Analytic optimality can be achieved under strong modeling assumptions, yet robustness may degrade when assumptions are violated. This trade off is increasingly relevant in modern games, where uncertainty models are rarely fully accurate and where opponent behavior may be nonstationary. Table 2 lists the representative corpus anchors for all six clusters.

Table 2. Representative anchors for each cluster in the survey taxonomy, drawn from the corpus bibliography

Cluster	Representative papers
C1: Pathfinding / Navigation	[10,11,17,18,23–27,31,41–45]
C2: Adversarial Game Tree	[1–4,6,7,20,29,34,56,59]
C3: MCTS / Bandits	[1,5,8,9,15,19,20,28,32,33,46–51,54,55,60]
C4: Metaheuristics / Rolling Horizon	[12,13,19,29,51–53]
C5: Learning Augmented Search	[1,2,5,17,20–22,34,35,38,45,50,54–57]
C6: Uncertainty / Search Games	[2,12–17,30,33,36–40,57–59,61]

6. Cross Cluster Comparative Analysis

The six clusters analyzed in Section 5 collectively trace the evolution of gaming search algorithms through distinct yet interconnected approaches to **decision quality**, **optimality**, and **robustness**. While each cluster addresses these dimensions through different mechanisms and assumptions, their historical interplay reveals recurring trade offs, migrations of ideas, and converging design principles. This section synthesizes insights across clusters, highlighting cross cutting patterns, hybridization trends, and paradigm shifts that have shaped modern game AI. Figure 6 schematically positions each cluster by its typical emphasis on decision quality, optimality, and robustness.

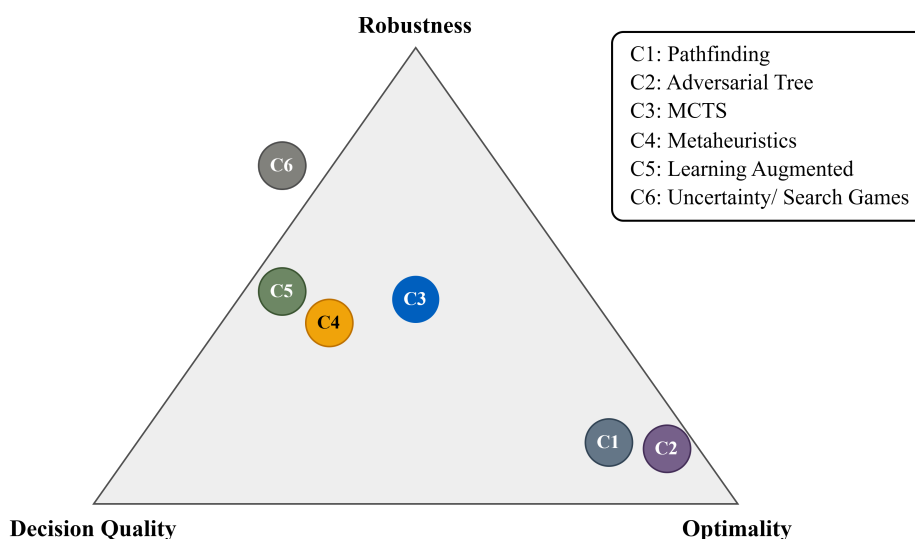


Figure 6. Schematic positioning of clusters by their typical emphasis on decision quality, optimality, and robustness. Positions summarize dominant tendencies; hybrids can occupy intermediate regions.

6.1. Decision Quality Across Clusters: From Structural Guarantees to Empirical Strength

Across clusters, decision quality evolves from **structural correctness** toward **empirical dominance**. Spatial pathfinding anchors decision quality in intrinsic measures such as path cost and feasibility, benefiting from clear problem structure and evaluation criteria [23,25]. Adversarial game-tree search reframes quality as relative strength against an opponent, where deeper search and better heuristics monotonically improve play under fixed assumptions [3,6]. In both clusters, decision quality is tightly coupled to formal structure and deterministic evaluation.

The emergence of MCTS marks a transition to **statistical decision quality**, where action selection quality is inferred from sampled outcomes rather than exact evaluation [8,9]. This shift decouples decision quality from handcrafted heuristics and enables scalability to larger games, but it also introduces variance and budget dependence. Optimization and metaheuristic methods extend this trend by treating decision quality as an explicit optimization objective under time constraints, often prioritizing short-horizon performance and adaptability [19].

Learning augmented search fundamentally redefines decision quality as a function of learned inductive bias and training processes. High quality decisions emerge from the interaction between learned priors and limited search, achieving levels of play unattainable by classical methods in complex domains [5,20,22]. In uncertainty driven search games, decision quality is conditioned on beliefs and distributions over hidden variables, emphasizing effectiveness under uncertainty rather than pointwise optimality [12–14].

A unifying insight is that **decision quality becomes increasingly contextual and empirical** as games grow in complexity. While early clusters emphasize correctness relative to a model, later clusters accept that quality must be demonstrated through performance across scenarios, opponents, and time budgets.

6.2. Optimality: From Exact Solutions to Conditional and Empirical Notions

Optimality exhibits the clearest divergence across clusters. Adversarial game tree search provides the strongest formal guarantees, with minimax optimality preserved through correctness preserving pruning and search control [6,29]. Spatial pathfinding similarly offers exact or bounded optimality under admissible heuristics and well defined cost functions [25]. These guarantees, however, are contingent on static environments, perfect information, and sufficient computation.

MCTS introduces **asymptotic optimality**, replacing exact guarantees with convergence properties that hold only in the limit of infinite simulation [8,9]. In finite time gaming contexts, optimality becomes approximate and probabilistic, dependent on rollout policies and exploration strategies. Optimization and metaheuristic approaches further relax optimality, treating it as a best so far outcome under resource constraints rather than a provable property [19].

Learning augmented methods represent a decisive shift: optimality is largely **empirical and benchmark driven**, defined by dominance over strong baselines rather than by convergence proofs [5,20,22]. In uncertainty and search games, optimality is often model relative, defined with respect to probabilistic assumptions or equilibrium concepts under incomplete information [13,30,37,58].

Across clusters, optimality thus transitions from a central design goal to a **conditional or aspirational property**, increasingly subordinate to scalability and robustness. This evolution reflects a broader acceptance that exact optimal play is unattainable in most modern games and that practical success requires reframing optimality in context sensitive terms.

6.3. Robustness: Expanding the Notion of Reliability

Robustness broadens significantly across clusters, evolving from narrow worst case reasoning to multi faceted resilience under uncertainty, variability, and constraint. In adversarial game tree search, robustness is achieved through pessimistic opponent modeling, ensuring performance against worst case play [3,6]. This form of robustness is powerful but limited to perfect information settings and can lead to conservative behavior.

Spatial pathfinding emphasizes robustness to environmental dynamics and time constraints, introducing replanning, abstraction, and anytime behavior to maintain acceptable performance under change [10,26,27]. MCTS enhances robustness through stochastic sampling and graceful degradation under reduced computation, performing reliably across stochastic dynamics and variable budgets [8,9,48].

Optimization and metaheuristic methods prioritize robustness through stochasticity and population diversity, reducing sensitivity to local optima and noisy evaluations [19]. Learning augmented approaches add robustness through exposure to diverse training scenarios and self play, yet simultaneously introduce new vulnerabilities related to distribution shift, model bias, and generalization [5,20–22].

The uncertainty and search games cluster foregrounds robustness as the primary objective, formalizing performance under hidden state, probabilistic beliefs, and strategic information asymmetry [12–14,16]. This cluster anticipates modern concerns about distributional evaluation and belief aware

decision making, highlighting that robustness must often be evaluated across ensembles of scenarios rather than through single point metrics.

6.4. Hybridization and Migration of Ideas

A salient cross cluster trend is the progressive hybridization of methods. MCTS inherits adversarial reasoning from minimax while abandoning deterministic evaluation; learning augmented MCTS fuses statistical planning with learned priors [5]. RTS systems integrate script guided MCTS to manage combinatorial explosion, blending domain knowledge with sampling based planning [28]. Learning based pathfinding and representation learning for navigation in partially observable environments demonstrate migration from the pathfinding cluster into learning augmented and uncertainty driven domains [17].

Similarly, concepts from search games such as belief dependent decision making and distributional robustness reappear implicitly in modern learning systems that must act under partial observability and opponent uncertainty [30]. These migrations suggest that no single cluster offers a complete solution; instead, progress arises from recombining principles across clusters to address emerging challenges.

6.5. Trade offs and Design Frontiers

Comparing clusters reveals persistent trade offs that define the design frontier of gaming search algorithms:

- **Decision quality vs. Optimality:** Exact guarantees often limit scalability, while empirical methods achieve higher quality in practice without proofs.
- **Optimality vs. Robustness:** Worst case guarantees can reduce adaptability, whereas robust empirical performance may sacrifice formal correctness.
- **Robustness vs. Efficiency:** Stochastic and population based methods improve resilience but may increase computational cost.
- **Learning driven quality vs. Generalization robustness:** Learned priors elevate performance but risk brittleness under distribution shift.

These trade offs are not static; they shift with game genre, information structure, and computational regime. Board games tolerate deeper deliberation and benefit from adversarial guarantees, while RTS and partially observable games demand rapid, robust decisions under uncertainty [17,21,28].

6.6. Synthesis

Viewed collectively, the clusters chart an evolution from **model centric** to **performance centric** gaming search. Early methods prioritize optimality within well defined models; later methods prioritize decision quality and robustness under real world constraints. The convergence of sampling, optimization, learning, and uncertainty modeling reflects an emerging consensus: effective gaming search must balance the three dimensions dynamically, guided by the structure and demands of the target game.

This cross cluster synthesis sets the stage for examining how evaluation practices evolved to accommodate these shifts and how future research might reconcile empirical success with principled guarantees in increasingly complex gaming environments. Table 3 summarizes the central cross cluster trade offs with respect to decision quality, optimality, and robustness.

Table 3. Cross cluster trade offs central to decision quality, optimality, and robustness in gaming search

Trade off	Cluster tendency A	Cluster tendency B	Survey insight
Optimality vs Scalability	C1/C2 (stronger guarantees)	C3/C5 (scales to huge spaces)	As games grow, strict guarantees become conditional or asymptotic; empirical dominance often replaces proof based optimality.
Worst case robustness vs Empirical robustness	C2 (worst case opponent)	C3/C5 (statistical + learned robustness)	Worst case models can be conservative; sampling/learning broaden robustness but create new failure modes (shift, model bias).
Quality vs Time budget (anytime behavior)	C1/C4 (explicit anytime design)	C2 (depth limited)	Real-time domains reward algorithms that degrade gracefully; quality must be measured as a function of compute, not a single score.
Model relative optimality vs Robustness under mismatch	C6 (optimal under model)	C5 (learns from data, risks shift)	C6 delivers principled belief conditional policies but depends on assumptions; learning can generalize but must be stress tested distributionally.
Handcrafted guidance vs Learned guidance	C1/C2 (heuristics/eval)	C5 (learned priors/models)	Learned priors can collapse branching factors and raise decision quality, but robustness demands evaluation across scenarios and scales.

7. Evolution of Evaluation Methodologies in Gaming Search

Evaluation methodology has played a decisive role in shaping how decision quality, optimality, and robustness are defined, pursued, and compared in gaming search algorithms. As game domains, algorithmic paradigms, and computational constraints evolved, so too did the metrics and experimental protocols used to assess progress. This section traces the historical evolution of evaluation practices across the clusters identified earlier, highlighting how shifts in evaluation both reflected and actively influenced the design of gaming search algorithms.

7.1. Early Analytic and Correctness Oriented Evaluation

In the earliest phases of gaming related search research, evaluation was primarily **analytic and model driven**. Foundational work in search theory assessed decision quality and optimality through closed form expressions and probabilistic models probability of detection, expected search time, or optimal effort allocation intrinsically tied to explicit assumptions about sensing, target motion, and uncertainty [12,13].

Early pathfinding research similarly emphasized correctness and optimality guarantees, centering evaluation on shortest path admissibility, heuristic properties, and computational complexity [23,25], while adversarial game playing work inferred quality from search depth and evaluation function

fidelity, framing optimality through minimax correctness [3]. Robustness in this era was largely implicit, assumed through worst case analysis rather than empirically tested.

7.2. Empirical Performance and Competitive Benchmarks

As games grew in complexity and computing power increased, evaluation shifted toward **empirical performance measurement**. In adversarial game tree search, head to head matches and tournament style evaluations became standard, providing direct evidence of decision quality relative to competing systems [6,7,29], culminating in large scale systems evaluated through competitive outcomes against human experts [4]. This competitive paradigm reframed decision quality as relative strength rather than intrinsic correctness, and optimality was inferred from dominance in match play rather than demonstrated through proof.

In pathfinding, comparative studies assessed algorithmic efficiency across different map types and sizes, with metrics such as runtime, memory usage, and node expansions complementing traditional path optimality measures [42]. These studies established the important principle that evaluation must account for **environmental diversity** rather than a single canonical instance.

7.3. Anytime Behavior and Real-Time Constraints

The introduction of real-time constraints marked a critical inflection point in evaluation methodology. Algorithms designed for real-time heuristic search and interactive game environments could no longer be evaluated solely on final solution quality; instead, performance had to be measured as a function of time [10]. Metrics such as time per decision, responsiveness, and graceful degradation under reduced computation became central to assessing decision quality.

Comparative evaluations between rolling horizon evolutionary methods and tree search explicitly highlighted the importance of anytime behavior, demonstrating that algorithms should be judged on how quickly they produce acceptable decisions, not just on asymptotic quality [19,51,52]. In commercial game AI contexts, evaluation further expanded to include qualitative criteria such as behavioral smoothness, predictability, and player experience [11], underscoring the widening scope of what constitutes meaningful evaluation in gaming.

7.4. Statistical Evaluation and Sampling Based Metrics

The rise of Monte Carlo Tree Search necessitated new evaluation practices aligned with **statistical decision making**. Comprehensive surveys of MCTS methods established that performance depends on simulation budgets, stochastic outcomes, and exploration parameters, making averages over multiple runs, confidence intervals, and convergence behavior standard evaluation tools [8,9,32,49]. Metrics such as number of simulations, rollouts per second, and convergence behavior became standard proxies for both decision quality and computational efficiency. More recent work on alternative selection strategies has further refined these criteria by requiring joint assessment of finite budget performance and asymptotic behavior across benchmarks [46].

In imperfect information games, evaluation complexity increased further. Performance could vary widely depending on hidden state realizations and opponent strategies, motivating repeated trials and aggregated statistics to assess robustness [15]. However, the lack of standardized benchmarks in such settings often made cross paper comparison difficult, highlighting a recurring challenge in evaluating robustness under uncertainty. Parallel and hardware accelerated MCTS introduced additional evaluation dimensions, including scalability with respect to compute resources and efficiency gains from parallelization [48].

7.5. Learning Centric and Distributional Evaluation

Learning augmented search fundamentally altered evaluation methodology by introducing **training dynamics** and **generalization** as core concerns. In neural guided and self play systems, evaluation typically focuses on win rates or scores against fixed baselines or across standardized test sets, often after extensive training [1,5,20,22]. Decision quality is thus evaluated at the system level

rather than at the level of individual decisions or search steps. Recent work has extended this further: unified systems evaluated across both perfect and imperfect information games [2], agents assessed under strict sample budgets [54], and systems achieving strong play without search at inference [34] each expose limitations of relying on a single outcome metric.

Crucially, learning based approaches expose the limitations of single number metrics. Performance can depend strongly on training data, hyperparameters, and evaluation protocol. As a result, robustness increasingly requires evaluation across multiple scenarios, opponents, and environment configurations. This issue is particularly pronounced in RTS games, where partial observability, multi agent interaction, and scale variation complicate evaluation [17,21,35,57]. Survey work in this domain emphasizes that differences in experimental setup can substantially affect reported performance, underscoring the need for more systematic robustness evaluation [35].

7.6. Evaluation in Uncertain and Search Game Domains

The uncertainty and search games cluster contributes a distinctive evaluation perspective centered on **distributional performance under hidden state**. Classical metrics such as probability of detection and expected capture time are explicitly distribution based, providing a natural measure of robustness across uncertainty realizations [12–14]. Survey work on search games consolidates this perspective by cataloging diverse evaluation approaches tailored to different uncertainty models and problem structures [16].

Extensions to multi player imperfect information games have introduced additional criteria such as expected value under opponent uncertainty [33] and convergence bounds under partial observability [39] while pursuit oriented domains have employed worst case capture time guarantees as primary robustness measures [40]. Modern partial observability game formulations extend this tradition by embedding evaluation within belief space reasoning and policy performance under information constraints [30], reinforcing the view that robust evaluation must account for uncertainty explicitly rather than treating it as noise around a nominal case.

7.7. Synthesis and Open Challenges

Across clusters and eras, evaluation methodologies in gaming search have evolved from **proof centric correctness to empirical, distributional, and system level assessment**. This evolution mirrors the field’s shifting priorities: from exact optimality in small, well defined games to robust decision quality under uncertainty, scale, and real-time constraints.

Several persistent challenges remain. First, evaluation practices are often tightly coupled to specific game genres, limiting cross domain comparability. Second, robustness claims frequently depend on narrow experimental settings, especially in learning based systems. Third, the coexistence of analytic, empirical, and qualitative metrics complicates synthesis across algorithm families.

Nevertheless, the historical trajectory is clear: meaningful evaluation in gaming search increasingly requires **multi dimensional, distribution aware, and resource sensitive metrics**. Understanding this evolution is essential for interpreting past results and for designing future evaluation frameworks that align with the intertwined goals of decision quality, optimality, and robustness in complex gaming environments. Table 4 provides a structured overview of the evaluation metrics employed across the surveyed literature, organized by the three core dimensions and resource efficiency.

Table 4. Evaluation metrics palette used in the surveyed gaming search literature, aligned with the survey’s three core dimensions plus resource efficiency

Metric Category	Examples	Clusters where common
Decision quality (outcome)	Win rate; tournament results; score/return; task success rate	C2, C3, C5 (also C4)

Table 4. Cont.

Metric Category	Examples	Clusters where common
Decision quality (solution quality)	Path cost/length; sub optimality gap; plan quality over horizon	C1, C4
Optimality / guarantees	Minimax correctness; admissibility; convergence (asymptotic); equilibrium notions (model relative)	C1, C2, C3, C6
Robustness (distributional)	Performance across maps/scales; scenario sweeps; robustness to hidden info; variance across seeds	C3, C5, C6
Robustness (search theoretic)	Detection probability; expected capture time; coverage under uncertainty	C6
Resource efficiency	Runtime; time per move; nodes expanded; rollouts/second; memory usage; GPU throughput	C1, C2, C3 (also C4, C5)

8. Game Type Driven Insights

While the preceding sections analyzed gaming search algorithms through taxonomy driven and cross cluster lenses, an equally important perspective is **game type driven analysis**. Different game genres impose distinct structural, informational, and temporal constraints that fundamentally shape how decision quality, optimality, and robustness can be achieved. This section synthesizes insights across clusters by examining how search algorithms behave in and are adapted to specific classes of games. Rather than introducing new categories, the discussion maps the existing taxonomy onto game types, revealing why particular algorithmic paradigms dominate in certain domains and how evaluation priorities shift accordingly.

8.1. Board Games: Perfect Information and Strategic Optimality

Classical board games such as Chess, Go, and Shogi represent the most extensively studied gaming domains in the literature. Decision quality is closely tied to strategic coherence and long horizon planning, with empirical strength correlating strongly with search depth and pruning efficiency [3,6,7]. Optimality is most sharply defined here: minimax provides a clear theoretical target, and correctness preserving pruning ensures this is not compromised by efficiency improvements [6], while large scale engineered systems demonstrated near optimal play through computational depth and domain specific heuristics [4].

Robustness is primarily adversarial. MCTS and learning augmented search rebalanced this dimension by sacrificing worst case guarantees in favor of empirical robustness across diverse opponents and positions [5,8,22]. Subsequent self play systems generalized this pattern across Chess, Shogi, and Go under a single algorithm [1], while work targeting general game playing extended these training and evaluation norms beyond fixed game families [55]. Notably, recent work achieving grandmaster level Chess without search at inference time challenges the assumption that decision quality and search depth are inseparable [34], and unified systems capable of strong play across both perfect and imperfect information games further broaden what optimality means in this domain [2].

8.2. Navigation and Spatial Games: Real-Time Optimality Under Environmental Dynamics

Navigation and spatial games impose stringent real-time and responsiveness constraints on otherwise deterministic, single agent problems. Decision quality is intrinsically defined through path cost and smoothness, but is strongly coupled to computational latency a slightly suboptimal path produced quickly often outperforms an optimal one computed too slowly [18,23,25]. Optimality is well defined in theory but selectively relaxed in practice through hierarchical abstraction and incremental

replanning, trading strict minimality for scalability and robustness [26,27]. Recent integration of large language model reasoning with incremental heuristic search has introduced a further trade off dimension, where global planning guidance reduces node expansions but shifts evaluation toward combined runtime and path quality metrics [45].

Robustness emphasizes adaptation to environmental change and computational variability rather than adversarial behavior. Real-time heuristic search, online pruning, and abstraction ensure stable performance across diverse maps and runtime budgets [18,24,43], foreshadowing similar concerns that reappear in real-time multi agent settings.

8.3. Real-Time Strategy Games: Multi Agent Complexity and Partial Observability

RTS games combine large state spaces, simultaneous multi agent decision making, partial observability, and strict real-time constraints, making them among the most demanding domains for search algorithms. Decision quality depends on coordinated behavior across agents and timely adaptation to hidden information, not on isolated action optimality. MCTS variants guided by domain scripts emerged as a practical compromise enabling limited lookahead under real-time constraints [28], while empirical analysis in multi agent settings has shown that behavioral characteristics under tight compute budgets explain performance differences between search paradigms more fully than outcome metrics alone [51].

Optimality in RTS games is fundamentally approximate; success is measured through win rates and task specific objectives rather than through formal guarantees. Learning augmented methods operationalize optimality as empirical dominance through self play and large scale training [17,21,35]. Robustness is the dominant concern, requiring belief aware decision making, generalization across scenarios, and resilience to incomplete information [17], with scale flexibility work underscoring that robustness must encompass competence across structurally different environments, not merely noise tolerance [21].

8.4. Stochastic and Imperfect Information Games: Belief Aware Decision Making

Games with stochastic dynamics or imperfect information foreground uncertainty as a central challenge. Decision quality is evaluated relative to beliefs about hidden state and expected outcomes rather than to known trajectories [12–14], while optimality is model relative conditional on assumed probability distributions or information structures and sensitive to model mismatch [37,58]. Superhuman performance in large scale multiplayer imperfect information settings such as poker demonstrated that search combined with self play can achieve robust decision quality even when beliefs over hidden state are approximate [33], and work on imperfect information chess variants has shown that learned policies can improve without explicit search when the information structure is appropriately modeled [38].

Robustness is explicitly formalized in these domains, assessed across distributions of uncertainty through expectation, worst case outcomes, or equilibrium behavior under incomplete information [58]. Heuristic search methods for partially observable stochastic games have introduced convergence bound metrics that characterize solution quality when exact solutions are intractable [39], while worst case pursuit strategies under partial observability ground robustness evaluation in provable capture time guarantees rather than average case performance [40]. These approaches anticipate distributional evaluation practices increasingly relevant for modern learning based systems under partial observability [30]. Figure 7 maps key superhuman performance milestones by game domain and dominant algorithmic cluster across historical eras.

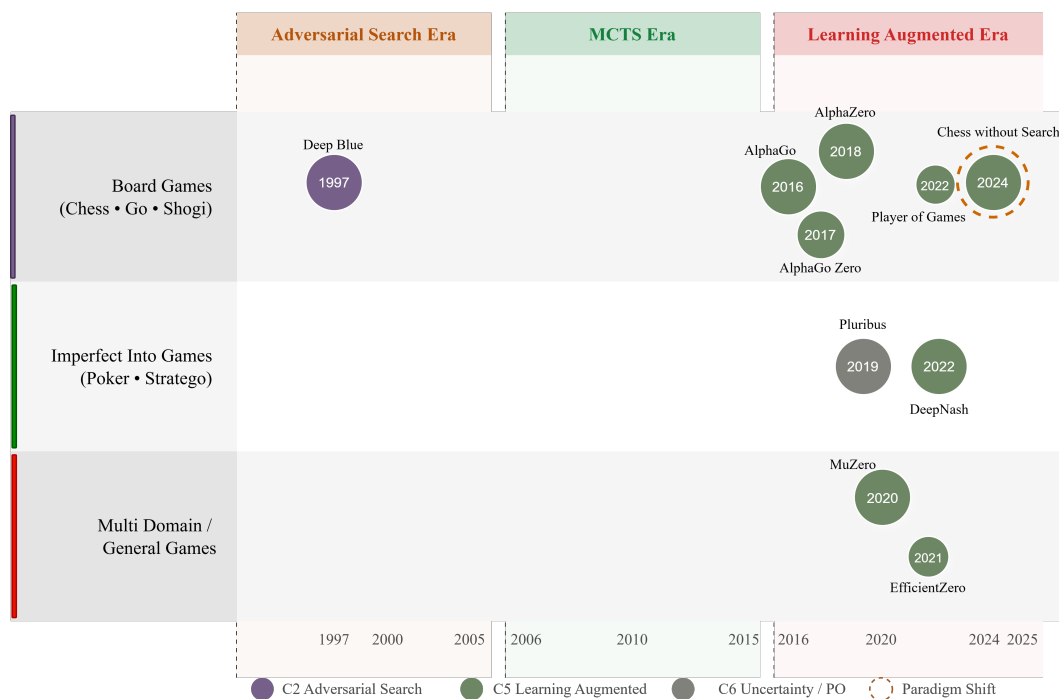


Figure 7. Superhuman performance milestones by game domain. Swim lanes separate game families. Circle size indicates landmark significance. Era bands show the dominant algorithmic cluster per period.

8.5. General and Multi Domain Games: Scalability and Transfer

Some modern research targets generality across multiple games or game families. Decision quality is defined by consistent performance across heterogeneous tasks rather than excellence in a single game [20,55], and optimality becomes an abstract benchmark evaluated relative to baselines within each domain. Robustness takes on a broader meaning—the ability to transfer knowledge, adapt representations, and maintain competence across structurally diverse environments connecting game AI to broader themes in general intelligence and highlighting the limitations of narrowly optimized, domain specific methods. Evaluation under strict data constraints further reveals that sample efficiency is an important dimension of robustness in general settings [54].

8.6. Synthesis: How Game Type Shapes Algorithmic Priorities

Across game types, a consistent pattern emerges: **the relative importance of decision quality, optimality, and robustness shifts with game structure.** Board games prioritize optimality and strategic depth; navigation games emphasize responsiveness and environmental robustness; RTS games demand adaptability under partial observability and scale; uncertain and search games foreground belief aware robustness.

These differences explain why no single cluster dominates all game types. Instead, effective gaming AI arises from aligning algorithmic design with the structural demands of the target game. Understanding these game type driven priorities is essential for selecting, evaluating, and advancing search algorithms in gaming and for appreciating why the field has repeatedly reinvented its methods as new game domains emerge.

This perspective sets the stage for identifying open challenges and future research directions, particularly those aimed at unifying decision quality, optimality, and robustness across increasingly diverse and complex game types. Table 5 summarizes the dominant clusters, primary challenges, and typical evaluation metrics for each game class.

Table 5. Game type driven view: dominant clusters and evaluation priorities by game class

Game type	Dominant clusters	Primary challenges	Typical metrics
Board games (Chess/Go/Shogi)	C2, C3, C5	Huge game trees; long horizon strategy; opponent strength; compute constraints	Win rate; Elo/rating; depth; rollouts; node expansions
Navigation / Grid maps	C1 (sometimes C5)	Large maps; dynamic obstacles; real-time constraints; smoothness and stability	Path cost/length; runtime; node expansions; replanning latency
Real-time strategy (RTS)	C3, C5 (plus hybrids)	Partial observability; multi agent coordination; large branching; real-time action selection; scale changes	Win rate; scenario success; time per decision; generalization across maps/scales
Stochastic / Imperfect information games	C3, C6 (sometimes C5)	Hidden state; randomness; belief maintenance; opponent unpredictability	Expected return; robustness across scenarios; variance; detection probability
Search / Pursuit evasion style domains	C6 (sometimes C1/C3 as tools)	Moving targets; sensing uncertainty; adversarial hiding; resource allocation	Detection probability; expected capture time; coverage; worst case guarantees

9. Open Challenges and Future Research Directions

Despite decades of progress, the evolution traced in this survey reveals that achieving a principled and unified balance among decision quality, optimality, and robustness in gaming search algorithms remains an open problem. As games continue to increase in complexity, scale, and diversity, existing approaches expose fundamental limitations that motivate new research directions. This section synthesizes the most salient open challenges identified across clusters and outlines promising avenues for future work, grounded in the patterns and gaps observed in the surveyed literature.

9.1. Reconciling Empirical Performance with Principled Optimality

A central challenge in modern gaming AI is the growing disconnect between **empirical dominance** and **formal optimality**. Learning augmented and MCTS based systems achieve exceptional performance in large games, yet their relationship to classical notions of optimal play is often unclear [5,20,22]. While asymptotic or conditional guarantees exist in some cases [8,9], finite time behavior where real gaming systems operate remains difficult to characterize theoretically.

Future research must address whether new notions of optimality can bridge this gap. Potential directions include:

- **Resource bounded optimality frameworks** that explicitly account for computation, uncertainty, and learning dynamics.
- **Policy level optimality criteria** for learned systems that connect empirical dominance to equilibrium concepts under realistic assumptions.
- **Formal analysis of hybrid systems**, where learned priors guide search but do not fully replace it, clarifying how guarantees degrade or persist under approximation.

Such efforts would help re-anchor optimality as a meaningful concept in modern gaming contexts without reverting to infeasible exact guarantees.

9.2. Robustness Under Distribution Shift and Model Mismatch

Robustness has emerged as a defining concern, particularly for learning augmented and uncertainty driven methods. While self play and sampling improve robustness within a training or

simulation regime, many systems remain vulnerable to **distribution shift**, **rule variations**, or **structural changes** in game environments [17,20,21]. This vulnerability is especially pronounced in partially observable and multi agent games, where hidden information and nonstationary opponents exacerbate mismatch between training and deployment conditions.

Key open challenges include:

- **Robust generalization guarantees** for learned policies and value functions in games with changing dynamics or information structures.
- **Belief aware and uncertainty calibrated learning**, integrating ideas from search games and partial observability more explicitly into modern learning pipelines [30,58].
- **Evaluation protocols that stress test robustness**, moving beyond narrow benchmarks to scenario ensembles that reveal brittleness.

Addressing these challenges requires tighter integration between classical robustness oriented formulations and contemporary learning based systems.

9.3. Unified Evaluation Methodologies Across Game Types

As shown in Section 7, evaluation practices have diversified alongside algorithmic approaches, resulting in fragmented metrics and protocols that hinder cross domain comparison. Decision quality, optimality, and robustness are often measured differently across clusters and game genres, making it difficult to assess progress holistically.

Future research should prioritize:

- **Multi dimensional evaluation frameworks** that jointly report quality, optimality approximation, robustness, and computational cost.
- **Distributional and variance aware metrics** that capture performance stability across uncertainty, opponents, and scenarios [35,58].
- **Standardized yet flexible benchmarks** that accommodate diverse game types while preserving comparability.

Without such evaluation advances, claims about algorithmic superiority risk remaining domain specific and difficult to generalize.

9.4. Scaling to Multi Agent, Real-Time, and Partially Observable Games

RTS and similar complex games expose the limitations of nearly all existing clusters. Classical search struggles with branching factor and time pressure; MCTS requires aggressive abstraction; learning based systems face generalization and coordination challenges [17,21,28,35]. Despite substantial progress, no approach yet provides a comprehensive solution that simultaneously ensures high decision quality, meaningful optimality approximations, and robust performance under uncertainty.

Open directions include:

- **Hierarchical and decomposed search learning hybrids** that separate strategic, tactical, and operational decision layers.
- **Coordination aware evaluation and training** for multi agent decision quality beyond independent or loosely coupled policies.
- **Belief space planning at scale**, integrating partial observability directly into search and learning architectures.

Progress in this area is likely to require principled combinations of ideas from multiple clusters rather than incremental refinements within a single paradigm.

9.5. Interpretable and Controllable Gaming Search Systems

As gaming AI systems become increasingly complex, interpretability and controllability emerge as underexplored dimensions of robustness. Handcrafted search algorithms offer transparency and

predictability, while learning augmented systems often behave as opaque black boxes whose failure modes are difficult to diagnose [20,21].

Future research may explore:

- **Interpretable search learning hybrids**, where learned components provide guidance without obscuring decision rationale.
- **Constraint aware and controllable learning**, allowing designers to impose safety, style, or fairness constraints on decision making.
- **Post hoc analysis tools** that connect observed game play behavior to underlying search and learning dynamics.

Such work would not only enhance robustness but also improve trust and usability in real world gaming applications.

9.6. Toward Unified Frameworks for Quality, Optimality, and Robustness

A recurring theme throughout this survey is that decision quality, optimality, and robustness have historically been treated as competing objectives, optimized in isolation or traded off implicitly. An important long term challenge is the development of **unified frameworks** that make these trade offs explicit and controllable.

Promising directions include:

- **Meta reasoning and adaptive computation allocation**, where algorithms dynamically balance search depth, exploration, and robustness based on game context and uncertainty.
- **Cross cluster synthesis**, combining worst case reasoning from adversarial search, statistical robustness from MCTS, and inductive bias from learning augmented methods.
- **Game aware algorithm selection and configuration**, where the structure of the target game guides the prioritization of quality, optimality, and robustness.

Achieving such unification would represent a significant conceptual advance, enabling gaming search algorithms to adapt their behavior systematically rather than through ad hoc design choices.

9.7. Outlook

The trajectory of gaming search algorithms reveals a field that continually redefines success in response to new challenges. Early emphasis on exact optimality gave way to empirical decision quality and, more recently, to robustness under uncertainty and scale. The next phase of research will likely focus on **making these trade offs explicit, principled, and measurable**, rather than accepting them as implicit consequences of algorithm choice.

By grounding future work in the lessons distilled across clusters and game types, the community can move toward gaming search systems that not only perform well but also offer clearer guarantees, stronger robustness, and deeper insight into the nature of intelligent decision making in games.

10. Conclusions

This taxonomy-based survey examined the evolution of **decision quality, optimality, and robustness** in gaming search algorithms, analyzing 61 works from 1946 to 2025 through a gaming-specific taxonomy and four-dimensional design space. The literature was organized into six thematic clusters — spatial pathfinding and navigation, adversarial game-tree search, Monte Carlo tree search and bandit-based planning, metaheuristic optimization, learning-augmented search, and search under uncertainty and partial observability enabling systematic cross-cluster comparison of methods that have historically been studied in isolation.

The evidence across all six clusters converges on a central finding: **no single algorithmic paradigm simultaneously maximizes decision quality, optimality, and robustness across all gaming contexts**. Early methods anchored in well-specified deterministic models achieved formal optimality guarantees but proved brittle as games grew in scale, interactivity, and uncertainty. Sampling-based planning decoupled decision quality from exact evaluation, recasting optimality as an asymptotic property and

robustness as graceful degradation under constrained computation. Learning-augmented methods subsequently achieved unprecedented empirical performance in large, complex games, yet traded formal guarantees for benchmark dominance and introduced new robustness vulnerabilities around generalization and distribution shift. Across all clusters, this pattern of rebalancing — rather than simultaneous improvement along all three dimensions has been the primary engine of progress.

This survey delivers four contributions. First, a **gaming-specific taxonomy** structures a fragmented literature into six coherent clusters grounded in how algorithms are actually deployed and evaluated in games, rather than in algorithmic technique alone. Second, a **four-dimensional design space** — covering interaction topology, information structure, computational regime, and source of search guidance provides a consistent framework for comparing methods across clusters, including hybrid approaches. Third, a **cross-cluster historical synthesis** traces the paradigm shifts from analytic correctness toward empirical benchmarking, sampling-based planning, and neural-guided search, documenting how evaluation methodologies evolved in parallel. Fourth, a **structured map of open challenges** identifies where principled frameworks for managing quality–optimality–robustness trade-offs remain absent, providing a research agenda for the field.

The corpus of 61 works was assembled through domain-scoped literature selection and is not exhaustive; papers at the boundary of gaming and adjacent fields such as robotic planning or combinatorial optimization were included only where they contributed directly to gaming search methodology. The temporal coverage extends to early 2025, and works published or released after this horizon are not reflected. Single-author classification of papers into taxonomy clusters introduces subjectivity at boundary cases, particularly for hybrid methods that span multiple clusters. These constraints bound the claims of comprehensiveness but do not affect the core analytical findings.

The open challenges identified across clusters — reconciling empirical performance with principled optimality, achieving robust generalization under distribution shift, unifying evaluation methodologies, scaling to multi-agent and partially observable environments, and developing interpretable search–learning hybrids are examined in depth. Addressing these challenges will require principled frameworks that make quality–optimality–robustness trade-offs explicit and controllable, rather than accepting them as implicit consequences of algorithm design. This survey provides artificial intelligence researchers and game developers with a structured foundation for navigating those trade-offs across the full landscape of gaming search algorithms.

Author Contributions: Conceptualization, V.S.S.; methodology, V.S.S.; software, V.S.S.; validation, V.S.S.; formal analysis, V.S.S.; investigation, V.S.S.; resources, V.S.S. and N.S.; data curation, V.S.S.; writing—original draft preparation, V.S.S.; writing—review and editing, V.S.S. and N.S.; visualization, V.S.S.; supervision, N.S.; project administration, V.S.S.; funding acquisition, V.S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
MCTS	Monte Carlo Tree Search
RTS	Real-Time Strategy
UCT	Upper Confidence bounds applied to Trees
FPS	First Person Shooter
HPA*	Hierarchical Pathfinding A*
LLM	Large Language Model

RL	Reinforcement Learning
MARL	Multi Agent Reinforcement Learning
PO-OSG	Partially Observable Off Switch Game
GPU	Graphics Processing Unit
FAB	(Washburn's) Find-A-moving-target via Backward induction Algorithm
LP	Linear Programming
DP	Dynamic Programming
R2PS	Worst Case Robust Real-Time Pursuit Strategies
LLM-A*	Large Language Model enhanced A*

Appendix A

Table A1. Full corpus mapping: all papers with complete taxonomy attributes. Cluster codes — C1: Spatial Pathfinding & Navigation Search; C2: Adversarial Game-Tree Search; C3: Monte Carlo Tree Search & Bandit Planning; C4: Optimization & Metaheuristic Search (incl. Rolling Horizon); C5: Learning Augmented Search & Neural Planning; C6: Uncertainty, Partial Observability & Search Games.

Title	Year	Cluster(s)	Game Type	Key Metrics (families)
Search and Screening [12]	1946	C6	Stochastic / Imperfect information	Runtime; Search effort; Robustness / Detection
Programming a Computer for Playing Chess [3]	1950	C2	Board games	Not specified
The theory of search III: the optimum distribution of searching effort [13]	1957	C6	Stochastic / Imperfect information	Path cost / Optimality; Runtime; Robustness / Detection
A Note on Two Problems in Connexion with Graphs [23]	1959	C1, C2	Navigation / Grid maps	Path cost / Optimality; Search effort
A Formal Basis for the Heuristic Determination of Minimum Cost Paths [25]	1968	C1, C2	Navigation / Grid maps	Path cost / Optimality; Search effort
An Analysis of Alpha-Beta Pruning [6]	1975	C1, C2	Board games	Path cost / Optimality; Search effort
Search for a Moving Target: The FAB Algorithm [14]	1983	C6	Stochastic / Imperfect information	Search effort; Robustness / Detection
Depth-First Iterative-Deepening: An Optimal Admissible Tree Search [10]	1985	C1, C2	Navigation / Grid maps	Path cost / Optimality; Search effort
The history heuristic and alpha-beta search enhancements in practice [29]	1989	C2	Board games	Search effort
Real-Time Heuristic Search [18]	1990	C1	Navigation / Grid maps	Runtime; Search effort
An optimal branch-and-bound procedure for the constrained path moving target search problem [36]	1990	C6	Stochastic / Imperfect information	Path cost / Optimality; Search effort; Robustness / Detection

Table A1. Cont.

Title	Year	Cluster(s)	Game Type	Key Metrics (families)
A linear programming approach to the search game on a network with a mobile hider [37]	1992	C6	Stochastic / Imperfect information	Search effort; Robustness / Detection
Optimal and Efficient Path Planning for Partially-Known Environments [31]	1994	C1, C6	Navigation / Grid maps	Path cost / Optimality; Runtime; Robustness / Detection
Best-first fixed-depth minimax algorithms [7]	1995	C2	Board games	Search effort
Differential Games (Book) [59]	1999	C1, C2, C6	Search / Pursuit Evasion	Path cost / Optimality; Runtime; Robustness / Detection
Deep Blue [4]	2002	C2	Board games	Not specified
D* Lite [26]	2002	C1	Navigation / Grid maps	Path cost / Optimality; Runtime; Search effort
Near Optimal Hierarchical Path-Finding [27]	2004	C1	Navigation / Grid maps	Path cost / Optimality; Runtime
Three States and a Plan: The A.I. of F.E.A.R. [11]	2005	C1	Navigation / Grid maps	Runtime; Robustness / Detection
A Two-Sided Optimization for Theater-Ballistic Missile Defense [61]	2005	C6	Search / Pursuit Evasion	Runtime; Robustness / Detection
Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search [8]	2006	C3, C5	Not specified	Search effort
Bandit Based Monte-Carlo Planning [9]	2006	C3, C5	Not specified	Search effort
Online Graph Pruning for Pathfinding on Grid Maps [43]	2011	C1	Navigation / Grid maps	Runtime; Search effort
Jump Point Search [24]	2011	C1	Navigation / Grid maps	Runtime; Search effort
Information Set Monte Carlo Tree Search [15]	2012	C3, C6	Stochastic / Imperfect information	Search effort; Robustness / Detection
A Survey of Monte Carlo Tree Search Methods [32]	2012	C3	General	Not specified
Rolling Horizon Evolution versus Tree Search for Navigation in Single-Player Real-Time Games [19]	2013	C1, C4	Navigation / Grid maps	Runtime; Search effort
SEARCH GAMES: LITERATURE AND SURVEY [16]	2015	C6	Stochastic / Imperfect information	Robustness / Detection

Table A1. Cont.

Title	Year	Cluster(s)	Game Type	Key Metrics (families)
A Search Allocation Game with Private Information of Initial Target Position [58]	2015	C6	Stochastic / Imperfect information	Robustness / Detection
Simulation and Comparison of Efficiency in Pathfinding Algorithms in Games [42]	2015	C1	Navigation / Grid maps	Runtime; Path cost / Optimality
Mastering the Game of Go with Deep Neural Networks and Tree Search [5]	2016	C3, C5	Board games	Win rate / Elo; Search effort
Mastering the game of Go without human knowledge [22]	2017	C5, C3	Board games	Win rate / Elo; Search effort; Robustness / Detection
A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play [1]	2018	C2, C3, C5	Chess, Shogi, Go	Win rate / Elo; Search effort (simulations)
Guiding Monte Carlo Tree Search by Scripts in Real-Time Strategy Games [28]	2013	C3, C5	RTS	Win rate / Elo; Runtime; Search effort
Superhuman AI for multiplayer poker [33]	2019	C3, C6	Poker (Texas Hold'em)	Win rate / Elo; Expected value (bb/100)
Analysis of Statistical Forward Planning Methods in Pommerman [51]	2019	C3, C4	Not specified	Score / Reward; Win rate
Action Guidance with MCTS for Deep Reinforcement Learning [50]	2019	C3, C5	Atari (ALE)	Score / Reward
Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model [20]	2020	C5, C2, C3	General / Multi-domain	Score / Reward; Search effort; Robustness / Detection
Rolling Horizon Evolutionary Algorithms for General Video Game Playing [52]	2020	C4	Not specified	Score / Reward; Win rate
Deep Reinforcement Learning for General Game Playing [55]	2020	C3, C5	Not specified	Win rate / Elo
A Review on Informed Search Algorithms for Video Games Pathfinding [41]	C1	Navigation / Grid maps	Path length / path cost; Node expansions; Runtime	
Path Finding and Map Feature Learning in RTS Games with Partial Observability [17]	2021	C5, C1, C6	RTS	Robustness / Detection; Runtime

Table A1. Cont.

Title	Year	Cluster(s)	Game Type	Key Metrics (families)
Solving Partially Observable Stochastic Shortest-Path Games [39]	2021	C6	General	Optimality gap; Convergence guarantees
Mastering Atari Games with Limited Data [54]	C3, C5	Atari (ALE)	Score / Reward; Sample efficiency (100k interactions)	
A Systematic Review and Analysis of Intelligence-Based Pathfinding Algorithms in Video Games [44]	2022	C1, C4, C5	General / Multi-domain	Not specified
Supervised and Reinforcement Learning from Observations in Reconnaissance Blind Chess [38]	2022	C5, C6	Chess (imperfect info variant)	Win rate / Elo
Player of Games [2]	2022	C1, C2, C6	Chess, Go, Poker (Texas Hold'em)	Win rate / Elo
Mastering the game of Stratego with model-free multiagent reinforcement learning [57]	2022	C5, C6	Stratego	Win rate / Elo
Generalized Entropy and Solution Information for Measuring Puzzle Difficulty [53]	2023	C4	Not specified	Robustness / Detection
Monte Carlo Tree Search: a review of recent modifications and applications [49]	2023	C3	Not specified	Not specified
Deep Reinforcement Learning in Real-Time Strategy Games: A Systematic Literature Review [35]	2024	C5, C4	RTS	Robustness / Detection
Optimizing Monte Carlo Tree Search for Parallel Computing on GPUs [48]	2025	C3	Not specified	Runtime; Search effort
LLM-A*: Large Language Model Enhanced Incremental Heuristic Search on Path Planning [45]	2024	C1, C5	Navigation / Grid maps	Path length /path cost; Node expansions; Runtime
Monte Carlo Tree Search with Boltzmann Exploration [46]	2024	C3	Go	Win rate / Elo; Rollouts / simulations
AlphaZero Neural Scaling and Zipf's Law: a Tale of Board Games and Power Laws [56]	2024	C2, C5	Chess, Go	Win rate / Elo; Compute scaling (FLOPs)

Table A1. Cont.

Title	Year	Cluster(s)	Game Type	Key Metrics (families)
Grandmaster-Level Chess Without Search [34]	2024	C2, C5	Chess	Win rate / Elo
Enhancing Deep Reinforcement Learning for Scale Flexibility in Real-Time Strategy Games [21]	2025	C5, C1	RTS	Robustness / Detection
Partially Observable Off-Switch Games [30]	2025	C6	Stochastic / Imperfect information	Robustness / Detection
Monte Carlo Tree Search for Knowledge Graph Reasoning [60]	2025	C3	Not specified	Search effort
Novelty in Monte Carlo Tree Search [47]	2025	C3	Not specified	Score / Reward; Search effort (simulations)
R2PS: Worst-Case Robust Real-Time Pursuit Strategies under Partial Observability [40]	2025	C6	Search / pursuit evasion games	Capture time; Robustness guarantees

References

1. Silver, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **2018**, *362*, 1140–1144.
2. Schmid, M.; Moravcik, M.; Burch, N.; Kadlec, R.; Davidson, J.; Waugh, K.; Bard, N.; Timbers, F.; Lanctot, M.; Holland, Z.; et al. Player of Games. *Science Advances* **2022**.
3. Shannon, C.E. XXII. Programming a computer for playing chess. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1950**, *41*, 256–275.
4. Campbell, M.; Hoane Jr, A.J.; Hsu, F.h. Deep blue. *Artificial intelligence* **2002**, *134*, 57–83.
5. Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; Van Den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; et al. Mastering the game of Go with deep neural networks and tree search. *nature* **2016**, *529*, 484–489.
6. Knuth, D.E.; Moore, R.W. An analysis of alpha-beta pruning. *Artificial intelligence* **1975**, *6*, 293–326.
7. Plaat, A.; Schaeffer, J.; Pijls, W.; De Bruin, A. Best-first fixed-depth minimax algorithms. *Artificial Intelligence* **1996**, *87*, 255–293.
8. Coulom, R. Efficient selectivity and backup operators in Monte-Carlo tree search. In Proceedings of the International conference on computers and games. Springer, 2006, pp. 72–83.
9. Kocsis, L.; Szepesvári, C. Bandit based monte-carlo planning. In Proceedings of the European conference on machine learning. Springer, 2006, pp. 282–293.
10. Korf, R.E. Depth-first iterative-deepening: An optimal admissible tree search. *Artificial intelligence* **1985**, *27*, 97–109.
11. Orkin, J. Three states and a plan: the AI of FEAR. In Proceedings of the Game developers conference. CMP Game Group SanJose, California, 2006, Vol. 2006, p. 4.
12. Koopman, B.O. *Search and screening*; Number no. 56 in OEG report, Operations Evaluation Group, Office of the Chief of Naval Operations, Navy Dept: Washington, D.C, 1946.
13. Koopman, B.O. The theory of search: III. The optimum distribution of searching effort. *Operations research* **1957**, *5*, 613–626.
14. Washburn, A.R. Search for a Moving Target: The FAB Algorithm. *Operations Research* **1983**, *31*, 739–751. <https://doi.org/10.1287/opre.31.4.739>.
15. Cowling, P.I.; Powley, E.J.; Whitehouse, D. Information set monte carlo tree search. *IEEE Transactions on Computational Intelligence and AI in Games* **2012**, *4*, 120–143.

16. Hohzaki, R. SEARCH GAMES : LITERATURE AND SURVEY. *Journal of The Operations Research Society of Japan* **2016**, *59*, 1–34.
17. Pan, H. Pathfinding and Map Feature Learning in RTS Games with Partial Observability. In Proceedings of the AIIDE Workshops, 2021.
18. Korf, R.E. Real-time heuristic search. *Artificial intelligence* **1990**, *42*, 189–211.
19. Perez, D.; Samothrakis, S.; Lucas, S.; Rohlfshagen, P. Rolling horizon evolution versus tree search for navigation in single-player real-time games. In Proceedings of the Proceedings of the 15th annual conference on Genetic and evolutionary computation, 2013, pp. 351–358.
20. Schrittwieser, J.; Antonoglou, I.; Hubert, T.; Simonyan, K.; Sifre, L.; Schmitt, S.; Guez, A.; Lockhart, E.; Hassabis, D.; Graepel, T.; et al. Mastering atari, go, chess and shogi by planning with a learned model. *Nature* **2020**, *588*, 604–609.
21. Lemos, M.L.H.D.; Tavares, A.R.; Marcolino, L.S.; Chaimowicz, L.; et al. Enhancing deep reinforcement learning for scale flexibility in real-time strategy games. *Entertainment Computing* **2025**, *52*, 100843.
22. Silver, D.; Schrittwieser, J.; Simonyan, K.; Antonoglou, I.; Huang, A.; Guez, A.; Hubert, T.; Baker, L.; Lai, M.; Bolton, A.; et al. Mastering the game of go without human knowledge. *nature* **2017**, *550*, 354–359.
23. Dijkstra, E.W. A note on two problems in connexion with graphs. In *Edsger Wybe Dijkstra: his life, work, and legacy*; 2022; pp. 287–290.
24. Harabor, D.; Grastien, A. The JPS pathfinding system. In Proceedings of the Proceedings of the International Symposium on Combinatorial Search, 2012, Vol. 3, pp. 207–208.
25. Hart, P.E.; Nilsson, N.J.; Raphael, B. A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics* **1968**, *4*, 100–107.
26. Koenig, S.; Likhachev, M. D* lite. In Proceedings of the Eighteenth national conference on Artificial intelligence, 2002, pp. 476–483.
27. Botea, A.; Müller, M.; Schaeffer, J. Near optimal hierarchical path-finding. *J. Game Dev.* **2004**, *1*, 1–30.
28. Yang, Z.; Ontanón, S. Guiding Monte Carlo tree search by scripts in real-time strategy games. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 2019, Vol. 15, pp. 100–106.
29. Schaeffer, J. The history heuristic and alpha-beta search enhancements in practice. *IEEE transactions on pattern analysis and machine intelligence* **2002**, *11*, 1203–1212.
30. Garber, A.; Subramani, R.; Luu, L.; Bedaywi, M.; Russell, S.; Emmons, S. The partially observable off-switch game. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2025, Vol. 39, pp. 27304–27311.
31. Stentz, A. Optimal and efficient path planning for partially-known environments. In Proceedings of the Proceedings of the 1994 IEEE international conference on robotics and automation. IEEE, 1994, pp. 3310–3317.
32. Browne, C.B.; Powley, E.; Whitehouse, D.; Lucas, S.M.; Cowling, P.I.; Rohlfshagen, P.; Tavener, S.; Perez, D.; Samothrakis, S.; Colton, S. A survey of Monte Carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games* **2012**, *4*, 1–43.
33. Brown, N.; Sandholm, T. Superhuman AI for multiplayer poker. *Science* **2019**, *365*, 885–890.
34. Ruoss, A.; Delétang, G.; Li, S.; Czarnecki, W.M.; Gretton, A.; Vinyals, O. Grandmaster-Level Chess Without Search. *arXiv preprint arXiv:2402.04494* **2024**.
35. Barros e Sa, G.C.; Madeira, C.A.G. Deep reinforcement learning in real-time strategy games: a systematic literature review: GC Barros e Sá et al. *Applied Intelligence* **2025**, *55*, 243.
36. Eagle, J.N.; Yee, J.R. An optimal branch-and-bound procedure for the constrained path, moving target search problem. *Operations research* **1990**, *38*, 110–114.
37. Anderson, E.J.; Aramendia, M. A linear programming approach to the search game on a network with mobile hider. *SIAM journal on control and optimization* **1992**, *30*, 675–694.
38. Bertram, T.; Fürnkranz, J.; Müller, M. Supervised and reinforcement learning from observations in reconnaissance blind chess. In Proceedings of the 2022 IEEE Conference on Games (CoG). IEEE, 2022, pp. 311–318.
39. Tomášek, P.; Horák, K.; Aradhye, A.; Bošanský, B.; Chatterjee, K. Solving partially observable stochastic shortest-path games. In Proceedings of the Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI), 2021, pp. 4182–4189.
40. Lu, R.; Shi, R.; Zhu, Y.; Zhao, D. R2PS: Worst-Case Robust Real-Time Pursuit Strategies under Partial Observability. *arXiv preprint arXiv:2511.17367* **2025**.

41. Kapi, A.Y. A Review on Informed Search Algorithms for Video Games Pathfinding. *International Journal of Advanced Trends in Computer Science and Engineering* **2020**, *9*, 2589–2598.
42. Noori, A.; Moradi, F. Simulation and Comparison of Efficiency in Pathfinding algorithms in Games. *Ciência e Natura* **2015**, *37*, 230–238.
43. Harabor, D.; Grastien, A. Online graph pruning for pathfinding on grid maps. In Proceedings of the Proceedings of the AAAI conference on artificial intelligence, 2011, Vol. 25, pp. 1114–1119.
44. Lawande, S.R.; Jasmine, G.; Anbarasi, J.; Izhar, L.I. A systematic review and analysis of intelligence-based pathfinding algorithms in the field of video games. *Applied Sciences* **2022**, *12*, 5499.
45. Meng, S.; Wang, Y.; Yang, C.F.; Peng, N.; Chang, K.W. LLM-A*: Large Language Model Enhanced Incremental Heuristic Search on Path Planning. In Proceedings of the Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 2024, pp. 10246–10263.
46. Painter, M.; Baioumy, M.; Hawes, N.; Lacerda, B. Monte Carlo Tree Search with Boltzmann Exploration. In Proceedings of the Advances in Neural Information Processing Systems, 2024, Vol. 37.
47. Baier, H.; Kaisers, M. Novelty in Monte Carlo Tree Search. *IEEE Transactions on Games* **2025**.
48. Klesk, P. MCTS-NC: A thorough GPU parallelization of Monte Carlo Tree Search implemented in Python via numba.cuda. *SoftwareX* **2025**, *30*, 102139. <https://doi.org/https://doi.org/10.1016/j.softx.2025.102139>.
49. Świechowski, M.; Godlewski, K.; Sawicki, B.; Mańdziuk, J. Monte Carlo tree search: a review of recent modifications and applications. *Artificial Intelligence Review* **2023**, *56*, 2497–2562.
50. Kartal, B.; Hernandez-Leal, P.; Taylor, M.E. Action Guidance with MCTS for Deep Reinforcement Learning. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 2019, Vol. 15, pp. 153–159.
51. Perez-Liebana, D.; Gaina, R.D.; Drageset, O.; Ilhan, E.; Balla, M.; Lucas, S.M. Analysis of statistical forward planning methods in Pommerman. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 2019, Vol. 15, pp. 66–72.
52. Gaina, R.D.; Devlin, S.; Lucas, S.M.; Pérez-Liebana, D. Rolling horizon evolutionary algorithms for general video game playing. *IEEE Transactions on Games* **2020**, *14*, 232–242.
53. Shen, J.; Sturtevant, N.R. Generalized entropy and solution information for measuring puzzle difficulty. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 2024, Vol. 20, pp. 117–126.
54. Ye, W.; Liu, S.; Kurutach, T.; Abbeel, P.; Gao, Y. Mastering Atari Games with Limited Data. In Proceedings of the Advances in Neural Information Processing Systems, 2021, Vol. 34, pp. 25476–25488.
55. Goldwasser, A.; Thielscher, M. Deep reinforcement learning for general game playing. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2020, Vol. 34, pp. 1701–1708.
56. Neumann, O.; Gros, C. AlphaZero Neural Scaling and Zipf's Law: a Tale of Board Games and Power Laws. *arXiv preprint arXiv:2412.11979* **2024**.
57. Perolat, J.; De Vylder, B.; Hennes, D.; Taz, E.; Strub, F.; Meunier, V.; Lanctot, M.; Munos, R.; Gruslys, A.; Lockhart, E.; et al. Mastering the game of Stratego with model-free multiagent reinforcement learning. *Science* **2022**, *378*, 990–996.
58. Hohzaki, R.; Joo, K. A search allocation game with private information of initial target position. *Journal of the Operations Research Society of Japan* **2015**, *58*, 353–375.
59. Isaacs, R. *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*; Courier Corporation, 1999.
60. Liu, L. Monte carlo tree search for graph reasoning in large language model agents. In Proceedings of the Proceedings of the 34th ACM International Conference on Information and Knowledge Management, 2025, pp. 4966–4970.
61. Brown, G.; Carlyle, M.; Diehl, D.; Kline, J.; Wood, K. A two-sided optimization for theater ballistic missile defense. *Operations research* **2005**, *53*, 745–763.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.