International Journal of Advances in Engineering and Computer Science
Volume 1, Issue 1, July-September 2025
(An Academians Publishers of Research & Academic Resources)

ISSN: XXXX-XXXX

Original Research Article

Received: 22-07-2025 **Accepted**: 28-08-2025 **Published**: 30-09-2025

Exploring Bounded Rationality Through Computational Intelligence: A Comprehensive Review

Magda Hussaon*[1]

Abstract: Bounded rationality departs from the traditional economic assumption of fully rational agents by highlighting the impact of cognitive and computational constraints on human decisions. This review synthesizes recent progress in computational intelligence that addresses how to model and enhance rationality within the bounds of these limitations. We discuss foundational theories, including Herbert Simon's bounded rationality and Ariel Rubinstein's algorithmic framework, alongside contemporary computational approaches involving heuristic search, machine learning, and multi-agent systems. Special attention is given to methods that bridge psychology, economics, and artificial intelligence, offering realistic models of decision-making and examining their consequences for economics, behavioral finance, and autonomous system design. The review concludes by identifying future research opportunities for creating more adaptable and robust agents capable of navigating complex environments under limited information and computational resources.

Keywords: *Bounded Rationality, Computational Rationality, Heuristics, Satisficing, Multi-Agent Systems, Reinforcement Learning, Behavioral Economics, Cognitive Models, Artificial Intelligence.*

[1]*Independent Scholar*

Introduction

The traditional economic model assumes that agents have complete information and infinite computational ability, allowing for entirely rational choices. In opposition to this ideal, Herbert Simon introduced the concept of bounded rationality, suggesting that actual decision-makers operate within cognitive and environmental limitations that limit their capacity to optimize (Simon, 1957). This essential change requires different models that encompass heuristics, satisficing, and adaptive reasoning in uncertain conditions

Ariel Rubinstein (1986) was the first to develop formal economic models that incorporate bounded rationality by depicting agents as algorithms limited by computational constraints, thereby engaging in strategic interactions. These viewpoints emphasize that assessing rationality should consider the computational resources available and the heuristics employed, situating economic actions within broader cognitive frameworks.

In recent decades, notable advancements have been made in computational intelligence, encompassing artificial intelligence (AI), machine learning, and heuristic search, which offer mechanistic structures for implementing bounded rationality. These advancements enable the creation of agents that demonstrate effective rationality: computationally viable approximations of optimal rational behavior suited to particular contexts.

This examination systematically investigates the convergence of bounded rationality and

Available Online:

*Corresponding Author: Magda Hussaon

computational intelligence, aiming to clarify theoretical foundations, assess methodological advancements, and pinpoint applications across economic modeling, behavioral sciences, and autonomous decision-making systems.

Foundational Theories of Bounded Rationality

Table 1: Foundational Theories of Bounded Rationality

Theory & Proponent	Core Concept	Key Reference	Distinctive Feature
Satisficing (Herbert Simon, 1957)	Agents settle for "good enough" solutions using aspiration-level heuristics rather than optimizing fully.	Simon (1957)	Balances aspiration levels and computational limits
Algorithmic Model (Rubinstein, 1986)	Agents modeled as algorithms with finite computational steps in strategic interactions.	Rubinstein (1986)	Formalizes cost of computation in decision processes
Prospect Theory (Kahneman & Tversky, 1979)	Humans evaluate gains and losses asymmetrically, deviating from expected-value maximization.	Kahneman (2011)	Captures framing effects, loss aversion
Dual-Process (Evans & Stanovich, 2013)	Decision making arises from interaction of fast, intuitive (System 1) and slow, deliberative (System 2).	Evans & Stanovich (2013)	Highlights interplay between automatic and controlled cognition

Herbert Simon and Satisficing

Simon's foundational research highlighted the challenges of optimization in intricate settings because of constraints in information processing, focus, and time (Simon, 1957). Instead, agents use satisficing, aiming for satisfactory—not always optimal—solutions by following heuristic search directed by aspiration levels. This procedure represents a harmony between ambition and practicality.

Rubinstein's Computational Rationality Model

Rubinstein (1986) proposed a model in which agents act as computational entities participating in strategic interactions limited by finite resources and restrictions. This framework clarifies how computational expenses affect strategy choice and belief development, redirecting rationality assessment to a procedural viewpoint that highlights algorithmic practicality.

Computational Intelligence Models of Effective Rationality

Table 2: Computational Intelligence Approaches to Effective Rationality

Method	Description	Applications	Strengths	Limitations
Heuristic Search	Approximate search (e.g., genetic algorithms, simulated annealing) guided by heuristics.	Combinatorial optimization, scheduling	Balances exploration/exploitation	May converge to local optima
Reinforcement Learning	Agents learn policies through trial-and-error interactions and reward signals.	Robotics, game playing, resource allocation	Adapts to dynamic environments	Requires large data, can be unstable
Multi-Agent Systems	Multiple bounded agents coordinate or compete within shared environments.	Traffic control, market simulations	Models emergent, collective behaviors	Complexity grows with agent count
Neural Heuristic Hybrids	Neural networks learn heuristics or value functions to guide search processes.	Real-time decision support, game AI	Learns problem-specific heuristics	Training costs and interpretability issues

Heuristic and Metaheuristic Search Algorithms

Methodological strategies, such as genetic algorithms, simulated annealing, and Monte Carlo tree search, have been employed to replicate bounded rationality, enabling agents to explore extensive solution domains through approximate, heuristic-driven search methods (Russell & Norvig, 2021). These techniques achieve a compromise between exploration and exploitation while adhering to computational limits

Machine Learning and Adaptive Behavior

Reinforcement learning and deep learning architectures enable agents to learn optimal policies from interaction without explicit programming, adapting flexibly to complex environments (Sutton & Barto, 2018). By integrating cost-aware exploration, these models instantiate practical bounded rationality.

Multi-agent Systems and Collective Rationality

Distributed AI systems model collective bounded rationality, where multiple agents with limited knowledge and computational power coordinate or compete within shared environments (Wooldridge, 2009). These frameworks elucidate how effective rationality emerges from localized interactions.

Behavioral and Economic Implications

Decision Making Under Uncertainty

Computational bounded rationality models illuminate the heuristics that underpin human economic behavior, including framing effects, loss aversion, and delayed gratification, providing mechanistic explanations that augment psychological theories (Kahneman, 2011). This promotes interdisciplinary collaboration, improving behavioral economics

Market Dynamics and Organizational Decision Processes

Bounded rationality limits organizational choices and market efficiency, highlighting phenomena such as information asymmetry, herd behavior, and institutional rigidity (Cyert & March, 1963). Computational modeling enables the simulation of intricate adaptive systems that represent these dynamics

Challenges and Future Directions

Although computational models enhance our grasp of effective rationality, obstacles persist in matching model complexity with interpretability, adapting to real-world issues, and incorporating diverse cognitive processes. Future research objectives should focus on developing benchmark datasets that mirror cognitively realistic constraints, improving human-AI interaction systems to facilitate bounded rational decision-making, and designing hybrid models that combine symbolic reasoning with sub-symbolic learning

Discussion

This review synthesizes the theory of bounded rationality and the practical advances enabled by computational intelligence. By combining these areas, decision-making research shifts from focusing only on ideal scenarios to exploring how people and machines actually make choices when faced with real-world limits. Here, I discuss what this means, the trade-offs involved, and the wider impact for both humans and AI.

6.1 The Paradigm Shift from Optimal to Effective Rationality

The most significant contribution of computational models is their ability to operationalize Simon's and Rubinstein's ideas. They provide a concrete answer to the question: "What does it mean to be rational when you cannot optimize?" The answer lies in the concept of **effective rationality**—achieving the best possible outcome given finite time, knowledge, and computational power. Models like metaheuristic search and reinforcement learning are not merely approximations of optimality; they are instantiations of a different kind of rationality, one that is context-dependent and resource-aware. This shifts the benchmark for success from finding a global optimum to robustly achieving satisfactory results across a wide range of scenarios, a criterion that often better matches real-world success.

6.2 The Dual Challenge of Fidelity and Scalability

A key challenge in this area is finding the right balance between models that closely reflect human thinking and those that can handle large, complex problems. Simple heuristics, such as the 'fast and frugal' rules from Gigerenzer and Todd (1999), are easy to understand and align with some human behaviors, but they may not adapt well to complex situations. On the other hand, advanced machine learning methods, such as deep reinforcement learning, can outperform humans in certain tasks; however, their decision-making is often difficult to explain. Moving forward, the field will likely focus on hybrid models that combine the clarity and human-like reasoning of symbolic approaches with the flexibility of machine learning. This way, we can build systems that not only work well but are also understandable and trustworthy.

6.3 Implications for the Principle of Rationality in AI and Economics

Advances in computational bounded rationality are transforming our understanding of intelligence in AI. Rather than aiming for one perfect form of intelligence, this view sees intelligence as having many sides, shaped by the environment and the resources available. This matches the idea of ecological rationality, where a strategy's value depends on the setting in which it is used.

In economics, these models provide a much-needed micro-foundation for behavioral phenomena. Rather than treating biases and heuristics as mere deviations from rationality, computational models can show how these behaviors emerge as adaptive responses to computational constraints. For instance, **loss aversion** in Prospect Theory can be reframed as a risk-management heuristic that is effective in environments where losses are more costly to recover from than equivalent gains are to build upon. This provides a mechanistic, rather than just a descriptive, account of economic behavior.

6.4 Ethical and Practical Considerations for Autonomous Systems

As more autonomous AI systems are deployed in areas such as finance, healthcare, and transportation, it is crucial to design them with real-world limitations in mind. No system has unlimited computing power, so we need to plan for these limits from the start. This means making clear choices about how to handle these boundaries.

1. **Making Trade-offs Explicit:** System designers must make conscious choices about how to allocate computational budget between speed, accuracy, and exploration.
2. **Ensuring Robust Satisficing:** Systems should be engineered to reliably find "good enough" solutions that are robust to uncertainty, rather than being fragile optimizers that fail catastrophically when their idealized assumptions are violated.
3. **Aligning with Human Bounds:** In human-AI collaboration, the AI's reasoning should be transparent and comprehensible within the bounds of human cognition to foster trust and effective teamwork.

6.5 Limitations and Avenues for Future Inquiry

While this field has made significant strides, several challenges remain. There is a lack of standardized **benchmark environments** for evaluating bounded rational agents that accurately reflect the resource constraints and complexities of real-world decision-making. Furthermore, most current models focus on a single agent or a homogeneous group; more research is needed on **heterogeneous multi-agent systems** where agents with different computational capabilities, goals, and cognitive architectures interact. Finally, bridging the gap between the neural implementation of cognitive control (e.g., O'Reilly & Frank, 2006) and abstract computational models remains a fertile ground for interdisciplinary research.

Conclusion

Computational intelligence provides robust methods to grasp and improve effective rationality within the limits of bounded rationality. Through the incorporation of algorithmic estimates, learning processes, and interactions among multiple agents, researchers can create decision-making models that better reflect human cognitive abilities. This review highlights the importance of multidisciplinary approaches to progress economic theory, artificial intelligence, and behavioral science, facilitating the creation of flexible, rational agents capable of operating in complex, uncertain environments

References

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Prentice-Hall.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Rubinstein, A. (1986). Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory*, 39(1), 83–96.

Simon, H. A. (1957). *Models of man: Social and rational*. Wiley.

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A modern approach* (4th ed.). Pearson.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

Wooldridge, M. (2009). *An introduction to multiagent systems* (2nd ed.). Wiley.

Ainslie, G. (1992). *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person*. Cambridge University Press.

Gigerenzer, G., & Todd, P. M. (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press.

Evans, J. St. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*(1), 1–20.

Boden, M. (2017). *AI: Its Nature and Future*. Oxford University Press.

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Prentice Hall.

Lou, Y., & Kadous, M. (2010). Learning bounded-rationality strategies in repeated games. In *Proceedings of AAMAS 2010*.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences, 43*.

Goel, V., & Dolan, R. J. (2004). The functional neuroanatomy of deductive and inductive reasoning. *Brain Research, 654*(1), 15–28.

Gigerenzer, G. (2008). *Rationality for Mortals: How People Cope with Uncertainty*. Oxford University Press.

Chater, N., & Oaksford, M. (2006). *The probabilistic mind: Prospects for Bayesian cognitive science*. Oxford University Press.

Tsang, E., & Lai, J. (2014). *Agent-based computational economics: An introduction*. Springer.

McCarthy, J. (2007). What is Artificial Intelligence? *Stanford Encyclopedia of Philosophy*.

O'Reilly, R., & Frank, M. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation, 18*(2), 283–328.

Sanfey, A. G. (2007). Social decision-making: Insights from game theory and neuroscience. *Science, 318*(5850), 598–602.