STUDIA UNIVERSITATIS BABEȘ-BOLYAI OECONOMICA
VOLUME 70, ISSUE 1, 2025, pp. 67-82
DOI: 10.2478/subboec-2025-0004

THE SECOND AXELROD TOURNAMENT: A MONTE CARLO EXPLORATION OF UNCERTAINTY ABOUT THE NUMBER OF ROUNDS IN ITERATED PRISONER'S DILEMMA

Gabriel POP*
Babeș-Bolyai University, Romania

Mircea MILENCIANU
Babeș-Bolyai University, Romania

Alexandra POP
Babeș-Bolyai University, Romania

ABSTRACT: Strategic decision-making in multi-agent interactions inside the Iterated Prisoner's Dilemma (IPD) is investigated in this work using Monte Carlo simulations. Building on Axelrod's work, we present a second-generation tournament with stochastic components, including unpredictable game lengths, to evaluate strategy adaptability and resilience. We analyze how uncertainty influences strategic performance by using a comparison between instances with fixed and uncertain times. We identify, using a descriptive approach, methods demonstrating important behavioral differences between deterministic and uncertain settings. The results provide understanding of adaptive learning, response dynamics, and strategic flexibility, so helping to build strong collaborative strategies for artificial intelligence and decision-making systems. Our results highlight the limitations of exclusively deterministic methods and suggest the necessity for adaptive approaches to improve long-term cooperative success.

JEL Classification: C70; C73; C79

Keywords: Prisoner's Dilemma; repeated games; Axelrod second Tournament; agent-based modeling; finite and infinite games

1. Introduction

An essential feature of multi-agent interactions is strategic decision-making, who game theory offers a strong structure for examining such choices. Being much investigated to understand cooperative and competitive behaviors in repeated

* Corresponding author. Address: Faculty of Economics and Business Administration, Babeş-Bolyai University, 58-60, Teodor Mihali Street, 400591, Cluj-Napoca, Romania.
E-mail: gabriel.pop@ubbcluj.ro

interactions, the Iterated Prisoner's Dilemma (IPD) is one of the most well-known models in this field of study (Chong et al., 2007). Robert Axelrod's innovative work revealed a collection of computational competitions evaluating the success of different approaches in the IPD, therefore generating important new perspectives on the development of collaboration.(Axelrod, 1984)

Building on this foundation, the current work investigates a second-generation Axelrod tournament (project, 2015) using improved computational frameworks and uncertainty factors to evaluate the adaptability and robustness of many agent strategies. Unlike the original tournament, which mostly concentrated on deterministic interactions, our work incorporates stochastic aspects, such unpredictable game lengths, to investigate how strategies function in non-deterministic situations.

Using Monte Carlo simulations, we methodically evaluate the performance of a diverse set of strategies under both fixed and uncertain conditions (Thomopoulos, 2013). We examine important indicators including win rates, utility differences, and adaptive resilience to find which tactics hold true when the game environment veers from strict determinism and which ones get vulnerable. The results establish the effects on perpetual success in repeated games of adaptive learning, a response dynamic, and strategic flexibility.

This work wants to develop and investigate a second-generation Axelrod tournament (Axelrod, 1984) including stochastic components, such unpredictable game lengths, so improving the realism of strategic interactions. It makes use of Monte Carlo simulations to evaluate agent adaptability of several approaches, therefore measuring agent resilience in uncertain environments. Finding important performance metrics—including win rates, utility variations, and adaptive resilience—that impact long-term strategic success under both deterministic and non-deterministic settings is a major focus.

Investigating how response dynamics and adaptive learning structures support cooperation or allow exploitation in iterative game environments is another objective. The study also compares conventional and most recently suggested approaches to evaluate their performance in handling uncertainty, therefore revealing important information on strategic resilience and adaptability (Chiong & Jankovic, 2008). Furthermore, it looks at how unpredictability in game parameters affects emergent strategy efficacy and how stochastic perturbations affect the stability and evolution of cooperative behaviors.

By using Monte Carlo methods to improve prediction accuracy and strategy evaluation, the study also seeks to maximize computational models for predicting recurrent interactions under uncertainty. At last, it aims to provide a basis for practical applications in artificial intelligence and decision-making by establishing rules for creating strong cooperative strategies that can resist different and changing strategic environment.

Through following these objectives, we want to improve our understanding of strategic decision-making in complex multi-agent interactions and provide a framework for future research on fostering cooperation in uncertain environments.

2. Theoretical background. Definitions. Notation

2. 1. Normal-form games

Normal-form games represent a fundamental model for studying strategic decisions. A normal-form game is defined as a tuple:

$$\boldsymbol{\Gamma} = (\boldsymbol{N}, \boldsymbol{A}, \boldsymbol{u}) \qquad (1)$$

where:

- $\boldsymbol{N}$ is a finite set of players, $|\boldsymbol{N}| = \boldsymbol{n}$, with $\boldsymbol{n} \geq \mathbf{2}$.
- $\boldsymbol{A_i}$ is the set of pure strategies for each player $\boldsymbol{i} \in \boldsymbol{N}$.
- $\boldsymbol{A} = \prod_{\boldsymbol{i} \in \boldsymbol{N}} \boldsymbol{A_i}$ is the Cartesian product of all strategies.
- $\boldsymbol{u_i}: \boldsymbol{A} \to \mathbb{R}$ is the utility function, assigning a numerical outcome to each strategy profile.

Games where players make decisions at the same time without knowing what other players are thinking work well with this structure.(Ben-Porath, 1990)

A Nash equilibrium is a fundamental concept in game theory, representing a strategy profile where no player can gain by unilaterally deviating from their strategy. Formally, a strategy profile (s_i^*, s_{-i}^*) a Nash equilibrium if:

$$\boldsymbol{u_i}(\boldsymbol{s_i^*}, \boldsymbol{s_{-i}^*}) \geq \boldsymbol{u_i}(\boldsymbol{s_i}, \boldsymbol{s_{-i}^*}), \quad \forall \boldsymbol{s_i} \in \boldsymbol{A_i}. \qquad (2)$$

In other words, given the strategies chosen by the other players, each player's strategy is optimal, meaning that no player has an incentive to deviate. Nash equilibria exist in all finite games with mixed strategies (Nash, 1950) The concept is applicable to various strategic contexts, ranging from basic two-player games to complex multi-agent systems. In the context of normal-form games, Nash equilibria can be classified as either pure or mixed. A pure strategy Nash equilibrium occurs when all players choose a single, deterministic strategy, while a mixed strategy Nash equilibrium involves players randomizing over their available strategies according to a probability distribution. Finding Nash equilibria in complex games often requires computational tools. One common method involves the use of Nash mappings, which are mathematical constructions that transform the problem of finding equilibria into solving fixed-point equations. Nash mappings are essential for identifying equilibria, particularly in games with infinite or continuous strategy spaces (Zhou et al., 2011). Zhou et al. explain that Nash mappings rely on fixed-point theorems, such as Tychonov's fixed-point theorem, which guarantee the existence of fixed points under certain conditions (Tychonov, 1935). The mapping ensures that for any starting point in the strategy space, there is a corresponding equilibrium point at which no player can independently enhance their payoff. This theoretical framework is essential for examining games with extensive strategy spaces and continuous payoff functions. Algorithms like the Lemke–Howson algorithm and best-response dynamics are frequently used to calculate Nash equilibria. Best-reaction dynamics involve the iterative adjustment of each player's strategy to optimize their response based on the dominant strategies of the other players. This process may converge to a Nash equilibrium in numerous instances, although this outcome is not invariably assured.

The importance of Nash equilibria extends simple theoretical examination. Nash equilibria provide a paradigm for forecasting the results of strategic interactions in practical applications throughout economics, political science, and computer science. They are utilized to simulate competitiveness in the market, negotiation scenarios, and multi-agent decision-making in distributed systems.

2. 2. Dominant strategies and Iterative elimination

Dominant strategies play a crucial role in decision-making within strategic games. A strategy is considered dominant if it provides a player with a higher or equal payoff regardless of the actions of other players (Rothe, 2010) In Axelrod's tournaments, iterative elimination of dominated strategies explains why inflexible strategies, such as constant defection, perform poorly in uncertain environments. Akin explains that zero-determinant (ZDS) strategies allow players to fix the relationship between their payoffs and those of their opponents, which can either stabilize cooperation or enable exploitation (Akin, 2016) ZDS strategies were extended by Press and Dyson (2012) and play a significant role in the evolutionary dynamics of repeated games.

2. 3. The classic example: Prisoner's Dilemma

The Prisoner's Dilemma is a classic example of a normal-form game and is central to studying cooperation and defection between players. Each player must simultaneously choose between cooperation (C) or defection (D), with outcomes depending on the combination of their decisions.

The payoff matrix can be represented as:

Table 1. The payoff matrix

X/Y	C	D
C	R	S
D	T	P

Where:

$\boldsymbol{R}$ is the reward for mutual cooperation;
$\boldsymbol{P}$ is the punishment for mutual defection;
$\boldsymbol{T}$ is the temptation payoff when a player defects while the other cooperates;
$\boldsymbol{S}$ is the sucker's payoff for cooperating when the other defects.

These payoff relationships must satisfy:

$$\boldsymbol{T} > \boldsymbol{R} > \boldsymbol{P} > \boldsymbol{S} \quad \text{and} \quad \boldsymbol{2R} > \boldsymbol{T} + \boldsymbol{S}. \tag{3}$$

Alternatively, the payoff vectors for each player can be expressed as:

$$\boldsymbol{S_X} = (\boldsymbol{R}, \boldsymbol{S}, \boldsymbol{T}, \boldsymbol{P}) \quad \text{and} \quad \boldsymbol{S_Y} = (\boldsymbol{R}, \boldsymbol{T}, \boldsymbol{S}, \boldsymbol{P}). \tag{4}$$

The expected payoffs s_X and s_Y for players X and Y, given a probability distribution v over the four possible outcomes, are:

$$s_X = \langle v \cdot S_X \rangle \quad \text{and} \quad s_Y = \langle v \cdot S_Y \rangle. \tag{5}$$

In the Prisoner's Dilemma, strategy C corresponds to cooperation, where both players receive the reward R when they cooperate. Conversely, strategy D (defection) leads to the temptation payoff T for the defector and the sucker's payoff S for the cooperating player. The mutual defection outcome results in the punishment payoff P for both players.

The condition $2R > T + S$ implies that mutual cooperation provides a payoff higher than splitting the total rewards of outcomes where one player defects and the other cooperates. The cooperative outcome (C, C) is a Pareto optimum, meaning that no other outcome can make one player better off without making the other worse off.

However, players face a dilemma due to the dominance of strategy D. Regardless of what the other player chooses, defection yields a higher payoff for the defector. As a result, both players rationally choose D, leading to the (D, D) outcome with a suboptimal payoff P for each.

In real-world scenarios, additional factors, such as reputation or external enforcement mechanisms, can modify these payoffs to promote cooperation. For instance, if one player defects and causes harm to the other, retaliation may reduce the desirability of defection below R. Anticipation of such consequences may encourage both players to honor agreements to cooperate.

2. 4. Repeated Prisoner's Dilemma

In a repeated version of the Prisoner's Dilemma, players encounter each other multiple times. The game can either have a fixed number of rounds (finite horizon) or continue indefinitely (infinite horizon). The difference in the horizon fundamentally affects the strategies employed by players.

In finitely repeated games, players know when the game will end, which often results in backward induction. Since defection is dominant in the final round, players reason backward to defect in all previous rounds (Myerson, 1991).

When the game is infinitely repeated or has an unknown ending, cooperation can emerge as a rational strategy. The incentive to cooperate arises from the potential future benefits of mutual cooperation. Two main concepts for evaluating payoffs in infinitely repeated games are:

1. Average reward: Given an infinite sequence of payoffs $r_i^1, r_i^2, \ldots$ for player i, the average reward is:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} r_i(j). \tag{6}$$

2. Discounted reward: Given an infinite sequence of payoffs $r_i^1, r_i^2, \ldots$ for player i, and a discount factor $0 \le \beta \le 1$, the discounted reward is:

$$\lim_{N\to\infty} \sum_{j=1}^{N} \beta^j r_i(j). \quad (7)$$

Strategies like Tit for Tat, which involves initial collaboration followed by imitating the opponent's previous action, effectively maintain cooperation in repeated games (Axelrod & Hamilton, 1981).

Tit for Tat is considered a Nash equilibrium under certain conditions, particularly when the discount factor is high enough, making future payoffs valuable.

The folk theorem provides further insight into repeated games, stating that if players are sufficiently patient, a wide range of outcomes, including mutual cooperation, can be sustained as Nash equilibria (Leyton-Brown & Shoham, 2008).

Good strategies for the iterated Prisoner's Dilemma have been characterized by Akin (2016) as those that stabilize cooperation and ensure Nash equilibria. When both players use a good strategy, neither player can gain by unilaterally changing their strategy. Markov strategies, including memory-one plans, play a central role in this context. For example, Tit for Tat is a memory-one strategy where a player's response depends on the opponent's previous action.

Meng Chen-Lu (2001) highlights that Tit for Tat and its variations, like Generous Tit for Tat, are essential for promoting ongoing cooperation in iterated games, in particular uncertain context. Nash mappings help identify stable strategies in complex situations, providing players with opportunities to optimize their long-term behaviour (Zhou et al., 2011).

Nachbar argues that through progressive learning and feedback-based adjustments, players can converge to Nash equilibria even when their initial strategies are suboptimal (Nachbar, 2005).

2. 5. Strategies in determined vs. uncertain environments

An essential component in analyzing repeated games is the distinction between deterministic and uncertain contexts. In determined environments, where the game's duration is known in advance, dominating strategies could be successful as participants anticipate the conclusion. On the other hand, in uncertain environments, adaptive strategies like Tit for Tat or exploration-exploitation models are generally more robust (Leonardos Stefanos, 2012).

Meng (2001) emphasizes that in complex environments with a large number of players or unpredictable interactions, an agent's performance depends on its ability to adapt. Adaptive algorithms, such as genetic algorithms, are crucial for optimizing behaviour in the face of diverse opponent strategies.

Ben-Porath (1990) highlights the computational complexity of establishing optimal response algorithms in recurrent games with mixed strategies. Computing and verifying optimal responds are frequently non-polynomial problems; however, they can become polynomial when the support size of mixed strategies is restricted.

2. 6. Axelrod's Second Tournament

Axelrod's second tournament extended the scope of his initial experiment by incorporating a larger set of strategies and introducing new conditions such as uncertainty regarding the number of game rounds. This allowed for a more comprehensive evaluation of strategies in the Iterated Prisoner's Dilemma (IPD) and provided deeper insights into the mechanisms that drive cooperation and competition (Axelrod, 1984) Unlike the first tournament, where Tit-for-Tat dominated due to its simplicity and reciprocity, the second tournament tested both reciprocal and adaptive strategies that adjusted dynamically to their opponents' behavior (Axelrod & Hamilton, 1981)

Key aspects of strategy that contribute to success in frequent experiences are highlighted in theoretical studies of Axelrod's tournaments. Among the key concepts that Axelrod's study has demonstrated are:

(i) **Niceness:** Strategies that avoid initial defection generally provide superior long-term payoffs. (Axelrod, 1984)
(ii) **Provocability:** Effective strategies must respond to defections in a way that discourages exploitation.
(iii) **Forgiveness:** The capacity to reestablish collaboration following the punishment of a defection helps in preventing extensive cycles of payback.
(iv) **Clarity:** Easily interpretable strategies allow opponents to predict the consequences of their actions, stabilizing interactions

In instances with a fixed number of rounds, inflexible strategies like Tit-for-Two-Tats, which delay retaliation until two successive defections occur, shown robust success. However, their inherent a lack in responding to defection made them especially vulnerable to manipulate by adaptive tactics like as Second by Tester. The latter exploited this tolerance by strategically defecting in specific cases, aware that immediate punishment would be missing. This opportunistic behavior enabled adaptive methods to achieve a greater cumulative payoff by exploiting the predictable responses of more rigid adversaries. Similarly, Grudger, which permanently punishes any opponent who defects, struggled in dynamic environments where adversaries could introduce strategic resets that neutralized its punitive approach (Tutzauer, 2007). On the other hand, strategies that adapted to changing conditions, such as Second by Tester and Second by Gladstein, thrived under uncertainty. These strategies adjusted their responses based on game history, allowing them to maximize payoffs in environments where the end-game was unpredictable. Vincent & Fryer (2021) further demonstrated that FSM-based strategies, which evolve through algorithmic optimization, could outperform even established heuristics like Tit-for-Tat by fine-tuning their behavior to different types of opponents.

Axelrod's second tournament findings, along with later theoretical improvements, highlight that adaptation is the essential factor for success in unpredictable situations. Strategies that strictly follow predetermined rules, such as Grudger or Tit-for-Two-Tats, are vulnerable to exploitation by adversaries capable of adapting their behavior (Axelrod, 1984).

3. Our approach to Monte Carlo simulations

3. 1. Definition of a Supergame

A supergame is defined as the repeated execution of N iterations of a simultaneous one-shot game (in our case, the Prisoner's Dilemma). The total utility is computed as the discounted sum of utilities across all one-shot games:

$$\boldsymbol{U}_{\text{total}} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{\beta}^{\boldsymbol{i-1}} \boldsymbol{u_i}, \tag{8}$$

where u_i is the utility at stage i and β is the discount factor with $0 \leq \beta \leq 1$.

Simulation of agent interactions

In a confrontation between two strategies or agents (e.g., Tit-for-Tat vs Tit-for-Two-Tats), the agents play K iterations of a supergame:

$\mathbf{vector_1}$**:** The proportion of wins, losses, and ties for each agent over $\boldsymbol{K}$ supergames.
$\mathbf{vector_2}$**: T**he average payoffs for each agent across $\boldsymbol{K}$ supergames.

To simulate uncertainty, the random variable $\boldsymbol{N}$ is generated from a distribution with mean $\boldsymbol{\mu_N}$ and standard deviation $\boldsymbol{\sigma_N}$. The distribution $\boldsymbol{\mathcal{D}}$ can be:

Normal distribution:

$$\boldsymbol{N} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu_N}, \boldsymbol{\sigma_N}) \tag{9}$$

where $\boldsymbol{\mu_N}$ is the mean and $\boldsymbol{\sigma_N}$ is the standard deviation.

Uniform distribution:

$$N \sim \mathcal{U}a,b \tag{10}$$

where the mean and standard deviation are given by the formulas:

$$\boldsymbol{\mu_N} = \frac{\boldsymbol{a+b}}{\boldsymbol{2}} \quad (11)$$
$$\boldsymbol{\sigma_N} = \frac{\boldsymbol{b-a}}{\sqrt{\boldsymbol{12}}} \quad (12)$$

3. 2. Comparison of fixed and uncertain cases

For each $\boldsymbol{k}$ in $\{\boldsymbol{1}, \boldsymbol{2}, \dots, \boldsymbol{K}\}$, the uncertain supergame is played by randomly selecting $\boldsymbol{N'}$ from the distribution in Equation (2). The corresponding $\mathbf{vector_1}$ and $\mathbf{vector_2}$ are computed and compared to the results from the fixed case.

Experimental setup

We select a set of $\boldsymbol{m}$ agents (strategies), denoted as $\text{Agent}_{\mathbf{1}}, \text{Agent}_{\mathbf{2}}, \dots, \text{Agent}_{\boldsymbol{m}}$. The results are stored as:

Case I: Fixed $\boldsymbol{N}$. Results are stored in a matrix $\boldsymbol{M}_{\text{case\_I}}$, where each entry above the diagonal represents $\mathbf{vector_1}$ and $\mathbf{vector_2}$ for the bilateral supergame between $\text{Agent}_{\boldsymbol{i}}$ and $\text{Agent}_{\boldsymbol{j}}$.

Case II: Uncertain $\boldsymbol{N}$. Results are stored in a matrix $\boldsymbol{M}_{\text{case\_II}}$ with the same structure as $\boldsymbol{M}_{\text{case\_I}}$.

Calculating distance metrics

The following metrics are used to measure discrepancies between Case I (fixed number of stages) and Case II (uncertain number of stages).

Discrepancy for "won" (victories)

$$\boldsymbol{D}_{\text{won}}(\boldsymbol{i},\boldsymbol{j}) = |\text{won}_{\text{case I}}(\boldsymbol{i},\boldsymbol{j}) - \text{won}_{\text{case II}}(\boldsymbol{i},\boldsymbol{j})|$$

Where:

$\boldsymbol{D}_{\text{won}}(\boldsymbol{i},\boldsymbol{j})$ represents the absolute difference between the number of victories for the pair of agents $\left(\text{Agent}_{\boldsymbol{i}}, \text{Agent}_{\boldsymbol{j}}\right)$ in the two cases.

The results are sorted in descending order to select the pairs of agents with the largest discrepancies, up to a threshold threshold $= \mathbf{10}$.

Discrepancy for "lost" (losses)

$$\boldsymbol{D}_{\text{lost}}(\boldsymbol{i},\boldsymbol{j}) = |\text{lost}_{\text{case I}}(\boldsymbol{i},\boldsymbol{j}) - \text{lost}_{\text{case II}}(\boldsymbol{i},\boldsymbol{j})|$$

Where:

$\boldsymbol{D}_{\text{lost}}(\boldsymbol{i},\boldsymbol{j})$ represents the absolute difference between the number of losses for the pair of agents $\left(\text{Agent}_{\boldsymbol{i}}, \text{Agent}_{\boldsymbol{j}}\right)$ in the two cases.

The results are sorted in descending order to select the pairs of agents with the largest discrepancies, up to a threshold threshold $= \mathbf{10}$.

Notation:

- $\text{won}_{\text{case I}}(\boldsymbol{i},\boldsymbol{j})$: the number of victories for the pair $\left(\text{Agent}_{\boldsymbol{i}}, \text{Agent}_{\boldsymbol{j}}\right)$ in case I (fixed number of stages).
- $\text{won}_{\text{case II}}(\boldsymbol{i},\boldsymbol{j})$: the number of victories for the pair $\left(\text{Agent}_{\boldsymbol{i}}, \text{Agent}_{\boldsymbol{j}}\right)$ in case II (uncertain number of stages).
- $\text{lost}_{\text{case I}}(\boldsymbol{i},\boldsymbol{j})$: the number of losses for the pair $\left(\text{Agent}_{\boldsymbol{i}}, \text{Agent}_{\boldsymbol{j}}\right)$ in case I.
- $\text{lost}_{\text{case II}}(\boldsymbol{i},\boldsymbol{j})$: the number of losses for the pair $\left(\text{Agent}_{\boldsymbol{i}}, \text{Agent}_{\boldsymbol{j}}\right)$ in case II.

Axelrod's tournament framework provides a valuable platform for studying the interactions and strategies of agents in repeated games, particularly the Iterated Prisoner’s Dilemma (IPD). This subsection focuses on the qualitative behavior of specific agent pairs under conditions of uncertainty about the number of game stages. By introducing uncertainty, we investigate whether strategies originally developed for fixed-stage environments demonstrate adaptability or vulnerabilities.

4. Results. Some relevant examples explained.

Second by Tester and Tit For 2 Tats

Behavior: Tit For 2 Tats, known for its leniency in tolerating a single defection, is exploited by the adaptive strategy of Second by Tester. Specifically, Second by Tester leverages the tolerance built into Tit For 2 Tats, which delays its retaliation and allows an opportunistic exploitation pattern.

Cause of discrepancy: The fundamental cause lies in the adaptability of Second by Tester. Unlike Tit For 2 Tats, which applies a rigid, predefined retaliation rule, Second by Tester dynamically adjusts its behavior to maximize payoffs under varying conditions. This adaptability is particularly advantageous in environments where the end of the game is uncertain.

Impact of uncertainty: Under conditions of uncertainty, Second by Tester's performance gains are amplified. This is due to the variability in game length, which increases the opportunities for exploitation. The increased discrepancies highlight the potential for adaptive strategies to thrive in variable conditions while exposing the limitations of rigid, rule-based strategies like Tit For 2 Tats.

Second by Gladstein and Tit For 2 Tats

Behavior: A similar pattern emerges in the interaction between Second by Gladstein and Tit For 2 Tats. Gladstein's strategy, characterized by its adaptive responses, takes advantage of Tit For 2 Tats' predictable tolerance for initial defections.

Observation: The introduction of uncertainty magnifies the discrepancy between these two strategies. Gladstein's ability to adjust to changing game dynamics enhances its effectiveness, while Tit For 2 Tats' static rule set makes it vulnerable.

Impact of uncertainty: The amplified discrepancy underscores the role of adaptability in uncertain environments. While Tit For 2 Tats performs reliably under fixed conditions, its lack of dynamic response mechanisms makes it susceptible to exploitation when the game's duration becomes unpredictable.

Grudger and Second by Colbert

Behavior: Grudger, a highly rigid strategy that perpetually punishes any defection, interacts with Second by Colbert, which introduces resets to destabilize its rigid retaliation mechanism. This interaction highlights the limitations of strict punitive strategies in dynamic environments.

Cause of discrepancy: The rigidity of Grudger is a critical weakness. Its unyielding approach fails to account for strategic resets, which Second by Colbert effectively uses to reset the game's cooperative dynamics. By doing so, Second by Colbert neutralizes Grudger's retaliatory strategy and destabilizes its long-term effectiveness.

Impact of uncertainty: In uncertain environments, Grudger's inability to adapt is further exposed. The resets introduced by Second by Colbert exploit the variability in game length, leading to significant performance discrepancies. This demonstrates that strategies overly reliant on fixed retaliation rules are ill-suited to uncertain conditions.

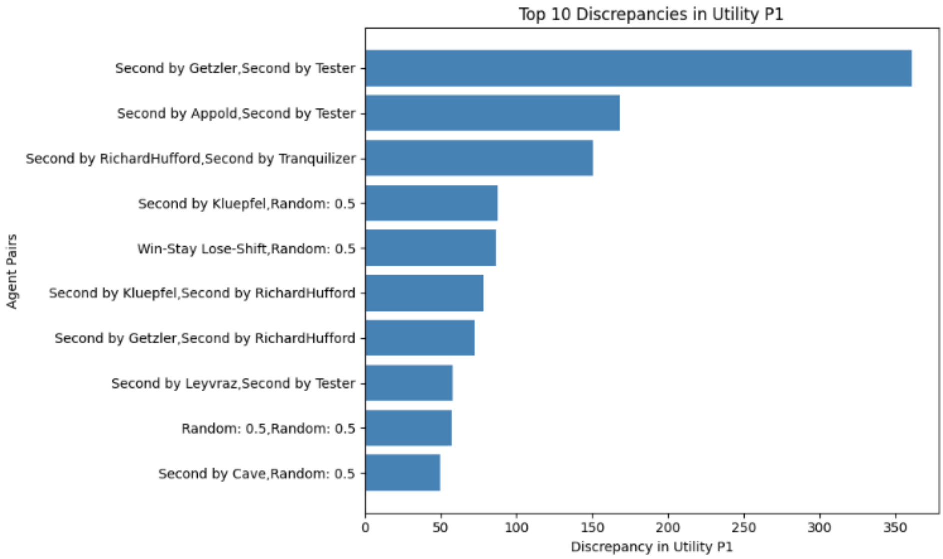
5. Discussions and conclusions

In this section, we analyze and interpret the results of the second-generation Axelrod tournament simulations conducted in environments with both certainty and uncertainty regarding the number of game stages. The main focus lies on understanding how agents adapt (or fail to adapt) to the uncertainty about the game's end stage and how this impacts their performance.

5. 1. Discussions about essential metrics and observations

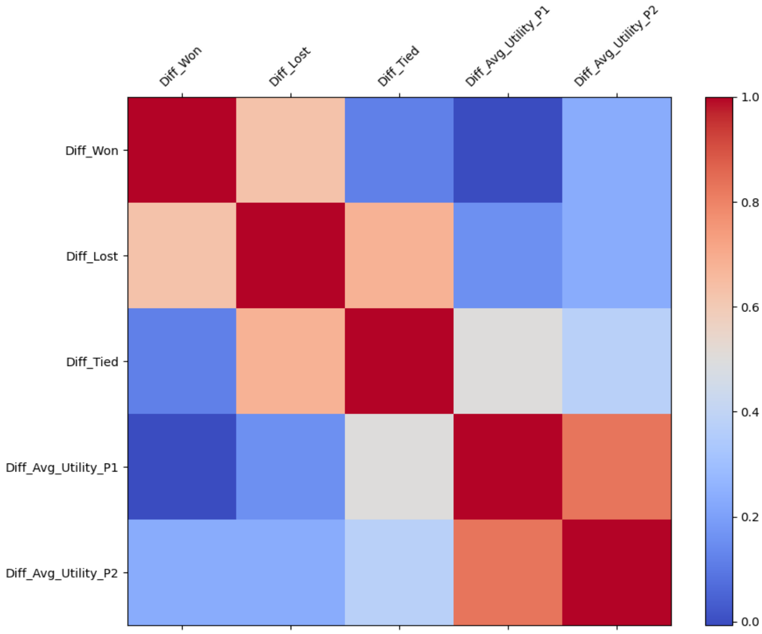
We evaluated agent behavior by computing critical metrics, especially looking at the differences in average utilities—referred to as "Diff P1" and "Diff P2"—between certain environments and those that introduce uncertainty regarding the end stage. These metrics establish a quantitative basis for identifying and evaluating performance differences in these distinct scenarios. The analysis presented in Figure 1 identifies the top 10 significant discrepancies in utility for P1. It demonstrates that certain agents, including "Second by Getzler" and "Second by Tester," show important sensitivity to the introduction of uncertainty, illustrating the influence of environmental conditions on their strategies.

Figure 1. Top 10 Discrepancies in Utility for P1 Across Certain and Uncertain Environments



A significant correlation was identified between the discrepancies in utilities and adaptability metrics, as illustrated in Figure 2. This suggests that agents' ability to adapt plays a pivotal role in their performance across varying environmental conditions. For example, strategies that are heavily dependent on end-game scenarios, such as “Second by Tester,” exhibit pronounced drops in utility when uncertainty is introduced, underscoring their reliance on predictable termination points. In contrast, more resilient strategies, such as “Grudger,” demonstrate remarkable stability and maintain consistent performance regardless of the level of uncertainty, reflecting their inherent robustness and less dependence on end-stage predictability.

Figure 2. Correlation Matrix of Wins, Losses, and Utilities



The difference in average utilities between the two environments is another important result that provides insight on how flexible and resilient various tactics are. Figure 3 illustrates these differences, highlighting specific patterns in performance. Strategies characterized by significant adaptability, such as "Tit for Tat," display minimal fluctuations in their average utility, indicating a reliable and stable reaction to environmental changes. This stability shows their intrinsic resilience and ability to function efficiently in both certain and uncertain conditions. In contrast, techniques characterized by rigid or deterministic decision-making frameworks, such as "Second by Tester," show significant reductions in utility when tested in unpredictable situations. These significant losses highlight their reliance on stable conditions and their restricted capacity to adapt to dynamic changes in the game framework. This disparity emphasizes the importance of flexibility and adaptability in strategy design, particularly in contexts where environmental conditions are unpredictable or volatile.

Figure 3. Average Utility Discrepancies Between Certainty and Uncertainty Environments

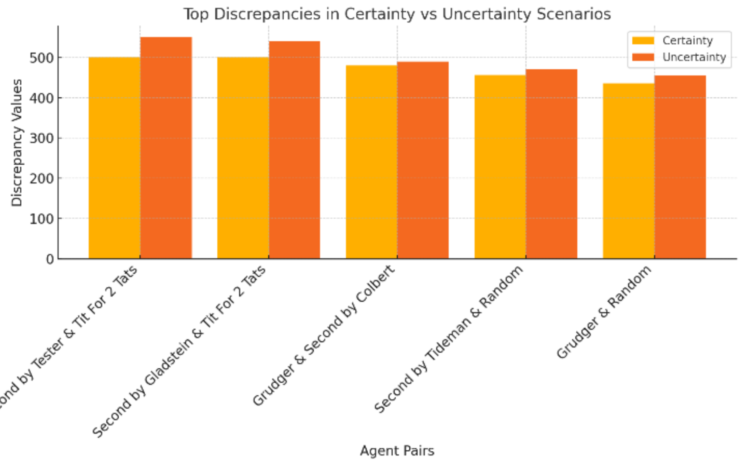


Figure 4 provides a thorough analysis of the distribution of utility values across scenarios defined by certainty and uncertainty, illustrating the fundamental dynamics of strategy performance. Strategies such as "Tit for Tat" show a compact utility distribution, indicating exceptional consistency and durability in their performance. Narrow distributions suggest that these techniques remain mostly unaffected by fluctuations in external variables, rendering them appropriate for both predictable and unexpected situations. In contrast, the "Second by Tester" strategy demonstrates a significantly wider utility distribution under uncertainty, indicating increased variability in its effectiveness. This wider distribution indicates a deficiency in robustness and a greater susceptibility to environmental variability. The difference in value distributions among different solutions highlights the essential importance of adaptability in reducing the negative impacts of uncertainty. Strategies characterized by narrower distributions usually show consistent decision-making frameworks, while those with broader distributions may lack mechanisms to stabilize performance in the face of unpredictable changes. This approach highlights the necessity of including both stability and adaptability in the formulation of strategies for situations characterized by various levels of predictability.

Figure 5 presents a comprehensive examination of the difficulties encountered by particular strategies in adjusting during transitions between environments characterized by certainty and uncertainty. This visualization highlights the differing levels of adaptation across strategies, providing essential insights regarding their performance resilience. The "Second by Tester" strategy demonstrates a substantial decline in efficacy while switching to an uncertain environment, indicating a possible overfitting to deterministic circumstances. This significant decline in performance indicates that the strategy is excessively dependent on predictable termination points and lacks the adaptability to respond to dynamic or less controlled scenarios.

These observations highlight the necessity for finding solutions that provide constant performance under different environmental situations. Strategies that demonstrate a significant decrease in utility under uncertainty may be useful in static or deterministic environments but are inappropriate for dynamic and unpredictable circumstances. This discrepancy analysis functions as an essential instrument for detecting strategy that are too specialized or restricted in their flexibility.

Figure 4. Utility Distributions in Certainty and Uncertainty Scenarios.

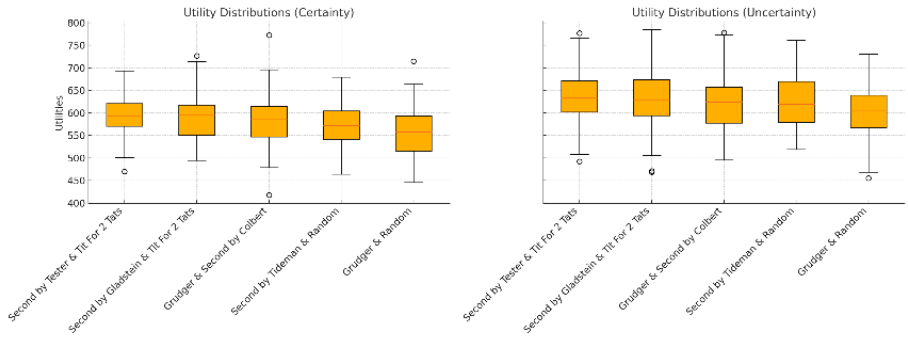
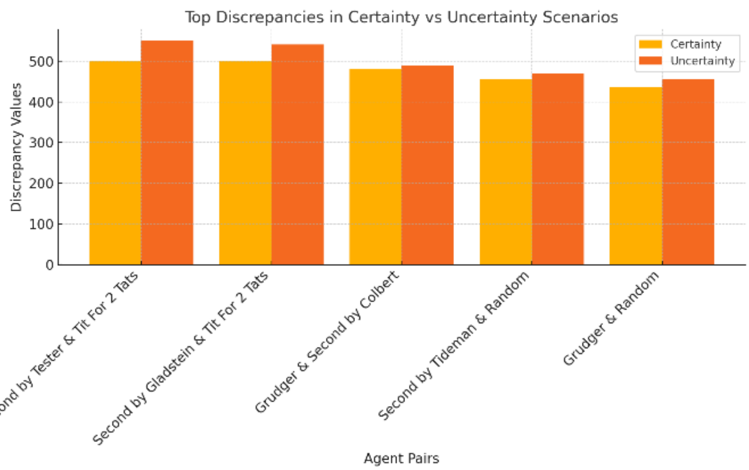


Figure 5. Discrepancy Analysis: Certainty vs. Uncertainty



5. 2. Conclusions

The above visualizations provide important insights into the performance dynamics of agents across different environmental situations, highlighting their adaptation, resilience, and limitations. Strategies like "Tit for Tat" and "Grudger" exemplify robustness, constantly providing stable and reliable performance regardless the volatility in their operating environments. These strategies demonstrate an effective combination of collaboration and flexibility, rendering them especially appropriate for volatile and uncertain situations where stability is important.

In contrast, strategy like “Second by Tester” and “Second by Getzler” reveal considerable weaknesses due to their evident dependence on deterministic frameworks and predictable game characteristics. These tactics demonstrate significant decreases in efficiency under uncertainty, indicating a fundamental inflexibility in their structure. Their incapacity to maintain constant performance under changing settings highlights essential areas for improvement, especially the necessity for reducing dependency on fixed ending points and deterministic frameworks.

An in-depth examination of these inconsistencies reveals opportunities for improving agent algorithms, highlighting the potential for innovation in various critical domains. Integrating probabilistic methods for making decisions may allow agents to navigate environments characterized by inherent randomness more effectively, enhancing adaptability. Moreover, the formulation of hybrid strategies—integrating resilient components from strategies such as “Tit for Tat” with innovative decision-making frameworks—presents a viable option to increase both flexibility and resilience. These improvements could allow agents to operate efficiently in both controlled environments and complex real-world situations marked by uncertainty and dynamic interactions.

Future research should focus on systematically refining these strategies to address their current limitations. This may include using sophisticated computational methods, such as machine learning, to model and improve agent behavior across diverse environmental situations. Furthermore, investigating methods to adapt strategic decision-making in response to real-time feedback might significantly enhance agents' capacity to react to evolving environments. By furthering these domains of investigation, we may improve the broader field of game theory and agent-based modeling, promoting developments that beyond academic study and reach practical applications in economics, artificial intelligence, and other disciplines.

References

Akin, E. (2016). The iterated Prisoner’s Dilemma: good strategies and their dynamics. In A. Idris (Ed.), *Ergodic Theory* (pp. 77-107). De Gruyter. https://doi.org/doi:10.1515/9783110461510-004

Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390-1396.

Ben-Porath, E. (1990). The Complexity of Computing a Best Response Automaton in Repeated Games with Mixed Strategies. *Games and Economic Behavior*, *2*, 1–12. https://doi.org/https://doi.org/10.1016/0899-8256(90)90010-R

Chiong, R., & Jankovic, L. (2008). Learning game strategy design through iterated Prisoner's Dilemma. *Int. J. Comput. Appl. Technol.*, *32*, 216-223.

Chong, S. Y., Humble, J., Kendall, G., Li, J., & Yao, X. (2007). Iterated Prisoner's Dilemma and Evolutionary Game Theory. In *The Iterated Prisoners' Dilemma* (Vol. Volume 4, pp. 23-62). WORLD SCIENTIFIC.

Leonardos Stefanos, P. G. (2012). Exploration-Exploitation Tradeoffs in Multi-Agent Systems: Catastrophe theory meets game theory. *Journal of Artificial Intelligence Research*, *304*, 103653. https://doi.org/https://doi.org/10.1016/j.artint.2021.103653

Leyton-Brown, K., & Shoham, Y. (2008). *Essentials of Game Theory: A Concise Multidisciplinary Introduction* (Vol. 2). Morgan and Claypool Publishers.

Meng Chen-Lu, P. R. (2001). The Iterated Prisoner's Dilemma: early experiences with Learning Classifier System-based simple agents. *Decision Support Systems*, *31*(4), 379-403. https://doi.org/https://doi.org/10.1016/S0167-9236(00)00137-8

Myerson, R. B. (1991). *Game Theory: Analysis of Conflict* (H. U. Press, Ed.). Harvard University Press.

Nachbar, J. (2005). Beliefs and Learning in Repeated Games. *Econometrica*, *73*(2), 459-480. http://www.jstor.org/stable/3598794

Nash, J. (1950). Equilibrium Points in n-Person Games. *Proceedings of the National Academy of Sciences*, *36*, 48–49.

Project, T. A. (2015). *Axelrod: second tournament*. https://axelrod.readthedocs.io/en/fix-documentation/reference/overview_of_strategies.html#axelrod-s-second-tournament

Press, W. H. & Dyson, F. J. (2012) Iterated Prisoners Dilemma contains strategies that dominate any evolutionary opponent. Proc. Natl. Acad Sci. USA 109, 10409–10413.

Rothe, J. (2010). *Economics and Computation - An Introduction to Algorithmic Game and Fair Division Theory, Computational Sociial Choice,and Fair Division*. Springer

Thomopoulos, N. (2013). *Essentials of Monte Carlo Simulation: Statistical Methods for Building Simulation Models*. Springer.

Tutzauer, F. (2007). A matrix approach to generalizing Axelrod-type tournaments. *Journal of Mathematical Sociology*, *31*(3), 249-266.

Tychonov, A. (1935). Ein Fixpunktsatz. *Mathematische Annalen*, *111*, 767–776.

Vincent, F. M., & Fryer, D. E. A. (2021). Top score in Axelrod Tournament. *arXiv preprint arXiv:2104.03347*.

Zhou, Y. H., Yu, J., & Wang, L. (2011). A new proof of existence of equilibria in infinite normal form games. *Applied Mathematics Letters*, *24*(3), 253-256. https://doi.org/https://doi.org/10.1016/j.aml.2010.09.014