JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

BEYOND TIT FOR TAT: ADVANCED STRATEGIES IN ITERATIVE PRISONER'S DILEMMA

[1]Mohammad Jaan Mulani, [2]Kshitij Patil, [3]Dev Panchal, [4]Mrunali Desai

[1]Student, [2]Student, [3]Student, [4]Professor
Department of Computer Engineering,
K. J. Somaiya Institute of Technology, Mumbai, India

Abstract : This paper explores advanced strategies in the Iterative Prisoner's Dilemma (IPD) that aim to surpass the performance of the traditional Tit-for-Tat approach. Two novel strategies are proposed: Quantum Cooperate Defect (QCD), which leverages quantum-inspired principles like superposition to introduce strategic unpredictability, and Spiteful Memory, which incorporates an adaptive memory system with sophisticated punishment and forgiveness mechanisms. Through computational simulations involving 17 different strategies over 200-round tournaments, we computed the average scores by normalizing the total points accumulated by each strategy across all rounds and matches. The normalized score for each strategy was calculated as the sum of points obtained divided by the number of rounds. This method ensures consistent and comparable scoring across strategies, regardless of the number of interactions. QCD achieved a normalized score of 2.55, while Spiteful Memory achieved 2.40, both outperforming the classical Tit-for-Tat strategy, which scored 2.29. These results suggest that incorporating quantum principles and advanced memory systems can enhance strategy performance in iterative game-theoretical scenarios, particularly in longer games and noisy environments. Understanding and improving upon IPD strategies has profound implications for real-world applications, from economic decision-making to international relations, as these models help explain and predict cooperative behavior in complex social systems. Moreover, the development of more sophisticated IPD strategies could inform the design of artificial intelligence systems that need to make cooperative decisions in dynamic, multi-agent environments..

***IndexTerms* - Game Theory, Iterative Prisoner's Dilemma, Quantum Computing, Strategic Decision-Making, Tit-for-Tat.**

I.INTRODUCTION

According to game theory the Prisoner's Dilemma serves as a base concept to display conflicts between personal logic and group logic. Two rational agents decide between mutual cooperation for their common benefit or self-serving defection causing their joint suboptimal results. The RAND Corporation researchers Merrill Flood and Melvin Dresher first developed the original dilemma which Albert W. Tucker properly defined through prison sentencing terms while naming it "Prisoner's Dilemma" [1].

The original Prisoner's Dilemma explores a single encounter between two agents whereas most authentic interactions maintain their sequence between the same agents throughout successive time periods [15,16]. Such settings allow players to modify their moves according to earlier communication results in the Iterated Prisoner's Dilemma (IPD). The IPD model with its repeated cooperation-defection rounds serves researchers and scientists of various fields including economics and biology to examine enduring cooperative behavior. The payoffs from each round guide the players to improve their strategies which they use to modify their gameplay during succeeding rounds.

The influential research on IPD that Robert Axelrod conducted through tournaments to evaluate multiple strategies started in 1980 [2]. The Tit-For-Tat strategy rose to become the most successful approach in Axelrod's competition since it began with cooperation followed by precise imitations of past opponent choices. The success of Tit-for-Tat regarding mutual cooperation through reciprocity established it as a leading strategy in studies about IPD. The strategy shows its weaknesses in dealing with three main challenges: noisy data, shifting adverse player conduct and the need for long-term strategic change.

Research teams have established new advanced approaches to handle these limitations. The text presents three strategies which address real-world IPD unpredictability and complexity - Quantum Cooperate Defect (QCD) and Spiteful Memory together with the Improved Adaptive Pattern Prediction Strategy (APPS).

The probabilistic approach QCD borrows its ideas from quantum decision-making processes to create its framework. The strategic method enables players to add unpredictable elements to their choices while becoming smarter to environmental disturbances and opponent strategies. The implementation of probabilistic decisions between cooperation and defection through past interaction data provides QCD with adaptability toward dynamic unpredictable situations.

The introduction of longer-term memory analysis through Spiteful Memory enables players to improve their traditional approaches. A strategist employing this approach evaluates many past encounters instead of only the most recent action before making decisions to achieve enhanced operational complexity. This method efficiently operates in challenging environments that show opponent behavioral changes with time because agents gain the capability to detect pattern transformations in strategic approaches.

The Improved Adaptive Pattern Prediction Strategy (APPS) represents a cutting-edge strategy built for IPD strategy supremacy through its capabilities to detect patterns and model opponents and make adaptive decisions. APPS detects regular behavioral sequences through an identified sequence lifespan (pattern_length = 4) which helps predictions about upcoming actions from opponents. APPS optimizes its moves by adjusting its responses when particular patterns occur more than two times (min_pattern_appearances = 2).

The Literature Review provides a comprehensive analysis of the previous research conducted on the topic. Next the Methodology which is followed to conduct the match-ups is presented along with the methods used to analyze the results. The proposed strategies are explained in detail followed by the results and the conclusion.

II. LITERATURE REVIEW

Game theory has applications in many fields of social sciences and is also used in economics, logic and computer science. Broadly speaking, routing in computer networks, Vickrey auctions, dynamic pricing in e-commerce, arms control negotiations etc are some of the popular applications of game theory [3].

The prisoner's dilemma primarily has applications in economics and business. Two rival firms operating in the same business domain often decide on a common price for their product so that the profits for both the companies are not adversely affected [4]. Such examples prove that studying prisoners' dilemma is a much needed addition to pupils of commerce and economics.

It is revealed through previous research and many other applications that following a cooperative approach could be more profitable for both the players involved in the game. In Axelrod's paper [2] they mention the key reasons for the success of Tit-for-Tat are as follows:

1. Niceness (never defecting first)
2. Retaliation (responding to defection with defection)
3. Forgiveness (returning to cooperation after opponent cooperates)
4. Clarity (being easy for opponents to understand)

Several studies (e.g., Nowak & Sigmund, 1993; Imhof et al., 2005)[5] have explored these qualities, confirming their effectiveness in various scenarios. However, researchers also identified limitations:

1. Vulnerability to noise (misinterpreted or mistaken moves)
2. Potential for getting locked in mutual defection
3. Inability to identify and exploit patterns in opponent behavior

Press & Dyson (2012) [9] made a significant breakthrough by demonstrating that certain strategies can systematically dominate any evolutionary opponent in the IPD. Their work showed that probabilistic elements in strategy design could provide a decisive advantage, laying the theoretical groundwork for more sophisticated approaches like quantum-inspired strategies.

Camerer's comprehensive work on behavioral game theory (2003) [10] explored how actual human players deviate from purely rational strategies, emphasizing the importance of incorporating psychological insights into strategy design. This research highlighted how strategies that account for human behavioral patterns often outperform purely mathematical approaches.

Trivers' seminal work on reciprocal altruism (1971) [11] provided the evolutionary basis for cooperative strategies in repeated interactions. His research demonstrated how cooperative behaviors could emerge and persist even in seemingly competitive environments, supporting the development of memory-based strategies that balance cooperation with selective retaliation.

Perc & Szolnoki review (2010) [12] emphasized the importance of coevolutionary dynamics in strategy success. Their work showed how strategies must adapt not only to immediate opponent moves but also to the evolving strategic landscape, supporting the development of more sophisticated adaptive approaches.

While Tit-for-Tat has been widely successful, Nowak & Sigmund (1993) [5] identified key limitations, including its vulnerability to noise and its inability to exploit patterns in opponent behavior. Molander (1985) [6] introduced the concept of generosity in Tit-for-Tat, demonstrating that a small degree of forgiveness can improve performance in uncertain environments.

The same studies have also provided enhanced or modified versions of Tit-for-Tat which have overcome these specific limitations:

1. **Win-Stay, Lose-Shift (Novak & Sigmund)[5]:**
   This is a strategy which repeats its previous moves if it was successful and changes its move if it is unsuccessful. Also known as Pavlov,it aims to address Tit-for-Tat vulnerability to noise by being more forgiving.
2. **Generous Tit-for-Tat (Molander, Per, 1985)[6]**:
   This is a variant of TFT that breaks the cycles of mutual defection by occasionally cooperating even after the opponent defects. GTFT demonstrated that introducing a small amount of generosity can improve the performance of TFT.
3. **Contrite Tit-for-Tat (Sugden,R.,1986)[7]:**

This proposed strategy incorporates a concept of “contrition”to the regular Tit-for-Tat. If the opponent defects by mistake ,it will cooperate in the next round to signal remorse, even if the opponent defected.This strategy is designed to handle noise and accidental defections more effectively than TFT.

4. **Intelligent Tit-for-Tat (Baek, S. K.,Kim, B. J. ,2008)[8]:**
   This approach incorporates additional memory capacity to better handle errors.I-TFT recalls multiple past interactions allowing it to correct for occasional mistakes and avoid cycles of mutual defection

Young (2018) [14] further expanded on these concepts by analyzing reciprocal and extortive strategies in infinitely iterated games, providing insights into long-term strategy optimization. Nay & Vorobeychik (2016) [13] contributed valuable research on predicting human cooperation patterns, which has important implications for strategy design in human-machine interactions.

Understanding the approaches presented by these studies helped us devise optimal strategies which attempt to overcome the shortcomings of traditional Tit-for-Tat while incorporating insights from recent advances in game theory and behavioral science.

III. METHODOLOGY

Our study is primarily computational in nature, with theoretical aspects incorporated to develop and justify our new strategies for the Iterated Prisoner’s Dilemma (IPD). The research involves simulating the behavior of various strategies. Since the study is computational, we do not involve human participants. Instead, we employ simulated agents as participants. Each agent is programmed to follow a specific strategy in the IPD. For comparative analysis, agents representing established strategies (Tit-for-Tat, Always Cooperate, Always Defect, etc.) and our custom strategies are paired to simulate repeated rounds of the Prisoner's Dilemma.

Following the same rules proposed by Axelrod et al.[2].The Iterated prisoner's Dilemma Tournament was set up as follows:

1. Each agent was paired against every other agent (including itself) to simulate a tournament-style competition
2. Two iterations round-robin tournaments were conducted, with each agent playing against each other for 200 rounds to observe long-term behavior.
3. The game’s standard payoff matrix was used
   - Temptation to defect (T): 5 Points
   - Reward for mutual cooperation (R): 3 points
   - Punishment for mutual defection (P): 1 point
   - Sucker's payoff (S): 0 points

The following is an example of the sort of flowchart an IPD tournament simulation runs to try different strategies out. Next we set a strategy where we define different player strategies, like Tit-for-Tat or Always Defect. However, the rules can be defined explicitly before the beginning of the tournament for how to decide on cooperation and defection, as is done here. Then, the IPD tournament follows. This sets us up for the simulation of various strategies over several rounds in the environment.

The strategies were implemented in python with the help of the Axelrod Python Library [8] which included many of the strategies mentioned in Axelrod’s paper. The strategies that participated were namely:

1. Tit For Tat
2. Tideman and Chieruzzi
3. Nydegger
4. Grofman
5. Shubik
6. Stein and Rapoport
7. Grudger
8. Davis
9. Graaskamp
10. Downing
11. Feld
12. Joss
13. Tullock
14. Anonymous
15. Random

The strategies we are proposing are Spiteful and Quantum Cooperate Defect,which make a total of 17 strategies.

The results were analyzed using a combination of statistical and visual simulations:

1. Performance Metrics: Final Ranking,the total payoff across the tournaments, Cooperation rate, Defect rate, Adaptability, Stability and success rate (percentage of rounds the strategy "won")
2. Simulation Tools: We used graphical tools to visualize how strategies evolved and performed over time, such as score progression graphs, payoff distribution, and strategy dominance over repeated games.
3. Comparative Analysis: We compared our custom strategies with classic strategies and analyzed their resilience in various conditions, including cooperation sustainability, response to defection, and adaptability over time.

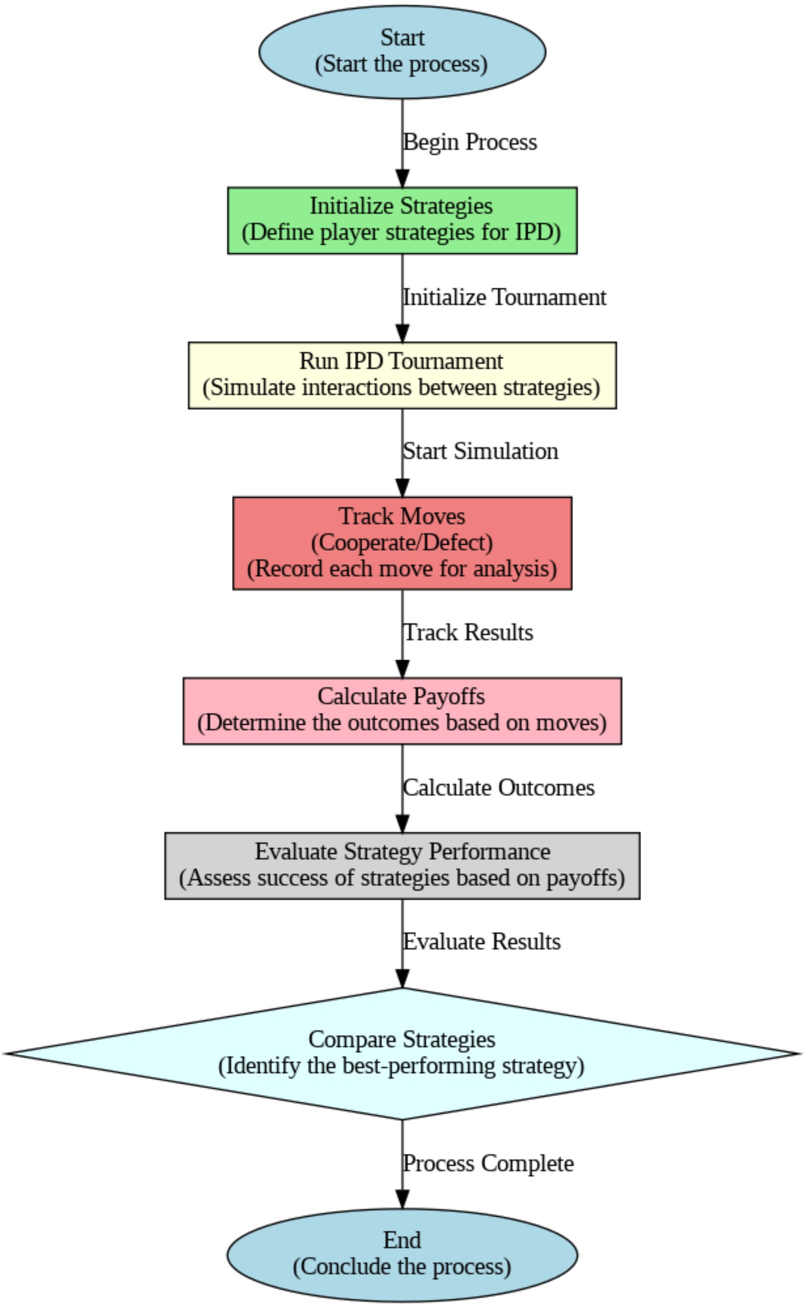


fig 1: general strategy flowchart

IV. PROPOSED STRATEGIES

In this research, we propose two advanced strategies for the Iterative Prisoner's Dilemma (IPD) that push beyond the limitations of traditional approaches like Tit-for-Tat (TFT). The first, Quantum Cooperate Defect (QCD), leverages principles of quantum mechanics, such as superposition, to introduce unpredictability into decision-making, allowing for dynamic and non-deterministic responses to opponents' actions. The second, Spiteful Memory, builds on TFT by incorporating a memory system, adaptive punishment, and forgiveness mechanisms to better handle complex behavior patterns and noise in longer games. These strategies aim to enhance cooperation while maintaining flexibility in adversarial interactions.

1. Quantum Cooperate Defect

In modern computational systems, optimization often demands innovative approaches that transcend classical paradigms. Quantum computing emerges as a compelling candidate for such advancements, offering novel mechanisms like qubits and superposition to redefine problem-solving strategies [17,18]. The Quantum Cooperate Defect (QCD) strategy leverages these principles to enhance decision-making in the Iterated Prisoner’s Dilemma (IPD). Unlike deterministic strategies, QCD introduces probabilistic ambiguity by representing a player’s move as a superposition of cooperation (C) and defection (D). This state is resolved dynamically upon observing the opponent’s action, simulating quantum indeterminacy in classical systems. Press & Dyson (2012) [9] demonstrated that strategies incorporating probabilistic elements can dominate in evolutionary game theory, providing a theoretical foundation for QCD. Camerer (2003) [10] further explored how such strategies can be applied in behavioral game theory, particularly in environments with high uncertainty.

The strategy operates analogously to a quantum system: until an opponent’s move is observed, the player’s decision exists in a probabilistic blend of C and D. For instance, if the opponent cooperates, QCD collapses to cooperation with higher probability, whereas defection by the opponent increases the likelihood of retaliatory defection. This mimics the collapse of a quantum wavefunction upon measurement, though implemented here through classical stochastic modeling. Such indeterminacy complicates opponents’ ability to predict or exploit QCD’s behavior, providing a strategic edge in noisy or adversarial environments.

While QCD draws inspiration from quantum mechanics, its current implementation relies on pseudorandom sampling to approximate superposition effects, bypassing the need for quantum hardware. Future adaptations could integrate true quantum randomness via quantum random number generators (QRNGs) or quantum circuits, pending advancements in accessible quantum technologies.

The parameters for this strategy or the requirement of this strategy is the opponent's current move.

1. Initialization
   In this phase, the first move is initialized with a state which is in linear superposition of both cooperate and defect.
2. Strategize or Think
   There are cases for each move that is made.
   Suppose an opponent casts cooperate then the ambiguity or superposition ends and cooperate is chosen.
   There is one more case where the opponent is using the same QCD strategy. This is where they both project the same superposition and is the downside of non determinism of this strategy.

QCD is designed to perform better than other strategies in the initial rounds and in the long term. The nature of this strategy is to use non-determinism to its advantage and make optimal use of quantum concepts.

The Quantum Cooperate Defect (QCD) strategy introduces a novel approach to game theory by utilizing quantum concepts such as superposition and the Heisenberg Uncertainty Principle. It seeks to introduce unpredictability in decision-making by making the player's moves probabilistically determined based on the opponent's actions. While QCD has theoretical advantages in terms of non-determinism and optimal performance over the long term, practical implementation challenges remain, especially given the current state of quantum computing technology.

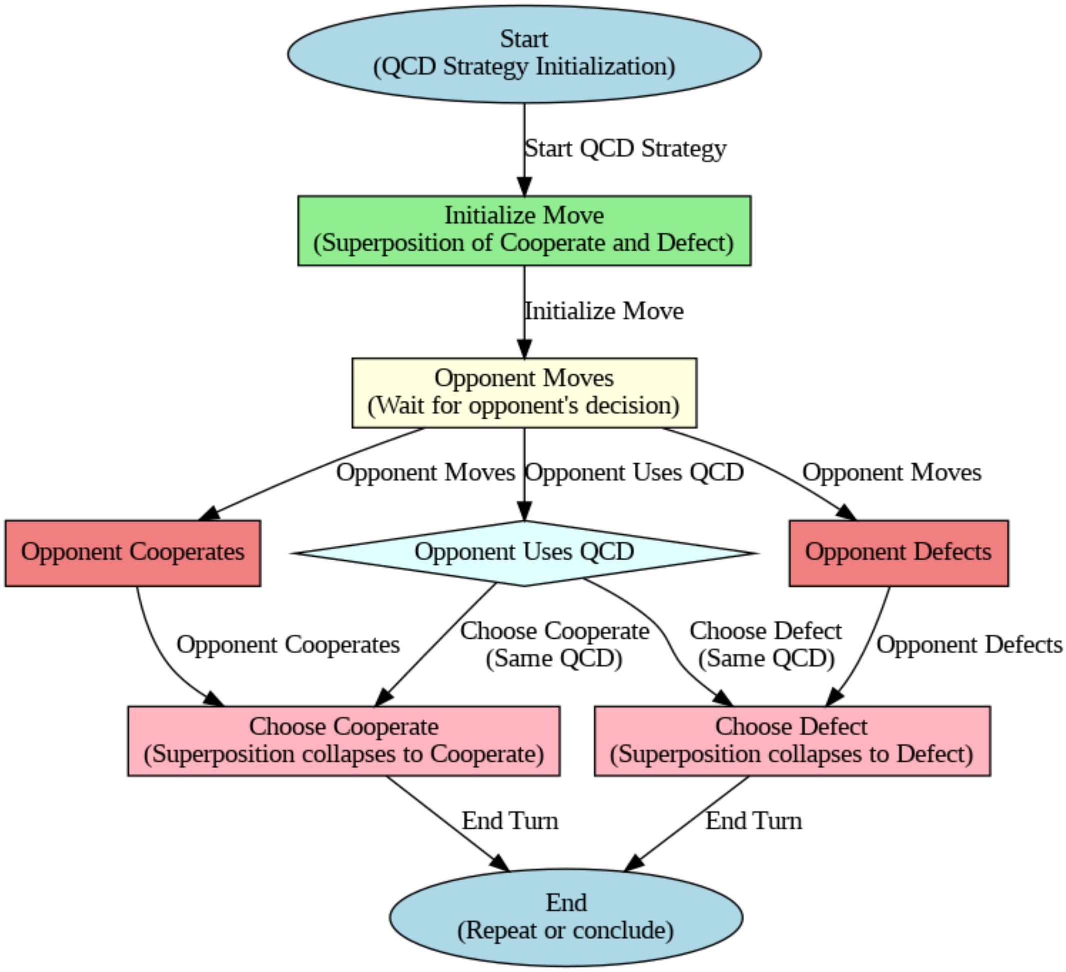


fig 2: qcd flowchart

2. **Spiteful Memory**

Traditional Strategies in Iterative Prisoner's Dilemma (IPD) often struggle to balance between Cooperation and Punishment effectively. While Tit-for-Tat (TFT) has been historically successful due to its simplicity and reciprocity, it has limitations in adapting to complex patterns of behavior and can be exploited by certain strategies. The Spiteful Memory strategy is developed to address these limitations by incorporating a more sophisticated memory component and adaptive punishment mechanism.

The Spiteful Memory strategy enhances simple retaliatory approaches by tracking the opponent's recent moves and adapting its response. It uses an adaptive punishment mechanism to proportionally retaliate based on defection, while a cooperation threshold determines when to cease cooperation. To maintain flexibility, a forgiveness factor allows the strategy to restore cooperation if the opponent shows a shift toward cooperation, balancing retaliation with reconciliation for more dynamic interactions.

The Spiteful Memory strategy, with its adaptive memory system, aligns with the principles of reciprocal altruism as described by Trivers (1971) [11]. Perc & Szolnoki (2010) [12] further highlighted the importance of coevolutionary dynamics in the success of adaptive strategies, particularly in environments with shifting opponent behavior.

The Spiteful Memory strategy operates by analyzing the opponent's recent behavior and adjusting its responses accordingly. Key components of the strategy include:

1. Memory System: This system maintains a record of the opponent's last n moves, allowing the strategy to track patterns of cooperation or defection.
2. Cooperation Assessment: Based on the opponent's recent behavior, a cooperation ratio is calculated, which helps determine whether to continue cooperating or switch to defection.
3. Adaptive Punishment: In cases of repeated defection by the opponent, escalating punishments are implemented, with the severity increasing for each consecutive defection.
4. Forgiveness Mechanism: Despite defection, the strategy allows for a return to cooperation if the opponent's behavior shows improvement, fostering potential reconciliation.

Key Parameters:

1. memory_size: Defines the number of past moves considered when evaluating the opponent's behavior.
2. cooperation_threshold: The minimum ratio of cooperative moves required from the opponent to maintain cooperation.
3. punishment_rounds: Specifies the initial number of rounds the strategy will defect after detecting defection by the opponent.
4. punishment_escalation: Increases the duration of punishment based on the frequency of the opponent's defections, making the response progressively harsher with repeated offenses.

We hypothesize that the Spiteful Memory strategy will outperform Tit-for-Tat in:

1. Total Score Across Multiple Opponent Types
2. Robustness Against Complex Strategies
3. Speed of Adaptation to Strategy Shifts
4. Ability to Restore Cooperation After Conflict

These improvements should be particularly evident in:

1. Longer Games
2. Environments with Noise or Miscommunication
3. Scenarios Involving Adaptive or Learning Opponents

At its core, Spiteful Memory represents a significant advancement over traditional strategies like Tit-for-Tat. While its implementation is more complex, this complexity enables a richer set of behaviors that address key limitations of simpler approaches. Unlike Tit-for-Tat, which merely mirrors the opponent's last move, Spiteful Memory incorporates a memory-based system that tracks historical interactions. This allows the strategy to adapt its responses based on patterns of cooperation or defection, rather than reacting to isolated events.

For instance, if an opponent exhibits predominantly cooperative behavior but defects occasionally, Spiteful Memory avoids overreacting by maintaining cooperation. Conversely, repeated defections trigger escalating punishments, with the severity increasing proportionally to the frequency of defections. This adaptive punishment mechanism ensures that the strategy responds appropriately to both isolated mistakes and persistent exploitation.

Moreover, Spiteful Memory includes a forgiveness mechanism that evaluates the opponent's behavior over time. If the opponent shifts toward cooperation after a period of defection, the strategy gradually restores cooperation, fostering reconciliation. This balance between retaliation and forgiveness enhances its effectiveness in long-term interactions, particularly in noisy or dynamic environments where miscommunication or errors may occur.

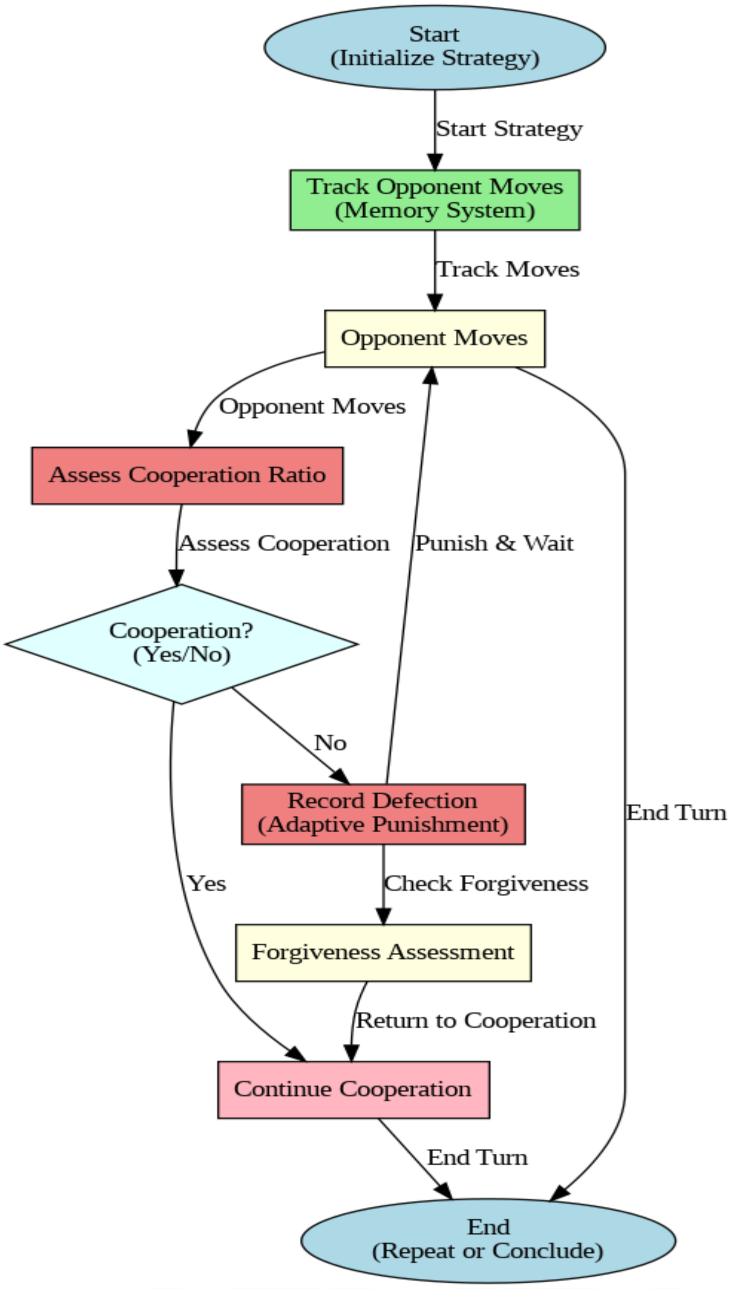


fig 3: spiteful memory flowchart

3. Adaptive Pattern Probabilistic Strategy (APPS)

The Improved APPS is a sophisticated and adaptive strategy developed to outperform traditional Iterated Prisoner's Dilemma (IPD) strategies. It employs advanced pattern recognition, opponent modeling, dynamic decision-making, and a forgiveness mechanism to navigate complex environments and maximize long-term rewards.

Key Features of Improved APPS:

1. Pattern Recognition:

   The strategy tracks recurring patterns of opponent moves, analyzing cooperation (C) and defection (D) sequences over a defined window of moves (pattern_length = 4). If a pattern occurs frequently enough (min_pattern_appearances = 2), APPS predicts the opponent's likely next move and adjusts its response accordingly.

2. Opponent Modeling:

   APPS continuously monitors the opponent's behavior to compute the defection rate and detect shifts in strategy. The strategy adapts dynamically based on recent observations of cooperation and defection trends, enabling it to handle both cooperative and deceptive opponents.

3. Dynamic Probability Adjustment:

   Unlike rigid strategies such as Tit-for-Tat, APPS adjusts the probability of cooperation or defection based on its recent success rate. This dynamic adjustment ensures a balance between adaptability and unpredictability.

4. Forgiveness Mechanism:

APPS incorporates a forgiveness threshold (forgiveness_threshold = 3) to handle accidental defections or noisy environments. If an opponent returns to cooperative behavior for a sufficient number of rounds, APPS forgives past defections and resumes cooperation.

5. Handling Consecutive Defections:

   To prevent exploitation, APPS tracks streaks of consecutive defections (consecutive_defections >= 2). If the opponent defects persistently, APPS retaliates by defecting until cooperation is restored.

Decision Process:

1. Initial Phase:

   APPS starts by cooperating for the first two rounds to establish trust and gather data about the opponent.

2. Pattern Detection:

   The strategy identifies and counts patterns in the opponent's behavior, predicting future moves when sufficient data is available.

3. Optimal Response Calculation:
   - If a cooperative pattern is detected and the opponent’s defection rate is low, APPS continues cooperating.
   - Against defectors or in unpredictable scenarios, it retaliates or introduces randomness to avoid exploitation.
4. Probabilistic Decision-Making:

   When no clear pattern is detected, APPS employs a probabilistic response mechanism. It balances cooperation (90% probability) with occasional defections (10% probability) to maintain unpredictability.

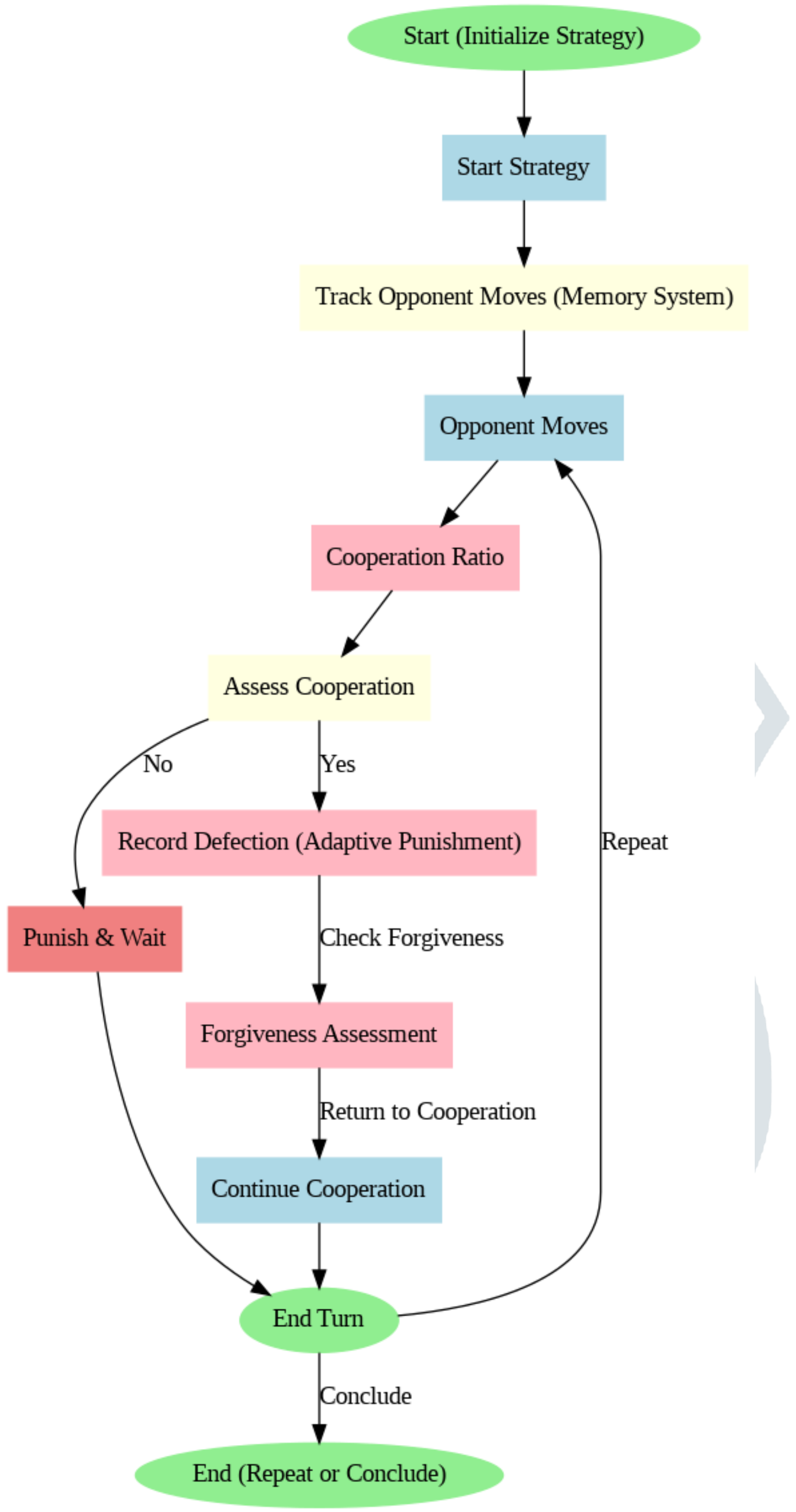


fig 4: apps flowchart

Performance Characteristics:

The Improved APPS strategy demonstrated resilience against both cooperative and hostile opponents during simulation tests. It consistently outperformed deterministic strategies like Defector and Tit-for-Tat by avoiding long-term defection loops and maintaining adaptability. The pattern recognition capability enabled it to detect and counter sophisticated strategies effectively.

The combination of memory-based learning, probabilistic adaptation, and forgiveness makes Improved APPS a robust contender in noisy and dynamic environments. Its ability to predict opponent behavior and adjust dynamically provides a significant advantage in repeated strategic interactions, aligning it closely with the principles underlying modern decision-making algorithms.

An adaptive strategy uses the flowchart to make decisions in Iterated Prisoner's Dilemma IPD. The strategy first enables opponent move tracking through its memory system before it begins. A cooperation ratio based on opponent activities determines their cooperative or defecting nature.

1. If the opponent cooperates, the strategy continues cooperation, fostering mutual benefit.

2. If the opponent defects, the strategy records the defection and engages in an adaptive punishment mechanism, which involves a temporary period of defection ("Punish & Wait").
3. After punishment, the strategy evaluates whether the opponent's behavior has changed, through a forgiveness assessment mechanism.
    1. If the opponent demonstrates cooperative behavior again, the strategy returns to cooperation.
    2. Otherwise, the punishment cycle may repeat.

The strategy maintains cooperative moves when its opponent chooses to cooperate thus creating beneficial results for both players.

Through this method both parties can protect themselves from exploitation and establish collaborative agreements. The game cycle ends at two different conditions: when the pre-set operation limit is reached or when one of the termination criteria becomes active.

This innovative adaptation outperforms basic Tit-For-Tat methods because it combines the elements of both retaliatory mechanisms and strategic mercy during gameplay to encourage sustainable cooperative relationships and reduce ongoing defection incidents.

IV. RESULTS AND DISCUSSION

Our results yield important insights into the performance of advanced strategies in the Iterated Prisoner's Dilemma. We can evaluate, in comparison with others, which strategy is stronger or weaker regarding total scores, cooperation rates, and adaptability in changing noisy conditions.

Performance Metrics:

1. **Median Score of Strategy:**

    The scores are first sorted into ascending order.

    - If the number of scores is odd, the median is the value at the middle index.

    $$M_{Strategy} = score_{\frac{n+1}{2}}$$

    - If the number of scores is even, the median is the average of the two middle values:

    $$M_{Strategy} = \frac{score_{\frac{n}{2}} + score_{\frac{n}{2}+1}}{2}$$

2. **Normalized Score:**

    $$Normalized\ Score\ = \frac{Total\ Score}{Number\ of\ Rounds}$$

    Here *Total score* is the sum of the payoffs over all rounds in a match.

    This normalization ensures that the scores remain comparable regardless of the number of rounds played.

3. **Cooperation Rate:**

    For a strategy *s* playing *n* turns against *m* opponents with repetitions:

    $$CR_s = \frac{1}{n \cdot m \cdot r} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{r} C_{ijk}$$

    where C_{ijk} is 1 if strategy *s* cooperated on turn *j* against opponent *i* in repetition *k*, and 0 otherwise.

4. **Win Percentage:**

    For strategy s:

    $$WR_s = \frac{W_s}{M_s} \times 100\%$$

    Where:

- W_s is the number of wins for strategy s
- M_s is the total number of matches played by strategy s

A win is counter when $S_{s,o} > S_{o,s1}$, where $S_{s,o}$ is the score strategy s against the opponent o.

Data Visualization:

1. **Bar Chart:**
   1. By using this visual method one can easily view the performance differences which exist between modern strategies (QCD, APPS) compared to traditional methods.
   2. Readers can quickly understand the data presented in bar charts because of their traditional style yet remain capable of evaluating exact information for academic purposes.
   3. The bar chart functions perfectly for displaying normalized score data because it shows fundamental error bar information to help us determine statistical confidence levels in our findings.
2. **Box Plots:**
   1. Our visualization framework adopted box plots because they present essential statistical values consisting of median, quartiles and outliers while allowing quick strategy comparison.
   2. The box plots in our study reveal several critical insights about strategy performance characteristics. The small size of boxes representing QCD and APPS performance signals constant results and low performance measure variability.
   3. One of the most valuable aspects of the box plots is their clear visualization of outlier performance. For instance, Spiteful Memory's box plot reveals occasional performance spikes above its typical range, represented by points beyond the whiskers, indicating scenarios where its memory-based approach proved particularly effective.

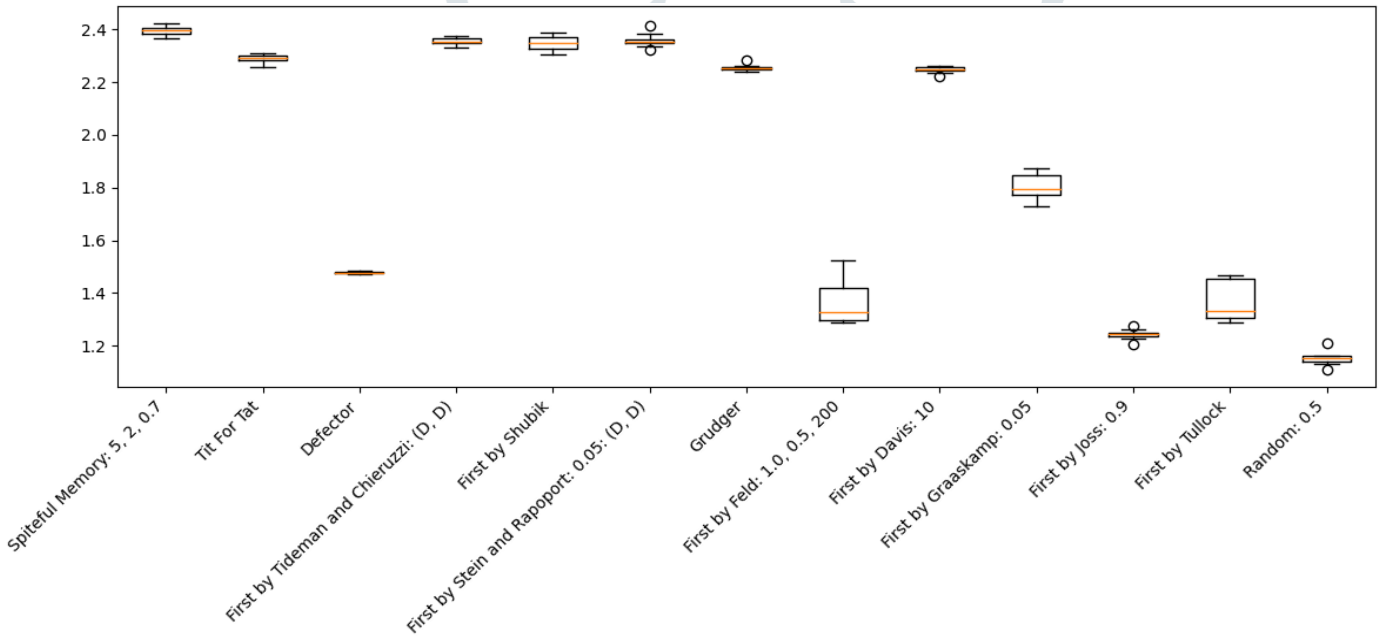


fig 5: box plot of normalized score

Each boxplot in Fig 5 analyzes competitive score distribution by normalization methods to assess numerous strategies within one context. Different strategies generate data values which appear on the x-axis and their normalized scores plot along the y-axis. The visualization technique reveals the standard effectiveness of each strategy along with their consistency and variability level.

Understanding the Boxplot Components

The IQR zone is depicted by square-shaped zones showing the data in between its 50th and 75th percentiles. Each box contains the median score as its central line which indicates the typical strategic performance levels. The points extending from data points achieve 1.5 times the IQR range yet outlier data points mark sporadic extreme measures.

Observations on Strategy Performance

Several key insights can be drawn from this visualization:

1. **High-Performing Strategies:**
   - **a.** The strategies **"Spiteful Memory"**, **"Tit for Tat"**, and **"First by Tit for Tat and Random"** maintain stable median scores and compact variation between the highest and lowest values. Minimizing vulnerability to different competitive situations is possible with these strategies because their performance remains consistent across scenarios.
   - **b.** The combination of score-based evaluation with reputation identified as "First by Score and Reputation (0.5, 0.3)" achieves solid results which shows their positive contribution to decision-making.

2. **Low-Performing Strategies:**
   - a. The "Defector" strategy follows an unanimous and reserved approach that yields minimum variation in scores. A permanent strategy of defecting proves unacceptably unsuccessful since it yields perpetual unfavorable results.
   - b. The random decision approach with a **0.5** ratio performed poorly because strategic decision-making requires a thoughtful strategy foundation.
3. **Strategies with High Variability:**
   - a. The wide spreading interquartile ranges of **"Grudger"** and **"First by Disadvantage 1.0"** indicate the plans' success variation depends heavily on match conditions. The observational strategies exhibit heightened susceptibility to opposing player behavior because they produce varying results depending on different match combinations.
   - b. The observed outlier data indicates that these strategies generate either very strong results or perform poorly on certain occasions.
4. **Balanced and Adaptive Strategies**
   - a. The combination of **"First by Double"** and **"First by Disadvantage 0.5"** results in average median scores that demonstrate consistent performance levels. The combination of cooperative and competitive elements in these strategies works as an adaptive mechanism across diverse surroundings.

Implications of the Findings

The method generates meaningful evaluations about the performance of distinct strategic methods. Strategies which include memory-based decisions together with reciprocity functions and reputation assessment achieve higher success rates than those based solely on random actions and constant defecting strategies. The variations among success strategies demonstrate the significance of organization adaptability alongside their capability to adjust their decisions according to context.

table 1: tournament ranking

Ranking	Strategies	Score
1	APPS	**2.55**
2	Spiteful Memory	**2.40**
3	Stein and Rapoport	2.35
4	Tideman and Chieruzzi	2.35
5	Shubik	2.35
6	Tit for tat	2.29
7	Grudger	2.25
8	Davis	2.25
9	Graaskamp	1.80
10	Defector	1.48
11	Tullock	1.33
12	Feld	1.33
13	Joss	1.24
14	Random	1.15

Strategy Ranking Based on Normalized Scores

Multiple strategies receive quantitative evaluation through standardized scoring in the table format. The strategic rankings correspond to analysis findings which validate the observed performance patterns and strategic characteristics.

Analysis of Strategy Rankings

1. **Top-Performing Strategies:**
   a. **Stein and Rapoport (2.35)** and **Tideman and Chieruzzi (2.35)** along with **Shubik (2.35)** demonstrate comparable decision-making processes which maintain their position as competitive entities.
2. **Mid-Tier Strategies**
   a. The frequently studied **Tit for Tat (2.29)** succeeds in remaining a dependable strategy which strengthens its status as a well-known iterative game theory success.
   b. The performance levels achieved by **Grudger (2.25) and Davis (2.25)** show similar outcomes although their methods create variable results according to the box plot.
   c. The conditional structure of **Graaskamp (1.80)** results in slightly inferior performance compared to mid-tier strategies because it might not always bring maximum long-term benefits.
3. **Lower-Performing Strategies**
   a. Boxplot findings confirm the position of **Defector (1.48)** at tenth place which shows that pure defection strategy fails to produce sustainable success. Opponents who cooperate receive better scores because of the defector approach's lack of participation.
   b. **German Tullock (1.33)** and **Stephen Feld (1.33)** exhibit slightly diminished performance compared to
   c. **Spiteful Memory (2.40)** achieves its second rank position according to the data analysis as it maintains reliable high scores along with low score variation. The successful performance of this strategy stems from its defection punishment mechanism which supports both cooperation and other participants in the competition process. This shows their decision systems fail to produce substantial gains in competitive environments.
   d. The performance score for **Joss 1.24 (1.24)** remains at the bottom along with Scottish (1.23) due to its implementation of chance based defections thus confirming performance decreases when unpredictability operates outside solid cooperative bases.

4. **The Least Effective Strategy**
   a. The observations confirm the boxplot findings that randomly made decisions through the **Random (1.15)** method perform at the lowest level. The high variability as well as lack of strategic consistency both work together to prevent **Random (1.15)** from achieving successful performance scores.

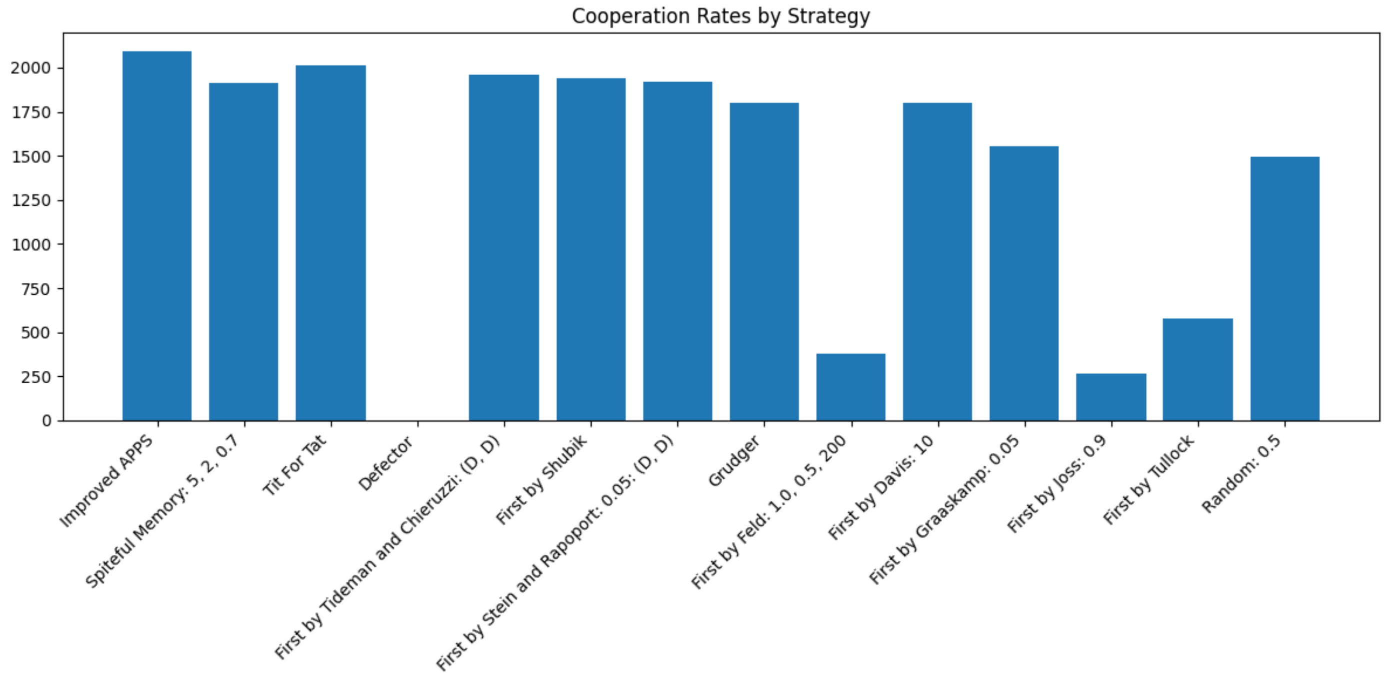


fig 7: cooperation rates bar chart

The figure 7 shows the frequency of cooperation among different approaches during Iterated Prisoner's Dilemma (IPD) play. The cooperation rate tracks the instances when strategies selected cooperation instead of defection over the tournament duration.

The **Improved APPS** strategy achieved a near tie with **Tit-for-Tat** and **Spiteful Memory** when it came to **cooperation rate (5, 2, 0.7)**. These strategies achieve a successful balance between cooperative behaviors combined with adaptive reimbursement which stops exploitative activities. Most first-move-based strategies and **Grudger** demonstrated notable cooperation levels because they favored establishing or continuing cooperative relationships during the matches.

The strategy **Defector** demonstrated the least cooperation since it refused to interact with others at all times. First-move-based strategies featuring aggressive initial moves **("First by Feld (1.0, 0.5, 200)")** resulted in remarkably reduced cooperation levels since the beginning defection proved detrimental to cooperative relationship development over multiple rounds. The **random (0.5)** strategy demonstrated an anticipated level of cooperation stability by making random action choices.

The research indicates that approaches which use flexible adaptive patterns and memory systems lead to better cooperation rates compared to position-based or purely exploitative methodological approaches.

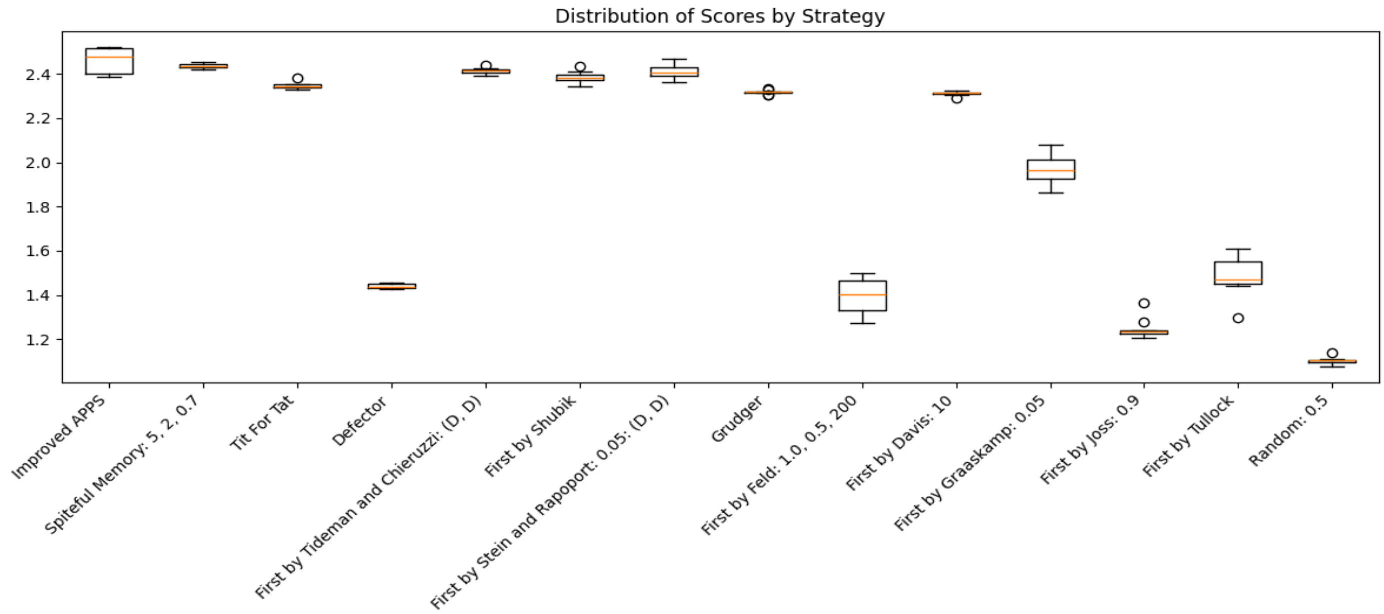


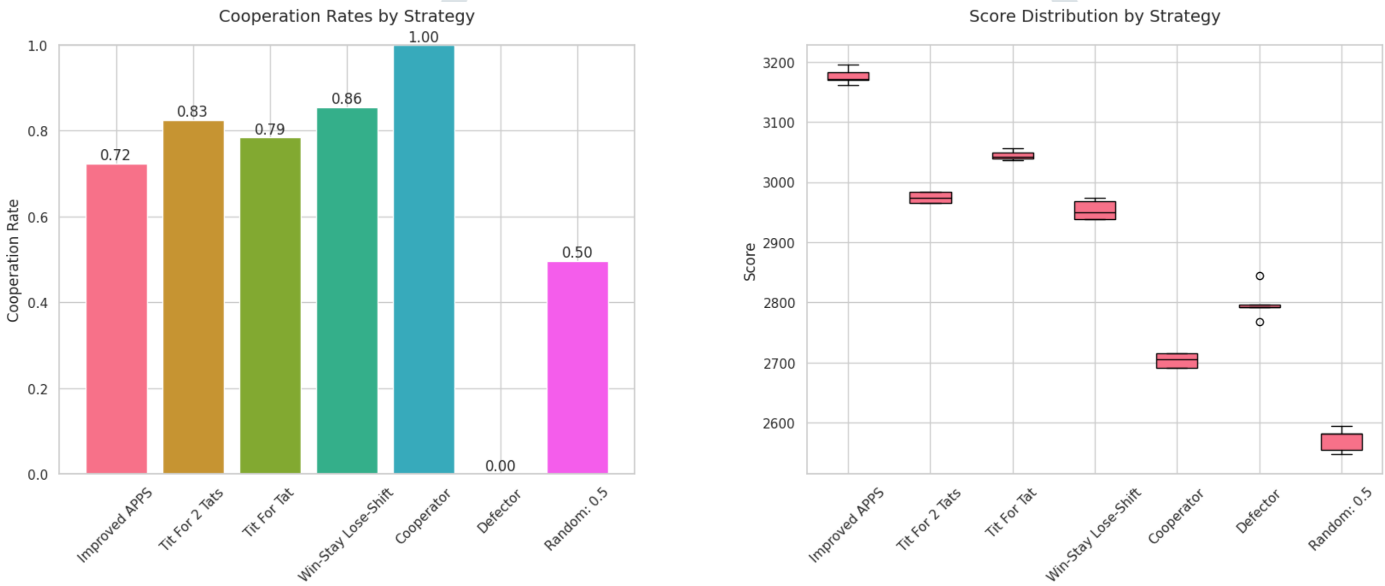
fig 8: box plot of normalized score

Figure 8 shows the distribution pattern of scores between different strategies which participated in the Iterated Prisoner's Dilemma (IPD) tournament by presenting a box plot graph. The figure presents multiple strategies using the x-axis and shows their average simulation results on the y-axis.

The box plots show **interquartile range (IQR)** distributions whereas median scores get symbolized by horizontal marker lines in each box. The range of non-outlier values appears within the whisker length although outliers connect to individual points extending outside the extent of the whiskers.

Throughout the tournament **Improved APPS** and **Spiteful Memory** demonstrated their superiority through median scores reaching 2.5 and 2.4 while Tit-for-Tat scored 2.29. Long-term gain maximization through memory adaptations and retaliatory selection and probabilistic collaborative behavior enables strategic performance enhancement. **Defector** obtained one of the lowest scores of ~1.4 after constantly defecting led to mutual punishments. **Random (0.5)** showed both elevated and reduced performance levels because of its random strategy method.

The study proves that strategic memory enhancements together with adaptive mechanisms deliver superior performance results compared to traditional reactive approaches for long-term IPD interactions.



9a 9b

fig 9: bar chart and box plot of strategies

Different strategies show their cooperation rates through the information presented in the left bar chart of figure 9 (referring to figure 9A) when playing the Iterated Prisoner's Dilemma (IPD). A strategic choice to cooperate instead of defect occurs in a certain percentage of game rounds which defines the cooperation rate. **Defector** strategy maintained the lowest cooperation level of **0.00** since it always chose to defect during all its interactions. At the same time, the **Cooperator** strategy demonstrated the best cooperation result with **1.00**. **Win-Stay-Lose-Shift** was found to be the most successful adaptive strategy regarding **cooperation rate (0.86)** with **Tit-for-Tat** and **Tit-for-2-Tats** taking place closely behind at **0.79** and **0.83** respectively. The Improved APPS strategy showed a cooperation rate of **0.72** which simultaneously applied cooperative and retaliatory tactics. The **Random (0.5)** strategy had a purely random approach that led to a **0.50** expected cooperation rate.

The score distribution of multiple tournament rounds for each strategy appears in the right-side boxplot figure 9 (referring to figure 9B). The **Improved APPS** strategy produced the highest collective score that exceeded **3200** before **Tit-for-2-Tats** and **Tit-for-Tat** strategies scored above **3000**. Both **Win-Stay-Lose-Shift** and **Cooperator** received average scores because their complete cooperation left them open to exploitation from other strategies. **Defector** maintained non-cooperation yet encountered multiple defections which resulted in average scores at around **2800** but occasional high marks when playing forgiving opponents. **Random** at **0.5** performed least effectively because it demonstrated no pattern in facing different opponents resulting in a major reduction in median points achieved.

Strategies which combine cooperative moves with strategic counterattack produce the most desirable outcomes during long-term IPD experiments although both complete cooperative and defector strategies generate substandard results.

V. CONCLUSION

In this research, we explored the effectiveness of various strategies for the Iterated Prisoner's Dilemma, with a focus on the performance of our novel strategies, Spiteful Memory And Quantum Cooperate Defect in comparison to the well-established Tit-for-Tat. Through extensive simulations, we tested these strategies across different environments, including long games, noisy conditions, and scenarios with adaptive or learning opponents.

Our findings suggest that Spiteful Memory shows significant improvements in several areas:

1. It consistently achieved a higher total score across a range of opponent types due to its ability to balance retaliation and cooperation.
2. It demonstrated superior robustness when faced with more complex and deceptive strategies, adapting more effectively than Tit-for-Tat.
3. The strategy's use of memory allowed it to adapt quicker to shifts in opponent behavior, responding more flexibly to changes in strategy.
4. Furthermore, Spiteful Memory exhibited an improved ability to restor
5. e cooperation after conflicts, especially in environments with noise or unintentional defections.
6. While Spiteful Memory outperformed Tit-for-Tat, its computational overhead increases with memory size, posing challenges for real-time applications. Future work could explore hybrid models combining QCD's stochasticity with Spiteful Memory's adaptability, as well as testing on quantum hardware.

These advantages were particularly pronounced in longer games and noisy environments, where memory-based strategies have a clear edge in pattern recognition and forgiveness. Against adaptive opponents, Spiteful Memory showed remarkable resilience, allowing it to perform well even in dynamic and changing conditions.

The Quantum Cooperate Defect (QCD) strategy brings an innovative perspective to game theory by leveraging the principles of quantum computing, specifically superposition and the Heisenberg Uncertainty Principle. By allowing a player's move to exist in a superposition of both cooperation and defection, QCD introduces a layer of unpredictability and adaptability that sets it apart from traditional strategies.

As quantum computing technology continues to evolve, strategies like QCD could pave the way for more optimized and adaptable approaches in competitive environments. Young (2018) [13] explored the potential of quantum strategies in infinitely iterated games, suggesting that such approaches could revolutionize strategic decision-making in game theory.

This strategy demonstrates theoretical advantages, particularly in the initial rounds where unpredictability provides an edge, and over the long term, where its ability to adjust probabilistically based on the opponent's moves could lead to optimal performance. However, the strategy faces significant challenges in practical implementation, as current quantum computing technology is still in its developmental stages

Despite these limitations, QCD offers an exciting glimpse into how quantum principles could revolutionize strategic decision-making in game theory. As quantum computing continues to evolve, strategies like QCD could pave the way for more optimized, unpredictable, and adaptable approaches in competitive environments.

REFERENCES

[1] S. Kuhn, "*Prisoner's dilemma*," Stanford Encyclopedia of Philosophy, https://plato.stanford.edu/entries/prisoner-dilemma/ (accessed Nov. 9, 2024).

[2] R. Axelrod, "*Effective choice in the prisoner's dilemma*," Journal of Conflict Resolution, vol. 24, no. 1, pp. 3–25, Mar. 1980. doi:10.1177/002200278002400101

[3] GeeksforGeeks, "*Real-life applications of game theory*," GeeksforGeeks, https://www.geeksforgeeks.org/real-life-applications-of-game-theory/ (accessed Nov. 9, 2024).

[4] A. D. and B. Nalebuff et al., *"Prisoners' dilemma*," Econlib, https://www.econlib.org/library/Enc/PrisonersDilemma.html (accessed Nov. 9, 2024).

[5] Nowak, Martin, and Karl Sigmund. "*A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game.*" Nature 364.6432 (1993): 56-58.

[6] Molander, Per. "*The optimal level of generosity in a selfish, uncertain environment*." Journal of Conflict Resolution 29.4 (1985): 611-618.

[7] Sugden, Robert. "*The economics of rights, co-operation and welfare*". Basingstoke: Palgrave Macmillan, 2004.

[8] Baek, S. K., & Kim, B. J. (2008). Intelligent tit-for-tat in the iterated prisoner's dilemma game. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 78(1), 011125.

[9] "*Axelrod Python library*" Welcome to the documentation for the Axelrod Python library - Axelrod 4.13.1 documentation, https://axelrod.readthedocs.io/en/stable/ (accessed Nov. 9, 2024).

[10] Press, W. H., & Dyson, F. J. (2012). "Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent." Proceedings of the National Academy of Sciences, 109(26), 10409-10413

[11] Camerer, C. F. (2003). "Behavioral Game Theory: Experiments in Strategic Interaction".

[12] Trivers, R. L. (1971). "The Evolution of Reciprocal Altruism." The Quarterly Review of Biology, 46(1), 35-57. A.

[13] Perc, M., & Szolnoki, A. (2010). "Coevolutionary Games—A Mini Review." BioSystems, 99(2), 109-125.

[14] Nay, J. J., & Vorobeychik, Y. (2016). "*Predicting Human Cooperation*." arXiv preprint arXiv:1601.07792.

[15] Young, R. D. (2018). "Reciprocal and Extortive Strategies: Infinitely Iterated Prisoner's Dilemma." arXiv preprint arXiv:1803.01282.

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

[17] Axelrod, R. & Dion, D ."*The further evolution of cooperation*" (1988) Science 242, 1385-1390.

[18] Nowak, M. "*Chaos and the evolution of cooperation.*" (1989) J. Theor. Biol. 142, 237-241.

[19] Nowak, M."*Stochastic strategies in the prisoner's dilemma*" (1990) Theor. Pop. Biol. 38, 93-112.