Professor W. Bentley MacLeod

Research Assistant: Hana Kwon

**Payoff Distribution by Strategy Category and Game Type**

**Introduction:**

This study examines the distribution of average payoffs per game across different strategy categories in both finite and infinite repeated Prisoner's Dilemma games. By analyzing players' behaviors and outcomes, we aim to understand how strategic choices influence cooperation and overall payoffs. The strategies are categorized into:

* **Only TFT (Tit-for-Tat):** Players whose behavior matches exclusively the Tit-for-Tat (TFT) strategy and no other.
* **TFTA:** Players whose behavior aligns with strategies consistent with TFT, including Grim strategies.
* **All Strategies:** The full set of players, irrespective of strategy adherence.

**Figures Overview:** The provided figures illustrate the payoff distribution across the three strategy categories for both finite and infinite games. These visualizations are designed to explore the behavior and performance of Tit-for-Tat (TFT) and its variants as strategies in repeated Prisoner's Dilemma games.

* **Axes:**
* **Horizontal Axis (X-axis):** Represents the average payoff per game, ranging from negative to positive values.
* **Vertical Axis (Y-axis):** Indicates the frequency of players achieving these average payoffs.
* **Bars:**
  + The histogram displays the distribution of average payoffs for each strategy category:
    - **Red Bars:** Only TFT
    - **Orange Bars:** TFTA
    - **Sky Blue Bars:** All Strategies

도표, 스크린샷, 텍스트, 그래프이(가) 표시된 사진

자동 생성된 설명

**Figure 1: Payoff Distribution for Finite Games**

**Description**: This histogram shows the average payoffs per game for players in finite games (horizon < 10). The x-axis represents the range of average payoffs, while the y-axis indicates the frequency of players achieving these payoffs.

**Insights**:

* **Only TFT**:
  + The distribution shows a concentration of payoffs around lower positive values. This reflects the performance of players strictly adhering to Tit-for-Tat in finite games, where the finite horizon can undermine the incentives for sustained cooperation.
  + In finite games, particularly as the end approaches, players may be tempted to defect, anticipating that there are limited future repercussions. This end-game effect can lead to a breakdown of cooperation, adversely affecting the payoffs of strictly cooperative strategies like TFT.
* **TFTA**:
  + Displays a broader spread of average payoffs, indicating variability in outcomes based on slight deviations from the pure Tit-for-Tat strategy.
  + Players in this category may employ strategies that are more forgiving or punitive, adjusting their behavior in response to their opponents. This adaptability can sometimes mitigate the negative impact of the end-game effect, leading to a wider range of payoffs.
* **All Strategies**:
  + Exhibits the widest distribution, capturing the full spectrum of strategies, including Always Cooperate (AC), Always Defect (AD), and random strategies.
  + The diversity in strategies results in a broad range of average payoffs, both negative and positive, reflecting the heterogeneous nature of player behaviors in finite games.
* **Negative Payoffs:** Some players receive negative payoffs, primarily due to repeated sucker outcomes (coop = 1, opponent coop = 0) or mutual defection penalties.
  + The presence of negative average payoffs is more noticeable in finite games due to the higher likelihood of defection as the game approaches its end. Players experiencing sustained unreciprocated cooperation or frequent mutual defection accumulate negative payoffs over the game.
* **Bins**: Current bin size of 0.2 allows for granular insights but could be adjusted for smoother visualization.

도표, 스크린샷, 그래프, 라인이(가) 표시된 사진

자동 생성된 설명

**Figure 2: Payoff Distribution for Infinite Games**

**Description**: This histogram presents the average payoffs per game for players in infinite games (with a horizon equal to 10, representing an indefinitely repeated game). The structure mirrors that of finite games but reflects differences in strategic behavior due to the longer, uncertain horizon.

**Insights**:

* **Improved Payoffs:** 
  + The distributions shift towards higher average payoffs compared to finite games, reflecting increased cooperation in an infinite horizon setting.
  + The shadow of the future encourages sustained cooperation, leading to better outcomes for cooperative strategies, as players anticipate ongoing interactions with their opponents. This anticipation discourages defection, as the long-term cost of losing cooperative gains outweighs short-term benefits.
* **Only TFT and TFTA**
  + Players adhering to Tit-for-Tat and similar strategies achieve higher average payoffs, evidenced by the concentration of payoffs in the higher range.
  + The effectiveness of these strategies in promoting mutual cooperation over the long term is more pronounced in infinite games, where the threat of future retaliation or the reward of continued cooperation sustains collaborative behavior.
* **All Strategies**:
  + While the distribution still spans a wide range, there is a higher frequency of positive average payoffs compared to finite games.
  + The infinite horizon mitigates the temptation to defect, even among players using varied strategies, as the potential long-term losses from defection are significant.
* **Negative Payoffs:** 
  + Negative average payoffs are less frequent in infinite games due to the reduced incentive to defect and the greater potential cost of breaking cooperation.
  + Players are more likely to maintain cooperation to avoid mutual losses, leading to overall higher average payoffs.

**Implications for the Study**

* **Impact of Game Horizon:**
  + The figures demonstrate that the game's horizon significantly impacts strategic behavior and outcomes.
  + In finite games, the approaching end diminishes the incentives for cooperation, leading to more frequent defection and lower average payoffs, especially for cooperative strategies.
  + In infinite games, the uncertainty about when the game will end fosters an environment where cooperation is the rational long-term strategy, resulting in higher average payoffs.
* **Effectiveness of Cooperative Strategies:**
  + Tit-for-Tat and similar strategies are more effective in infinite games, where they can sustain mutual cooperation and maximize payoffs.
  + The adaptability of TFTA strategies allows players to respond to their opponents' behavior, enhancing cooperation even when facing occasional defection.
* **Risks of Unreciprocated Cooperation:**
  + The presence of negative payoffs underscores the risks associated with unreciprocated cooperation.
  + Players who consistently cooperate without reciprocation from their opponents incur losses, highlighting the importance of strategy selection and the ability to respond appropriately to defection.
* **Strategic Diversity:**
  + The wide range of average payoffs in the ‘All Strategies’ category reflects the diversity of human decision-making in strategic settings.
  + Understanding the distribution of strategies and their outcomes can inform the development of models that better predict behavior in repeated games.