Beyond Tit-for-Tat Proposal - Melt

Liam Andrew Beattie[a], Abdul Qaadir Cassiem[a]

[a] *Microeconomics 871, Stellenbosch University, South Africa*

Abstract

Beyond Tit-for-tat: Set up a repeated prisoner's dilemma computer tournament, in which strategies compete against each other. Write a report on your findings.

1. Introduction

The phrase horses for courses alludes to the fact that a racehorse performs best on a racecourse to which it is specifically suited. More generally this idiom is used to express that certain tools and strategies are better suited over others depending on the task or situations at hand. In the context of the repeated prisoners' dilemma, the strategy of tit-for-tat, where one mimics their opponent's previous move, reigns supreme and is best suited over others for the situation at hand[1].

The question this paper aims to answer is as to which situations is tit-for-tat not the dominant strategy. To do this we have to venture down two potential avenues. The first is the adjustment of pay-off values within games, and the second is adjusting pay-off values from games. Consider a standard prisoners' dilemma pay-off table:

Player 1 / Player 2	**C (Cooperate)**	**D (Defect)**
C (Cooperate)	(R, R)	(S, T)
D (Defect)	(T, S)	(P, P)

Table 1.1: Prisoner's Dilemma Payoff Matrix with R, P, S, and T Outcomes

Adjusting the values of R (Reward for mutual cooperation), P (Punishment for mutual defection), S (Sucker's pay-off for cooperating while the other defects), and T (Temptation to defect when the other cooperates) is an example of within game pay-off adjustments. These adjustments might produce a new dominant strategy and our analysis aims to find if it does.

Email addresses: 22562435@sun.ac.za (Liam Andrew Beattie), 20863667@sun.ac.za (Abdul Qaadir Cassiem)

[1] The tic-for-tat strategy is the dominant strategy in Axelrod (1980)

From-game adjustments are a bit different and it considers the utility a player gets from the payoffs of its opponent.

Player 1 / Player 2	C (Cooperate)	D (Defect)
C (Cooperate)	$(R(1-p)+Rp, R(1-p)+Rp)$	$(S(1-p)+Tp, T(1-p)+Sp)$
D (Defect)	$(T(1-p)+Sp, S(1-p)+Tp)$	$(P(1-p)+Pp, P(1-p)+P)$

Table 1.2: Prisoner's Dilemma Payoff Matrix

The level p here is adapted from Charness & Rabin (2002) who created a utility function that captures various social preferences. In essence, p is how much you care about your opponent's pay-offs as well as your own. In standard prisoners' dilemma games, this is 0 and thus people are purely self-interested. If we let our pay-offs be R = 3, T = 5, S = 0, and P = 3, then this situation in strategic form would look like:

Player 1 / Player 2	C (Cooperate)	D (Defect)
C (Cooperate)	$(3,3)$	$(0,5)$
D (Defect)	$(5,0)$	$(1,1)$

Table 1.3: Prisoner's Dilemma Payoff Matrix for $p = 0$ (Self-interested person)

However we can adjust the value of p for people who are partially considerate of other people's outcomes, or we can make people egalitarian who care just as much for others as they do for themselves.

Player 1 / Player 2	C (Cooperate)	D (Defect)
C (Cooperate)	$(3,3)$	$(1,4)$
D (Defect)	$(4,1)$	$(1,1)$

Table 1.4: Prisoner's Dilemma Payoff Matrix for $p = 0.2$ (Partially considers others' outcomes)

Player 1 / Player 2	C (Cooperate)	D (Defect)
C (Cooperate)	$(3,3)$	$(2.5,2.5)$
D (Defect)	$(2.5,2.5)$	$(1,1)$

Table 1.5: Prisoner's Dilemma Payoff Matrix for $p = 0.5$ (Egalitarian person)

p could also take a negative value, which indicates a person is status-seeking and actively wants to bring down their opponent.

Player 1 / Player 2	C (Cooperate)	D (Defect)
C (Cooperate)	$(3.6, 3.6)$	$(-1, 6)$
D (Defect)	$(6, -1)$	$(0.8, 0.8)$

Table 1.6: Prisoner's Dilemma Payoff Matrix for $p = -0.2$ (Negative influence by others' outcomes)

It would be interesting to see under which values of p the dominant strategy changes.

2. Literature Review

We aim to do a short literature review and provide insight from the following sources: Lange & Baylor (2007), Farrell & Ware (1989), Kreps, Milgrom, Roberts & Wilson (1982), Romero & Rosokha (2018), Bó & Fréchette (2019), Breitmoser (2015), Gaudesi, Piccolo, Squillero & Tonda (2016), García & Veelen (2018), Embrey, Fréchette & Yuksel (2018).

Most importantly we aim to structure our output in tables similar to Axelrod (1980).

3. Game Construction

1. Always Strategies:

- Always Cooperate
- Always Defect

2. Tit for Tat Variants:

- Tit for Tat
- Tit for Two Tats
- Tit for Tat with Forgiveness
- Tit for Tat with Randomization

3. Win-Stay/Lose-Switch Strategies:

- Pavlov

4. Punishment-Based Strategies:

- Grim Trigger
- Bully
- Retaliatory Defector

5. Adaptive/Adjusting Strategies:

- Adaptive Defector
- Adaptive Peacekeeper
- Probing Adjuster
- Forgiving Tester
- Prober
- Cautious Rebuilder

6. Gradient/Probability-Based Strategies:

- Progressive Cooperator (Gradually increases cooperation)
- Diminishing Cooperator (Gradually decreases cooperation)
- Bounded Gradient (Probability based on the opponent's entire history)
- Recent Gradient (Probability based on recent actions of the opponent)

7. Random Strategies:

- Random 10%
- Random 25%
- Random 50%
- Random 75%
- Random 90%

Game tournaments will take place in a round-robin format where all strategies play each other for N number of games. Total utility is calculated over the whole tournament and the strategy with the greatest value will be the dominant strategy. The tournaments will not be evaluated on games won but this metric will be tracked.

We have not limited ourselves to the number of strategies just yet, but we aim to include most of the following and potentially we create more along the way.

Basic Strategies:

- Always Cooperate: This strategy always cooperates, regardless of the opponent's previous moves.
- Always Defect: This strategy always defects, regardless of the opponent's previous moves.
- Tit-for-Tat (TFT): Cooperates on the first move, then mimics the opponent's last move in subsequent rounds.
- Grim Trigger: Cooperates until the opponent defects once, then defects forever.
- Random: Randomly chooses to cooperate or defect with some probability.
- Tit-for-Two-Tats: Similar to Tit-for-Tat but defects only after two consecutive defections by the available player.
- Pavlov (Win-Stay, Lose-Shift): Cooperates if the last round was a success (mutual cooperation or mutual defection), otherwise defects.

More Complex Strategies:

- Generous Tit-for-Tat: Similar to Tit-for-Tat, but occasionally forgives a defection.
- Tit-for-Tat with Randomisation: A variant of Tit-for-Tat where the player may defect or cooperate with a certain probability after the opponent defects.
- Tit-for-Tat with Forgiveness: Like TFT but occasionally forgives a defection, returning to cooperation.

4. Feedback

Any feedback would be greatly appreciated.

Will be powerful to give a snapshot of one turnoment

Table 4.1: Strategy Rankings Across Different p Values

Strategy	-0.1	-0.05	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
Always Cooperate	61	58.75	63	58.5	60	61.5	65	66.25	69	66.25	69	72	72
Always Defect	115	110	105	103.75	91.5	90	85	82.75	72.5	65.5	65	56.5	55
Tit for Tat	57.5	58.75	60	58.5	65	61.5	63	66.25	67.5	68.75	69	71.25	73
Tit for Two Tats	61	52.25	54	58.5	62.5	66	65	66.25	67.5	68.75	69	69.75	72.5
Tit for Tat with Forgiveness	57.5	55.5	54	64	60	63.75	67	66.25	67.5	67.5	67	70.5	72
Tit for Tat with Randomisation	54	55.5	57	64	62.5	66	65	66.25	64.5	68.75	68	70.5	72
Pavlov	50.5	52.25	60	61.25	62.5	61.5	65	66.25	67.5	66.25	70	70.5	72
Grim/Trigger	54	55.5	60	55.75	62.5	63.75	63	68	66	70	70	70.5	72.5
Bully	110.5	110	105	100	95	83.5	85	77.25	75	70	61	56.5	55
Retaliatory Defector	57.5	58.75	57	64	57.5	61.5	63	62.75	67.5	70	70	71.25	72
Adaptive Defector	57.5	55.5	60	58.5	62.5	59.25	65	66.25	69	67.5	69	70.5	73
Adaptive Peacekeep	57.5	58.75	60	58.5	60	66	65	66.25	66	68.75	70	72	72.5
Probing Adjuster	61	62	60	61.25	62.5	66	67	68	67.5	67.5	69	70.5	73
Forgiving Tester	57.5	58.75	57	61.25	60	61.5	63	62.75	64.5	67.5	71	69.75	72
Prober	57.5	55.5	60	58.5	65	63.75	65	64.5	66	67.5	69	71.25	72
Cautious Rebuilder	54	58.75	57	58.5	60	63.75	61	62.75	66	67.5	70	71.25	71.5
Progressive Cooperator	115	110	101	96.25	98.5	86.75	80	80	74.5	67.75	61	58.25	56.5
Deminishing Cooperator	54	62	60	59.25	60	63.75	67	62.75	64.5	68.75	69	71.25	72
Bounded Gradient	57.5	58.75	60	58.5	62.5	63.75	63	64.5	69	67.5	69	72	72
Recent Gradient	54	58.75	60	58.5	62.5	59.25	69	64.5	67.5	67.5	70	69.75	72.5
Random 10%	113.5	103.25	105	94.75	91.5	89.75	80	75	75.5	68.25	64	58.75	57.5
Random 25%	93.5	98.25	91	85	90	81	77	76.25	73.5	68.25	69	62.25	58.5
Random 50%	87.5	70.75	83	80.75	77	80.75	72	72.25	66.5	69	67	64.5	60.5
Random 75%	77	74.75	69	69.5	66	69.5	67	69.75	67	66.5	69	68.5	69
Random 90%	65	66	65	63.75	59	70.25	67	65.75	67.5	69	70	68.25	71.5

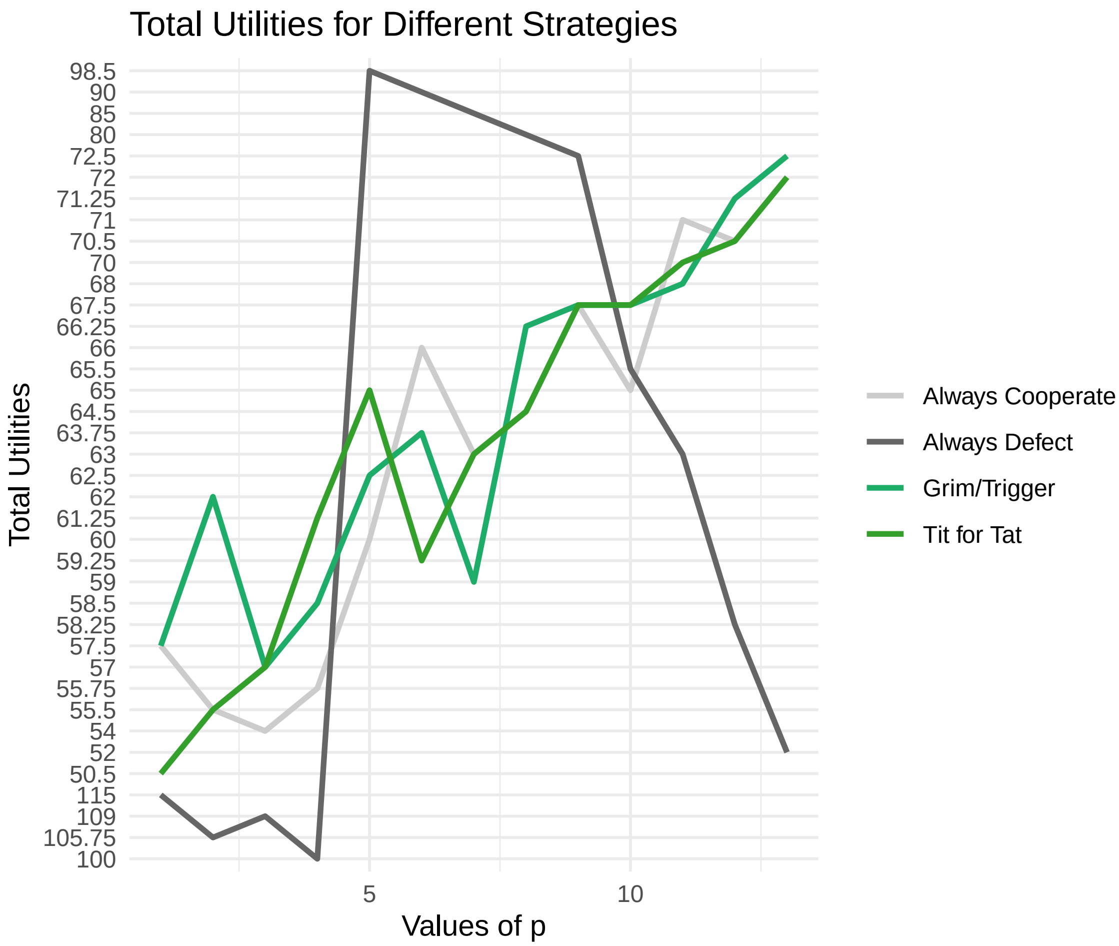

Figure 4.1: Strategy Ranks Across Different Values of p

Defector Strategies Definition: These strategies lean towards defection more often than cooperation. They may defect unconditionally, strategically exploit others, or only cooperate under very specific conditions. Defector strategies prioritize short-term gains, aiming to maximize their own payoff without much regard for fostering mutual cooperation.

Cooperative Strategies Definition: These strategies favour cooperation and are designed to foster mutual cooperation over multiple rounds. They aim to achieve better outcomes for both players through sustained cooperation, but they often include some mechanisms for retaliation to avoid being exploited.

Increased Temptation: c(3, 8, 0, 1) Reduced Punishment: c(3, 5, 0, 2) Increased Reward for Cooperation: c(5, 3, 0, 1)

Increased Temptation: c(3, 8, 0, 1) This structure provides a strong test of a strategy's ability to resist short-term gains for long-term benefits. It is particularly useful for identifying strategies that are susceptible to exploitation.

Reduced Punishment: c(3, 5, 0, 2) This structure will encourage more risk-taking behavior, as the consequences of mutual defection are less severe. It is a good way to evaluate how strategies respond to a less punishing environment.

Increased Reward for Cooperation: c(5, 3, 0, 1) This structure will test a strategy's ability to incentivize and maintain cooperation, even when the temptation to defect remains. It is especially useful for identifying strategies that are effective in promoting mutual benefit.

5. Bump Chart

References

Axelrod, R. 1980. Effective choice in the prisoner's dilemma. *The Journal of Conflict Resolution*. 24(1):3–25. [Online], Available: http://www.jstor.org/stable/173932.

Bó, P.D. & Fréchette, G.R. 2019. Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review*. 109(11):3929–3952.

Breitmoser, Y. 2015. Cooperation, but no reciprocity: Individual strategies in the repeated prisoner's dilemma. *American Economic Review*. 105(9):2882–2910.

Embrey, M., Fréchette, G.R. & Yuksel, S. 2018. Cooperation in the finitely repeated prisoner's dilemma. *The Quarterly Journal of Economics*. 133(2):509–551.

Farrell, J. & Ware, R. 1989. Evolutionary stability in the repeated prisoner's dilemma. *Journal of Economic Theory*. 47(1):1–12.

García, J. & Veelen, M. van. 2018. No strategy can win in the repeated prisoner's dilemma: Linking game theory and computer simulations. *Frontiers in Robotics and AI*. 5:102.

Gaudesi, M., Piccolo, E., Squillero, G. & Tonda, A. 2016. Exploiting evolutionary modeling to prevail in iterated prisoner's dilemma tournaments. *IEEE Transactions on Computational Intelligence and AI in Games*. 8(3):235–247.

Kreps, D.M., Milgrom, P., Roberts, J. & Wilson, R. 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*. 27(2):245–252.

Lange, C. & Baylor, A.L. 2007. Teaching the repeated prisoner's dilemma with a computerized tournament. *The Journal of Economic Education*. 38(4):407–418.

Romero, J. & Rosokha, Y. 2018. Constructing strategies in the indefinitely repeated prisoner's dilemma game. *European Economic Review*. 104:185–219.