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: Tournament Payoff Matrix for p= 0

	Payoff Against Other Strategies																									Total	Rank
	AC	AD	TfT	Tf2T	TfTF	TfTR	P	G/T	B	RD	ADe	APe	PA	FT	P	CR	PC	DC	BG	RG	R0.1	R0.25	R0.5	R0.75	R0.9		
Always Cooperate	600	0	600	600	600	600	600	600	147	600	600	501	3	453	600	483	291	261	600	600	69	156	300	420	519	**10803**	**24**
Always Defect	1000	200	204	208	276	204	204	204	396	992	204	212	208	212	1000	212	664	600	204	1000	276	392	648	792	896	**11408**	**20**
Tit for Tat	600	199	600	600	600	600	600	600	346	600	600	567	399	551	600	561	441	426	600	600	250	344	432	531	580	**12827**	**4**
Tit for Two Tats	600	198	600	600	600	600	600	600	296	600	600	501	298	453	600	483	372	370	600	600	244	288	394	480	548	**12125**	**15**
Tit for Tat with Forgiveness	600	183	600	600	600	600	600	600	322	600	600	561	370	539	600	553	408	411	600	600	249	321	403	527	571	**12618**	**8**
Tit for Tat with Randomisation	600	199	600	600	600	600	600	600	346	600	600	567	399	551	600	561	422	436	600	600	252	357	478	542	577	**12887**	**2**
Pavlov	600	199	600	600	600	600	600	600	346	600	600	567	399	551	600	561	453	434	600	600	243	325	442	519	582	**12821**	**5**
Grim/Trigger	600	199	600	600	600	600	600	600	395	600	600	223	205	219	600	221	607	575	600	600	263	399	551	823	933	**12813**	**6**
Bully	902	151	351	551	399	351	351	155	298	894	155	683	355	751	902	631	533	554	348	902	230	319	556	700	847	**12869**	**3**
Retaliatory Defector	600	2	600	600	600	600	600	600	149	600	600	505	9	457	600	487	263	283	600	600	74	122	319	463	541	**10874**	**23**
Adaptive Defector	600	199	600	600	600	600	600	600	395	600	600	501	205	453	600	483	448	420	600	600	276	387	457	487	548	**12459**	**10**
Adaptive Peacekeep	666	197	567	666	576	567	567	213	263	660	666	534	256	503	666	543	418	400	572	666	223	285	393	483	600	**12150**	**14**
Probing Adjuster	998	198	404	553	436	404	404	205	345	989	205	591	400	557	998	643	587	544	325	998	264	369	549	719	890	**13575**	**1**
Forgiving Tester	698	197	551	698	563	551	551	209	246	692	698	583	297	502	698	572	452	431	565	698	239	275	393	543	631	**12533**	**9**
Prober	600	0	600	600	600	600	600	600	147	600	600	501	3	453	600	483	324	303	600	600	63	156	318	492	540	**10983**	**21**
Cautious Rebuilder	678	197	561	678	576	561	561	211	276	672	678	573	238	522	678	522	413	385	554	678	229	270	402	541	641	**12295**	**11**
Progressive Cooperator	798	102	439	557	484	417	414	105	271	794	444	546	203	503	790	504	417	467	289	798	159	247	506	658	742	**11654**	**18**
Deminishing Cooperator	784	109	438	536	454	432	447	141	263	799	410	543	235	487	818	522	499	431	619	798	176	262	435	653	715	**12006**	**16**
Bounded Gradient	600	199	600	600	600	600	600	600	342	600	600	564	245	520	600	545	513	325	600	600	273	374	442	540	566	**12648**	**7**
Recent Gradient	600	0	600	600	600	600	600	600	147	600	600	501	3	453	600	483	285	312	600	600	84	150	291	462	549	**10920**	**22**
Random 10%	960	184	255	340	354	251	267	182	381	966	184	325	266	375	938	351	589	560	242	976	238	381	542	812	882	**11801**	**17**
Random 25%	886	149	367	452	415	353	354	156	329	900	233	533	314	406	900	508	553	498	353	904	216	364	564	695	809	**12211**	**13**
Random 50%	794	100	428	654	497	460	469	105	260	803	480	634	259	576	806	565	427	467	448	806	188	257	445	625	710	**12263**	**12**
Random 75%	714	59	538	656	538	525	536	77	217	692	643	609	151	536	726	574	377	358	530	704	107	224	325	532	660	**11608**	**19**
Random 90%	630	21	570	634	581	572	577	64	169	628	631	541	94	471	620	517	319	324	555	634	76	147	329	460	591	**10755**	**25**

Table 4.2: Strategy Rankings Across Different p Values

	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	25	24	25	22	19	19	14	10	4	4	4	1	3
Always Defect	15	18	20	25	25	25	25	25	25	25	25	25	25
Tit for Tat	5	4	3	3	3	1	2	4	5	8	8	11	9
Tit for Two Tats	18	16	15	13	11	7	6	5	9	6	6	6	6
Tit for Tat with Forgiveness	10	8	7	5	4	4	4	3	6	7	7	7	8
Tit for Tat with Randomisation	4	6	5	4	1	2	1	2	7	9	10	8	11
Pavlov	6	5	2	2	2	3	3	1	8	10	9	9	10
Grim/Trigger	3	3	4	7	9	9	15	18	18	18	19	19	20
Bully	2	2	8	8	10	15	20	21	22	21	21	21	21
Retaliatory Defector	23	21	22	20	17	16	10	9	1	3	2	3	4
Adaptive Defector	12	10	11	9	7	8	7	11	13	13	13	13	14
Adaptive Peacekeep	17	14	13	14	13	10	12	13	15	14	14	14	13
Probing Adjuster	1	1	1	1	5	12	19	20	21	22	22	22	22
Forgiving Tester	9	9	9	10	8	6	8	12	14	15	15	16	16
Prober	21	22	24	24	21	17	13	8	3	1	1	2	1
Cautious Rebuilder	14	12	12	12	12	11	9	16	16	16	16	15	15
Progressive Cooperator	19	20	17	18	23	22	22	22	20	20	20	20	19
Deminishing Cooperator	16	17	16	15	16	20	21	19	19	19	18	18	18
Bounded Gradient	8	7	6	6	6	5	5	6	10	11	12	12	12
Recent Gradient	22	25	23	21	18	18	11	7	2	2	3	4	2
Random 10%	11	15	18	19	24	24	24	24	24	24	24	24	24
Random 25%	7	13	14	16	19	23	23	23	23	23	23	23	23
Random 50%	13	11	10	11	14	13	18	17	17	17	17	17	17
Random 75%	20	19	19	17	15	14	16	15	12	12	11	10	7
Random 90%	24	23	21	23	22	21	17	14	11	5	5	5	5

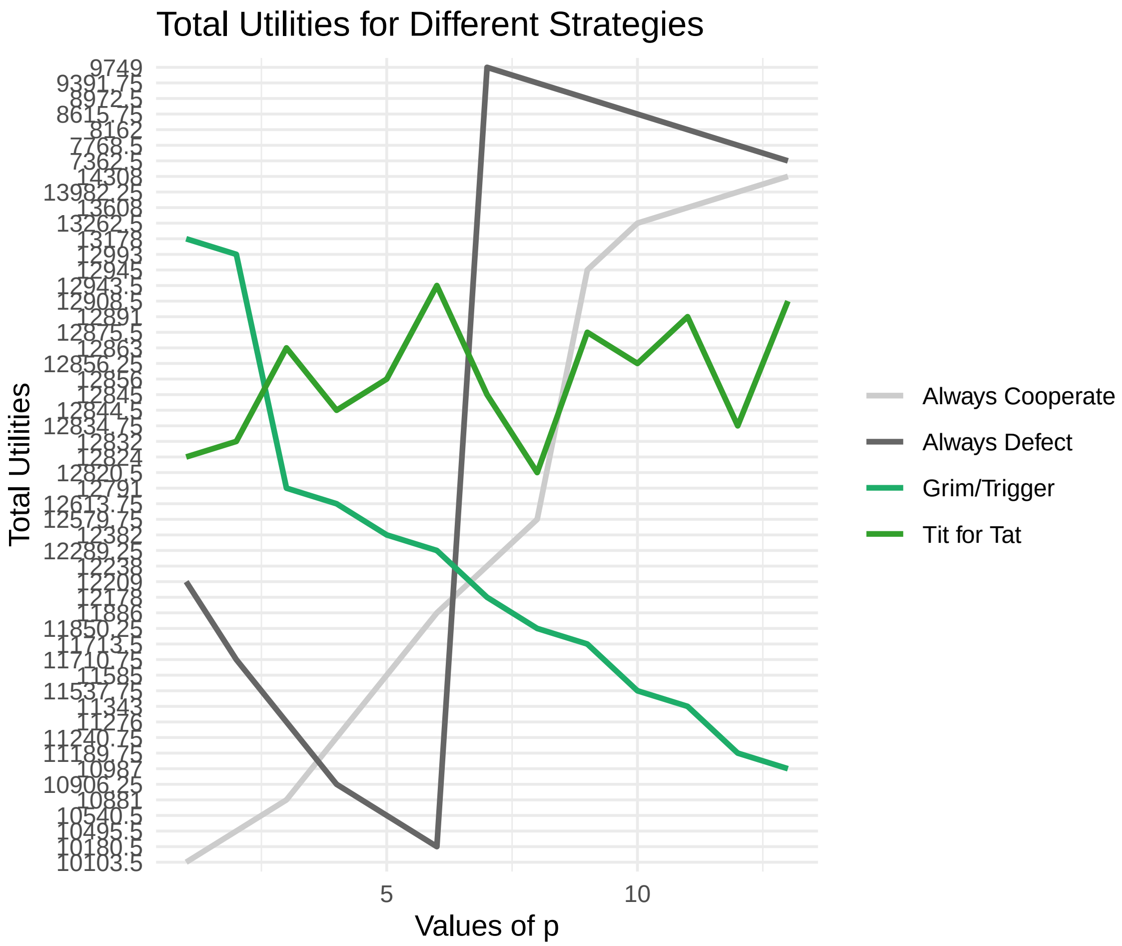

Figure 4.1: Strategy Ranks Across Different Values of p

Table 4.3: Strategy Values Across Different p Values

	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	10103.5	10495.5	10881	11240.75	11585	11886	12238	12579.75	12945	13262.5	13608	13982.25	14308
Always Defect	12209	11710.75	11276	10906.25	10540.5	10180.5	9749	9391.75	8972.5	8615.75	8162	7768.5	7362.5
Tit for Tat	12824	12832	12863	12844.5	12856	12943.5	12845	12820.5	12875.5	12856.25	12891	12834.75	12908.5
Tit for Two Tats	11897.5	11983.5	12067	12221.25	12304	12453.5	12584	12701.5	12786.5	12878.25	13021	13159.75	13310.5
Tit for Tat with Forgiveness	12551	12573.75	12692	12711.5	12787.5	12766	12809	12878	12863.5	12874.5	12936	12926	13004.5
Tit for Tat with Randomisation	12883.5	12801.25	12772	12837.25	12864.5	12836.5	12884	12886.75	12854	12843.25	12806	12920.5	12813.5
Pavlov	12818	12807.5	12902	12866.5	12864	12787.25	12815	12891.5	12851	12797.25	12875	12919.5	12849.5
Grim/Trigger	13178	12993	12791	12613.75	12382	12289.25	12178	11850.25	11713.5	11537.75	11343	11189.75	10987
Bully	13447.5	13160	12685	12513.75	12312	11926.25	11688	11410	11103	10778.75	10552	10240.25	9950
Retaliatory Defector	10192.5	10593	10972	11274.25	11629.5	11915.75	12296	12581	12963.5	13270.25	13616	13955.5	14294.5
Adaptive Defector	12417.5	12400.25	12365	12433	12444.5	12452	12465	12511.5	12474.5	12522.75	12666	12661.75	12520.5
Adaptive Peacekeep	12071.5	12127	12145	12154.5	12277	12286.75	12271	12382	12370	12455.5	12393	12479.75	12549
Probing Adjuster	14349.5	13907	13434	13067	12773.5	12220.5	11878	11552.25	11306	10737.25	10455	10093.5	9762.5
Forgiving Tester	12574	12528	12549	12374.5	12437	12464.5	12427	12404	12399.5	12352.25	12372	12251.25	12264.5
Prober	10313.5	10580	10884	11166.5	11537.5	11910.75	12248	12593.75	12948	13302.5	13637	13962	14313.5
Cautious Rebuilder	12268.5	12276.25	12152	12263.5	12287	12258.25	12329	12265.5	12303.5	12299	12363	12403	12482.5
Progressive Cooperator	11735.5	11501.25	11803	11680.25	11437.5	11518.25	11505	11369	11387.5	11359.5	11262	11146.25	11245.5
Deminishing Cooperator	12135	11963.25	11882	12028	11763.5	11817	11676	11703	11622.5	11501.75	11485	11439	11315.5
Bounded Gradient	12641.5	12580.75	12724	12688.75	12752.5	12726.5	12737	12661.25	12728.5	12713.75	12727	12690	12741.5
Recent Gradient	10240	10414.25	10887	11246.25	11602.5	11904	12272	12642.75	12951	13291.25	13613	13953.75	14311
Random 10%	12488	12072.5	11518	11561	11011	10884.75	10554	10170.75	9691.5	9407	9083	8798.25	8374.5
Random 25%	12705.5	12206.75	12123	11906	11585	11484	11173	10777.75	10706	10359.75	10201	9980.25	9710
Random 50%	12391.5	12384.25	12371	12330.25	12168	12000	11914	11968.5	11827.5	11699.25	11690	11614.5	11613.5
Random 75%	11350	11521	11512	11756.75	11776	11989.25	12117	12323	12484	12661.25	12782	12911	13015.5
Random 90%	10159.5	10523.5	10980	11167.75	11493	11775.25	12019	12379.75	12656.5	12909.5	13192	13475	13753.5

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.