Prisoner's Dilemma Revisited

Tit-for-Tat is not the best strategy

Stuart Ferguson · Follow

5 min read · Jan 19, 2024

YouTuber *Veritasium* has an excellent video about prisoner's dilemma.

I've written about this myself. We reported the same conclusion — that *tit-for-tat* is the best strategy for iterated prisoner's dilemma. We conclude more generally that successful strategies will be: nice, retaliatory but also forgiving.

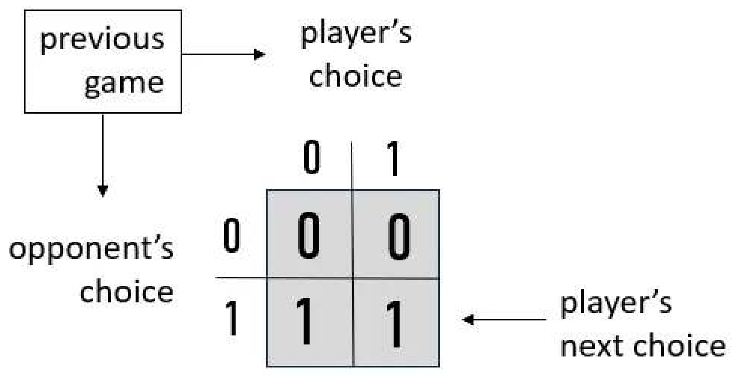
Problem is it's not true.

I have long been fascinated by this result, and yet all of my attempts to replicate it have failed. I'm not claiming that Axelrod's results are fake or that his program had bugs. The problem is that his tournament approach very much depends on the combination of strategies that are playing against each other. Axelrod basically picked a set of strategies at random — things that his friends thought might be good — and used those to seed his tournament.

That's neither complete nor objective.

Completing the Set

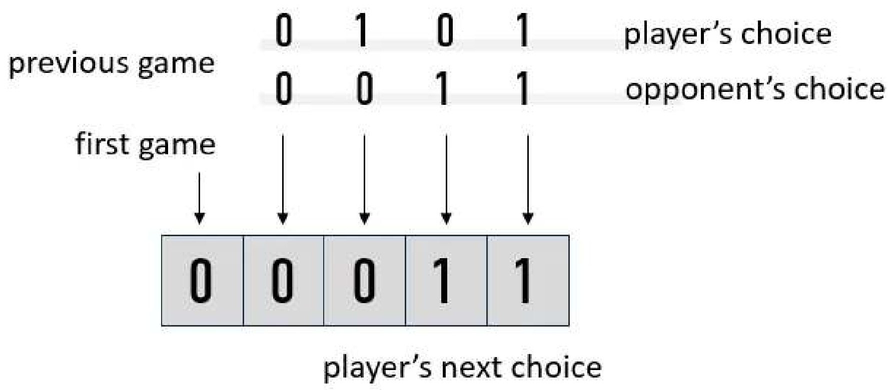
My approach was to test tit-for-tat against all other possible similar strategies. Tit-for-tat decides how to play based entirely on the previous game. Here's the truth-table for tit-for-tat, where 0 means cooperation and 1 means defection.



Truth-table for player's next choice based on previous game

The next choice the strategy makes echos the last choice that the opponent made, but there are other possible truth-tables: 16 in total. We also need to know the first play the strategy makes since that's not based on a previous game, which doubles the number to 32.

Here's the complete description of tit-for-tat, which starts by cooperating, expressed as a 5-digit binary number.



Tit-for-tat strategy as five-digit binary number

While there are 32 possible strategies from 00000 (always cooperate) to 11111 (always defect), it's not clear that they are all unique. To classify strategies by what they do in practice I compute a signature. I play the strategy against the first three moves of an opponent and record its four choices. Here is what tit-for-tat looks like against an opponent that defects for the first three rounds (the fourth round doesn't matter).

round	1	2	3	4		
opponent	1	1	1	-		
tit-for-tat	0	1	1	1	=	7

Four rounds of tit-for-tat as hexadecimal digit

These four bits are a single hexadecimal digit — 7 in this case. If we try all 8 possible openers we can record 8 such digits, and so the signature for tit-for-tat is 7654–3210.

```
1 1 1 -         0 1 1 -
0 1 1 1   7     0 0 1 1   3

1 1 0 -         0 1 0 -
0 1 1 0   6     0 0 1 0   2

1 0 1 -         0 0 1 -
0 1 0 1   5     0 0 0 1   1

1 0 0 -         0 0 0 -
0 1 0 0   4     0 0 0 0   0
```

All possible 4-round games for tit-for-tat

Here are all the signatures for all 32 possible strategies.

	signature	111-	110-	101-	100-	011-	010-	001-	000-
00000	0000-0000	0000	0000	0000	0000	0000	0000	0000	0000
00001	0000-0000	0000	0000	0000	0000	0000	0000	0000	0000
00010	5454-2210	0101	0100	0101	0100	0010	0010	0001	0000
00011	7654-3210	0111	0110	0101	0100	0011	0010	0001	0000
00100	0000-0000	0000	0000	0000	0000	0000	0000	0000	0000
00101	0000-0000	0000	0000	0000	0000	0000	0000	0000	0000
00110	5467-2310	0101	0100	0110	0111	0010	0011	0001	0000
00111	7777-3310	0111	0111	0111	0111	0011	0011	0001	0000
01000	0122-4545	0000	0001	0010	0010	0100	0101	0100	0101
01001	0132-7645	0000	0001	0011	0010	0111	0110	0100	0101
01010	5555-5555	0101	0101	0101	0101	0101	0101	0101	0101
01011	7655-7655	0111	0110	0101	0101	0111	0110	0101	0101
01100	0123-4567	0000	0001	0010	0011	0100	0101	0110	0111
01101	0133-7777	0000	0001	0011	0011	0111	0111	0111	0111
01110	5567-5567	0101	0101	0110	0111	0101	0101	0110	0111
01111	7777-7777	0111	0111	0111	0111	0111	0111	0111	0111

Signatures for 16 strategies that start with cooperation

	signature	111-	110-	101-	100-	011-	010-	001-	000-
10000	8888-8888	1000	1000	1000	1000	1000	1000	1000	1000
10001	fecc-8888	1111	1110	1100	1100	1000	1000	1000	1000
10010	aa98-aa98	1010	1010	1001	1000	1010	1010	1001	1000
10011	fedc-ba98	1111	1110	1101	1100	1011	1010	1001	1000
10100	8888-ccef	1000	1000	1000	1000	1100	1100	1110	1111
10101	ffff-ffff	1111	1111	1111	1111	1111	1111	1111	1111
10110	ab98-dcef	1010	1011	1001	1000	1101	1100	1110	1111
10111	ffff-ffff	1111	1111	1111	1111	1111	1111	1111	1111
11000	89aa-89aa	1000	1001	1010	1010	1000	1001	1010	1010
11001	fecd-89ba	1111	1110	1100	1101	1000	1001	1011	1010
11010	aaaa-aaaa	1010	1010	1010	1010	1010	1010	1010	1010
11011	fedd-baba	1111	1110	1101	1101	1011	1010	1011	1010
11100	89ab-cdef	1000	1001	1010	1011	1100	1101	1110	1111
11101	ffff-ffff	1111	1111	1111	1111	1111	1111	1111	1111
11110	abab-ddef	1010	1011	1010	1011	1101	1101	1110	1111
11111	ffff-ffff	1111	1111	1111	1111	1111	1111	1111	1111

Signatures for 16 strategies that start with defection

Four strategies are equivalent to always defect (0000–0000), and 4 are equivalent to always cooperate (ffff-ffff). Other than that they are all distinct, leaving us with 26 unique strategies.

Many of these are not particularly good, nor should we expect them to be. An interesting example are 01010 (5555–5555) and 11010 (aaaa-aaaa), which always play the opposite of what they played last. They alternate back and forth no matter what the opponent does, and the only difference is whether they start by defecting or cooperating. As we'll see these end up right in the middle of the pack with identical scores.

Tournament Results

My tournament rules were the same as Axelrod's: every strategy plays every strategy including itself for 200 rounds. Here are the scores.

1.	00111	(7777-3310)	16731
2.	11110	(abab-ddef)	15678
3.	11111	(ffff-ffff)	15600
4.	01111	(7777-7777)	15561
5.	10110	(ab98-dcef)	15177
6.	11100	(89ab-cdef)	14490
7.	00110	(5467-2310)	14222
8.	01110	(5567-5567)	14168
9.	00011	(7654-3210)	12613
10.	10100	(8888-ccef)	12213
11.	01011	(7655-7655)	12097
12.	00010	(5454-2210)	11880
13.	11010	(aaaa-aaaa)	11700
14.	01010	(5555-5555)	11700
15.	01101	(0133-7777)	11382
16.	10010	(aa98-aa98)	11304
17.	11011	(fedd-baba)	11226
18.	10011	(fedc-ba98)	10527
19.	11000	(89aa-89aa)	9231
20.	01100	(0123-4567)	9170
21.	11001	(fecd-89ba)	8949
22.	01001	(0132-7645)	8452
23.	01000	(0122-4545)	8016
24.	10000	(8888-8888)	7839
25.	00000	(0000-0000)	7800
26.	10001	(fecc-8888)	6474

Scores for all 26 unique strategies

The winning strategy is 00111 (7777–3310). This might be described as *permanent retaliation*. It starts by cooperating and continues to cooperate against a cooperating opponent, but as soon as it sees the opponent defect once, it defects after that and continues to defect no matter what the opponent does later.

It's tit-for-tat without the forgiveness.

The second place strategy is 11110 (abab-ddef), which always defects except in the case where it and its opponent both defected in the previous game. This is strangely better than always defecting when playing against this mix of other strategies, and it nets a lot of points from cooperation.

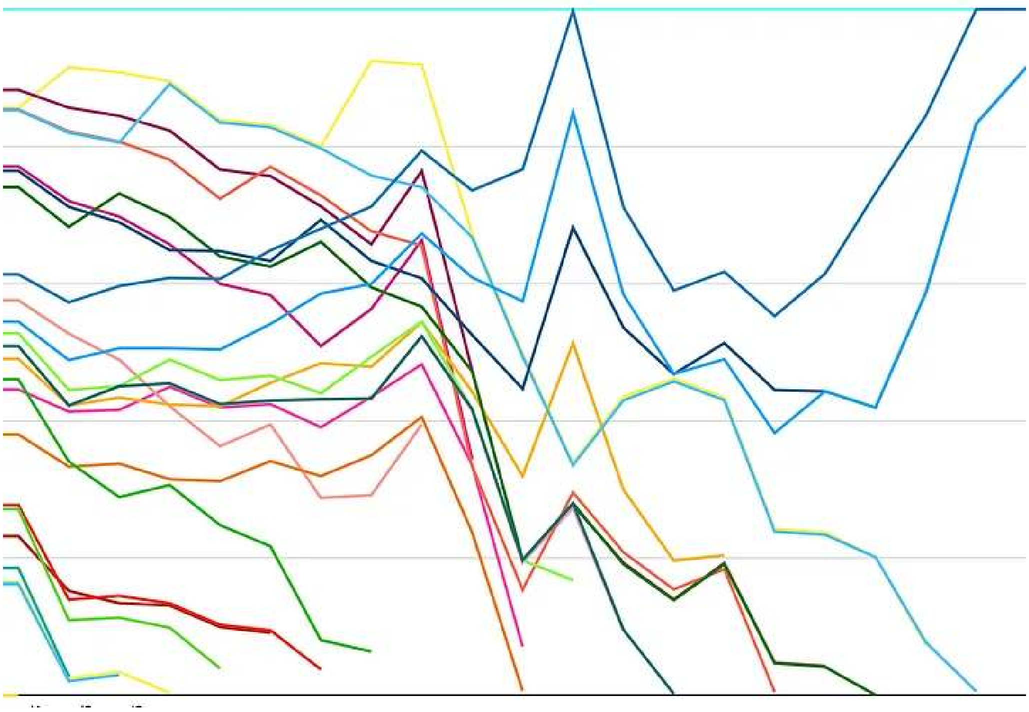
Third place is 11111 (ffff-ffff) — always defect. This scores highly because there are many strategies that cooperate even without getting cooperation in return, like the second place strategy above for example.

There are five better strategies before we get to tit-for-tat — 00011 (7654–3210) — in ninth place. It's solidly in the middle of the top strategies, but it's by no means best or even close to best.

The very worst strategy is 10001 (fecc-8888) which is essentially the opposite of the winning strategy. It starts by defecting but if the opponent ever cooperates then this strategy cooperates forever after that. If 00111 is a cynical cooperator, then 10001 is an ever-hopeful defector.

Testing Robustness

To see if I could find robust winners I removed the lowest-scoring strategy and computed new scores with the reduced pool, and repeated that until there were only 4 strategies remaining. The process is quite chaotic.



Rankings of the dwindling pool during elimination

As rival strategies are eliminated, the ranking of other strategies rearrange themselves significantly. The only one that remains consistently at the top, the light teal line, is 00111.

The four winners were 00111 (7777–3310), 00011 (7654–3210), 00110 (5467–2310), and 00010 (5454–2210). All of these strategies retaliate to varying degrees, but always cooperate when faced with an opponent that always cooperates. This is indicated by the zero at the end of the signature. These are the only strategies with that signature, and they are the four that come out on top.

Conclusions

Pitting all 26 unique single-game strategies against each other (also true for the non-unique 32) contradicts the common wisdom about iterated prisoner's dilemma. 00111 — the strategy I've called permanent retaliation — is either best or very near the top, and above the much lauded tit-for-tat. At least in this deterministic tournament there's not much benefit to forgiveness.

Prisoners Dilemma | Simulation | Cooperation | Game Theory | Experiment

Written by Stuart Ferguson

254 Followers · 26 Following

Following

3D graphics pioneer, entrepreneur, maker, champion of science and reason, and philosophical gadfly

Responses (6)

What are your thoughts?

Respond

Roberto A. Foglietta You
Just now

In my opinion the biggest difference between your results and Axelrod's results rely on "random" strategy players plus inheritance, possibly. However, in the first instance forget about inheritance which is a broader context. The "random" players, whom are equivalent to 50%-50% cooperation or defection without any memory which can be impersonated by the random unknown guy make who we met a single time among all the others unknown people, like it often happens in a big city but not in a small country village, make such a difference moving the outcome in favor of the tic-for-tac (single time retaliation, or retaliate immediately and forgive). Which is equivalent in human terms to not losing the trust in the random unknown guy in a big city. While the opposite would be: I do not know you, then I defect. Which is a characteristic that only "loser" strategies have. R-

Reply

Scott Anderson
Jan 21, 2024

Very interesting and thorough! One variation would be to consider a small touch of randomness. (In real life, someone might intend to cooperate, but make a mistake, or their actions were misinterpreted.) The never-forgive strategy then endlessly... more

3 1 reply Reply

Drew H
Aug 12, 2024

I understand that Axelrod's strategies were collected from his friends but they were -human- strategies. I'm not convinced that exhaustively enumerating possible strategies produces a result that's applicable in the sociological/philosophical way... more

1 1 reply Reply

Bob Snyder
Apr 29, 2024

Does your simulation have randomness(mistakes)? And have you tried your strategy against Pavlov?

Prisoner’s Dilemma’s Dilemma

Some math applies to real life, some of the time

Stuart Ferguson · Follow
6 min read · Jan 15, 2023

Young people always complain that they are never going to need the abstract math they are being taught in school. Of course there are many professions that require quite advanced math, not just science and engineering but business, finance and public policy. Many crafts are surprisingly math-adjacent, finding their foundations in linear algebra and geometry. Despite that, the sentiment isn’t entirely false. After all, what can math tell us about getting along with one another? That has to be learned in the school of hard knocks.

Actually math does a have a field that purports to be relevant to human behavior — game theory. Game theory is the study of various kinds of formal games, and how to unlock reliable solutions. Games are often used to model more complex relationships in the hope of finding principles that can be generalized. The mother of all of those game theory problems is *prisoner’s dilemma*.

The setup is simple. Two people have been arrested for a crime. They are separated and are each independently offered the chance to defect — to implicate the other person in the crime. If both cooperate and refuse to defect, they both go free; if one defects that person goes free while the other goes to prison; if both defect then they both go to prison.

Ideally the prisoners cooperate with each other, no one gets implicated and no one goes to prison. The problem is that neither prisoner knows what the other one is doing. If you add up all the possible outcomes the best decision is to defect. Since both of the prisoners use the same logic to arrive at the same decision, both go to prison. Even though acting together they could both succeed, the fact that they are acting independently means they both fail.

This is the crux of the dilemma.

Now if you paid close attention you may have noticed that, no, the setup I described implies no such thing. If we just want to avoid jail, then either cooperating or defecting are equally possible to succeed depending on what the other person does. There's no obvious reason why defecting leads to the better outcome. In fact game theorists have to stack the game to capture the behavior they want: for example, they might score going to prison alone as 0, going to prison together as 2, going free together as 3, and going free alone as 5. Thus cooperation gets you 0 or 3, while defection gets you 2 or 5, giving defection the better payoff if the other person acts randomly.

It's hard to quite see how the narrative can support those numbers. How is going to prison together that much better than going to prison alone?

Likewise in messy reality there are social factors to consider. Many communities look down on anyone who helps the authorities. From gangs and organized crime, to marginalized people, to the lawless wasteland of the schoolyard, everyone knows that snitches get stitches. In that circumstance the social pressure against defection could easily tip the balance to successful cooperation. People also have a reputation they are concerned about, which could be tainted by a defection here. Even a simple commitment to justice and fairness should leave a person outraged enough by this blatant abuse of police power to override the math of self-interest.

The inequalities in the payoff matrix really define prisoner's dilemma, not the strained metaphor. Purely as a game where two players secretly pick their move and then earn a payout based on the table described above, it does manage to model an interesting situation. It's a bit like a collective action problem, where parties need to work together for all to succeed, but there are substantial rewards — and no penalties — for one party to back out of the arrangement. Thus everyone betrays each other, like the tragic ending of a heist movie.

This result is formalized as the *Nash equilibrium*, which is any outcome in which none of the parties can improve their position by making a unilateral change, and therefore stable. “Both defect” is the Nash equilibrium for prisoner’s dilemma. In all other outcomes a cooperating player can always improve their position by defecting.

To game theorists this is considered the rational solution.

It’s easy to generalize. We could conclude from this well-studied example that any collective action is impossible. Perhaps this illustrates that we cannot rely on individual action to maintain common goods. Any compact founded on trust will be bound to fail because individuals will find exploits that benefit themselves at the expense of those who followed the compact. Is prisoner’s dilemma a mathematical proof of the *tragedy of the commons*? There are social sanctions that disapprove of non-cooperation in ones social duties, but what do those really mean anymore? How strong can that be against the promise of personal enrichment?

Fortunately there’s more to the story than just the Nash equilibrium. Human interactions are rarely one-time events, and normal social behavior may be better modeled by what’s called iterated prisoner’s dilemma. Instead of a single game, the same two players play a series of games where they can adjust their strategy as they go. Purely logical analysis would conclude that “always defect” is still the strongest strategy, because it’s the logical single-game strategy so it should dominate in iterated games as well. Indeed a strong strategy includes defecting when necessary, but there’s a complication.

In iterated prisoner’s dilemma what counts is the cumulative score. Two players who always defect come away from the contest each with middling scores, because the “both defect” outcome has marginal returns. In order to do better a player could play a predatory game trying to trick the other player into cooperation when they intend to defect. This isn’t sustainable. The other way to get higher scores is for two players to cooperate. Not just once or twice but many times.

The only way to consistently beat the default score is to learn to recognize opponents who will reliably cooperate.

It may seem counterintuitive — after all we've been informed by game theory that defection is the rational choice — but this result is borne out by experiments. There's no closed-form solution, but strategies that did better than the baseline in iterative prisoner's dilemma tended to have these features:

- **Nice** — start with cooperation and don't be the first to defect
- **Retaliation** — meet defection with defection
- **Forgiving** — after retaliation allow a return to cooperation
- **Non-envious** — don't try to score more than the other player

Cooperation doesn't mean being a pushover. Strategies that find a way to cooperate do better in long run because there are objective rewards for cooperating. Even given the miserly payoff schedule of prisoner's dilemma, collective action is worthwhile. The cost is that sometimes the other player does better than you.

It makes you wonder, is the Nash equilibrium actually even the rational solution to this kind of game? Turns out the answer is no, but that's a story for another time.

Written by Stuart Ferguson

254 Followers · 26 Following

3D graphics pioneer, entrepreneur, maker, champion of science and reason, and philosophical gadfly