Prisoner’s Dilemma

A Game Theory Problem

Introduction

Two criminals, A and B, are confined without communication. With insufficient evidence for the main charge, the police propose a deal: testify against your partner to go free while they serve three years. However, if both remain silent, they face a lesser charge and a year in jail each. The twist comes if both testify against each other, resulting in a two-year sentence for both.

1. If A and B both remain silent, they will each serve one year in prison.
2. If A testifies against B but B remains silent, A will be set free while B serves three years in prison.
3. If A remains silent but B testifies against A, A will serve three years in prison and B will be set free.
4. If A and B testify against each other, they will each serve two years.

Let's explore the outcome when our prisoners make decisions based on complete rationality (picks what works out best for them).

Prisoner B / Prisoner A	Prisoner B stays silent (*cooperates*)	Prisoner B testifies (*defects*)
Prisoner A stays silent (*cooperates*)	Each serves 1 year	Prisoner A: 3 years Prisoner B: goes free
Prisoner A testifies (*defects*)	Prisoner A: goes free Prisoner B: 3 years	Each serves 2 years

This leads us to a scenario known in game theory as a 'Nash Equilibrium,' where both players choose to defect, and any deviation from this strategy would result in a suboptimal outcome for either player.

In General

We redefine our scenario involving two players, Blue and Red, as an optimization problem where the goal is to maximize the points earned. To qualify as a Prisoner's Dilemma, certain conditions must be met: mutual cooperation being preferable to mutual defection (R > P), and defection always being the dominant strategy (T > R and P > S).

Where T > R > P > S

Canonical prisoner's dilemma payoff matrix

Blue \ Red	Cooperate	Defect
Cooperate	*R*, *R*	*S*, *T*
Defect	*T*, *S*	*P*, *P*

Examples

Cigarette Advertising:

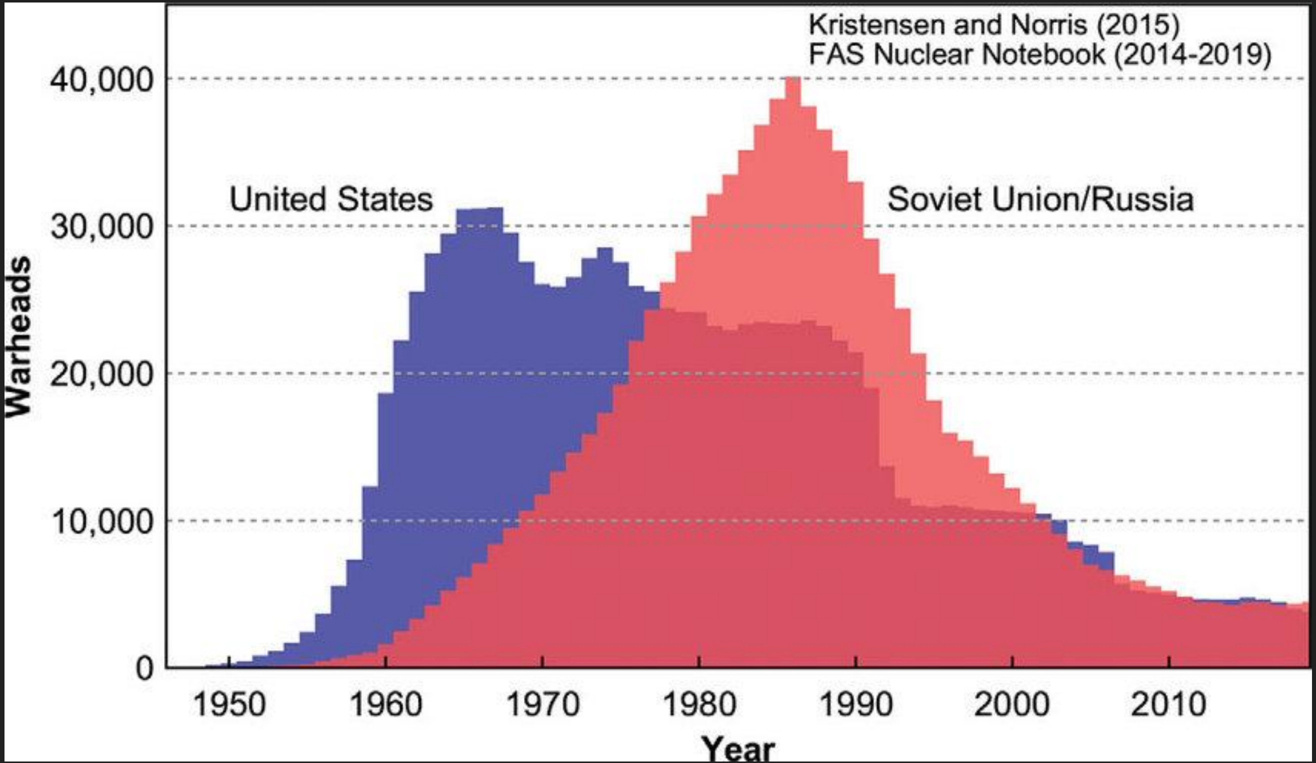
Firm B / Firm A	Firm B does not advertise (*cooperates*)	Firm B advertisers (*defects*)
Firm A does not advertise (*cooperates*)	Each profit $100	Firm A profits $40 Firm B makes $160, with a $30 cost of advertising for $130 profit
Firm A advertises (*defects*)	Firm A makes $160, with a $30 cost of advertising for $130 profit Firm B profits $40	Each makes $100, with a $30 cost of advertising for $70 profit

Cold War:

Soviet Union / United States	Soviet Union does not develop more nuclear weapons (*cooperates*)	Soviet Union develops more nuclear weapons (*defects*)
United States does not develop more nuclear weapons (*cooperates*)	Both have no risk of nuclear destruction	Soviet Union dominates the US in a hypothetical war
United States develops more nuclear weapons (*defects*)	United States dominates the Soviet Union in a hypothetical war	Both are at risk of nuclear destruction

Is there hope?

Initially, our Cold War leaders, acting as rational agents, determined that defection was the optimal move. However, it's crucial to highlight a shift in their perspective. This change is attributed to the nature of the scenario—it's not a single-instance Prisoner's Dilemma but rather a repeated iteration of the dilemma.



Iterated Prisoner’s Dilemma

The game is straightforward: two players engage in the repeated Prisoner's Dilemma, taking into account past actions. At first look we can think of many ways this influencers our game such as a concept like “trust”, where consistent cooperation from the opponent can influence one's strategy to maintain a cooperative approach, so that you don’t change your opponents mind. Additionally, we impose the condition 2R > T + S, ensuring that mutual cooperation yields a better outcome than alternating between cooperation and defection.

Canonical prisoner's dilemma payoff matrix

Blue \ Red	Cooperate	Defect
Cooperate	*R* *R*	*T* *S*
Defect	*S* *T*	*P* *P*

Nash Equilibrium for Iterated Prisoner's Dilemma

If we can identify a Nash equilibrium for the iterated Prisoner's Dilemma, involving cooperation, it would solve so many problems! Consider a scenario where both players are aware that they are engaged in N repeated Prisoner's Dilemma games. In the Nth game, rational players understand that their choice won't have repercussions, reducing it to the classic Prisoner's Dilemma, leading both to defect. Since both rational players anticipate this, the N-1th game essentially becomes the last, perpetuating a cycle where the rational strategy still remains constant: repeated defection. 😭 (Draw out explanation)

Axelrod’s Contest of 200-step Iterative Prisoner Dilemma

Axelrod invited academic colleagues from around the world to devise computer strategies to compete in an iterated prisoner's dilemma tournament. The programs that were entered varied widely in algorithmic complexity, initial hostility, capacity for forgiveness, and so forth.

He played them all against each other, and ranked them by who had the most total points

Some example contestants:

- Friedman, cooperates repeatedly until opponent defects once then repeatedly defects
- Joss, cooperates then copies the opponent’s last move (90% of the time) and 10% of the time pulls a sneaky defect
- Grasskamp, cooperates then copies the opponent’s last move except always defecting in the 50th round
- Tit for Tat, cooperates then copies the opponent last move.

Axelrod's payoff matrix

Red / Blue	Cooperate	Defect
Cooperate	3 / 3	5 / 0
Defect	0 / 5	1 / 1

TABLE 2
Tournament Scores

Player \ Other Players	TIT FOR TAT	TIDEMAN AND CHIERUZZI	NYDEGGER	GROFMAN	SHUBIK	STEIN AND RAPOPORT	FRIEDMAN	DAVIS	GRAASKAMP	DOWNING	FELD	JOSS	TULLOCK	(Name Withheld)	RANDOM	*Average Score*
1. TIT FOR TAT (Anatol Rapoport)	600	595	600	600	600	595	600	600	597	597	280	225	279	359	441	504
2. TIDEMAN AND CHIERUZZI	600	596	600	601	600	596	600	600	310	601	271	213	291	455	573	500
3. NYDEGGER	600	595	600	600	600	595	600	600	433	158	354	374	347	368	464	486
4. GROFMAN	600	595	600	600	600	594	600	600	376	309	280	236	305	426	507	482
5. SHUBIK	600	595	600	600	600	595	600	600	348	271	274	272	265	448	543	481
6. STEIN AND RAPOPORT	600	596	600	602	600	596	600	600	319	200	252	249	280	480	592	478
7. FRIEDMAN	600	595	600	600	600	595	600	600	307	207	235	213	263	489	598	473
8. DAVIS	600	595	600	600	600	595	600	600	307	194	238	247	253	450	598	472
9. GRAASKAMP	597	305	462	375	348	314	302	302	588	625	268	238	274	466	548	401
10. DOWNING	597	591	398	289	261	215	202	239	555	202	436	540	243	487	604	391
11. FELD	285	272	426	286	297	255	235	239	274	704	246	236	272	420	467	328
12. JOSS	230	214	409	237	286	254	213	252	244	634	236	224	273	390	469	304
13. TULLOCK	284	287	415	293	318	271	243	229	278	193	271	260	273	416	478	301
14. (Name Withheld)	362	231	397	273	230	149	133	173	187	133	317	366	345	413	526	282
15. RANDOM	442	142	407	313	219	141	108	137	189	102	360	416	419	300	450	276

Results!

Out of all complex strategies devised to compete in Axelrod’s Tournament. Tit for Tat won!

Analyzing the top contenders some patterns emerged, among all top contenders:

- Nice (never defects first, as opposed to Nasty)
- Able to retaliate
- Forgiving (can retaliate, but does not let defections from before the previous round influence it’s decisions)
- Non-envious (does not care about the opponents score, note that Tit for Tat can never win in a one on one, only draw or lose)

Among the 15 strategies, 8 were characterized as nice, while 7 were deemed nasty. Surprisingly, the top-performing 8 were all classified as nice!

Friedman's strategy, although nice, proved to be unforgiving, placing him towards the lower end of the top 8 nice strategies, securing the 7th position. A

Axelrod observed that submitting 'Tit for Two Tats,' even more forgiving than 'Tit For Tat,' would have been the winning move.

Axelrod’s Second Tournament

Addressing the previously mentioned issue regarding the Nash Equilibrium of defecting in a N-step game, Axelrod implemented a solution by setting the average number of rounds in the iterative games to 200, but having the precise N per iterative matchup unknown preventing players from knowing the final round. This adjustment holds more relevance in real-world scenarios.

Numerous participants opted for nice and forgiving strategies, including popular choices like 'Tit for Tat' and 'Tit for Two Tats.' On the other hand, some submitted strategies aiming to exploit forgiving approaches in the tournament. For instance, 'Tester' begins with defection but switches to 'Tit For Tat' as a form of apology if the opponent retaliates. However, if the opponent does not retaliate, 'Tester' continues to defect every other move after that.

Results:

Among the meticulously crafted strategies from 63 contestants, the familiar champion emerges – Tit for Tat, despite all contestants knowing Tit for Tat was gonna be playing. Remarkably, among the top 15, only one strategy was not classified as nice, while among the bottom 15, only one was not considered nasty.

From this competition, Axelrod identified two additional traits shared by top contenders:

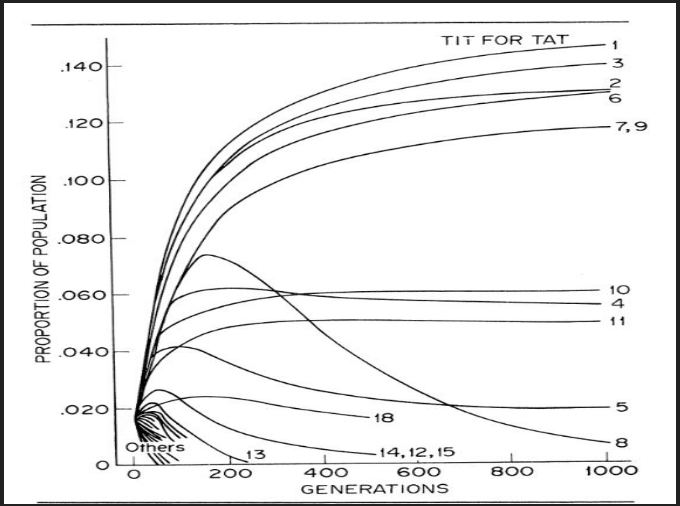
- Retaliatory (if your opponent defects, strike back immediately)
- Clear: Ensure that opponents can comprehend your strategy. Without this clarity, the iterations lose significance, resembling independent single Prisoner's Dilemma games, where defecting against you is rational.

These traits combined with the other two encourage some type of “eye for an eye” philosophy

Some notes and Ecological Tournament

We observe that although Tit for Two Tats could hypothetically be considered a winner in the first tournament, it ended up securing the 23rd position in the second one. This underscores a crucial realization: there is no universally optimal strategy. The effectiveness of a strategy is contingent upon the choices made by opponents. AKA the environment matters.

We note that a Tit for Tat in an environment full of always defect strategies, always gets last place.



To ensure that Tit for Tat's success was not a mere fluke, Axelrod conducted multiple iterations of the tournament, introducing multiple instances of each strategy per generation. The best-performing strategies increased in number, while the underperforming ones decreased and eventually became extinct.

Axelrod noted an interesting pattern: most nasty strategies quickly fade away, but one persistent nasty bot earlier in the top 15 manages to outlast the others. Its longevity, however, is contingent on preying on some of the worst-performing nice bots. Once these bots go extinct, the persistent nasty bot also declines. After thousands of generations, the proportions stabilize, and Tit for Tat emerges as the leader, constituting 14.5% of the population.

There is a nice ecological analogy here that this mirrors how animals interact, and the concept of trophic cascade.

The evolutionary analogy is not one-to-one since we do not have mutations here.

Geographic comments

Imagine a scenario where, instead of interacting with everyone, you only engage with those in your vicinity. If a majority of players consistently defect, but there's a cluster of Tit for Tat players who cooperate with each other, the simulation results in a substantial increase in their numbers in the next generation. Surprisingly, a small island of cooperation can flourish and eventually dominate the entire population.

This is intriguing as it highlights that, even when acting rationally, cooperative behaviors can naturally emerge within a system.

Some posit that this phenomenon elucidates the cooperative behaviors observed in nature, such as animals engaging in grooming activities. Evolutionary biology supports the idea that these beneficial cooperative behaviors can be encoded in the DNA of organisms, eventually dominating and influencing entire populations.

Returning to our Cold War example, the gradual process of starting with cooperative measures and mutual checks, akin to the repeated Prisoner's Dilemma games, proved more effective in slowly decreasing nuclear arms. This approach contrasts with a single Prisoner's Dilemma option, which would likely lead to a rational defection.

Noise and Signal Errors

Another aspect under scrutiny is the concept of noise, where occasional cooperation may be inaccurately perceived as defection. A notable real-world instance is the 1983 incident when the Soviets mistakenly detected the launch of a U.S. ICBM (defection). Fortunately, the Soviet officer on duty dismissed the alarm, underscoring the potential costs of noise and signal errors.

In a noisy environment, when Tit for Tat plays against another Tit for Tat, the initial cooperation continues until a false defection triggers a chain of alternating retaliation. This pattern persists until another false defection occurs, leading to an average point outcome that is approximately one-third of what they would achieve in a perfect environment.

To address this issue, introducing more forgiveness to break out of retaliatory cycles proves effective. Tit for Tat with approximately 10% generosity, where it refrains from retaliating, emerges as the most effective strategy in such scenarios.

Zero-Sum game? No

A curious observation emerges: Tit for Tat consistently fails to win in a one-on-one iterated Prisoner's Dilemma game, while always defecting guarantees avoidance of defeat. However, despite this, Tit for Tat is a superior strategy.

The rationale behind this lies in the nature of the iterated Prisoner's Dilemma, which differs from zero-sum games like Poker or Chess, where one player's gain is directly offset by the other player's loss.

Our best strategies are non-envious, they do not strive to score more than their opponent in a given game.

Conclusions

- Be Nice (it pays off)
- Don't be a pushover (retaliate)
- However, sometimes be generous to stop perpetual cycles
- Be forgiving (don't hold grudges)
- And be clear about your actions (don't overcomplicate your responses)

AKA Tit for Tat with some generosity

And somehow these lessons beneficial to the group can come out of a completely selfish analysis of the world.

References

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. Journal of conflict resolution, 24(1), 3-25. - https://ve42.co/Axelrod1980a

Axelrod, R. (1980). More effective choice in the prisoner's dilemma. Journal of conflict resolution, 24(3), 379-403. - https://ve42.co/Axelrod1980b