Project AXLRD: Understanding Cooperation in Games

Jonathan L. Graves

Last updated: April 30, 2021

Learning Objectives

After completing this lesson, you will be able to:

1. Describe how we can use repeated games to describe repeated interactions between individuals in strategic situations
2. Explain how a Prisoner's Dilemma describes the tension between cooperation and self-interest
3. Analyse this model and use theory to predict individual behaviour in finitely repeated Prisoner's Dilemma games
4. Understand how this prediction relates to the beliefs individuals have about others playing the game and what we have found in experiments
5. Design your own strategies to take part in similar experiments

Background and Motivation

Cooperation versus Selfishness?

In strategic situations, we often imagine that individuals are fundamentally self-interested:

- They're willing to cost other people large amounts of welfare for relatively small gains
- In biological terms, they display selfish behaviour:
  - Animals are "red in tooth and claw" who "[w]ith ravine, shriek'd against [love]" (Tennyson, 1850)
- In human terms, they are antisocial, looking out for themselves over others

However, this description is often untrue. Cooperation and altruism are typically the rule, not the exception.

Amartya Sen (1977) - *Rational Fools*

"Central to this problem is the assumption that when asked a question, the individual gives that answer which will maximize his personal gain. How good is this assumption? I doubt that in general it is very good.

("Where is the railway station?" he asks me. "There," I say, pointing at the post office, "and would you please post this letter for me on the way?" "Yes," he says, determined to open the envelope and check whether it contains something valuable.)"

Discussion Activity

Can you think of some reasons why people or animals cooperate even though they might be self-interested?

Discussion Activity

Can you think of some reasons why people or animals cooperate even though they might be self-interested?

Some reasons might be:

- Kin selection in animals
- Reciprocity
- Altruism or other-regarding preferences
- Reputation
- Moral arguments (e.g. Kantian imperative)

Prisoner's Dilemma

A famous model of the tension between cooperation and self-interest is the Prisoner's Dilemma (PD, for short).

- In PD games, two players make simultaneous decisions
- They can either work together (cooperate, C) or be selfish (defect, D).
- If both players work together (C, C) they both do better than if they are both selfish (D, D)
- However, if one player is selfish and the other is not, the selfish player does better than if they worked together, while the other player does worse.

Therefore, players have an incentive to mutually work together, but also to be individually selfish.

PD Matrix Representation

In a matrix, this game looks like:

Player 1		Player 2: C	Player 2: D
	C	3,3	0,5
	D	5,0	1,1

Figure: A Prisoner's Dilemma

Equilibrium in the PD

- Both the Nash equilibrium and dominance equilibrium of the PD are for both Players to choose defect (D)
- The “prediction” of the PD is that there should be no cooperation
- In the real world, however, cooperation is very common.

Equilibrium in the PD

- Both the Nash equilibrium and dominance equilibrium of the PD are for both Players to choose defect (D)
- The "prediction" of the PD is that there should be no cooperation
- In the real world, however, cooperation is very common.

Question

Think back to our discussion of reasons to cooperate; what is the Prisoner's Dilemma failing to capture?

Repeated Prisoner's Dilemma

Key Insight: Repetition

In the real world, one-off-interactions are exceptionally rare.

- We normally interact with the same individuals over and over again
- This creates incentives for long-term thinking:
  - If we help them today, they may help us tomorrow
  - If we are mean to them today, they may be mean to us tomorrow
- Reciprocity becomes the key consideration

To describe this, we use a repeated version of the PD

- Two player game
- The game takes place over T rounds
- Each round consists of a single PD
  - Actions are still C or D
  - After round is over, both players observe actions and outcomes
- Total payoff in game is sum of the payoffs earned in each round

Example: Two Rounds ($T = 2$)

- Suppose in Round 1, both Players choose C
- They would earn 3 each, and game moves to round 2
  - Both Players see the choices in Round 1 before making Round 2 decisions
- Suppose they both choose D.
- Game ends; both players have earned $3 + 1$ in total.

Notice: Players Can React

The repeated nature of the RPD allows players to observe and react

- They can reward cooperative behaviour by cooperating
- They can punish selfish behaviour by defecting
- They can take advantage of cooperative behaviour by cooperating
- They can “test” their opponent’s behaviour
- Etc.

Many (complex) possibilities for behaviour.

Solving the RPD: Backwards Induction

In order to understand how people behave in the RPD, let's imagine working backwards from the end of the game $\rightarrow$ this is called backwards induction.

- Notice that in Round T, the game is now just a regular (one-off) Prisoner's Dilemma
- Therefore, you should always defect in the last period
  - Nothing you do can come back and hurt you, so you maximize your payoff $\rightarrow$ choose D
- Therefore, any strategy should choose D in the last round no matter what.

Backward Induction: Round $T - 1$

Now, let's imagine how we should behave in the second-to-last round.

- **Remember**: in the next round, everyone will choose D no matter what you do in this round.
- In other words, your choices today won't change the behaviour in the future.

Backward Induction: Round $T - 1$

Now, let's imagine how we should behave in the second-to-last round.

- **Remember**: in the next round, everyone will choose D no matter what you do in this round.
- In other words, your choices today won't change the behaviour in the future.

- So, this is just (again) like a one-off PD
- If nothing you do this round will affect the future, you should maximize your payoff right now $\rightarrow$ choose D
- Therefore, no matter what you should always choose D in round $T - 1$

Equilibrium in the RPD

However, you can see that this logic applies to any round.

- Since nothing you do can affect the choices in later rounds, you should always maximize your current payoff $\rightarrow$ choose D.

Theorem (Result)

In any finitely repeated Prisoner's Dilemma, the unique equilibrium is for both Players to always defect.

- This kind of equilibrium is called a subgame perfect Nash equilibrium (or SPNE)

Cooperation in the RPD?

We can see that repetition alone does not suggest that cooperation will occur $\rightarrow$ the equilibrium is for everyone to always choose "defect"

- However, this requires you to believe that your opponent will always act selfishly in every future period
- If you mutually believe this, we get the result $\rightarrow$ however, this belief is very strong!

What happens if we do not believe we will play an opponent who always defects? Why might we want to cooperate?

Beliefs in RPD

Beliefs about your Opponent

This highlights the very important role beliefs play in informing decision making in games $\rightarrow$ especially games like the RPD.

- When you play a game, you will have expectations about how your opponent will play
- If you play many opponents, or opponents at random, these beliefs might involve quite complicated strategies you think they could follow
- Although we have seen the equilibrium strategy is "always defect" this **does not mean** that it's always the best strategy.

In particular, you can find strategies that do better against some opponents $\rightarrow$ this is a key difference between the RPD and the regular PD.

Example: `GRIM-TRIGGER`

Suppose you believe your opponent is following the strategy `GRIM-TRIGGER`, which is described by the following rules:

1. If it is the first round, player C
2. Otherwise, play C unless in *any* past round a player has played D.
3. Otherwise, D.

This is a type of trigger strategy because it has an event which “set it off”.

Questions for Discussion

1. What kind of real-world behaviour is `GRIM-TRIGGER` like? How could your describe it?
2. How long does `GRIM-TRIGGER`'s “memory” need to be?
3. Is `GRIM-TRIGGER` a “cooperative” or “selfish” strategy?

Best Responses to GRIM-TRIGGER

Consider the game of length T and let's find the best response to GRIM-TRIGGER:

- If you never choose D you will get get a payoff of $3 \cdot T$ in total.
- If you choose D at round t things get a little more complicated.
  - For the $t-1$ rounds before you chose D, you get $(t-1) \cdot 3$
  - For the $T-t$ rounds after you chose D, you get $(T-t) \cdot 1$
  - At round t you get a payoff of 5

This is easier to visualize in a picture.

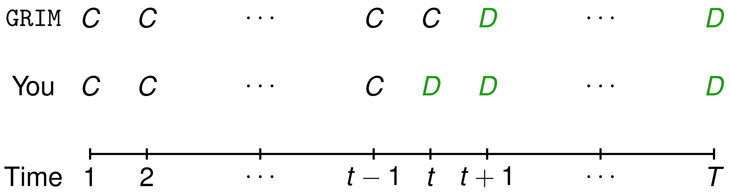

Figure: Actions in RPD versus GRIM-TRIGGER

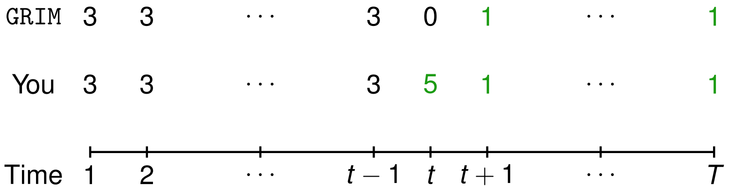

Figure: Payoffs in RPD versus GRIM-TRIGGER

What if you chose D at $t+1$?

At $t+1$	3	3	$\cdots$	3	3	5	1	$\cdots$	1
At t	3	3	$\cdots$	3	5	1	1	$\cdots$	1
Time	1	2	$\cdots$	$t-1$	t	$t+1$	$t+2$	$\cdots$	T

Figure: Comparison of Defection Timing

What if you chose D at $t+1$?

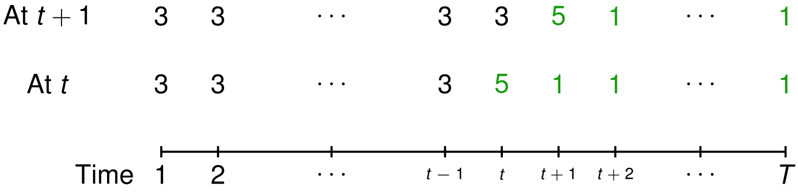

Figure: Comparison of Defection Timing

Discussion Question

Which option is better? Can you see why?

Solution?

As we can see, by moving the period t where you choose to “set-off” trigger, you trade one period of the cooperative payoff (3) for one period of the non-cooperative payoff (1).

- This is a net positive, because mutual cooperation is better than mutual selfishness
- You will only *not* want to do this when you can’t “trade” any more periods $\rightarrow$ when $t = T$

In other words, you should wait until `GRIM-TRIGGER` cannot retaliate against you any-more, then take advantage of them!

Best Responses to GRIM-TRIGGER

What this has shown is that your best response if you believe your opponent is playing GRIM-TRIGGER is to play C until the very last period.

- This is a **very cooperative** situation $\rightarrow$ you will cooperate $\frac{T-1}{T}$% of the time!
  - For instance, if $T = 100$, that's a 99% cooperation rate.
- Compare this to the best response to "always defect", which has a 0% cooperation rate.

This shows how important beliefs are to your best responses $\rightarrow$ and how they can create strong incentives to cooperate.

Equilibrium in the RPD?

However, you might have noticed one thing: is your belief correct?

- If your opponent was truly playing `GRIM-TRIGGER`, you could expect they might anticipate you will play *D* at the last period.
- In that case, `GRIM-TRIGGER` isn't the best strategy possible.
- They should change their strategy to improve their outcomes.

However, what they believe about you, and how far their reasoning has gone may not be clear! That's what makes this game so interesting.

The Axelrod Experiment

Robert Axelrod's Idea

In the 1970s the political scientist Robert Axelrod was fascinated by these tensions: between cooperation and selfishness, between beliefs and actions.

- His idea was to model competition and cooperation in society using the combination of the RPD and the new "computer" technology.
- He reached out to his friends in the emerging community surrounding computers and asked them to write a program to play the RPD against one another.
- The competitors would be all of the strategies submitted.

His intuition was that repeated interaction might lead to the emergence of stable, cooperative behaviour - after all, if we interact repeatedly, we might learn how to be "friends".

Axelrod's Tournament: A Surprising Result!

The idea caught on! It was clever, interesting, and kind of fun to think about.

- Dozens of people took part: economists, political scientists, computer programmers, mathematicians, military experts, and philosophers.
- They wrote all kinds of strategies - complicated, simple, elegant, and exhaustive. Axelrod collected them all, and ran the tournament.
- When he looked at the result, something very surprising happened.

Axelrod's Tournament: A Surprising Result!

The idea caught on! It was clever, interesting, and kind of fun to think about.

- Dozens of people took part: economists, political scientists, computer programmers, mathematicians, military experts, and philosophers.
- They wrote all kinds of strategies - complicated, simple, elegant, and exhaustive. Axelrod collected them all, and ran the tournament.
- When he looked at the result, something very surprising happened.

What kind of strategy do you think won?

Do you think the strategy was complicated or simple? Was it selfish or cooperative? Did it use complex programming or intuitive rules? Did it need a long memory or short?

The Winner: `TIT-FOR-TAT`

The winner (submitted by psychologist Anatol Rapoport) was very simple (only 4 lines of code) and was called tit for tat (or "copycat"). It works as follows:

- In the first period, play C
- In any other period, play what your opponent played last period.

This strategy has consistently been among one of the best performances, and nearly perfect captures the tension between being nice and being mean.

Improvements?

It has been very difficult to find strategies that outperform `TIT-FOR-TAT`:

- Strategies that try to “improve” have to make other sacrifices
- For instance, “tit for two tats” solves the problem where mistakes can trap tit-for-tat into a cycle of retaliation but it also is easier to exploit.

One way to defeat it is to use a *set* of strategies which “sacrifice” themselves.[1]

[1] see Reading 1

Discussion Activity

Why do you think `TIT-FOR-TAT` did so well? What kinds of strategies might it perform well against? What kinds of strategies might it perform poorly about?

Project AXLRD

Project AXLRD

Your task is to investigate and design your own strategy to take part in a tournament just like the one Axelrod created.

- While you do this, you will work through a worksheet which will help you explore how to build strategies, and what you meed to think about when designing a strategy.
- You will also be able to test your strategy against other opponents, and improve your design.
- You will be competiting against all the strategies your classmates (in this class!) will also be submitting

As we saw earlier, it is critical to anticipate what your classmates will design, and plan accordingly.

Rules

- Payoffs are as in the standard Prisoner's dilemma
- Each strategy will play every other strategy once, in a 100 round RPD
- Your score is the total across all of your matches
- You can only submit a single strategy, so make it a good one.

There are more details about the game on your worksheet.

Have Fun and Good Luck!

References and Readings

Suggested Readings

1. Wired Staff. (2017, June 05). New tack wins prisoner's dilemma. Retrieved February 1, 2021, from `https://www.wired.com/2004/10/new-tack-wins-prisoners-dilemma/`
2. Dawkins, R. (2016). The selfish gene. Oxford university press.
3. Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. science, 211(4489), 1390-1396.
4. Axelrod, R. (1997). The complexity of cooperation: Agent-based models of competition and collaboration (Vol. 3). Princeton university press.

References

1. Sen, Amartya K. "Rational fools: A critique of the behavioral foundations of economic theory." *Philosophy & Public Affairs* (1977): 317-344.
2. Tennyson, Alfred Tennyson Baron. *In Memoriam AHH*. Bankside Press, 1900.

Image Credits

- Slide 5: Karen Arnold (CC0): **Meerkat looking up in close-up** (n.d.)
- Slide 6: Peter Flint (CC0): **The Old Man** (n.d.)