Cooperation in Repeated Games

Jaden Zhu

July 13, 2025

Why study game theory?

In the cold war, both countries faced a dilemma:

- US and Soviet Union had enough nuclear weapons to end civilization
- Both had incentive to strike first for advantage
- But, they gradually reduced amounts of nukes and avoided war

Our game theory answer:

- Real-world interactions aren't all about immediate gain
- Future consequences of retaliation create cooperation where unexpected
- Small actions can build trust over time (nuclear disarmament)

Other applications:

- International relations
- Business and trade agreements
- Social cooperation
- Wildlife survival

So, when does it pay to be nice?

Prisoner's Dilemma

The Scenario:

- Two prisoners suspected of crime
- They are then questioned separately
- Police offer a deal to both for them to confess
- The prisoners cannot communicate and must decide independently

Outcomes of the Scenario:

- Both Cooperate (remain silent): light sentence
- One confesses and the other cooperates: confessor released, silent one gets a harsh sentence
- Both confess: moderate sentences for both

Prisoner's Dilemma Cont.

Classic Payoff Matrix:

- (C,C): (3,3)
- (C,D): (0,5)
- (D,C): (5,0)
- (D,D): (1,1)

Analysis:

- Defection is the dominant strategy for both players
- Nash equilibrium is mutual defection
- Paradox: both would benefit more from cooperating

Repeated Games vs. One Shot

One Shot Games:

- Players play for immediate gain
- No memory of past actions
- Games are played once.

Repeated Games:

- Cooperation is possible
- Players interact more than once
- Use of punishment strategies

Types of Repeated Games

Finitely Repeated Games:

- Fixed number of rounds
- Use backward induction (work from last stage to first)
- For games with one equilibrium, result in final stage is the same for every stage.

Infinitely Repeated Games:

- No endpoint
- Cooperation is possible
- Players weigh long-term consequences with immediate gain

Discount Factor:

- "Shadow of the future"
- Discounts future payoffs, represents player patience, range is $0 < \delta < 1$

Types of Strategies

Basic Strategies

- Always Defect: Defect every round (ruthless, no trust)
- Always Cooperate: Cooperate every round (trusting but vulnerable)

Trigger Strategies

- Grim Trigger: Cooperation, but permanent defection if anyone defects once.
- Naive Grim Trigger: Only defect permanently in response to opponent's defection (not stable)

Types of Strategies Pt. 2

Retaliation Strategies

- Limited Retaliation: Punish defection for k periods then resume cooperation
- Tit-For-Tat: Mirror opponent's previous move.
- Generous TFT: Similar to Tit-For-Tat but random chance to cooperate after opponent defects
- Tit-for-2-Tats: Only retaliates after two defections in a row

Adaptive Strategies

- Win-Stay, Lose-Shift: Start with cooperation, repeat same action if last outcome was "good" (C,C) or (D,D)
- Otherwise, switch action

Conditions for Cooperation

One Shot Deviation Principle:

- A way to **simplify** looking for equilibria
- Strategy is optimized if the payoff can't be improved by changing one move

Sustained Cooperation:

- The discount factor plays a huge role in sustaining cooperation
- High δ, players have more patience. Low δ, players prioritize immediate gain
- Higher $\delta =$ more likely cooperation

Analysis of Strategies

Grim Trigger:

- $\delta \geq \frac{1}{2}$
- Players need **moderate** patience.

Limited Retaliation:

- $\delta^{K+1} \geq 2\delta + 1$
- A value for K is necessary for a concrete value
- When K $= 2$: $\delta \geq 0.63$
- As K approaches ∞, δ nears $\frac{1}{2}$.

Analysis of Strategies Pt. 2

Tit For Tat:

- $\delta = \frac{1}{4}$
- Not a subgame perfect equilibrium.
- There is no incentive to resume cooperation after defection
- Extremely fragile discount factor

Win-Stay, Lose-Shift:

- $\delta \geq 0.25$.

In summary, the only strategy that is **not** a subgame-perfect equilibria is Tit-For-Tat.

The Folk Theorem

Definition

Feasible Payoffs are the set of payoffs achievable through some combination of strategies

Definition

Minmax Payoff is the minimum payoff a player can guarantee themselves no matter what the opponent plays

Definition

Strictly-enforceable is a payoff vector where every round the player gets more than their minmax payoff

The Folk Theorem

Any feasible and enforceable payoff can be sustained as an equilibrium given a sufficiently high δ value.

The Folk Theorem continued

What this means is that virtually any outcome is possible in infinitely repeated games.

Conclusions:

- Repetition opens a lot of new possibilities
- Higher δ means more sustainable equilibria

Limitations:

- Perfect monitoring: in the real-world, players can't perfectly observe actions
- Imperfect information: Observation errors and accidents can lower the probability of cooperation arising
- Human limitations: People aren't perfectly rational and cannot keep track of infinite payoffs

Challenges to Cooperation

There are a couple of key differences that need to be addressed to apply game theory to real-world situations.
Issues:

- Noise: Accidental deviations could trigger infinite punishment
- Imperfect monitoring: Players can't confirm what their opponent chose
- Solution: Forgiving strategies (GTFT, TF2T) work better

Like as mentioned before, humans have many intellecual limitations like:

- Humans can't calculate infinite payoffs
- Humans have limited memory
- Human biases: Overconfidence and revenge which affect decisions

Real World Applications

Now, let's see how these skills and theorems learned from game theory shape up to the real-world.

- Game Theory rose to prominence during the Cold War because the US government turned to game theorists to solve the US-Soviet strategic dilemma
- Nuclear disarmament is like a repeated game, where both countries gradually started cooperation to remove nuclear weapons and verified that the other was following aswell.
- Cuban Missile Crisis: 'shadow of the future' that helped drive cooperation

QnA and Conclusion

Thank you all for listening!