

Game Theory 08-01

Finite Repeated Games and History-Based Coordination

BCSE Game Theory

Nov. 25, 2025

Learning Goals

- ▶ Review stage games such as the Teamwork example and understand the tension between individual incentives and cooperation.
- ▶ Define finite repeated games, histories, and how strategy profiles depend on past play.
- ▶ Apply backward induction to finite horizons and explain the unraveling argument.
- ▶ Study how stage games with multiple equilibria allow threats to sustain cooperative paths.
- ▶ Prepare for the infinite repetition setting and the dynamic programming perspective.

Setting the Stage

Stage games we repeat

Why repeat (stage) games?

- ▶ Most strategic interactions continue beyond one shot, so players care about how today's move shapes tomorrow. (e.g., Teamworks (limited periods), competition in a market)
- ▶ Module 07 taught us SPNE in extensive-form games; repeating a stage game reuses that machinery with longer histories.
- ▶ Repetition opens new enforcement tools: reputations, threats, and history-dependent punishments that are unavailable in single-shot play.

Formal Stage Game

Definition: Stage Game

A stage game is a strategic-form game $G = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$ where:

- ▶ Players $i \in N$ choose actions $s_i \in S_i$ simultaneously.
- ▶ Payoffs $u_i(s)$ are realized immediately after the action profile $s = (s_1, \dots, s_n)$ is selected.

- ▶ We will repeat this base game T times.
- ▶ A Nash equilibrium of G is a profile s^* where $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$ for all $s_i \in S_i$.

Stage Game Example: Teamwork

	C	D
C	(3, 3)	(1, 4)
D	(4, 1)	(2, 2)

- ▶ There is a unique Nash equilibrium: (D,D).
- ▶ (C,C) is Pareto superior, but (D,D) is the only stable outcome in a one-shot game.

Stage Game Example with Multiple Nash Equilibria

	Stag	Hare
Stag	(4, 4)	(0, 3)
Hare	(3, 0)	(3, 3)

- ▶ Both (Stag, Stag) and (Hare, Hare) are Nash equilibria.
- ▶ (Stag, Stag) is payoff dominant but requires mutual commitment.
- ▶ The Stage game alone offers no mechanism to pick the Pareto-optimal equilibrium.

Equilibrium selection cues

- ▶ Payoff dominance versus risk dominance guides players when multiple equilibria exist.
- ▶ Without repetition, players rely on focal points, norms, or communication to coordinate.
- ▶ Repetition gives us an enforcement mechanism that can privilege the desired equilibrium.

Finite Repeated Games

Definitions

Finite repeated game definition

- ▶ Let G be a stage game. The finite repeated game G^T runs T periods.
- ▶ In each period t , players simultaneously select actions s^t and observe the outcome.
- ▶ The entire sequence of actions (s^1, \dots, s^T) determines the payoffs.

History-dependent strategies

- ▶ The history before period t is $h^t = (s^1, \dots, s^{t-1})$.
- ▶ Strategies are functions $\sigma_i^t : H^t \rightarrow S_i$ mapping the realized history into an action.
- ▶ Players possess perfect recall, so they remember all past profiles.

History notation example

- ▶ If period 1 outcome is (C,C) and period 2 is (C,D) , then $h^3 = ((C,C), (C,D))$.
- ▶ Strategies can condition on such histories to switch from cooperation to punishment.
- ▶ Specifying off-path behavior is necessary to define subgame perfection.

Strategy as history mapping

- ▶ A strategy specifies what to do for every possible past, including off-path histories.
- ▶ History-dependent strategies enable punishments: a deviation can change future play.
- ▶ The set of feasible continuations is richer than the stage game alone.

Payoff evaluation in finite repetition

- ▶ Each player's total payoff is the sum of stage utilities:

$$U_i = \sum_{t=1}^T u_i(s^t).$$

- ▶ Equivalently, we can consider the average payoff $(1/T) \sum_t u_i(s^t)$ when comparing different strategies.
- ▶ Non-discounted sums are standard for finite horizons, but we keep discounting in mind for the limit case.

Teamwork stage payoffs

	C	D
C	(3, 3)	(1, 4)
D	(4, 1)	(2, 2)

- ▶ The unique Nash equilibrium is (D,D) because defecting is a dominant strategy.
- ▶ Cooperation cannot be sustained in the single-shot game.

Finite repeated Teamwork: short horizon

- ▶ Players play the stage game T times and observe the outcome before moving to the next period.
- ▶ The only Nash equilibrium in the stage game is (D,D).
- ▶ To find the equilibrium of the repeated game, we use backward induction.
- ▶ We start from the final period and reason backwards to the first period.

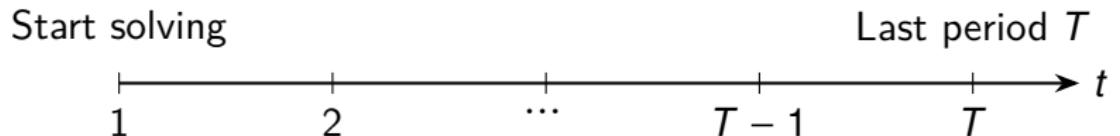
Solving Finite Repeated Games

Backward induction & Unraveling

Backward induction recipe

1. Solve the last period T : players play a stage Nash equilibrium.
2. Given the equilibrium payoffs of period T , solve period $T - 1$ with the continuation payoffs fixed.
3. Repeat until reaching period 1. Each period is solved as a stage game with continuation payoffs added.
4. The resulting profile is a subgame perfect equilibrium for the repeated game.

Timeline of backward induction



- ▶ The analysis begins at the right-hand side (period T) and moves backward.
- ▶ The final period's equilibrium determines the continuation payoffs used earlier.

Unraveling definition

- ▶ Unraveling occurs when the unique equilibrium in the last period forces that equilibrium earlier.
- ▶ If every continuation game has a unique subgame perfect equilibrium, cooperation collapses backward to $t = 1$.
- ▶ The finite horizon prevents players from using out-of-equilibrium threats that rely on future cooperation beyond T .

Unraveling in the Teamwork Game

- ▶ Period T : (D,D) is the only Nash equilibrium, so both defect.
- ▶ Period $T - 1$: knowing that T is doomed, immediate defection is optimal (no future threat remains).
- ▶ By induction, every period contains defection; cooperation cannot be sustained.

Theorem: Unique Equilibrium in Finite Repeated Games

Theorem: Unique Equilibrium Payoff

Let G be a stage game with a unique Nash equilibrium s^* yielding payoff vector $u(s^*)$. Then every subgame perfect equilibrium of the finite repeated game G^T yields the average payoff vector $u(s^*)$.

- ▶ **Proof sketch:** By backward induction.
- ▶ Period T : Players must play s^* (unique Nash equilibrium), yielding $u(s^*)$.
- ▶ Period $T - 1$: Given that period T yields $u(s^*)$, the continuation payoff is fixed. Players again play s^* .
- ▶ By induction, every period $t \in \{1, \dots, T\}$ plays s^* , so the average payoff is $u(s^*)$.

General unraveling condition

- ▶ Whenever the continuation game has a unique subgame perfect equilibrium, the same path repeats in earlier periods.
- ▶ Unique last-period equilibrium removes any leverage for threats in period $T - 1$, and so on.
- ▶ We need multiple equilibrium or viable punishments to break the chain of unraveling.

Teamwork repeated: building threats

- ▶ We can use the inferior equilibrium (D,D) as a punishment for deviating from (C,C).
- ▶ The punishment path must be credible: after a deviation, players switch to a Nash equilibrium of the remaining game.
- ▶ The history of play determines whether cooperation continues or the punishment path (D,D) is triggered.

Threat structure for Teamwork

- ▶ Start by playing (C,C) for $t = 1, \dots, T - 1$.
- ▶ If any player deviates, revert to (D,D) for the remaining periods.
- ▶ (D,D) is a stage Nash equilibrium, so the threat is subgame perfect.

Punishment path as credible equilibrium

- ▶ The punishment path itself must be an equilibrium of the continuation game.
- ▶ (D,D) remains an equilibrium after a deviation, so it can serve as a credible punishment.
- ▶ Without such a stage equilibrium, the punishment loses credibility.

Incentive inequality for cooperation

- ▶ The short-term gain from deviating must not exceed the loss from the punishment phase.
- ▶ For Teamwork: deviating to D yields 3 instead of 4, but the punishment replaces future (4, 4) payoffs with (2, 2).
- ▶ Choosing the horizon and punishment length appropriately keeps the inequality in favor of cooperation.

Verifying Subgame Perfection

- ▶ SPNE requires the strategy to be optimal in **every** subgame, not just on the equilibrium path.
- ▶ **On-path (Cooperation):** Checked by the incentive inequality (future loss > current gain).
- ▶ **Off-path (Punishment):** If a deviation occurs, players switch to (D,D).
 - ▶ Is playing D optimal given the opponent plays D?
 - ▶ Yes, because (D,D) is a stage Nash equilibrium.
 - ▶ Best response to D is D ($2 > 1$).
 - ▶ Thus, players have no incentive to deviate from the punishment.
- ▶ This confirms the threat is credible and the strategy is an SPNE.

Example play path with punishment

1. Periods 1 to t^* : both players choose Work, believing cooperation continues.
2. If a deviation occurs at t^* , both switch to D for $t^* + 1, \dots, T$.
3. The deviator's short-term gain is offset by the guaranteed lower payoff in the punishment phase.

Multi-equilibrium stage games give better targets

- ▶ Stage games like the Stag-Hare offer both payoff-dominant and risk-dominant equilibria.
- ▶ Repetition enables us to target the Pareto-superior equilibrium by threatening to play the safety equilibrium in case of deviation.
- ▶ The ability to switch between equilibria is what expands the repeated equilibrium set.

Credible punishment requirement

- ▶ A punishment is credible if the punishment path is a Nash equilibrium of the continuation game.
- ▶ Using (D,D) after a deviation satisfies credibility because it is a stage Nash equilibrium.
- ▶ Players must prefer sticking to the cooperative path rather than deviating and facing punishment.

Finite horizon strategy families

- ▶ One can design two-phase strategies: cooperate for the first m periods and play a stage equilibrium afterwards.
- ▶ The punishment can be permanent or limited to a fixed number of periods.
- ▶ The longer the horizon, the more elaborate the punishment structures we can sustain.

Tradeoff: horizon length and enforcement

- ▶ As T increases, future payoffs gain weight and deviations become less attractive.
- ▶ However, the final period still unravels unless there are alternative punishment equilibria.
- ▶ Finite repetition cannot match the richness of the infinite horizon, but it does enlarge the equilibrium set compared to the single-shot game.

Comparison: repeated NE vs stage NE

- ▶ Repeated NE can condition on histories, while stage NE cannot.
- ▶ The availability of credible punishment equilibria defines what repeated NE are feasible.
- ▶ Finite repetition allows partial cooperation even when stage cooperation is not an equilibrium.

Playing Non-Nash Action Profiles

- ▶ **Key insight:** Repetition allows players to sustain action profiles that are **not** stage Nash equilibria.
- ▶ Example (Stag Hunt): (Hare,Stag) is not a Nash equilibrium of the stage game.
 - ▶ Player 2 would prefer to deviate to Hare (payoff 4 instead of 0).
- ▶ **Can we enforce (Hare,Stag) through repetition?** Consider $T = 2$:
 - ▶ Period 1: Play (Hare,Stag). If Player 2 deviates, play (Hare,Hare) in period 2.
 - ▶ Period 2: Play (Stag,Stag) if no deviation occurred.
 - ▶ Player 2's payoff from cooperating: $(0 + 3)/2 = 1.5$.
 - ▶ Player 2's payoff from deviating in period 1: $(4 + 2)/2 = 3.0$.
 - ▶ This doesn't work! Player 2 prefers to deviate.
- ▶ **General principle:** Future punishment can make non-equilibrium actions rational, but requires:
 - ▶ Sufficient horizon length (larger T).
 - ▶ Credible and severe enough punishment.

Axelrod's Tournament

The Evolution of Cooperation

The Tournament Setup

- ▶ Robert Axelrod (1980) invited game theorists to submit programs for a repeated Prisoner's Dilemma tournament.
- ▶ Strategies played against each other in a round-robin format.
- ▶ Each match consisted of exactly 200 rounds—a finite repeated game.
- ▶ Payoffs were standard PD values (similar to our Teamwork game).

The Winner: Tit-for-Tat

- ▶ The winner was the simplest program submitted by Anatol Rapoport.
- ▶ **Tit-for-Tat (TFT):**
 1. Cooperate on the first move.
 2. Thereafter, copy the opponent's previous move.
- ▶ It beat complex strategies that tried to be "sneaky" or "exploitative."

Why Tit-for-Tat Won? (The 4 Properties)

Axelrod identified four properties of successful strategies:

1. **Nice**: Never defect first. (Prevents unnecessary conflict).
2. **Retaliatory**: Punish defection immediately. (Discourages exploitation).
3. **Forgiving**: Return to cooperation if the opponent stops defecting. (Avoids endless vendettas).
4. **Clear**: Be easy for the opponent to understand. (Encourages cooperation).

Evolutionary Stability

- ▶ In evolutionary simulations, strategies that score poorly "die out," while successful ones replicate.
- ▶ "Nasty" strategies might win initially by exploiting naive cooperators, but they eventually run out of victims and fight each other.
- ▶ Cooperative strategies like TFT thrive when they can interact with other cooperators.
- ▶ Cooperation can emerge and persist even in a selfish world.

Exercises preview

- ▶ Compute the defection path of the Teamwork Game when $T = 4$ and identify all subgame perfect equilibria.
- ▶ Design a threat strategy for the Teamwork stage game that sustains (C,C) for $T = 5$ and compare payoffs.
- ▶ Analyze a stage game with multiple Nash equilibria and describe which equilibrium serves as punishment.

Common pitfalls

- ▶ Using punishments that are not themselves Nash equilibria, which destroys credibility.
- ▶ Ignoring off-path histories when specifying strategies.
- ▶ Assuming cooperation survives without ensuring the threat is sufficiently costly.

Takeaways & looking ahead

- ▶ Finite repetition enlarges the equilibrium set by leveraging history-dependent strategies.
- ▶ Stage games with multiple Nash equilibria provide ready-made targets for punishment paths.
- ▶ Next time we move to infinite repetition, introduce dynamic programming, and study Folk theorems.