Finitely Repeated Games

Game Theory · Module 5 — Repeated Games

Node 5.1

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Repeated Games **Node:** 5.1

A finitely repeated game is the simplest way to lift a one-shot strategic interaction into a temporally extended one: fix a stage game G, pick a horizon T, and play G in each of T periods while letting each player's total payoff be a function (usually the sum or average) of the stage payoffs. The move seems gentle — we are merely iterating a game — but the consequences are sharp. When G has a unique Nash equilibrium and T is common knowledge, backward induction (Module 3) collapses the entire repeated game to a unique subgame-perfect equilibrium in which each player defects in every period. The shadow of the future buys you nothing, because the last period is always a one-shot game and the rot propagates backward. This node establishes the finitely repeated setting, derives the backward-induction collapse, and sets up the contrast with infinite repetition (5.2) where the collapse does not occur. Module 5 is the structural home of the "shadow of the future" — the idea that the prospect of continued interaction disciplines present behavior — and 5.1 is where we learn under what conditions the shadow actually falls.

Formal Definition

Let $G = (N, \{S_i\}, \{u_i\})$ be a finite stage game in normal form (the strategic form triple of 1.1). A **finitely repeated game** with horizon T and stage game G, written G^T, is the extensive-form game obtained by playing G in each period $t \in \{1, 2, \ldots, T\}$ with **perfect monitoring**: at the end of period t, every player observes the realized stage-profile $s^t = (s_1^t, \ldots, s_n^t)$.

Histories. A period-t **history** is a sequence $h^t = (s^1, s^2, \ldots, s^{t-1}) \in S^{t-1}$, the profile of play realized in all prior periods. The set of period-t histories is $H^t = S^{t-1}$, with $H^1 = \{\emptyset\}$.

Strategy. A (pure) strategy for player i is a collection of maps

$$\sigma_i = (\sigma_i^1, \sigma_i^2, \ldots, \sigma_i^T), \qquad \sigma_i^t : H^t \to S_i,$$

assigning a stage action to every period-t history. Mixed strategies replace S_i with $\Delta(S_i)$. A strategy is a **complete contingent plan**: at every history I could reach, what would I do?

Total payoff. The total payoff to player i from a play path $(s^1, \ldots, s^T)$ is typically

$$U_i(\sigma) = \sum_{t=1}^{T} u_i(s^t) \qquad \text{(undiscounted)} \qquad \text{or} \qquad U_i(\sigma) = \sum_{t=1}^{T} \delta^{t-1} u_i(s^t) \qquad \text{(discounted)},$$

where $\delta \in (0, 1]$ is the discount factor and the right-hand sum is the **present-value** total. The average payoff is $\frac{1}{T} U_i$ (undiscounted) or $\frac{1-\delta}{1-\delta^T} U_i$ (discounted).

Subgame-perfect equilibrium (SPE) of G^T. A strategy profile σ^* is an SPE if for every period t and every history h^t, the continuation play from period t onward — played from h^t — is a Nash equilibrium of the continuation game. Equivalently (by the one-shot deviation principle), σ^* is SPE iff no player can improve by a single-period deviation at any history.

Backward-induction theorem (finite horizon, unique stage NE). If the stage game G has a **unique** Nash equilibrium s^*, then G^T has a unique subgame-perfect equilibrium: the profile σ^* in which every player plays s^* in every period regardless of history.

This theorem is the structural heart of 5.1. Its proof is a direct backward induction: in period T, the continuation is just G and the unique NE must be played; in period $T-1$, the continuation payoff from any period-T outcome is the same $u_i(s^*)$, so period-$T-1$ play is again a one-shot game whose unique NE is s^*; iterating backward gives s^* in every period.

Intuition

The central lesson is that repetition alone does not buy cooperation; what buys cooperation is the **open-endedness** of the future. If the horizon is finite and commonly known, the last period is a one-shot game. Whatever threats players have issued about future punishment have no bite in the last period — there is no future. So last-period play is Nash of the stage game. But then the second-to-last period has no future either, because the last period is fixed and its outcome is independent of what happens now. And so on. The backward-induction argument is the formal expression of the intuition "everything I could threaten to do tomorrow has no effect today, because I will do the same thing tomorrow regardless."

Two things can rescue cooperation in finite repetition, and both of them matter enormously:

First, if the stage game has **multiple** Nash equilibria, then the continuation payoff in each period is not pinned down by one-shot analysis. Players can use the threat of switching from a "good" continuation NE to a "bad" continuation NE as a disciplining device in earlier periods. This is the **Benoit–Krishna construction** (1985) and it generates non-trivial SPE of G^T even for finite T. The payoffs of stage NE must differ across equilibria for this to work, and the horizon must be long enough for the disciplining to cover the one-shot gain from deviation.

Second, if players are **uncertain** about the horizon, or if there is even a small probability that the game continues, the finite game effectively becomes infinite. This is the modeling move of 5.2: a geometric continuation probability δ is formally identical to a discount factor on an infinite-horizon game. In practice, almost every real strategic situation has this kind of continuation probability; the pure finite-horizon model is the extreme where the end is commonly known and fixed.

A third, weaker rescue is **reputation** in the Kreps–Wilson–Milgrom–Roberts sense: if there is a small probability that one player is a "commitment type" who plays cooperatively unconditionally, then even in a finite horizon the rational player may mimic the commitment type for many periods to build reputation. This connects to 5.5 and the chain-store paradox and is where finite repetition becomes interesting again despite the backward-induction pressure.

For now, 5.1 establishes the baseline: under perfect information, known finite horizon, unique stage NE, the repeated game collapses to the stage NE. Everything in the rest of Module 5 is a departure from one or more of these assumptions.

Structural Role

5.1 is the negative result of Module 5, and like all good negative results it tells you what the positive theorems require. The backward-induction collapse shows that finite horizon + unique stage NE + common knowledge is the triple that kills cooperation. Each of the subsequent nodes relaxes one of the three:

- **5.2 Infinitely repeated games and discounting** drops the finite horizon and shows that the set of equilibrium payoffs expands dramatically.

- **5.3 Trigger strategies** introduces the canonical class of infinite-horizon equilibrium strategies, whose threats have bite only because the future has no end.
- **5.4 Folk theorem** characterizes the enormous set of sustainable payoffs in infinitely repeated games.
- **5.5 Chain store paradox** relaxes common knowledge (of rationality) and shows how small reputation types can support cooperation in finite horizons.
- **5.6 Cooperation in repeated Prisoner's Dilemma** is the canonical case study.

5.1 is also the natural link back to Module 3 (extensive form and subgame perfection) — the entire analysis uses SPE as the solution concept, and the backward-induction argument is just 3.5 applied to G^T.

Mathematical Development

Backward induction on G^T: proof of the collapse

Claim. If G has a unique Nash equilibrium s^*, every SPE of G^T prescribes s^* in every period along every history.

Proof. Induct on the number of remaining periods. Let σ^* be any SPE of G^T.

Base case: consider period T. The continuation from any history h^T is a copy of G (there are no further periods), so SPE requires the continuation strategy to be a Nash equilibrium of G. By uniqueness, $\sigma_i^*(h^T) = s_i^*$ for every i and every h^T.

Induction: suppose $\sigma_i^*(h^\tau) = s_i^*$ for every $\tau > t$ and every h^τ. Consider period t. The continuation from period $t+1$ onward is independent of the action chosen in period t — it is always s^* forever. So player i's total continuation payoff from period t is

$$u_i(s_i, s_{-i}^t) + \sum_{\tau=t+1}^{T} \delta^{\tau-t} u_i(s^*),$$

and the second sum does not depend on s_i. Player i therefore maximizes only the first term, which is player i's one-shot payoff in the stage game. The maximizers are i's best responses to s_{-i}^t, and in a Nash equilibrium of the period-t continuation this gives $s^t = s^*$. By uniqueness of s^*, this is the only period-t equilibrium outcome, and SPE requires it at every history. Complete the induction.

The proof uses two things: **(i)** the continuation payoff from period $t+1$ onward is independent of period-t actions (because of the inductive fix), and **(ii)** this makes period t a one-shot game. Both fail in infinite horizon: the continuation payoff can depend on past play through non-trivial strategies like trigger strategies, and there is no last period from which to start inducting.

Benoit–Krishna: multiple stage equilibria rescue cooperation

When the stage game has $k \geq 2$ Nash equilibria with distinct payoffs — call them $s^{(1)}, s^{(2)}, \ldots$ with $u(s^{(1)}) \neq u(s^{(2)})$ — the backward-induction argument no longer forces a unique continuation payoff. We can support a non-equilibrium profile $\bar{s}$ in early periods with threats of switching from a good continuation equilibrium to a bad one.

Construction sketch. Let $\bar{s}$ be the target cooperative profile. Let s^g be the good stage equilibrium (high payoffs) and s^b the bad one. Consider the strategy: "play $\bar{s}$ in periods $1, \ldots, T-K$. In periods $T-K+1, \ldots, T$, play s^g. If anyone ever deviates from this prescription, play s^b in every remaining period."

For this to be SPE, the one-period gain from deviating at any period $t \leq T-K$ must be no larger than the cost of switching the last K periods from s^g to s^b:

$$\max_{s_i} u_i(s_i, \bar{s}_{-i}) - u_i(\bar{s}) \leq K \cdot \big[u_i(s^g) - u_i(s^b)\big].$$

If K is large enough (and T is at least $T - K + K = T$, so we just need $T \geq K + 1$), this can be satisfied. And the "punishment phase" itself is credible because s^b is a stage Nash equilibrium — playing it is already a best response, so no player has incentive to deviate from the punishment.

Benoit–Krishna theorem (1985). If the stage game has at least two Nash equilibria that are **strict** with distinct payoffs for each player, then as $T \to \infty$, the set of SPE payoffs of G^T converges to the set of feasible individually rational payoffs of G — the same set characterized by the folk theorem for infinite horizons.

This theorem is the first hint that the backward-induction collapse is fragile: even in finite horizons, multiplicity of stage equilibria is enough to recover the folk theorem asymptotically.

One-shot deviation principle

The one-shot deviation principle says: to check that σ^* is SPE, it suffices to verify that no player can improve by deviating at a single history and conforming to σ^* everywhere else. This reduces the SPE check from "consider all possible alternative strategies" (infinitely many) to "consider all possible one-period deviations at all histories" (finitely many for finite games). It is the workhorse of repeated-game analysis and will be used throughout Module 5.

Statement. Let σ^* be a strategy profile in G^T. Then σ^* is SPE iff for every player i, every history h^t, and every one-period deviation $s_i' \neq \sigma_i^*(h^t)$, the deviation does not strictly improve i's continuation payoff given that all other players conform to σ^* and that i conforms to σ^* from period $t+1$ onward.

The proof is straightforward: if there were a profitable multi-period deviation, there would be a last period in which i deviates, and the one-period deviation at that last-deviation history would be profitable in isolation. The one-shot deviation principle fails for infinite-horizon games without continuity at infinity, which is why 5.2 introduces discounting.

Undiscounted vs discounted finite horizons

In a finite horizon, the choice between undiscounted total payoff and discounted present value is mostly cosmetic — both produce the same SPE when the stage game has a unique NE (backward induction is insensitive to the discount factor in finite horizons). The discount factor matters more when stage equilibria are multiple (it adjusts the Benoit–Krishna calculus) and massively in 5.2 when the horizon becomes infinite.

Worked Example

Finitely repeated Prisoner's Dilemma

Stage game, payoffs (u_1, u_2):

	C	D
C	$(3,3)$	$(0,5)$
D	$(5,0)$	$(1,1)$

The unique Nash equilibrium of the stage game is (D, D) with payoff $(1, 1)$. The cooperative profile (C, C) is not a stage NE but dominates (D, D) in payoffs.

Horizon $T = 1$: The game is just G. Unique NE: (D, D). Total payoff $(1, 1)$.

Horizon $T = 2$: Back-induct. Period 2: play the unique stage NE (D, D) regardless of history, since this is the last period. Period 2 payoff: $(1, 1)$. Period 1: knowing that period 2 will be (D, D) regardless, period 1 is effectively a one-shot game. Unique NE: (D, D). Total: $(2, 2)$.

Horizon $T = 100$: Same backward induction. Period 100: (D, D). Period 99: continuation is fixed at (D, D) regardless of period-99 actions, so period 99 is a one-shot game, (D, D). And so on backward. Total: $(100, 100)$.

What tit-for-tat does in this setting. Suppose player 1 commits to the tit-for-tat strategy: "cooperate in period 1; in every subsequent period, play whatever player 2 played in the previous period." Is the profile (tit-for-tat, tit-for-tat) an SPE? No. In the last period T, tit-for-tat's prescription depends on period $T-1$ play, but the optimal response is always D because this is the last period and there is no future punishment. The profile is not SPE because tit-for-tat is not a best response in period T. Backward induction unravels tit-for-tat.

What the Benoit–Krishna construction cannot do here. The stage game has a single Nash equilibrium, so there is no "bad" continuation equilibrium with which to threaten players. The construction fails; the backward-induction collapse holds.

Finitely repeated game with multiple stage equilibria

Stage game:

	L	M	R
L	$(4,4)$	$(0,5)$	$(0,0)$
M	$(5,0)$	$(3,3)$	$(0,0)$
R	$(0,0)$	$(0,0)$	$(1,1)$

This stage game has two pure Nash equilibria: (M,M) with payoff $(3,3)$ and (R,R) with payoff $(1,1)$. (Check: at (M,M), unilateral deviations give $L \to (0,3)$, $R \to (0,3)$ for row, so no profitable deviation; similarly for column. At (R,R), deviations to L or M give 0 to the deviator, so it is strict NE.) Neither is Pareto optimal: $(L,L) = (4,4)$ dominates both but is not a stage NE (row can deviate to M for payoff 5).

Horizon $T = 2$, sustaining (L,L) in period 1. Consider: "play L in period 1; in period 2, play M if period 1 was (L,L), else play R."

Check period 1: the one-shot gain from deviating to M in period 1 is $5-4=1$. The cost is the period-2 continuation switch from $(M,M)=(3,3)$ to $(R,R)=(1,1)$, which is $3-1=2$ per player. Since cost (2) exceeds gain (1), no profitable one-shot deviation. ✓

Check period 2: both (M,M) and (R,R) are stage Nash equilibria, so the continuation strategy is a best response whether the history triggers the good or the bad one.

Subgame perfect equilibrium in which period 1 plays the non-equilibrium profile (L,L). Horizon $T=2$ suffices because the one-shot gain is small and the equilibrium-payoff gap is large. If the gain-to-gap ratio were larger, we would need a longer horizon — the Benoit–Krishna bound.

This construction is the simplest example of how multiplicity of stage equilibria rescues cooperation in finite horizons. It also exhibits the trade-off the folk theorem will formalize in 5.4: the more room you have in the continuation payoff space, the richer the set of sustainable outcomes.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Stage game G	A single hand of poker in isolation
Horizon T	Number of hands in the session/match
History h^t	All prior hands: who played what, who won how much
Strategy σ_i	Your complete plan for the session, conditional on every possible history
Stage NE	The GTO strategy for a single hand

Formal object	Poker instantiation
Backward-induction collapse	In a known-length heads-up match, nothing you do today changes how your opponent plays the last hand
Benoit–Krishna construction	Using "I will switch to my exploitable weak strategy if you don't conform" threats in multi-hand contexts
Discount factor δ	Probability the session continues (fatigue, being kicked, running out of time)

Concrete situation: the heads-up match of known length

You are playing a heads-up no-limit hold'em match against a weaker opponent, and both of you know there will be exactly 100 hands. Your opponent is exploitable: they call too much on the river. You could try to make a "non-aggression pact" — fewer bluffs from both sides, smaller pots, less variance. Can this arise in equilibrium?

Answer: No, at least not via backward induction on the stage game. The last hand is a one-shot game. Whatever you do in hands 1–99 has no effect on hand 100; both players play their one-shot best response in hand 100. Working backward, hand 99's continuation is fixed (hand 100 will be GTO regardless), so hand 99 is also a one-shot game, and so on. In every hand you should play the stage-game best response to your opponent's strategy. If they call too much, you value-bet more and bluff less — the pure exploitation strategy in every hand.

What this does not mean: It does not mean that playing GTO in every hand is optimal against this opponent. Against an exploitable opponent, the stage-game best response *is* an exploitation strategy, and backward induction says "play the exploitation strategy in every hand." The theorem says nothing about GTO; it says that the optimal strategy in every hand equals the optimal single-hand strategy.

Where the assumption breaks: If the opponent might quit when losing, the match is no longer of known length. The continuation probability introduces an effective infinite horizon, and 5.2's analysis kicks in. Reputation considerations (opponent learning about you and adapting) also break the backward-induction argument — that is 5.5.

Concrete situation: short tournament

A three-handed sit-and-go where the top two get paid. At the bubble (three players), the game has multiple equilibria: "fold everything marginal" and "jam wide." The existence of multiple equilibria means the backward-induction collapse does not hold in the simple form; players can make implicit threats about continuation play that discipline current behavior. This is the Benoit–Krishna setting. The theoretical implication is that cooperative non-aggression bubble play is supportable as SPE with the right continuation, even in a finite tournament. Practical solvers (ICMIZER, HoldemResources) do not work this out formally — they take ICM payoffs as given and compute the single-hand Nash — but the structural reason bubble play "feels cooperative" is that the multiplicity of continuation equilibria allows coordination.

What this understanding buys you

Finite-length poker matches collapse to one-shot play. If a match has a commonly known length and the stage game has a unique NE, nothing you do in early hands disciplines opponent behavior in late hands, and therefore nothing disciplines your own early-hand play. Every hand is independently a best-response problem.

The interesting cases are the violations of the assumption. Reputation, unknown horizons, multiple stage equilibria, exploration, and learning all break the backward-induction argument. Real poker sits in this region. The finite-horizon baseline tells you what to expect in the maximally simple case, and everything richer is a departure from that baseline.

Pro players' "session" framing is meaningful only with a continuation probability. When pros talk about "playing the session" or "building up a table image over the course of a night," they are implicitly assuming $\delta < 1$ — that the session continues with some probability into indeterminate future. If the session were commonly known to be exactly 500 hands and no more, the shadow of the future would not actually discipline anyone on a single-hand basis; the value of image-building would be zero in hand 500 and propagate backward. In practice, sessions are unbounded, which lets Module 5.2's infinite-horizon results apply.

Cross-Links

- **3.5 Subgame perfect equilibrium** — the solution concept used throughout Module 5.
- **2.1 Nash equilibrium** — the stage NE that backward induction locks in.
- **5.2 Infinitely repeated games and discounting** — the next step, where the finite-horizon collapse fails to happen.
- **5.3 Trigger strategies and punishment** — the canonical strategies of infinite-horizon repeated games.
- **5.4 Folk theorem** — the characterization of sustainable payoffs.
- **5.5 Reputation and chain-store paradox** — how reputation restores cooperation in finite horizons.
- **9.1 Fictitious play** — learning dynamics that also run over repeated play, with different structural assumptions.

Structural Compression

Invariant: In a finitely repeated game with commonly known finite horizon and a stage game with a unique Nash equilibrium, backward induction forces the unique stage NE to be played in every period of every subgame-perfect equilibrium; the shadow of the future does not discipline present behavior.

Mechanism: The last period is a one-shot game (there is no future), so its SPE prescribes the unique stage NE; the continuation from any period is therefore fixed and independent of current actions; each earlier period becomes effectively a one-shot game; induction backward to period 1 yields the collapse.

Cross-System Echo: The finite-horizon collapse is a textbook example of how **terminal boundary conditions propagate through a dynamic system**. Compare to the Hamilton–Jacobi–Bellman equation in optimal control: the value function at the terminal time is given, and the value in earlier periods is determined by backward recursion. In both cases, the entire temporal trajectory is determined by the endpoint plus the one-period optimization. The difference is that in optimal control the endpoint is exogenous and given, whereas in finite repeated games the "endpoint" is the one-shot Nash, which is itself a fixed point. Removing the endpoint — in control, by taking infinite horizon; in games, by taking infinite repetition — opens up the space of non-trivial time-consistent policies (discounted infinite-horizon HJB, trigger-strategy SPE). The parallel is deep and will return in 5.2 and again in Module 9 on learning dynamics.

Exercises

1. Prove the backward-induction collapse for the stage game G with unique NE s^* when the horizon is $T = 3$. Write out each period's argument explicitly.

2. Consider the stage game with two pure Nash equilibria s^g with payoffs (a, a) and s^b with payoffs (b, b), $a > b$. Let $\bar{s}$ be a non-equilibrium profile with payoffs (c, c) where $c > a$, and let d be the one-shot gain from deviating at $\bar{s}$, so the deviator's payoff is $c + d$. Show that for $T \geq 1 + d/(a - b)$, there is an SPE of G^T in which $\bar{s}$ is played in period 1 and the continuation is s^g if period 1 is $\bar{s}$ and s^b otherwise.

3. Show that the one-shot deviation principle fails in an infinite-horizon repeated game with $\delta = 1$ (no discounting) by constructing a strategy profile that is immune to one-period deviations but can be beaten by a multi-period deviation.
4. In the finitely repeated Prisoner's Dilemma with $T = 10$, compute the total payoff of each player under: (a) both players always defect; (b) both players play tit-for-tat (cooperating in period 1, then copying the opponent's previous action); (c) one plays tit-for-tat and the other always defects. Which of these is SPE?
5. In the finitely repeated coordination game G with two Pareto-ranked pure NE (U, U) with payoff $(2, 2)$ and (D, D) with payoff $(1, 1)$ (and with a single round of "coordination failure" payoff $(0, 0)$ when players mismatch), show that every SPE of G^T has the property that play in each period is a stage NE, but that not all SPE coincide — the selection among (U, U) and (D, D) in different periods can vary.
6. Construct a finitely repeated game where the Benoit–Krishna approximate folk theorem cannot be tight because the payoff from the cooperative profile is so much higher than any stage NE payoff that T would need to be impractically large. Interpret.
7. **Argue, structurally,** why the backward-induction collapse of 5.1 is the negative result that motivates the entire rest of Module 5. What specific assumptions of 5.1 does each subsequent node relax, and why does relaxing each one restore the possibility of cooperation?

Game Theory · Module 5 — Repeated Games · Node 5.1
Concepts: nash-equilibrium, repeated-interaction
Prerequisites: 1.1, 2.1, 3.5
Enables: 5.2, 5.3, 5.4, 5.5, 5.6

hbar.university — own.your.cognition