

Game Theory 08-02

Infinite Repeated Games, Folk Theorems, and Monitoring

BCSE Game Theory

Nov. 25, 2025

Learning Goals

- ▶ Introduce the infinite repeated game model and the discounted payoff evaluation.
- ▶ Define continuation payoffs, dynamic programming, and the Bellman equation for enforcement.
- ▶ Present the Folk theorem for perfect monitoring and its implications for equilibrium payoffs.
- ▶ Examine imperfect public monitoring and why standard triggers can reduce average payoffs as players become more patient.

Infinite Repetition

Strategy and payoffs

Why infinite repetition?

- ▶ Finite repetitions still unravel because the last period often provides no future punishment.
- ▶ An infinite horizon eliminates a final date, allowing more room for threats and cooperation.
- ▶ We now need to track discounted sums and recursive values to characterise equilibria.

Infinite repeated game definition

- ▶ The stage game G is played in every period $t = 1, 2, \dots$ with no predetermined end.
- ▶ Each player chooses actions based on the entire history of past signals or profiles.
- ▶ The strategy set includes plans for every conceivable history, including off-path events.

Discounted payoff formula

- ▶ Players evaluate infinite streams using a discount factor $\delta \in (0, 1)$.
- ▶ Normalised payoff:

$$U_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(s^t).$$

- ▶ The prefactor $(1 - \delta)$ keeps payoffs bounded and comparable as $\delta \rightarrow 1$.
- ▶ **Interpretations of δ :**
 - ▶ Inverse of growth rate (Net Present Value discount rate)
 - ▶ Probability game continues:
 $\delta = \Pr(\text{game continues to next period})$
 - ▶ Pure time preference: higher δ means more patient

Histories and strategies

- ▶ The history at time t is $h^t = (s^1, \dots, s^{t-1})$. Infinite histories run forever, but strategies only depend on finite prefixes.
- ▶ Strategies map any finite history into an action: $\sigma_i(h^t) \in S_i$.
- ▶ Perfect recall ensures players can condition on all past actions or signals.

Strategy categories

- ▶ Stationary strategies depend only on the current public signal or last action profile.
- ▶ Trigger strategies switch state once a deviation/signal occurs and stay in punishment thereafter.
- ▶ Finite automaton strategies keep a compact history summary, which suffices for most repeated-game paths.

Trigger strategy concept

- ▶ A trigger strategy starts by cooperating but switches to punishment if it detects a deviation.
- ▶ The threat is typically a stage Nash equilibrium that delivers lower payoffs.
- ▶ The hope is that the future loss outweighs the temptation to deviate today.

Teamwork Game (Reminder)

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

- ▶ C = Cooperate, D = Defect
- ▶ Unique Nash equilibrium: (D,D) with payoff (2, 2)
- ▶ (C,C) is Pareto superior with payoff (3, 3)

Grim trigger definition

- ▶ Play a cooperative action (e.g., (C,C)) as long as everyone cooperated before.
- ▶ Once a deviation is observed, permanently switch to the worst stage Nash equilibrium (e.g., (D,D)).
- ▶ Grim trigger relies on infinite punishment, so the weight on the future is crucial.

Comparing grim and tit-for-tat

- ▶ Grim trigger punishes forever, while tit-for-tat retaliates only once.
- ▶ Grim is harsher and requires higher δ to maintain cooperation, but it deters any deviation.
- ▶ Tit-for-tat is forgiving, making it more robust when noise or monitoring errors exist.

Incentive constraint for Grim trigger

- ▶ Let R be reward, T temptation, and P punishment payoffs in the Teamwork Game.
- ▶ **Teamwork Game values:** $R = 3$ (cooperate), $T = 4$ (defect when opponent cooperates), $P = 2$ (mutual defection).
- ▶ The inequality for cooperation is:

$$R \geq (1 - \delta)T + \delta P.$$

- ▶ Substituting: $3 \geq (1 - \delta) \cdot 4 + \delta \cdot 2 = 4 - 2\delta$.
- ▶ Rearranged: $2\delta \geq 1$, so cooperation is sustainable if $\delta \geq \frac{1}{2}$.
- ▶ General formula: $\delta \geq \frac{T-R}{T-P} = \frac{4-3}{4-2} = \frac{1}{2}$.

Verifying Optimality Off-Path

- ▶ Grim trigger specifies playing D forever after any deviation. Is this optimal?
- ▶ Suppose we are in the punishment phase (history includes a deviation).
- ▶ Opponent plays D forever.
- ▶ **If I play D:** Payoff is $2 + 2\delta + 2\delta^2 + \dots = \frac{2}{1-\delta}$.
- ▶ **If I deviate to C:** Payoff is $1 + 1\delta + 1\delta^2 + \dots = \frac{1}{1-\delta}$.
- ▶ Since $2 > 1$, playing D is strictly better than C.
- ▶ Therefore, carrying out the punishment is optimal (credible threat).
- ▶ Note: This relies on (D,D) being a stage Nash equilibrium.

Feasibility region for δ

- ▶ The condition $\delta \geq (T - R)/(T - P)$ defines the patience needed for cooperation.
- ▶ As the punishment payoff P approaches R , the fraction shrinks and the inequality tightens.
- ▶ When stage payoffs are farther apart, the required δ becomes smaller.

Intuition for δ threshold

- ▶ Higher δ means future payoffs weigh more; the punishment phase becomes more costly.
- ▶ When δ is small, the loser immediately recovers from punishment, so the threat loses bite.
- ▶ The threshold $\frac{T-R}{T-P}$ captures how patient players must be before cooperation pays.

Continuation payoff definition

- ▶ After seeing history h , the continuation payoff $\gamma_i(h)$ is the discounted value of future play.
- ▶ Formally:

$$\gamma_i(h) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(s^{t+|h|} | h).$$

- ▶ Continuation payoffs act as the state variable in our recursive enforcement problem.

Dynamic programming overview

- ▶ Dynamic programming breaks the infinite-horizon problem into a stage payoff plus a continuation value.
- ▶ The value function satisfies a recursive equation (Bellman equation) linking today and tomorrow.
- ▶ We use DP to ensure a given path is incentive compatible by adjusting continuation payoffs.

Bellman equation for repeated games

- ▶ The Bellman equation reads:

$$V_i(h) = \max_{a_i} \{u_i(a_i, a_{-i}) + \delta \mathbb{E}[V_i(h, a_i, a_{-i})]\},$$

- ▶ The expectation accounts for how others respond and how public signals evolve.
- ▶ Solving this equation gives us continuation payoffs that support desired equilibria.

Recursive enforcement example

- ▶ Suppose we aim to support (C,C) with a target payoff v .
- ▶ Choose a stage action delivering $u(C,C)$ and set continuation payoff γ so that $u(C,C) + \delta\gamma = v$.
- ▶ The punishment continuation payoff must satisfy the Bellman inequality for deviators.

Continuation state as future value

- ▶ The continuation payoff $\gamma_i(h)$ serves as the state variable in DP.
- ▶ When players reach history h , they compare the current payoff with the future value $\gamma_i(h)$.
- ▶ Properly chosen continuation values ensure that deviating today reduces the total payoff.

Using continuation payoffs to support paths

- ▶ To support a target path, split its payoff into the current stage reward and continuation payoff.
- ▶ Deviations must yield no higher total payoff when combined with the punishment continuation value.
- ▶ This decomposition is what makes the folk theorem constructive.

Folk Theorem

Perfect monitoring enforcement

Folk Theorem (Informal Statement)

Theorem: Folk Theorem for Infinite Repeated Games

Let G be a stage game with minmax payoff vector $(\underline{v}_1, \underline{v}_2)$. For any feasible payoff vector (v_1, v_2) with $v_i > \underline{v}_i$ for all i (strictly above minmax), there exists $\bar{\delta} < 1$ such that for all $\delta \geq \bar{\delta}$, there is a subgame perfect equilibrium of the infinitely repeated game with average payoff (v_1, v_2) .

- ▶ Key: achievable payoffs are in the **interior** of the feasible region above minmax
- ▶ Boundary points (including minmax itself) can only be approximated

Implications of the Folk theorem

- ▶ The equilibrium set is huge: any individually rational payoff can be implemented.
- ▶ Cooperation can be sustained even if the stage game has a unique Nash equilibrium.
- ▶ This flexibility is obtained by carefully designing continuation payoffs and punishment paths.

Constructing SPNE via continuation payoffs

- ▶ Pick a target payoff vector v that lies strictly above the minmax values (in the interior).
- ▶ Choose stage actions to deliver the current reward, and assign continuation payoffs that satisfy incentive constraints.
- ▶ The Bellman recursion ensures these continuation payoffs themselves come from equilibrium paths.

Example: Extended Teamwork Game

	C	D	B
C	(3, 3)	(1, 4)	(0, 0)
D	(4, 1)	(2, 2)	(1, 1)
B	(0, 0)	(1, 1)	(-1, -1)

- ▶ C = Cooperate, D = Defect, B = Bully (punish opponent)
- ▶ B is a costly punishment strategy that hurts both players
- ▶ This game extends the Teamwork game with an additional punishment option

Minmax vs Nash Equilibrium

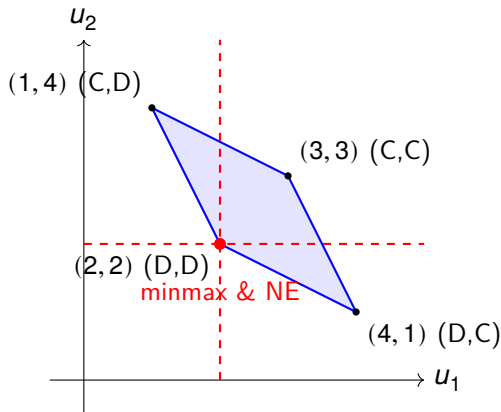
- ▶ **Nash Equilibrium:** (D,D) with payoff (2,2)
- ▶ **Minmax for Player 1:** Player 2 plays B, holding Player 1 to payoff 0
- ▶ **Minmax for Player 2:** Player 1 plays B, holding Player 2 to payoff 0
- ▶ **Minmax vector:** (0,0)
- ▶ **Pareto Dominance:** Nash equilibrium (2,2) strictly Pareto dominates minmax (0,0)
- ▶ **Folk Theorem:** For high δ , any payoff strictly above (0,0) is achievable, e.g., (3,3)

Teamwork Game (Reminder)

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

- ▶ C = Cooperate, D = Defect
- ▶ Unique Nash equilibrium: (D,D) with payoff (2, 2)
- ▶ (C,C) is Pareto superior with payoff (3, 3)

Folk Theorem: Achievable Payoff Region



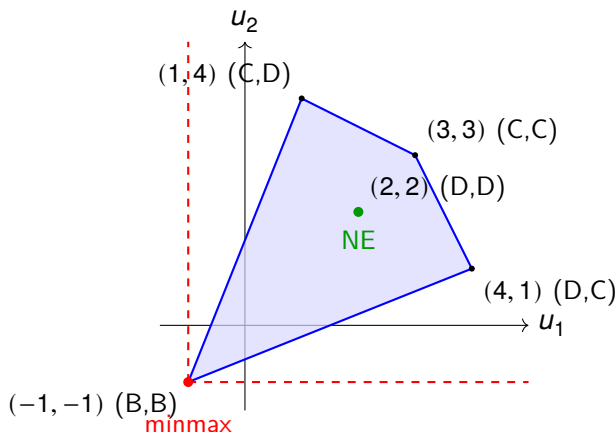
- Minmax: $(2, 2)$ (same as NE). Any payoff in feasible set strictly above $(2, 2)$ achievable for high δ

Example: Extended Teamwork Game

	C	D	B
C	(3, 3)	(1, 4)	(0, 0)
D	(4, 1)	(2, 2)	(1, 1)
B	(0, 0)	(1, 1)	(-1, -1)

- ▶ C = Cooperate, D = Defect, B = Bully (punish opponent)
- ▶ Nash equilibrium: (D,D) with payoff (2, 2)
- ▶ Minmax: $(-1, -1)$ - Player can hold opponent to -1 by playing B

Extended Teamwork: Minmax vs Nash Equilibrium



- NE (2,2) Pareto dominates minmax (-1, -1). Folk Theorem: payoffs strictly above (-1, -1) achievable

Monitoring and signals

When triggers misfire

Public imperfect monitoring

- ▶ Players observe a public signal y_t that depends probabilistically on the action profile.
- ▶ The signal may be noisy, so identical signals can arise from both cooperation and defection.
- ▶ Imperfect monitoring complicates the detection of deviations.

Trigger strategies with noisy signals

- ▶ A trigger strategy may punish when the signal indicates defection, even if the signal is a false alarm.
- ▶ With noise, even faithful cooperation can trigger the punishment with positive probability.
- ▶ The expected time spent in punishment grows with δ , which can hurt payoffs.

High δ and reduced average payoffs

- ▶ As δ increases, each punishment spell lasts longer in present-value terms.
- ▶ False positives (due to monitoring errors) cause repeated punishments, dragging down the average payoff.
- ▶ Paradoxically, being more patient can reduce average welfare if the monitoring is too noisy.

Illustrating the noise effect

- ▶ Suppose noise forces a punishment with probability ε every period.
- ▶ The expected cost of punishment is proportional to $\frac{\delta}{1-\delta}$, which explodes as $\delta \rightarrow 1$.
- ▶ Designers therefore prefer forgiving punishments that limit the length of punishment spells.

Expected Payoff under Grim Trigger with Noise

Setup: Teamwork game with Grim trigger. Noise causes false alarm with probability ε per period.

- ▶ Cooperation payoff: $R = 3$. Punishment payoff: $P = 2$.
- ▶ Expected periods until false alarm: $\frac{1}{\varepsilon}$.
- ▶ Once punishment triggered, it lasts forever.
- ▶ Expected payoff:

$$\begin{aligned} U &= (1 - \delta) \left[\sum_{t=1}^{\infty} \{ \delta (1 - \varepsilon) \}^{t-1} R + \delta^{t-1} \{ 1 - (1 - \varepsilon)^{t-1} \} P \right] \\ &= (1 - \delta) \sum_{t=1}^{\infty} \{ \delta (1 - \varepsilon) \}^{t-1} (R - P) + P \\ &= \frac{1 - \delta}{1 - \delta(1 - \varepsilon)} (R - P) + P \end{aligned}$$

- ▶ As $\delta \rightarrow 1$: punishment phase dominates, $U \rightarrow P = 2$ (worse than cooperation $R = 3$)

Forgiving strategies as a remedy

- ▶ Generous tit-for-tat and win-stay lose-shift forgive occasional defections or noise-induced punishments.
- ▶ They reduce the time spent in punishment while still deterring persistent deviators.
- ▶ These strategies are easier to analyse with continuation payoffs that reflect lenient responses.

Continuation payoffs with public signals

- ▶ Now the continuation payoff depends on both the history and the current signal: $\gamma_i(h, y)$.
- ▶ The state updates after each signal and action profile, keeping track of where we stand in the enforcement path.
- ▶ DP methods generalise by conditioning on public information as well as past actions.

Bellman equation with signals

- ▶ The equation becomes:

$$V_i(h, y) = \max_{a_i} \left\{ u_i(a_i, a_{-i}) + \delta \sum_{y'} \pi(y' \mid a_i, a_{-i}) V_i(h, y') \right\}.$$

- ▶ The transition probability π captures the noisy public signal.
- ▶ Solving this recursion yields continuation payoffs that account for imperfect monitoring.

Alternative lenient punishments

- ▶ Generous tit-for-tat occasionally forgives defections, limiting punishment duration.
- ▶ Win-stay lose-shift cooperates after mutual cooperation and switches after a bad outcome.
- ▶ Bounded-memory strategies that forget old deviations can also reduce the cost of false alarms.
- ▶ The key is balancing deterrence with forgiveness to avoid endless punishment cycles.

Private Monitoring: Additional Challenges

- ▶ **Private monitoring:** Each player observes a private signal about others' actions, not a public outcome.
- ▶ Players cannot coordinate on a common history, making punishment harder to enforce.
- ▶ **Costly observation:** Acquiring information may require effort or resources, creating a trade-off between monitoring accuracy and cost.
- ▶ Private signals can lead to disagreement about what happened, complicating equilibrium construction.
- ▶ Folk Theorems for private monitoring require more complex belief-updating and communication protocols.

Advanced Topics in Repeated Games

- ▶ **Alternating games:** Play different stage games in even/odd periods to effectively reduce the discount factor and expand equilibrium set.
- ▶ **Review strategies:** Maintain the same action for multiple periods to accumulate information and improve signal precision before adjusting behavior.
- ▶ **Public randomization:** Use publicly observable random devices (e.g., sunspots) to coordinate on mixed-strategy equilibria or break symmetry.
- ▶ **Communication:** Allow cheap talk or costly signaling to share private information and coordinate on better outcomes.
- ▶ **Mediators:** Introduce a trusted third party to receive reports, recommend actions, and facilitate coordination in complex environments.

Takeaways & looking ahead

- ▶ Infinite repetition enables cooperation via credible threats and continuation payoffs.
- ▶ Folk theorems show that patience (high δ) expands the equilibrium set dramatically.
- ▶ Imperfect monitoring complicates enforcement and may require forgiving strategies.
- ▶ Next time we explore applications and extensions: reputation, renegotiation, and mechanism design.

Group Homework

Group Homework: The Evolution of Trust

- ▶ **Play:** [The Evolution of Trust](#) interactive game, exploring repeated game strategies and their outcomes.
- ▶ **Report (A4, one page):** Summarize your key insights from the game, including:
 - ▶ Which strategies performed best and why
 - ▶ How the game relates to concepts from this module (Grim Trigger, Tit-for-Tat, etc.)
 - ▶ What you learned about cooperation and trust in repeated interactions
- ▶ **Contribution log:** Include a short paragraph naming who contributed what to the report.
- ▶ **Submission:** Email a single PDF per group to y.hino@vju.ac.vn by Nov. 30, Sunday 23:59.

Group Homework: The Evolution of Trust

Evaluation focus

1. **Content Comprehension:** Demonstrate understanding of the game's mechanics and how they relate to repeated game theory concepts covered in class.
2. **Organization & Clarity:** Present your insights in a well-structured, easy-to-follow format within the one-page limit.
3. **Critical Insight:** Identify key takeaways about cooperation, punishment, and forgiveness strategies, and explain why certain strategies succeed or fail.
4. **Application:** Connect the game's lessons to real-world scenarios or course concepts, showing depth of understanding.