Repeated Games

IE 5545: Decision Analysis
Prof. Krishnamurthy Iyer*

May 2, 2026

Having developed the concept of SPNE for dynamic games of complete information, we will now look at the properties of SPNE in some of the most common forms of dynamic games, namely repeated games. A repeated game arises when a game G is played repeatedly among a group of players. For example, two players may play a "best-of-three" Rock-Papers-Scissors (RPS), or two firms might be repeatedly engaged in quantity competition (repeated Cournot competition). A natural question that arises in such repeated games is how do the SPNE of the repeated game related to the NE of the game G. We will briefly explore this question next.

1 Finitely repeated games

We start by formally defining what we mean by a repeated game. Fix a static game G among N players, where player i's payoff function is given by π_i^G. We first consider finitely repeated games, where the game G is repeated a finite (commonly known) number of times.

Definition 1. *For any $T \geq 1$, the game $G(T)$ is a dynamic game where the players play T rounds of G, at times $t = 0, \cdots, T-1$, observing the outcome of each round before playing the next round. The payoff of each player i is equal to the sum of the payoffs they receive in each of the T rounds of G:*

$$\pi_i = \sum_{t=0}^{T-1} \pi_i^G(a_i^t, a_{-i}^t),$$

where (a_i^t, a_{-i}^t) is the action profile that is played at time t.

Note that according to this definition, the players observe what happened the first time G is played before choosing their actions the second time G is played. Thus, the players can choose their action in the second instance of G depending on what happened the first time. Thus, if G is Rock-Paper-Scissors (RPS), then in $G(2)$, player 1 may choose to play P in the second instance of G only if (R, S) was played in the first time, and otherwise play S.

1.1 Information sets, strategies, and subgames

To analyze the repeated game, let us first understand its structure. We would like to know the information sets and the pure strategies for each player. Similarly, before identifying SPNE, we would like to understand the subgames.

Fix a static game G, and focus on the finitely repeated game $G(T)$. Recall that in a repeated game, each player observes the outcomes of the previous rounds before she has to choose her action for the next round. This directly leads us to the following observation:

Observation 1. *In the repeated game $G(T)$, one can associate each information set of a player i after t rounds with a unique sequence of plays in the first t rounds (including the empty sequence).*

Thus, if G is RPS, then in $G(2)$, player 1 has an information set in the first round corresponding to the empty sequence, and 9 different information sets in the second round corresponding to each possible play in the first round. (Same for player 2.) Similarly, in $G(3)$, in addition to the above information sets, there are 81 information sets for the player in the third round, corresponding to each possible combination of outcomes in the first two rounds (e.g., one such information set corresponds to the outcome in the first two rounds being $((R, R), (S, P))$).

The preceding observation implies that we can specify a strategy for player i in $G(T)$ by specifying what player i will do after any sequence of plays of length at most $T - 1$:

Observation 2. *In $G(T)$, one can specify a player i's strategy by specifying what player i will play after any sequence of plays of length at most $T - 1$.*

Finally, one must describe the subgames in $G(T)$. Recall that G is a static game and hence has no proper subgame. Then, the only subgames in $G(T)$ are the subgames that start after players have already played G until time t, for some $0 \leq t \leq T - 1$. For $t = 0$, we get the original game $G(T)$. For any other $t \leq T - 1$, the game G has already been played for t rounds, and remains to be played for $T - t$ more rounds. Note that this subgame is just $G(T - t)$. In particular, there is a subgame identical to $G(T - t)$ after each sequence of outcomes of length t.

2 SPNEs in repeated games

An SPNE of $G(T)$ requires the players to play a Nash equilibrium in every subgame of $G(T)$, even those that will not be played under the SPNE. As mentioned earlier, our goal is to relate the SPNEs of $G(T)$ to the NEs of G.

We start by considering a simple setting where G has a unique Nash equilibrium σ. What are the SPNE of $G(T)$? A natural contender is a strategy profile where every player plays their corresponding strategy in σ in every round. The following result shows that indeed this intuition is correct:

Theorem 1. *Suppose G has a unique NE σ. Then the* only *SPNE of $G(T)$ is $\sigma^T = (\sigma_1^T, \sigma_2^T, \cdots, \sigma_N^T)$, where σ_i^T involves player i playing σ_i in each of her information sets. In other words, each player i plays the strategy σ_i in each round of $G(T)$.*

For intuition, suppose G is the prisoner's dilemma, which has the unique NE (B, B), where both players betray. Then, the theorem suggests that if prisoner's dilemma is repeated for a finite number of rounds, then the only SPNE of the repeated game involves the two players betraying each other in every round. This is true no matter how many times the prisoner's dilemma is repeated, as long as it is finite.

We will provide a brief proof of this result:

Proof. Note that since we are focused on the SPNE of $G(T)$, we require that in any subgame after $T - 1$ rounds, the players must play a Nash equilibrium. Now, the subgame after $T - 1$ rounds is just $G(1) = G$, and hence has a unique NE σ. Thus, from this, we obtain that in any subgame after $T-1$ rounds (i.e., in the last round), the players have to play σ in any SPNE.

Now consider the subgame after $T-2$ rounds. This subgame is same as $G(2)$, i.e., G repeated twice. For an SPNE, we require that the players play an SPNE in this game $G(2)$. Now note that from previous argument, the players are going to play σ in round T, no matter what players play at round $T-1$. Thus, essentially, what happens in round $T-1$ has no bearing on what happens in round T. It follows that the strategy of the players in the two rounds are independent of each other. Thus, we can think of the play in the subgame as the players playing two independent rounds of G. Since G has a unique NE, the play in each of this round must be σ. Thus, in any SPNE $G(T)$, the play in the last two rounds must be σ.

We can repeat the argument in the preceding paragraph inductively for the subgame after t rounds, for $t = T-2, T-3, \cdots, 1, 0$. Since the subgame after 0 rounds is just $G(T)$, the only SPNE of $G(T)$ is where each player i plays σ_i in each of her information sets. $\square$

The above argument can be easily modified to obtain the following result:

Theorem 2. *Suppose G has an NE σ. Then $\sigma^T = (\sigma_1^T, \sigma_2^T, \cdots, \sigma_N^T)$ is an SPNE of $G(T)$.*

Proof. Exercise. $\square$

Also, if G has multiple NE, then the preceding backwards-induction argument can be easily adapted to show that there is an SPNE $\hat{\sigma}^T$ where the players play a prespecified NE at each *time* period (but the same NE strategies at all information sets of a fixed time period). For example, in a finitely repeated co-ordination game, the players alternating between (S, S) and (ST, ST) NE at consecutive rounds would also constitute an SPNE. Note that here, the choice of which NE to play at any round is *time-dependent*.

The backward-induction argument fails if the choice of NE strategies to be played in different rounds is *history-dependent*: that is, the players decide which NE to play depending on the entire sequence of past outcomes. This is because, in this case, a player's action can influence which NE is played in the future. This influence on future implies that we can no longer argue, for example, that "what happens in round $T-1$ has no bearing on what happens in round T". Thus, we cannot use backward induction to conclude such strategy profiles must constitute SPNEs. In particular, it is straightforward to construct strategy profiles where each player plays the NE strategy in each round, but the strategy profile is not an SPNE.

(To emphasize, this does not mean that no history-dependent strategy profile can be an SPNE — just that the backward induction argument fails. In fact, by carefully choosing which NE is to be played after each sequence of outcomes, one can indeed construct a history-dependent strategy profile that is an SPNE. The previous discussion only states that not every assignment of NE strategies to different rounds must be an SPNE.)

Theorem 1 proves that in an SPNE of a finitely repeated game $G(T)$ where the stage game G has a unique Nash equilibrium, the players have no choice but to repeatedly play the Nash equilibrium in each round. This suggests that in such settings, no new or interesting behavior arises in a repeated game – players play the repeated game as if it were independent copies of the finite game. Note that proof of Theorem 1 requires two assumptions:

1. The game G has a unique NE. Without this assumption, one could selectively choose different NE in different subgames, based on what has happened earlier. Thus, as discussed above, if G has two different NEs, then the NE chosen after $T-1$ rounds can depend on the action chosen at round $T-1$, implying that the two rounds are not independent.

2. The players know exactly how many rounds the game is repeated. Note that this is necessary to start the inductive argument in the proof above.

In the following, we will investigate what happens when the second assumption fails: does there exist a non-trivial SPNE if the players do not know how many times they will play the game (or if they play the game infinitely)? (In the assignment, you investigate what happens when the first assumption fais; in particular, you show that there exist non-trivial SPNE in a twice repeated game where players play a strictly dominated action in the first round.)

3 SPNE in infinitely repeated games

We have the following definition of infinitely repeated games:

Definition 2. *Fix a stage game G, and $\delta \in (0, 1)$. Then the repeated game $G(\infty, \delta)$ is a dynamic game of complete information, where players play infinitely many rounds of G, observing the outcome of each round before playing the next round. The payoff of each player i is equal to the discounted sum of the payoffs they receive in each round, where the payoff at time t is discounted by δ^t:*

$$\pi_i = \sum_{t=0}^{\infty} \delta^t \pi_i^G(s_i^t, s_{-i}^t).$$

There are two interpretations of $G(\infty, \delta)$, as follows:

1. The players are playing infinite rounds of G, however future payoffs are not worth as much as current payoffs. Essentially, a payoff of 1 at time t is worth only a payoff of δ^t at time 0. This is similar to discounting future income in finance, since a \$1 today is worth more to you than a \$1 a year from now, since you can invest that dollar today to make more money in a year.

   One can think of δ as a measure of patience of players. When δ is small and close to zero, players do not care very much about payoffs that are obtained later. More precisely, when $\delta = 1/100$, a payoff of 1 at time $t = 1$ is worth more to the player than a payoff of 99 at time $t = 2$. In other words, the player is very impatient. On the other hand, for values of δ close to 1, the players can be thought of being very patient.

   Remark 1. *This is an assumption about how players care about future payoffs. Under this discounting, a payoff of 1 at time t is worth only δ^{t-1} at time 1. This constant rate of discounting is called "geometric discounting". Essentially, under this assumption, you have the following property: For any t, s, a, b, you prefer a payoff of a at time t over a payoff of b at time $s > t$, if and only if, you prefer a payoff of a at time 1 over a payoff of b at time $s - t + 1$. This is called "time consistent preference" – your preference between two options does not change if they are both equally postponed.*

   There are other notions of discounting, especially "hyperbolic discounting", where this time-consistency does not hold true. Under such discounting, you may prefer a payoff of a at time 1 over a payoff of b at time 2, but prefer b at time 101 over payoff of a at time 100. One example of such a discounting is when the payoffs at time t are discounted by $\frac{1}{t}$ instead of δ^t. Hyperbolic discounting is often used to model procrastination*, which is often a time-incosistent behavior.*

2. The players are playing the game $G(\tilde{T})$, where $\tilde{T}$ is a random variable unknown to the players. Moreover, the random variable $\tilde{T}$ has a $\mathsf{Geom}(1-\delta)$ distribution, meaning $\mathbf{P}(\tilde{T} = t) =$

		2	
		S	B
1	S	$-1,-1$	$-10,0$
	B	$0,-10$	$-5,-5$

Figure 1: Prisoner's dilemma

$(1-\delta)\delta^t$. In other words, each time G is played, the game ends (no more rounds of G) with probability $1-\delta$, and the game continues with probability δ.

In this case, the payoff of player i under any strategy profile σ is given by

$$\pi_i(\sigma) = \mathbf{P}(\tilde{T} \geq 0)\pi_i^G(\sigma^0) + \mathbf{P}(\tilde{T} \geq 1)\pi_i^G(\sigma^1) + \mathbf{P}(\tilde{T} \geq 2)\pi_i^G(\sigma^2) + \cdots + \mathbf{P}(\tilde{T} \geq t)\pi_i^G(\sigma^t) + \cdots .$$

Here σ^t is the strategy profile played in round t. Thus, $\pi_i^G(\sigma^t)$ is the payoff the player i gets the t^{th} time G is played. This is multiplied by the probability that the game will last for at least t rounds, which is equal to $\mathbf{P}(\tilde{T} \geq t)$. Since $\mathbf{P}(\tilde{T} \geq t) = \delta^t$, we get

$$\pi_i(\sigma) = \pi_i^G(\sigma^0) + \delta\pi_i^G(\sigma^1) + \delta^2\pi_i^G(\sigma^2) + \cdots + \delta^t\pi_i^G(\sigma^t) + \cdots ,$$

which is the same as the discounted payoff.

Finally, note that, just as in the finitely repeated game setting, in an infinitely repeated game, we can associate each information set of a player i with a unique sequence of play until that round. Hence, again, a strategy of player i can be stated by specifying her choice of action after every finite sequence of play in the game. Similarly, in $G(\infty, \delta)$, there exists a subgame for each finite sequence of play in the game. (Note that this subgame is identical to a copy of $G(\infty, \delta)$.)

3.1 Repeated prisoner's dilemma

Now we focus on infinitely repeated games $G(\infty, \delta)$ where G has a unique Nash equilibrium, and ask whether we can support repeated plays of actions that are not NEs of G as SPNEs of $G(\infty, \delta)$. For concreteness suppose G is the prisoner's dilemma with payoff matrix as in Figure 1.

Clearly, in this game, if the players could cooperate, both players prefer playing (S, S) over (B, B). The worry for player 1 is that player 2 will deviate and play B instead of S if she plays S. Note that since the game is played for infinite rounds, player 1 can "punish" such deviation by playing B from thereon. We want to find conditions under which such a punishment ensures co-operation from both players. We begin by giving a formal description of the punishment strategy.

Formally, we define a *grim trigger* strategy s_1^{grim} for player 1 as follows:

1. At time $t = 0$, start by playing S.
2. At time $t \geq 1$, if the play has been (S, S) throughout, then play S. Otherwise play B.

We can similarly define the grim-trigger strategy s_2^{grim} for player 2. Note that the grim trigger strategy starts by co-operating and playing S. As long as no player has deviated from S, the grim trigger strategy continues cooperating by playing S. If either of the player has deviated and played B, then grim-trigger strategy stops cooperating and forever plays B. In other words, all it takes is one B for the grim trigger strategy to stop playing S and playing B throughout. (It is in this sense is this a trigger strategy. It is grim in the sense that after one betrayal, no amount of cooperation from the other player can convince the strategy to start cooperating.)

Note that under the strategy profile $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$, the play in the game $G(\infty, \delta)$ is $(S,S),(S,S),(S,S),\cdots$. Thus the payoff of player 1 under this strategy profile is given by

$$\pi_1(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}}) = -1 + \delta(-1) + \delta^2(-1) + \cdots = -1(1 + \delta + \delta^2 + \cdots) = \frac{-1}{1-\delta}.$$

(Note that since the game is symmetric, player 2's payoff is also $\frac{-1}{1-\delta}$ under $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$.)

We want to find conditions under which $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ is an SPNE of $G(\infty, \delta)$. This requires that this strategy profile induces an NE in every subgame of $G(\infty, \delta)$.

We first ask when is $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ an NE of the $G(\infty, \delta)$? Due to symmetry, this is equivalent to determining conditions under which

$$\pi_1(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}}) \geq \pi_1(s_1, s_2^{\mathsf{grim}})$$

for any strategy s_1 of player 1.

Note that since we have an infinitely repeated game, there are infinitely many possible strategies s_1 for player 1. For example, one possible deviation s_1 could be play B forever. Another one could be play B in round t if t is a prime, and play S otherwise. Other deviation could be play B twice, and do what s_1^{grim} does from thereafter. Checking that s_1^{grim} indeed performs better than all such deviations seems to be difficult, if not impossible.

However, since player 2 is playing a grim-trigger strategy, the first time player 1 stops playing S, player 2 will start playing B after that, and hence the only best response for player 1 subsequent to a deviation is to play B forever after. Thus, in order to check whether s_1^{grim} is a best-response to s_2^{grim}, it is enough for us to consider deviations s_1 for player 1, where player 1 plays S until some time $t-1$, deviates at time t (to play B), and subsequent to that plays B forever.

For such a deviation, the play will be (S,S) until time $t-1$, (B,S) at time t and (B,B) from then on. Thus, the payoff to player 1 under such a deviation is given by

$$\begin{aligned}\pi_1(s_1, s_2^{\mathsf{grim}}) &= -1(1 + \delta + \delta^2 + \cdots \delta^{t-1}) + 0\delta^t + -5(\delta^{t+1} + \delta^{t+2} + \cdots) \\ &= -1\left(\frac{1-\delta^t}{1-\delta}\right) - \frac{5\delta^{t+1}}{1-\delta}.\end{aligned}$$

From these expressions, one can check that $\pi_1(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}}) \geq \pi_1(s_1, s_2^{\mathsf{grim}})$ if and only if $1 \leq (1 - \delta^t) + 5\delta^{t+1}$, which holds if and only if $\delta \geq \frac{1}{5}$. By symmetry, we conclude that s_2^{grim} is a best-response to s_1^{grim} if and only if $\delta \geq 1/5$. Thus, $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ is an NE.

To conclude that $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ is an SPNE, we must also verify that this strategy profile is an NE in every subgame, not just the original game. To see this, observe that we can split all the subgames into two categories: in the first category, the play has always been (S,S) until then, and in the second category, at least one player has played B at least once prior to the subgame. For every subgame in the first category, the strategy profile $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ induces a game exactly same as the original game, and hence the strategy profile is an NE for subgames in this category. For every subgame in the second category, the strategy profile $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ always plays (B,B), also an NE for the subgame. Thus, $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ also induces an NE for such subgames. Thus, indeed $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ induces an NE in every subgame, and hence is an SPNE.

From this analysis, we see that if players are sufficiently patient, i.e. $\delta \geq 1/5$, the we can sustain co-operation (the actions (S,S)) as an SPNE in the infinitely repeated prisoner's dilemma, even though this is not possible in any finitely repeated prisoner's dilemma.

Given our interpretation of $G(\infty,\delta)$ as $G(\tilde{T})$ for a random $\tilde{T}$, we can conclude that if the total number of rounds that the prisoner's dilemma is repeated is unknown to the players, then again one could sustain cooperation indefinitely, as long as $\delta \geq 1/5$. (Since $\tilde{T} \sim \mathsf{Geom}(1-\delta)$, we have $\mathbf{E}[T] = \frac{1-\delta}{\delta} = 4$. Thus, for these payoff values, cooperation can be sustained if players play a random number (with a geometric distribution) of games with a mean of $\bar{T}+1 = 5$.)

3.2 One-step deviation principle

In the previous analysis for showing that $(s_1^{\mathsf{grim}}, s_2^{\mathsf{grim}})$ is an SPNE, we were able to directly check against all deviations and see none could do better if δ were large enough. This was possible primarily because the strategy s_i^{grim} was quite simple. If the strategies are more complex, it is unclear how to go about checking if they are indeed best responses in every subgame.

To overcome this difficulty, we will use the following principle, known as the *one-step deviation principle*. (We will not prove this. However, it follows from dynamic programming.) To state the principle, we need to define the concept of a *one-step deviation*.

Definition 3 (one-step deviation). *Fix a strategy s_i for player i. A strategy $\widehat{s}_i$ is a* one-step deviation *from s_i in a subgame E of $G(\infty,\delta)$, if under $\widehat{s}_i$, player i deviates from the action played by s_i in her first information set in E, and follows the recommendation of s_i subsequently.*

As an example, consider the "tit-for-tat" strategy s_i^{tft}, where player i starts by playing S at time $t=0$, and at any time $t \geq 1$, plays whatever her opponent played the previous time period. Suppose both players play the tit-for-tat strategy. Then, the sequence of play is given by $(S,S),(S,S),(S,S),\ldots$. The payoff to player 1 is $\frac{-1}{1-\delta}$.

A one-step deviation for player 1 at $t=0$ is a strategy $\widehat{s}_1$ that plays B at $t=0$ (the deviation), and at $t \geq 1$ does what s_1^{tft} recommends: play the opponent's strategy. So, if player 1 plays this strategy and player 2 still plays s_2^{tft}, the sequence of play would be $(B,S),(S,B),(B,S),\ldots$. The payoff to player 1 is $(0+(-10)\delta+0\delta^2+(-10)\delta^3+\cdots) = \frac{-10\delta}{1-\delta^2}$.

An strategy that is not a one-step deviation at $t=0$ is the strategy $\widetilde{s}_1$ for player 1 which plays B in the first two periods, and then does what s_1^{tft} recommends. (This will be a two-step deviation.) If player 1 plays $\widetilde{s}_1$ and player 2 plays s_2^{tft}, the sequence of play will be $(B,S),(B,B),(B,B),(B,B),\ldots$.

The one-step deviation principle provides a relatively simpler way to verify that a strategy profile is an SPNE: ensure that there is no-one *profitable* step deviation in any subgame for any player. In particular, using this principle, we need to concern with two or multi-step deviations.

Lemma 1 (One-step deviation principle). *The one-step deviation principle states that in any game $G(\infty,\delta)$, if for a given strategy profile $(s_1,s_2,\cdots,s_N)$, no player i has a profitable one-step deviation from s_i in **any** subgame, then $(s_1,\cdots,s_N)$ is an SPNE.*

Thus, in order to check if a strategy profile is an SPNE, it suffices to check whether there exists a profitable one-step deviation for any player in any subgame. Note that this is what we did in the case of the grim-trigger strategies, where the deviation we considered for player i were indeed one-step deviations, where player 1 deviates from s_1^{grim} in the first information set of the subgame starting at time t, but follows the recommendation of s_1^{grim} (which is B) subsequently.

The one-step deviation principle allows us to check whether a broader class of strategy profiles constitute an SPNE.

Exercise: Using the one-step deviation principle, identify if there exists a $\delta \in (0, 1)$ for which both players playing the "tit-for-tat" strategy constitutes an SPNE. One must verify that there is no profitable one-step deviation for either player in every subgame (Hint: there are *four* types of subgames.)