Chapter 8

Repeated Games

In the games considered so far the players took just a single decision: a strategy they selected. In this chapter we consider a natural idea of playing a given strategic game repeatedly. We assume that the outcome of each round is known to all players before the next round of the game takes place.

8.1 Finitely repeated games

In the first approach we shall assume that the same game is played a fixed number of times. The final payoff to each player is simply the sum of the payoffs obtained in each round.

Suppose for instance that we play the Prisoner's Dilemma game, so

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

twice. It seems then that the outcome is the following game in which we simply add up the payoffs from the first and second round:

	CC	CD	DC	DD
CC	$4,4$	$2,5$	$2,5$	$0,6$
CD	$5,2$	$3,3$	$3,3$	$1,4$
DC	$5,2$	$3,3$	$3,3$	$1,4$
DD	$6,0$	$4,1$	$4,1$	$2,2$

However, this representation is incorrect since it erronously assumes that the decisions taken by the players in the first round have no influence on their decisions taken in the second round. For instance, the option that the first player chooses C in the second round if and only iff the second player chose C in the first round is not listed. In fact, the set of strategies available to each player is much larger.

In the first round each player has two strategies. However, in the second round each player's strategy is a function $f : \{C, D\} \times \{C, D\} \to \{C, D\}$. So in the second round each player has $2^4 = 16$ strategies and consequently in the repeated game each player has $2 \times 16 = 32$ strategies. Each such strategy has two components, one of each round. It is clear how to compute the payoffs for so defined strategies. For instance, if the first player chooses in the first round C and in the second round the function

$$f_1(s) := \begin{cases} C & \text{if } s = (C, C) \\ D & \text{if } s = (C, D) \\ C & \text{if } s = (D, C) \\ D & \text{if } s = (D, D) \end{cases}$$

and the second player chooses in the first round D and in the second round the function

$$f_2(s) := \begin{cases} C & \text{if } s = (C, C) \\ D & \text{if } s = (C, D) \\ D & \text{if } s = (D, C) \\ C & \text{if } s = (D, D) \end{cases}$$

then the corresponding payoffs are:

- in the first round: $(0, 3)$ (corresponding to the joint strategy (C, D)),
- in the second round: $(1, 1)$ (corresponding to the joint strategy (D, D)).

So the overall payoffs are: $(1, 4)$, which corresponds to the joint strategy (CD, DD) in the above bimatrix.

Let us consider now the general setup. The strategic game that is repeatedly played is called the ***stage game***. Given a stage game $(S_1, \ldots, S_n, p_1, \ldots, p_n)$ the ***repeated game with*** k ***rounds*** (in short: a repeated game), where $k \geq 1$, is defined by first introducing the set of ***histories***. This set $\mathcal{H}$ is defined inductively as follows, where ε denotes the empty sequence, $t \geq 1$, and, as usual, $S = S_1 \times \ldots \times S_n$:

$$\begin{aligned}
&\mathcal{H}^0 := \{\varepsilon\}, \\
&\mathcal{H}^1 := S, \\
&\mathcal{H}^{t+1} := \mathcal{H}^t \times S, \\
&\mathcal{H} := \textstyle\bigcup_{t=0}^{k-1} \mathcal{H}^t.
\end{aligned}$$

So $h \in \mathcal{H}^0$ iff $h = \varepsilon$ and for $t \in \{1, \ldots, k-1\}$, $h \in \mathcal{H}^t$ iff $h \in S^t$. That is, a history is a (possibly empty) sequence of joint strategies of the stage game of length at most $k - 1$.

Then a ***strategy*** for player i in the repeated game is a function $\sigma_i : \mathcal{H} \to S_i$. In particular $\sigma_i(\varepsilon)$ is a strategy in the stage game chosen in the first round.

We denote the set of strategies of player i in the repeated game by Σ_i and the set of joint strategies in the repeated game by Σ.

The ***outcome*** of the repeated game corresponding to a joint strategy $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Sigma$ of the players is the history that consists of k joint strategies selected in the consecutive stages of the underlying stage game. This history $(o^1(\sigma), \ldots, o^k(\sigma)) \in \mathcal{H}^k$ is defined as follows:

$$\begin{aligned}
&o^1(\sigma) := (\sigma_1(\varepsilon), \ldots, \sigma_n(\varepsilon)), \\
&o^2(\sigma) := (\sigma_1(o^1(\sigma)), \ldots, \sigma_n(o^1(\sigma))), \\
&\ldots \\
&o^k(\sigma) := (\sigma_1(o^1(\sigma), \ldots, o^{k-1}(\sigma)), \ldots, \sigma_n(o^1(\sigma), \ldots, o^{k-1}(\sigma)).
\end{aligned}$$

In particular $o^k(\sigma)$ is obtained by applying each of the strategies $\sigma_1, \ldots, \sigma_n$ to the already defined history $(o^1(\sigma), \ldots, o^{k-1}(\sigma)) \in \mathcal{H}^{k-1}$.

Finally, the ***payoff function*** P_i of player i in the repeated game is defined as

$$P_i(\sigma) := \sum_{t=1}^{k} p_i(o^t(\sigma)).$$

So the payoff for each player is the sum of the payoffs he received in each round.

Now that we defined formally a repeated game let us return to the Prisoner's Dilemma game and assume that it is played k rounds. We can now define the following natural strategies:[1]

[1]These definitions are incomplete in the sense that the strategies are not defined for all

- ***cooperate***: select at every stage C,
- ***defect***: select at every stage D,
- ***tit for tat***: first select C, then repeatly select the last strategy played by the opponent,
- ***grim*** (or ***trigger***): select C as long as the opponent selects C; if he selects D select D from now on.

For example, it does not matter if one chooses tit for tat or grim strategy against a grim strategy: in both cases each player repeatedly selects C. However, if one selects C in the odd rounds and D in the even rounds, then against the tit for tat strategy the following sequence of stage strategies results:

- for player 1: $C, D, C, D, C, \ldots$,
- for player 2: $C, C, D, C, D, \ldots$

while against the grim strategy we obtain:

- for player 1: $C, D, C, D, C, \ldots$,
- for player 2: $C, C, D, D, D, \ldots$

Using the concept of strictly dominant strategies we could predict that the outcome of the Prisoner's dilemma game is (D, D). A natural question arises whether we can also predict the outcome in the repeated version of this game. To do this we first extend the relevant notions to the repeated games.

Given a stage game G we denote the repeated game with k rounds by $G(k)$. After the obvious identification of $\sigma_i : \mathcal{H}^0 \to S_i$ with $\sigma_i(\varepsilon)$ we can identify $G(1)$ with G.

In general we can view $G(k)$ as a strategic game $(\Sigma_1, \ldots, \Sigma_n, P_1, \ldots, P_n)$, where the strategy sets Σ_i and the payoff functions P_i are defined above. This allows us to apply the basic notions, for example that of Nash equilibrium, to the repeated game.

As a first result we establish the following.

histories. However, the specified parts completely determine the outcomes that can arise against any strategy of the opponent.

Theorem 24 *Consider a stage game G and $k \geq 1$.*

(i) If s is a Nash equilibrium of G, then the joint strategy σ, where for all $i \in \{1, \ldots, n\}$ and $h \in \mathcal{H}$

$$\sigma_i(h) := s_i,$$

is a Nash equilibrium of $G(k)$.

(ii) If s is a unique Nash equilibrium of G, then for each Nash equilibrium of $G(k)$ the outcome corresponding to it consists of s repeated k times.

Proof.

(i) The outcome corresponding to σ consists of s repeated k times. That is, in each round of $G(k)$ the Nash equilibrium is selected and the payoff to each player i is $p_i(s)$, where $G := (S_1, \ldots, S_n, p_1, \ldots, p_n)$.

Suppose that σ is not a Nash equilibrium in $G(k)$. Then for some player i a strategy τ_i yields a higher payoff than σ_i when used against σ_{-i}. So in some round of $G(k)$ player i receives a strictly larger payoff than $p_i(s)$. But in this (and every other) round every other player j selects s_j. So the strategy of player i selected in this round yields a strictly higher payoff against s_{-i} than s_i, which is a contradiction.

(ii) We proceed by induction on k. Since we identified $G(1)$ with G, the claim holds for $k = 1$. Suppose it holds for $k \geq 1$.

Take a Nash equilibrium σ in $G(k+1)$. Consider the last joint strategy of the considered outcome. It constitutes the Nash equilibrium of the stage game. Indeed, otherwise some player i did not select a best response in the last round and thus can obtain a strictly higher payoff in G by switching in the last round of $G(k+1)$ to another strategy s'_i in G. The corresponding modification of σ_i according to which in the last round s'_i is selected yields against σ_{-i} a strictly higher payoff than σ_i. This contradicts the assumption that σ is a Nash equilibrium in $G(k+1)$.

Now redistribute for each player his payoff in the last round evenly over the previous k rounds, by modifying appropriately the payoff functions, and subsequently remove this last round. The resulting game is a repeated game $G'(k)$ such that s is a unique Nash equilibrium of G', so we can apply to it the induction hypothesis. Moreover, by the above observation each Nash equilibrium of $G(k+1)$ consists of a Nash equilibrium of $G'(k)$ augmented with s selected in the last round (i.e., in each Nash equilibrium of $G(k+1)$ each player i selects s_i in the last round). So the claim holds for $k+1$.

The claim now holds by induction. □

The definition of a strategy in a repeated game determines player's choice for *each* history, in particular for histories that cannot be outcomes of the repeated game. As a result the joint strategy from item (i) is not a unique Nash equilibrium of $G(k)$ when players have two or more strategies in the stage game.

As an example consider the Prisoner's Dilemma game played twice. Then the pair of defect strategies is a Nash equilibrium. Moreover, the pair of strategies according to which one selects D in the first round and C in the second round iff the first round equals (C, C) is also a Nash equilibrium. These two pairs differ though they yield the same outcome.

Note, further that if a player has a strictly dominant strategy in the stage game then he does not necessarily have a strictly dominant strategy in the repeated game. In particular, choosing in each round the strictly dominant strategy in the stage game does not need to yield a maximal payoff in the repeated game.

Example 16 Take the Prisoner's Dilemma game played twice.

Consider first a best response against the tit for tat strategy. In it C is selected in the first round and D in the second round. In contrast, in each best response against the cooperate strategy in both rounds D is selected. So for each player no single best response strategy exists, that is, no player has a strictly dominant strategy.

In contrast, in the stage game strategy D is strictly dominant for both players. Note also that in our first, incorrect, representation of the Prisoner's Dilemma game played twice strategy DD is strictly dominant for both players, as well. □

In the one shot version of the Prisoner's Dilemma game we could predict that both players will select the defect (D) strategy on the basis that it is a strictly dominant strategy. The above theorem shows that when Prisoner's Dilemma game is played repeatedly cooperation still won't occur. However, this prediction is weaker, in the sense that selecting D repeatedly is not anymore a strictly dominant strategy. We can only conclude that in any Nash equilibrium in every round each player selects D.

The above theorem actually shows more: when the stage game has exactly one Nash equilibrium, then in each Nash equilibrium of the repeated game the players select their equilibrium strategies. So in each round their payoff is simply their payoff in the Nash equilibrium of the stage game.

However, when the stage game has more than one Nash equilibrium the situation changes. In particular, players can achieve in a Nash equilibrium of the repeated game a higher average payoff than the one achieved in any Nash equilibrium of the stage game.

Example 17 Consider the following stage game:

	A	B	C
A	$5,5$	$0,0$	$12,0$
B	$0,0$	$2,2$	$0,0$
C	$0,12$	$0,0$	$10,10$

This game has two Nash equilibria (A, A) and (B, B). So when the game is played once the highest payoff in a Nash equilibrium is 5 for each player. However, when the game is played twice a Nash equilibrium exists with a higher average payoff. Namely, consider the following strategy for each player:

- select C in the first round,
- if the other player selected C in the first round, select A in the second round and otherwise select M.

If each player selects this strategy, they both select in the first round C and A in the second round. This yields payoff 15 for each player.

We now prove that this pair of strategies forms a Nash equlibrium. The only way a player, say the first one, can receive a larger payoff than 15 is by selecting A in the first round. But then the second player selects B in the second round. So in the first round the first player receives the payoff 12 but in the second round he receives the payoff of at most 2. Consequently by switching to another strategy the first player can secure at best payoff 14. □

The above example shows that playing a given game repeatedly can lead to some form of coordination that can be beneficial to all players. This coordination is possible because crucially the choices made by the players in the previous rounds are commonly known.

Exercise 10 Compute the strictly and weakly dominated strategies in the Prisoner's Dilemma game played twice. □

8.2 Infinitely repeated games

In this section we consider infinitely repeated games. To define them we need to modify appropriately the approach of the previous section.

First, to ensure that the payoffs are well defined we assume that in the underlying stage game the payoff functions are bounded (from above and below). Then we redefine the set of ***histories*** by putting

$$\mathcal{H} := \bigcup_{t=0}^{\infty} \mathcal{H}^t,$$

where each $\mathcal{H}^t$ is defined as before.

The notion of a ***strategy*** of a player remains the same: it is a function from the set of all histories to the set of his strategies in the stage game. An ***outcome*** corresponding to a joint strategy σ is now the infinite set of joint strategies of the stage game $o^1(\sigma), o^2(\sigma), \ldots$ where each $o^t(\sigma)$ is defined as before.

Finally, to define the ***payoff function*** we first introduce a ***discount***, which is a number $\delta \in (0,1)$. Then we put

$$P_i(\sigma) := (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} p_i(o^t(\sigma)).$$

This definition requires some explanations. First note that this payoff function is well-defined and always yields a finite value. Indeed, the original payoff functions are assumed to be bounded and $\delta \in (0,1)$, so the sequence $(\sum_{t=1}^{t} \delta^{t-1} p_i(o^t(\sigma)))_{t=1,2,\ldots}$ converges.

Note that the payoff in each round t is discounted by δ^{t-1}, which can be viewed as the accumulated depreciation. So discounted payoffs in each round are summed up and subsequently multiplied by the factor $1-\delta$. Note that

$$\sum_{t=1}^{\infty} \delta^{t-1} = 1 + \delta \sum_{t=1}^{\infty} \delta^{t-1},$$

hence

$$\sum_{t=1}^{\infty} \delta^{t-1} = \frac{1}{1-\delta}.$$

So if in each round the players select the same joint strategy s, then their respective payoffs in the stage game and the repeated game coincide. This explains the adjustment factor $1-\delta$ in the definition of the payoff functions. Further, since the payoffs in the stage game are bounded, the payoffs in the repeated game are finite.

Given a stage game G and a discount δ we denote the infinitely repeated game defined above by $G(\delta)$.

We observed in the previous section that in each Nash equilibrium of the finitely repeated Prisoner's Dilemma game the players select in each round the defect (D) strategy. So finite repetition does not allow us to induce cooperation, i.e., the selection of the C strategy. We now show that in the infinitely repeated game the situation dramatically changes. Namely, the following holds.

Theorem 25 (Prisoner's Dilemma) *Take as G the Prisoner's Dilemma game. Then for all $\delta \in (\frac{1}{2}, 1)$ the pair of trigger strategies forms a Nash equilibrium of $G(\delta)$.*

Note that the outcome corresponding to the pair of trigger strategies consists of the infinite sequence of (C, C), that is, in the claimed Nash equilibrium of $G(\delta)$ both players repeatedly select C, i.e., always cooperate.

Proof. Suppose that, say, the first player deviates from his trigger strategy while the other player remains at his trigger strategy. Let t be the first stage in which the first player selects D. Consider now his payoffs in the consecutive rounds of the stage game:

- in the rounds $1, \ldots, t-1$ they equal 2,
- in the round t it equals 3,
- in the rounds $t+1, \ldots,$ they equal at most 1.

So the payoff in the repeated game is bounded from above by

$$\begin{array}{rl} & (1-\delta)(2\sum_{j=1}^{t-1} \delta^{j-1} + 3\delta^{t-1} + \sum_{j=t+1}^{\infty} \delta^{j-1}) \\ = & (1-\delta)(2\frac{1-\delta^{t-1}}{1-\delta} + 3\delta^{t-1} + \frac{\delta^t}{1-\delta}) \\ = & 2(1-\delta^{t-1}) + 3\delta^{t-1}(1-\delta) + \delta^t \\ = & 2 + \delta^{t-1} - 2\delta^t. \end{array}$$

Since $\delta > 0$, we have

$$\delta^{t-1} - 2\delta^t < 0 \text{ iff } 1 - 2\delta < 0 \text{ iff } \delta > \tfrac{1}{2}.$$

So when the first player deviates from his trigger strategy and $\delta > \frac{1}{2}$, his payoff in the repeated game is less than 2. In contrast, when he remains at the trigger strategy, his payoff is 2.

This concludes the proof. $\square$

This theorem shows that cooperation can be achieved by repeated interaction, so it seems to carry a positive message. However, repeated selection of the defect strategy D by both players still remains a Nash equilibrium and there is an obvious coordination problem between these two Nash equilibria.

Moreover, the above result is a special case of a much more general theorem. To formulate it we shall use the $minmax_i$ value introduced in Section 5.2 that, given a game $(S_1, \ldots, S_n, p_1, \ldots, p_n)$ was defined by

$$minmax_i := \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} p_i(s_i, s_{-i}).$$

The following result is called Folk theorem since some version of it has been known before it was recorded in a journal paper. From now on we abbreviate $(p_1(s), \ldots, p_n(s))$ to $p(s)$ and similarly with the P_i payoff functions.

Theorem 26 (Folk Theorem) *Consider a stage game $G := (S_1, \ldots, S_n, p_1, \ldots, p_n)$ with the bounded payoff functions.*

Take some $s' \in S$ and suppose $r := p(s')$ is such that for $i \in \{1, \ldots, n\}$ we have $r_i > minmax_i$. Then $\delta_0 \in (0, 1)$ exists such that for all $\delta \in (\delta_0, 1)$ the repeated game $G(\delta)$ has a Nash equilibrium σ with $P(\sigma) = r$.

Note that this theorem is indeed a generalization of the Prisoner's Dilemma Theorem 25 since for the Prisoner's Dilemma game we have $minmax_1 = minmax_2 = 1$, while for the joint strategy (C, C) the payoff to each player is 2. Now, the only way to achieve this payoff for both players in the repeated game is by repeatedly selecting C.

Proof. The argument is analogous to the one we used in the proof of the Prisoner's Dilemma Theorem 25. Let the strategy σ_i consist of selecting in each round s'_i. Note that $P(\sigma) = r$.

We first define an analogue of the trigger strategy. Let s^*_{-i} be such that $\max_{s_i} p_i(s_i, s^*_{-i}) = minmax_i$. That is, s^*_{-i} is the joint strategy of the opponents of player i that when selected by them results in a minimum possible payoff to player i. The idea behind the strategies defined below is that the opponents of the deviating player jointly switch forever to s^*_{-i} to 'inflict' on player i the maximum 'penalty'.

Recall that a history h is a finite sequence of joint strategies in the stage game. Below a *deviation in* h refers to the fact that a specific player i did not select s'_i in a joint strategy from h.

Given $h \in \mathcal{H}$ and $j \in \{1, \ldots, n\}$ we put

$$\sigma_j(h) := \begin{cases} s'_j & \text{if no player } i \neq j \text{ deviated in } h \text{ from } s'_i \text{ unilaterally} \\ s^*_j & \text{otherwise, where } i \text{ is the first player who deviated in } h \text{ from} \\ & s'_i \text{ unilaterally} \end{cases}$$

We now claim that σ is a Nash equilibrium for appropriate δs. Suppose that some player i deviates from his strategy σ_i while the other players remain at σ_{-i}. Let t be the first stage in which player i selects a strategy s''_i different from s'_i. Consider now his payoffs in the consecutive rounds of the stage game:

- in the rounds $1, \ldots, t-1$ they equal r_i,
- in the round t it equals $p_i(s''_i, s'_{-i})$,
- in the rounds $t+1, \ldots,$ they equal at most $minmax_i$.

Let $r^*_i > p_i(s)$ for all $s \in S$. The payoff of player i in the repeated game $G(\delta)$ is bounded from above by

$$\begin{array}{rl} & (1-\delta)(r_i \sum_{j=1}^{t-1} \delta^{j-1} + r^*_i \delta^{t-1} + minmax_i \sum_{j=t+1}^{\infty} \delta^{j-1}) \\ = & (1-\delta)(r_i \frac{1-\delta^{t-1}}{1-\delta} + r^*_i \delta^{t-1} + minmax_i \frac{\delta^t}{1-\delta}) \\ = & r_i - \delta^{t-1} r_i + (1-\delta)\delta^{t-1} r^*_i + \delta^t minmax_i \\ = & r_i + \delta^{t-1}(-r_i + (1-\delta) r^*_i + \delta\, minmax_i). \end{array}$$

Since $\delta > 0$ and $r^*_i \geq r_i > minmax_i$, we have

$$\begin{array}{rl} & \delta^{t-1}(-r_i + (1-\delta) r^*_i + \delta\, minmax_i) < 0 \\ \text{iff} & r^*_i - r_i - \delta(r^*_i - minmax_i) < 0 \\ \text{iff} & \frac{r^*_i - r_i}{r^*_i - minmax_i} < \delta. \end{array}$$

But $r^*_i > r_i > minmax_i$ implies that $\delta_0 := \frac{r^*_i - r_i}{r^*_i - minmax_i} \in (0,1)$. So when $\delta > \delta_0$ and player i selects in some round a strategy different than s'_i, while every other player j keeps selecting s'_j, player's i payoff in the repeated game is less than r_i. In contrast, when he remains selecting s'_i his payoff is r_i.

So σ is indeed a Nash equilibrium. $\square$

The above result can be strengthened to a much larger set of payoffs. Recall that a set of points $A \subseteq \mathbb{R}^n$ is called ***convex*** if for any $\mathbf{x}, \mathbf{y} \in A$ and $\alpha \in [0, 1]$ we have $\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in A$. Given a subset $A \subseteq \mathbb{R}^k$ denote the smallest convex set that contains A by $conv(A)$.

Then the above theorem holds not only for $r \in \{p(s) \mid s \in S\}$, but also for all $r \in conv(\{p(s) \mid s \in S\})$. In the case of the Prisoner's Dilemma game G we get that for *any*

$$r \in conv(\{(2, 2), (3, 0), (0, 3), (1, 1)\}) \cap \{r' \mid r'_1 > 1, r'_2 > 1\}$$

there is $\delta_0 \in (0, 1)$ such that for all $\delta \in (\delta_0, 1)$ the repeated game $G(\delta)$ has a Nash equilibrium σ with $P(\sigma) = r$. In other words, cooperation can be achieved in a Nash equilibrium, but equally well many other outcomes.

Such results belong to a class of similar theorems collectively called Folks theorems. The considered variations allow for different sets of payoffs achievable in an equilibrium, different ways of computing the payoff, different forms of equilibria, and different types of repeated games.