Chapter 12

Repeated Games

In real life, most games are played within a larger context, and actions in a given situation affect not only the present situation but also the future situations that may arise. When a player acts in a given situation, she takes into account not only the implications of her actions for the current situation but also their implications for the future. If an analyst does not take the broader context into consideration, her analysis may be misleading.

In particular, the players often have repeated interactions: firms compete with each other day after day, adjusting their productions and prices in response to what happened in the past; many countries coexist for centuries, choosing policies that affect each other (e.g. trade and environmental policies); coworkers, employees and employers interact repeatedly in workplaces, and families need to make household decisions repeatedly where family members have differing priorities, and so on. In such repeated interactions, players' actions have implications to future, as they may affect the other players' future behavior. A price cut by a firm can trigger a price war; a short-sighted tariff can lead to a long-lasting trade war, or a nice gesture in a work-place may win over fellow co-workers, and so on. Players then have long-term and short-term incentives, where the long term incentives are determined by how other players are expected to react to the player's action (endogenously), while the short-term incentives are determined by the payoff structure of the underlying game that is played at the moment. If the players are patient and the current actions have significant implications for the future, then the considerations about the future may take over. This may lead to a rich set of behavior that may seem to be irrational when one considers the current situation alone.

Repeated games epitomize the repeated interactions and the issues that come with such interactions. In these game, players play a fixed "stage-game", such as the Prisoners' Dilemma game, repeatedly, adjusting their stage-game strategies as they observe what players play. The stage game is repeated regardless of what has been played in the previous repetitions. The players are forward-looking, in that their overall payoffs depend on the stage-game payoffs they will get in all future interactions. This chapter explores the basic ideas in the theory of repeated games and applies them in a variety of economic problems.

As it turns out, it is important whether the game is repeated finitely or infinitely many times. Imagine that there is a fixed deadline at which the game ends exogenously. If the stage game has a unique subgame-perfect Nash equilibrium, the players will not have any long-term incentive and play the game as if there is not future to consider, regardless of how far the deadline is. Indeed, in the last period, there is no future, and the past actions have been played already. Hence, the players play the stage-game as if it is an isolated game, for which there is a unique solution. Thus, in the last period players play according to the unique solution regardless of what they have played in the past. Now, the day before the last day, the players foresee that they will play according to the unique solution in the next period regardless of what they do in this period. Thus, although there is future interaction, they cannot influence it by their actions today. They cannot influence the past either. Therefore, they play according to their short-term incentive, playing the unique solution of the stage game. By repeating this argument repeatedly, one sees that the players can never influence the future and the players do not have any long-term incentives.

This backward-induction logic breaks down when there is no known deadline. At any given period, there is considerable future interactions to take into account, and one cannot start the backward induction argument. In that case, when the players are sufficiently patient, there may be subgame-perfect Nash equilibria in which the players react harshly for some "unwanted" behavior, and the players avoid such behavior fearing that those actions will trigger a harsh reaction. For example, in the Prisoners' Dilemma game, players may cooperate because they expect that the other player will cooperate until somebody deviates and switch to defecting forever as soon as somebody defects. Similarly, in a repeated Bertrand oligopoly, the firms may charge high prices—as if they are in an implicit cartel—fearing that any price cut will trigger them to charge competitive prices. Or more counterintuitively, in a Cournot oligopoly, the firms may flood the market with goods and charge very low prices, hoping that this will lead them to charge high prices later, where they are charging high prices in order to avoid another episode of low prices.

In fact, the main theorem of this chapter will be what is known as the Folk Theorem: in an infinitely repeated game, every "individually rational" and feasible payoff vector is achieved by some subgame-perfect Nash equilibrium when players are sufficiently patient. This suggests that when players are in a repeated interaction, the analysis of the stage game in isolation may be misleading. One should rather analyze carefully players' concerns about the future interactions, the same way one analyzes players' beliefs about the underlying payoffs in a given game. This does not mean that anything goes. The subgame-perfect Nash equilibria will have considerable structure, and equilibrium analysis will still shed light to how players will behave and react to various actions of other players.

12.1 Finitely-repeated games

This section illustrates the formulation of the repeated games and some of the main issues focusing on the repeated interactions for which there is a known deadline. After describing the formulation in detail, it illustrates the backward-induction logic above in detail on finite repetition of the Prisoners' Dilemma game and provides an application of this idea to entry-deterrence problem, an important problem in Industrial Organization. When the underlying game has multiple equilibria, one can construct a large set of subgame-perfect Nash equilibria for the repeated game. In these equilibria, players may take non-equilibrium actions in early rounds of repetition as their actions in the earlier rounds determine which equilibrium of the stage game will be played in later rounds.

12.1.1 Formulation

There are $T+1$ periods: $t = 0, 1, \ldots, T$. In each period t, players play a "stage game" G, observing what happened in the previous periods up to that point. At the end of period T, the game ends and each player gets sum of the stage-game payoffs she has

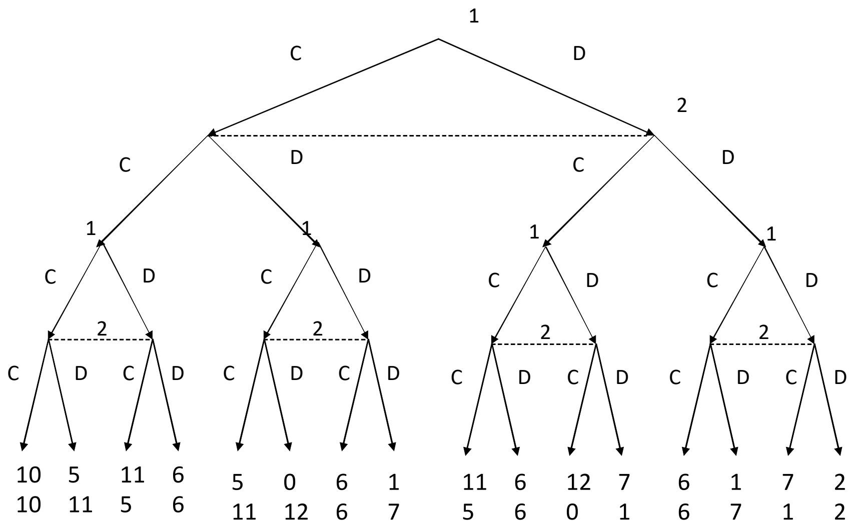


Figure 12.1: Twice-repeated Prisoners' Dilemma game

accumulated. The larger multi-stage game is called a finitely repeated game and denoted by G^T.

For a concrete example, suppose that there are two periods, $t = 0, 1$, and the stage-game G is the Prisoners' Dilemma game

$$\begin{array}{c|c|c|} & C & D \\ \hline C & 5,5 & 0,6 \\ \hline D & 6,0 & 1,1 \\ \hline \end{array} \tag{12.1}$$

The repeated game, G^T, is the extensive-form game depicted in Figure 12.1. After they play the Prisoners' Dilemma game at period 0, each player observes what each of them played and play the Prisoners' Dilemma game one more time (at period 1). They could have played the Prisoners' Dilemma game in four different ways in the initial period: (C, C), (C, D), (D, C), and (D, D). For each of these possible plays, there is a subgame that represents the game in the final period following the initial play. In the figure, these are the smaller subgames that look like the extensive-form representation of the Prisoners' dilemma game. For example, the subgame in Figure 12.2 follows the play of (C, C) in the initial period. In this subgame, one adds 5 to each players' payoff from Prisoners' Dilemma game, accounting for the payoff each player has gotten from both cooperating in the initial period. As in this example, the payoff of each player at the end

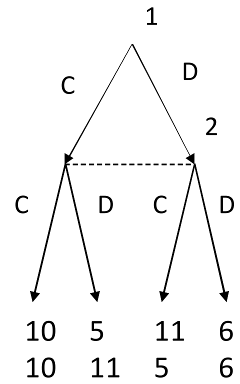


Figure 12.2: A subgame of the twice-repeated Prisoners' Dilemma game

is computed by adding the payoffs that she would get in the two stages. For example, the payoff vector associated with play $(C, D), (C, C)$ is $(5, 11)$, which is computed by adding the payoff vector $(0, 6)$ associated with the first-period play (C, D) and the payoff vector $(5, 5)$ associated with the second-period play (C, C).

At the beginning of period $t = 1$, a player can remember whether the period 0 play was (C, C), (C, D), (D, C), and (D, D). Hence, a strategy for a player prescribes what she plays in period 0 and also what she plays for each of these four contingencies in period 1. Accordingly, in the extensive-form game in Figure 12.1, each player has five information sets, one for the initial period, and one of each of the four contingencies above. A strategy picks a move for each of these five information sets.

The stage game can also be a dynamic game. For example, stage game can be the "Entry-Deterrence" game in Figure 12.3. Twice-repeated game with this stage game is depicted in Figure 12.4. In this game, the period 0 play can be X (i.e. Player 1 plays X), EA (i.e., Player 1 plays E and Player 2 plays A) and EF (i.e., Player 1 plays E and Player 2 plays F). Accordingly, there are three smaller subgames, each corresponding one of these contingencies. In particular, if Player 1 plays E, players would not know what Player 2 would have done if Player 1 played E.

Formally, in a repeated game at the beginning of each period t, each player recalls a sequence $(a_0, \ldots, a_{t-1})$ that describes the outcome of the play in each period up to that period. When the stage-game is a simultaneous action game (as in the Prisoners' Dilemma game), each of these outcomes is a strategy profile of the stage game. When the stage game is dynamic (as in the Entry-Deterrence game), each outcome here is a terminal history of the stage game, describing what happened within the stage. Each such sequence $(a_0, \ldots, a_{t-1})$ is called a *history*—or a period-t history for clarity. A strategy in the repeated game prescribes a strategy of the stage game for *each* history $(a_0, \ldots, a_{t-1})$ at *each* date t. For example, in the twice-repeated Prisoners' Dilemma game, a possible strategy is: play C at the beginning and mimic the other player's initial move in the second period, playing C if she played C and D if she played D.

Exercise 12.1. How many strategies are there in twice-repeated prisoners dilemma game?

Exercise 12.2. Suppose that the stage game is a two-player games in which each player i has m_i strategies. How many strategies each player has in an T-times repeated game?

12.1.2 Backward-Induction Logic

When the stage game has a unique subgame-perfect Nash equilibrium, the repeated game also has a unique subgame-perfect Nash equilibrium in which the subgame-perfect Nash equilibrium of the stage game is played in each period regardless of the history.

For an illustration, consider the twice-repeated Prisoners' Dilemma game. As discussed above, this game has four proper subgames, each corresponding to the last-round game after a history of plays in the initial round. As in the subgame in Figure 12.2, in each of these subgames, a constant has been added to each player's Prisoners' Dilemma payoff from the final period, accounting for the payoff accrued in the initial period. But adding a constant to a player's payoff does not change the preferences in a game. Hence, the set of equilibria in each of these subgames is the same as the set of Nash equilibria of the original Prisoners' Dilemma game. Since the Prisoners' dilemma game has a unique Nash equilibrium of (D, D), each of these subgames has a unique Nash equilibrium: (D, D). This can be seen vividly in Figure 12.2 for the subgame after both cooperate.

Therefore, the actions in the last round are independent of what is played in the initial round. Hence, the players will ignore the future and play the game as if there is no future game, each playing D. Indeed, given the behavior in the last round, the game in the initial round reduces to

	C	D
C	$6, 6$	$1, 7$
D	$7, 1$	$2, 2$

where 1 is added to each payoff, accounting for the payoff from (D, D) in the last round. The unique equilibrium of this reduced game is (D, D). Therefore, the repeated game has a unique subgame-perfect Nash equilibrium: *at each history, each player plays D.*

The analysis remains the same for arbitrary T. In the last day, T, independent of what has been played in the previous rounds, there is a unique Nash equilibrium for the resulting subgame: each player plays D. Hence, the actions at day $T - 1$ do not have any effect in what will be played in the next day. Then, we can consider the subgame as a separate game of the Prisoners' Dilemma. Indeed, the augmented game for any subgame starting at $T - 1$ is

	C	D
C	$5 + 1 + \pi_1, 5 + 1 + \pi_2$	$0 + 1 + \pi_1, 6 + 1 + \pi_2$
D	$6 + 1 + \pi_1, 0 + 1 + \pi_2$	$1 + 1 + \pi_1, 1 + 1 + \pi_2$

where π_1 is the sum of the payoffs of i from the previous plays at dates $0, \ldots, T-2$. Here one adds π_i for these payoffs and 1 for the last round payoff. Both the payoff π_i from past repetition and the payoff 1 from the future repetition are independent of what happens at date $T - 1$. A player cannot influence the past as it has passed already and cannot influence the future because rationality in the future imposes a particular play that is independent of the actions at period $T - 1$. This is another version of the Prisoner's Dilemma game, which has the unique Nash equilibrium of (D, D). Proceeding in this way all the way back to date 0, one finds out that there is a unique subgame-perfect Nash equilibrium: *at each t and for each history of previous plays, each player plays D.*

That is to say, although there are many repetitions in the game and the stakes in the future may be high, any plan of actions other than playing myopically D everywhere unravels, as players cannot commit to any plan of action in the last round. This is indeed a general result.

Theorem 12.1. *For any finite T and any stage game G with a unique subgame-perfect Nash equilibrium s^*, the repeated game G^T has a unique subgame-perfect Nash equi-*

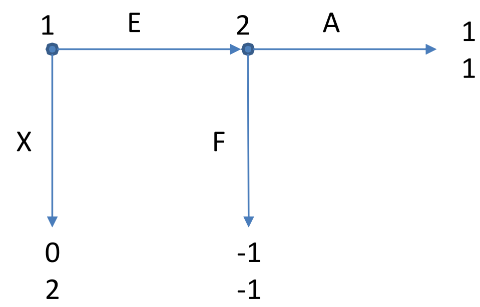


Figure 12.3: Entry-Deterrence game

librium, and according to this equilibrium s^* is played at each date independent of the history of the previous plays.

Exercise 12.3. Prove Theorem 12.1.

12.1.3 Chain-Store Paradox

The next example provides another illustration of this general fact in an important application in Industrial Organization. An incumbent firm may be enjoying high profits that may entice some potential entrepreneurs to enter the market and compete with the incumbent firm. When there is an entry, the incumbent firm faces a dilemma. It can either engage a costly fight with the entrant, for example by upgrading its product at a high cost, in the hopes that this will force the entrant exit the market. Alternatively, it can accommodate the entrant, for example by keeping the existing product. Accommodation is more profitable than fighting, but it would be best for the incumbent if there were no entry to start with.

This situation is modeled abstractly by the Entry-Deterrence game in Figure 12.3. In this game Player 1 is a potential entrant, and Player 2 is a monopoly (incumbent). Player 1 either chooses exit (X), in which case the incumbent firm maintains its monopoly, or enters the market (E). If it enters the market, the incumbent firm decides whether to fight the entry (F) or to accommodate it (A). Entry is profitable for the entrant only if the incumbent accommodates the entry, and although incumbent prefers that the entrant does not enter, it is costly to fight entry, and it would rather accommodate when there is entry. Backward induction leads to a unique solution. When there is an

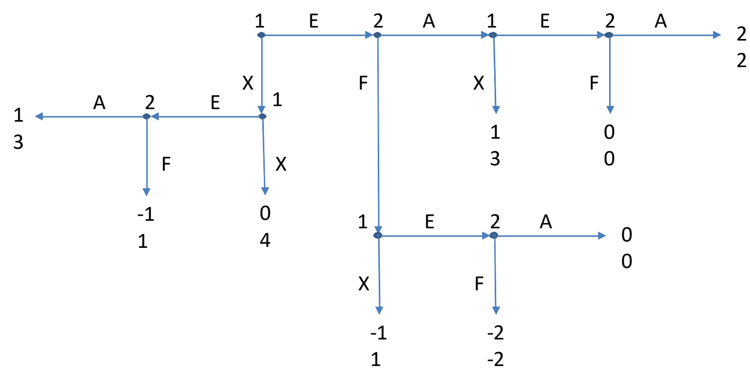


Figure 12.4: Twice repeated Entry-Deterrence Game

entry, the incumbent firm accommodates the entry (as fighting is costly), and foreseeing this Player 1 enters.

Now imagine that incumbent firm is a reputable company that faces with potential entrants year after year, and the potential entrants see what happened in the past when somebody entered the market.[1] Assuming that each potential entrant lives for only one period, represent them as one single entrant deciding whether to enter or stay out each period. The game is a finitely repeated game with the Entry-Deterrence game in Figure 12.3 as its stage game. For example, twice repeated version of this game is as in Figure 12.4. The repeated game is a perfect-information game with finite horizon, and hence one can apply backward induction. In the last period, if the entrant enters the market, the incumbent accommodates. Foreseeing this, the entrant enters. Then, in the next to last period, the incumbent again accommodates the entrant, as fighting would not deter the entry in the next period, and once again the entrant foresees this and entrants. Repeating this argument inductively, one concludes that the entrant enters and the incumbent accommodates at every history.

This is found paradoxical because one would have thought that, as a large reputable company, the incumbent would have tried to form a reputation for being tough on entry by fighting entries early on in order to deter future entries. The above backward-

[1] An alternative interpretation is that the incumbent is a giant company that operates in many markets and faces with potential entrants in those markets, where the entrants in those markets decide whether to enter the market according to a given order. This is what gives the name "chain-store paradox".

induction argument shows that this plan will unravel because the incumbent will accommodate the entry in the last period even if it fought entry all the way. Its threat to fight an entry is not credible, so to speak.

The rest of this chapter will show that these are very peculiar examples. In general, in many subgame-perfect equilibria, the patient players will take a long-term view, and their decisions will be determined mainly by the future considerations.

12.1.4 Stage Games with Multiple Equilibria

If the stage game has more than one equilibrium, then the equilibrium played in the last round may depend on how the game is played in earlier rounds. Then, in earlier rounds, the players may take actions that are not played in any subgame-perfect Nash equilibrium of the stage game.

For example, consider the stage game[2]

	L_2	M_2	R_2
L_1	$1,1$	$5,0$	$0,0$
M_1	$0,5$	$4,4$	$0,0$
R_1	$0,0$	$0,0$	$3,3$

In this game, strategy M_i is strictly dominated and cannot be played rationally—although (M_1, M_2) yields a high payoff to each player. There are two pure strategy Nash equilibria, (L_1, L_2) and (R_1, R_2), and there is also a mixed strategy Nash equilibrium in which each player i plays L_i and R_i with probabilities $3/4$ and $1/4$, respectively.

Now consider the repeated game in which this stage game is repeated twice. A repeated-game strategy prescribes what the player plays at $t = 0$ and what she plays at $t = 1$ conditional on the history of the play at $t = 0$. There are 10 actions in total, one for $t = 0$ and 9 for $t = 1$. Consider the following strategy profile: for each player i,

> play M_i at $t = 0$; at $t = 1$, play R_i if (M_1, M_2) played at $t = 0$, and play L_i otherwise.

Acoording to this strategy profile, at $t = 0$, the players play (M_1, M_2) even though (M_1, M_2) is not a Nash equilibrium of the stage game. This is a subgame-perfect Nash

[2] This stage game is taken from Gibbons, and the discussion builds on the deiscussion therein.

equilibrium. Indeed, in period $t = 1$, they play a stage-game Nash equilibrium: they play (L_1, L_2) or (R_1, R_2), depending on the period 0 play. Moreover, given the behavior in period $t = 1$, the game in period $t = 0$ reduces to

	L_2	M_2	R_2
L_1	$2,2$	$6,1$	$1,1$
M_1	$1,6$	$7,7$	$1,1$
R_1	$1,1$	$1,1$	$4,4$

Here, a payoff of 3 is added to the stage-game payoffs at (M_1, M_2) (accounting for the payoff from (R_1, R_2) in the second round) and a payoff of 1 is added to the payoffs at the other strategy profiles (accounting for the payoff from (L_1, L_2) in the second round). Clearly, (M_1, M_2) is a Nash equilibrium in the reduced game, showing that the above strategy profile is a subgame-perfect Nash equilibrium.

In summary, players can coordinate on different stage-game Nash equilibria in the second round conditional on the behavior in the first round. Then, in the first round, anticipating that their current actions will affect the future play, the players may play a strictly dominated stage-game strategy, if that strategy leads to a better equilibrium later.

When there are multiple subgame-perfect Nash equilibria in the stage game, a large number of outcome paths can result in a subgame-perfect Nash equilibrium of the repeated game even if it is repeated just twice. But not all outcome paths can be a result of a subgame-perfect Nash equilibrium. In the following, I will illustrate why some of the paths can and some paths cannot emerge in an equilibrium in the above example.

Can $((M_1, M_2), (M_1, M_2))$ be an outcome of a subgame-perfect Nash equilibrium? The answer is No. This is because in any Nash equilibrium, the players must play a Nash equilibrium of the stage game in the last period on the path of equilibrium. Since (M_1, M_2) is not a Nash equilibrium of the stage game, the outcome $((M_1, M_2), (M_1, M_2))$ cannot emerge in any Nash equilibrium, let alone in a subgame-perfect Nash equilibrium.

Can $((M_1, M_2), (L_1, L_2))$ be an outcome of a subgame-perfect Nash equilibrium in pure strategies? The answer is No. Although (L_1, L_2) is a Nash equilibrium of the stage game, in any subgame-perfect Nash equilibrium, a Nash equilibrium of the stage game must be played after every play in the first round. In particular, after (L_1, M_2), the play is either (L_1, L_2) or (R_1, R_2), yielding 6 or 8, respectively, for Player 1. Since she gets only 5 from $((M_1, M_2), (L_1, L_2))$, she has an incentive to deviate to L_1 in the first period. Intuitively, since they play the worst possible pure strategy equilibrium after (M_1, M_2), there is no future loss due to any deviation, and hence players have incentive to deviate and realize the short-term gain of 1.

Can $((M_1, L_2), (R_1, R_2))$ be an outcome of a subgame-perfect Nash equilibrium in pure strategies? As it must be clear from the previous discussion the answer would be Yes if and only if (L_1, L_2) is played after every play of the period except for (M_1, L_2). In that case, the game in the first period reduces to

	L_2	M_2	R_2
L_1	$2,2$	$6,1$	$1,1$
M_1	$3,8$	$5,5$	$1,1$
R_1	$1,1$	$1,1$	$4,4$

Since (M_1, L_2) is indeed a Nash equilibrium of the reduced game, the answer is Yes. It is the outcome of the following subgame-perfect Nash equilibrium: Play (M_1, L_2) in the first round; in the second round, play (R_1, R_2) if (M_1, L_2) is played in the first round and play (L_1, L_2) otherwise. Intuitively, a deviation in the first round costs a payoff of 2 in the future round, ofseeting the short-term gain of 1 in the first period.

As the game gets longer, more an more stage-game strategy profiles can be supported in earlier rounds by a subgame-perfect Nash equilibrium of the repeated game. For example, (R_1, L_2) cannot be played in any subgame-perfect Nash equilibrium of the twice repeated game. This is because Player 2 can gain 3 in the first period and this short-term gain exceeds any possible future cost, which is bounded by 2, the payoff difference between (R_1, R_2) and (L_1, L_2).[3] However, (R_1, L_2) can be played in period 0 in a subgame-perfect Nash equilibrium of the thrice repeated game. Indeed, one can construct an equilibrium with outcome path $((R_1, L_2), (M_1, M_2), (R_1, R_2))$, by playing (L_1, L_2) for the rest of the repeated game whenever they deviate from this path.

[3] The future cost can be as high as 2.25 if one considers the mixed strategy equilibria; this is the payoff difference between (R_1, R_2) and the mixed-strategy equilibrium. Once again, the short-term gain exceeds the future cost.

12.2 Infinitely repeated games

In most repeated interactions in real world, there is no set deadline at which the interactions end. We will all be dead in the long run, and all interactions will come to end eventually. But the time at which the interaction ends is usually not known by the players when they play the game. At any given period, there is some possibility that the game will continue one more period. Such repeated interactions are modeled by infinitely repeated games, formally described as follows.

There are infinitely many periods: $t = 0, 1, \ldots$. In each period t, players play a stage game G, observing what happened in the previous periods up to that point. The payoff of a player in the repeated game is the discounted sum of the payoffs that she gets in stage games throughout; the discounted sum will be defined formally momentarily.

It is implicitly assumed throughout the chapter that in the stage game G, either the strategy sets are all finite, or the strategy sets are convex subsets of $\mathbb{R}^n$ and the utility functions are continuous in all strategies and quasiconcave in players' own strategies.

Present Value calculations In an infinitely repeated game, one cannot simply add the stage-game payoffs, as the sum becomes infinite. For these games, the payoff of a player in the repeated game is the discounted sum of the payoffs she gets from the stage games. The (discounted) *present value* of any given payoff stream $\pi = (\pi_0, \pi_1, \ldots, \pi_t, \ldots)$ is computed by

$$PV(\pi; \delta) = \sum_{t=0}^{\infty} \delta^t \pi_t = \pi_0 + \delta\pi_1 + \cdots + \delta^t \pi_t + \cdots,$$

where $\delta \in (0, 1)$ is the *discount factor*.

For example, in the repeated Prisoners' Dilemma game with stage payoffs in (12.1), imagine that the players cooperate throughout, so that the path is $(C, C), (C, C), \ldots$, yielding stage-game payoff of 5 to each player at each day. For each player, the present value of this payoff stream is

$$5 + 5\delta + \cdots + 5\delta^t + \cdots = \frac{5}{1 - \delta},$$

where one uses the well-known formula for geometric series to compute the sum $1 + \delta + \cdots + \delta^t + \cdots = 1/(1 - \delta)$. Similarly, if they both cooperate at periods $t = 0, \ldots, T$ and defect at periods $t = T+1, T+2, \ldots$, then each player will get 5 in first $T+1$ periods and 1 thereafter, and the present value of this will be

$$5 + 5\delta + \cdots + 5\delta^T + \delta^{T+1} + \delta^{T+2} \cdots = \frac{5\left(1-\delta^{T+1}\right) + \delta^{T+1}}{1-\delta},$$

where one computes the geometric sums as in the previous case.

As in these examples, one often gets $1-\delta$ in the denominator in present-value calculations, and it is often more convenient to normalize the payoffs by multiplying it with $(1-\delta)$. The normalization yields the *average (discounted) value*:

$$(1-\delta)\, PV(\pi;\delta) \equiv (1-\delta) \sum_{t=0}^{\infty} \delta^t \pi_t,$$

the present value multiplied by $(1-\delta)$. Note that, for a constant payoff stream (i.e., $\pi_0 = \pi_1 = \cdots = \pi_t = \cdots$), the average value is simply the stage payoff (namely, π_0). In the above example, if players cooperate throughout, the average value of each player will be simply 5. On the other hand, if they cooperate at periods $t = 0, \ldots, T$ and defect at periods $t = T+1, T+2, \ldots$, then the average value of each player will be

$$5\left(1-\delta^{T+1}\right) + \delta^{T+1},$$

a convex combination of the average value 5 for the first $T+1$ periods and 1 for the rest, where the weight of 1 is simply the discount factor that corresponds to the period at which the stream of low payoffs starts.

The present and the average values can be computed with respect to the current date. That is, given any t, *the present value at t* is

$$PV_t(\pi;\delta) = \sum_{s=t}^{\infty} \delta^{s-t} \pi_s = \pi_t + \delta\pi_{t+1} + \cdots + \delta^k \pi_{t+k} + \cdots,$$

and the *average value at t* is $(1-\delta)\, PV_t(\pi;\delta)$. Clearly,

$$PV(\pi;\delta) = \pi_0 + \delta\pi_1 + \cdots + \delta^{t-1}\pi_{t-1} + \delta^t PV_t(\pi;\delta).$$

Hence, the analysis does not change whether one uses PV or PV_t, but using PV_t is simpler. In repeated games considered here, each player maximizes the present value of the payoff stream she gets from the stage games, which will be played indefinitely. Since the average value is simply a linear transformation of the present value, one can also use average values instead of present values. Such a choice sometimes simplifies the expressions without affecting the analyses.

Remark 12.1. There are at least two interpretations for infinitely repeated games with discounting. The first interpretation is that the game is repeated literally infinitely many times, and the players discount the payoffs in the future. A dollar earned tomorrow is not as valuable as a dollar earned today for these players. Typically, the interest rates in the real world are positive, and there are individuals who are willing to borrow money at these interest rates, paying larger sums in the future. Under this interpretation, the discount factor corresponds to the rate at which the players discount future payoffs. In a typical classroom experiment, such a discount factor should be 1 because the payoffs are paid at the end as a sum. There is a second interpretation, which is alluded to at the beginning of this section. Under this interpretation, the game ends stochastically. After each period t, the game ends with probability $1-p$ and each player gets a constant payoff thereafter. With probability p, they play one more round. Normalizing the payoff after termination to zero, one gets the present value formula for discount factor $\delta = p$. In a classroom experiment, one can reasonably assume that this is the discount factor. In real life, there can be a combination of time discounting—with discount factor $\bar{\delta}$—and termination—with probability p. In that case, the discount factor will be $\delta = \bar{\delta}p$.

Histories and strategies The histories and strategies are defined as in the finite-horizon case. A period-t *history* is a sequence of the outcomes of the play at dates $0, \ldots, t-1$. A period-t history is denoted by $h = (a_0, \ldots, a_{t-1})$, where $a_{t'}$ is the outcome of stage game in round t'; h is empty when $t = 0$. Examples of histories are $(X, X, EA, X, \ldots, X)$ in repeated Entry-Deterrence game and $((C, C), (C, D))$ in repeated Prisoners' Dilemma game.

A *strategy* in a repeated game is a mapping that assigns a stage-game strategy to each period-t history for *each* t. The key here is that the stage-game strategy can vary by histories. For a given strategy profile s in the repeated game, the *outcome path* is defined as the infinite history realized if all players players play according to their strategies in s. A history $h = (a_0, \ldots, a_{t-1})$ is *on the path* of a strategy profile if the play up to period t is $a_0, \ldots, a_{t-1}$ according to the strategy profile; it is *off the path* otherwise.[4]

For an illustration of these definitions, consider the following strategies in the re-

[4] This assumes that there is no chance move in the stage game. If there are chance moves in the stage game or the strategy profile is mixed, then the outcome is a probility distribution on paths, and a history is on the path if it occurs with positive probability.

peated Prisoner's Dilemma game (the first two strategies are very important while the last one is given for comparison):

Grim Trigger Play C at $t = 0$; thereafter play C if the players have always played (C, C) in the past, play D otherwise (i.e., if anyone ever played D in the past).

Tit-for-Tat Play C at $t = 0$, and at each $t > 0$, play whatever the other player played at $t - 1$.

Naively Cooperate Play always C (no matter what happened in the past).

Strategy profiles (Grim Trigger, Grim Trigger), (Tit-for-Tat, Tit-for-Tat), and (Naively Cooperate, Naively Cooperate) all lead to the same outcome path:

$$((C, C), (C, C), (C, C), \ldots).$$

For example according to (Grim Trigger, Grim Trigger), each player play C at the beginning. At $t = 1$, the history is (C, C) and they must play C according to the strategy profile. The history at the beginning of $t = 2$ is $((C, C), (C, C))$, and they must play C again according to the strategy profile. An outside observer, such as a journalist or a statistician, will see that the players keep cooperating throughout.

Nevertheless, these strategy profiles are quite distinct, each telling a distinct story, and each giving a different explanation for perpetual cooperation. The strategy profile (Grim Trigger, Grim Trigger) tells us that any deviation from this path will result in players switching to their myopic stage-game dominant strategy of D and playing it myopically forever. According to this story, the explanation for cooperation is the fear of social break-down that leads to perpetual defection. On the other hand, Tit-for-Tat tells us that players are reciprocating what the other players have done in the previous period, and the fear of reciprocation prevents them from defecting. Finally, Naive Cooperation simply states that players play cooperate regardless; a naive observer could have concluded this by extrapolating from their observation of perpetual cooperation. Here, (Grim Trigger, Grim Trigger) is a subgame-perfect Nash equilibrium for large values of δ; (Tit-for-Tat, Tit-for-Tat) is a Nash equilibrium for large values of δ, but it is not subgame-perfect, while (Naively Cooperate, Naively Cooperate) is clearly not a Nash equilibrium. All these will be clear momentarily.

12.2.1 One-Shot Deviation Principle

In infinitely repeated games, one uses the *One-Shot Deviation Principle* in order to check whether a strategy profile is a subgame-perfect Nash equilibrium. Here, One-Shot Deviation Principle takes a simple form, as the augmented stage game is the same as the stage game G but one simply augments the payoffs in the stage game by adding the present value of future payoffs under the purported equilibrium.

Augmented Stage Game Formally consider a strategy profile $s^* = (s_1^*, s_2^*, \ldots, s_n^*)$ in the repeated game. Consider any date t and any history $h = (a_0, \ldots, a_{t-1})$, where $a_{t'}$ is the outcome of the play at date t'. *Augmented stage game for s^* and h* is the same game as the stage game in the repeated game except that the payoff of each player i from each terminal history z of the stage game is

$$U_i(z|s^*, h) = u_i(z) + \delta PV_{i,t+1}(h, z, s^*)$$

where $u_i(z)$ is the stage-game payoff of player i at z in the original stage game, and $PV_{i,t+1}(h, z, s^*)$ is the present value of player i at $t+1$ from the payoff stream that results when all players follow s^* starting with the history $(h, z) = (a_0, \ldots, a_{t-1}, z)$, which is a history at the beginning of date $t+1$. Note that $U_i(z|s^*, h)$ is the time t present value of the payoff stream that results when the outcome of the stage game is z in round t and everybody sticks to the strategy profile s^* from the next period on. Note also that the only difference between the original stage game and the augmented stage game is that the payoff in the augmented game is $U_i(z|s^*, h)$ while the payoff in the original game is $u_i(z)$.

One-Shot Deviation Principle now states that a strategy profile in the repeated game is subgame-perfect if it always yields a subgame-perfect Nash equilibrium in the augmented stage game:

Theorem 12.2 (One-Shot Deviation Principle)**.** *Strategy profile s^* is a subgame-perfect Nash equilibrium of the repeated game if and only if $(s_1^*(h), \ldots, s_n^*(h))$ is a subgame-perfect Nash equilibrium of the augmented stage game for s^* and h for every date t and every history $h = (a_0, \ldots, a_{t-1})$.*

Note that $s_i^*(h)$ is what player i is supposed to play at the stage game after history h at date t according to s^*. Hence, $s_i^*(h)$ is a strategy in the stage game as well as a strategy in the augmented stage game. Therefore, $(s_1^*(h), \ldots, s_n^*(h))$ is a strategy profile in the augmented stage game, and a potential subgame-perfect Nash equilibrium. Note also that, in order to show that s^* *is* a subgame-perfect Nash equilibrium, one must check *for all* histories h and dates t that s^* yields a subgame-perfect Nash equilibrium in the augmented stage game. Conversely, in order to show that s^* is *not* a subgame-perfect Nash equilibrium, one only needs to find *one* history (and date) for which s^* does not yield a subgame-perfect Nash equilibrium in the augmented stage game. Finally, although the above result considers pure strategy profile s^* the same result is true for mixed strategies. The result is stated that way for clarity. The rest of this section is devoted to illustration of One-Shot Deviation Principle on infinitely repeated Entry Deterrence and Prisoners' Dilemma games, and step-by-step recipe for this task in general stage games.

Infinitely Repeated Entry Deterrence Towards illustrating the One-Shot Deviation Principle when the stage game is dynamic, consider the infinitely repeated Entry-Deterrence game, where the stage game as in Figure 12.3. Consider the following strategy profile.

Switching Strategy Profile The entrant enters the market if an only if the incumbent has accommodated the entrant sometimes in the past. The incumbent accommodates the entrant if an only if she has accommodated the entrant before.[5]

One applies the One-Shot Deviation Principle to show that this strategy profile is a subgame-perfect Nash equilibrium for large values of δ—as follows. The strategy profile puts the histories in two groups:

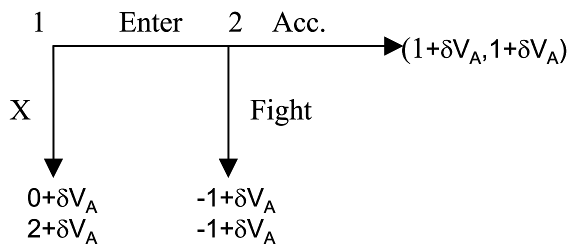
1. The histories at which there was an entry and the incumbent has accommodated; the histories that contain an entry EA, and
2. all the other histories, i.e., the histories that do not contain the entry EA at any date.

[5] This is a switching strategy, where initially incumbent fights whenever there is an entry and the entrant never enters. If the incumbent happens to accommodate an entrant, they switch to the new regime where the entrant enters the market no matter what the incumbent does after the switching, and incumbent always accommodates the entrant.

Consequently, in the application of One-Shot Deviation Principle, one puts histories in the above two groups, depending on whether the incumbent has ever accommodated any entrant. First take any date t and any history $h = (a_0, \ldots, a_{t-1})$ in the first group, where incumbent has accommodated some entrants. Now, independent of what happens at t, the histories at $t + 1$ and later will contain a past instance of accommodation EA (before t), and according to the strategy profile, at $t + 1$ and on, entrant will always enter and incumbent will accommodate, each player getting the constant stream of 1s. The present value of this at $t + 1$ is

$$V_A = 1 + \delta + \delta^2 + \cdots = 1/(1 - \delta).$$

That is, for every outcome $z \in \{X, EA, EF\}$, $PV_{i,t+1}(h, z, s^*) = V_A$. Hence, the augmented stage game for h and s^* is



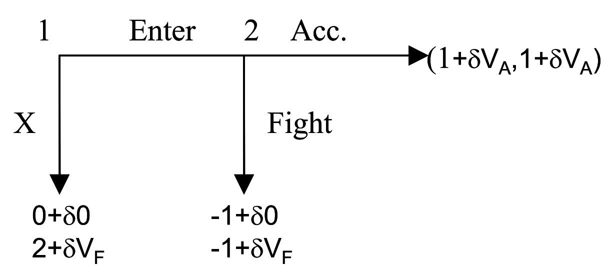
For example, if the incumbent accommodates the entrant at t, her present value (at t) will be $1 + \delta V_A$; and if she fights her present value will be $-1 + \delta V_A$, and so on. This is another version of the Entry-Deterrence game, where the constant δV_A is added to the payoffs. The strategy profile s^* yields (Enter, Accommodate) for round t at h. According to One-Shot Deviation Principle, (Enter, Accommodate) must be a subgame-perfect Nash equilibrium of the augmented stage game here. This is indeed the case, and s^* passes the one-shot deviation test for such histories.

Note that future behavior is independent of the actions taken today at the above histories. Consequently, the augmented game is equivalent to the stage game. Then, a strategy profile passes the One-Shot Deviation Principle at any such history if and only if the strategy profile prescribes a subgame-perfect Nash equilibrium of the stage game at that history. This is true in general.

Now, for an arbitrary date t, consider a history $h = (a_0, \ldots, a_{t-1})$ in the second group, where the incumbent has never accommodated the entrant before, i.e., $a_{t'}$ differs from EA for all t'. Towards constructing the augmented stage game for h, first consider the outcome $z = EA$ at t. In that case, at the beginning of $t+1$, the history is (h, EA), which includes EA as in the previous paragraph. Hence, according to s^*, Player 1 enters and Player 2 accommodates at $t+1$, yielding a history that contains EA for the next period. Therefore, in the continuation game, all histories are in the first group (containing EA), and the play is (Enter, Accommodate) at every $t' > t$, resulting in the outcome path $(h, EA, EA, \ldots)$. Starting from $t+1$, each player gets 1 for each date, resulting the present value of $PV_{i,t+1}(h, z, s^*) = V_A$. Now consider another outcome $z \in \{X, EF\}$ in period t. The continuation play for other outcomes is quite different now. At the beginning of $t+1$, the history (h, z) is either (h, X) or (h, EF). Since h does not contain EA, neither does (h, z). Hence, according to s^*, at $t+1$, Player 1 exits, and Player 2 would have chosen Fight if there were an entry, yielding outcome X for period $t+1$. Consequently, at any $t' > t+1$, the history is $(h, z, X, X, \ldots, X)$, and Player 1 chooses to exit at t' according to s^*. This results in the outcome path $(h, z, X, X, \ldots)$. Therefore, starting from $t+1$, Player 1 gets 0 and Player 2 gets 2 every day, yielding present values of $PV_{1,t+1}(h, z, s^*) = 0$ and

$$PV_{2,t+1}(h, z, s^*) = V_F = 2 + 2\delta + 2\delta^2 + \cdots = 2/(1-\delta),$$

respectively. Therefore, the augmented stage game for h and s^* is now



At this history the strategy profile prescribes $(X, Fight)$, i.e., the entrant does not enter, and if she enters, the incumbent fights. One-Shot Deviation Principle requires then that $(X, Fight)$ is a subgame-perfect Nash equilibrium of the above augmented stage game. Since X is a best response to Fight, we only need to ensure that Player 2 weakly prefers Fight to Accommodate after the entry in the above game. For this, we must have

$$-1 + \delta V_F \geq 1 + \delta V_A.$$

Substitution of the definitions of V_F and V_A in this inequality shows that this is equivalent to[6]

$$\delta \geq 2/3.$$

We have considered all possible histories, and when $\delta \geq 2/3$, the strategy profile has passed the one-shot deviation test. Therefore, when $\delta \geq 2/3$, the strategy profile is a subgame-perfect Nash equilibrium.

On the other hand, when $\delta < 2/3$, s^* is not a subgame-perfect Nash equilibrium. To show this it suffices to consider one history at which s^* fails the one-shot deviation test. For a history h in the second group, the augmented stage game is as above, and $(X, Fight)$ is not a subgame-perfect Nash equilibrium of this game, as $1 + \delta V_A > -1 + \delta V_F$.

A Step-by-Step Recipe The above application of One-Shot Deviation Principle illustrates a general step-by-step recipe:

Step 1 A strategy profile puts histories in groups, summarizing the relevant aspect of the history for the continuation game according to the strategy profile. These groups can be viewed as *states*. *The first step is to identify these states and the transition rule between the state—from the given strategy profile.* In the above example, there are two states: "deterrence", at which (X, F) is played, and "accommodation", at which (E, A) is played. The game starts at "deterrence" state and switches to accommodation state if and only if EF is played; once in accommodation state, the game remains there forever. Strategy profile are often described directly by the states, transition rule, and the stage-game strategy profile played in each state. In that case, one skip this step.

Step 2 *For each state, compute the outcome path (aka continuation play) and the present values (or equivalently the average values).* In the above example, in the "deterrence" state, the outcome path is $X, X, X, X, \ldots$, yielding the average value of $(0, 2)$; in the "accommodation" state, the outcome path is $EA, EA, \ldots$, yielding the average payoff of $(1, 1)$.

[6] The inequality is $\delta (V_F - V_A) \geq 2$. Substituting the values of V_F and V_A, we obtain $\delta / (1 - \delta) \geq 2$, i.e., $\delta \geq 2/3$.

Step 3 *For each state, construct the augmented stage game, by simply augmenting the payoffs at each terminal history taking into account which state that will be transitioned to in the next period.* If one uses the average values, this is done simply taking the weighted average of the stage game payoff—with weight $1-\delta$—and the average value in the transitioned state—with weight δ. For example, in the deterrence state, the transitioned state is "accommodation" after EA. Hence, in the augmented stage game, the payoff vector at node EA is

$$(1-\delta)(1,1)+\delta(1,1)=(1,1),$$

where the first $(1,1)$ is the payoff vector at EA in the stage game and the second $(1,1)$ is the vector of average values at state "accommodation". On the other hand, at EF the next state remains to be "deterrence". Hence, in the augmented stage game, the payoff vector at node EF is

$$(1-\delta)(-1,-1)+\delta(0,2)=(\delta-1,3\delta-1),$$

where $(-1,-1)$ is the stage-game payoff of EF and $(0,2)$ is the average value at the transitioned state. Alternatively, as in the above example, one can use present values and simply add the present values in the transitioned state, multiplied by δ, to the stage game payoffs.

Step 4 *For each state, check that the stage-game strategy profiled induced by the repeated game at that state is indeed a subgame-perfect Nash equilibrium of the augmented stage game for that state.* For example, in the above example, in state "deterrence", the induced stage-game strategy profile is (X,F), and it is a subgame-perfect Nash equilibrium of the augmented stage game if and only if $\delta\geq 2/3$. If this is the case for all states, the repeated game strategy profile is a subgame-perfect Nash equilibrium of the repeated game; it is not a subgame-perfect Nash equilibrium if it is not the case for a state.

Infinitely Repeated Prisoners' Dilemma When the stage game is a simultaneous action game, there is no distinction between subgame-perfect Nash equilibrium and Nash equilibrium. Hence, for one-shot deviation test, one simply checks whether $s^*(h)$ is a Nash equilibrium of the augmented stage game for h for every history h. This simplifies the analysis substantially because one only needs to compute the payoffs without deviation and with unilateral deviations in order to check whether the strategy profile is a Nash equilibrium.

As an example, consider the infinitely repeated Prisoner's dilemma game in (12.1). Consider the strategy profile (Grim Trigger,Grim Trigger). There are two kinds of histories we need to consider separately for this strategy profile.

1. Cooperation: Histories in which D has never been played by any player.

2. Defection: Histories in which D has been played by some player at some date.

There are two states: Cooperation and Defection. According to (Grim Trigger,Grim Trigger), at Cooperation state, players play (C, C) and the state remains Cooperation as long as (C, C) is played, and it switches to Defection state otherwise. At the Defection state, (D, D) is played and the state remains remains as is forever. At Cooperation state, the outcome path is $(C, C), (C, C), \ldots$, yielding payoff of 5 to each player every period. Hence, at this state, the average value of each player is

$$V_C = 5.$$

At the Defection state, the outcome path is (D, D) forever. Hence, at this state, the average value of each player is

$$V_D = 1.$$

The next two steps are to construct an augmented stage game and check that (C, C) is a Nash equilibrium of the augmented stage game in Cooperation state and (D, D) is a Nash equilibrium of the augmented stage game in Defection state some. To save time, one only needs to compute the payoffs necessary to check that players do not have incentive to deviate in these profiles.

In the Coordination state, with average values, the payoff from (C, C) is 5, as computed above; one writes $(1 - \delta) 5 + \delta 5 = 5$ to compute this payoff. The payoff of Player 1 from (D, C) is

$$6 (1 - \delta) + \delta V_D = 6 - 5\delta$$

because she gets 6 at the current period and the state transitions to Defection in the next period, where her average value is $V_D = 1$. Player 1 has no incentive to deviate if and only if

$$5 \geq 6 - 5\delta,$$

and this is the case if and only if

$$\delta \geq 1/5.$$

If $\delta < 1/5$, the strategy profile fails the One-Shot Deviation Test, and this shows that it is not a subgame-perfect Nash equilibrium. If $\delta \geq 1/5$, it passes the test at Cooperation state, and one still needs to apply One-Shot Deviation Test at Defection state.

At Defection state, the next state is always Defection. Hence, the payoffs in the augmented stage game are obtained by adding a constant (and multiplying with $1 - \delta$ if one uses average values). Since such affine transformations do not change the game, one simply checks that (D, D) is a Nash equilibrium of the stage game, which is of course the case.

When $\delta \geq 1/5$, (Grim Trigger, Grim Trigger) passes the one-shot deviation test at each history, showing that it is a subgame-perfect Nash equilibrium.

One can use the same technique to show that (Tit-for-tat,Tit-for-tat) is *not* a subgame-perfect Nash equilibrium (except for the knife-edge case $\delta = 1/5$). First observe that Tit-for-tat strategies at $t+1$ only depends on what is played at t not any previous play. Hence, the play in the previous period serves as the state in the "step-by-step recipe" above. There are four states: $(C,C), (C,D), (D,C), (D,D)$, and the game starts at state (C,C). At any state, the play in the current period becomes the state for the next period, and hence the payoffs in the augmented stage game do not depend on the state: *all states have identical augmented stage games.* Since each state prescribes a different stage-game strategy profile, (Tit-for-tat,Tit-for-tat) is a subgame-perfect Nash equilibrium if and only if every stage-game strategy profile is a Nash equilibrium of the augmented stage game. This can happen only in knife-edge cases, and (Tit-for-tat,Tit-for-tat) can be a subgame-perfect Nash equilibrium only at those knife-edge cases.

Formally, at state (C,C), each player plays C, and the outcome path is $(C,C), (C,C), \ldots$ forever, yielding average payoff of 5 for each player. Similarly, at state (D,D), each player plays D, and the outcome path is $(D,D), (D,D), \ldots$ forever, yielding average payoff of 1 for each player. At state (C,D), they alternate between (D,C) and (C,D), yielding the payoff stream

$$(6,0), (0,6), (6,0), (0,6), \ldots.$$

Thus, if (C, D) is played at the current period, the above path will played starting from the next period, yielding the stream of payoff vectors

$$\underbrace{(0,6)}_{\text{current period}} \quad \underbrace{(6,0), (0,6), (6,0), (0,6), \ldots}_{\text{future periods}}.$$

Therefore, in the augmented stage game, the payoff vector at (C, D) is

$$\left(\frac{6\delta}{1+\delta}, \frac{6}{1+\delta}\right).$$

Therefore, for every state, the augmented stage game is

	C	D
C	$5, 5$	$\frac{6\delta}{1+\delta}, \frac{6}{1+\delta}$
D	$\frac{6}{1+\delta}, \frac{6\delta}{1+\delta}$	$1, 1$

Now at $t = 0$, the players are supposed to play (C, C), and One-Shot Deviation Principle requires that (C, C) is a Nash equilibrium of the augmented stage game. This is the case if and only if

$$5 \geq \frac{6}{1+\delta},$$

which is equivalent to $\delta \geq 1/5$. If $\delta < 1/5$, the above inequality fails and this will be enough to show that Tit-for-Tat strategy profile above is not a subgame-perfect Nash equilibrium. It passes the test at this state if $\delta \geq 1/5$. But suppose (C, D) was played in the last period. Now they are supposed to play (D, C), and One-Shot Deviation Principle requires that (D, C) is a Nash equilibrium of the augmented stage game. This is the case if and only if

$$5 \leq \frac{6}{1+\delta},$$

i.e., $\delta \leq 1/5$. If $\delta > 1/5$, the above inequality fails and this will be enough to show that Tit-for-Tat strategy profile above is not a subgame-perfect Nash equilibrium. Therefore, Tit-for-Tat strategy profile is *not* a subgame-perfect Nash equilibrium—unless $\delta = 1/5$.[7]

[7] At that knife-edge case $\delta = 1/5$, the augmented stage game is

	C	D
C	$5, 5$	$1, 5$
D	$5, 1$	$1, 1$

and every profile is a Nash equilibrium of the augmented stage game. Tit-for-Tat strategy profile is a subgame-perfect Nash equilibrium at this knife-edge case.

12.2.2 Folk Theorem

A main objective of studying repeated games is to explore the interaction between the short-term incentives (within a single period) and the long-term incentives (within the broader repeated game). Conventional wisdom in game theory suggests that, when players are patient, their long-term incentives will overwhelm their short-term incentives, and a large set of behavior may emerge as a result in equilibrium. Indeed, for any given feasible and "individually rational" payoff vector and for sufficiently large values of δ, there exists some subgame perfect equilibrium that yields the payoff vector as the average value of the payoff stream. This fact is called Folk Theorem. There are many folk theorems that establish this fact under various technical conditions and setups. This section is devoted to presenting a basic version of folk theorem and illustrating its proof.

Throughout this section, it is assumed that the stage game is a **simultaneous action** game (N, A, u) where set $N = \{1, \dots, n\}$ is the set of players, $A = A_1 \times \cdots \times A_n$ is a **finite** set of strategy profiles, and $u_i : A \to \mathbb{R}$ is the stage-game utility functions.

Feasible Payoffs Imagine that the players collectively randomize over stage game strategy profiles $a \in A$. Which payoff vectors could they get if they could choose any probability distribution $p : A \to [0, 1]$ on A? (Recall that $\sum_{a \in A} p(a) = 1$.) The answer is: the set V of payoff vectors $v = (v_1, \dots, v_n)$ such that

$$v = \sum_{a \in A} p(a) \left(u_1(a), \dots, u_n(a) \right)$$

for some probability distribution $p : A \to [0, 1]$ on A. Note that V is the smallest convex set that contains all payoff vectors $(u_1(a), \dots, u_n(a))$ from pure strategy profiles in the stage game. A payoff vector v is said to be *feasible* if $v \in V$. Throughout this section, V is assumed to be n-dimensional.

For a visual illustration consider the Prisoners' Dilemma game in (12.1). The set V is plotted in Figure 12.5. Since there are two players, V contains pairs $v = (v_1, v_2)$. The payoff vectors from pure strategies are $(1, 1)$, $(5, 5)$, $(6, 0)$, and $(0, 6)$. The set V is the diamond shaped area that lies between the lines that connect these four points.

For every strategy profile s in the repeated game, the vector of average discounted

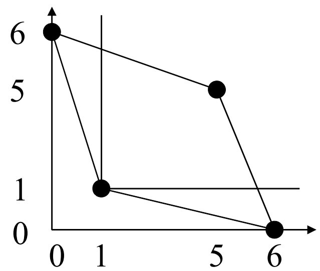


Figure 12.5: Feasible payoffs in Prisoners' Dilemma

values from s is in V.[8] Since the average values from mixed strategy profiles are convex combinations of those from pure strategy profiles, the average values from mixed strategy profiles are also contained in V. Conversely, if the players can collectively randomize on strategy profiles in the repeated games, all vectors $v \in V$ could be obtained as average payoff vectors.

Individual Rationality—MinMax payoffs There is a lower bound on how much a player gets in equilibrium. For example, in the repeated prisoners' dilemma, if one keeps playing defect everyday no matter what happens, she gets at least 1 every day, netting an average payoff of 1 or more. Then, she must get at least 1 in any Nash equilibrium because she could otherwise profitably deviate to the above strategy.

Towards finding a lower bound on the payoffs from pure-strategy Nash equilibria, for each player i define *pure-strategy minmax payoff* as

$$m_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i}). \tag{12.2}$$

[8] Indeed, the average payoff vector can be written as

$$U(s) = \sum_{a \in A} p_s(a)(u_1(a), \ldots, u_n(a))$$

where

$$p_s(a) = (1-\delta) \sum_{t \in T_{a,s}} \delta^t$$

and $T_{a,s}$ is the set of dates at which a is played on the outcome path of s. Clearly,

$$\sum_{a \in A} p_s(a) = (1-\delta) \sum_{a \in A} \sum_{t \in T_{a,s}} \delta^t = (1-\delta) \sum_{t \in T} \delta^t = 1.$$

Here, the other players try to minimize the payoff of player i by choosing a pure strategy s_{-i} for themselves, knowing that player i will play a best response to a_{-i}. Then, the harshest punishment they can inflict on i is m_i. For example, in the prisoners' dilemma game, $m_i = 1$ because i gets maximum of 6 if the other player plays C and gets maximum of 1 if the other player plays D.

Observe that in any pure-strategy Nash equilibrium s^* of the repeated game, the average payoff of player i is at least m_i. To see this, suppose that the average payoff of i is less than m_i in s^*. Now consider the strategy $\hat{s}_i$, such that for each history h, $\hat{s}_i(h)$ is a stage-game best response to $s^*_{-i}(h)$, i.e.,

$$u_i\left(\hat{s}_i(h), s^*_{-i}(h)\right) = \max_{a_i \in A_i} u_i\left(a_i, s^*_{-i}(h)\right).$$

Since

$$\max_{a_i \in A_i} u_i\left(a_i, s^*_{-i}(h)\right) \geq m_i$$

for every h, this implies that the average payoff from $\left(\hat{s}_i, s^*_{-i}\right)$ is at least m_i, giving player i an incentive to deviate.

A lower bound for the average payoff from a mixed strategy Nash equilibrium is given by *minmax payoff*, defined as

$$\mu_i = \min_{\alpha_j, j \neq i} \max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} \prod_{j \neq i} \alpha_j(a_j) u_i(a_i, a_{-i}), \tag{12.3}$$

where α_j is a mixed strategy of j in the stage game. Similarly to pure strategies one can show that the average payoff of player i is at least μ_i in any Nash equilibrium (mixed or pure). Note that, by definition, $\mu_i \leq m_i$. The equality can be strict. For example, in the matching-penny game

	Head	Tail
Head	$-1, 1$	$1, -1$
Tail	$1, -1$	$-1, 1$

the pure-strategy minmax payoff m_i is 1 while minmax payoff μ_i is 0. (This is obtained when $\alpha_j(Head) = \alpha_j(Tail) = 1/2$.)

Definition 12.1. A payoff vector v is said to be *individually rational* if $v_i \geq \mu_i$ for every $i \in N$.

Exercise 12.4. Show that in any Nash equilibrium σ^* of the repeated game, the average payoff of player i is at least μ_i.

Folk Theorem I will next present a general folk theorem and illustrate the main idea of the proof for a special case.

Theorem 12.3 (Folk Theorem)**.** *Let $v \in V$ be such that $v_i > \mu_i$ for every player i. Then, there exists $\bar{\delta} \in (0,1)$ such that for every $\delta > \bar{\delta}$ there exists a subgame-perfect Nash equilibrium of the repeated game under which the average value of each player i is v_i. Moreover, if $v_i > m_i$ for every i, then the subgame-perfect Nash equilibrium above is in pure strategies.*

The Folk Theorem states that any strictly individually rational and feasible payoff vector can be supported in subgame-perfect Nash equilibrium when the players are sufficiently patient. Since all equilibrium payoff vectors need to be individually rational and feasible, the Folk Theorem provides a rough characterization of the equilibrium payoff vectors when players are patient: the set of all feasible and individually rational payoff vectors.

I will next illustrate the main idea of the proof for a special case. Assume that, in the theorem, $v = (u_1(a^*), \ldots, u_n(a^*))$ for some $a^* \in A$ and there exists a Nash equilibrium $\hat{a}$ of the stage game such that $v_i > u_i(\hat{a})$ for every i. In the prisoners' dilemma example, the unique stage-game Nash equilibrium is $\hat{a} = (D, D)$ with payoff vector $(1, 1)$. In this example, one can then take $v = (5, 5)$ and $a^* = (C, C)$. As it has been shown above, in this example, strategy profile (Grim Trigger, Grim Trigger) yields $(a^*, a^*, \ldots)$ as its outcome path, and it is a subgame-perfect Nash equilibrium when $\delta > 1/5$. The main idea here is a generalization of Grim Trigger strategy. Consider the following strategy profile:

s^* : Play a^* until somebody deviates, and play $\hat{a}$ thereafter.

Clearly, under s^*, the average value of each player i is $u_i(a^*) = v_i$. Moreover, s^* is a subgame-perfect Nash equilibrium when δ is large. To see this, note that s^* passes the one-shot deviation test at histories with previous deviation because $\hat{a}$ is a Nash equilibrium of the stage game. Now consider a history in which a^* is played throughout. In the augmented stage game (with average payoffs), the payoff from a^* is v because they will keep playing a^* forever after that play. The payoff from any other $a \in A$ is

$$(1 - \delta)\, u(a) + \delta u(\hat{a})$$

because the players will switch to $\hat{a}$ after any such play. Then, a^* is a Nash equilibrium of the augmented stage game if and only if

$$v_i \geq (1-\delta) \max_{a_i} u\left(a_i, a_{-i}^*\right) + \delta u_i(\hat{a}) \tag{12.4}$$

for every player i. Let

$$\delta_i = \frac{\max_{a_i} u\left(a_i, a_{-i}^*\right) - v_i}{\max_{a_i} u\left(a_i, a_{-i}^*\right) - u_i(\hat{a})}$$

be the discount rate for which (12.4) becomes equality; such $\delta_i < 1$ exists because $\max_{a_i} u\left(a_i, a_{-i}^*\right) \geq u_i(a^*) = v_i > u_i(\hat{a})$. Take $\bar{\delta} = \max\{\delta_1, \ldots, \delta_n\}$. Then, for every $\delta > \bar{\delta}$, inequality (12.4) holds, and hence a^* is a Nash equilibrium of the augmented stage game. Therefore, s^* is a subgame-perfect Nash equilibrium whenever $\delta > \bar{\delta}$. Note that in the case of prisoners' dilemma, $\bar{\delta} = (6-5)/(6-1) = 1/5$.

The economic intuition of the proof is as follows. There is a social norm (or social convention), according which players play a^*. The players anticipate that any deviation from the social norm will trigger the other players to deviate from the social norm, leading to a breakdown of the social norm. Once the social norm breaks down, the players play perpetually according to the myopic Nash equilibrium $\hat{a}$. Then, each player faces a fundamental trade off between the short-term gain from deviation from the social norm and the long-term cost of such deviation due to the breakdown of social norm. The short-term gain is

$$\max_{a_i} u\left(a_i, a_{-i}^*\right) - v_i$$

for one period while the long-term loss is

$$\max_{a_i} u\left(a_i, a_{-i}^*\right) - u_i(\hat{a})$$

for every period starting from next period on. The player has no incentive to deviate if and only if the short-term gain, multiplied by $1-\delta$, is offset by the long term gain, multiplied by δ. This is the case, when δ exceeds the ratio

$$\delta_i = \frac{\text{Short-term Gain}}{\text{Long-term Loss}}.$$

In the above illustration, the vector v is obtained from playing the same a^*. What if this is not possible, i.e., v is a convex combination of payoff vectors from $a \in A$ but $v \neq u(a)$ for any $a \in A$. In that case, one can use time averaging to obtain v from pure strategy in the repeated game. For an illustration, consider $(2,2)$ in the repeated Prisoners' Dilemma game. Note that

$$(2,2) = \frac{1}{4}(5,5) + \frac{3}{4}(1,1) = (1,1) + \frac{1}{4}(4,4).$$

One can obtain average payoff vectors near $(2,2)$ in various ways. For example, consider the path

$$(C,C)\,(D,D)\,(D,D)\,(D,D)\,(C,C)\,(D,D)\,(D,D)\,(D,D)\cdots(C,C)\,(D,D)\,(D,D)\,(D,D)\cdots$$

The average value of each player from this path is

$$1 + \frac{1-\delta}{1-\delta^4}4 = 1 + \frac{4}{1+\delta+\delta^2+\delta^3}.$$

As $\delta \to 1$, this value approaches 2. Another way to approximate $(2,2)$ would be first to play (D,D) then switch to (C,C). For example, let t^* be the smallest integer for which $\delta^{t^*} \leq 1/4$. Note that when δ is large, $\delta^{t^*} \cong 1/4$. Now consider the path on which (D,D) is played for every $t < t^*$ and (C,C) is played for every $t \geq t^*$. The average value is

$$\left(1-\delta^{t^*}\right)\cdot 1 + \delta^{t^*}5 \cong 2.$$

Here, v is *approximated* by time averaging. When δ is large, one can obtain each v *exactly* by time averaging.[9]

12.2.3 Reward and Punishment Strategies

Subgame-perfect Nash equilibria are typically supported by a fear that a deviation leads to a reactive path that yields a very low average payoff to the deviating player. The fear of such punishment is what deters the players from deviating from the path. The heavier and the longer is the punishment, the more deterrent it is. The trigger strategies studied so far have the advantage of punishing the deviations forever using a stage-game Nash equilibrium. However, they are also limited by how bad a stage-game Nash equilibrium can be. In some games, such as the Prisoners' Dilemma game and Bertrand competition, the stage-game Nash equilibrium payoffs are very low and the trigger strategies are highly effective self-enforcement mechanisms. Indeed in these games, the minmax payoffs

[9] For mathematically oriented students: imagine writing each weight $p(a) \in [0,1]$ in base $1/\delta$.

are achieved by a Nash equilibrium of the stage game, and a player's average payoff cannot be below her minmax payoff after any deviation. In those games, the trigger strategies are the most effective self-enforcement mechanisms. But in many games, such as Cournot duopoly, the Nash equilibrium payoffs can be very high. In that case, the trigger strategies will have limited deterrence, and there can be more effective ways to enforce a behavior.

Once one goes outside the trigger strategies, new difficulties arise. After a player deviates, the other players may not want to follow through the punishments that were meant to deter players from deviation because those punishments may be too costly for the punishing players. In order for them to follow through the punishments, they must have an incentive; they may need to be rewarded in the future for carrying out these punishments and also they may need to be punished for failing to punish the other players. If they do not have incentive to follow through the punishments, then the players foresee that the punishments will not be carried out and deviate from the path. (That is why one checks incentive to deviate at all histories, including those that are not supposed to arise if players stick to their strategies.)

When the minmax payoffs are below Nash equilibrium payoffs in the stage game, one can enforce a wide range of equilibrium behavior using versions Carrot & Stick strategies that reward "good" behavior and punish "bad" behavior. For an illustration of these strategies, consider the following Prisoners' Dilemma game with an additional strategy P that involves self-punishment:

	C	D	P
C	$5,5$	$0,6$	$0,x$
D	$6,0$	$1,1$	$0,x$
P	$x,0$	$x,0$	x,x

(12.5)

where $x < 0$. Observe that the minmax strategy is

$$\underline{v}_i = 0,$$

which is achieved by other player playing P and player i playing D. Since P is strictly dominated by D, the unique Nash equilibrium is (D, D) as in the Prisoners' Dilemma game, yielding the payoff of 1 for each player. The lowest payoff that can be achieved by a trigger strategy is then 1, as player will get 1 after deviation. However, by the

Folk Theorem, one can achieve any positive payoff as the average value in a subgame-perfect Nash equilibrium when players are sufficiently patient. I will next use Carrot & Stick strategies to construct such equilibria. First consider the following Carrot & Stick strategy profile:

Carrot & Stick There are two states: Carrot and Stick. The game starts at Carrot state. In Carrot state, the players play (C, C); they remain in Carrot state if they play (C, C), and switch to Stick state otherwise. In Stick state, they play (P, P); they switch to Carrot if they play (P, P) and remain in Stick otherwise.

Observe that Carrot state is always used as a reward for the desired behavior in the previous period: the state is Carrot if they played what they were supposed to play in the previous period. The Stick is used as punishment, as any unwanted play leads to Stick in the next period.

The necessary and sufficient conditions for Carrot & Stick being a subgame-perfect Nash equilibrium are as follows. First, in the Carrot state, the average value of each player is

$$V_C = 5$$

because they will play (C, C) repeatedly forever. In the Stick state, the average value of each player is

$$V_S = (1 - \delta)\, x + \delta V_C$$

because they will play (P, P) and switch to Carrot state in the next period. Now, in the Carrot state, the best deviation is to play D. If one deviates, she gets 6 at that period, x in the next period, and goes back to the Carrot state after that. If she sticks to the strategy, she gets 5 in both of those periods and remains in the Carrot state afterwards. Therefore, a player does not have an incentive to deviate at Carrot state if and only if

$$5 + \delta 5 \geq 6 + \delta x,$$

i.e.,

$$\delta\, (5 - x) \geq 1. \tag{12.6}$$

(In general, under Carrot & Stick strategies, one checks incentive to deviate in the Carrot state by comparing the payoff in the next two periods.) One checks the incentive to deviate in the Stick state as follows. If a player sticks to the strategy at Stick state, she gets V_S as her average value. If she deviates in best possible way, she plays D and gets payoff of zero that period, but then the game starts at stick again, yielding the average value of $(1-\delta)\times 0+\delta V_S$. The player has no incentive to deviate if and only if

$$V_S \geq (1-\delta)\times 0+\delta V_S.$$

That is,

$$V_S \geq 0. \tag{12.7}$$

In other words, a player does not have an incentive to deviate if and only if her average value V_S in Stick state is at least as high as her payoff from best deviation at the stick state. This will always be the condition in the Stick state. Of course, the condition $V_S \geq 0$ is equivalent to

$$x \geq -5\delta/(1-\delta) \tag{12.8}$$

in this particular example. The above strategy profile is a SPNE if and only if (12.6) and (12.8) hold.

In the above example, the game starts at the Carrot state, but it could also start at the Stick state, yielding possibly very low average value of V_S. The only condition for V_S is that $V_S \geq 0$. Since 0 is the minmax value, when x is near $-5\delta/(1-\delta)$, one can achieve nearly minmax value using this strategy with Stick as its initial state. One can then use this strategy to enforce a wide range of behavior.

For a more concrete example, take $\delta = 0.9$ and $x = -42$. The average value for Carrot & Stick strategy with Stick as the initial state is $V_S = 0.3$. This is substantially lower than the stage-game equilibrium payoff. As discussed above, since Player 2 cannot get less than 1 in a trigger strategy, Player 1 cannot get more that 5.8 as the average value in a trigger strategy equilibrium. But using the Carrot & Stick as the deterrent, one can get much higher payoff for Player 1 as follows. Take $\hat{t} = 20$ so that $\delta^{\hat{t}} \cong 0.12$. Consider the path in which they play (D, C) at period $t = 0, 1, \ldots, \hat{t}-1$ and (C, C) at period $\hat{t}$ and thereafter. The average values of Players 1 and 2 are $6\left(1-\delta^{\hat{t}}\right)+5\delta^{\hat{t}} = 6-\delta^{\hat{t}} \cong 5.88$ and $5\delta^{\hat{t}} \cong 0.6$. On this path Player gets very high payoff while Player 2 gets very low payoff. One can achieve this path as an equilibrium path using the following strategy profile.

A highly asymmetric SPNE There are three "states": P1, Carrot and Stick. The game starts at P1. If the state is P1, the play depends on the date: they play (D, C) for $t < \hat{t}$ and (C, C) for $t \geq \hat{t}$. They remain at P1 as long as they both stick to this path. If anybody deviates, the state transitions to Stick. In Stick state, they play (P, P); they switch to Carrot if they play (P, P) and remain in Stick otherwise. In Carrot state, the players play (C, C); they remain in Carrot state if they play (C, C), and switch to Stick state otherwise.

The calculations above show that players do not have an incentive to deviate in Carrot and Stick states, and the average value at the Stick state is $V_S = 0.3$. At state P1, clearly, Player 1 has no incentive to deviate; Player 1 gets the highest possible payoff for $t < \hat{t}$, and we are in Carrot state of the original Carrot & Stick equilibrium for $t \geq \hat{t}$. Player 2 has highest incentive to deviate on the path at $t = 0$ because her average value improves over time on the path and remains constant at 0.3 after any deviation. Hence, it suffices to check that Player 2 has no incentive to deviate at $t = 0$. At $t = 0$, her average value is $5\delta^{\hat{t}} \cong 0.6$. If Player 2 deviates and plays D, her average value is $(1 - \delta) + \delta V_S = 0.37$, which is lower. Therefore, she does not have an incentive to deviate.

What if x is not very small? In that case, one can achieve similarly small payoffs by considering multiple rounds of punishments. For a concrete example, take $x = -1$. Now, in the Carrot & Stick equilibrium above, the average value at Stick state is $V_S = 4.4$, higher than the myopic Nash equilibrium payoff. In that case, one can use the following version of Carrot & Stick strategy profile to get very low payoffs for each player:

Carrot & m Sticks There are $m + 1$ states: Carrot, Stick$_0$, Stick$_1$,…,Stick$_{m-1}$. The game starts at Stick$_0$ state. In Carrot state, the players play (C, C); they remain in Carrot state if they play (C, C), and switch to Stick$_0$ state otherwise. In each Stick$_k$ state, they play (P, P); if they play (P, P), they switch to Stick$_{k+1}$ for $k < m - 1$ and to Carrot if $k = m - 1$; they switch to Stick$_0$ if anybody deviates from P.

Observe that, once a punishment starts, it lasts m periods, and any deviation restarts the m periods of punishments. The average value is $V_C = 5$ at Carrot state and

$$V_k = \left(1 - \delta^{m-k}\right) x + 5\delta^{m-k}$$

at any Stick$_k$ state. Players do not have an incentive to deviate at Carrot state if and only if

$$5 + \delta 5 + \delta^2 5 + \cdots + \delta^m 5 \geq 6 + \delta x + \delta^2 x + \cdots + \delta^m x,$$

which can be rewritten as

$$(5 - x)\,\delta\,(1 - \delta^m) \geq 1 - \delta.$$

For $x = -1$ and $\delta = 0.9$, this condition is satisfied for any $m \geq 1$. For any Stick$_k$ state, players do not have an incentive to deviate if and only if

$$V_k \geq \delta V_0$$

because they get 0 when they deviate and the deviation restarts the punishment. Since

$$V_{m-1} \geq V_{m-2} \geq \cdots \geq V_0,$$

all of these incentive conditions $V_k \geq \delta V_0$ are satisfied if and only if

$$V_0 \geq 0,$$

as in the usual Carrot & Stick strategy. That is,

$$V_0 = (1 - \delta^m)\,x + 5\delta^m = x + (5 - x)\,\delta^m \geq 0,$$

which can be re-written as

$$\delta^m \geq -\frac{x}{5 - x},$$

bounding the number of stick periods one can have. Since V_0 can be arbitrarily close to the minmax value of 0 for sufficiently large m, one can use this strategy to enforce a wide range of payoffs as in the previous example.

More generally, subgame-perfect Nash equilibrium requires that the continuation game is also a subgame-perfect Nash equilibrium of the repeated game starting from any deviation. For example, a player's payoff cannot be lower than her minmax payoff at any history. This limits how bad a player's payoff can be after a deviation, limiting what can be enforced in equilibrium. If one knew the worst subgame-perfect Nash equilibrium path for each player, one could enforce any possible subgame-perfect Nash equilibrium outcome path by switching to the first day of the worst subgame-perfect Nash equilibrium for the deviating player whenever there is a unilateral deviation. For example, in the

Prisoners' Dilemma game and Bertrand competition, the minmax payoffs are achieved by the unique Nash equilibrium, and the worst subgame-perfect Nash equilibrium repeats the stage-game Nash equilibrium forever. In those games, one can focus on the triggers strategies to determine what can be supported by a subgame-perfect Nash equilibrium. More generally, it is difficult to compute the worst equilibria of the players, and one can use versions of Carrot and Stick strategies to enforce a wide range of outcome paths.

12.3 Exercises with Solutions

Exercise 12.5. Consider the $\bar{t}$-times repeated game with the following stage game

	a	b	c
a	$3,3$	$0,0$	$0,0$
b	$0,0$	$2,2$	$1,0$
c	$0,0$	$0,1$	$0,0$

1. Find a lower bound π for the *average* payoff of each player in all pure strategy Nash equilibria. Prove indeed that the payoff of a player is at least $\pi\bar{t}$ in every pure-strategy Nash equilibrium.

2. Construct a pure-strategy subgame-perfect Nash equilibrium in which the payoff of each player is at most $\bar{t}+1$. Verify that the strategy profile is indeed a subgame-perfect Nash equilibrium.

Solution. Part 1: Note that the pure strategy minmax payoff of each player is 1. Hence, the payoff of a player cannot be less than $\bar{t}$. Indeed, if a player mirrors what the other player is supposed to play in any history at which the other player plays a or b according to the equilibrium and play b if the other player is supposed to play c at the history, then her payoff would be at least $\bar{t}$. Since she plays a best response in equilibrium, her payoff is at least that amount. This lower bound is tight. For $\bar{t} = 2k > 1$, consider the strategy profile

`Play` (c,c) `for the first` k `periods and` (b,b) `for the last` k `periods; if any player deviates from this path, play` (c,c) `forever.`

Note that the payoff from this strategy profile is $\bar{t}$. To check that this is a Nash equilibrium, note that the best possible deviation is to play play b forever, which yields

$\bar{t}$, giving no incentive to deviate. Note also that the equilibrium here is not subgame-perfect.

Part 2: Recall that $T = \{0, \ldots, \bar{t} - 1\}$. For $\bar{t} = 1$, (b, b) is the desired equilibrium. Towards a mathematical induction, now take any $\bar{t} > 1$ and assume that for every $m < \bar{t}$, the m-times repeated game has a pure-strategy subgame-perfect Nash equilibrium $s^* [m]$ in which each player gets $m + 1$. For $\bar{t}$-times repeated game, consider the path

$$\underbrace{(c, c) \cdots (c, c)}_{(\bar{t} - 1)/2 \text{ times}} \quad \underbrace{(b, b) \cdots (b, b)}_{(\bar{t} + 1)/2 \text{ times}}$$

if $\bar{t}$ is odd and the path

$$\underbrace{(c, c) \cdots (c, c)}_{\bar{t}/2 \text{ times}} \quad \underbrace{(b, b) \cdots (b, b)}_{\bar{t}/2 - 1 \text{ times}} \quad (a, a)$$

if $\bar{t}$ is even. Note that the total payoff of each player from this path is $\bar{t} + 1$. Consider the following strategy profile.

`Play according to the above path; if any player deviates from this path at any` $t \leq \bar{t}/2 - 1$, `switch to` $s^* [\bar{t} - t - 1]$ `for the remaining` $(\bar{t} - t - 1)$`-times repeated game; if any player deviates from this path at any` $t > \bar{t}/2$, `remain on the path.`

This is a subgame-perfect Nash equilibrium. There are three classes of histories to check. First, consider a history in which some player deviated from the path at some $t' \leq \bar{t}/2$. In that case, the strategy profile already prescribes to follow the subgame-perfect Nash equilibrium $s^* [\bar{t} - t' - 1]$ of the subgame that starts from $t' + 1$, which remains subgame perfect at the current subgame as well. Second, consider a history in which no player has deviated from the path at any $t' \leq \bar{t}/2$ and take $t > \bar{t}/2$. In the continuation game, the above strategy profile prescribes: play (b, b) every day if $\bar{t}$ is odd and play (b, b) every day but the last day and play (a, a) on the last day if $\bar{t}$ is even. Since (a, a) and (b, b) are Nash equilibria of the stage game, this is clearly a subgame-perfect Nash equilibrium of the remaining game. Finally, take $t \leq \bar{t}/2$ and consider any on-path history. Now, a player's payoff is $\bar{t} + 1$ if she follows the strategy profile. If she deviates at t, she gets at most 1 at t and $(\bar{t} - t - 1) + 1 \leq \bar{t}$ from the next period on, where $(\bar{t} - t - 1) + 1$ is her payoff from $s^* [\bar{t} - t - 1]$. His total payoff cannot exceed $\bar{t} + 1$, and she has no incentive to deviate.

Exercise 12.6. Consider the infinitely repeated prisoners' dilemma game of (12.1) with discount factor $\delta = 0.999$.

1. Find a subgame-perfect Nash equilibrium in pure strategies under which the average payoff of each player is in between 1.1 and 1.2. Verify that your strategy profile is indeed a subgame-perfect Nash equilibrium.

2. Find a subgame perfect Nash equilibrium in pure strategies under which the average payoff of player 1 is at least 5.7. Verify that your strategy profile is indeed a subgame-perfect Nash equilibrium.

3. Can you find a subgame-perfect Nash equilibrium under which the average payoff of player 1 is more than 5.8?

Solution. Take any $\hat{t}$ with $\left(1-\delta^{\hat{t}}\right)+5\delta^{\hat{t}} = 1+4\delta^{\hat{t}} \in (1.1, 1.2)$, e.g., any $\hat{t}$ between 2994 and 3687. Consider the strategy profile

`Play` (D, D) `at any` $t < \hat{t}$ `and` (C, C) `at` $\hat{t}$ `and thereafter. If any player deviates from this path, play` (D, D) `forever.`

Note that the average value of each player is $\left(1-\delta^{\hat{t}}\right)+5\delta^{\hat{t}} \in (1.1, 1.2)$. To check that it is a subgame-perfect Nash equilibrium, first take any on-path history with date $t \geq \hat{t}$. At that history, the average value of each player is 5. If a player deviates, then her average value is only $6(1-\delta)+\delta = 1.05$. Hence, she has no incentive to deviate. For $t < \hat{t}$, the average value is

$$\left(1-\delta^{\hat{t}-t}\right)+5\delta^{\hat{t}-t} \geq \left(1-\delta^{\hat{t}}\right)+5\delta^{\hat{t}} > 1.1.$$

If she deviates, her average value is only δ. Therefore, she does not have an incentive to deviate, once again. Since they play static Nash equilibrium after switch, there is no incentive to deviate at such a history, either. Therefore, the strategy profile above is a subgame-perfect Nash equilibrium.

Part 2: Take any $\hat{t}$ with $\left(1-\delta^{\hat{t}}\right)6+5\delta^{\hat{t}} = 6-\delta^{\hat{t}} \in (5.7, 5.8)$, i.e., $\delta^{\hat{t}} \in (0.2, 0.3)$. The possible values for $\hat{t}$ are the natural numbers from 1204 to 1608. Consider the strategy profile

`Play` (D, C) `at any` $t < \hat{t}$ `and` (C, C) `at` $\hat{t}$ `and thereafter. If any player deviates from this path, play` (D, D) `forever.`

Note that the average value of Player 1 is $6-\delta^{\hat{t}}$, taking values between $6-0.999^{1204}=5.7002$ and $6-0.999^{1608}=5.7999$. Note also that the strategy profile coincides with the one in part (a) at all off-the-path histories and at all on-the-path histories with $t \geq \hat{t}$. Hence, to check whether it is a subgame-perfect Nash equilibrium, it suffices to check for on-the-path histories with $t < \hat{t}$. At any such history, clearly, Player 1 does not have an incentive to deviate (as in part (a)). For Player 2, the average value is

$$5\delta^{\hat{t}-t} \geq 5\delta^{\hat{t}} \geq 0.999^{1608}5 \cong 1.0006.$$

If she deviates, her average value is only 1 (getting 1 instead of 0 on the first day and getting 1 forever thereafter). Therefore, she does not have an incentive to deviate. Therefore, the strategy profile above is a subgame-perfect Nash equilibrium.

Part 3: While the average payoff of Player 1 can be as high as 5.7999, it cannot be higher than 5.8. This is because $v_2 < 1$ for any feasible v with $v_1 > 5.8$. Such an individually irrational payoff cannot result in equilibrium because Player 2 could do better by simply playing D at every history (as discussed in the text).

Exercise 12.7. Consider the infinitely repeated game in which the stage game is the game in Exercise 10.4 and the discount factor is $\delta > 0$. Fix some $\hat{r} \in (v_A, v_B)$. Assuming that δ is sufficiently high, find a subgame-perfect Nash equilibrium with the following outcome path: $r = \hat{r} = b_C > b_B$ at even dates $t \in \{0, 2, \ldots\}$ and $r = \hat{r} = b_B > b_C$ at odd dates $t \in \{1, 3, \ldots\}$. Specify the range of δ under which your strategy profile is a subgame-perfect Nash equilibrium, and verify that this is indeed the case.

Solution. Consider the following strategy profile. For any t and any h at the beginning of t in which no bidder deviated from the following rule, they follow the following rule:

$$\begin{aligned} r(h) &= \hat{r} \\ b_C(h,r) &= \min\{r,\hat{r}\}, b_B(h,r) < \min\{r,\hat{r}\} \qquad \text{(if } t \text{ is even)} \\ b_B(h,r) &= \min\{r,\hat{r}\}, b_C(h,r) < \min\{r,\hat{r}\} \qquad \text{(if } t \text{ is odd)}; \end{aligned}$$

for any other h, the players play according to SPNE in part (a) of Problem 3 in Section 10.7. (That is, if the bidders deviate from the path, they switch to the stage game SPNE; they do not switch if Alice deviates.) After the switch, Alice sets $r = v_C$, and Caroll buys it at price v_C, yielding payoff v_C for Alice and 0 for Bob and Caroll.

Now, after the switch, they play a SPNE of the stage game forever; hence the strategy profile is a SPNE after the switch. Before the switch, given the bidders' strategy, Alice's move does not affect her future payoff and she can sell the painting at price $\hat{r}$ at most. Hence, in the augmented stage game, setting the reservation price $r = \hat{r}$ is a best response to the bidders' strategies. Now consider any h before the switch and any reserve price r. In the augmented stage game, the payoff of a bidder i is as follows, assuming that the other bidder follows the equilibrium strategy:

case	payoff from strategy	best deviation payoff
$b_i(h,r) = \min\{r,\hat{r}\}, r \leq \hat{r}$	$(v_i - r)(1-\delta) + \delta^2 V_i$	$(v_i - r)(1-\delta)$
$b_i(h,r) = \min\{r,\hat{r}\}, r > \hat{r}$	$\delta^2 V_i$	$(v_i - r)(1-\delta)$
otherwise	δV_i	$(v_i - \max\{r,\hat{r}\})(1-\delta)$

where $V_i = (v_i - \hat{r})/(1+\delta)$. In order for the strategy profile to be a SPNE, i must not have an incentive to deviate in all cases above. A sufficient condition is

$$\delta^2 (v_i - \hat{r})/(1+\delta) \geq (v_i - \hat{r})(1-\delta),$$

i.e.,

$$\delta^2 \geq 1 - \delta^2,$$

i.e., $\delta \geq 1/\sqrt{2}$.

Exercise 12.8. Alice and Bob are a couple, playing the infinitely repeated game with the following stage game and discount factor δ. Every day, simultaneously, Alice and Bob spend $x_A \in [0,1]$ and $x_B \in [0,1]$ fraction of their time in their relationship, respectively, receiving the stage payoffs $u_A = \ln(x_A + x_B) + 1 - x_A$ and $u_B = \ln(x_A + x_B) + 1 - x_B$, respectively. (Alice and Bob are denoted by A and B, respectively.) For each of the strategy profiles below, find the conditions on the parameters for which the strategy profile is a subgame-perfect Nash equilibrium.

1. Both players spend all of their time in their relationship (i.e. $x_A = x_B = 1$) until somebody deviates; the deviating player spends 1 and the other player spends 0 thereafter. (Find the range of δ.)

2. There are 4 states: E (namely, Engagement), M (namely, Marriage), D_A and D_B. The game starts at state E, in which each player spends $\hat{x} \in (0,1)$. If both spends

$\hat{x}$, they switch to state M; they remain in state E otherwise. In state M, each spends 1. They remain in state M until one player $i \in \{A, B\}$ spends less than 1 while the other player spends 1, in which case they switch to D_i state. In D_i state, player i spends $\tilde{x}_i$ and the other player spends $1 - \tilde{x}_i$ forever. (Find the set of inequalities that must be satisfied by the parameters δ, $\hat{x}$, $\tilde{x}_A$, and $\tilde{x}_B$.)

Solution. It is useful to note that $(x, 1-x)$ is a Nash equilibrium of the stage game for every $x \in [0, 1]$.

Part 1: Since $(1, 0)$ and $(0, 1)$ are Nash equilibria of the stage game, there is no incentive to deviate at any history with previous deviation by one player. Now consider any other history, in which they both are supposed to spend 1. If a player i follows the strategy, her average payoff is

$$\ln 2.$$

Suppose she deviates and spends $x_i < 1$. Then, since the other player is supposed to spend 1, in the continuation game, player i spends 1 and the other player spends 0. This yields 0 for player i. Hence, the average value of player i from deviation is

$$(\ln(1 + x_i) + 1 - x_i)(1 - \delta).$$

The best possible deviation is $x_i = 0$, yielding the payoff of

$$1 - \delta.$$

Hence, the strategy profile is a subgame-perfect Nash equilibrium if and only if

$$\ln 2 \geq 1 - \delta,$$

where the values on left and right hand sides of inequality are the average values from following the strategy profile and best deviation, respectively. One can write this as a lower bound on the discount factor:

$$\delta \geq 1 - \ln 2.$$

Part 2: Since $(\tilde{x}_i, 1 - \tilde{x}_i)$ is a Nash equilibrium of the stage game, there is no incentive to deviate at state D_i for any $i \in \{A, B\}$. In state M, the average payoff from following the strategy profile is $\ln 2$. If a player i deviates at state M, the next state is D_i (as in part (a)), which gives the average payoff of $1-\tilde{x}_i$ to i. Hence, as in part (a), the average payoff from best deviation is $1-\delta+\delta(1-\tilde{x}_i)=1-\delta\tilde{x}_i$. Therefore, there is no incentive to deviate at state M if and only if $\ln 2 \geq 1-\delta\tilde{x}_i$, i.e.

$$\delta\tilde{x}_i \geq 1-\ln 2. \tag{12.9}$$

On the other hand, in state E, the average payoff from following the strategy is

$$\begin{aligned} V_E &= (1-\delta)(\ln(2\hat{x})+1-\hat{x})+\delta\ln 2 \\ &= \ln 2+(1-\delta)(\ln\hat{x}+1-\hat{x}). \end{aligned}$$

By deviating and playing $x_i \neq \hat{x}$, player i can get

$$(1-\delta)(\ln(\hat{x}+x_i)+1-x_i)+\delta V_E.$$

The best deviation is $x_i = 1-\hat{x}$ and yields the maximum average payoff of

$$(1-\delta)\hat{x}+\delta V_E.$$

There is no incentive to deviate at E if and only if

$$V_E \geq (1-\delta)\hat{x}+\delta V_E,$$

which simplifies to

$$V_E \geq \hat{x}.$$

By substituting the value of V_E, one can write this condition as

$$\ln 2+(1-\delta)(\ln\hat{x}+1-\hat{x}) \geq \hat{x}. \tag{12.10}$$

The strategy profile is a SPNE if and only if (12.9) and (12.10) are satisfied.

Remark 12.2. One can make strategy profile above a subgame-perfect Nash equilibrium by varying all three parameters $\hat{x}$, $\tilde{x}_1$, $\tilde{x}_2$, and δ. For a fixed $(\hat{x},\tilde{x}_1,\tilde{x}_2)$, both conditions bound the discount factors from below, yielding

$$\delta \geq \max\left\{\frac{1-\ln 2}{\tilde{x}_1}, \frac{1-\ln 2}{\tilde{x}_2}, 1-\frac{\hat{x}-\ln 2}{\ln\hat{x}+1-\hat{x}}\right\}.$$

(To see this, observe that $\ln \hat{x} + 1 - \hat{x} < 0$.) Of course, when δ is fixed, the above conditions can also be interpreted as bounds on $\tilde{x}_i$ and $\hat{x}$. First, the contribution of the guilty party i in the divorce state D_i cannot be too low:

$$\tilde{x}_i \geq \frac{1 - \ln 2}{\delta}.$$

For otherwise, the parties deviate and marriage cannot be sustained. Second, the above lower bound on δ also gives an absolute upper bound on the effort level during the engagement. Since $\delta < 1$ and $\ln \hat{x} + 1 - \hat{x} < 0$, the condition on δ implies that

$$\hat{x} < \ln 2 \cong 0.693.$$

For otherwise, the lower bound on δ would exceed 1. That is, one must start small, as engagement may never turn into marriage otherwise. Of course, one could also skip the engagement altogether.

Exercise 12.9. This question is about a milkman and a customer. At any day, with the given order,

- Milkman puts $m \in [0, 1]$ liter of milk and $1 - m$ liter of water in a container and closes the container, incurring cost cm for some $c > 0$;
- Customer, without knowing m, decides on whether or not to buy the liquid at some price p. If she buys, her payoff is $vm - p$ and the milkman's payoff is $p - cm$. If she does not buy, she gets 0, and the milkman gets $-cm$. If she buys, then she learns m.

1. Assume that this is repeated for 100 days, and each player tries to maximize the sum of her or her stage payoffs. Find all subgame-perfect equilibria of this game.
2. Now consider the infinitely repeated game with the above stage game and with discount factor $\delta \in (0, 1)$. What is the range of prices p for which there exists a subgame perfect equilibrium such that, everyday, the milkman chooses $m = 1$, and the customer buys on the path of equilibrium play?

Solution. Part 1: The stage game has a unique Nash equilibrium, in which $m = 0$ and the customer does not buy. Therefore, the finitely repeated game has a unique subgame-perfect Nash equilibrium, in which the stage equilibrium is repeated.

Part 2: The milkman can guarantee himself 0 by always choosing $m = 0$. Hence, his continuation value at any history must be at least 0. Hence, in the worst equilibrium, if he deviates customer should not buy milk forever, giving the milkman exactly 0 as the continuation value. Hence, the SPNE we are looking for is *the milkman always chooses $m = 1$ and the customer buys until anyone deviates, and the milkman chooses $m = 0$ and the customer does not buy thereafter.* If the milkman does not deviate, his average value is

$$V = p - c.$$

The best deviation for her (at any history on the path of equilibrium play) is to choose $m = 0$ (and not being able to sell thereafter). In that case, his average value is

$$V_d = p(1 - \delta) + \delta 0 = p(1 - \delta).$$

In order this to be an equilibrium, we must have $V \geq V_d$; i.e.,

$$p - c \geq p(1 - \delta),$$

i.e.,

$$p \geq c/\delta.$$

In order for the customer to buy on the equilibrium path, it must also be true that $p \leq v$. Therefore,

$$v \geq p \geq c/\delta.$$

Exercise 12.10. Since the British officer had a thick pen when he drew the border, the border of Iraq and Kuwait is disputed. Unfortunately, the border passes through an important oil field. In each year, simultaneously, each of these countries decide whether to extract high (H) or low (L) amount of oil from this field. Extracting high amount of oil from the common field hurts the other country. In addition, Iraq has the option of attacking Kuwait (W), which is costly for both countries. The stage game is as follows:

	H	L
H	$2, 2$	$4, 1$
L	$1, 4$	$3, 3$
W	$-1, -1$	$-1, -2$

where Iraq is the row player and Kuwait is the column player. Consider the infinitely repeated game with this stage game and with discount factor $\delta = 0.9$.

1. Find a subgame perfect Nash equilibrium in which each country extracts low (L) amount of oil every year on the equilibrium path.[10]

2. Find a subgame perfect Nash equilibrium in which Iraq extracts high (H) amount of oil and Kuwait extracts low (L) amount of oil every year on the equilibrium path.

Solution. Part 1: Consider the strategy profile

`Play` (L, L) `until somebody deviates and play` (H, H) `thereafter.`

This strategy profile is a subgame-perfect Nash equilibrium whenever $\delta \geq 1/2$. (You should be able to verify this at this stage.)

Part 2: Consider the following ("Carrot and Stick") strategy profile[11]

`There are two states: War and Peace. The game starts at state Peace. In state Peace, they play` (H, L)`; they remain in Peace if` (H, L) `is played and switch to War otherwise. In state War, they play` (W, H)`; they switch to Peace if` (W, H) `is played and remain in War otherwise.`

This strategy profile is a subgame-perfect Nash equilibrium whenever $\delta \geq 3/5$. The vector of average values is $(4, 1)$ in state Peace and $(-1, -1)(1 - \delta) + \delta(4, 1) = (5\delta - 1, 2\delta - 1)$ in War. Note that both countries strictly prefer Peace to War. In state Peace, Iraq clearly has no incentive to deviate. On the other hand, Kuwait can have a short term gain of 1 by deviating to H. Deviation to H leads to payoffs of 2 and -1 for the present and the next period, going back to Peace after that. By sticking to equilibrium strategy, Kuwait would get 1 for each of those periods. Hence, Kuwait has no incentive to deviate if and only if

$$2 - \delta \geq 1 + \delta,$$

i.e., $\delta \geq 1/2$, which is indeed the case. In state War, Kuwait clearly has no incentive to deviate. In that state, Iraq could possibly benefit from deviating to H, the best deviation at that state. As is typical in the Stick state of a Carrot-Stick strategy, Iraq does not have an incentive to deviate if and only if its average value at War state, $5\delta - 1$, is at least as high as the payoff of 2 from the best deviation H at that state:

$$5\delta - 1 \geq 2,$$

[10] That is, an outside observer would observe that each country extracts low amount of oil every year.

[11] See the next chapter for more on Carrot and Stick strategies.

i.e., $\delta \geq 3/5$, which is clearly the case.

Exercise 12.11. Below, there are pairs of stage games and strategy profiles. For each pair, check whether the strategy profile is a subgame-perfect Nash equilibrium of the infinitely repeated game with the given stage game and discount factor $\delta = 0.99$.

1. **Stage Game:**

	S	R
S	$6, 6$	$0, 4$
R	$4, 0$	$4, 4$

   Strategy profile: Each player plays S in the first round and in the following rounds she plays what the other player played in the previous round (i.e., at each $t > 0$, she plays what the other player played at $t - 1$).

2. **Stage Game:**

	L	M	R
T	$3, 1$	$0, 0$	$-1, 2$
M	$0, 0$	$0, 0$	$0, 0$
B	$-1, 2$	$0, 0$	$-1, 2$

   Strategy profile: Until some player deviates, Player 1 plays T and Player 2 plays L. If anyone deviates, then each plays M thereafter.

3. **Stage Game:**

	L	M	R
T	$2, -1$	$0, 0$	$-1, 2$
M	$0, 0$	$0, 0$	$0, 0$
B	$-1, 2$	$0, 0$	$2, -1$

   Strategy profile: Until some player deviates, Player 1 plays T and Player 2 alternates between L and R. If anyone deviates, then each play M thereafter.

4. **Stage Game:**

	A	B
A	$2, 2$	$1, 3$
B	$3, 1$	$0, 0$

> **Strategy profile:** The play depends on three states. In state S_0, each player plays A; in states S_1 and S_2, each player plays B. The game start at state S_0. In state S_0, if each player plays A or if each player plays B, they stay at S_0, but if a player i plays B while the other is playing A, then they switch to state S_i. At any S_i, if player i plays B, they switch to state S_0; otherwise they stay at state S_i.

Solution. Part 1: This is a version of Tit–for–tat; it is not a subgame perfect Nash equilibrium. (Make sure that you can show this quickly! at this point.)

Part 2: This is a subgame perfect Nash equilibrium. After the deviation, the players play a Nash equilibrium forever. Hence, we only need to check that no player has any incentive to deviate on the path of equilibrium. Player 1 has clearly no incentive to deviate. If Player 2 deviates, she gets 2 in the current period and gets zero thereafter. If she sticks to her equilibrium strategy, then she gets 1 forever. The present value of this is $1/(1-\delta) > 2$. Therefore, Player 2 doesn't have any incentive to deviate, either.

Part 3: It is subgame perfect. Since (M, M) is a Nash equilibrium of the stage game, we only need to check if any player wants to deviate at a history in which Player 1 plays T and Player 2 alternates between L and R throughout. In such a history, the average value of Player 1 is

$$V_{1L} = 2 - \delta = 1.01$$

if Player 2 is to play L and

$$V_{1R} = 2\delta - 1 = 0.98$$

if Player 2 is to play R. In the case Player 2 is to play L, Player 1 cannot gain by deviating. In the case Player 2 is to play R, Player 1 can get at most gets

$$2(1-\delta) + 0 = 0.02$$

by deviating to B. Since $0.02 < 0.98$, she has no incentive to deviate. The only possible profitable deviation for Player 2 is to play R when she is supposed to play L. In that contingency, if she follows the strategy she gets $V_{1R} = 0.98$; if she deviates, she gets only $2(1-\delta) + 0 = 0.02$.

Part 4: It is not subgame-perfect. At state S_2, Player 2 is to play B, and the state in the next round is S_0 no matter what Player 1 plays. In that case, Player 1 would gain by deviating and playing A (in state S_2).

12.4 Exercises

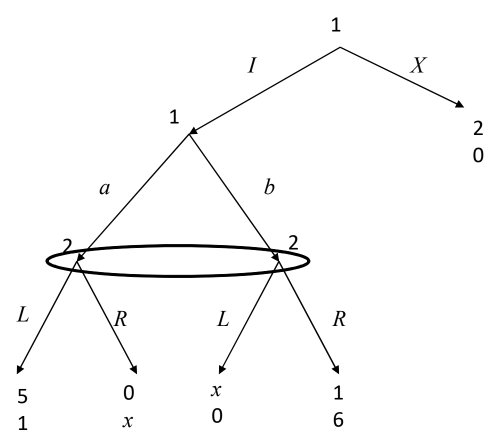
Exercise 12.12. Consider the infinitely repeated game with discount factor $\delta = 0.99$ and the following stage game (in which the players are trading favors):

	Give	Keep
Give	$1, 1$	$-1, 2$
Keep	$2, -1$	$0, 0$

1. Find a subgame-perfect Nash equilibrium under which the average expected payoff of Player 1 is at least 1.33. Verify that your strategy profile is indeed a subgame-perfect Nash equilibrium.

2. Find a subgame-perfect Nash equilibrium under which the average expected payoff of Player 1 is at least 1.49. Verify that your strategy profile is indeed a subgame-perfect Nash equilibrium.

Exercise 12.13. Consider the 100-times repeated game with the following stage game:



where x is either 0 or 6.

1. Find the set of **pure-strategy** subgame-perfect equilibria of the **stage game** for each $x \in \{0, 6\}$.

2. Take $x = 6$. What is the highest payoff Player 2 can get in a subgame-perfect Nash equilibrium of the **repeated game**?

3. Take $x = 0$. Find a subgame-perfect Nash equilibrium of the **repeated game** in which Player 2 gets more than 300 (i.e. more than 3 per day on average)?

Exercise 12.14. Consider an infinitely repeated game in which the stage game is as in the previous problem. Take the discount factor $\delta = 0.99$ and $x = 6$. For each strategy profile below, check whether it is a subgame-perfect Nash equilibrium.

1. They play (Ia, L) everyday until somebody deviates; they play (Xb, R) thereafter.

2. There are three states: A, $P1$, and $P2$, where the play is (Ia, L), (Ia, R), and (Ib, L), respectively. The game starts at state A. After state A, it switches to state $P1$ if the play is (Ib, L) and to state $P2$ if the play is (Ia, R); it stays in state A otherwise. After states $P1$ and $P2$, it switches back to state A regardless of the play.

Exercise 12.15. Consider an infinitely repeated game in which the discount factor is $\delta = 0.9$ and the stage game is

	a	b	c
w	$4, 4$	$0, 5$	$0, 0$
x	$5, 0$	$3, 3$	$-1, 0$
y	$2, 2$	$1, 1$	$-2, 0$
z	$0, 0$	$0, -1$	$-3, -2$

For each payoff vector below (u, v), find a subgame-perfect Nash equilibrium of the repeated game in which the average discounted payoff is (u, v). Verify that the strategy profile you identified is indeed a subgame-perfect Nash equilibrium.

1. $(u, v) = (4, 4)$.

2. $(u, v) = (2, 2)$.

Exercise 12.16. Consider the infinitely repeated game with the stage game in the previous problem and the discount factor $\delta \in (0, 1)$. Find the conditions on the discount factor for which the following strategy profile is a subgame-perfect Nash equilibrium.

There are 4 states: (w, a), (x, a), (w, b), and (z, c). At each state (s_1, s_2), the play is (s_1, s_2). The game starts at state (w, a). For any t with (s_1, s_2), the state at $t + 1$ is

$$\begin{array}{ll} (w,a) & \text{if the play at } t \text{ is } (s_1, s_2) \\ (x,a) & \text{if the play at } t \text{ is } (s_1, s_2') \text{ for some } s_2' \neq s_2 \\ (w,b) & \text{if the play at } t \text{ is } (s_1', s_2) \text{ for some } s_1' \neq s_1 \\ (z,c) & \text{if the play at } t \text{ is } (s_1', s_2') \text{ for some } s_1' \neq s_1 \text{ and } s_2' \neq s_2. \end{array}$$

Exercise 12.17. Consider the infinitely repeated game with discount factor $\delta \in (0, 1)$ and the stage game

	x	y	z
x	$0, 0$	$5, 1$	$2, 0$
y	$1, 5$	$4, 4$	$1, 0$
z	$0, 2$	$0, 1$	$0, 0$

For each strategy profile s below, specify the range of δ under which s is a subgame-perfect Nash equilibrium, and verify that this is indeed the case. (The range may be empty.)

1. There are two states Y and Z. The game starts at state Y. Each player is to play y in state Y and z in state Z. If the state is Y and the play is (y, y) or the state is Z and the play is (z, z), then the state in the next date is Y; otherwise the state in the next date is Z.

2. Each player is to play y at the beginning and play what the other player played in the previous period thereafter.

Exercise 12.18. Consider the infinitely repeated game with discount factor δ and the stage game

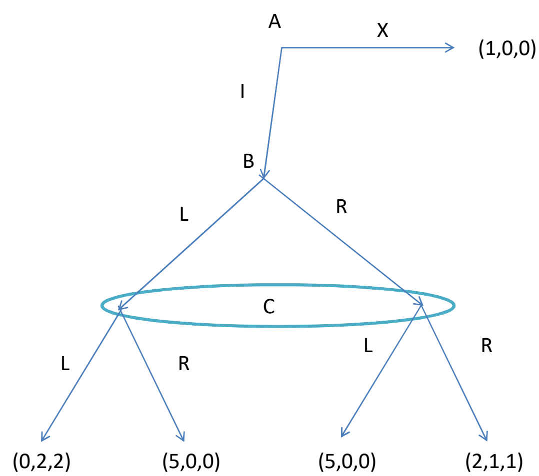
	a	b	c
a	$3, 1$	$-1, 2$	$0, 0$
b	$-1, 2$	$2, -1$	$0, 0$
c	$0, 0$	$0, 0$	$-1, -1$

For each strategy profile below, determine the range of δ (and possibly the other parameters specified in the strategy profile) under which the strategy profile is a subgame-perfect Nash equilibrium; verify that it is indeed a SPNE for each such δ:

1. There are two states: a and c. In any state s, each player is to play s. The game starts at state a. At any state s, if both players play s, then the state in the next period is a. Otherwise, the state in the next period is c. (Determine δ.)

2. Play (a, a) until somebody deviates; play mixed stage-game strategy profile σ thereafter. (Determine δ and σ.)

Exercise 12.19. Consider the $\bar{t}$-times repeated game with the following stage game.



1. For $\bar{t} = 2$, what is the largest payoff A can get in a subgame-perfect Nash equilibrium in pure strategies?

2. For $\bar{t} > 2$, find a subgame-perfect Nash equilibrium in which the payoff of A is at least $5\bar{t} - 6$.

Exercise 12.20. For the stage game in the previous problem, consider the infinitely repeated game with discount factor $\delta \in (0, 1)$. For each of the strategy profile below, find the range of δ under which the strategy profile is a subgame-perfect Nash equilibrium. (The range may be empty.)

1. The play is (I, L, R) if (I, L, R) has been played at all previous days, and (X, L, L) is played otherwise.

2. A plays I and B and C rotate between (L, R), (R, L), and (R, R) until somebody deviates; they play (X, L, L) thereafter.

3. A plays I, B plays L, and C plays L on even dates and R on odd dates until somebody deviates from this; they play (X, L, L) thereafter.

4. A plays I, B plays L, and C plays L on even dates and R on odd dates until somebody deviates from this; A plays X while B and C play the mixed strategy that puts probability $1/10$ on L (and probability $9/10$ on R) thereafter.

5. A always plays I. B and C both play R until somebody deviates and play L thereafter.

Exercise 12.21. Seagulls love shellfish. In order to break the shell, they need to fly high up and drop the shellfish. The problem is the other seagulls on the beach are kleptoparasites, and they steal the shellfish if they can reach it first. This question tells the story of two seagulls, named Irene and Jonathan, who live in a crowded beach where it is impossible to drop the shellfish and get it before some other gull steals it. The possible dates are $t = 0, 1, 2, 3, \ldots$ with no upper bound. Everyday, simultaneously Irene and Jonathan choose one of the two actions: "Up" or "Down". Up means to fly high up with the shellfish and drop it next to the other sea gull's nest, and Down means to stay down in the nest. Up costs $c > 0$, but if the other seagull is down, it eats the shellfish, getting payoff $v > c$. That is, we consider the infinitely repeated game with the following stage game

	Up	Down
Up	$-c, -c$	$-c, v$
Down	$v, -c$	$0, 0$

and discount factor $\delta \in (0, 1)$.[12] For each strategy profile below, find the set of discount factors δ under which the strategy profile is a subgame-perfect Nash equilibrium.

1. Irrespective of the history, Irene plays Up in the even dates and Down in the odd dates; Jonathan plays Up in the odd dates and Down in the even dates.

2. Irene plays Up in the even dates and Down in the odd dates while Jonathan plays the other way around until someone fails to go Up in a day that he or she is supposed to do so. They both stay Down thereafter.

[12] Evolutionarily speaking, the discounted sum is the fitness of the genes, which determine the behavior.

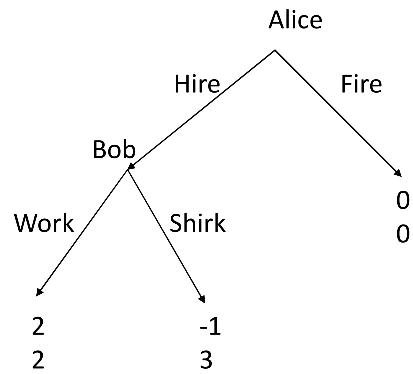


Figure 12.6:

3. For n days Irene goes Up and Jonathan stays Down; in the next n days Jonathan goes Up and Irene stays Down. This continues back and forth until someone deviates. They both stay Down thereafter.

4. Irene goes Up on "Sundays", i.e., at $t = 0, 7, 14, 21, \ldots$, and stays Down on the other days, while Jonathan goes up everyday except for Sundays, when he rests Down, until someone deviates; they both stay Down thereafter.

5. At $t = 0$, Irene goes Up and Jonathan stays Down, and then they alternate. If a seagull i fails to go Up at a history when i is supposed to go Up, then the next day i goes Up and the other seagull stays Down, and they keep alternating thereafter until someone fails to go Up when it is supposed to do so. (For example, given the history, if Irene is supposed to go Up at t but stays Down, then Irene goes Up at $t + 1$, Jonathan goes Up at $t + 2$, and so on. If Irene stays down again at $t + 1$, then she is supposed to go up at $t + 2$, and Jonathan is supposed to go at $t + 3$, etc.)

Exercise 12.22. Consider the infinitely repeated game, between Alice and Bob, with the stage game in Figure 12.6. The discount factor is $\delta = 0.9$. (Fire does not mean that the game ends.) For each strategy profile below, check if it is a subgame-perfect Nash equilibrium. If it is not a SPNE for $\delta = 0.9$, find the set of discount factors δ under which it is a SPNE.

1. Alice Hires if and only if there is no Shirk in the history. Bob Works if and only if there is no Shirk in the history.

2. Alice Hires unless Bob (was hired and) Shirked in the previous period, in which case she Fires. Bob always Works.

3. There are three states: Employment, Punishment for Alice, and Punishment for Bob. In the Employment state, Alice Hires and Bob Works. In the Punishment state for Alice, Alice Hires but Bob Shirks. In the Punishment state for Bob, Alice Fires, and Bob would have worked if Alice Hired him. The game starts in Employment state. If only one player fails to play the action prescribed to her in a given period, then the the state in the next period is the Punishment state for that player; it is Employment otherwise.

Exercise 12.23. Consider the infinitely repeated game with the following stage game

	Hawk	Dove
Hawk	$0,0$	$4,1$
Dove	$1,4$	$3,3$

and discount factor $\delta = 0.99$. For each strategy profile below check if it is a subgame-perfect Nash equilibrium.

1. There are two states: Cooperation and Fight. The game starts in the Cooperation state. In the Cooperation state, each player plays Dove. If both players play Dove, then they remain in the Cooperation state; otherwise they go to the Fight state in the next period. In the Fight state, both play Hawk, and they go back to the Cooperation state in the following period (regardless of the actions).

2. There are three states: Cooperation, P_1 and P_2. The game starts in the Cooperation state. In the Cooperation state, each player plays Dove. If they play (Dove, Dove) or (Hawk, Hawk), then they remain in the Cooperation state in the next period. If player i plays Hawk while the other player plays Dove, then in the next period they go to the state P_i. In state P_i, player i plays Dove while the other player plays Hawk; they then go back to Cooperation state (regardless of the actions).

Exercise 12.24. Alice has two sons, Bob and Colin. Every day, she is to choose between letting them play with the toys ("Play") or make them visit their grandmother ("Visit"). If she make them visit their grandmother, each of them gets 1. If she lets them play, then Bob and Colin simultaneously choose between Grab and Share, which leads to the payoffs as in the following table, where Bob and Colin choose rows and columns, respectively, and the payoffs are ordered alphabetically:

	Grab	Share
Grab	$-1, -1, -1$	$-2, 3, -2$
Share	$-2, -2, 3$	$2, 2, 2$

Consider the infinitely repeated game with the above game is the stage game and the discount factor is $\delta = 0.9$. For each strategy profile below check if it is a subgame-perfect Nash equilibrium. Show your work.

1. There are three states: Share, P_{BOB} and P_{COLIN}. In state Share, Alice lets them play, and Bob and Colin both share. In state P_i, Alice lets them play, and player i shares and his brother grabs. The game starts in state Share. In any state, if one son i deviates while the other plays what he is supposed to play, then the state in the next period is P_i; the state in the next period is Share otherwise.

2. There are two states: Play and Visit. The game starts in the Play state. In the Play state, Alice lets them play, and both sons share. In the Play state, if everybody does what they are supposed to do, they remain in the Play state; they go to the Visit state next day otherwise. In the Visit state, Alice makes them visit their grandmother, and they would both Grab if she let them play. In the Visit state, they automatically go back to the Play state next day.

Exercise 12.25. Alice has a restaurant, and Bob is a potential customer. Each day Alice is to decide whether to use high quality supply (High) or low quality supply (Low) to make the food, and Bob is to decide whether to buy food from Alice or skip (without knowing the quality of the food).[13] The payoffs for a given day are as follows

	Buy	Skip
High	$p-1, 3-p$	$-1, 0$
Low	$p, -p$	$0, 0$

where $p \in [1, 3]$ is a known number, representing the price of the food. The discount factor is $\delta = 0.99$. For each strategy profile below, find the range of $p \in [1, 3]$ for which the strategy profile is a subgame-perfect Nash equilibrium.

1. There are two states: Trade and No-trade. The game starts at the Trade state. In the Trade state, Alice uses High quality supply, and Bob Buys. If in the Trade state Alice uses Low quality supply, then they go to the No-Trade state, in which for n days Alice uses Low quality supply and Bob Skips. At the end of n days, independent of what happens, they go back to the Trade state.

2. Alice is to use High quality supply in the even days, $t = 0, 2, 4, \ldots$, and Low quality supply in the odd days, $t = 1, 3, 5, \ldots$; Bob is to Buy everyday. If anyone deviates from this program, then, for the rest of the game, Alice uses Low quality and Bob Skips.

3. Everyday Alice uses High quality supply. Bob buys the product in the first day. Afterwards, Bob buys the product if and only if Alice has used High quality supply in the previous day.

4. There are two states: Trade and Punishment. The game starts at Trade state. In the Trade state, Alice uses High quality supply, and Bob Buys. In the Trade state if Alice uses Low quality, then they go to the Punishment state. In the Punishment state, Alice uses High quality supply, and Bob Skips. In the Punishment state, if Alice uses Low quality supply or Bob Buys, then they remain in the Punishment state; otherwise they go to the Trade state.

Exercise 12.26. In an eating club, there are $n > 2$ members. Each day, each member i is to decide how much to eat, denoted by y_i, and the payoff of i for that day is

$$\sqrt{y_i} - \frac{y_1 + \cdots + y_n}{n}.$$

[13] Bob knows the quality of the food in the previous days even if he did not buy the food in those days.

For the discount factor $\delta = 0.99$, check if either of the following strategy profiles is a subgame-perfect Nash equilibrium.

1. Each player eats $y = 1/4$ units until somebody eats more than $1/4$; each eats $y = n^2/4$ units thereafter.

2. Each player eats $y = 1/4$ units until somebody eats more than $1/4$; each eats $y = n^2$ units thereafter.

Exercise 12.27. Each day Alice and Bob receive 1 dollar. Alice makes an offer x to Bob, and Bob accepts or rejects the offer, where $x \in \{0.01, 0.02, \ldots, 0.98, 0.99\}$. If Bob accepts the offer Alice gets $1-x$ and Bob gets x. If Bob rejects the offer, then they both get 0. Find the values of δ for which the following is a subgame-prefect equilibrium, where $\bar{x} \in \{0.01, 0.02, \ldots, 0.98, 0.99\}$ is fixed.

`At` $t = 0$, `Alice offers` $\bar{x}$ `and Bob accepts Alice's offer,` x`, if and only if` $x \geq \bar{x}$. `They keep doing this until Bob deviates from this program (i.e. until Bob accepts an offer` $x < \bar{x}$`, or Bob rejects an offer` $x \geq \bar{x}$`). Thereafter, Alice offers` $x = 0.01$ `and Bob accepts any offer.`

Exercise 12.28. Consider a Firm and a Worker. The firm first decides whether to pay a wage $w > 0$ to the worker (hire him), and then the worker is to decide whether work, which costs him $c > 0$ and produces π to the firm where $\pi > w > c$. The payoffs are as follows:

	Firm	Worker
pay, work	$\pi - w$	$w - c$
pay, shirk	$-w$	w
don't pay, work	π	$-c$
don't pay, shirk	0	0

1. Find all Nash equilibria.

2. Now consider the game this stage game is repeated infinitely many times and the players discounts the future with δ. The following are strategy profiles for this repeated game. For each of them, Check if it is a subgame-perfect Nash equilibrium for large values of δ, and if so, find the lowest discount rate that makes the strategy profile a subgame-perfect Nash equilibrium.

(a) No matter what happens, the firm always pays and the worker works.

(b) At any time t, the worker works if he is paid at t, and the firm always pays.

(c) At $t = 0$, the firm pays and the worker works. At any time $t > 0$, the firm pays if and only if the worker worked at all previous dates, and the worker works if and only if he has worked at all previous dates.

(d) At $t = 0$, the firm pays and the worker works. At any time $t > 0$, the firm pays if and only if the worker worked at all previous dates at which the firm paid, and the worker works if and only if he is paid at t and he has worked at all previous dates at which he was paid.

(e) There are two states: Employment, and Unemployment. The game starts at Employment. In this state, the firm pays, and the worker works if and only if he has been paid at this date. If the worker shirks we go to Unemployment state; otherwise we stay in Employment. In Unemployment the firm does not pay and the worker shirks. After $T > 0$ days of Unemployment we always go back to Employment. (Your answer should cover each $T > 0$.)

Exercise 12.29. Consider the following Stage Game: Alice and Bob simultaneously choose contributions $a \in [0,1]$ and $b \in [0,1]$, respectively, and get payoffs $u_A = 2b - a$ and $u_B = 2a - b$, respectively.

1. Find the set of rationalizable strategies in the Stage Game above.

2. Consider the infinitely repeated game with the Stage Game above and with discount factor $\delta \in (0,1)$. For each δ, find the maximum (a^*, b^*) such that there exists a subgame-perfect Nash equilibrium of the repeated game in which Alice and Bob contribute a^* and b^*, respectively, on the path of equilibrium.

3. In Part 2, now assume that at the beginning of each period t one of the players (Alice at periods $t = 0, 2, 4, \ldots$ and Bob at periods $t = 1, 3, 5, \ldots$) offers a stream of contributions $\vec{a} = (a_t, a_{t+1}, \ldots)$ and $\vec{b} = (b_t, b_{t+1}, \ldots)$ for Alice and Bob, respectively, and the other player accepts or rejects. If the offer is accepted then the game ends leading the automatic contributions $\vec{a} = (a_t, a_{t+1}, \ldots)$ and $\vec{b} = (b_t, b_{t+1}, \ldots)$ from period t on. If the offer is rejected, they play the Stage Game and proceed

to the next period. Find (a_A, b_A), (a_B, b_B), and $\left(\hat{a}, \hat{b}\right)$ such that the following is a subgame-perfect Nash equilibrium:

s^* : When it is Alice's turn, Alice offers $(a_A, a_A, \ldots)$ and $(b_A, b_A, \ldots)$ and Bob accepts an offer $(\vec{a}, \vec{b})$ if and only if $(1-\delta)\left[2a_t - b_t + \delta\left(2a_{t+1} - b_{t+1}\right) + \cdots\right] \geq 2a_A - b_A$. When it is Bob's turn, Bob offers $(a_B, a_B, \ldots)$ and $(b_B, b_B, \ldots)$ and Alice accepts an offer $(\vec{a}, \vec{b})$ if and only if $(1-\delta)\left[2b_t - a_t + \delta\left(2b_{t+1} - a_{t+1}\right) + \cdots\right] \geq 2b_A - a_A$. If there is no agreement, in the stage game they play $\left(\hat{a}, \hat{b}\right)$.

Verify that s^* is a subgame perfect equilibrium for the values that you found. (If you find it easier, you can consider only the constant streams of contributions $\vec{a} = (a, a, \ldots)$ and $\vec{b} = (b, b, \ldots)$.)

Exercise 12.30. Below, there are pairs of stage games and strategy profiles. For each pair, check whether the strategy profile is a subgame-perfect Nash equilibrium of the game in which the stage game is repeated infinitely many times. Each agent tries to maximize the discounted sum of his expected payoffs in the stage game, and the discount rate is $\delta = 0.99$. (Clearly explain your reasoning in each case.)

1. **Stage Game:** There are $n > 2$ players. Each player, simultaneously, decides whether to contribute \$1 to a public good production project. The amount of public good produced is $y = (x_1 + \cdots + x_n)/2$, where $x_i \in \{0, 1\}$ is the level of contribution for player i. The payoff of a player i is $y - x_i$.

   Strategy profile: Each player contributes, choosing $x_i = 1$, if and only if the amount of public good produced at each previous date is greater than $n/4$; otherwise each chooses $x_i = 0$. (According to this strategy profile, each player contributes in the first period.)

2. **Stage Game:**

	S	R
S	$6, 6$	$0, 4$
R	$4, 0$	$4, 4$

   Strategy profile: Each player plays S until someone deviates. If a player deviates, then he is to keep playing S and the other player plays R forever.

3. **Stage Game:**

	S	R
S	$6,6$	$0,4$
R	$4,0$	$4,4$

Strategy profile: Each player plays S until someone deviates. If a player deviates, then each player plays R forever.

4. **Stage Game:** Player 1 decides whether to give a \$100 to Player 2. If Player 1 gives \$100, then Player 2 decides whether to provide a service to Player 1, which is worth \$200 for Player 1 and costs \$50 to Player 2.

   Strategy Profile: There are two states: Trade and No trade. The game starts in Trade state. If Player 1 pays 100, and Player 2 does not provide the service, then they go to No trade state and stay there for <u>two</u> periods. In No trade period, Player 1 does not give any money, and Player 2 does not provide service (if Player 1 pays him \$100).

Exercise 12.31. Consider the infinitely repeated game with the Prisoners' Dilemma game

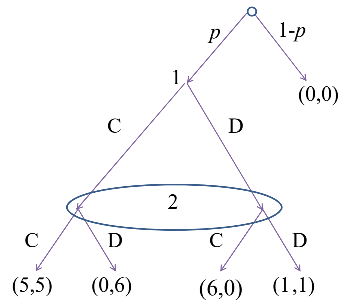
	C	D
C	$4,4$	$0,5$
D	$5,0$	$1,1$

as its stage game and with discount factor δ.

1. What is the lowest discount rate δ such that there exists a subgame perfect equilibrium in which each player plays C on the path of equilibrium play? [Hint: Note that a player can always guarantee himself an average payoff of 1 by playing D forever.]

2. For sufficiently large values of δ, construct a subgame-perfect Nash equilibrium in which any agent's action at any date t only depends on the play at dates $t-1$ and $t-2$, and in which each player plays C on the path of equilibrium play.

Exercise 12.32. Consider the infinitely repeated game with discount factor $\delta \in (0,1)$ and the stage game



(Note that, in the stage game, the players play a prisoners' dilemma game with probability p, where p is independent of the past history; they simply get 0 in the alternative case.) For each strategy profile s below, specify the range of δ under which s is a subgame-perfect Nash equilibrium, and verify that this is indeed the case. (The range may be empty.)

1. On the path, they play (C, D) for the first 100 rounds and then they play (C, C) forever. If any player deviates, then they play (D, D) forever.

2. For this part, take $p = 1$. For each player i and for each history, player i plays D if and only if the number of times i has played C is strictly larger than the number of times the other player has played C. (He plays C at the beginning.)

Exercise 12.33. Consider the infinitely repeated game with discount factor $\delta \in (0,1)$ and the following stage game with 2 players. Simultaneously each player i contributes $x_i \geq 0$ towards producing a public good of amount $y = x_1 + x_2$. The payoff of player i is $y - x_i^2$.

1. Compute the largest $\hat{x}$ for which the following is a subgame-perfect Nash equilibrium; verify your answer.

   `Each player contributes` $\hat{x}$ `until somebody deviates; each produces` $\bar{x}$ `thereafter.`

2. Find the conditions on a and b under which the following is a subgame-perfect Nash equilibrium; show your work.

   `There are three states:` C, P_1, `and` P_2. `At state` C, `each player is to contribute` a. `At state` P_i, `player` i `is to contribute` b `while the other` `player contributes 0. The game starts at state` C. `Given any state, the next state is` C `if either both players contribute the amount they are supposed to contribute at that state or both players deviate; the next state is` P_i `if player` i `unilaterally deviates.`

Exercise 12.34. Consider the game in Figure 12.7, as the stage game.

1. Consider twice-repeated game. Find all action profiles played in the first round in a pure-strategy subgame-perfect Nash equilibrium; identify the SPNE for each profile.

2. Consider the infinitely repeated game with discount factor $\delta \in (0, 1)$.

   (a) For each player, find the smallest average value he can get in a subgame-perfect Nash equilibrium.

   (b) Find the range of δ for which (M, m) is played in every period on the path of a SPNE.

   (c) For $\delta = 0.99$, find a SPNE in which the average value of player 2 is at least 5.5.

Exercise 12.35. Consider the infinitely repeated Prisoners' Dilemma game with discount factor $\delta = 0.9$ and the stage game

	C	D
C	$2, 2$	$0, 6$
D	$6, 0$	$1, 1$

1. Find a subgame perfect Nash equilibrium s under which both players play C on the path throughout. Verify that s is a subgame perfect Nash equilibrium.

2. Find a subgame perfect Nash equilibrium s under which each player's average payoff is at least 2.8. Verify that s is a subgame perfect Nash equilibrium.

Exercise 12.36. Consider the infinitely repeated game with discount factor $\delta \in (0,1)$ and the stage game

	a	b	c
a	$5,5$	$0,6$	$0,0$
b	$6,0$	$3,3$	$2,0$
c	$0,0$	$0,2$	$0,0$

1. Compute the pure-strategy minmax payoff for each player.

2. For each strategy profile s below, find the range of δ for which the following is a subgame-perfect Nash equilibrium; show your work. (The range can be empty.)

   s^* : Both players play (a,a) until somebody deviates; they play (c,c) thereafter.

   s^{**} : There are two states: A and C. Each player plays a in state A and c in state C. The game starts at state A; at any later period, the state is A if in the previous period both player played what he was supposed to play, and the state is C otherwise.

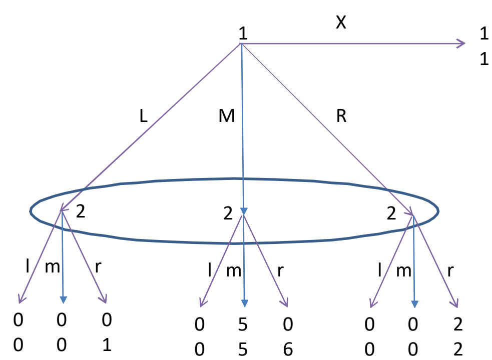


Figure 12.7: