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, he takes into account not only the implications of his actions for the current situation but also their implications for the future. 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. Such ideas are captured in the repeated games, in which a "stage game" is played repeatedly. The stage game is repeated regardless of what has been played in the previous games. 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.

12.1 Finitely-repeated games

Let $T = \{0, 1, \ldots, n\}$ be the set of all possible dates. Consider a game in which at each $t \in T$ players play a "stage game" G, knowing what each player has played in the past. Assume that the payoff of each player in this larger game is the sum of the payoffs that he obtains in the stage games. Denote the larger game by G^T.

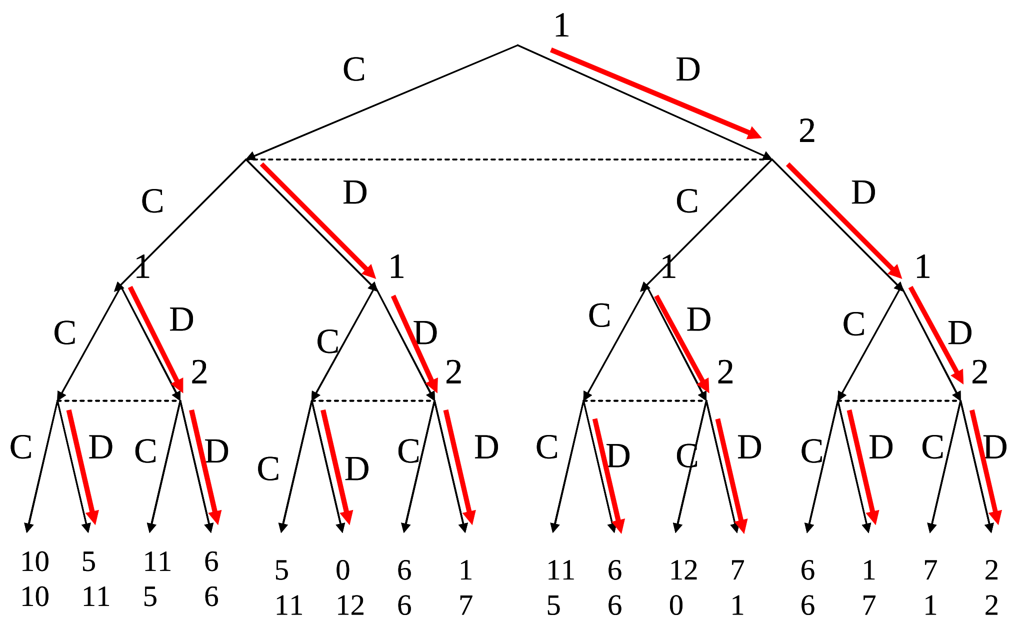
Note that a player simply cares about the sum of his payoffs at the stage games. Most importantly, at the beginning of each repetition each player recalls what each player has played in each previous play. A strategy then prescribes what player plays at each t as a function of the plays at dates 0, ..., $t-1$. More precisely, let us call the outcomes of the previous stage games a *history*, which will be a sequence $(a_0, \ldots, a_{t-1})$. 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, consider a situation in which two players play the Prisoners' Dilemma game,

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

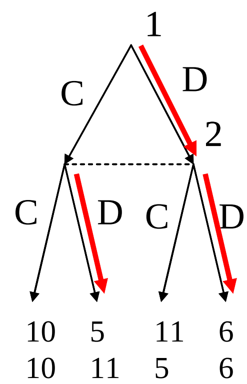
, (12.1)

twice. In that case, $T = \{0, 1\}$ and G is the Prisoners' Dilemma game. The repeated game, G^T, can be represented in the extensive-form as



Now at date $t = 1$, a history is a strategy profile of the Prisoners' Dilemma game, indicating what has been played at $t = 0$. There are four histories at $t = 1$: (C, C), (C, D), (D, C), and (D, D). A strategy is to describe what the player plays at $t = 0$, and what he plays at each of these four histories. (There are 5 actions to be determined.) This is rather clear in the extensive-form game above.

Let us compute the subgame-perfect equilibrium of G^T; the equilibrium is depicted in the figure. G^T has four proper subgames, each corresponding to the last-round game after a history of plays in the initial round. For example, after (C, C) in the initial round, we have subgame



where we add 5 to each player's payoffs, corresponding to the payoff that he gets from playing (C, C) in the first round. Recall that adding a constant to a player's payoff does not change the preferences in a game, and hence the set of equilibria in this game is the same as the original Prisoners' Dilemma game, which possesses the unique Nash equilibrium of (D, D). This equilibrium is depicted in the figure. Likewise, in each proper subgame, we add some constant to the players' payoffs, and hence we have (D, D) as the unique Nash equilibrium at each of these subgames.

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 we add 1 to each player's payoffs, accounting for his payoff in the last round. The unique equilibrium of this reduced game is (D, D). This leads to a unique subgame-perfect equilibrium: *At each history, each player plays D.*

What would happen for arbitrary n? The answer remains the same. In the last day, n, 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 $n - 1$ 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 reduced game for any subgame starting at $n-1$ is

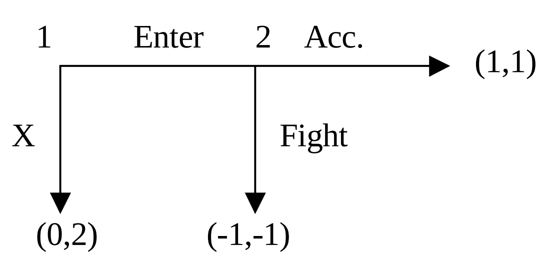
	C	D
C	$5+1+\pi_1, 5+1+\pi_2$	$0+1+\pi_1, 6+1+\pi_2$
D	$6+1+\pi_1, 0+1+\pi_2$	$1+1+\pi_1, 1+1+\pi_2$

where π_1 is the sum of the payoffs of i from the previous plays at dates $0, \ldots, n-2$. Here we add π_i for these payoffs and 1 for the last round payoff, all of which are independent of what happens at date $n-1$. This is another version of the Prisoner's dilemma, which has the unique Nash equilibrium of (D, D). Proceeding in this way all the way back to date 0, we find out that there is a unique subgame-perfect 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.

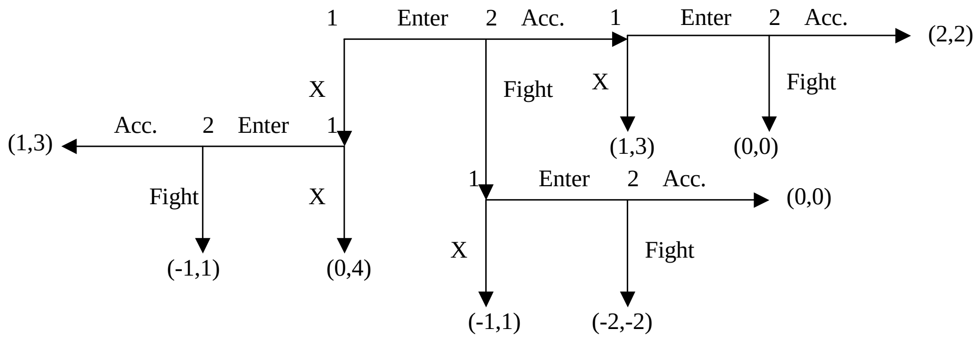
Theorem 12.1 *Let n be finite and assume that G has a unique subgame-perfect equilibrium s^*. Then, G^T has a unique subgame-perfect equilibrium, and according to this equilibrium s^* is played at each date independent of the history of the previous plays.*

The proof of this result is left as a straightforward exercise. The result can be illustrated by another important example. Consider the following Entry-Deterrence game, where an entrant (Player 1) decides whether to enter a market or not, and the incumbent (Player 2) decides whether to fight or accommodate the entrant if he enters.



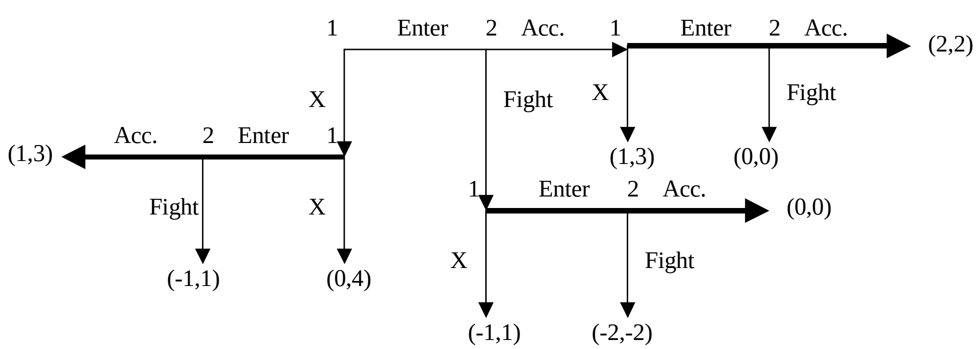
(12.2)

Consider the game where the Entry-Deterrence game is repeated twice, and all the previous actions are observed. This game is depicted in the following figure.

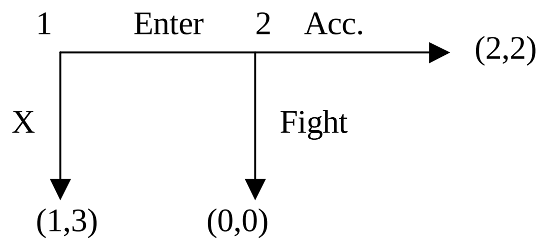


As depicted in the extensive form, in the repeated game, at $t = 1$, there are three possible histories: X, $(Enter, Acc)$, and $(Enter, Fight)$. A strategy of Player 1 assigns an action, which has to be either Enter or X, to be played at $t = 0$ and action to be played at $t = 1$ for each possible outcome at $t = 0$. In total, we need to determine 4 actions in order to define a strategy for Player 1. Similarly for Player 2.

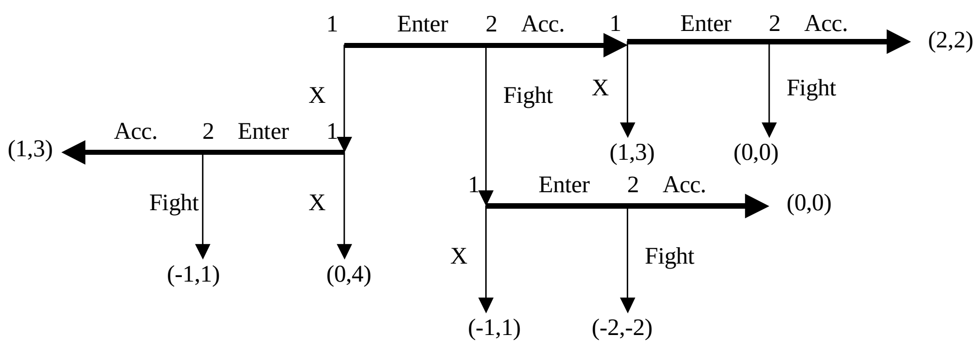
Note that after the each outcome of the first play, the Entry-Deterrence game is played again, where the payoff from the first play is added to each outcome. Since a player's preferences do not change when we add a number to his utility function, each of the three games played on the second "day" is the same as the stage game (namely, the Entry-Deterrence game above). The stage game has a unique subgame perfect equilibrium, where the incumbent accommodates the entrant and the entrant enters the market. In that case, each of the three games played on the second day has only this equilibrium as its subgame perfect equilibrium. This is depicted in the following.



Using backward induction, we therefore reduce the game to the following.



Notice that we simply added the unique subgame-perfect equilibrium payoff of 1 from the second day to each payoff in the stage game. Again, adding a constant to a player's payoffs does not change the game, and hence the reduced game possesses the subgame-perfect equilibrium of the stage game as its unique subgame perfect equilibrium. Therefore, the unique subgame perfect equilibrium is as depicted below.



This can be generalized for arbitrary n as above. All these examples show that in certain important games, no matter how high the stakes are in the future, the considerations about the future will not affect the current actions, as the future outcomes do not depend on the current actions. In the rest of the lectures we 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.

Indeed, if the stage game has more than one equilibrium, then in the repeated game we may have some subgame-perfect equilibria where, in some stages, players play some actions that are not played in any subgame-perfect equilibrium of the stage game. This is because the equilibrium to be played on the second day can be conditioned to the play on the first day, in which case the "reduced game" for the first day is no longer the same as the stage game, and thus may obtain some different equilibria. I will now illustrate this using an example in Gibbons. (See Exercises 1 and 2 at the end of the chapter before proceeding.)

Take $T = \{0, 1\}$ and the stage game G be

	L	M	R
A	$1,1$	$5,0$	$0,0$
B	$0,5$	$4,4$	$0,0$
C	$0,0$	$0,0$	$3,3$

Notice that a strategy in a stage game prescribes what the player plays at $t = 0$ and what he plays at $t = 1$ conditional on the history of the play at $t = 0$. There are 9 such histories, such as (A, L), (B, R), etc. A strategy of Player 1 is defined by determining an action (A,B, or C) for $t = 0$, and determining an action for each of these histories at $t = 1$ (There will be 10 actions in total.) Consider the following strategy profile:

> Player 1: play B at $t = 0$; at $t = 1$, play C if (B, M) played at $t = 0$, and play A otherwise.
>
> Player 2: play M at $t = 0$; at $t = 1$, play R if (B, M) played at $t = 0$, and play L otherwise.

According to this equilibrium, at $t = 0$, players play (B, M) even though $(B.M)$ is not a Nash equilibrium of the stage game. Notice that in order for a strategy profile to be subgame perfect, after each history, at $t = 1$, we must have a Nash equilibrium. Since (A, L) and (C, R) are both Nash equilibria of the stage game, this is in fact the case. Given this behavior, the first round game reduces to

	L	M	R
A	$2,2$	$6,1$	$1,1$
B	$1,6$	$7,7$	$1,1$
C	$1,1$	$1,1$	$4,4$

Here, we add 3 to the payoffs at (B, M) (for it leads to (C, R) in the second round) and add 1 for the payoffs at the other strategy profiles, for they lead to (A, L) in the second round. Clearly, (B, M) 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 equilibria in the second round conditional on the behavior in the first round, and the players may play a non-equilibrium (or even irrational) strategies in the first round, if those strategies lead 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 $((B,M),(B,M))$ 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 (B,M) is not a Nash equilibrium of the stage game $((B,M),(B,M))$ cannot emerge in any Nash equilibrium, let alone in a subgame-perfect Nash equilibrium.

Can $((B,M),(A,L))$ be an outcome of a subgame-perfect Nash equilibrium in pure strategies? The answer is No. Although (A,L) is a Nash equilibrium of the stage game, in a subgame-perfect Nash equilibrium, a Nash equilibrium of the stage game must be played after every play in the first round. In particular, after (A,M), the play is either (A,L) or (C,R), yielding 6 or 8, respectively for Player 1. Since he gets only 5 from $((B,M),(A,L))$, he has an incentive to deviate to A in the first period. (What about if we consider mixed subgame-perfect Nash equilibria or non-subgame-perfect Nash equilibria?)

Can $((B,L),(C,R))$ 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 (A,L) is played after every play of the period except for (B,L). In that case, the reduced game for the first period is

	L	M	R
A	$2,2$	$6,1$	$1,1$
B	$3,8$	$5,5$	$1,1$
C	$1,1$	$1,1$	$4,4$

Since (B,L) 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 (B,L) in the first round; in the second round, play (C,R) if (B,L) is played in the first round and play (A,L) otherwise.

As an exercise, check also if $((C,L),(C,R))$ or $((C,L),(B,M),(C,R))$ can be an outcome of a subgame-perfect Nash equilibrium in pure strategies (in twice and thrice repeated games, respectively).

12.2 Infinitely repeated games with observed actions

Now consider the infinitely repeated games where all the previous moves are common knowledge at the beginning of each stage. That is, in the previous section take $T = \{0,1,2,\ldots\}$ as the set of natural numbers instead of $T = \{0,1,\ldots,n\}$. The game continues indefinitely regardless of what players play along the way.

For the technically oriented students, the following must be noted. It is implicitly assumed throughout the chapter that in the stage game, 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.

12.2.1 Present Value calculations

In an infinitely repeated game, one cannot simply add the payoffs of each stage, as the sum becomes infinite. For these games,assume instead that players maximize the discounted sum of the payoffs from the stage games. The *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*. The *average value* is simply

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

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). 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 he 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.

12.2.2 Histories and strategies

Once again, in a repeated game, a *history* at the beginning of a given date t is the sequence of the outcomes of the play at dates $0,\ldots,t-1$. For example, in the Entry-Deterrence game, the possible outcomes of the stage game are X, $EA = (Enter, Acc)$, and $EF = (Enter, Fight)$, and the possible histories are t-tuples of these three outcomes for each t. Examples of histories are

$$XXEAXEFXXEA\cdots X \text{ and } EAXEFXEFEAXEA\cdots EA.$$

In the repeated Prisoner's Dilemma, the possible histories are t-tuples of (C,C), (C,D), (D,C), and (D,D), such as

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

where t varies. A history at the beginning of date t 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 $t = 0$. For example, in the repeated prisoners' dilemma, $((C,C),(C,D))$ is a history for $t = 2$. In the repeated entry-deterrence game, (X, EF) is a history for $t = 2$.

A strategy in a repeated game, once again, determines a strategy in the stage game for *each* history and for *each* t. The important point is that the strategy in the stage game at a given date can vary by histories. Here are some possible strategies in the repeated Prisoner's Dilemma game:

> **Grim:** 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).

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

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

Note that strategy profiles (Grim, Grim), (Naively Cooperate, Naively Cooperate) and (Tit-for-Tat, Tit-for-Tat) all lead to the same outcome path:[1]

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

Nevertheless, they are quite distinct strategy profiles. Indeed, (Naively Cooperate, Naively Cooperate) is not even a Nash equilibrium (why?), while (Grim, Grim) is a subgame-perfect Nash equilibrium for large values of δ. On the other hand, while (Tit-for-Tat, Tit-for-Tat) is a Nash equilibrium for large values of δ, it is not subgame-perfect. All these will be clear momentarily.

12.2.3 Single-deviation principle

In an infinitely repeated game, one uses the *single-deviation principle* in order to check whether a strategy profile is a subgame-perfect Nash equilibrium. In such a game, single-deviation principle takes a simple form and is applied through *augmented stage games.* Here, augmented refers to the fact that one simply augments the payoffs in the stage game by adding the present value of future payoffs under the purported equilibrium. One may also use the term *reduced game* instead of augmented stage game, interchangeably.

Augmented Stage Game (aka Reduced 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

[1] Make sure that you can compute the outcome path for each strategy profile above.

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 $u_i(z)$.

Single-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 (Single-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 at the beginning of t.*

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 single-deviation principle on infinitely repeated Entry Deterrence and Prisoners' Dilemma games.

Infinitely Repeated Entry Deterrence Towards illustrating the single-deviation principle when the stage game is dynamic, consider the infinitely repeated Entry-Deterrence game in (12.2). Consider the following strategy profile.

> At any given stage, 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 he has accommodated the entrant before.[2]

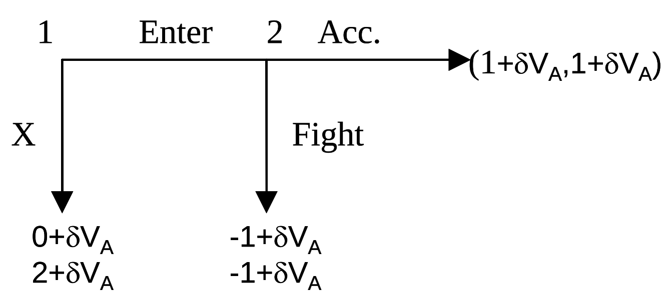
Using the single-deviation principle, we will now show that for large values of δ, this a subgame-perfect Nash equilibrium. The strategy profile puts the histories in two groups:

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.

Consequently, in the application of single-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



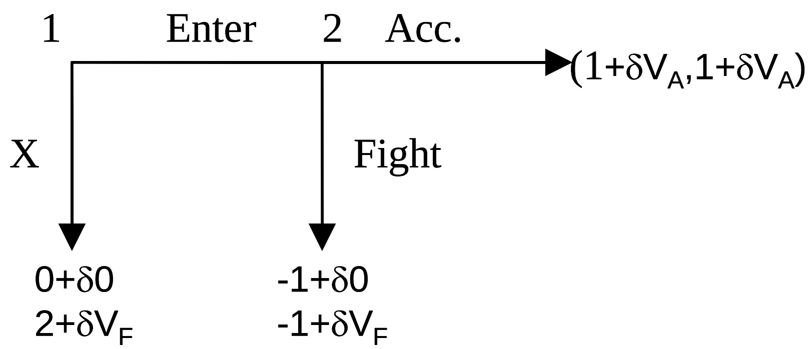
[2] This is a switching strategy, where initially incumbent fights whenever there is an entry and the entrant never enters. If the incumbent happens to accommodate an entrant, they switch to the new regime where the entrant enters the market no matter what the incumbent does after the switching, and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

For example, if the incumbent accommodates the entrant at t, his present value (at t) will be $1+\delta V_A$; and if he fights his 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 single-deviation principle, (Enter, Accommodate) must be a subgame-perfect equilibrium of the augmented stage game here. This is indeed the case, and s^* passes the single-deviation test for such histories.

Now for some 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 he enters, the incumbent fights. Single-deviation principle requires then that $(X, Fight)$ is a subgame-perfect 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[3]

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

We have considered all possible histories, and when $\delta \geq 2/3$, the strategy profile has passed the single-deviation test. Therefore, when $\delta \geq 2/3$, the strategy profile is a subgame-perfect 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 single-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 equilibrium of this game, as $1 + \delta V_A > -1 + \delta V_F$.

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 single-deviation test, one simply checks whether $s^*(h)$ is a Nash

[3] 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., $\delta \geq 2/3$.

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 (12.2). Consider the strategy profile (Grim,Grim). 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 one at some date.

First consider a Cooperation history for some t. Now if both players play C, then according to (Grim,Grim), from $t+1$ on each player will play C forever. This yields the present value of

$$V_C = 5 + 5\delta + 5\delta^2 + \cdots = 5/(1-\delta)$$

at $t+1$. If any player plays D, then from $t+1$ on, all the histories will be Defection histories and each will play D forever. This yields the present value of

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

at $t+1$. Now, at t, if they both play C, then the payoff of each player will be $5+\delta V_C$. If Player 1 plays D while Player 2 is playing C, then Player 1 gets $6+\delta V_D$, and Player 2 gets $0+\delta V_D$. Hence, the augmented stage game at the given history is

	C	D
C	$5+\delta V_C, 5+\delta V_C$	$0+\delta V_D, 6+\delta V_D$
D	$6+\delta V_D, 0+\delta V_D$	$1+\delta V_D, 1+\delta V_D$

To pass the single-deviation test, (C,C) must be a Nash equilibrium of this game.[4] (That is, we fix a player's action at C and check if the other player has an incentive to deviate.)

[4] It is important to note that we do *not* need to know all the payoffs in the reduced game. For example, for this history we only need to check if (C,C) is a Nash equilibrium of the reduced game, and hence we do not need to compute the payoffs from (D,D). In this example, it was easy to compute. In general, it may be time consuming to compute the payoffs for all strategy profiles. In that case, it will save a lot of time to ignore the strategy profiles in which more than one player deviates from the prescribed behavior at t.

This is the case if and only if

$$5 + \delta V_C \geq 6 + \delta V_D,$$

i.e.,

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

We also need to consider Defection histories. Consider a Cooperation history for some t. Now, independent of what is played at t, according to (Grim,Grim), from $t+1$ on we will have defection histories and each player will play D forever. The present value of payoffs from $t+1$ on will always be V_D. Then, the augmented stage game at this history is

	C	D
C	$5 + \delta V_D, 5 + \delta V_D$	$0 + \delta V_D, 6 + \delta V_D$
D	$6 + \delta V_D, 0 + \delta V_D$	$1 + \delta V_D, 1 + \delta V_D$

.

Single-deviation test for (Grim,Grim) requires that (D, D) is a Nash equilibrium of this game,[5] and in fact (D, D) is the only Nash equilibrium.

Since (Grim,Grim) passes the single-deviation test at each history, it is a subgame-perfect Nash equilibrium when $\delta \geq 1/5$.[6]

We will now use the same technique to show that (Tit-for-tat,Tit-for-tat) is *not* a subgame-perfect Nash equilibrium (except for the degenerate case $\delta = 1/5$). Tit-for-tat strategies at $t+1$ only depends on what is played at t not any previous play. If (C, C) is played at t, then starting at $t+1$ and we will have (C, C) throughout, and hence the vector of present values at $t+1$ will be

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

If (C, D) is played at t, then according to (Tit-for-tat,Tit-for-tat) the sequence of plays starting at $t+1$ will be

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

with $t+1$-present value of

$$\left(\frac{6}{1-\delta^2}, \frac{6\delta}{1-\delta^2}\right) = (6,0) + \delta\,(0,6) + \delta^2\,(6,0) + +\delta^3\,(0,6) + \delta^4\,(6,0) \cdots .$$

[5] Once again, to check this, we do not need to know the payoffs for (C, C).

[6] Once again, it is not a subgame-perfect Nash equilibrium when $\delta < 1/5$. In that case, it suffices to show that (C, C) is not a Nash equilibrium of the augmented stage game for a Cooperation history.

Similarly if (D,C) is played at t, then $t+1$-present value will be

$$\left(\frac{6\delta}{1-\delta^2}, \frac{6}{1-\delta^2}\right)$$

After (D,D) at t, we will have (D,D) throughout, yielding $t+1$-present value $(1/(1-\delta), 1/(1-\delta))$.

In order to show that (Tit-for-tat,Tit-for-tat) is *not* a subgame-perfect Nash equilibrium, we will consider two histories. [To show a strategy profile is not subgame-perfect, one only needs to find a case where it fails the single-deviation principle.] Given the above continuation games, the reduced game at any t for any previous history is

	C	D
C	$5+\delta\frac{5}{1-\delta}, 5+\delta\frac{5}{1-\delta}$	$0+\delta\frac{6}{1-\delta^2}, 6+\delta\frac{6\delta}{1-\delta^2}$
D	$6+\delta\frac{6\delta}{1-\delta^2}, 0+\delta\frac{6}{1-\delta^2}$	$1+\delta/(1-\delta), 1+\delta/(1-\delta)$

.

1. Consider $t=0$, when (Tit-for-tat,Tit-for-tat) prescribes (C,C). Single-deviation principle then requires that (C,C) is a Nash equilibrium of the reduced game above. That is, we must have
$$5+\delta\frac{5}{1-\delta} \geq 6+\delta\frac{6\delta}{1-\delta^2}.$$

2. Consider a history in which (C,D) is played at $t-1$. Now according to (Tit-for-tat,Tit-for-tat) we must have (D,C) at t. Single-deviation principle now requires that (D,C) is a Nash equilibrium of the above game. That is, we must have
$$6+\delta\frac{6\delta}{1-\delta^2} \geq 5+\delta\frac{5}{1-\delta},$$
the opposite of the previous requirement.

Hence, (Tit-for-tat,Tit-for-tat) is *not* a subgame-perfect Nash equilibrium, unless $6+\delta\frac{6\delta}{1-\delta^2} = 5+\delta\frac{5}{1-\delta}$, or equivalently $\delta = 1/5$.

12.3 Folk Theorem

A main objective of studying repeated games is to explore the relation between the short-term incentives (within a single period) and long term incentives (within the broader repeated game). Conventional wisdom in game theory suggests that when players are patient, their long-term incentives take over, and a large set of behavior may 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 the Folk Theorem. 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 $N = \{1, \ldots, n\}$ 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.

12.3.1 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 $v = (v_1, \ldots, v_n)$ such that

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

for some probability distribution $p : A \to [0, 1]$ on A. Note that V is the smallest convex set that contains all payoff vectors $(u_1(a), \ldots, u_n(a))$ from pure strategy profiles in the stage game. A payoff vector v is said to be *feasible* iff $v \in V$. 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 12.1. 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.

Note that for every strategy profile s in the repeated game, the average payoff vector from s is in V.[7] This also implies that the same is true for mixed strategy profiles in the repeated game. Conversely, if the players can collectively randomize on strategy profiles

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

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

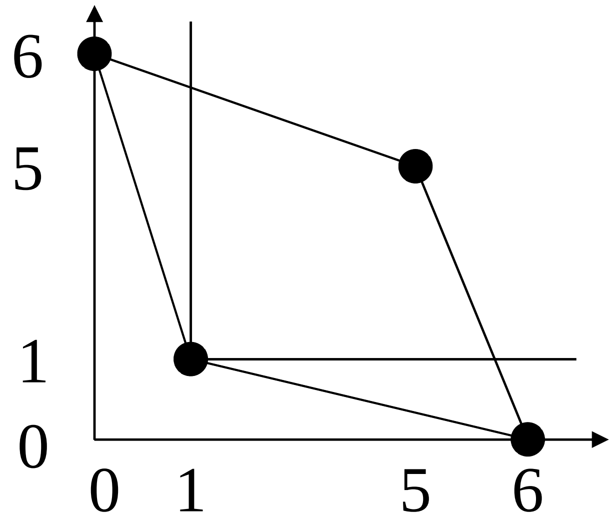


Figure 12.1: Feasible payoffs in Prisoners' Dilemma

in the repeated games, all vectors $v \in V$ could be obtained as average payoff vectors. (See also the end of the section.)

12.3.2 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, he gets at least 1 every day, netting an average payoff of 1 or more. Then, he must get at least 1 in any Nash equilibrium because he could otherwise profitably deviate to the above strategy.

Towards finding a lower bound on the payoffs from pure-strategy Nash equilibria, for

---

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,

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

each player i define *pure-strategy minmax payoff* as

$$m_i = \min_{a_{-i}\in A_{-i}} \max_{a_i\in A_i} u_i\left(a_i, a_{-i}\right). \tag{12.3}$$

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 could 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, $\hat{s}_i\left(h\right)$ is a stage-game best response to $s^*_{-i}\left(h\right)$, i.e.,

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

Since

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

for every h, this implies that the average payoff from $\left(\hat{s}_i, s^*_{-i}\right)$ is at least 1, 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

$$\mu_i = \min_{\alpha_j, j\neq i} \max_{a_i\in A_i} \sum_{a_{-i}\in A_{-i}} \prod_{j\neq i} \alpha_j\left(a_j\right) u_i\left(a_i, a_{-i}\right), \tag{12.4}$$

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 $\alpha_j\left(Head\right) = \alpha_j\left(Tail\right) = 1/2$.) For the sake of exposition, it is assumed that $\left(\mu_1, \ldots, \mu_n\right) \in V$.

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

12.3.3 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 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 above, then the subgame-perfect equilibrium above is in pure strategies.*

he 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, $a^* = (C, C)$, yielding $v = (5, 5)$, and $\hat{a} = (D, D)$, yielding payoff vector $(1, 1)$. Recall that in that case one could obtain v from strategy profile (Grim, Grim), which is a subgame-perfect Nash equilibrium when $\delta > 1/5$. The main idea here is a generalization of Grim strategy. Consider the following strategy profile s^* of the repeated game:

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 single-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

$$v_i \geq (1-\delta) \max_{a_i} u\left(a_i, a_{-i}^*\right) + \delta u_i\left(\hat{a}\right) \tag{12.5}$$

for every player i. Let

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

be the discount rate for which (12.5) becomes equality; such $\delta_i < 1$ exists because $\max_{a_i} u\left(a_i, a_{-i}^*\right) \geq u_i\left(a^*\right) = v_i > u_i\left(\hat{a}\right)$. Take $\bar{\delta} = \max\{\delta_1, \ldots, \delta_n\}$. Then, for every $\delta > \bar{\delta}$, inequality (12.5) 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, $\bar{\delta} = (6-5)/(6-1) = 1/5$.

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

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

We could 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, I *approximated* v by time averaging. When δ is large, one can obtain each v *exactly* by time averaging.[8]

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

12.4 Exercises with Solutions

1. Consider the n-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	$0, 0$	$0, 1$	$0, 0$

(a) 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 πn in every pure-strategy Nash equilibrium.

Solution: Note that the pure strategy minmax payoff of each player is 1. Hence, the payoff of a player cannot be less than n. 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 his payoff would be at least n. Since he plays a best response in equilibrium, his payoff is at least that amount. This lower bound is tight. For $n = 2k > 1$, 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 n. To check that this is a Nash equilibrium, note that the best possible deviation is to play play b forever, which yields n, giving no incentive to deviate. Note also that the quilibrium here is not subgame-perfect.

(b) Construct a pure-strategy subgame-perfect Nash equilibrium in which the payoff of each player is at most $n + 1$. Verify that the strategy profile is indeed a subgame-perfect Nash equilibrium.

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

n-times repeated game, consider the path

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

if n is odd and the path

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

if n is even. Note that the total payoff of each player from this path is $n+1$. Consider the following strategy profile.

Play according to the above path; if any player deviates from this path at any $t \leq n/2-1$, switch to $s^*[n-t-1]$ for the remaining $(n-t-1)$-times repeated game; if any player deviates from this path at any $t > n/2$, 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 $t' \leq n/2$. In that case, the strategy profile already prescribes to follow the subgame-perfect Nash equilibrium $s^*[n-t'-1]$ 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 $t' \leq n/2$ and take $t > n/2$. In the continuation game, the above strategy profile prescribes: play (b,b) every day if n is odd and play (b,b) every day but the last day and play (a,a) on the last day if n is even. Since (a,a) and (b,b) are Nash equilibria of the stage game, this is clearly a subgame-perfect equilibrium of the remaining game. Finally, take $t \leq n/2$ and consider any on-the path history. Now, a player's payoff is $n+1$ if he follows the strategy profile. If he deviates at t, he gets at most 1 at t and $(n-t-1)+1 \leq n$ from the next period on, where $(n-t-1)+1$ is his payoff from $s^*[n-t-1]$. His total payoff cannot exceed $n+1$, and he has no incentive to deviate.

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

(a) 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.

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 his average value is only $6(1-\delta)+\delta=1.05$. Hence, he 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 he deviates, his average value is only δ. Therefore, he 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.

(b) 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.

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

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

Answer: While the average payoff of Player 1 can be as high as 5.7999, it cannot be higher than 5.8. This is because $v_1 < 1$ 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).

3. [Midterm 2, 2006] Two firms, 1 and 2, play the following infinitely repeated game in which all the previous plays are observed, and each player tries to maximize the discounted sum of his or her profits at the stage games where the discount rate is $\delta = 0.99$. At each date t, simultaneously, each firm i selects a price $p_i \in \{0.01, 0.02, \ldots, 0.99, 1\}$. If $p_1 = p_2$, then each firm sells 1 unit of the good; otherwise, the cheaper firm sells 2 units and the more expensive firm sells 0 units. Producing the good does not cost anything to firms. Find a subgame-perfect equilibrium in which the *average value* of Firm 1 is at least 1.4. (Check that the strategy profile you construct is indeed subgame-perfect equilibrium.)

Solution: (There are several such strategy profiles; I will show one of them.) In order for the average value to exceed 1.4, the present value must exceed 140. We can get average value of approximately 1.5 for player 1 by alternating between $(0.99, 1)$, which yields $(1.98, 0)$, and $(1, 1)$, which yields $(1, 1)$. The average value of that payoff stream for player 1 is

$$\frac{1.98 + \delta}{1 + \delta} \cong 1.49.$$

Here is a SPE with such equilibrium play: *At even dates play* $(0.99, 1)$ *and at odd* *dates play* $(1,1)$*; if any player ever deviates from this scheme, then play* $(0.01, 0.01)$ *forever.*

We use the single-deviation principle, to check that this is a SPE. First note that in "deviation" mode, they play a Nash equilibrium of the stage game forever, and it passes the single-deviation test. Now, consider an even t and a history where there has not been any deviation. Player 1 has no incentive to deviate: if he follows the strategy, he will get the payoff stream 1.98, 1, 1.98, 1, 1.98, ...; if he deviates, he will get x, 0.01, 0.01, ... where $x \leq 1.96$ ($x = 1$ for upward deviation). For player 2: if he plays according to the strategy, he will get the payoff stream of 0, 1, 0, 1, 0, 1, ... with present value of

$$\delta / \left(1 - \delta^2\right) \cong 49.75.$$

If he deviates, he will get x, 0.01, 0.01, ... where $x \leq 1.96$. (The best deviation is $p_2 = 0.98$.) This yields present value of

$$x + 0.01 \cdot \delta / (1 - \delta) = x + 1 \leq 2.96 \ll 49.75.$$

He has no incentive to deviate. We also need to check an even date t with no previous deviation. Now the best deviation is to set price 0.99 and get 1.98 today and get 0.01 forever, which yields the present value of 2.98. This is clearly lower than what each player gets by sticking to their strategies (148.5 for player 1, and 50.25 for player 2).

4. [Midterm 2, 2011] 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 equilibrium.

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

(a) 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.)

Solution: 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, his he average payoff is

$$\ln 2.$$

Suppose he 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

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

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

$$1 - \delta.$$

Hence, the strategy profile is a subgame-perfect Nash equilibrium iff

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

where the valeus 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:

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

(b) 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 $\tilde{x}_B$.)

Hint: The following facts about logarithm may be useful:

$$\frac{d}{dx}\ln(x) = 1/x; \quad \ln(x) \leq x - 1; \quad \ln(xy) = \ln x + \ln y.$$

Solution: 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 iff $\ln 2 \geq 1-\delta\tilde{x}_i$, i.e.

$$\delta\tilde{x}_i \geq 1 - \ln 2. \tag{12.6}$$

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 iff

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

which simplifies to

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

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

$$\ln 2 + (1-\delta)(\ln \hat{x} + 1 - \hat{x}) \geq \hat{x}. \tag{12.7}$$

The strategy profile is a SPE iff (12.6) and (12.7) are satisfied.

Remark 12.1 *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 discound 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 interpeted 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.

5. [Final, 2001] 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.

(a) Assume that this is repeated for 100 days, and each player tries to maximize the sum of his or her stage payoffs. Find all subgame-perfect equilibria of this game.

Solution: 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 equilibrium, in which the stage equilibrium is repeated.

(b) 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: 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 SPE 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 him (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,

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

6. [Midterm 2 Make up, 2006] 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	$-1,-1$	$-1,-2$

Consider the infinitely repeated game with this stage game and with discount factor $\delta = 0.9$.

(a) Find a subgame perfect Nash equilibrium in which each country extracts low (L) amount of oil every year on the equilibrium path.[9]

Solution: 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.)

(b) 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: Consider the following ("Carrot and Stick") strategy profile[10]

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

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

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

Peace to War. In state Peace, Iraq clearly has no incentive to deviate. On the other hand, Kuwait gets $2(1-\delta)+\delta[2\delta-1]$ from H. Hence, it has no incentive to deviate if

$$2(1-\delta)+\delta[2\delta-1]\leq 1,$$

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, getting $2(1-\delta)+\delta(2\delta-1)$. It does not have an incentive to deviate if

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

i.e.,

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

This is equivalent to $\delta \geq 3/5$, which is clearly the case.

7. [Selected from Midterms 2 in years 2001 and 2002] 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 $\delta = 0.99$.

   (a) **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 he plays what the other player played in the previous round (i.e., at each $t>0$, he plays what the other player played at $t-1$).

   Solution: 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.)

   (b) **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.

Solution: 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, he gets 2 in the current period and gets zero thereafter. If he sticks to his equilibrium strategy, then he gets 1 forever. The present value of this is $1/(1-\delta) > 2$. Therefore, Player 2 doesn't have any incentive to deviate, either.

(c) **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.

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

(d) **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: 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).

12.5 Exercises

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

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 n-times repeated game?

3. Prove Theorem 12.1.

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

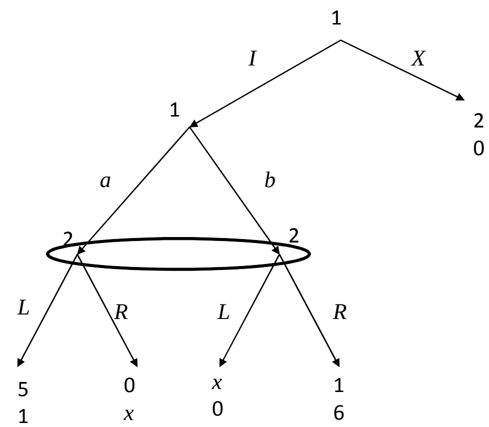
5. [Homework 4, 2011] Consider the infinitely repeated game with discount factor $\delta = 0.99$ and the following stage game (in which the players are trading favors):

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

(a) Find a subgame perfect 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.

(b) Find a subgame-perfect 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.

6. [Midterm 2, 2011] Consider the 100-times repeated game with the following stage game:



where x is either 0 or 6.

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

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

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

7. [Midterm 2, 2011] 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.

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

(b) 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.

8. [Midterm 2 Make Up, 2011] 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 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 equilibrium.

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

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

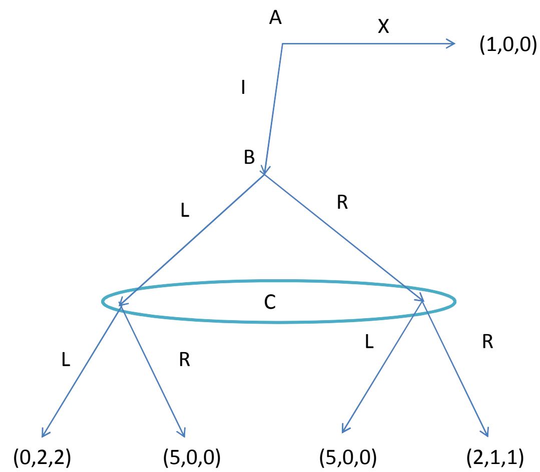
9. [Midterm 2 Make Up, 2011] Consider the infinitely repeated game with the stage game in the previous problem and the discount factor $\delta \in (0,1)$. For each of the strategy profiles below, find the conditions on the discount factor for which the strategy profile is a subgame-perfect equilibrium.

(a) At $t = 0$, they play (w,a). At each t, they play (w,a) if the play at $t-1$ is (w,a) or if the play at $t-2$ is **not** (w,a). Otherwise, they play (z,c).

(b) 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

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

10. [Homework 4, 2011] Consider the n-times repeated game with the following stage game.



(a) For $n = 2$, what is the largest payoff A can get in a subgame-perfect Nash equilibrium in pure strategies?

(b) For $n > 2$, find a subgame-perfect Nash equilibrium in which the payoff of A is at least $5n - 6$.

11. [Homework 4, 2011] Consider the infinitely repeated game with discount factor $\delta \in (0, 1)$ and the stage game in the previous problem. For each of the strategy profile below, find the range of δ under which the strategy profile is a subgame-perfect Nash equilibrium.

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

(b) 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.

(Note that the outcome is $(I, L, R), (I, R, L), (I, R, R), (I, L, R), (I, R, L), \ldots$.)

12. [Homework 4, 2007] 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 $\delta \in (0, 1)$.[11] For each strategy profile below, find the set of discount factors δ under which the strategy profile is a subgame-perfect equilibrium.

(a) 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.

(b) 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 is supposed to do so. They both stay Down thereafter.

(c) 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.

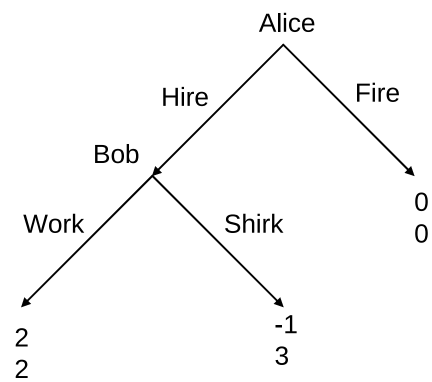
(d) Irene goes Up on "Sundays", i.e., at $t = 0, 7, 14, 21, \ldots$, and stays Down on the other days, while Jonathan goes up everyday except for Sundays, when he rests Down, until someone deviates; they both stay Down thereafter.

(e) 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

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

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.)

13. [Homework 4, 2007] Consider the infinitely repeated game, between Alice and Bob, with the following stage game:



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 equilibrium. If it is not a SPE for $\delta = 0.9$, find the set of discount factors δ under which it is a SPE.

(a) 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.

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

(c) 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. At any state, if only one player fails to play what s/he is supposed to play at that state, then we go to the Punishment state for that player in the next period; otherwise we go to the Employment state in the next period.

14. [Midterm 2, 2007] Consider the infinitely repeated game with the following stage game

	Chicken	Lion
Chicken	$3,3$	$1,4$
Lion	$4,1$	$0,0$

and discount factor $\delta = 0.99$. For each strategy profile below check if it is a subgame-perfect equilibrium. (You need to state your arguments clearly; you will not get any points for Yes or No answers.)

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

   (b) There are three states: Cooperation, P1 and P2. The game starts in the Cooperation state. In the Cooperation state, each player plays Chicken. If they play (Chicken, Chicken) or (Lion, Lion), then they remain in the Cooperation state in the next period. If player i plays Lion while the other player plays Chicken, then in the next period they go to Pi state. In Pi state player i plays Chicken while the other player plays Lion; they then go back to Cooperation state (regardless of the actions).

15. [Midterm 2 Make Up, 2007] 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 the third entry is the payoff of Alice:

Bob\Colin	Grab	Share
Grab	$-1,-1,-1$	$3,-2,-2$
Share	$-2,3,-2$	$2,2,2$

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 equilibrium. Show your work.

(a) There are three states: Share, P_{BOB} and P_{COLIN}. In Share state Alice lets them play and Bob and Colin both share. In P_{BOB} state (resp. P_{COLIN} state), Alice lets them play, and Bob (resp. Colin) shares while the other brother grabs. The game starts in Share state. If Bob (resp. Colin) does not play what he is supposed to play while the other player plays what he is supposed to play, then the next day we go to P_{BOB} (resp. P_{COLIN}) state; we go to Share state next day otherwise.

(b) 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 Play state, if everybody does what they are supposed to do, we remain in Play state; we go to 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 Play state next day.

16. [Homework 4, 2006] 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 or not at price $p \in [1, 3]$. (At the time Bob buys the food, he cannot tell if it is of high quality, but after buying he knows whether it was high or low quality.) The payoffs for a given day is as follows.

Alice\Bob	Buy	Skip
High	$p-1, 3-p$	$-1, 0$
Low	$p, -p$	$0, 0$

The discount rate is $\delta = 0.99$. For each of the following strategy profiles, find the range of $p \in [1, 3]$ for which the strategy profile is a subgame-perfect equilibrium

(a) There are two states: Trade and No-trade. The game starts at Trade state. In 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 day, independent of what happens, they go back to the Trade state.

(b) 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 in the rest of the game Alice uses Low quality and Bob Skips.[12]

17. [Homework 4, 2006] In question 1, take $p = 2$, and check whether each of the following is a subgame-perfect equilibrium. [We assume here that Bob somehow can check whether the food was good in the previous day even if did not buy it.]

    (a) 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.

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

18. [Homework 4, 2006] 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

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

    For $\delta = 0.99$, check if either of the following strategy profiles is a subgame-perfect equilibrium. [If you solve the problem for $n = 3$, you will get 80%.]

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

[12] That is, at any t', Alice will use Low quality supply and Bob wil Skip in either of the following cases: (i) Alice used Low quality supply at an even date $t < t'$, or (ii) she used High quality supply at an odd date $t < t'$, or (iii) Bob Skipped at some date $t < t'$.

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

19. [Homework 4, 2006] 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.

20. [Homework 3, 2004] 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

(a) Find all Nash equilibria.

(b) 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 equilibrium.

i. No matter what happens, the firm always pays and the worker works.

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

iii. 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.

iv. 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.

v. 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 $T > 0$.)

21. **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.

(a) (5 points) Find the set of rationalizable strategies in the Stage Game above.

(b) (10 points) 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 equilibrium of the repeated game in which Alice and Bob contribute a^* and b^*, respectively, on the path of equilibrium.

(c) (10 points) In part (b), 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 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 $\vec{b} = (b, b, \ldots)$.)

22. [Selected from Midterms 2 and make up exams in years 2002 and 2004] Below, there are pairs of stage games and strategy profiles. For each pair, check whether the strategy profile is a subgame-perfect 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.)

    (a) **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.)

    (b) **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.

(c) **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.

(d) **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).

23. [Midterm 2 Make Up, 2001] 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. Each agent tries to maximize the discounted sum of his expected payoffs in the stage game with discount rate δ.

(a) 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.]

(b) For sufficiently large values of δ, construct a subgame-perfect 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.