then either he gets 0 forever, with probability β, or he gets 2 forever, with probability $1 - \beta$. If player 1 chooses x_1 in round 1, he gets a current expected payoff of $2\beta + 0(1 - \beta)$, and thereafter the game continues in state 1 at round 2. So equations (7.2) and (7.3) imply that

$$z = 0\beta + 2(1 - \beta) = (1 - \delta)2\beta + \delta z.$$

Thus, $\beta = \frac{1}{2}$, and $z = 1$. Similarly, to make player 2 willing to randomize between x_2 and y_2, we must have

$$\begin{aligned} -z &= \alpha((1 - \delta)(-2) + \delta(-z)) + (1 - \alpha)(0) \\ &= \alpha((1 - \delta)(0) + \delta(-z)) + (1 - \alpha)(-2); \end{aligned}$$

so $\alpha = 1/(2 - \delta)$. Notice that, as δ approaches 1, player 1's equilibrium strategy approaches the strategy in which he always chooses x_1; but that cannot be an equilibrium because player 2 would respond to it by always choosing y_2.

Blackwell and Ferguson (1968) studied the Big Match under the limit of average payoffs criterion and showed that, in this two-person zero-sum game, player 1's minimax value is $+1$ and player 2's minimax value (or security level) is -1, just as it was under the discounted average payoff criterion. Player 2 can guarantee that player 1 cannot expect more than $+1$ by randomizing between x_1 and y_2, each with probability $\frac{1}{2}$, at every round while the state of nature is 1. However, there is no equilibrium under the limit of average payoffs because player 1 has no minimax strategy against player 2. Blackwell and Ferguson showed that, for any positive integer M, player 1 can guarantee himself an expected limit of average payoffs that is not lower than $M/(M + 1)$ (so player 2's limit of average payoffs is not more than $-M/(M + 1)$) by using the following strategy: at each round, choose y_1 with probability $1/(1 + M + x^* - y^*)^2$, where x^* is the number of times that player 2 has chosen x_2 in the past and y^* is the number of times that player 2 has chosen y_2 in the past.

7.4 Repeated Games with Standard Information: Examples

A *repeated game with standard information*, or a *standard repeated game* (also sometimes called a *supergame*), is a repeated game in which there is only one possible state of nature and the players know all of each other's past moves. That is, in a standard repeated game $|\Theta| = 1$, $S_i = \times_{j \in N} D_j$ for every i, and

$$p(d, \ldots, d, \theta \mid d, \theta) = 1, \quad \forall d \in D.$$

Repeated games with standard information represent situations in which a group of individuals face exactly the same competitive situation infinitely often and always have complete information about each other's past behavior. Thus, the standard information structure maximizes the players' abilities to respond to each other.

The stationary equilibria of a standard repeated game are just the equilibria of the corresponding one-round game, repeated in every round. However, the nonstationary equilibria of standard repeated games are generally much larger sets, because of the players' abilities to respond to each other and punish each other's deviations from any supposed equilibrium path. In fact, under weak assumptions, almost any feasible payoff allocation that gives each player at least his minimax value can be achieved in an equilibrium of a standard repeated game. A general statement of this result is developed and proved in the next section; in this section, we develop the basic intuition behind this important result by detailed consideration of two examples.

To illustrate the analysis of standard repeated games, consider the repeated game of *Chicken*, where the payoffs to players 1 and 2 are as shown in Table 7.3. Each player i has two moves to choose between at each round: the "cautious" move a_i and the "bold" move b_i. Each player would most prefer to be bold while the other is cautious, but for both to be bold is the worst possible outcome for both. The best symmetric outcome is when both are cautious.

If the players play this game only once, then there are three Nash equilibria of this game, one giving payoffs (6,1), one giving payoffs (1,6), and one randomized equilibrium giving expected payoffs (3,3). When communication is allowed, there is a correlated equilibrium of this game in which the probability of (a_1,a_2) is equal to .5 and the expected payoff allocation is (3.75,3.75), but no higher symmetric ex-

Table 7.3 Payoffs at any round, for all move profiles, in the repeated game of Chicken

	D_2	
D_1	a_2	b_2
a_1	4,4	1,6
b_1	6,1	−3,−3

pected payoff allocation can be achieved in a correlated equilibrium of the one-round game. In particular, (a_1,a_2) is not an equilibrium of the one-round game, because each player i would prefer to be bold (choose b_i) if he expected the other player to be cautious (choose a_{-i}).

Suppose now that this game is repeated infinitely, with standard information, and each player uses a δ-discounted average payoff criterion, for some number δ that is between 0 and 1. A strategy for a player in the repeated game is a rule for determining his move at every round as a function of the history of moves that have been used at every preceding round. For example, one celebrated strategy for repeated games like this one (or the repeated Prisoners' Dilemma) is the strategy called *tit-for-tat*. Under the tit-for-tat strategy, a player chooses his cautious move in the first round, and thereafter chooses the same move as his opponent chose in the preceding round.

If both players follow the tit-for-tat strategy, then the actual outcome will be (a_1,a_2) in every round, giving each player a discounted average payoff of 4. However, tit-for-tat should not be confused with the strategy "always choose a_i," under which a player would choose a_i at every round no matter what happened in the past. To see why, suppose first that player 1 is following the strategy of "play a_1 at every round." Then player 2's best response is to always play b_2 and get a discounted average payoff of 6. On the other hand, if player 1 is following the tit-for-tat strategy and the discount factor δ is close to 1, then player 2 will never want to choose b_2, because she will lose more from player 1's reprisal next round than she will gain from choosing b_2 in this round. Thus, the part of a strategy that specifies how to punish unexpected behavior by the opponent may be very important, even if such punishments are not carried out in equilibrium.

Let us check to see how large δ has to be for tit-for-tat to deter bold behavior by one's opponent. Suppose that player 1 is implementing the tit-for-tat strategy and that player 2 is considering whether to be bold at the first round and thereafter go back to being cautious. If she does so, her payoff will be 6 in the first round, 1 in the second round, and 4 at every round thereafter, so her δ-discounted average payoff is

$$(1 - \delta)(6 + 1\delta + \sum_{k=3}^{\infty} 4\delta^{k-1}) = 6 - 5\delta + 3\delta^2.$$

On the other hand, if she is always cautious against player 1's tit-for-tat, she will get the discounted average payoff of 4. So, to deter player 2 from being bold, we need

$$4 \geq (1 - \delta)(6 + 1\delta + \sum_{k=3}^{\infty} 4\delta^{k-1}),$$

or $\delta \geq 2/3$. In fact, if $\delta \geq 2/3$, then neither player can gain by being the first to deviate from the tit-for-tat strategy, if the other is expected to always use tit-for-tat; so tit-for-tat is an equilibrium.

Although it is an equilibrium for both players to use the tit-for-tat strategy, the equilibrium is not subgame perfect. To be a subgame-perfect equilibrium, the players must always choose sequentially rational moves, even after histories of moves that have zero probability under the given strategies. Consider the situation that would be faced by player 1 at the second round if player 2 accidentally chose b_2 at the first round. Under the assumption that both players will follow tit-for-tat hereafter, the outcome will be (b_1,a_2) in every even-numbered round and (a_1,b_2) in every odd-numbered round; and the discounted average payoff to player 1 will be

$$(1 - \delta)(1 + 6\delta + 1\delta^2 + 6\delta^3 + \ldots) = (1 + 6\delta)/(1 + \delta).$$

On the other hand, if player 1 deviates from tit-for-tat by choosing a_1 at round 2 and both players then follow tit-for-tat thereafter, then the outcome will be (a_1,a_2) at every round after round 1; so the discounted average payoff to player 1 will be

$$(1 - \delta)(1 + 4\delta + 4\delta^2 + 4\delta^3 + \ldots) = 1 + 3\delta.$$

With some straightforward algebra, it can be shown that

$$1 + 3\delta > (1 + 6\delta)/(1 + \delta) \text{ when } \delta > 2/3.$$

Furthermore, for any δ, either player could gain from unilaterally deviating from the tit-for-tat strategy after both players chose (b_1,b_2), because alternating between payoffs 1 and 6 is better than always getting -3. So it would be irrational for a player to implement the punishment move in tit-for-tat if the other player is also expected to implement tit-for-tat hereafter. That is, the scenario in which both play tit-for-tat is not a subgame-perfect equilibrium.

There is a way to modify tit-for-tat to get around this difficulty, however. Consider the following strategy for player i: at each round, i chooses a_i unless the other player has in the past chosen b_{-i} strictly more times than i has chosen b_i, in which case i chooses b_i. We call this strategy *getting-even*. It can be shown that, as long as δ is greater than $2/3$, it is a subgame-perfect equilibrium for both players to follow the getting-even strategy. For example, if player 2 deviates from the strategy and chooses b_2 at round 1 but is expected to apply the strategy thereafter, then it is rational for player 1 to retaliate and choose b_1 at round 2. Player 1 expects that player 2, under the getting-even strategy, will treat player 1's retaliatory b_1 as a justified response; so player 2 should not reply with a counterretaliatory b_2 in the third round. The condition $\delta \geq 2/3$ is needed only to guarantee that neither player wants to choose b_i when the number of past bold moves is equal for the two players.

The distinction between tit-for-tat and getting-even is very fine, because they differ only after a mistake has been made. Let $r_i(k)$ denote the number of rounds at which player i chooses b_i before round k. If player 1 applies the getting-even strategy correctly, then, no matter what strategy player 2 uses, at any round k, $r_2(k) - r_1(k)$ will always equal either 0 or 1 and will equal 1 if and only if player 2 chose b_2 last round. Thus, according to the getting-even strategy, player 1 will choose b_1 if and only if player 2 chose b_1 in the preceding round. The distinction between the two strategies becomes evident only in cases where player 1 himself has made some accidental deviation from his own strategy but then goes back to implementing it.

Neither tit-for-tat nor getting-even can be sustained bilaterally as an equilibrium of this game if $\delta < 2/3$. However, as long as $\delta \geq .4$, there are other subgame-perfect equilibria of this repeated game that achieve the outcome (a_1,a_2) in every round with probability 1.

When both players are supposed to choose a_i in equilibrium, player 1 can deter player 2 from choosing b_2 only if he has a credible threat of some retaliation or punishment that would impose in the future a greater cost (in terms of the discounted average) than player 2 would gain from getting 6 instead of 4 in the current round. As the discount factor becomes smaller, however, losses in later payoffs matter less in comparison with gains in current payoffs, so it becomes harder to deter deviations to b_2. To devise a strategy that deters such deviations for the lowest possible discount factor, we need to find the credible threat that is most costly to player 2.

At worst, player 2 could guarantee herself a payoff of $+1$ per round, by choosing a_2 every round. So the discounted average payoff to player 2 while she is being punished for deviating could not be lower than 1. Such a payoff would actually be achieved if player 1 chose b_1 and player 2 chose a_2 forever after the first time that player 2 chose b_2. Furthermore, it would be rational for player 1 to choose b_1 forever and player 2 to choose a_2 forever if each expected the other to do so. Thus, the worst credible punishment against player 2 would be for player 1 to choose b_1 forever while player 2 responds by choosing a_2.

Consider the following *grim strategy* for player i: if there have been any past rounds where one player was bold while the other player was cautious (outcome (b_1,a_2) or (a_1,b_2)) and, at the *first* such round, i was the cautious player, then i chooses b_i now and hereafter; otherwise i chooses a_i. That is, under the grim strategy, i plays cooperatively (that is, "cautiously") until his opponent deviates and is bold, in which case i punishes by being bold forever. As long as $\delta \geq .4$, it is a subgame-perfect equilibrium for both players to follow the grim strategy.

Such infinite unforgiving punishment may seem rather extreme. So we might be interested in finding other strategies that, like tit-for-tat, have only one-round punishments and can sustain (a_1,a_2) forever as the actual outcome in a subgame-perfect equilibrium when the discount factor is smaller than $2/3$. Such a strategy does exist. We call it *mutual punishment*, and it can be described for player i as follows: i chooses a_i if it is the first round, or if the moves were (a_1,a_2) or (b_1,b_2) last round; i chooses b_i if the moves were (b_1,a_2) or (a_1,b_2) last round. If both players are supposed to be following this strategy but one player deviates in a particular round, then in the next round the players are supposed to choose (b_1,b_2), which is the worst possible outcome for both of them (hence the name "mutual punishment"). If player 2 is following this strategy correctly and player 1 deviates to choose b_1, then player 2 will choose b_2 in retaliation the next round and will continue to do so until player 1 again chooses b_1. The idea is that player 2 punishes player 1's deviation until player 1 participates in his own punishment or bares his chest to the lash.

To check whether it is a subgame-perfect equilibrium for both players to follow the mutual-punishment strategy, we must consider two possible deviations: choosing b_i when the strategy calls for a_i, and choosing a_i when the strategy calls for b_i. If the strategy calls for a_i but player i deviates for one round while the other player follows the strategy, then i will get payoff $+6$ this round and -3 the next, whereas he would have gotten $+4$ in both rounds by not deviating. Because the next round is discounted by an extra factor of δ, the deviation is deterred if $4 + 4\delta \geq 6 + -3\delta$, that is, if $\delta \geq 2/7$. On the other hand, if the strategy calls for b_i but i deviates for one round while the other player follows the strategy, then i will get payoff $+1$ this round and -3 the next, whereas he would have gotten -3 in this round and $+4$ in the next by not deviating. So this deviation is deterred if $-3 + 4\delta \geq 1 + -3\delta$, that is, if $\delta \geq 4/7$. Thus, it is a subgame-perfect equilibrium for both players to follow the mutual-punishment strategy as long as the discount factor is at least $4/7$.

Thus far, we have been discussing only equilibria in which the actual outcome is always the cooperative (a_1,a_2) at every round, unless someone makes a mistake or deviates from the equilibrium. There are many other equilibria to this repeated game, however. Any equilibrium of the original one-round game would be (when repeated) an equilibrium of the repeated game. For example, there is an equilibrium in which player 1 always chooses b_1 and player 2 always chooses a_2. This is the best equilibrium for player 1 and the worst for player 2.

There are also equilibria of the repeated game that have very bad welfare properties and are close to being worst for both players. Furthermore, some of these bad equilibria have a natural or logical appeal that may, in some cases, make them the focal equilibria that people actually implement. To see the logic that leads to these bad equilibria, notice that the getting-even and grim equilibria both have the property that the player who was more bold earlier is supposed to be less bold later. Some people might suppose that the opposite principle is more logical: that the player who has been more bold in the past should be the player who is expected to be more bold in the future.

So consider a strategy, which we call the *q-positional* strategy, that may be defined for each player i as follows: i chooses b_i if i has chosen b_i strictly more times than the other player has chosen b_{-i}; i chooses a_i if the other player has chosen b_{-i} strictly more times than i has chosen b_i; and if b_i and b_{-i} have been chosen the same number of times, then i chooses b_i now with probability q and chooses a_i now with probability $1 - q$. The intuitive rationale for this strategy is that the player who has established a stronger reputation for boldness can be bold in the future, whereas the player who has had a more cautious pattern of behavior should conform to the cautious image that he has created. When neither player has the more bold reputation, then they may independently randomize between bold and cautious in some way.

Given any value of the discount factor δ between 0 and 1, there is a value of q such that it is a subgame-perfect equilibrium for both players to follow the q-positional strategy. To compute this q, notice first that, although there are no alternative states of nature intrinsically defined in this repeated game, the players' positional strategies are defined in terms of an implicit state: the difference between the number of past rounds in which player 1 played b_1 and the number of past rounds in which player 2 played b_2. At any round, if this implicit state is positive and both players follow the q-positional strategy, then the payoffs will be (6,1) at every round from now on. Similarly, if the implicit state is negative, then the payoffs will be (1,6) at every round from now on. Let z denote the expected δ-discounted average payoff expected by each player (the same for both, by symmetry) when the implicit state is 0 at the first round (as, of course, it actually is) and both apply the q-positional strategy. Then player 1 is willing to randomize between b_1 and a_1 when the implicit state is 0 if and only if

$$\begin{aligned} &q((1 - \delta)(-3) + \delta z) + (1 - q)((1 - \delta)(6) + \delta(6)) \\ &\quad = q((1 - \delta)(1) + \delta(1)) + (1 - q)((1 - \delta)(4) + \delta z) = z. \end{aligned}$$

These equalities assert that, when player 2's probability of being bold is q, player 1's expected δ-discounted average payoff in the truncated game is the same, whether he is bold (first expression) or cautious (second expression) in the first round, and is in fact equal to z (as our definition of z requires). Given δ, these two equations can be solved for q and z. The results are shown in Table 7.4. Thus, if the players have long-term objectives, so δ is close to 1, then the positional equilibrium gives each player an expected δ-discounted average payoff that is only slightly better than the minimum that he could guarantee himself in the worst equilibrium. In effect, the incentive to establish a reputation for boldness pushes the two players into a *war of attrition* that is close to the worst possible equilibrium for both.

The various equilibria discussed here may be taken as examples of the kinds of behavior that can develop in long-term relationships. When people in long-term relationships perceive that their current behavior will have an influence on each other's future behavior, they may ration-

Table 7.4 Expected δ-discounted average payoffs (z) and initial boldness probabilities (q) for four discount factors (δ), in positional equilibrium of repeated Chicken

δ	q	z
.99	.992063	1.0024
.90	.925	1.024
.667	.767	1.273
.40	.582	1.903

Table 7.5 Payoffs at any round, for all move profiles, in a repeated game

	D_2	
D_1	a_2	b_2
a_1	8,8	1,2
b_1	2,1	0,0

ally become more cooperative (as in the getting-even equilibrium) or more belligerent (as in the positional equilibrium), depending on the kind of linkage that is expected between present and future behavior. Qualitatively, the more cooperative equilibria seem to involve a kind of reciprocal linkage (e.g., "expect me tomorrow to do what you do today"), whereas the more belligerent equilibria seem to involve a kind of extrapolative linkage ("expect me tomorrow to do what I do today").

For a simple example in which repeated-game equilibria may be worse for both players than any equilibrium of the corresponding one-round game, consider the game in Table 7.5. It is easy to see that the unique equilibrium of the one-round game is (a_1,a_2), which gives payoffs (8,8). For the repeated game, however, consider the following scenario: if the total number of past rounds when (a_1,a_2) occurred is even, then player 1 chooses a_1 and player 2 chooses b_2; if the total number of past rounds when (a_1,a_2) occurred is odd, then player 1 chooses b_1 and player 2 chooses a_2. When the players implement these strategies, (a_1,a_2) never occurs; so (a_1,b_2) is the outcome at every round (because 0 is even) and payoffs are (1,2). This scenario is a subgame-perfect equilibrium if $\delta \geq 6/7$. For example, if player 2 deviated from this scenario by choosing a_2 at round 1, then her discounted average payoff would be

$$(1 - \delta)(8 + 1\delta + 1\delta^2 + 1\delta^3 + \ldots) = 8 - 7\delta.$$

Notice that $2 \geq 8 - 7\delta$ if $\delta \geq 6/7$.

7.5 General Feasibility Theorems for Standard Repeated Games

The general intuition that we take from the examples in the preceding section is that, in standard repeated games, when players are sufficiently patient, almost any feasible payoff allocation that gives each player at least his minimax security level can be realized in an equilibrium of the repeated game. That is, the payoff allocations that are feasible in equi-