CHAPTER 14: REPEATED PRISONER’S DILEMMA

In this chapter, we consider infinitely repeated play of the Prisoner’s Dilemma game. We denote the possible actions for $P_i$ by $C_i$ for cooperating with the other player and $D_i$ for defecting from the other player. (Earlier, these actions were called quiet and fink respectively.) The payoff matrix for the game is assumed to be as follows:

$$
egin{pmatrix}
C_2 & D_2 \\
C_1 & (2, 2) & (0, 3) \\
D_1 & (3, 0) & (1, 1)
\end{pmatrix}
$$

To simplify the situation, we consider the players making simultaneous moves with the current move unknown to the other player. This is defined formally on page 206. We use a game graph rather than a game tree to represent this game. See Figure 1.

Let $a_t = (a_1^{(t)}, a_2^{(t)})$ be the action profile at the $t$th stage. The one step payoff is assumed to depend on only the action profile at the last stage, $u_i(a^{(t)})$. There is a discount factor $0 < \delta < 1$ to bring this quantity back to an equivalent value at the first stage, $\delta^{t-1} u_i(a^{(t)})$. There are two ways to understand the discounting. (i) If the payoff is in money and $r > 0$ is an interest rate, then capital $V_1$ at the first stage is worth $V_t = (1 + r)^{t-1} V_1$ at the $t$th stage ($t - 1$ steps later). Thus, the value of $V_t$ at the first stage is $V_t/(1 + r)^{t-1}$. In this context, the discounting is $\delta = 1/(1 + r)$. (ii) If the payoff is not money but satisfaction, then $\delta$ is a measure of the extent the player wants rewards now, i.e., how impatient the player is. See the book for further explanation.

For a finitely repeated game of $T$ stages (finite horizon), the total payoff for $P_i$ is

$$
U_i(a^{(1)}, \ldots, a^{(T)}) = u_i(a^{(1)}) + \delta u_i(a^{(2)}) + \cdots + \delta^{T-1} u_i(a^{(T)})
$$

For a finitely repeated prisoner’s dilemma game with payoffs as above, at the last stage, both players optimize their payoff by selecting $D_i$. Given this choice, then the choice that optimizes the payoff at the $T - 1$ stage is again $D_i$. By backward induction, both players will select $D$ at each stage. See Section 14.4.

For the rest of this chapter, we consider an infinitely repeated game starting at stage one (infinite horizon). The discounted payoff for player $P_i$ is given by

$$
U_i(\{a_i\}_{i=1}^\infty) = \sum_{i=1}^{\infty} \delta^{i-1} u_i(a^{(i)}).
$$
Some Nash Equilibria Strategies for Infinitely Repeated Games

We consider some strategies as reactions to action of the other player that have gone before. We only analyze situations where both players use the same strategy and check for which δ this strategy is a Nash equilibrium. In describing the strategy for \( P_i \), we let \( P_j \) be the other player. Thus, if \( i = 1 \) then \( j = 2 \), and if \( i = 2 \) then \( j = 1 \).

**Defection Strategy.** In this strategy, both players select \( D \) in response to any history of actions. It is easy to check that this is a Nash equilibrium.

**Grim Trigger Strategy.** (page 426) The strategy for \( P_i \) is given by the rule

\[
s_i(\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(t-1)}) = \begin{cases} 
  C_i & \text{if } t = 1 \text{ or } a_j^{(\ell)} = C_j \text{ for all } 1 \leq \ell \leq t - 1 \\
  D_i & a_j^{(\ell)} = D_j \text{ for some } 1 \leq \ell \leq t - 1.
\end{cases}
\]

We introduce the concept of states of the two players to give an alternative description of this strategy. The states depend on the strategy and are defined so that the action of the strategy for player \( P_i \) depends only on the state of \( P_i \). For the grim trigger strategy, we define the following two states for \( P_i \):

- \( \mathcal{S}_i = \{t = 1\} \cup \{(\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(t-1)}): a_j^{(\ell)} = C_j \text{ for all } 1 \leq \ell \leq t - 1\} \)
- \( \mathcal{D}_i = \{(\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(t-1)}): a_j^{(\ell)} = D_j \text{ for some } 1 \leq \ell \leq t - 1\} \).

These states determine a new game tree that has a vertex at each stage for a pair of states for the two players. Figure 2 presents a partial game graph. The transitions between the states depend only on the action of the other player at the last stage. Where either action results in the same next state, we put a star for the action.

The grim trigger strategy can easily be given in terms of these states: The strategy of \( P_i \) is to select \( C_i \) if the state is \( \mathcal{S}_i \) and to select \( D_i \) if the state is \( \mathcal{D}_i \).

Rather than giving a game graph, it is simpler to give a figure presenting the states and transitions. Each box is labeled with the state of the player and the next action taken by that player in that state according to the strategy. The arrows represent the transitions between states determined by the action of the other player. See Figure 3 for the states and transitions for the grim trigger strategy. The double box is the starting state.

We next check that if both players use the grim trigger strategy the result is a Nash equilibrium. Since the game starts in state \( (\mathcal{S}_1, \mathcal{S}_2) \), applying the strategy will keep both players in the same states. The one step payoff at each stage is 2. Assume that \( P_2 \) maintains the strategy and \( P_1 \) deviates at stage \( T \) by selecting \( D_1 \).
Then, $P_2$ selects $C_2$ for $t = T$ and selects $D_2$ for $t > T$. The greatest payoff for $P_1$ results from selecting $D_1$ for $t > T$. Thus, if $P_1$ selects $D_1$ for $t = T$, then the greatest payoff from that stage onward is

$$3\delta^{T-1} + \delta^T + \delta^{T+1} + \cdots = 3\delta^{T-1} + \delta^T \left(1 + \delta + \delta^2 + \cdots\right) = 3\delta^{T-1} + \frac{\delta^T}{1 - \delta}.$$ 

If $P_1$ plays the original strategy, the payoff from the $T^{th}$ stage onward is

$$2\delta^{T-1} + 2\delta^T + 2\delta^{T+1} + \cdots = \frac{2\delta^{T-1}}{1 - \delta}.$$ 

Therefore, the grim trigger strategy is a Nash equilibrium provided that

$$\frac{2\delta^{T-1}}{1 - \delta} \geq 3\delta^{T-1} + \frac{\delta^T}{1 - \delta},$$

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

$$2\delta \geq 1,$$

$$\delta \geq \frac{1}{2}.$$ 

This shows that if both players are patient enough so that $\delta \geq 1/2$, then the grim trigger strategy is a Nash equilibrium.

**Tit-for-tat Strategy.** (Section 14.7.3) We describe the tit-for-tat strategy in terms of states of the players. For the tit-for-tat strategy, there are two states for $P_i$ that only depend on the action of $P_j$ in the last period:

$$\mathcal{C}_i = \{ t = 1 \} \cup \{ (a_1^{(1)}, \ldots, a_t^{(t-1)}): a_j^{(t-1)} = C_j \}$$

$$\mathcal{D}_i = \{ (a_1^{(1)}, \ldots, a_t^{(t-1)}): a_j^{(t-1)} = D_j \}.$$ 

For the tit-for-tat strategy, player $P_i$ chooses $C_i$ in state $\mathcal{C}_i$ and $D_i$ in state $\mathcal{D}_i$. The transitions between states caused by actions of the other player are given in Figure 4.

![Figure 3. States and transitions for grim trigger](image)

![Figure 4. States and transitions for tit-for-tat](image)
We next check that the tit-for-tat strategy by both players is also a Nash equilibrium for \( \delta \geq \frac{1}{2} \). Assume that \( P_2 \) maintains the strategy and \( P_1 \) deviates by selecting \( D_1 \) at the \( T^{th} \)-stage. The other possibilities for actions by \( P_1 \) include (a) \( D_1 \) for all \( t \geq T \), (b) \( D_1 \) and then \( C_1 \), and (c) \( D_1 \) for \( k \) times and then \( C_1 \). In cases (b) and (c), player \( P_2 \) returns to the original state \( C_2 \); so it is enough to calculate this segment of the payoffs. Note that the book ignores the last case.) We check these three cases in turn.

(a) If \( P_1 \) uses \( D_1 \) for all \( t \geq T \), then \( P_2 \) uses \( C_2 \) for \( t = T \) and \( D_2 \) for \( t > T \). The payoff for these choices is

\[
3 \delta^{T-1} + \delta^T + \delta^{T+1} + \cdots = 3 \delta^{T-1} + \frac{\delta^T}{1 - \delta}.
\]

The payoff for the original tit-for-tat strategy starting at the \( T^{th} \)-stage is \( \frac{2 \delta^{T-1}}{1 - \delta} \), so for it to be a Nash equilibrium, we need

\[
\frac{2 \delta^{T-1}}{1 - \delta} \geq 3 \delta^{T-1} + \frac{\delta^T}{1 - \delta}.
\]

\[
2 \geq 3(1 - \delta) + \delta = 3 - 2 \delta
\]

\[
2 \delta \geq 1
\]

\[
\delta \geq \frac{1}{2}.
\]

(b) If \( P_1 \) selects \( D_1 \) and then \( C_1 \), then \( P_2 \) selects \( C_2 \) and then \( D_2 \). The payoff for \( P_1 \) is \( 3 \delta^{T-1} + (0) \delta^T \) versus the original payoff of \( 2 \delta^{T-1} + 2 \delta^T \). In order for tit-for-tat to be a Nash equilibrium, we need

\[
2 \delta^{T-1} + 2 \delta^T \geq 3 \delta^{T-1}
\]

\[
2 \delta^T \geq \delta^{T-1}
\]

\[
\delta \geq \frac{1}{2}.
\]

We get the same condition on \( \delta \) as in (a).

(c) If \( P_1 \) selects \( D_1 \) for \( k \) stages and then \( C_1 \), then \( P_2 \) will select \( C_2 \) and then \( D_2 \) for \( k \) stages. At the end, \( P_2 \) is back in state \( C_2 \). The payoffs for these \( k + 1 \) stages of the original strategy and the deviation are

\[
2 \delta^{T-1} + \cdots + 2 \delta^{T+k-1} \quad \text{and} \quad 3 \delta^{T-1} + \delta^T + \cdots + \delta^{T+k-2} + (0) \delta^{T+k-1}.
\]

Thus, we need

\[
2 \delta^{T-1} + \cdots + 2 \delta^{T+k-1} \geq 3 \delta^{T-1} + \delta^T + \cdots + \delta^{T+k-2} \quad \text{or} \quad -1 + \delta + \cdots + \delta^{k-1} + 2 \delta^k \geq 0.
\]

If \( \delta \geq \frac{1}{2} \), then

\[
2 \delta^k + \delta^{k-1} + \cdots + \delta - 1 \geq 2 \left( \frac{1}{2} \right)^k + \left( \frac{1}{2} \right)^{k-1} + \cdots + \frac{1}{2} - 1
\]

\[
= \left( \frac{1}{2} \right)^{k-1} + \left( \frac{1}{2} \right)^{k-1} + \cdots + \frac{1}{2} - 1
\]

\[
= 2 \left( \frac{1}{2} \right)^{k-1} + \left( \frac{1}{2} \right)^{k-2} + \cdots + \frac{1}{2} - 1
\]

\[
= 2 \left( \frac{1}{2} \right) - 1
\]

\[
= 0.
\]

Thus, the condition is satisfied. For \( \delta < \frac{1}{2} \) the inequalities go the other direction and it is less than zero. This checks all the possible deviations, so the tit-for-tat strategy is a Nash equilibrium for \( \delta \geq \frac{1}{2} \).
**Limited punishment Strategy.** (Section 14.7.2) For the limited punishment strategy, each player has \( k + 1 \) states for some \( k \geq 2 \). For \( P_i \), starting in state \( \mathcal{P}_{i,0} \), if the other player selects \( D_j \), then there is a transition to \( \mathcal{P}_{i,1} \), then a transition to \( \mathcal{P}_{i,2}, \ldots, \mathcal{P}_{i,k} \), and then back to \( \mathcal{P}_{i,0} \). The transitions from \( \mathcal{P}_{i,\ell} \) for \( 1 \leq \ell \leq k \) do not depend on the actions of either player. For the limited punishment strategy, the actions of \( P_i \) are \( C_i \) in state \( \mathcal{P}_{i,0} \) and \( D_i \) in states \( \mathcal{P}_{i,\ell} \) for \( 1 \leq \ell \leq k \). See Figure 5 for the case of \( k = 2 \). See Figure 427.2 in Osborne for the case of \( k = 3 \).

![Figure 5](image1.png)

**Figure 5.** States and transitions for limited punishment for \( k = 2 \)

If \( P_1 \) select \( D_1 \) the \( (T + 1)^{th} \) stage, then \( P_2 \) will select \( C_2 \) and then \( D_2 \) for the next \( k \) stages. The maximum payoff for \( P_1 \) is obtained by selecting \( D_1 \) for all of these \( k + 1 \) stages. The payoffs for \( P_1 \) are

\[
2\delta^T + 2\delta^{T+1} + \cdots + 2\delta^{T+k} \leq 2\delta^T + 2\delta^{T+1} + \cdots + 2\delta^{T+k},
\]

where

\[
1 - \delta - \delta^k = \delta \left( \frac{1 - \delta^k}{1 - \delta} \right),
\]

or

\[
g_k(\delta) = 1 - 2\delta + \delta^{k+1} \leq 0.
\]

We check that this inequality is valid for \( \delta \) large enough. The derivatives of \( g_k \) are

\[
g'_k(\delta) = -2 + (k+1)\delta^k \quad \text{and} \quad g''_k(\delta) = k(k+1)\delta^{k-1} > 0 \quad \text{for} \quad \delta > 0.
\]

Some values of \( g_k \) are as follows:

\[
g_k \left( \frac{1}{2} \right) = 1 - 1 + \left( \frac{1}{2} \right)^{k+1} > 0,
\]

\[
g_k \left( \frac{3}{2} \right) = 1 - \frac{3}{2} + \left( \frac{3}{2} \right)^{k+1} \leq -\frac{1}{2} + \frac{27}{64} = -\frac{5}{64} < 0,
\]

\[
g_k(1) = 0.
\]

By the Intermediate Value Theorem, there must be a \( \frac{1}{2} < \delta_k^* < \frac{3}{2} \) such that \( g_k(\delta_k^*) = 0 \). As stated in the book, \( \delta_2^* \approx 0.618 \) and \( \delta_3^* \approx 0.544 \). See Figure 6. The function is concave up (convex) so \( g_k(\delta) \leq 0 \) for \( \delta_k^* \leq \delta < 1 \), and the limited punishment strategy is a Nash equilibrium for \( \delta_k^* \leq \delta < 1 \).

![Figure 6](image2.png)

**Figure 6.** Plot of \( g_2(\delta) \)
CHAPTER 14: REPEATED PRISONER’S DILEMMA

Subgame Perfect Equilibria: Sections 14.9 & 14.10

The following is a criterion for a subgame perfect equilibrium.

**Definition. One deviation property:** No player can increase her payoff by changing her action at the start of any subgame, given the other player’s strategy and the rest of her own strategy. Notice that the rest of the strategy is fixed, not the rest of the actions.

The point is that the deviation needs only be checked at one stage at a time.

**Proposition (438.1).** A strategy in an infinitely repeated game with discount factor 0 < \( \delta < 1 \) is a subgame perfect equilibrium iff it satisfies the one deviation property.

**Defection Strategy.** This is obviously a subgame perfect strategy since the same choice is made at every vertex and it is a Nash equilibrium.

**Grim Trigger Strategy.** (Section 14.10.1) This is not subgame perfect as given. Starting at the state \((C_1, D_2)\), it is not a Nash equilibrium. Since \( P_2 \) is playing the grim trigger, she will pick \( D_2 \) at every stage. Player \( P_1 \) will play \( C_1 \) and then \( D_1 \) for every other stage. The payoff for \( P_1 \) is

\[
0 + \delta + \delta^2 + \cdots.
\]

However, if \( P_1 \) changes to always playing \( D_1 \), then the payoff is

\[
1 + \delta + \delta^2 + \cdots,
\]

which is larger. Therefore, this is not a Nash equilibrium on a subgame with root pair of states \((C_1, D_2)\).

A slight modification leads to a subgame perfect equilibrium. Keeping the same states for \( C_i \) and \( D_i \), change the transitions to depend on the state of both players. If the action of either player is \( D \), then there is a transition from \((C_1, C_2)\) to \((D_1, D_2)\). See Figure 7. This strategy is a subgame perfect equilibrium for \( \delta \geq \frac{1}{2} \).

![Figure 7. States and transitions for the modified grim trigger](image)

**Limited punishment Strategy.** (Section 14.10.2) This can also be modified to make a subgame perfect equilibrium: Make the transition from \((P_{1,0}, P_{2,0})\) to \((P_{1,1}, P_{2,1})\) when either player takes the action \( D \). The rest is the same.

**Tit-for-tat Strategy.** (Section 14.10.3) The four combinations of states for the two players are \((C_1, C_2)\), \((C_1, D_2)\), \((D_1, C_2)\), and \((D_1, D_2)\). We need to check that the strategy is a Nash equilibrium on a subgame starting at any of these four state profiles.

(i) \((C_1, C_2)\): The analysis we gave to show that it was a Nash equilibrium applies and shows that it is true for \( \delta \geq \frac{1}{2} \).

(ii) \((C_1, D_2)\): If both players adhere to the strategy, then the actions will be

\((C_1, D_2), (D_1, C_2), (C_1, D_2), \cdots\)

with payoff

\[
0 + 3 \delta + (0) \delta^2 + 3 \delta^3 = 3 \delta(1 + \delta^2 + \delta^4 + \cdots) = \frac{3 \delta}{1 - \delta^2}.
\]
If $P_1$ instead starts by selecting $D_1$, then the actions will be

$$(D_1, D_2), (D_1, D_2), \cdots$$

with payoff

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

So we need

$$\frac{3 \delta}{1 - \delta^2} \geq \frac{1}{1 - \delta}$$

$$3 \delta \geq 1 + \delta$$

$$2 \delta \geq 1$$

$$\delta \geq \frac{1}{2}.$$ 

(iii) ($\mathcal{G}_1, \mathcal{G}_2$): If both players adhere to the strategy, then the actions will be

$$(D_1, C_2), (C_1, D_2), (D_1, C_2), \cdots$$

with payoff

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

If $P_1$ instead starts by selecting $C_1$, then the actions will be

$$(C_1, C_2), (C_1, C_2), \cdots$$

with payoff

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

So we need

$$\frac{3}{1 - \delta^2} \geq \frac{2}{1 - \delta}$$

$$3 \geq 2 + 2 \delta$$

$$1 \geq 2 \delta$$

$$\frac{1}{2} \geq \delta.$$ 

(iv) ($\mathcal{G}_1, \mathcal{G}_2$): If both players adhere to the strategy, then the actions will be

$$(D_1, D_2), (D_1, D_2), (D_1, D_2), \cdots$$

with payoff

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

If $P_1$ instead starts by selecting $C_1$, then the actions will be

$$(C_1, D_2), (D_1, C_2), \cdots$$

with payoff

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

So we need

$$\frac{1}{1 - \delta} \geq \frac{3 \delta}{1 - \delta^2}$$

$$1 + \delta \geq 3 \delta$$

$$\frac{1}{2} \geq \delta.$$ 

For all four of these conditions to hold, we need $\delta = \frac{1}{2}$. 
Prevalence of Nash equilibria

It is possible to realize many different payoffs with Nash equilibrium; in particular, there are uncountably many different payoffs for different Nash equilibrium. The payoffs that are possible for Nash equilibrium are stated in terms of what is called the discounted average payoff which we now define.

If \( \{w_t\}_{t=1}^{\infty} \) is the stream of payoffs (for one of the players), then the discounted sum is

\[
U(\{w_t\}_{t=1}^{\infty}) = \sum_{t=1}^{\infty} \delta^{t-1} w_t.
\]

If all the payoffs are the same value, \( w_t = c \) for all \( t \), then

\[
U(\{c\}_{t=1}^{\infty}) = \sum_{t=1}^{\infty} \delta^{t-1} c = \frac{c}{1 - \delta}, \quad \text{so} \quad c = (1 - \delta) U(\{c\}_{t=1}^{\infty}).
\]

For this reason, we call the quantity

\[
\tilde{U}(\{w_t\}_{t=1}^{\infty}) = (1 - \delta) U(\{w_t\}_{t=1}^{\infty})
\]

is called the discounted average. This quantity \( \tilde{U}(\{w_t\}_{t=1}^{\infty}) \) is such that if the same quantity is repeated infinitely many times then the same quantity is returned by \( \tilde{U} \). Applying this to actions, the quantity

\[
\tilde{U}_i(\{a_t\}_{t=1}^{\infty}) = (1 - \delta) U_i(\{a_t\}_{t=1}^{\infty})
\]

is called the discounted average payoff of the action stream.

**Definition.** The set of feasible payoff profiles of a strategic game is the set of all weighted averages of payoff profiles in the game. For the the Prisoner’s Dilemma game we are considering, the feasible payoff profiles are the weighted averages (convex combinations) of \( u(C_1, D_2) = (2, 2) \), \( u(C_1, D_2) = (0, 3) \), \( u(D_1, C_2) = (3, 0) \), and \( u(D_1, D_2) = (1, 1) \). See Figure 433.1 in the book.

Clearly any discounted average payoff profile for a game must lie in the set of feasible payoff profiles. We want to see what other restrictions there are on the discounted average payoff profiles. We start with the Prisoner’s Dilemma.

**Theorem (Subgame Perfect Nash Equilibrium Folk Theorem, 435.1 & 447.1).** Consider an infinitely repeated Prisoner’s Dilemma, \( G \).

a. For any discount factor \( 0 < \delta < 1 \), the discounted average payoff of each player \( P_i \) for a (subgame perfect) Nash equilibrium is at least \( u_i(D_1, D_2) \). (In addition, the discounted average payoff profile must lie in the set of feasible payoff profiles.)

b. For any discount factor \( 0 < \delta < 1 \), the infinitely repeated game of \( G \) has a subgame perfect equilibrium in which the discounted average payoff is \( u_i(D_1, D_2) \) for each for each player \( P_i \).

c. Let \( (x_1, x_2) \) be a feasible pair of payoffs in \( G \) for which \( x_i > u_i(D, D) \) for \( i = 1, 2 \). There exists a \( 0 < \delta < 1 \) such that if \( \delta < \delta < 1 \), then the infinitely repeated game of \( G \) has a subgame perfect equilibrium in which the discounted average payoff for each player \( P_i \) is \( x_i \).

Part (a) follows since \( P_i \) could insure the payoff of at least \( u_i(D, D) \) by always selecting \( D_i \). For part (b), if both players select \( D_i \) at every stage, then the discounted average payoff profile is exactly \( (u_1(D_1, D_2), u_2(D_1, D_2)) \). The idea of the proof of part (c) is to find a sequence of actions whose discounted average is close to the desired payoff. Then a strategy that punishes the other player who deviates from this sequence of actions makes it into a subgame perfect equilibrium. See the discussion in the book on pages 435-436 and 446-447.
For a game other than a Prisoner’s Dilemma, a way of determining the minimum payoff for a Nash equilibrium must be given. We give the value for a two person strategic game where \( P_i \) is the person under consideration with set of possible actions \( A_i \) and \( P_j \) is the other person with set of possible actions \( A_j \).

Player \( P_i \)’s minmax payoff in a strategic game is
\[
m_i = \min_{a_j \in A_j} \max_{a_i \in A_i} u_i(a_i, a_j).
\]

Parts (a) and (c) of folk theorem are now the same where the value \( u_i(D_1, D_2) \) is replaced by the minmax for \( P_i \). For part (b), if the one time strategic game \( G \) has a Nash equilibrium in which each player’s payoff is her minmax payoff, then for any discount factor the infinitely repeated game of \( G \) has a subgame perfect Nash equilibrium in which the discounted average payoff of each player \( P_i \) is her minmax payoff. Note that the game \[
\begin{pmatrix}
(2, 1) & (0, 0) \\
(0, 0) & (1, 2)
\end{pmatrix}
\]
has minmax payoff of \((1, 1)\) and there is not strategy that realizes this payoff.

**Theorem (Subgame Perfect Nash Equilibrium Folk Theorem, 454.1 & 458.2).** Let \( G \) be a two-player strategic game in which each player has finitely many actions and let \( m_i \) be the minimax payoff for player \( P_i \).

a. For any discount factor \( 0 < \delta < 1 \), the discounted average payoff of each player \( P_i \) for a (subgame perfect) Nash equilibrium of the infinitely repeated game \( G \) is at least \( m_i \).

b. If the one time game \( G \) has a Nash equilibrium in which each player’s payoff is \( m_i \), then for any discount factor \( 0 < \delta < 1 \), the infinitely repeated game of \( G \) has a subgame perfect equilibrium in which the discounted average payoff for each player \( P_i \) is \( m_i \).

c. Let \((x_1, x_2)\) be a feasible pair of payoffs in \( G \) for which \( x_i > m_i \) for \( i = 1, 2 \). There exists a \( 0 < \delta \leq 1 \) such that if \( \bar{\delta} < \delta < 1 \), then the infinitely repeated game of \( G \) has a subgame perfect equilibrium in which the discounted average of the payoff for each player \( P_i \) is \( x_i \).