REPEATED STRATEGIC GAME

Consider the prisoner's dilemma game with possible actions C_i for P_i cooperating (with the other player) and D_i for P_i 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:

$$\begin{array}{cc} & \begin{array}{cc} C_2 & D_2 \end{array} \\ \begin{array}{c} C_1 \\ D_1 \end{array} & \begin{pmatrix} (2,2) & (0,3) \\ (3,0) & (1,1) \end{pmatrix} \end{array}$$

We want to consider repeated play of this game for several or an infinite number of times. 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.

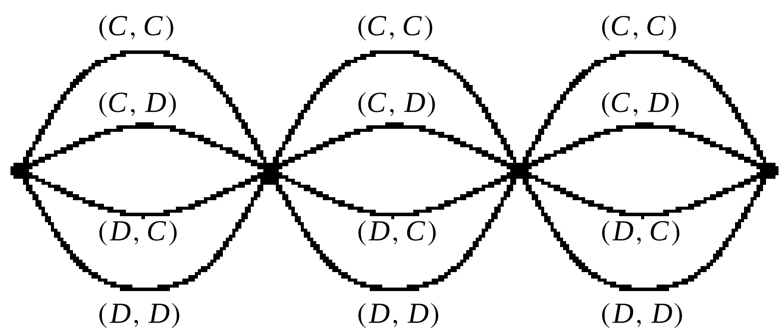

FIGURE 1. Game tree for repeated prisoner's dilemma

Let $\mathbf{a}^{(t)} = (a_1^{(t)}, a_2^{(t)})$ 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, $u_i(\mathbf{a}^{(\ell)})$. There is a discount factor $) < \delta < 1$ to bring this quantity back to an equivalent value at the first stage, $\delta^{t-1} u_i(\mathbf{a}^{(t)})$. For a finitely repeated game of T stages (finite horizon), the total payoff for P_i is

$$\begin{aligned} U_i(\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(T)}) &= u_i(\mathbf{a}^{(1)}) + \delta\, u_i(\mathbf{a}^{(2)}) + \cdots + \delta^{T-1}\, u_i(\mathbf{a}^{(T)}) \\ &= \sum_{t=1}^{T} \delta^{t-1}\, u_i(\mathbf{a}^{(t)}). \end{aligned}$$

There are a couple of ways to understand the discounting. If $r > 0$ 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 $\delta = {}^{1}\!/_{(1+r)}$. If the payoff is not money but satisfaction, then δ 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 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.

INFINITELY REPEATED GAMES (INFINITE HORIZON)

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

$$U_i(\{\mathbf{a}_t\}_{t=1}^{\infty}) = \sum_{t=1}^{\infty} \delta^{t-1} u_i(\mathbf{a}^{(t)}).$$

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

$$\begin{aligned} U(\{c\}_{t=1}^{\infty}) &= \sum_{t=1}^{\infty} \delta^{t-1} c \\ &= c \sum_{k=0}^{\infty} \delta^{k} \\ &= \frac{c}{1-\delta}, \qquad \text{so} \\ c &= (1-\delta)\, U(\{c\})_{t=1}^{\infty}). \end{aligned}$$

Thus, 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

$$\tilde{U}_i(\{\mathbf{a}_t\}_{t=1}^{\infty}) = (1-\delta)\, U_i((\mathbf{a}_t)_{t=1}^{\infty})$$

is the discounted average payoff of the action stream.

SOME NASH EQUILIBRIA STRATEGIES

We describe some strategies as reactions to action profiles that have gone before. We only describe situations where both players use the same rules to define their strategies. In describing the strategy for P_i, we let j be the other player. Thus, if $i = 1$ then $j = 2$, and if $i = 2$ then $j = 1$. We then describe a manner in which to understand these strategies in terms of a modified game graph.

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

$$s_i(\mathbf{a}^{(1}, \dots, \mathbf{a}^{(t-1)}) = \begin{cases} C_i & \text{if } t = 1 \text{ or } a_j^{(\ell)} = C \text{ for all } 1 \le \ell \le t-1 \\ D_i & a_j^{(\ell)} = D \text{ for some } 1 \le \ell \le t-1. \end{cases}$$

We are next going to decibel this strategy in terms of states of the two players. The states are defined so that the action of the strategy for player P_i depends only on the state of P_i. These states can be used to determine a new game tree that has a vertex at each stage for a pair of states for the two players.

For the grim trigger strategy, there are two states for P_i:

$$\begin{aligned} \mathscr{C}_i &= \{t=1\} \cup \{(\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(t-1)}) : a_j^{(\ell)} = C_j \text{ for all } 1 \le \ell \le t-1\} \\ \mathscr{D}_i &= \{(\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(t-1)}) : a_j^{(\ell)} = D_j \text{ for some } 1 \le \ell \le t-1\}. \end{aligned}$$

The strategy of P_i is to select C_i if the state is $\mathscr{C}_i$ and to select D_i if the state is $\mathscr{D}_i$. The transitions between the states depend only on the action of the other player at the last stage. This situation can be represented by the game tree in Figure 2.

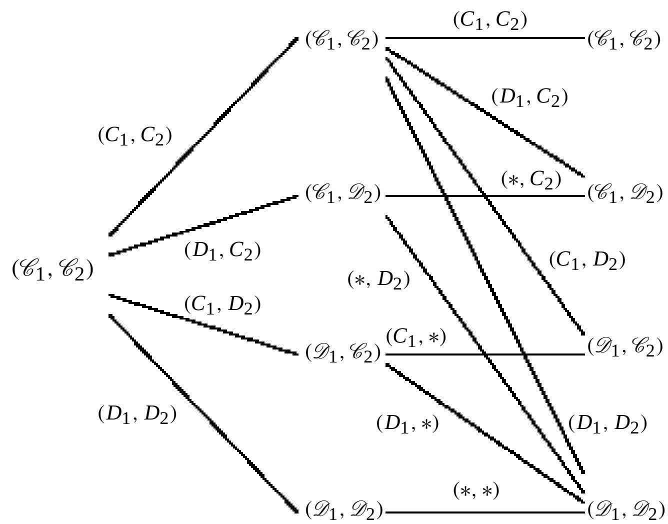

FIGURE 2. Game tree for grim trigger

As given in the book, rather than giving a game tree, it is easier to give a figure presenting the transitions and states (of only one player). See Figure 3.

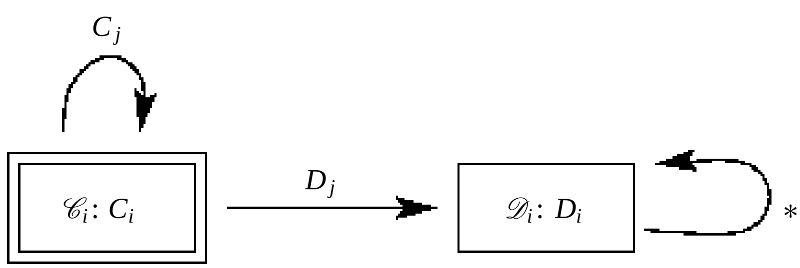

FIGURE 3. States and transitions for grim trigger

We next check that if both players use the grim trigger strategy the result is a Nash equilibrium. Since we start in state $(\mathscr{C}_1, \mathscr{C}_2)$, 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

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

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

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

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

$$\begin{aligned}\frac{2\,\delta^T}{1-\delta} &\geq 3\,\delta^T + \frac{\delta^{T+1}}{1-\delta}\\ 2 &\geq 3(1-\delta)+\delta = 3-2\,\delta\\ 2\,\delta &\geq 1\\ \delta &\geq \tfrac{1}{2}.\end{aligned}$$

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. (page 427, Section 14.7.3) We describe this 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:

$$\begin{aligned}\mathscr{C}_i &= \{t=1\} \cup \{(\mathbf{a}^{(1)},\dots,\mathbf{a}^{(t-1)}) : a_j^{(t-1)} = C_j\,\}\\ \mathscr{D}_i &= \{(\mathbf{a}^{(1)},\dots,\mathbf{a}^{(t-1)}) : a_j^{(t-1)} = D_j\,\}.\end{aligned}$$

The transitions between states are given in Figure 4.

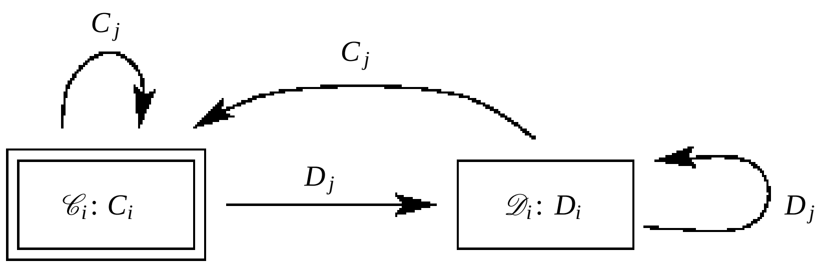

FIGURE 4. States and transitions for tit-for-tat

We next check that the tit-for-tat strategy by both players is also a Nash equilibrium for $\delta \geq 1/2$. Assume that P_2 maintains the strategy and P_1 deviates by selecting D_1 at the T^{th}-stage. The payoff for the original strategy starting at the T^{th}-stage is

$$\frac{2\,\delta^T}{1-\delta}.$$

The other possibilities for actions by P_1 include (a) D_1 for $t \geq T$, (b) alternating D_1 and C_1 forever, and (c) D_1 for k times and then C_1. (The latter returns P_2 to the original state, so it is enough to calculate this segment of the payoffs. Note that the book ignores the last case.) We check these three case in turn.

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

$$3\,\delta^T + \delta^{T+1} + \delta^{T+2} + \cdots = 3\,\delta^T + \frac{\delta^{T+1}}{1-\delta}.$$

For tit-for-tat to be a Nash equilibrium, we need

$$\begin{aligned}\frac{2\,\delta^T}{1-\delta} &\geq 3\,\delta^T + \frac{\delta^{T+1}}{1-\delta}\\ 2 &\geq 3(1-\delta)+\delta = 3-2\,\delta\\ 2\,\delta &\geq 1\\ \delta &\geq \tfrac{1}{2}.\end{aligned}$$

(**b**) If P_1 alternates D_1 and C_1, then P_2 alternates C_2 and D_2. The payoff for P_1 is

$$\begin{aligned} 3\,\delta^T + (0)\,\delta^{T+1} + 3\,\delta^{T+2} + \cdots &= 3\,\delta^T\left(1 + \delta^2 + \delta^4 + \cdots\right) \\ &= \frac{3\,\delta^T}{1-\delta^2}. \end{aligned}$$

In order for tit-for-tat to be a Nash equilibrium, we need

$$\begin{aligned} \frac{2\,\delta^T}{1-\delta} &\geq \frac{3\,\delta^T}{1-\delta^2} \\ 2\,(1+\delta) &\geq 3 \\ 2\,\delta &\geq 1 \\ \delta &\geq \tfrac{1}{2}. \end{aligned}$$

We get the same condition on δ as in case (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 $\mathscr{C}_2$. The payoffs for these $k+1$ stages of the original strategy and the the deviation are

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

Thus, we need

$$\begin{aligned} 2\delta^T + \cdots + 2\delta^{T+k} &\geq 3\delta^T + \delta^{T+1} + \cdots + \delta^{T+k-1} \qquad \text{or} \\ -1 + \delta + \cdots + \delta^{k-1} + 2\delta^k &\geq 0. \end{aligned}$$

If $\delta \geq 1/2$, then

$$\begin{aligned} 2\delta^k + \delta^{k-1} + \cdots + \delta - 1 &\geq 2\left(\tfrac{1}{2}\right)^k + \left(\tfrac{1}{2}\right)^{k-1} + \cdots + \tfrac{1}{2} - 1 \\ &\geq \left(\tfrac{1}{2}\right)^{k-1} + \left(\tfrac{1}{2}\right)^{k-1} + \cdots + \tfrac{1}{2} - 1 \\ &\geq 2\left(\tfrac{1}{2}\right)^{k-1} + \left(\tfrac{1}{2}\right)^{k-2} + \cdots + \tfrac{1}{2} - 1 \\ &\;\;\vdots \\ &\geq 2\left(\tfrac{1}{2}\right) - 1 \\ &= 0. \end{aligned}$$

Thus, the condition is satisfied. This checks all the possible deviations, so the tit-for-tat strategy is a Nash equilibrium for $\delta \geq 1/2$.

Limited punishment Strategy. (Section 14.7.2) In this strategy, each player has $k+1$ states for some $k \geq 2$. For P_i, starting in state $\mathscr{P}_{i,0}$, if the other player selects D_j, then there is a transition to $\mathscr{P}_{i,1}$, then a transition to $\mathscr{P}_{i,2} \ldots, \mathscr{P}_{i,k}$, and then back to $\mathscr{P}_{i,0}$. The transitions from $\mathscr{P}_{i,\ell}$ 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 $\mathscr{P}_{i,0}$ and D_i in states $\mathscr{P}_{i,\ell}$ for $1 \leq \ell \leq k$. See Figure 5 for the case of $k = 2$. See the book for the case of $k = 3$.

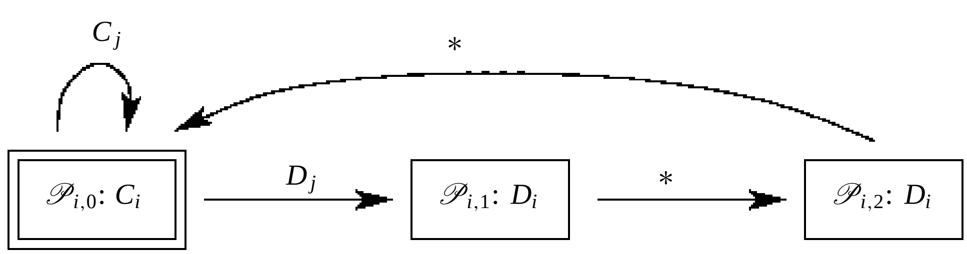

FIGURE 5. States and transitions for limited punishment

If P_1 select D_1 at some stage, the 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 $2+2\delta+\cdots+2\delta^k$ for the limited punishment strategy that results in all C for both players, and $3+\delta+\cdots+\delta^k$ for the deviation. Therefore, we need

$$\begin{aligned} 3+\delta+\cdots+\delta^k &\leq 2+2\delta+\cdots+2\delta^k, \\ 1 &\leq \delta+\cdots+\delta^k = \delta\left(\frac{1-\delta^k}{1-\delta}\right), \\ 1-\delta &\leq \delta-\delta^{k+1}, \qquad \text{and} \\ g_k(\delta) &= 1-2\,\delta+\delta^{k+1} \leq 0. \end{aligned}$$

To check that this is true for δ large enough, we use calculus.

$$\begin{aligned} g_k(1) &= 0, \\ g_k\left(\tfrac{1}{2}\right) &= 1-1+\left(\tfrac{1}{2}\right)^{k+1} > 0, \\ g_k'(\delta) &= -2+(k+1)\delta^k, \qquad \text{and} \\ g_k'(1) &= -2+k+1 > 0 \qquad \text{since } \ k \geq 2. \end{aligned}$$

There is only one $\bar{\delta}$ such that $g_k'(\bar{\delta}) = 0$:

$$\begin{aligned} \bar{\delta}^k &= \frac{2}{k+1} \\ \bar{\delta}^k &= \left(\frac{2}{k+1}\right)^{\frac{1}{k}}. \end{aligned}$$

Therefore, there is a $\frac{1}{2} \leq \delta_k^* \leq \bar{\delta} < 1$ such that $g_k(\delta) \leq 0$ for $\delta_k^* \leq \delta < 1$. For this range of δ, the limited punishment strategy is a Nash equilibrium.

The book mentions that $\delta_2^* \approx 0.62$ and $\delta_3^* \approx 0.55$.

Existence of many Nash equilibrium. The book states that it is possible to realize many different payoffs with Nash equilibrium. See Theorem 435.1. In particular, there are uncountably many different payoffs for different Nash equilibrium.

Subgame Perfect Equilibria: Sections 14.9 & 14.10

The following is a criterion for a subgame perfect equilibrium.

Definition 1. **One deviation property:** No player can increase her payoff by changing her action at the start of any subgame in which she is the first mover, given the other players' strategy *and* the rest of her own strategy.

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 $(\mathscr{C}_1, \mathscr{D}_2)$, it is not a Nash equilibrium. Since P_2 is playing the grim trigger, she will pick D_2 at every stage after. 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 $(\mathscr{C}_1, \mathscr{D}_2)$.

A slight modification leads to a subgame perfect equilibrium. Keep the states the same, but make a transition from $\mathscr{C}_i$ to $\mathscr{D}_i$ if the action of either player is D. See Figure 6. This gives a subgame perfect equilibrium for $\delta \geq 1/2$.

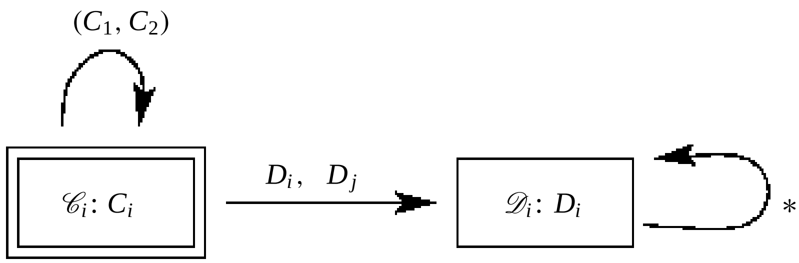

FIGURE 6. States and transitions for the modified grim trigger

Limited punishment Strategy. (Section 14.10.2) This can also be modified to make a subgame perfect equilibrium: Make the transition from $\mathscr{P}_{i,0}$ to $\mathscr{P}_{i,1}$ 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 $(\mathscr{C}_1, \mathscr{C}_2)$, $(\mathscr{C}_1, \mathscr{D}_2)$, $(\mathscr{D}_1, \mathscr{C}_2)$, and $(\mathscr{D}_1, \mathscr{D}_2)$. We need to check that the strategy is a Nash equilibrium on a subgame starting at any of these four state profiles.

(i) $(\mathscr{C}_1, \mathscr{C}_2)$: The analysis we gave to show that it was a Nash equilibrium applies and shows that it is true for $\delta \geq 1/2$.

(ii) $(\mathscr{C}_1, \mathscr{D}_2)$: 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

$$\begin{aligned} \frac{3\,\delta}{1 - \delta^2} &\geq \frac{1}{1 - \delta} \\ 3\,\delta &\geq 1 + \delta \\ 2\,\delta &\geq 1 \\ \delta &\geq \tfrac{1}{2}. \end{aligned}$$

(iii) $(\mathscr{D}_1, \mathscr{C}_2)$: 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

$$\begin{aligned} \frac{3}{1-\delta^2} &\geq \frac{2}{1-\delta} \\ 3 &\geq 2 + 2\,\delta \\ 1 &\geq 2\,\delta \\ \delta &\leq \tfrac{1}{2}. \end{aligned}$$

(iv) $(\mathscr{D}_1, \mathscr{D}_2)$: 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

$$\begin{aligned} \frac{1}{1-\delta} &\geq \frac{3\,\delta}{1-\delta^2} \\ 1 + \delta &\geq 3\,\delta \\ 1 &\geq 2\,\delta \\ \delta &\leq \tfrac{1}{2}. \end{aligned}$$

For all four of these conditions to hold, we need $\delta = 1/2$.