

Perfect Equilibria in Stochastic Games¹

F. THUIJSMAN,² S. H. TIJS,³ AND O. J. VRIEZE⁴

Communicated by R. Rishel

Abstract. We examine stochastic games with finite state and action spaces. For the β -discounted case, as well as for the irreducible limiting average case, we show the existence of trembling-hand perfect equilibria and give characterizations of those equilibria. In the final section, we give an example which illustrates that the existence of stationary limiting average equilibria in a nonirreducible stochastic game does not imply the existence of a perfect limiting average equilibrium.

Key Words. Stochastic games, equilibria, trembling-hand perfect equilibria.

1. Introduction

To keep notations simple, we restrict the analysis to 2-person stochastic games; the results can easily be extended to N -person games with $N > 2$.

A 2-person stochastic game is a finite set of matrices $\{M_1, M_2, \dots, M_z\}$ corresponding with the set of states $S = \{1, 2, \dots, z\}$. The matrix M_s has size $m_s \times n_s$, and the entry (i, j) of M_s is given as

$$\begin{array}{|c|} \hline r^1(s, i, j), r^2(s, i, j) \\ \hline p(s, i, j) \\ \hline \end{array}.$$

¹ Support was provided by the Netherlands Organization for Scientific Research NWO via the Netherlands Foundation for Mathematics SMC, Project 10-64-10.

² Assistant Professor, Department of Mathematics, Faculty of General Sciences, University of Limburg, Maastricht, The Netherlands.

³ Professor, Department of Mathematics, Catholic University, Nijmegen, The Netherlands.

⁴ Professor, Department of Mathematics, Faculty of General Sciences, University of Limburg, Maastricht, The Netherlands.

Here,

$$p(s, i, j) = (p(1|s, i, j), p(2|s, i, j), \dots, p(z|s, i, j)) \in \Delta^z, \\ r^1(s, i, j), r^2(s, i, j) \in \mathbf{R}, \quad \Delta^z = \left\{ q \in \mathbf{R}^z: q \geq 0, \sum_{i=1}^z q_i = 1 \right\}.$$

The interpretation is as follows: Play can start in any state of S , and transitions occur at stages $n \in \mathbf{N}$. If play is in state s at stage n , then independently and simultaneously player 1 chooses $i \in \{1, 2, \dots, m_s\}$ and player 2 chooses $j \in \{1, 2, \dots, n_s\}$. The triplet (s, i, j) determines a direct payoff $r^k(s, i, j)$ to player $k \in \{1, 2\}$ as well as a probability vector $p(s, i, j)$, whose meaning is that, at stage $n+1$, play will be in state $t \in S$ with probability $p(t|s, i, j)$. In this way, the play proceeds from stage to stage and from state to state.

We allow the players to randomize over their pure actions, which means that, in state s , player 1 [2] uses some mixed action $x_s \in \Delta^{m_s}[y_s \in \Delta^{n_s}]$. Now, player 1 [2] will choose action $i[j]$ with probability $x_s(i)[y_s(j)]$.

Each player's interest is to maximize his total income, without making binding agreements. To achieve this goal, each player uses some strategy: this is a plan that, at any stage of play, given the current state as well as the history, tells the player what mixed action to use. The history of play at stage n is the sequence

$$((s_1, i_1, j_1), (s_2, i_2, j_2), \dots, (s_{n-1}, i_{n-1}, j_{n-1}))$$

of past triplets of states and actions that occurred.

A stationary strategy is a strategy for which only the current state decides what mixed action is to be used, and neither stage nor history play a role. Hence, a stationary strategy for player 1 is simply some

$$x \in X := \prod_{s=1}^z \Delta^{m_s};$$

for player 2, it is some

$$y \in Y := \prod_{s=1}^z \Delta^{n_s}.$$

Each player wants to maximize his own expected total income. In this paper, we deal with two interpretations of expected total income: the β -discounted reward and the limiting average (or undiscounted) reward. Let R_n^k be the random variable denoting the direct payoff to player k at stage n , and let $E_{s\sigma\tau}$ denote expectation with respect to starting in s and players using (σ, τ) . Then,

$$\gamma_\beta^k(s, \sigma, \tau) := E_{s\sigma\tau} \left((1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} R_n^k \right),$$

with $\beta \in [0, 1)$, is the β -discounted reward for play k , and

$$\gamma_1^k(s, \sigma, \tau) := E_{s\sigma\tau} \left(\liminf_{T \rightarrow \infty} (1/T) \sum_{n=1}^T R_n^k \right)$$

is the limiting average (or undiscounted) reward for player k .

For $\beta \in [0, 1]$, we write $\gamma_\beta^k(\sigma, \tau)$ for

$$(\gamma_\beta^k(1, \sigma, \tau), \gamma_\beta^k(2, \sigma, \tau), \dots, \gamma_\beta^k(z, \sigma, \tau)).$$

In any stochastic game, either both players want to maximize their β -discounted rewards, or they both want to maximize their limiting average rewards.

Clearly, in general a pair of strategies need not exist which maximizes both player 1's and player 2's income at the same time. However, if the players have independently selected their strategies to play the game, then each of them will be satisfied if his strategy is a best reply to the strategy of his opponent, for otherwise, a better strategy could be chosen. Hence, a pair of strategies which are best replies to one another seems to be a reasonable and stable solution for the game, stable since no player has an incentive to deviate unilaterally.

Definition 1.1. A pair of strategies (σ^*, τ^*) is a β -discounted equilibrium ($\beta < 1$), or a limiting average equilibrium ($\beta = 1$) if, for all strategies σ and τ ,

$$\gamma_\beta^1(\sigma, \tau^*) \leq \gamma_\beta^1(\sigma^*, \tau^*),$$

$$\gamma_\beta^2(\sigma^*, \tau) \leq \gamma_\beta^2(\sigma^*, \tau^*).$$

In this paper, we are mainly concerned with stationary equilibria, i.e., equilibria consisting of stationary strategies. The set of stationary equilibria of a stochastic game Γ will be denoted by $E(\Gamma_\beta)$ for the β -discounted case ($\beta < 1$) as well as for the limiting average case ($\beta = 1$). Here, the emphasis is on stationary strategies, because such strategies have attractive properties. If one uses a stationary strategy, then one can disregard stage numbers as well as the full history of play at all stages. For any fixed pair of stationary strategies, play will take place as a Markov process on the set of states, which implies that the course of play is well structured. This Markov structure simplifies a great deal the computation of rewards. For nonstationary strategies, computation of rewards will often be very hard, even if the number of states is small. Another, more specific reason for focusing on stationary strategies is given in the concluding remarks (Section 4).

For bimatrix games, stochastic games with just one state and only one stage to play, the following theorem is well known (see Nash, Ref. 1).

Theorem 1.1. For every bimatrix game, there is at least one equilibrium.

For bimatrix games, Selten (Ref. 2) introduced the perfectness concept. The idea is as follows. Suppose that you are playing an equilibrium, but your opponent has a trembling hand and might make small mistakes. Then, your strategy should not only be a best reply to his equilibrium strategy, but to certain small perturbations of his strategy as well.

Selten (Ref. 2) introduced the notion of perturbed game to investigate such perfect equilibria. We will now extend Selten's definitions to stochastic games.

Definition 1.2. Let $\Gamma = \{M_1, M_2, \dots, M_z\}$ be a stochastic game, with mixed action space $\Delta^{m_s}[\Delta^{n_s}]$ for player 1 [2] in state s . For each $s \in S$, let $\epsilon_s^1 \in \mathbf{R}^{m_s}[\epsilon_s^2 \in \mathbf{R}^{n_s}]$ such that $\epsilon_s^1(i) > 0$ for all i and $\sum_{i=1}^{m_s} \epsilon_s^1(i) < 1$ [$\epsilon_s^2(j) > 0$ for all j and $\sum_{j=1}^{n_s} \epsilon_s^2(j) < 1$]. Such a pair (ϵ^1, ϵ^2) is called a mistake pair. The (ϵ^1, ϵ^2) -perturbed game is defined as the stochastic game $\Gamma(\epsilon^1, \epsilon^2)$ given by $\{M_1, M_2, \dots, M_z\}$, but where, for all s , player 1 [2] is restricted to mixed actions in

$$\Delta_{\epsilon^1}^{m_s} := \{x_s \in \Delta^{m_s} : x_s(i) \geq \epsilon_s^1(i) \text{ for all } i\}$$

[resp., $\Delta_{\epsilon^2}^{n_s}$ with similar definition].

Let $X_{\epsilon^1}[Y_{\epsilon^2}]$ be the set of stationary strategies of player 1 [2] in $\Gamma(\epsilon^1, \epsilon^2)$.

Definition 1.3. A pair of stationary strategies $(x, y) \in X \times Y$ is called perfect if there exists a sequence of mistake pairs $\{(\epsilon^{1n}, \epsilon^{2n}) : n \in \mathbf{N}\}$ converging to 0, and a sequence $\{(x^n, y^n) \in E(\Gamma_\beta(\epsilon^{1n}, \epsilon^{2n})) : n \in \mathbf{N}\}$ converging to (x, y) . Thus, for $\beta < 1$, we speak of perfect β -discounted pairs; and, for $\beta = 1$, we speak of perfect limiting average pairs. The set of perfect equilibria, i.e., perfect pairs which are equilibria of a stochastic game Γ_β , $\beta \in [0, 1]$, will be denoted by $PE(\Gamma_\beta)$.

Theorem 1.2. For every bimatrix game, there is at least one perfect equilibrium.

This theorem has been proved by Selten (Ref. 2).

In Section 3, we show the existence of perfect equilibria for β -discounted stochastic games as well as for irreducible limiting average stochastic games; we also give characterizations of perfect equilibria for these cases. A stochastic game is called irreducible if, for all pairs of stationary strategies, the related stochastic process on the set of states is an irreducible process. Recall that the idea behind the trembling-hand perfect

equilibrium is that players not only want to play a best reply to each other's strategy, but want to play a best reply to small perturbations thereof as well. Implicitly, this means that the reward function, β -discounted or limiting average, should be continuous on the set of stationary strategies. The β -discounted reward function is continuous for any stochastic game, while the limiting average reward function certainly is not. To have continuity of the limiting average reward function, one needs the irreducibility property (or at least that for any pair of stationary strategies there is precisely one ergodic set).

In Section 4, we give an example to illustrate phenomena that may occur in nonirreducible stochastic games.

First, we state some useful facts about stationary strategies in Section 2.

2. Preliminaries

For $(x, y) \in X \times Y$ and $k \in \{1, 2\}$, the following notations, facts, and lemmas will be used in the sequel. For Remarks 2.1 and 2.2, we refer to Kemeny and Snell (Ref. 3); for Lemma 2.1, we refer to Shapley (Ref. 4); and for Lemmas 2.2 and 2.3, we refer to Blackwell (Ref. 5).

Notation 2.1.

$$r^k(x, y) := (r^k(1, x_1, y_1), r^k(2, x_2, y_2), \dots, r^k(z, x_z, y_z)),$$

$$r^k(s, x_s, y_s) := \sum_{i=1}^{m_s} \sum_{j=1}^{n_s} x_s(i) r^k(s, i, j) y_s(j).$$

Notation 2.2. $P(x, y)$ is the $z \times z$ matrix in which the entry (s, t) is $p(t|s, x_s, y_s)$, which is given by

$$p(t|s, x_s, y_s) := \sum_{i=1}^{m_s} \sum_{j=1}^{n_s} x_s(i) p(t|s, i, j) y_s(j).$$

Notation 2.3.

$$Q(x, y) := \lim_{T \rightarrow \infty} (1/T) \sum_{n=1}^T (P(x, y))^n.$$

Remark 2.1. $P(x, y)$ and $Q(x, y)$ are stochastic matrices; i.e., all entries are nonnegative and all row sums are 1.

Remark 2.2.

$$Q(x, y)P(x, y) = Q(x, y) = P(x, y)Q(x, y).$$

Lemma 2.1. $\gamma_\beta^k(x, y)$ is the unique solution $v^k(x, y) \in \mathbf{R}^z$ of the equation

$$v^k(x, y) = (1 - \beta)r^k(x, y) + \beta P(x, y)v^k(x, y).$$

Lemma 2.2.

$$\gamma_1^k(x, y) = Q(x, y)r^k(x, y) = \lim_{\beta \rightarrow 1} \gamma_\beta^k(x, y).$$

Lemma 2.3. There exists a unique $w^k(x, y)$ such that

$$Q(x, y)w^k(x, y) = 0,$$

$$\gamma_1^k(x, y) + w^k(x, y) = r^k(x, y) + P(x, y)w^k(x, y).$$

Remark 2.3. If the stochastic game is irreducible, then all entries of $Q(x, y)$ are strictly positive and all rows of $Q(x, y)$ are equal to the unique stationary distribution of the stochastic process with transition matrix $P(x, y)$.

From Lemma 2.2 and Remark 2.3, one directly obtains the following remark.

Remark 2.4. For an irreducible stochastic game,

$$\gamma_1^k(s, x, y) = \gamma_1^k(t, x, y), \quad \text{for all } s, t \in S.$$

We also introduce the following notations (cf. Lemmas 2.1 and 2.3), which will enable us to relate best strategy choices in the multiple-state and infinite-horizon problem to best action choices in single-state and single-stage problems.

Notation 2.4. Let $(x, y) \in X \times Y$, $s \in S$, $a \in \Delta^{m_s}$, and $b \in \Delta^{n_s}$.

(i) For $\beta < 1$,

$$\varphi_\beta^1(a, s, x, y) := (1 - \beta)r^1(s, a, y_s) + \beta \sum_{t=1}^z p(t|s, a, y_s) \gamma_\beta^1(t, x, y),$$

$$\varphi_\beta^2(b, s, x, y) := (1 - \beta)r^2(s, x_s, b) + \beta \sum_{t=1}^z p(t|s, x_s, b) \gamma_\beta^2(t, x, y),$$

and

$$\varphi_1^1(a, s, x, y) := r^1(s, a, y_s) + \sum_{t=1}^z p(t|s, a, y_s) w^1(t, x, y),$$

$$\varphi_1^2(b, s, x, y) := r^2(s, x_s, b) + \sum_{t=1}^z p(t|s, x_s, b) w^2(t, x, y).$$

(ii) For $\beta \in [0, 1]$,

$$B_\beta^1(s, x, y) := \{\underline{a} \in \Delta^{m_s}: \varphi_\beta^1(\underline{a}, s, x, y) = \max_a \varphi_\beta^1(a, s, x, y)\},$$

$$B_\beta^2(s, x, y) := \{\underline{b} \in \Delta^{n_s}: \varphi_\beta^2(\underline{b}, s, x, y) = \max_b \varphi_\beta^2(b, s, x, y)\}.$$

(iii) For $\beta \in [0, 1]$, $x^* \in X$ and $y^* \in Y$,

$$\begin{aligned} & \varphi_\beta^1(x^*, x, y) \\ &:= (\varphi_\beta^1(x_1^*, 1, x, y), \varphi_\beta^1(x_2^*, 2, x, y), \dots, \varphi_\beta^1(x_z^*, z, x, y)), \\ & \varphi_\beta^2(y^*, x, y) \\ &:= (\varphi_\beta^2(y_1^*, 1, x, y), \varphi_\beta^2(y_2^*, 2, x, y), \dots, \varphi_\beta^2(y_z^*, z, x, y)). \end{aligned}$$

Observe that Lemma 2.1 and Notation 2.4 directly imply the following remark.

Remark 2.5.

$$\begin{aligned} \varphi_\beta^1(x_s, s, x, y) &= \gamma_\beta^1(s, x, y), \\ \varphi_\beta^2(y_s, s, x, y) &= \gamma_\beta^2(s, x, y), \quad \text{for all } \beta \in [0, 1]. \end{aligned}$$

The importance of these definitions becomes clear in the following lemmas.

Lemma 2.4. For a stochastic game Γ , let $x, x^* \in X$, $y \in Y$, and $\beta < 1$.

- (i) If $\varphi_\beta^1(x^*, x, y) \leq \gamma_\beta^1(x, y)$, then $\gamma_\beta^1(x^*, y) \leq \gamma_\beta^1(x, y)$.
- (ii) If $\varphi_\beta^1(x^*, x, y) \not\leq \gamma_\beta^1(x, y)$, then $\gamma_\beta^1(x^*, y) \not\leq \gamma_\beta^1(x, y)$.
- (iii) If $\varphi_\beta^1(x^*, x, y) \geq \gamma_\beta^1(x, y)$, then $\gamma_\beta^1(x^*, y) \geq \gamma_\beta^1(x, y)$.
- (iv) If $\varphi_\beta^1(x^*, x, y) \not\geq \gamma_\beta^1(x, y)$, then $\gamma_\beta^1(x^*, y) \not\geq \gamma_\beta^1(x, y)$.

The discerned cases in the above lemma can be proved using an iteration argument.

Lemma 2.5. For an irreducible stochastic game, let $x, x^* \in X$ and $y \in Y$.

- (i) If $\varphi_1^1(x^*, x, y) \leq \gamma_1^1(x, y) + w^1(x, y)$, then $\gamma_1^1(x^*, y) \leq \gamma_1^1(x, y)$.
- (ii) If $\varphi_1^1(x^*, x, y) \not\leq \gamma_1^1(x, y) + w^1(x, y)$, then $\gamma_1^1(x^*, y) \not\leq \gamma_1^1(x, y)$.
- (iii) If $\varphi_1^1(x^*, x, y) \geq \gamma_1^1(x, y) + w^1(x, y)$, then $\gamma_1^1(x^*, y) \geq \gamma_1^1(x, y)$.
- (iv) If $\varphi_1^1(x^*, x, y) \not\geq \gamma_1^1(x, y) + w^1(x, y)$, then $\gamma_1^1(x^*, y) \not\geq \gamma_1^1(x, y)$.

Each of the cases in the above lemma can be proved by multiplying both sides of the inequality with $Q(x^*, y)$ and using Lemmas 2.2, 2.3, Remarks 2.3, 2.4, and Notation 2.4.

The next lemma follows directly from Lemmas 2.4, 2.5, and Definition 1.1. It uses the functions φ_β^k to characterize stationary equilibria in the stochastic game.

Lemma 2.6. Let again $(x, y) \in X \times Y$. For $\beta < 1$, as well as for Γ irreducible and $\beta = 1$, (x, y) is an equilibrium if and only if, for all $s \in S$, we have:

$$\varphi_\beta^1(x_s, s, x, y) = \max_a \varphi_\beta^1(a, s, x, y),$$

$$\varphi_\beta^2(y_s, s, x, y) = \max_b \varphi_\beta^2(b, s, x, y).$$

In the following lemma, we characterize stationary equilibria in a perturbed stochastic game $\Gamma_\beta(\epsilon^1, \epsilon^2)$, using the sets B_β^k . Verbally, the lemma says that, if a pair of stationary strategies is an equilibrium, then there can only be surplus weight on actions that are among the best; conversely, if for a pair of stationary strategies all actions that have surplus weight are among the best, then this pair of strategies is an equilibrium. We use $e_i[f_j]$ to denote the degenerate mixed action: choose action $i[j]$ with probability 1.

Lemma 2.7. Let Γ be a stochastic game, let (ϵ^1, ϵ^2) be a mistake pair, and let $(x, y) \in X_{\epsilon^1} \times Y_{\epsilon^2}$. For $\beta < 1$, as well as for Γ irreducible and $\beta = 1$, the following statement holds: $(x, y) \in E(\Gamma_\beta(\epsilon^1, \epsilon^2))$, if and only if, for all s, i, j ,

$$[x_s(i) > \epsilon_s^1(i) \Rightarrow e_i \in B_\beta^1(s, x, y)],$$

$$[y_s(j) > \epsilon_s^2(j) \Rightarrow f_j \in B_\beta^2(s, x, y)].$$

Proof. The proof is analogous for the two discerned cases, so let $\beta \in [0, 1]$.

(i) The “only if” part. Suppose that $(x, y) \in E(\Gamma_\beta(\epsilon^1, \epsilon^2))$ and, for some s^* and i^* ,

$$x_{s^*}(i^*) > \epsilon_{s^*}^1(i^*) \quad \text{and} \quad e_{i^*} \notin B_\beta^1(s^*, x, y).$$

Take $i' \in B_\beta^1(s^*, x, y)$, and define $x' \in X_{\epsilon^1}$ by

$$\begin{aligned} x'_s &:= x_s, & \text{for } s \neq s^*, \\ x'_{s^*}(i) &:= x_{s^*}(i), & \text{for } i \notin \{i^*, i'\}, \\ x'_{s^*}(i') &:= x_{s^*}(i') + x_{s^*}(i^*) - \epsilon_{s^*}^1(i^*), \\ x'_{s^*}(i^*) &:= \epsilon_{s^*}^1(i^*). \end{aligned}$$

Using the linearity of $\varphi_\beta^1(\cdot, s, x, y)$, we derive that

$$\begin{aligned} \varphi_\beta^1(x'_{s^*}, s^*, x, y) &> \varphi_\beta^1(x_{s^*}, s^*, x, y), \\ \varphi_\beta^1(x'_s, s, x, y) &= \varphi_\beta^1(x_s, s, x, y), \quad \text{for all } s \neq s^*, \end{aligned}$$

which contradicts $(x, y) \in E(\Gamma_\beta(\epsilon^1, \epsilon^2))$; cf. Lemma 2.6.

(ii) The “if” part. Suppose that, for all s, i and j ,

$$\begin{aligned} [x_s(i) > \epsilon_s^1(i) &\Rightarrow e_i \in B_\beta^1(s, x, y)], \\ [y_s(j) > \epsilon_s^2(j) &\Rightarrow f_j \in B_\beta^2(s, x, y)]. \end{aligned}$$

Then, using the fact that

$$\varphi_\beta^1(a, s, x, y) = \sum_{i=1}^{m_s} a_i \varphi_\beta^1(e_i, s, x, y), \quad \text{for all } a \in \Delta^{m_s},$$

we derive that

$$\varphi_\beta^1(a, s, x, y) \leq \varphi_\beta^1(x_s, s, x, y), \quad \text{for all } a \in \Delta_{\epsilon_s^1}^{m_s}.$$

Similarly, one can show that

$$\varphi_\beta^2(b, s, x, y) \leq \varphi_\beta^2(y_s, s, x, y), \quad \text{for all } b \in \Delta_{\epsilon_s^2}^{n_s}.$$

In view of Lemma 2.6, this implies that

$$(x, y) \in E(\Gamma_\beta(\epsilon^1, \epsilon^2)). \quad \square$$

3. Perfect Equilibria in Stochastic Games

In this section, we will treat β -discounted stochastic games and irreducible limiting average stochastic games simultaneously. Therefore, Γ_β should be interpreted as a general β -discounted stochastic game for $\beta < 1$ and as an irreducible limiting average stochastic game for $\beta = 1$.

Theorem 3.1. For each stochastic game Γ and for every $\beta \in [0, 1]$, $E(\Gamma_\beta) \neq \emptyset$.

The proof of Theorem 3.1 is based on the facts that γ_β^k is a continuous function on $X \times Y$, for each k and β , and that X and Y are closed convex polyhedra; see Fink (Ref. 6) and Rogers (Ref. 7). Because for any mistake pair (ϵ^1, ϵ^2) , the sets X_{ϵ^1} and Y_{ϵ^2} are also closed convex polyhedra, the next lemma immediately follows.

Lemma 3.1. Let $\Gamma(\epsilon^1, \epsilon^2)$ be a perturbed stochastic game. Then,

$$E(\Gamma_\beta(\epsilon^1, \epsilon^2)) \neq \emptyset, \quad \text{for all } \beta \in [0, 1].$$

Lemma 3.2. If $(x, y) \in X \times Y$ is a perfect pair for Γ_β , $\beta \in [0, 1]$, then $(x, y) \in \text{PE}(\Gamma_\beta)$.

Proof. There exist a sequence of mistake pairs $\{(\epsilon^{1n}, \epsilon^{2n}): n \in \mathbb{N}\}$ converging to 0 and a sequence $\{(x^n, y^n) \in E(\Gamma_\beta(\epsilon^{1n}, \epsilon^{2n})): n \in \mathbb{N}\}$ converging to (x, y) . We have to show that, for all $(x^*, y^*) \in X \times Y$,

$$\gamma_\beta^1(x^*, y) \leq \gamma_\beta^1(x, y), \quad \gamma_\beta^2(x, y^*) \leq \gamma_\beta^2(x, y).$$

So, let $x^* \in X$, and take a sequence $\{x^{n*} \in X_{\epsilon^{1n}}: n \in \mathbb{N}\}$ converging to x^* . Then,

$$\gamma_\beta^1(x^{n*}, y^n) \leq \gamma_\beta^1(x^n, y^n), \quad \text{for all } n \in \mathbb{N}.$$

Using the continuity of γ_β^1 , we get

$$\gamma_\beta^1(x^*, y) \leq \gamma_\beta^1(x, y).$$

Similarly, one can show that

$$\gamma_\beta^2(x, y^*) \leq \gamma_\beta^2(x, y), \quad \text{for all } y^* \in Y. \quad \square$$

Lemma 3.3. $\text{PE}(\Gamma_\beta) \neq \emptyset$, for all $\beta \in [0, 1]$.

Proof. Let $\beta \in [0, 1]$. Take a sequence of mistake pairs $\{(\epsilon^{1n}, \epsilon^{2n}): n \in \mathbb{N}\}$ converging to 0. By Lemma 3.1, there exists a sequence $\{(x^n, y^n) \in E(\Gamma_\beta(\epsilon^{1n}, \epsilon^{2n})): n \in \mathbb{N}\}$. Because $(x^n, y^n) \in X \times Y$, for all $n \in \mathbb{N}$, and $X \times Y$ is a compact set, we may assume, without loss of generality, that the sequence $\{(x^n, y^n) \in E(\Gamma_\beta(\epsilon^{1n}, \epsilon^{2n})): n \in \mathbb{N}\}$ converges in $X \times Y$. Then, $\lim_{n \rightarrow \infty} (x^n, y^n)$ is a perfect equilibrium in Γ_β by Lemma 3.2. $\square$

In Theorem 3.2, we give two characterizations of perfect equilibria. First, we need some definitions. We define completely mixed strategies as stationary strategies that use all actions with positive probability. We also

define ϵ -perfect pairs as pairs of completely mixed strategies where, for a specific mistake pair (constant ϵ for all actions), any surplus weight of the strategies is on best actions; see Lemma 2.7.

Definition 3.1. A stationary strategy $x \in X$ [$y \in Y$] is called completely mixed if, for all s, i, j ,

$$x_s(i) > 0 \text{ } [y_s(j) > 0].$$

Definition 3.2. Let $\epsilon > 0$ and $\beta \in [0, 1]$. A pair of completely mixed strategies (x, y) is called ϵ -perfect in Γ_β if, for all s, i, j ,

$$[x_s(i) > \epsilon \Rightarrow e_i \in B_\beta^1(s, x, y)],$$

$$[y_s(j) > \epsilon \Rightarrow f_j \in B_\beta^2(s, x, y)].$$

Notice from Lemma 2.7 that a pair of stationary strategies is an ϵ -perfect pair if and only if it is an equilibrium in the specific perturbed game where all weights are perturbed by the same $\epsilon > 0$.

The next lemma tells us that, in order to decide whether or not a pair of stationary strategies is a perfect equilibrium, it is sufficient to consider equilibria of such constant ϵ -perturbed games, where ϵ tends to zero. Hence, one can restrict the analysis to a relatively small class of perturbed games in order to investigate perfectness for a given pair of stationary strategies.

Theorem 3.2. Let $(x, y) \in X \times Y$ and let $\beta \in [0, 1]$. The following three statements are equivalent:

- (i) (x, y) is a perfect equilibrium in Γ_β .
- (ii) There is a sequence $\{\epsilon^n \in (0, 1): n \in \mathbb{N}\}$ converging to 0 and a sequence $\{(x^n, y^n) \in X \times Y: n \in \mathbb{N}\}$ converging to (x, y) such that (x^n, y^n) is ϵ^n -perfect in Γ_β for all $n \in \mathbb{N}$.
- (iii) There is a sequence of completely mixed pairs $\{(x^n, y^n) \in X \times Y: n \in \mathbb{N}\}$ converging to (x, y) such that, for all s, i, j, n

$$[x_s(i) > 0 \Rightarrow e_i \in B_\beta^1(s, x^n, y^n)],$$

$$[y_s(j) > 0 \Rightarrow f_j \in B_\beta^2(s, x^n, y^n)].$$

Proof. We successively show that (i) $\rightarrow$ (ii), (ii) $\rightarrow$ (iii), (iii) $\rightarrow$ (i).

(i) $\rightarrow$ (ii). Suppose that $(x, y) \in \text{PE}(\Gamma_\beta)$. Then, there is a sequence of mistake pairs $\{(\epsilon^{1n}, \epsilon^{2n}): n \in \mathbb{N}\}$ converging to 0 as well as a sequence $\{(x^n, y^n) \in E(\Gamma_\beta(\epsilon^{1n}, \epsilon^{2n})): n \in \mathbb{N}\}$ converging to (x, y) . Let

$$\epsilon^n := \max_{s, i, j} \{\epsilon_s^{1n}(i), \epsilon_s^{2n}(j)\}, \quad \text{for each } n \in \mathbb{N}.$$

Then, by Lemma 2.7 and by Definition 3.2, each pair (x^n, y^n) is ϵ^n -perfect.

(ii) $\rightarrow$ (iii). Suppose that ϵ^n and (x^n, y^n) , $n \in \mathbb{N}$, are as in (ii). Then, by definition, (x^n, y^n) is completely mixed for all n . Since

$$\lim_{n \rightarrow \infty} (x^n, y^n) = (x, y),$$

it holds that, if $x_s(i) > 0$, then $x_s^n(i) > \epsilon^n$ for large n . Thus, by Definition 3.2,

$$e_i \in B_\beta^1(s, x^n, y^n), \quad \text{for large } n.$$

(iii) $\rightarrow$ (i). Let $\{(x^n, y^n): n \in \mathbb{N}\}$, and let (x, y) be as in (iii). For all $n \in \mathbb{N}$, for all s, i, j , define

$$\epsilon_s^{1n}(i) := n^{-1}, \text{ if } x_s(i) > 0; \quad \epsilon_s^{1n}(i) := x_s^n(i), \text{ otherwise;}$$

$$\epsilon_s^{2n}(j) := n^{-1}, \text{ if } y_s(j) > 0; \quad \epsilon_s^{2n}(j) := y_s^n(j), \text{ otherwise.}$$

Then, $\Gamma_\beta(\epsilon^{1n}, \epsilon^{2n})$ is a well-defined perturbed game for large n . For large n , for $s \in S$ and i such that $x_s(i) > 0$, we have $x_s^n(i) > n^{-1} = \epsilon_s^{1n}(i)$; and, for s and i such that $x_s(i) = 0$, we have $x_s^n(i) = \epsilon_s^{1n}(i)$. Hence, $x^n \in X_{\epsilon^{1n}}$ for large n ; and, by (iii), we have that $x_s^n(i) > \epsilon_s^{1n}(i) > 0$ implies $e_i \in B_\beta^1(s, x^n, y^n)$. Using the same argument for player 2 and using Lemma 2.7, we have shown (i). $\square$

4. Concluding Remarks

One of our reasons to consider this notion of trembling-hand perfect equilibrium for stochastic games was to be able to work with particular sequences of stationary β -discounted equilibria, for β tending to 1. In Vrieze and Thuijsman (Ref. 8) and more generally in Thuijsman (Ref. 9), it is shown that, for certain classes of stochastic games, one can derive limiting average ϵ -equilibria by examining arbitrary sequences of stationary β -discounted equilibria. However, the existence of limiting average ϵ -equilibria in general stochastic games is still an open problem. Now, being able to use sequences of perfect stationary β -discounted equilibria may bring us closer to a solution for this challenging problem.

To illustrate the fact that trembling-hand perfectness does not make much sense in nonirreducible limiting average stochastic games, we examine

the following example:

2,2 1	0,0 2	0,0 2
0,0 2	1,0 1	0,1 1
0,0 2	0,1 3	1,0 4
state 1	state 2	state 3

Here, the notation for state 3, for example, means a payoff 0 [1] to player 1 [2] and a transition to state 3 with probability 1. For this game, a stationary strategy is completely determined by the mixed action for state 1. If we let $x = e_1$ and let $y = f_1$, then $(x, y) \in E(\Gamma_1)$. So there is at least one stationary limiting average equilibrium for this game, but is there also a perfect one?

Let (ϵ^1, ϵ^2) be a mistake pair for state 1, and let

$$(x', y') \in E(\Gamma_1(\epsilon^1, \epsilon^2)).$$

Then,

$$y'(2) = 1 - \epsilon^2(1) - \epsilon^2(3),$$

because player 2's limiting average reward is maximal if column 2 gets maximal weight, as to maximize the probability of absorption in state 3. So, if (x^*, y^*) is a perfect limiting average equilibrium, then $y^* = f_2$. This, however, would imply that $x^* = e_2$ in case (x^*, y^*) is an equilibrium. But then the only best answer to x^* for player 2 is $f_3 \neq y^*$, which contradicts the assumption that (x^*, y^*) is an equilibrium. Hence, there are no perfect limiting average equilibria in this stochastic game, even though a stationary limiting average equilibrium exists. One can also check that (x, y) is a perfect β -discounted equilibrium for all $\beta < 1$.

It should be observed that any stationary equilibrium in a β -discounted or irreducible limiting average stochastic game is necessarily a subgame-perfect equilibrium.

Characterizations in this paper for perfect equilibria in stochastic games are similar to characterizations for perfect equilibria in bimatrix games (see Ref. 10).

References

1. NASH, J. F., *Equilibrium in a Stochastic n-Person Game*, Proceedings of the National Academy of Sciences USA, Vol. 36, pp. 48-49, 1950.
2. SELTEN, R., *Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games*, International Journal of Game Theory, Vol. 4, pp. 25-55, 1975.
3. KEMENY, J. G., and SNELL, J. L., *Finite Markov Chains*, Van Nostrand, Princeton, New Jersey, 1961.
4. SHAPLEY, L. S., *Stochastic Games*, Proceedings of the National Academy of Sciences USA, Vol. 39, pp. 1095-1100, 1953.
5. BLACKWELL, D., *Discrete Dynamic Programming*, Annals of Mathematical Statistics, Vol. 33, pp. 719-726, 1962.
6. FINK, A. M., *Equilibrium in a Stochastic n-Person Game*, Journal of Science of Hiroshima University, Series A-I 28, pp. 89-93, 1964.
7. ROGERS, P. D., *Nonzero-Sum Stochastic Games*, PhD Thesis and Report No. ORC-69-8, University of California, Berkeley, 1969.
8. VRIEZE, O. J., and THUIJSMAN, F., *On Equilibria in Repeated Games with Absorbing States*, International Journal of Game Theory, Vol. 18, pp. 293-310, 1989.
9. THUIJSMAN, F., *Optimality and Equilibria in Stochastic Games*, PhD Thesis, University of Limburg, Maastricht, The Netherlands, 1989.
10. TIJS, S. H., *Noncooperative Game Theory*, Lecture Notes, Catholic University, Nijmegen, The Netherlands, 1987.