A Survey on Optimality and Equilibria in Stochastic Games

Frank Thuijsman

Department of Mathematics, Maastricht University[1]

Abstract

In this paper we discuss the main existence results on optimality and equilibria in two-person stochastic games with finite state and action spaces. Several examples are provided to clarify the issues.

1 The Stochastic Game Model

In this introductory section we give the necessary definitions and notations for the two-person case of the stochastic game model and we briefly present some basic results. In section 2 we discuss the main existence results for zero-sum stochastic games, while in section 3 we focus on general-sum stochastic games. In each section we discuss several examples to illustrate the most important phenomena.

It all started with the fundamental paper by Von Neumann [1928] in which he proves the so called minimax theorem which says that for each finite matrix of reals $A = [a_{ij}]_{i=1,j=1}^{m\ \ n}$ there exist probability vectors $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m)$ and $\bar{y} = (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$ such that for all x and y it holds that[2] $xA\bar{y} \leq \bar{x}A\bar{y} \leq \bar{x}Ay$. In other words: $\max_x \min_y xAy = \min_y \max_x xAy$. This theorem can be interpreted to say that each matrix game has a value. A matrix game A is played as follows. Simultaneously, and independent from each other, player 1 chooses a row i and player 2 chooses a column j of A. Then player 2 has to pay the amount a_{ij} to player 1. Each player is allowed to randomize over his available actions and we assume that player 1 wants to maximize his expected payoff, while player 2 wants to minimize the expected payoff to player 1. The minimax theorem tells us that, for each matrix A there is a unique amount val(A), which player 1 can guarantee as his minimal expected payoff, while at the same time player 2 can guarantee that the expected payoff to player 1 will be at most this amount.

Later Nash [1951] considered the n-person extension of matrix games, in the sense that all n players, simultaneously and independently choose actions that determine a payoff for each and every one of them. Nash [1951] showed that in such games there always exists at least one (Nash-)equilibrium: a tuple of strategies such that each player is playing a best reply against the joint strategy of his opponents. For the two-player case this boils down to a “bimatrix game” where players 1 and 2 receive a_{ij} and b_{ij} respectively in case their choices determine entry (i, j). The result of Nash says that there exist $\bar{x}$ and $\bar{y}$ such that for all x and y it holds that $\bar{x}A\bar{y} \geq xA\bar{y}$ and $\bar{x}B\bar{y} \geq \bar{x}By$, where $A = [a_{ij}]$ and $B = [b_{ij}]$ are finite matrices of the same size.

Shapley [1953] introduced dynamics into game theory by considering the situation that at discrete stages in $I\!N$ the players play one of finitely many matrix games, where the choices of the players determine a payoff to player 1 (by player 2) as well as a stochastic transition to go to a next matrix game. He called these games “stochastic games”, which brings us to the topic of this paper. Formally, a two-person stochastic game with finite state and action spaces can be represented by a finite set of matrices $A^1, A^2, \ldots, A^z$ corresponding to the set of states $S = \{1, 2, \ldots, z\}$. For $s \in S$ matrix A^s has size $m_s \times n_s \in I\!N \times I\!N$ and entry (i, j) of A^s contains:

a) a payoff $r^k(s, i, j) \in I\!R$ for each player $k \in \{1, 2\}$

b) a transition probability vector $p(s, i, j) = (p(1|s, i, j), p(2|s, i, j), \ldots, p(z|s, i, j))$ where $p(t|s, i, j)$ is the probability of a transition from s to t whenever entry (i, j) of A^s is selected.

Play can start in any state of S and evolves by players independently choosing actions i_n and j_n of A^{s_n}, where s_n denotes the state visited at stage n. In case $r^1(s, i, j) + r^2(s, i, j) = 0$, the game is called zero-sum, otherwise it is called general-sum. In zero-sum games players have strictly opposite interests, since they are paying each other.

A strategy for a player is a rule that tells him for any history $h_n = (s_1, i_1, j_1, s_2, i_2, j_2, \ldots, s_{n-1}, i_{n-1}, j_{n-1}, s_n)$ up to stage n, what mixed action to use in state s_n at stage $n \in I\!N$. Such behavior strategies will be denoted by π for player 1 and by σ for player 2.

[1]Further address: P.O. Box 616, 6200 MD Maastricht, The Netherlands (frank@math.unimaas.nl)

[2]Note that we do not distinguish row vectors from column vectors. In the matrix products this should be clear from the context.

For initial state s and any pair of strategies (π, σ) the limiting average reward and the β-discounted reward, $\beta \in (0,1)$, to player $k \in \{1,2\}$ are respectively given by

$$\gamma^k(s, \pi, \sigma) = E_{s\pi\sigma}\left(\liminf_{T\to\infty} \tfrac{1}{T}\sum_{n=1}^{T} r^k(S_n, I_n, J_n)\right) \tag{1}$$

$$\gamma^k_\beta(s, \pi, \sigma) = E_{s\pi\sigma}\left((1-\beta)\sum_{n=1}^{\infty} \beta^{n-1} r^k(S_n, I_n, J_n)\right) \tag{2}$$

where S_n, I_n, J_n are random variables for the state and actions at stage n. Let $\gamma^k(\pi,\sigma)$ and $\gamma^k_\beta(\pi,\sigma)$ denote vectors of rewards with coordinates corresponding to the initial states.
A stationary strategy for a player consists of a mixed action for each state, to be used whenever that state is being visited, regardless of the history. Stationary strategies for player 1 are denoted by $x = (x_1, x_2, \ldots, x_z)$, where x_s is the mixed action to be used in state s. For player 2's stationary strategies we write y. A pair (x,y) of stationary strategies determines a Markov-chain (with transition matrix) $P(x,y)$ on S, where entry (s,t) of $P(x,y)$ is $p(t|s,x_s,y_s) = \sum_{i=1}^{m_s}\sum_{j=1}^{n_s} x_s(i)p(t|s,i,j)y_s(j)$. If we use the notation $r^k(x,y) = (r^k(1,x_1,y_1),\ r^k(2,x_2,y_2), \ldots, r^k(z,x_z,y_z))$ with $r^k(s,x_s,y_s) = \sum_{i=1}^{m_s}\sum_{j=1}^{n_s} x_s(i)r^k(s,i,j)y_s(j)$ then

$$\gamma^k_\beta(x,y) = (1-\beta)(I - \beta P(x,y))^{-1} r^k(x,y) \tag{3}$$

where I is the identity matrix, and

$$\gamma^k(x,y) = Q(x,y) r^k(x,y) \tag{4}$$

with

$$Q(x,y) = \lim_{T\to\infty} \tfrac{1}{T}\sum_{n=1}^{T}(P(x,y))^n \tag{5}$$

It is well-known (cf. Blackwell [1962]) that

$$Q(x,y)P(x,y) = Q(x,y) \tag{6}$$

$$Q(x,y) = \lim_{\beta\uparrow 1} (1-\beta)(I-\beta P(x,y))^{-1} \tag{7}$$

and hence (3), (4) and (7) give

$$\gamma^k(x,y) = \lim_{\beta\uparrow 1} \gamma^k_\beta(x,y) \tag{8}$$

Notice that (5) and (6) imply that row s of $Q(x,y)$ is the unique stationary distribution for the Markov chain $P(x,y)$ starting in state s. A stationary strategy x is called pure if $\#\{i : x_s(i) > 0\} = 1$ for all s. Pure stationary strategies shall be denoted by f and g for players 1 and 2 respectively. It is well-known (cf. Hordijk et al. [1983]) that, when playing against a fixed stationary strategy, a player always has a pure stationary best reply:

$$\forall\beta\ \forall y\ \exists f^1, f^1_\beta\ \forall\pi\ \left[\gamma^1(f^1,y) \geq \gamma^1(\pi,y) \text{ and } \gamma^1_\beta(f^1_\beta,y) \geq \gamma^1_\beta(\pi,y)\right] \tag{9}$$

Obviously, for player 2's best replies an analogon of (9) holds.

Finally, we wish to mention one more type of strategies, namely Markov strategies. These are strategies that, at any stage of play, prescribe actions that only depend on the current state and stage. Thus, the past actions of the opponent are not being taken into account. Strategies for which these choices do depend on those past actions shall be called history dependent.

2 Zero-sum Stochastic Games

In zero-sum stochastic games it is customary to consider only the payoffs to player 1, which player 1 wishes to maximize and player 2 wants to minimize. Since in the sequel we also consider the zero-sum situation where player 2 is the maximizer and player 1 the minimizer, we shall incorporate the player number in the definitions of value and optimal strategies. Thus, for $k = 1,2$, the k-zerosum game is the stochastic game determined by player k's payoffs, where player k is maximizing his reward while the other player is minimizing player k's reward.

In his ancestral paper on stochastic games Shapley [1953] shows

$$\forall\beta\ \exists v^1_\beta\ \exists x^1_\beta, y^1_\beta\ \forall\pi,\sigma\ \left[\gamma^1_\beta(x^1_\beta,\sigma) \geq v^1_\beta \geq \gamma^1_\beta(\pi, y^1_\beta)\right] \tag{10}$$

The vector v^1_β is called the β-discounted 1-value and the strategies x^1_β, y^1_β are called stationary β-discounted optimal strategies in the 1-zerosum game. Shapley's proof is based on the observation that v^1_β is the unique solution of the following system of equations:

$$\alpha_s = \ \mathrm{val}[(1-\beta)r^1(s,i,j) + \beta\sum_t p(t|s,i,j)\alpha_t]_{i=1,j=1}^{m_s\ \ n_s} \quad ,s\in S \tag{11}$$

where val denotes the matrix game value operator.

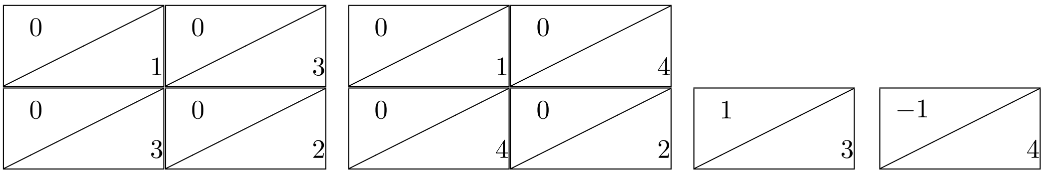
Everett [1957] and Gillette [1957] were the first to consider undiscounted rewards. Everett [1957] examined recursive games, which can be defined as stochastic games where the only non-zero payoffs can be obtained in absorbing states, i.e. states that have the property that once play gets there, it remains there forever. Although optimal strategies need not exist for such games, Everett [1957] shows that for each recursive game the limiting average value v^1 exists, and can be achieved by using stationary ε-optimal strategies $x^1_\varepsilon, y^1_\varepsilon$. Precisely:

$$\exists v^1 \ \forall \varepsilon > 0 \ \exists x^1_\varepsilon, y^1_\varepsilon \ \forall \pi, \sigma \left[\gamma^1(x^1_\varepsilon, \sigma) + \varepsilon 1_z \geq v^1 \geq \gamma^1(\pi, y^1_\varepsilon) - \varepsilon 1_z \right] \tag{12}$$

Here 1_z denotes the vector $(1, 1, \ldots, 1)$ in IR^z.

Example 2.1

Consider the following recursive game.



1 2 3 4

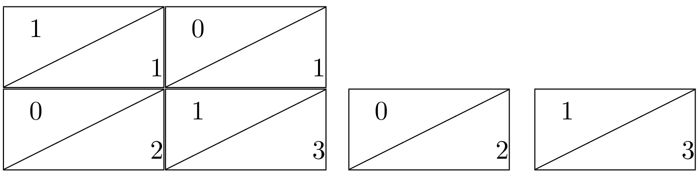
To explain this notation: Player 1 chooses rows; player 2 chooses columns; for each entry the above diagonal number is the payoff to player 1 and the below diagonal number is the state at which play is to proceed; in case of a stochastic transition we write the transition probability vector at this place.

States 3 and 4 are absorbing and obviously states 1 and 2 are the only interesting initial states. For this game the limiting average value is $v^1 = (1, -1, 1, -1)$. For player 1 a stationary limiting average ε-optimal strategy is given by $((1-\varepsilon, \varepsilon), (1, 0))$ for states 1 and 2 respectively (clearly, in states 3 and 4 he can only choose the one available action). As can be verified using (11), the β-discounted value is $v^1_\beta = (\frac{1-\sqrt{1-\beta^2}}{\beta}, \frac{-1+\sqrt{1-\beta^2}}{\beta}, 1, -1)$ and for player 1 the unique stationary β-discounted optimal strategies are given by playing Top, his first action, with probability $\frac{1-\beta^2-\sqrt{1-\beta^2}}{\beta-\beta^2-\beta\sqrt{1-\beta^2}}$ in state 1 as well as in state 2.

An elementary proof for Everett's [1957] result is given by Thuijsman & Vrieze [1992], where for the recursive game situation a stationary limiting average ε-optimal strategy is constructed from an arbitrary sequence of stationary β_n-discounted optimal strategies, with $\lim_{n\to\infty} \beta_n = 1$.

Example 2.2

This famous game is the so called big match introduced by Gillette [1957].



1 2 3

For this game the unique stationary β-discounted optimal strategies are given by $x^1_\beta = (\frac{1}{2-\beta}, \frac{1-\beta}{2-\beta})$ and $y^1_\beta = (\frac{1}{2}, \frac{1}{2})$ for players 1 and 2 respectively, and $v^1_\beta = \frac{1}{2}$ for initial state 1. However, it was not clear for a long time, whether or not the limiting average value would exist. The problem was that against any Markov strategy for player 1 and for any $\varepsilon > 0$ player 2 has a Markov strategy such that player 1's limiting average reward is less than ε. On the other hand, player 2 can guarantee that he has to pay a limiting average reward of at most $\frac{1}{2}$, but he cannot guarantee anything less than $\frac{1}{2}$. Hence there is an apparent gap between the amounts the players can guarantee using only Markov strategies. The matter was settled by Blackwell & Ferguson [1968], who formulated, for arbitrary $\varepsilon > 0$, a history dependent strategy for player 1 which guarantees a limiting average reward of at least $\frac{1}{2} - \varepsilon$ against any strategy of player 2. This history dependent limiting average ε-optimal stratey is of the following type. At stage n suppose that play is still in state 1 where player 2 has chosen left $l(n)$ times, while he has chosen right $r(n)$ times. Then, player 1 should play Bottom (his second action) with probability $\varepsilon^2(1-\varepsilon)^{k(n)}$, where $k(n) = \max\{0, l(n) - r(n)\}$.

Later, this result on the big match was generalized by Kohlberg [1974], who showed that every repeated game with absorbing states has a limiting average value. A repeated game with absorbing states is a stochastic game in which, just like in the big match, all states but one are absorbing.

Finally, by an ingeneous proof Mertens & Neyman [1981] showed that for every stochastic game the limiting average value exist. Their proof exploits the remarkable observation by Bewley & Kohlberg [1976] that the β-discounted value as well as the stationary β-discounted optimal strategies can be expanded as Puiseux series in powers of $1-\beta$. For example, for the above big match we have that $x^1_\beta = (1,0) + (-1,1)(1-\beta) + (1,-1)(1-\beta)^2 + (-1,1)(1-\beta)^3 + \ldots$

Apart from these general results, specially structured stochastic games have been examined. We already discussed recursive games and repeated games with absorbing states, but we should also mention the following classes: irreducible/unichain stochastic games (cf. Rogers [1969], Sobel [1971], Federgruen [1978]), i.e. stochastic games for which for any pair of stationary strategies the related Markov chain is irreducible/unichain; single controller stochastic games (cf. Parthasarathy & Raghavan [1981]), i.e. games in which the transitions only depend on the actions of one and the same player for all states; switching control stochastic games (cf. Filar [1981], Vrieze et al. [1983]), i.e. games with transitions for each state depending on the action of only one player; perfect information stochastic games (cf. Liggett & Lippman [1969]), where in each state one of the players has only one action available; stochastic games with additive rewards and additive transitions ARAT (cf. Raghavan et al. [1985]), i.e. there are r^k_1, r^k_2, p_1, p_2 such that $r^k(s,i,j) = r^k_1(s,i) + r^k_2(s,j)$ and $p(s,i,j) = p_1(s,i) + p_2(s,j)$ for all s,i,j; stochastic games with separable rewards and state independent transitions (cf. Parthasarathy et al. [1984]), i.e. there are r^k_a, r^k_b, p_a such that $r^k(s,i,j) = r^k_a(s) + r^k_b(i,j)$ and $p(s,i,j) = p_a(i,j)$ for all s,i,j. All these classes admit stationary limiting average optimal strategies. Later, in Thuijsman & Vrieze [1991, 1992] and in Thuijsman [1992] new (and far more simple) proofs were provided for the existence of stationary solutions in several of these classes. Characterizations, in terms of game properties, for the existence of stationary limiting average optimal strategies are provided in Vrieze & Thuijsman [1987], Filar et al. [1991] and Thuijsman [1992].

Before closing this section on optimality we mention the result by Tijs & Vrieze [1986] (also see Vrieze [1987]) who show that for every stochastic game there is for each player a non-empty set of initial states for which a stationary limiting average optimal strategy exists. This proof relies on the Puiseux series work by Bewley & Kohlberg [1976]. A new and direct proof for the same result is given in Thuijsman & Vrieze [1991], Thuijsman [1992]. A detailed study of the possibilities for limiting average optimality by means of stationary strategies can be found in Thuijsman & Vrieze [1993], while in Flesch et al. [1996b] it is shown that the existence of a limiting average optimal strategy implies the existence of stationary limiting average ε-optimal strategies.

3 General-sum Stochastic Games

One of the first persons to examine non-zerosum stochastic games was Fink [1964], who showed the existence of stationary β-discounted equilibria for stochastic games:

$$\forall\beta\ \exists x,y\ \forall\pi,\sigma\ \left[\gamma^1_\beta(\pi,y) \leq \gamma^1_\beta(x,y) \text{ and } \gamma^2_\beta(x,\sigma) \leq \gamma^2_\beta(x,y)\right] \tag{13}$$

Since, by its definition, for the zero-sum situation an equilibrium can only consist of a pair of optimal strategies, the big match (cf. example 2.2) immediately shows that limiting average equilibria do not always exist. Where we introduced ε-optimal strategies for the zero-sum case, we now have to introduce ε-equilibria for the general-sum case. A pair of strategies $(\pi_\varepsilon, \sigma_\varepsilon)$ is called a limiting average ε-equilibrium ($\varepsilon > 0$) if neither player 1 nor player 2 can gain more than ε by a unilateral deviation. To put it precisely

$$\forall\pi,\sigma\ [\gamma^1(\pi,\sigma_\varepsilon) \leq \gamma^1(\pi_\varepsilon,\sigma_\varepsilon) + \varepsilon 1_z \text{ and } \gamma^2(\pi_\varepsilon,\sigma) \leq \gamma(\pi_\varepsilon,\sigma_\varepsilon) + \varepsilon 1_z] \tag{14}$$

The existence of limiting average ε-equilibria for arbitrary general-sum stochastic games has not yet been established. Neither do we know of any counterexample to their existence. The most general results on the existence of equilibria are the following. First it was observed that in every stochastic game there is a non-empty set of initial states for which ε-equilibria exist (cf. Thuijsman & Vrieze [1991], Thuijsman [1992] or Vieille [1993]). Our proof of this result was based on ergodicity properties of a converging sequence of stationary β_n-discounted equilibra, with $\lim_{n\to\infty}\beta_n = 1$. However, the equilibrium strategies are of a behavioral type: at all stages players must take into account the history of past moves of their opponent. Nevertheless, a side-result of this approach was a simple and straightforward proof for the existence of stationary limiting average equilibria for irreducible/unichain stochastic games (which was earlier derived by Rogers [1969], Sobel [1971], Federgruen [1978]).

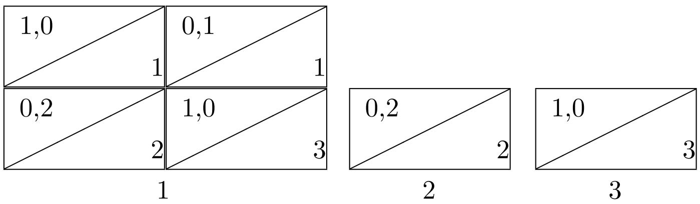
Concerning the (simultaneous) existence of limiting average ε-equilibria for all initial states, sufficient conditions have been formulated in Thuijsman [1992], which are based on properties of a converging sequence of stationary β_n-discounted equilibra, with $\lim_{n\to\infty}\beta_n = 1$, while in Thuijsman & Vrieze [1997] quite general sufficient conditions have been formulated in terms of stationary strategies, and of observability and punishability of deviations. We call this the threat approach, since the players are constantly checking after each other, and any "wrong" move of the opponent will immediately trigger a punishment. Thus the threats are the stabilizing force in the limiting average equilibria. Using this threat approach existence of ε-equilibria has been shown for repeated games with absorbing states (cf. Vrieze & Thuijsman [1989], where a prototype threat approach is being used), as well as for stochastic games with state independent transitions (cf. Thuijsman [1992]), as well as for stochastic games with three states (cf. Vieille [1993]), as well as for stochastic games with switching control (cf. Thuijsman & Raghavan [1997]), and existence of pure 0-equilibria has been shown for stochastic games with additive rewards and additive transitions (ARAT, cf. Thuijsman & Raghavan [1997]), which includes the perfect information games.

We remark that previous to our threat approach for none of these classes, the existence of limiting average ε-equilibria was known, even though the zero-sum solutions had been derived a long time ago. Also note that even for perfect information stochastic games stationary limiting average equilibria generallly do not exist, although for the zero-sum case pure stationary limiting average optimal strategies are available (cf. Liggett & Lippman [1969]). Example 3.2 below will illustrate this point.

For recursive repeated games with absorbing states (cf. Flesch et al. [1996a]) and for ARAT repeated games with absorbing states (cf. Evangelista et al. [1997]) the existence of stationary limiting average ε-equilibria has been shown (without threats).

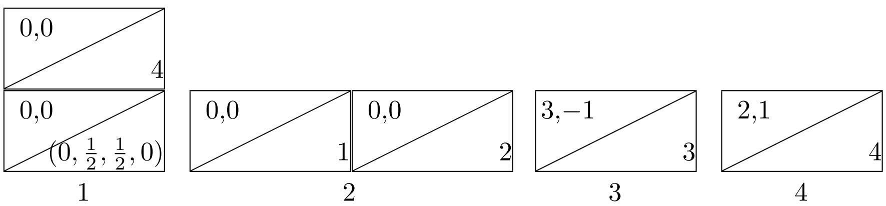
We conclude this paper with three very special examples. In example 3.1 we examine a repeated game with absorbing states for which there is a gap between the β-discounted equilibrium rewards and the limiting average equilibrium rewards. In example 3.2 we discuss a perfect information stochastic game which does not have stationary limiting average ε-equilibria, but where the only equilibria known to us, are of the threat type. In example 3.3 we discuss a three person recursive repeated game with absorbing states for which the only limiting average equilibria consist of cyclic Markov strategies. This is very remarkable since, in that game, neither history dependent nor stationary limiting average ε-equilibria do exist.

Example 3.1



This is an example of a repeated game with absorbing states, where play remains in the initial state 1 as long as player 1 chooses Top, but play reaches an absorbing state as soon as player 1 ever chooses Bottom. Sorin [1986] examined this example in great detail. The 1-zero-sum and 2-zero-sum limiting average values (for initial state 1) are given by $(v^1, v^2) = (\frac{1}{2}, \frac{2}{3})$. Clearly then, there can be no stationary limiting average ε-equilibrium, because against any stationary strategy of player 1, player 2 can get at least 1, and by doing so player 1 would get $0 < v^1$, which he can always achieve by playing limiting average ε-optimal in the 1-zero-sum game. However, for each pair in $\text{Conv}\{(\frac{1}{2}, 1), (\frac{2}{3}, \frac{2}{3})\}$, where Conv stands for convex hull, Sorin [1986] gives history dependent limiting average ε-equilibria that yield this pair as an equilibrium reward. Besides, he shows that any limiting average ε-equilibrium corresponds to a reward in $\text{Conv}\{(\frac{1}{2}, 1), (\frac{2}{3}, \frac{2}{3})\}$, while all β-discounted equilibria yield $(\frac{1}{2}, \frac{2}{3})$. Although this observation suggests that the limiting average general-sum case can not be approached from the β-discounted general-sum case, by studying this example Vrieze & Thuijsman [1989] discovered a general principle to construct, starting from any arbitrary sequence of stationary β_n-discounted equilibria with $\lim_{n\to\infty} \beta_n = 1$, a limiting average ε-equilibrium.

Example 3.2



This game is a recursive perfect information game for which there is no stationary limiting average ε-equilibrium. One can prove this as follows. Suppose player 2 puts positive weight on Left in state 2, then player 1's only stationary limiting average ε-best replies are those that put weight at most $\frac{\varepsilon}{2-\varepsilon}$ on Top in state 1; against any of these strategies, player 2's only stationary limiting average ε-best replies are those that put weight 0 on Left in state 2. So there is no stationary limiting average ε-equilibrium where player 2 puts positive weight on Left in state 2. But there is neither a stationary limiting average ε-equilibrium where player 2 puts weight 0 on Left in state 2, since then player 1 should put at most 2ε weight on Bottom in state 1, which would in turn contradict player 2's putting weight 0 on Left. Following the construction of Thuijsman & Raghavan [1997], where existence of limiting average 0-equilibria is shown for arbitrary n-person games with perfect informtion, we can find an equilibrium by the following procedure. Take a pure stationary limiting average optimal strategy f^1 for player 1 (this exists by Liggett & Lippman [1969]); let g^1 be pure stationary limiting average optimal for player 2 in the 1-zero-sum game; let g^2 be a pure stationary limiting average best reply for player 2 against f^1 in the 2-zero-sum game (which exists by (9)). Now define g^* for player 2 by: play g^2 unless at some stage player 1 has ever deviated from playing f^1, then play g^1. Here, $f^1 = (1, 0) = g^2$ and $g^1 = (0, 1)$. Now it can be verified that (f^1, g^*) is a limiting average equilibrium.

Example 3.3

N

	L	R
T	0,0,0	0,1,3 *
B	1,3,0 *	1,0,1 *

F

3,0,1 *	1,1,0 *
0,1,1 *	0,0,0 *

This is a three-person recursive repeated game with absorbing states, where an asterisk in any particular entry denotes a transition to an absorbing state with the same payoff as in this particular entry. There is only one entry for which play will remain in the non-trivial initial state. One should picture the game as a $2 \times 2 \times 2$ cube, where the layers belonging to the actions of player 3 (Near and Far) are represented separately. As before, player 1 chooses Top or Bottom and player 2 chooses Left or Right. The entry (T, L, N) is the only non-absorbing entry for the initial state. Hence, as long as play is in the initial state the only possible history is the one where entry (T, L, N) was played at all previous stages. This rules out the use of any non-trivial history dependent strategy for this game. Therefore, the players only have Markov strategies at their disposal. In Flesch et al. [1997] it is shown that, although (cyclic) Markov limiting average 0-equilibria exist for this game, there are no stationary limiting average ε-equilibria. Moreover, the set of all limiting average equilibria is being characterized completely. An example of a Markov equilibrium for this game is (π, σ, τ), where π is defined by: at stages 1, 4, 7, 10, ... play T with probability $\frac{1}{2}$ and at all other stages play T with probability 1. Similarly, σ is defined by: at stages 2, 5, 8, 11, ... play L with probability $\frac{1}{2}$ and at all other stages play L with probability 1. Likewise, τ is defined by: at stages 3, 6, 9, 12, ... play N with probability $\frac{1}{2}$ and at all other stages play N with probability 1. The limiting average reward corresponding to this equilibrium is $(1, 2, 1)$.

References

Bewley T, Kohlberg E [1976] The asymptotic theory of stochastic games. *Math Oper Res* **1**: 197–208.

Blackwell D [1962] Discrete dynamic programming. *Ann Math Statist* **33**: 719–726.

Blackwell D, Ferguson TS [1968] The big match. *Ann Math Statist* **39**: 159–163.

Evangelista FS, Raghavan TES, Vrieze OJ [1997]: ARAT repeated games with absorbing states. In: Volume in Honor of D Blackwell (forthcoming).

Everett H [1957] Recursive games. In: M. Dresher, A.W. Tucker & P. Wolfe (eds), *Contributions to the Theory of Games* **III**, *Annals of Mathematical Studies* **39**, Princeton University Press, Princeton, 47–78.

Federgruen A [1978] On N-person stochastic games with denumerable state space. *Adv Appl Prob* **10**: 452–471.

Filar JA [1981] Ordered field property for stochastic games when the player who controls transitions changes from state to state. *J Opt Theory Appl* **34**: 503–515.

Filar JA, Schultz TA, Thuijsman F, Vrieze OJ [1991] Non-linear programming and stationary equilibria in stochastic games. *Math Progr* **50**: 227–237.

Fink AM [1964] Equilibrium in a stochastic n-person game. *J Sci Hiroshima Univ*, Series A-I **28**, 89–93.

Flesch J, Thuijsman F, Vrieze OJ [1996a] Recursive repeated games with absorbing states. *Math Oper Res* **21**: 1016–1022.

Flesch J, Thuijsman F, Vrieze OJ [1996b] Simplifying optimal strategies in stochastic games. Report M96-11, Department of Mathematics, Maastricht University.

Flesch J, Thuijsman F, Vrieze OJ [1997] Cyclic Markov equilibria in a cubic game. *Int J Game Theory* (forthcoming).

Gillette D [1957] Stochastic games with zero stop probabilities. In: Dresher M, Tucker AW, Wolfe P (eds), *Contributions to the Theory of Games* **III**, *Annals of Mathematical Studies* **39**, Princeton University Press, Princeton, pp 179–187.

Hordijk A, Vrieze OJ, Wanrooij GL [1983] Semi-Markov strategies in stochastic games. *Int J Game Theory* **12**: 81–89.

Kohlberg E [1974] Repeated games with absorbing states. *Annals of Statistics* **2**: 724–738.

Liggett TM, Lippman SA [1969] Stochastic games with perfect information and time average payoff. *SIAM Review* **11**: 604–607.

Mertens JF, Neyman A [1981] Stochastic games. *Int J Game Theory* **10**: 53–66.

Nash J [1951] Non-cooperative games. *Annals of Mathematics* **54**: 286–295.

Parthasarathy T, Raghavan TES [1981] An orderfield property for stochastic games when one player controls transition probabilities. *J Opt Theory Appl* **33**: 375–392.

Parthasarathy T, Tijs SH, Vrieze OJ [1984] Stochastic games with state independent transitions and separable rewards. In: Hammer G, Pallaschke D (eds), *Selected Topics in Operations Research and Mathematical Economics*, Springer Verlag, Berlin, pp 262–271.

Raghavan TES, Tijs SH, Vrieze OJ [1985] On stochastic games with additive reward and transition structure. *J Opt Theory Appl* **47**: 451–464.

Rogers PD [1969] Non-zerosum stochastic games. PhD thesis, report ORC 69-8, Operations Research Center, University of California, Berkeley.

Shapley LS [1953] Stochastic games *Proc Nat Acad Sci USA* **39**: 1095–1100.

Sobel MJ [1971] Noncooperative stochastic games. *Ann Math Statist* **42**: 1930–1935.

Sorin S [1986] Asymptotic properties of a non-zerosum stochastic game. *Int J Game Theory* **15**: 101–107.

Thuijsman F [1992] *Optimality and Equilibria in Stochastic Games*. CWI-tract **82**, Center for Mathematics and Computer Science, Amsterdam.

Thuijsman F, Raghavan TES [1997] Perfect information stochastic games and related classes. *Int J Game Theory* (forthcoming).

Thuijsman F, Vrieze OJ [1991] Easy initial states in stochastic games. In: Raghavan TES et al (eds), *Stochastic Games and Related Topics*, Kluwer Academic Publishers, Dordrecht, pp 85–100.

Thuijsman F, Vrieze OJ [1992] Note on recursive games. In: Dutta B et al (eds), *Game Theory and Economic Applications, Lecture Notes in Economics and Mathematical Systems* **389**, Springer, Berlin, pp 133–145.

Thuijsman F, Vrieze OJ [1993] Stationary ε-optimal strategies in stochastic games. *OR Spektrum* 15, 9–15.

Thuijsman F, Vrieze OJ [1997] The power of threats in stochastic games. In: Bardi et al (eds), *Stochastic Games and Numerical Methods for Dynamic Games*, Birkhauser, Boston (forthcoming).

Tijs SH, Vrieze OJ [1986] On the existence of easy initial states for undiscounted stochastic games. *Math Oper Res* **11**: 506–513.

Vieille N [1993] Solvable states in stochastic games. *Int J Game Theory* **21**: 395–404.

Von Neumann J [1928] Zur Theorie der Gesellschafsspiele. *Mathematische Annalen* **100**: 295–320.

Vrieze OJ [1987] *Stochastic Games with Finite State and Action Spaces*. CWI-tract **33**, Center for Mathematics and Computer Science, Amsterdam.

Vrieze OJ, Thuijsman F [1987] Stochastic games and optimal stationary strategies, a survey. In: Domschke W et al (eds), *Methods of Operations Research* **57**:513–529.

Vrieze OJ, Thuijsman F [1989] On equilibria in repeated games with absorbing states. *Int J Game Theory* **18**: 293–310.

Vrieze OJ, Tijs SH, Raghavan TES, Filar JA [1983] A finite algorithm for the switching control stochastic game. *OR Spektrum* **5**: 15–24.