Statistics, Probability and Game Theory
IMS Lecture Notes – Monograph Series (1996) Volume 30

REPEATED ARAT GAMES

F. EVANGELISTA T. E. S. RAGHAVAN, O. J. VRIEZE
University of Wisconsin, University of Illinois, University of Limburg

Abstract

We show the existence of ε-equilibria in stationary strategies for the class of repeated games with Additive Reward and Additive Transition (ARAT) structure. A new approach to existence questions for stochastic games is introduced in the proof — the strategy space of one player is perturbed so that the player uses any pure action with at least probability ε. For this ε-perturbed game, equilibria in stationary strategies exist. By analyzing the limit properties of these strategies, the existence of stationary ε-equilibria in the original game follows.

1. Introduction. When Shapley [1953] first defined stochastic games and proved the existence of value and stationary optimal strategies, he essentially gave a complete result for zero-sum games with discounted payoffs. Independently, Blackwell [1962, 1965] initiated the systematic study of Markovian decision processes. He investigated the nature of optimal policies for discounted payoffs and the relation between the discounted and the limiting average (undiscounted) case. Together with Ferguson (Blackwell and Ferguson [1968]), he analyzed the stochastic game called "The Big Match," revealing an important difference between Markovian decision processes and stochastic games. This zero-sum game with limiting average payoff has a value. But unlike Markovian decision processes where the player has a stationary optimal strategy, only a behavioral ε-optimal strategy exists for the maximizer. In the Big Match, while one player has no power to terminate the game, the opponent can terminate the game any time by choosing the absorbing row. It is this freedom to terminate the game that complicates the player's near optimal strategy. Generalizing the Big Match as a single loop stochastic game which can be terminated by one player, Filar extended the game and showed that such games admit once again, an epsilon optimal but only a behavioral strategy for the controlling player [Filar (1981)].

Attempts by Kohlberg [1974], and Bewley and Kohlberg [1976], to extend this result to general undiscounted payoffs led to them to study the Puiseux expansion of the value function. Using ideas from Blackwell and Ferguson, and Bewley and Kohlberg, Mertens and Neyman [1981] proved the existence of value for all limiting average zero-sum stochastic games.

Fink [1964] extended Shapley's result to nonzero-sum games, showing the existence of equilibria in stationary strategies for games with discounted payoffs. However, there is as yet no result analogous to that of Mertens and Neyman: the existence of ε-equilibria in nonzero-sum games with limiting average payoff is still an open problem.

Another direction in the study of stochastic games with limiting average payoff is to identify subclasses of games that have ε-equilibria in stationary strategies by imposing conditions on the structure of the payoffs and/or transition probabilities. The following examples from the literature are games with ε-equilibria in stationary strategies for both the zero-sum and nonzero-sum cases: irreducible/unichain stochastic games (Rogers [1969], Stern [1978], Thuijsmann and Vrieze [1991]); single controller stochastic games (Parthasarathy and Raghavan [1981]); stochastic games with state independent transitions and separable rewards (Parthasarathy, Tijs and Vrieze [1984]). Filar and Raghavan [1991] provides a detailed survey of these games.

In this paper we add another class to the list of nonzero-sum games with ε-equilibria in stationary strategies. We show that repeated games with absorbing states for which the additive reward and additive transition property (ARAT) holds have ε-equilibria, and sometimes even equilibria ($\varepsilon = 0$) in stationary strategies.

"The Big Match" is an example of a repeated game with absorbing states. It has only one non-absorbing state; whenever the game exits from this state to one of the other states, the game is essentially over since it will remain forever in one of the absorbing states with the same payoffs at every decision moment. Using the idea of threat strategies, Vrieze and Thuijsman [1989] showed that repeated games with absorbing states have ε-equilibria, generally in non-stationary strategies.

In a stochastic game with an ARAT structure, the rewards at each decision moment can be written as a sum of a term dependent on player 1's action and the state variable and a term dependent on player 2's action and the state variable. Moreover, the same holds for the transitions. Zero-sum ARAT games have optimal stationary strategies (Raghavan, Tijs and Vrieze [1986]) but this is not true in the nonzero-sum case (Evangelista [1993]).

We end this introduction with the remark that Thuijsman and Raghavan [1994] showed that both switching control and ARAT stochastic games possess ε-equilibria which are stationary-like; that is, a retaliating player may need to use non-stationary strategies only in the case of a retaliation as a consequence of a deviation of the other player. Recently, Flesch [1995] gave an example of a perfect information game that does not have ε-equilibria in stationary strategies.

2. Preliminaries. In this section we will give some properties of repeated ARAT games which will be used in the proof of the main result of the paper. First we need some notations and definitions.

A stochastic game is identified by a six-tuple $\{S, \{A^1(s); s \in S\}, \{A^2(s); s \in S\}, r^1, r^2, p\}$. Here, $S = \{1, \dots, K\}$ is the state space of the game; $A^k(s), k \in \{1,2\}, s \in S$, is the action set of player k in state s; r^k denotes the payoff to player k, defined on the set of triples (s,i,j) with $i \in A^1(s)$ and $j \in A^2(s)$. When the players choose i and j in state s, the payoff to player k will be $r^k(s,i,j)$, and the probability that the next state will be s' will be $p(s'|s,i,j)$ with $\sum_{s'} p(s'|s,i,j) = 1$.

For a repeated game with absorbing states, the non-absorbing state will be called state 1 and the absorbing states will be states $2, 3, \dots, K$. Strategies for the game are completely defined by the choices the players make in state 1. The dimension of this state is taken to be $M \times N$, i.e. player I has M pure actions and player II has N pure actions. A stationary strategy for player I will be denoted by $x \in \Delta_M := \{x \in \Re^M \ : \ x(i) \geq 0 \ , \sum_{i=1}^{M} x(i) = 1\}$. A stationary strategy for player II will be denoted by $y \in \Delta_N := \{y \in \Re^N \ : \ y(j) \geq 0 \ , \sum_{j=1}^{N} y(j) = 1\}$. When a player uses a stationary strategy, he chooses an action according to the same randomized selection at every decision moment; that is, when player I uses $x \in \Delta_M$ then x_i equals the probability that action i will be chosen, $i = 1, 2, \dots, M$; likewise for $y \in \Delta_N$. A pure strategy in which action i is chosen with probability 1, will be denoted by e_i.

An ARAT game is defined by the structural properties:

$$r_s^k(i,j) = a_s^k(i) + b_s^k(j) \ , \ k = 1,2 \ \text{ and}$$

$$p(s|i,j) = p(s,i) + \ p(s,j), \ \ s = 1, 2, \dots, K \ \text{ and all } i, j.$$

Since both players have just one action in the absorbing states, we will drop the parameters i and j in the payoffs when $s = 2, 3, \dots K$. For $s = 1$, the subscript "1" will be dropped. We will suppose that the actions of the players are arranged such that for some integers $\tilde{M}$ and $\tilde{N}$,

$$\begin{aligned}
&p(s,i) = 0 \ , \quad \forall s \geq 2 \quad , \quad i \in \{1, 2, \dots, \tilde{M}\} \\
&\sum_{s=2}^{K} p(s,i) > 0 \ , \quad i \in \{\tilde{M}+1, \dots, M\} \\
&p(s,j) = 0 \ , \quad \forall s \geq 2 \ , \quad j \in \{1, 2, \dots, \tilde{N}\} \\
&\sum_{s=2}^{K} p(s,j) > 0 \ , \quad j \in \{\tilde{N}+1, \dots, N\}.
\end{aligned}$$

The actions $\tilde{M}+1, \dots, M$ and $\tilde{N}+1, \dots, N$ will be called absorbing actions. A randomized strategy x that puts positive weight on some absorbing action will be called absorbing. Note that when (x, y) is a strategy pair for which either x or y is absorbing, then the game will eventually reach one of the absorbing states. We will also use the term "absorbing" to denote a pair of strategies of this type. For a pair of absorbing strategies (x,y) we denote by $r^k(x,y)$ the one-step expected payoff to player k, given that absorption occurs. The one-step absorption probability is denoted by $p(x,y)$. Hence with probability $1-p(x,y)$ the play will remain at state 1 and go on.

We define $C(x) := \{i : x(i) > 0\}$ and $C(y) := \{j : y(j) > 0\}$. If (x,y) is a strategy pair for which $C(x) \subset \{1,2,\ldots,\tilde{M}\}$ and $C(y) \subset \{1,2,\cdots,\tilde{N}\}$, then the game will forever remain in state 1. We will call (x,y) non-absorbing. At every decision moment the one-step expected payoff for player k equals $a^k(x)+b^k(y) := \sum_{i=1}^{\tilde{M}} a^k(i)x(i) + \sum_{j=1}^{\tilde{N}} b^k(j)y(j)$.

In general, the limiting average expected payoff to player k when the game starts at state s and when strategies (Π,Γ) are used by players I and II with an immediate reward $I^k(s_\tau,i_\tau,j_\tau)$ to player k on the τ-th day will be denoted by

$$\Phi^k(\Pi,\Gamma)(s) = \liminf_{N\to\infty} \frac{1}{N} \mathrm{E}_{\,\Pi,\Gamma}[\sum_{\tau=1}^{N} I_\tau^k(s_\tau,i_\tau,j_\tau)]$$

In our model, state 1 is the only relevant state and from now on Φ^k will represent the payoff starting at state 1.

Lemma 2.1

(i) If stationary (x,y) is absorbing, then

$$\Phi^k(x,y) = \frac{r^k(x,y)}{p(x,y)}$$

$$= \frac{\sum_{i=\tilde{M}+1}^{M}\sum_{s=2}^{K} p(s,i)x(i)a_s^k + \sum_{j=\tilde{N}+1}^{N}\sum_{s=2}^{K} p(s,j)y(j)b_s^k}{\sum_{i=\tilde{M}+1}^{M}\sum_{s=2}^{K} p(s,i)x(i) + \sum_{j=\tilde{N}+1}^{N}\sum_{s=2}^{K} p(s,j)y(j)} \tag{2.1}$$

(ii) If (x,y) is non-absorbing, then

$$\Phi^k(x,y) = a^k(x) + b^k(y). \tag{2.2}$$

Proof: After the game moves to an absorbing state, the expected payoff equals $r^k(x,y)$ at every decision moment. Thus the expected one-step absorption payoffs for every decision moment count fully in the limiting average payoff. For decision moment t, the contribution equals $(1-p(x,y))^t r^k(x,y)$. Hence, when the pair (x,y) is absorbing, then

$$\Phi^k(x,y) = \sum_{t=0}^{\infty}(1-p(x,y))^t r^k(x,y) = \frac{r^k(x,y)}{p(x,y)}.$$

By the definition of the ARAT structure, the righthand side of (2.1) follows.

When (x,y) is non-absorbing, then player k's expected payoff equals $a^k(x) + b^k(y)$ at every decision moment. □

Corollary 2.2 *If x is absorbing and y is non-absorbing, then the limiting average payoff is independent of y:*

$$\Phi^k(x,y) = \frac{\sum_{i=\widetilde{M}+1}^{M} \sum_{s=2}^{K} p(s,i)x(i)a_s^k}{\sum_{i=\widetilde{M}+1}^{M} \sum_{s=2}^{K} p(s,i)x(i)}, \quad k = 1,2 \; .$$

Similarly, if y is absorbing and x is non-absorbing, then the payoff is independent of x.

To construct the perturbed game, we first let

$$Y_n \; := \{y \in \Delta_N \; ; \; \sum_{j=\widetilde{N}+1}^{N} y(j) \; \geq \frac{1}{n}\}.$$

Observe that

(i) each $y \in Y_n$ is absorbing

(ii) $e_j \in Y_n$ for $j \in \{\tilde{N}+1, \ldots, N\}$.

Define the perturbed game, Γ_n, as the repeated game with payoffs Φ^1 and Φ^2 and strategy spaces Δ_M and Y_n.

Lemma 2.3 *For every n, the game Γ_n has at least one equilibrium.*

Proof : By Corollary 2.2, an equilibrium for the repeated game Γ_n is a pure ε-equilibrium in the (one-step) two-person nonzero-sum game with payoffs Φ^1 and Φ^2 and strategy spaces Δ_M and Δ_N. Since each stationary strategy pair (x,y) is absorbing, the payoff functions Φ^1 and Φ^2 are continuous on the compact, convex set $\Delta_M \times Y_n$. Furthermore, since all a_s^k and b_s^k may be supposed to be positive it can be shown (see Evangelista [1993]) that Φ^1 and Φ^2 are quasi-concave on $\Delta_M \times Y_n$, i.e. for any real number α, the sets $\{(x,y)| \; \Phi^k(x,y) > \alpha\}$ are convex. By a theorem of Glicksberg [1952] or Fan [1952], the nonzero-sum game $[\Phi^1, \Phi^2, \Delta_M, Y_n]$ has a pure Nash equilibrium which in turn is an equilibrium of the repeated game Γ_n. □

The proof of the main theorem will be based on a sequence of equilibria (x_n, y_n) of the game $\Gamma_n, n = 1, 2, \cdots$.

The following definition is essential for our approach. Let

$$X_{abs} \; := \; \{x \in \Delta_M : \; \sum_{i=1}^{\widetilde{M}} x(i) < 1\}.$$

Thus, X_{abs} consists of all the absorbing strategies of player I. Define for $j \in \{1, 2, \ldots, N\}$,

$$E_j \ := \ \{x \in X_{abs} \ ; \ \max_{y \in \Delta_N} \Phi^2(x, y) \ = \ \Phi^2(x, e_j)\}$$

Hence, E_j contains those $x \in X_{abs}$, for which pure strategy e_j of player II is a best reply. By Corollary 2.2 it follows that $E_{j_1} = E_{j_2}$ for all $j_1, j_2 \in \{1, 2, \ldots, \tilde{N}\}$, and so

$$\Delta_M \ = \ \Delta_{\widetilde{M}} \ \cup E_1 \cup \ (\cup_{j=\widetilde{N}+1}^{N} E_j). \tag{2.3}$$

3. The Existence Proof. In accordance with (2.3) we distinguish three cases which cover all possibilities. If in the sequel we need an accumulation point of a bounded sequence, we suppose that an appropriate subsequence is chosen.

Case 1. There exists $x_n \in \cup_{j=\widetilde{N}+1}^{N} E_j$ for some n.
Case 2. Case 1 does not hold and $x_n \in E_1$, all n.
Case 3. Cases 1 and 2 do not hold and $x_n \in \Delta_{\widetilde{M}}$, for all $n \in I\!N$.

For each of these cases we will prove the existence of an ε-equilibrium.

Case 1.

Theorem 3.1 *If $x_n \in E_{\tilde{j}}$ for some $\tilde{j} \in \{\tilde{N}+1, \ldots, N\}$, then (x_n, y_n) is an equilibrium in the original game Γ.*

Proof : Since (x_n, y_n) is an equilibrium of Γ_n, we have

$$\Phi^1(x_n, y_n) \geq \Phi^1(x, y_n), \ \forall x \in \Delta_M \tag{3.1}$$

Since $e_{\tilde{j}} \in Y_n$ and $Y_n \subset \Delta_N$,

$$\Phi^2(x_n, y_n) = \max_{y \in Y_n} \Phi^2(x_n, y) \geq \Phi^2(x_n, e_{\tilde{j}}).$$

Also, $x_n \in E_{\tilde{j}}$ implies that

$$\Phi^2(x_n, e_{\tilde{j}}) = \max_{y \in \Delta_n} \Phi^2(x_n, y) \geq \max_{y \in Y_n} \Phi^2(x_n, y),$$

and so

$$\Phi^2(x_n, y_n) = \max_{y \in \Delta_N} \Phi^2(x_n, y) \tag{3.2}$$

The combination of (3.1) and (3.2) proves the theorem. □

Case 2. We start with a lemma that states that the total weight that y_n puts outside $\{1, 2, \ldots, \tilde{N}\}$ is minimal, namely $\frac{1}{n}$. We use the following abbrevations:

$$\begin{aligned}
A^k(x) :&= \sum_{i=\widetilde{M}+1}^{M} \sum_{s=2}^{K} p(s,i)x(i)a_s^k \\
a(x) :&= \sum_{i=\widetilde{M}+1}^{M} \sum_{s=2}^{K} p(s,i)x(i) \\
B^k(y) :&= \sum_{j=\widetilde{N}+1}^{N} \sum_{s=2}^{K} p(s,j)y(j)b_s^k \\
b(y) :&= \sum_{j=\widetilde{N}+1}^{N} \sum_{s=2}^{K} p(s,j)y(j)
\end{aligned}$$

Lemma 3.2 *If $x_n \in E_1$ and $x_n \notin \cup_{j=\widetilde{N}+1}^{N} E_j$, then $\sum_{j=\widetilde{N}+1}^{N} y_n(j) = \frac{1}{n}$.*

Proof: If $x_n \in E_1$ and $x_n \notin E_j$ for each $j \in \{\tilde{N}+1, \ldots, N\}$, then for every $y \in \Delta_N$,

$$\frac{A^2(x_n)}{a(x_n)} = \Phi^2(x_n, e_1) > \Phi^2(x_n, y) = \frac{A^2(x_n) + B^2(y)}{a(x_n) + b(y)},$$

which gives

$$\frac{A^2(x_n)}{a(x_n)} > \frac{B^2(y}{b(y)}, \ \forall y \in \Delta_N. \tag{3.3}$$

Suppose that the total weight that y_n puts on $\{\tilde{N}+1, \ldots, N\}$ is strictly more than $\frac{1}{n}$. Let $\tilde{j} \in C(y_n) \bigcap \{\tilde{N}+1, \ldots, N\}$ be such that

$$\frac{B^2(e_{\tilde{j}})}{b(e_{\tilde{j}})} = \min_{j \in C(y_n) \cap \{\widetilde{N}+1, \ldots, N\}} \frac{B^2(e_j)}{b(e_j)}.$$

Consider the strategy $\widetilde{y}_n$ defined as:

$$\widetilde{y}_n(1) = y_n(1) + \delta$$

$$\widetilde{y}_n(\tilde{j}) = y_n(\tilde{j}) - \delta$$

$$\widetilde{y}_n(j) = y_n(j), \ j \neq 1, \ j \neq \tilde{j}$$

where $\delta > 0$ is small enough so that

$$\sum_{j=\tilde{N}+1}^{N} \tilde{y}_n(j) \geq \frac{1}{n}.$$

It follows from (3.3) and the definition of $\tilde{j}$ that:

$$\frac{A^2(x_n) + B^2(y_n)}{a(x_n) + b(y_n)} > \frac{B^2(y_n)}{b(y_n)} \geq \frac{B^2(e_{\tilde{j}})}{b(e_{\tilde{j}})}.$$

This implies that

$$\begin{aligned}\Phi^2(x_n, \tilde{y}_n) &= \frac{A^2(x_n) \ + \ B^2(y_n) - \delta B^2(e_{\tilde{j}})}{a(x_n) + b(y_n) - \delta b(e_{\tilde{j}})} \\ &> \frac{A^2(x_n) + B^2(y_n)}{a(x_n) + b(y_n)} \ = \ \Phi^2(x_n, y_n)\end{aligned}$$

which contradicts the equilibrium point assumption of (x_n, y_n). Hence the assumption $\sum_{j=\tilde{N}+1}^{N} y_n(j) > \dfrac{1}{n}$ is incorrect. □

Corollary 3.3 *In case 2, we have $C(y_0) \subset \{1, \ldots, \tilde{N}\}$ where $y_0 = \lim_{n\to\infty} y_n$.*

By the assumption of Case 2, $x_n \in E_1$, so $x_n \in X_{abs}$. Define $\tilde{x}_n$ as:

$$\begin{aligned}\tilde{x}_n(i) &= 0, \ i = 1, 2, \ldots, \tilde{M} \\ \tilde{x}_n(i) &= x_n(i)/X_n, \ i = \tilde{M}+1, \ldots, M \\ \text{with } X_n &= \sum_{i=\tilde{M}+1}^{M} x_n(i)\end{aligned}$$

So $\tilde{x}_n(i)$ puts weight only on absorbing actions.

Theorem 3.4 *For $\varepsilon > 0$, the pair $(\tilde{x}_n, y_n)$ is an ε-equilibrium point in the original game Γ, for n large enough.*

Proof. Part I: In the first part of the proof we will show that $\Phi^1(\tilde{x}_n, y_n) = \Phi^1(x_n, y_n)$. This will prove that for player I, $\tilde{x}_n$ is a best reply to y_n. First, observe that

$$\frac{A^k(\tilde{x}_n)}{a(\tilde{x}_n)} = \frac{(1/X_n) \ A^k(x_n)}{(1/X_n) \ a(x_n)} = \frac{A^k(x_n)}{a(x_n)} \tag{3.4}$$

Next, we show that

$$\frac{A^1(x_n)}{a(x_n)} = \frac{B^1(y_n)}{b(y_n)}$$

a) Since (x_n, y_n) is an equilibrium point, we have for $i \in \{1, \ldots, \tilde{M}\}$:

$$\frac{A^1(x_n) + B^1(y_n)}{a(x_n) + b(y_n)} = \Phi^1(x_n, y_n) \geq \Phi^1(e_i, y_n) = \frac{B^1(y_n)}{b(y_n)}.$$

Hence

$$\frac{A^1(x_n)}{a(x_n)} \geq \frac{B^1(y_n)}{b(y_n)} \tag{3.5}$$

b) Also, $\Phi^1(x_n, y_n) \geq \Phi^1(\tilde{x}_n, y_n)$ or

$$\begin{aligned} \frac{A^1(x_n) + B^1(y_n)}{a(x_n) + b(y_n)} &\geq \frac{(1/X_n)\ A^1(x_n) + B^1(y_n)}{(1/X_n)\ a(x_n) + b(y_n)} \\ &= \frac{A^1(x_n) + B^1(y_n) + (1/X_n - 1)A^1(x_n)}{a(x_n) + b(y_n) + (1/X_n - 1)\ a(x_n)}. \end{aligned}$$

Hence

$$\frac{A^1(x_n) + B^1(y_n)}{a(x_n) + b(y_n)} \geq \frac{(1/X_n - 1)\ A^1(x_n)}{(1/X_n - 1)\ a(x_n)} = \frac{A^1(x_n)}{a(x_n)}.$$

And so

$$\frac{B^1(y_n)}{b(y_n)} \geq \frac{A^1(x_n)}{a(x_n)}. \tag{3.6}$$

Combining (3.5) and (3.6) gives

$$\frac{A^1(x_n)}{a(x_n)} = \frac{B^1(y_n)}{b(y_n)}.$$

But then

$$\frac{A^1(x_n) + B^1(y_n)}{a(x_n) + b(y_n)} = \frac{(1/X_n)\ A^1(x_n) + B^1(y_n)}{(1/X_n)\ a(x_n) + b(y_n)}$$

and so $\Phi^1(\tilde{x}_n, y_n) = \Phi^1(x_n, y_n) \geq \Phi^1(x, y_n)$, $\forall x \in \Delta_M$.

Part II: In the second part of the proof we will show that for player II, y_n is an ε-best reply to $\tilde{x}_n$, for n large enough.

Since $x_n \in E_1$,

$$\frac{A^2(x_n)}{a(x_n)} \geq \frac{A^2(x_n) + B^2(y_n)}{a(x_n) + b(y_n)}. \tag{3.7}$$

Because (x_n, y_n) is an equilibrium point of the game Γ_n,

$$\frac{A^2(x_n) + B^2(y_n)}{a(x_n) + b(y_n)} \geq \frac{A^2(x_n) + B^2(y)}{a(x_n) + b(y)}\ ,\ \forall y \in Y_n. \tag{3.8}$$

By (3.7) and (3.8),

$$\frac{A^2(x_n)}{a(x_n)} \geq \frac{A^2(x_n)+B^2(y)}{a(x_n)+b(y)}, \ \forall y \in Y_n.$$

Hence

$$\frac{A^2(x_n)}{a(x_n)} \geq \frac{B^2(y)}{b(y)}\ , \ \forall y \in Y_n.$$

And also

$$\frac{A^2(\tilde{x}_n)}{a(\tilde{x}_n)} \ = \ \frac{(1/X_n)A^2(x_n)}{(1/X_n)a(x_n)} \geq \frac{B^2(y)}{b(y)}\ , \ \forall y \in Y_n.$$

It follows that

$$\Phi^2(\tilde{x}_n, y) \ = \ \frac{A^2(\tilde{x}_n)+B^2(y)}{a(\tilde{x}_n)+b(y)} \leq \frac{A^2(\tilde{x}_n)}{a(\tilde{x}_n)}\ , \forall y \in Y_n. \tag{3.9}$$

Observe that, since $\tilde{x}_n$ puts all its weight on the absorbing actions, while y_n puts only weight $1/n$ on the absorbing actions we have for n large enough:

$$\frac{A^2(\tilde{x}_n)}{a(\tilde{x}_n)} \leq \frac{A^2(\tilde{x}_n)+B^2(y_n)}{a(\tilde{x}_n)+b(y_n)} + \varepsilon \ = \Phi^2(\tilde{x}_n, y_n) + \varepsilon \tag{3.10}$$

Further, for $j = 1, \ldots, \tilde{N}$:

$$\Phi^2(\tilde{x}_n, e_j) \ = \ \Phi^2(x_n, e_j) \ = \ \frac{A^2(x_n)}{a(x_n)} \ = \ \frac{A^2(\tilde{x}_n)}{a(\tilde{x}_n)} \tag{3.11}$$

By (3.9), (3.10) and (3.11), $\Phi^2(\tilde{x}_n, y_n) \geq \Phi^2(\tilde{x}_n, y) - \varepsilon$, $\forall y \in \Delta_N$. Parts I and II prove that the pair $(\tilde{x}_n, y_n)$ is an ε-equilibrium point. □

Case 3. Recall that this is the case where cases 1 and 2 do not hold and x_n is non-absorbing, i.e. $x_n \in \Delta_{\tilde{M}}$.

Let $\tilde{j} \in \{\tilde{N}+1, \ldots, N\}$ be such that:

1. $$\frac{B^2(e_{\tilde{j}})}{b(e_{\tilde{j}})} \geq \frac{B^2(e_j)}{b(e_j)}\ , \ \forall j \in \{\tilde{N}+1, \ldots, N\} \tag{3.12}$$

2. $B^2(e_{\tilde{j}}) \geq B^2(e_j)$

$$\text{whenever } \frac{B^2(e_{\tilde{j}})}{b(e_{\tilde{j}})} \ = \ \frac{B^2(e_j)}{b(e_j)}\ , \ \forall j \in \{\tilde{N}+1, \ldots, N\} \tag{3.13}$$

Case 3 itself is separated into 2 cases. We distinguish between the cases $E_j \neq \emptyset$ for some $j \in \{\tilde{N}+1, \ldots, N\}$ and $E_j = \emptyset$ for all $j \in \{\tilde{N}+1, \ldots, N\}$.

Case 3A: $E_j \neq \emptyset$ for some $j \in \{\tilde{N}+1, \ldots, N\}$

Theorem 3.5 *Let $x \in E_{\bar{j}}$ for some $\bar{j} \in \{\tilde{N}+1,\dots,N\}$. Then, for λ small enough, the pair $((1-\lambda)x_n + \lambda x, y_n)$ is an ε-equilibrium point.*

Proof: Recall that $E_{\bar{j}} \subset X_{abs}$, so $x \in X_{abs}$. By the equilibrium property of (x_n, y_n):

$$\frac{B^2(y_n)}{b(y_n)} = \Phi^2(x_n, y_n) \geq \Phi^2(x_n, e_{\bar{j}}) = \frac{B^2(e_{\bar{j}})}{b(e_{\bar{j}})}.$$

On the other hand it follows from (3.12) that

$$\frac{B^2(y_n)}{b(y_n)} \leq \frac{B^2(e_{\bar{j}})}{b(e_{\bar{j}})}$$

and so

$$\frac{B^2(y_n)}{b(y_n)} = \frac{B^2(e_{\bar{j}})}{b(e_{\bar{j}})} \tag{3.14}$$

Now for fixed $\epsilon > 0$, fixed n, we can choose $\lambda > 0$ small enough such that the inequalities

$$\begin{aligned} \Phi^2((1-\lambda)x_n + \lambda x, y_n) &= \frac{B^2(y_n) + \lambda A^2(x)}{b(y_n) + \lambda a(x)} \\ &\geq \frac{B^2(y_n)}{b(y_n)} - \varepsilon/2 \end{aligned} \tag{3.15}$$

and for each $j \in \{\tilde{N}+1,\dots,N\}$:

$$\begin{aligned} \Phi^2((1-\lambda)x_n + \lambda x, e_j) &= \frac{B^2(e_j) + \lambda A^2(x)}{b(e_j) + \lambda a(x)} \\ &\leq \frac{B^2(e_j)}{b(e_j)} + \varepsilon/2 \end{aligned} \tag{3.16}$$

are both satisfied. Combining (3.14), (3.15) and (3.16) yields:

$$\Phi^2((1-\lambda)x_n + \lambda x, e_j) \leq \Phi^2((1-\lambda)x_n + \lambda x, y_n) + \varepsilon \tag{3.17}$$

for each $j \in \{\tilde{N}+1,\dots,N\}$. Then $\Phi^2(x, e_{\bar{j}}) \geq \Phi^2(x, e_1)$, or:

$$\frac{A^2(x) + B^2(e_{\bar{j}})}{a(x) + b(e_{\bar{j}})} \geq \frac{A^2(x)}{a(x)}$$

Hence

$$\frac{B^2(e_{\bar{j}})}{b(e_{\bar{j}})} \geq \frac{A^2(x)}{a(x)}$$

By definition of $e_{\tilde{j}}$, we obtain

$$\frac{B^2(e_{\tilde{j}})}{b(e_{\tilde{j}})} \geq \frac{A^2(x)}{a(x)},$$

which in combination with (3.14) leads to

$$\frac{B^2(y_n)}{b(y_n)} \geq \frac{A^2(x)}{a(x)}.$$

Hence, by (3.15) we have for each $j \in \{1, \ldots, \tilde{N}\}$:

$$\Phi^2((1-\lambda)x_n + \lambda x, e_j) = \frac{A^2(x)}{a(x)} \leq \Phi^2((1-\lambda)x_n + \lambda x, y_n) + \varepsilon \quad (3.18)$$

The inequalities (3.17) and (3.18) show that y_n is an ε-best answer to $(1-\lambda)x_n + \lambda x$.

For player I things are easier. Since (x_n, y_n) is absorbing and $((1-\lambda)x_n + \lambda x, y_n)$ is absorbing for every $\lambda > 0$ it follows that:

$$\lim_{\lambda \downarrow 0} \Phi^1((1-\lambda)x_n + \lambda x, y_n) = \Phi^1(x_n, y_n)$$

Since x_n is a best reply to y_n, it is obvious that, for $\lambda > 0$ small enough, $(1-\lambda)x_n + \lambda x$ is an ε-best reply to y_n. This proves theorem 3.5. □

Case 3B: $E_j = \emptyset$, all $j \in \{\tilde{N}+1, \ldots, N\}$.

Lemma 3.6 *If $E_j = \emptyset$, all $j \in \{\tilde{N}+1, \ldots, N\}$, then $E_j = X_{abs}$ for all $j \in \{1, \ldots, \tilde{N}\}$.*

Proof: The lemma is trivially implied by its condition, since each $x \in X_{abs}$ belongs to at least one E_j, $j \in \{1, \ldots, N\}$, and the E_j's for $j \in \{1, \ldots, \tilde{N}\}$ are identical. □

Theorem 3.7 *Let $x_n \in \Delta_{\tilde{M}}$ for all n and suppose that $E_j = \emptyset$, all $j \in \{\tilde{N}+1, \ldots, N\}$. Then there exists an ε-equilibrium point for the original game Γ.*

Proof : First suppose that

$$\max_{i \in \{\tilde{M}+1, \ldots, M\}} \frac{A^1(e_i)}{a(e_i)} \geq \min_{j \in \{\tilde{N}+1, \ldots, N\}} \frac{B^1(e_j)}{b(e_j)} \quad (3.19)$$

Let $\tilde{j}$ be such that:

$$\frac{B^1(e_{\tilde{j}})}{b(e_{\tilde{j}})} = \min_{j \in \{\tilde{N}+1, \ldots, N\}} \frac{B^1(e_j)}{b(e_j)} \quad (3.20)$$

Let $\tilde{\imath}$ be such that :

$$\frac{A^1(e_{\tilde{\imath}})}{a(e_{\tilde{\imath}})} = \max_{i \in \{\tilde{M}+1,\ldots,M\}} \frac{A^1(e_i)}{a(e_i)} \tag{3.21}$$

Let $y_0 \in \Delta_{\tilde{N}}$ be arbitrary. Consider the pair $(e_{\tilde{\imath}}, (1-\lambda)y_0 + \lambda e_{\tilde{\jmath}})$ for $\lambda > 0$ small enough. Since

$$\lim_{\lambda \downarrow 0} \Phi^2(e_{\tilde{\imath}}, (1-\lambda)y + \lambda e_{\tilde{\jmath}}) = \Phi^2(e_{\tilde{\imath}}, y)$$

then for player II, $(1-\lambda)y + \lambda e_{\tilde{\jmath}}$ is an ε-best reply to $e_{\tilde{\imath}}$ for λ small enough.

For player 1, note that for $i \in \{1, \ldots, \tilde{M}\}$,

$$\Phi^1(e_i, (1-\lambda)y + \lambda e_{\tilde{\jmath}}) = \frac{B^1(e_{\tilde{\jmath}})}{b(e_{\tilde{\jmath}})},$$

and that

$$\Phi^1(e_i, (1-\lambda)y + \lambda e_{\tilde{\jmath}}) = \frac{A^1(e_i) + \lambda B^1(e_{\tilde{\jmath}})}{a(e_i) + \lambda b(e_{\tilde{\jmath}})}$$

for $i \in \{\tilde{M}+1, \ldots, M\}$. It follows from (3.19), (3.20) and (3.21) that $e_{\tilde{\imath}}$ is an ε-best reply to $(1-\lambda)y_0 + \lambda e_{\tilde{\jmath}}$ for λ small enough.

The only remaining case is:

$$\max_{i \in \{\tilde{M}+1,\ldots,M\}} \frac{A^1(e_i)}{a(e_i)} < \min_{j \in \{\tilde{N}+1,\ldots,N\}} \frac{B^1(e_j)}{b(e_j)} \tag{3.22}$$

By the ARAT property it follows that the game restricted to $\{1, 2, \ldots, \tilde{M}\} \times \{1, 2, \ldots, \tilde{N}\}$ has a pure Nash equilibrium, say (k, l). If

$$a^1(k) + b^1(l) \geq \max_{i \in \{\tilde{M}+1,\ldots,M\}} \frac{A^1(e_i)}{a(e_i)}$$

and if

$$a^2(k) + b^2(l) \geq \max_{j \in \{\tilde{N}+1,\ldots,N\}} \frac{B^2(e_j)}{b(e_j)},$$

then this pure equilibrium is also an equilibrium in the original repeated game.

Let $\tilde{\imath}$ be such that

$$\frac{A^1(e_{\tilde{\imath}})}{a(e_{\tilde{\imath}})} = \max_{i \in \{\tilde{M}+1,\ldots,M\}} \frac{A^1(e_i)}{a(e_i)},$$

and let $\tilde{\jmath}$ be such that

$$\frac{B^2(e_{\tilde{\jmath}})}{b(e_{\tilde{\jmath}})} = \max_{j \in \{\tilde{N}+1,\ldots,N\}} \frac{B^2(e_j)}{b(e_j)}.$$

If $\dfrac{A^1(e_{\bar{i}})}{a(e_{\bar{i}})} > a^1(i) + b^1(j)$. then we claim that $(e_{\bar{i}}, e_j)$ is an equilibrium point. By the definition of $e_{\bar{i}}$:

$$\Phi^1(e_{\bar{i}}, e_j) \geq \Phi^1(e_i, e_j), \ \forall i \in \{1, \ldots, N\},$$

and

$$\Phi^2(e_{\bar{i}}, e_j) \geq \Phi^2(e_{\bar{i}}, y) \ , \ \forall y \in \Delta_N \ ,$$

since $e_{\bar{i}} \in E_1 = E_j$ by lemma 3.6.

If $\dfrac{B^2(e_{\bar{j}})}{b(e_{\bar{j}})} > a^2(i) + b^2(j)$ then we claim that $(e_i, e_{\bar{j}})$ is an equilibrium point. By the definition of $e_{\bar{j}}$:

$$\Phi^2(e_i, e_{\bar{j}}) \geq \Phi^2(e_i, e_j) \ , \ \forall j \in \{1, \ldots, M\},$$

and by assumption (3.22) we have

$$\frac{A^1(x)}{a(x)} < \frac{B^1(e_{\bar{j}})}{b(e_{\bar{j}})} \ , \ \forall x \in X_{abs},$$

so

$$\Phi^1(x, e_{\bar{j}}) \ = \ \frac{A^1(x) + B^1(e_{\bar{j}})}{a(x) + b(e_{\bar{j}})} \ < \ \frac{B^1(e_{\bar{j}})}{b(e_{\bar{j}})} \ = \ \Phi^1(e_i, e_{\bar{j}}).$$

Further $\Phi^1(e_{i_1}, e_{\bar{j}}) \ = \Phi^1(e_{i_2}, e_{\bar{j}}) \ , \ \forall i_1, i_2 \in \{1, \ldots, \tilde{M}\}$. This leads us to the conclusion that e_i is a best answer to $e_{\bar{j}}$.

All the possible cases have been considered which completes the proof. □

Summarizing Theorems 3.1, 3.4, 3.5 and 3.7 we have:

Theorem 3.8 *Every repeated ARAT game possesses an ε-equilibrium point in stationary strategies.*

REFERENCES

Bewley, T. and Kohlberg, E. (1976). The asymptotic theory of stochastic games. *Operations Research* **1** 197-208.

Blackwell, D. (1962). Discrete dynamic programming. *Ann. Math. Statist.* **33** 719-726.

Blackwell, D. (1965). Discounted dynamic programming. *Ann. Math. Statist.* **36** 226-235.

Blackwell, D. and Ferguson, T. S. (1968). The big match. *Ann. Math. Statist.* **39** 159-163.

EVANGELISTA, FE (1993). Equilibrium in bimatrix games and in repeated games with additive rewards and transitions. PhD thesis, University of Illinois at Chicago.

FAN, K. (1952). Fixed point and minimax theorems in locally convex topological linear spaces. *Proc. Nat. Acad. Sciences* **38** 121-126.

FILAR, J. A. (1981). A single loop stochastic game which one player can terminate. *Opsearch* **18** 185-203.

FILAR, J. A. AND RAGHAVAN, T. E. S. (1991). Algorithms for stochastic games - a survey. *Methods and Models of Operations Research* **35** 437-472.

FINK, A. M. (1964). Equilibrium points of stochastic noncoooperative games. *J. Sci. Hiroshima University, Series A* **28** 89-93.

FLESCH, F. (1995). Tech. Rept 09-1995, Dept. Math. University of Limburg, Mastricht.

GLICKSBERG, I. L. (1952). A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points. *Proc Amer Math Soc* **38** 170-174.

KOHLBERG, E. (1974). Repeated games with absorbing states. *Ann. Statistics* **2** 724-738.

MERTENS, J. F. AND NEYMAN, A. (1991). Stochastic games. *Int. J. of Game Theory* **10** 53-66.

PARTHASARATHY, T. AND RAGHAVAN, T. E. S. (1971). *Some topics in two-person games*. American Elsevier Publishers, New York.

PARTHASARATHY, T. AND RAGHAVAN, T. E. S. (1981). An orderfield property for stochastic games when one player controls transition probabilities. *J. Optimization Theory and Appl.* **33** 375-392.

PARTHASARATHY, T., TIJS, S. H. AND VRIEZE, O. J. (1984). Stochastic games with state independent transitions and reparable rewards. In Hammer G., Pallasche D. (eds), *Selected Topics in Operations Research and Mathematical Economics.* Springer Verlag, Berlin, 262-271.

RAGHAVAN, T. E. S., TIJS, S. H. AND VRIEZE, O. J. (1985). On stochastic games with additive reward and additive transition structure. *J. Optimization Theory and Appl.* **47** 451-464.

ROGERS, P. D. (1969). Nonzero-sum stochastic games. Report ORC 69-8, Operations Research Centre, University of California, Berkeley.

STERN, M. A. (1975). On stochastic games with limiting average payoff. Ph.D. thesis, University of Illinois, Chicago.

THUIJSMAN, F. AND RAGHAVAN, T. E. S. (1994). Stochastic games with switching control or ARAT structure. Report M94-06, Department of Mathematics, University of Limburg, Maastricht.

THUIJSMAN F. AND VRIEZE, O. J. (1991). Easy initial states in stochastic games. In: Raghavan T.E.S., et.al. (eds). *Stochastic Games and Related Topics*. Kluwer Academic Publishers, Dordrecht: 85-100.

VRIEZE, O. J. AND THUIJSMAN, F. (1989). On equilibria in repeated games with absorbing states. *Int. J. of Game Theory* **18** 293-310.

FE EVANGELISTA
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN AT WAUSAU
WAUSAU, WI 54401
`fevangel@uwcmail.uwc.edu`

O. J. VRIEZE
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LIMBURG
P.O. BOX 616, MAASTRICHT
THE NETHERLANDS
`OJ.Vrieze@math.rulimburg.nl`

T. E. S. RAGHAVAN
DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS
CHICAGO, IL 60607-7045
`ter@uic.edu`