JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 34, No. 4, AUGUST 1981

Ordered Field Property for Stochastic Games When the Player Who Controls Transitions Changes from State to State[1]

J. A. FILAR[2]

Communicated by G. Leitmann

Abstract. In this paper, we consider a zero-sum stochastic game with finitely many states restricted by the assumption that the probability transitions from a given state are functions of the actions of only one of the players. However, the player who thus controls the transitions in the given state will not be the same in every state. Further, we assume that all payoffs and all transition probabilities specifying the law of motion are rational numbers. We then show that the values of both a β-discounted game, for rational β, and of a Cesaro-average game are in the field of rational numbers. In addition, both games possess optimal stationary strategies which have only rational components. Our results and their proofs form an extension of the results and techniques which were recently developed by Parthasarathy and Raghavan (Ref. 1).

Key Words. Stochastic games, discounting, undiscounted stochastic games, stationary strategies, Cesaro-average payoff, probability transitions, Archimedean field.

1. Introduction

In a recent paper, Parthasarathy and Raghavan (Ref. 1) proved that the value and at least one pair of optimal strategies of a stochastic game lie in the same ordered Archimedean field as the data describing the game, provided that only one player controls the transition probabilities in all states.

The results of Parthasarathy and Raghavan include an algorithm for solving *β-discounted games* in which one player controls the *law of motion*,

[1] The author wishes to thank Professor T. E. S. Raghavan for introducing him to this problem and for discussing stochastic games with him on many occasions. This research was supported in part by AFOSR Grant No. 78-3495B.

[2] Assistant Professor, Department of Applied Statistics, School of Statistics, University of Minnesota, St. Paul, Minnesota. Current address, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, Maryland.

and they suggest that an algorithm should also exist for solving the *undiscounted games* or *Cesaro-average games*. In fact, Filar and Raghavan (Ref. 2) just proposed a finite-step algorithm which does precisely that.

A natural generalization of the above class of games is one in which one player controls transition probabilities in some states, while the other player controls these transitions in the remaining states. This generalization was first suggested by Maschler during the Game Theory Workshop at Cornell University (1978); it has intuitive appeal, since we can easily imagine situations where a player may be tempted to enter a state of the game with possible high rewards, but at the cost of losing the ability to control future transitions. In this paper, we show that, in zero-sum discounted and undiscounted stochastic games in which the control of transition probabilities changes from player to player, depending on the state, the value of the game and at least one pair of optimal stationary strategies exist and lie in the same ordered Archimedean field as the data describing these games.

It must be mentioned that the existence part of the above statement can be derived from Bewley and Kohlberg's results (Ref. 3). The proofs given here are quite unrelated to Bewley and Kohlberg's work, but they are an adaptation of Parthasarathy and Raghavan's approach to this more general class of games.

There are two basic reasons why the original proofs of Ref. 1 cannot be extended immediately to our class of games. First, the linear programs used in Ref. 1 to solve the β-discounted games are no longer linear, since the transition probabilities now depend on actions of different players in different states. For the same reason, the probability transition matrix (which determines the game when stationary strategies are used) will, in our situation, depend on the strategies of both players, which invalidates some limiting operations (such as $\beta \to 1^-$), that were crucial to the arguments of Ref. 1 for the undiscounted games. Secondly, Parthasarathy and Raghavan relied on the fact that one player possessed a uniformly discount optimal stationary strategy in all the states, which will no longer hold under our generalization.

Fortunately, it turns out that, by exploiting the special structure of the probability transition matrices that can occur in our games, we can, with the help of some results from Blackwell (Ref. 4), extend the basic line of argument of Ref. 1 to this new situation.

2. Definitions and Notation

A *stochastic game*, as formulated by Shapley (Ref. 5), is played in stages. At each stage, the game is in one of finitely many states, $s =$

$1, 2, \ldots, S$, in which players I and II are obliged to play a matrix game,

$$A^s = (a_{ij}^s)_{i,j=1}^{m_s, n_s},$$

once. The *law of motion* is defined by the probabilities $q(s'/s, i, j)$, where the event $\{s'/s, i, j\}$ is the event that the game will enter state s' at the next stage, given that, at the current stage, the state of the game is s, and given that players I and II choose the ith row and the jth column of A^s, respectively. In general, the players' strategies will depend on complete past histories. In this paper, however, we shall be concerned only with *stationary strategies.* We may represent a typical stationary strategy f for player I by a *composite vector,*

$$f = (f(1), f(2), \ldots, f(S)),$$

where each $f(s)$ is a probability vector[3] given by

$$f(s) = (f_1(s), f_2(s), \ldots, f_{m_s}(s)).$$

Here, $f_i(s)$ is the probability that player I chooses the ith row of A^s whenever the game is in state s. Player II's stationary strategies are similarly defined.

Once we specify the initial state and a strategy pair (f, g) for players I and II, we implicitly define a probability distribution over all sequences of states and actions which can occur during the game and consequently over all sequences of payoffs to player I. Let $\pi_n(f, g)(s)$ denote the expected income to player I at the nth stage when players I and II use the strategy pair (f, g) and the game begins in s. The two types of stochastic games which we shall consider are determined by the manner in which the players evaluate a stream of payoffs $(\pi_1, \pi_2, \ldots)$. They are given below.

β-Discounted Games. Here,

$$\Gamma_\beta = \{\Gamma_\beta(1), \Gamma_\beta(2), \ldots, \Gamma_\beta(S)\},$$

$\beta \in (0, 1)$, and $\Gamma_\beta(s)$ refers to the game beginning in state s. In such games, $\Phi_\beta(f, g)(s)$, the expected income to player I in $\Gamma_\beta(s)$ when the strategy pair (f, g) is used is defined by

$$\Phi_\beta(f, g)(s) = \sum_{n=1}^{\infty} \beta^{n-1} \pi_n(f, g)(s). \tag{1}$$

[3] Throughout this paper, we shall not differentiate between n-component row and column vectors. This is intended to simplify the already complicated notation and should not confuse the reader.

Undiscounted Games or Cesaro-Average Games.[4] Here,

$$\Gamma = \{\Gamma(1), \Gamma(2), \ldots, \Gamma(S)\}.$$

In such games, $\Phi(f, g)(s)$ is defined by

$$\Phi(f, g)(s) = \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \pi_n(f, g)(s). \tag{2}$$

Note that $\Phi(f, g)(s)$ has an analogous meaning to $\Phi_\beta(f, g)(s)$. To show that a number $v_\beta(s)$ is the *value* of $\Gamma_\beta(s)$, $s = 1, 2, \ldots, S$, it is sufficient to show that there exists a stationary pair of strategies (f^β, g^β) such that, for each s,

$$\Phi_\beta(f, g^\beta)(s) \leq \Phi_\beta(f^\beta, g^\beta)(s) = v_\beta(s) \leq \Phi_\beta(f^\beta, g)(s), \tag{3}$$

for any stationary f for player I and for any stationary g for player II. For the undiscounted game, $v(s)$, $s = 1, 2, \ldots, S$, and an optimal stationary pair (f^0, g^0) is defined similarly.

All the stochastic games considered below will be constrained by the following hypotheses.

Hypothesis (H1). There exists an integer S_1, $S_1 < S$, for which the law of motion satisfies

$$q(s'/s, i, j) = \begin{cases} q(s'/s, i), & \text{if } s \leq S_1, \\ q(s'/s, j), & \text{if } S_1 < s \leq S. \end{cases}$$

Hypothesis (H2). All entries of the matrices A^s, $s = 1, 2, \ldots, S$, and all transition probabilities $q(s'/s, i)$, $q(s'/s, j)$ are rational numbers.

Hypothesis (H1) simply states that player I controls the law of motion in states $1, 2, \ldots, S_1$, while player II controls the law of motion in the remaining states. Of course, it is irrelevant which set of states is controlled by which player, since it is always possible to relabel the states. Furthermore, all of the results derived under Hypothesis (H2) will extend naturally from rational numbers to any ordered Archimedean field.

It should be clear that a stationary pair (f, g) determines an $S \times S$ Markov matrix

$$Q(f, g) = [q(s'/s, f, g)]_{s,s'=1}^{S},$$

[4] It must be mentioned that the payoff criterion (2) is only one of a number of criteria which may be used when discounting is not appropriate. Bewley and Kohlberg (Ref. 3) consider as many as six alternative criteria [including (2)]; however, they show that all six are equivalent in games which possess optimal stationary strategies.

where, due to Hypothesis (H1), we have

$$q(s'/s, f, g) = \begin{cases} \sum_{i=1}^{m_s} q(s'/s, i) f_i(s), & \text{if } s \leq S_1, \\ \sum_{j=1}^{n_s} q(s'/s, j) g_j(s), & \text{if } s > S_1. \end{cases} \tag{4}$$

Now, if (f, g) is any pair of stationary strategies, we define a *current payoff vector* associated with this pair by

$$r = r(f, g) = (r(f, g)(1), r(f, g)(2), \ldots, r(f, g)(S)),$$

where

$$r(f, g)(s) = f(s) A^s g(s) = \sum_{i=1}^{m_s} \sum_{j=1}^{n_s} a_{ij}^s f_i(s) g_j(s). \tag{5}$$

Further, since $Q(f, g)$ is a Markov matrix, it is known that there exists a Markov matrix $Q^*(f, g)$, such that

$$Q^*(f, g) = \lim_{N \to \infty} \left[(1/(N+1)) \sum_{n=0}^{N} Q^n(f, g) \right], \tag{6}$$

where

$$Q^0(f, g) = I$$

is the identity matrix. The proof of the following lemma is almost identical to Blackwell's proof of Theorem 4(a) of Ref. 4.

Lemma 2.1. (a) For any stationary pair of strategies (f, g) and any $\beta \in (0, 1)$, we have

$$\begin{aligned} \Phi_\beta(f, g) &= \sum_{n=0}^{\infty} \beta^n Q^n(f, g) r(f, g) \\ &= [1/(1-\beta)] Q^*(f, g) r(f, g) + y(f, g) + E(\beta, f, g), \end{aligned}$$

where $E(\beta, f, g) \to 0$ as $\beta \to 1^-$ and where the components of the $S \times 1$ vector $y(f, g)$ are bounded.

(b) Under conditions (a),

$$\Phi(f, g) = \lim_{\beta \to 1^-} (1-\beta) \Phi_\beta(f, g) = Q^*(f, g) r(f, g).$$

3. Technical Preliminaries

We shall need four lemmas proved in Parthasarathy and Raghavan (Ref. 1).

Lemma 3.1. Let $r(t)=p(t)/q(t)$ be a rational function, well defined for $t\in(\beta_0, 1)$. If $r(t)$ is a rational number for all rational t in $(\beta_0, 1)$, then $r(t)=p^*(t)/q^*(t)$, where p^* and q^* are polynomials with rational coefficients.

Lemma 3.2. Let $f(\beta)$ be a rational function bounded in some interval $(\beta_0, 1)$. Further, if $f(\beta)$ is rational when β is rational, then $\lim_{\beta\to 1^-} f(\beta)$ is a rational number.

Lemma 3.3. Let $v(\beta)$ be a continuous vector function for $\beta\in(0, 1)$. Let $u_j(\beta)$, $j=1, 2, \ldots, k$, be k vector functions which are rational functions of β componentwise. If, for each $\beta\in(0, 1)$, $v(\beta)$ coincides with one of these rational functions, then there exists some $\beta_0\in(0, 1)$ such that

$$v(\beta)\equiv u_j(\beta), \qquad \text{for all } \beta\in(\beta_0, 1),$$

for some fixed j.

Lemma 3.4. Let

$$A=(a_{ij}+a_i)_{i,j=1}^{n,n}$$

be a nonsingular matrix, with $a_{ij}>0$ for all (i, j). Further, let

$$xA=\alpha\underline{1}$$

have a nonnegative solution x, with

$$\sum_{i=1}^{n} x_i=1;$$

here,

$$\underline{1}=(1, 1, \ldots, 1).$$

Then, the matrix

$$\bar{A}=(a_{ij})_{i,j=1}^{n,n}$$

is nonsingular, and

$$x\bar{A}=\theta\underline{1}, \qquad \text{for some } \theta.$$

Analogous result holds if A is of the form

$$A=(a_{ij}+b_j)_{i,j=1}^{n,n}.$$

4. β-Discounted Games

With each state payoff matrix A^s, we shall associate a Shapley *dummy matrix* for every $\beta \in (0, 1)$:

$$A^s(\beta) = \left[a^s_{ij} + \beta \sum_{s'=1}^{S} v_\beta(s') q(s'/s, i, j) \right]_{i,j=1}^{m_s, n_s}. \tag{7}$$

Note that Hypothesis (H1) ensures that, in our case,

$$\sum_{s'=1}^{S} v_\beta(s') q(s'/s, i, j)$$

is a function of only one of the indices i and j, depending on whether $s \leq S_1$ or not.

Shapley (Ref. 5) proved that, if

$$(f^\beta(s), g^\beta(s))$$

is an optimal strategy pair in the matrix game $A^s(\beta)$ for each s, then

$$f^\beta = (f^\beta(1), \ldots, f^\beta(S)) \qquad \text{and} \qquad g^\beta = (g^\beta(1), \ldots, g^\beta(S))$$

are optimal stationary for players I and II, respectively, in the stochastic game Γ_β.

Now, we may apply the Snow–Shapley theorem (Ref. 6) to the matrix games $A^s(\beta)$, for each s and any $\beta \in (0, 1)$. This guarantees the existence of a nonsingular submatrix $\dot{A}^s(\beta)$ of $A^s(\beta)$ and a pair of extreme optimal strategies $(f^\beta(s), g^\beta(s))$ which, when appropriately truncated [by deletion of 0 entries corresponding to rows and columns not present in $\dot{A}^s(\beta)$], satisfy[5]

$$f^\beta(s)\dot{A}^s(\beta) = v_\beta(s)\underline{1} \qquad \text{and} \qquad \dot{A}^s(\beta) g^\beta(s) = v_\beta(s)\underline{1}. \tag{8}$$

However, by Lemma 3.4, the submatrix $\dot{A}^s$ of A^s, which suppresses the same rows and columns as $\dot{A}^s(\beta)$, is also nonsingular and satisfies

$$\begin{aligned} f^\beta(s)\dot{A}^s &= \dot{\theta}^\beta(s)\underline{1}, \qquad \text{if } s \leq S_1, \\ \dot{A}^s g^\beta(s) &= \dot{\theta}^\beta(s)\underline{1}, \qquad \text{if } s > S_1. \end{aligned} \tag{9}$$

Since $f^\beta(s)$, $g^\beta(s)$ (and their truncations) are probability vectors, we find from (9) that their entries and the value of $\dot{\theta}^\beta(s)$ depend only on the submatrix $\dot{A}^s$. However, $\dot{A}^s$ might vary among the finitely many square

[5] We shall not differentiate between $f^\beta(s)$, $g^\beta(s)$ and their truncations in order not to complicate the notation even more. However, this ambiguity must be remembered.

submatrices of A^s, depending on β. Thus, we have that

$$\dot{\theta}^\beta(s) = \dot{\theta}(s) = \mathrm{Det}(\dot{A}^s) \Big/ \Big(\sum_i \sum_j \dot{A}^s_{ij}\Big), \tag{10}$$

$$f^\beta_i(s) = f_i(s) = \Big(\sum_j \dot{A}^s_{ij}\Big) \Big/ \Big(\sum_i \sum_j \dot{A}^s_{ij}\Big), \qquad \text{if } s \le S_1, \tag{11}$$

$$g^\beta_j(s) = g_j(s) = \Big(\sum_i \dot{A}^s_{ij}\Big) \Big/ \Big(\sum_i \sum_j \dot{A}^s_{ij}\Big), \qquad \text{if } s > S_1, \tag{12}$$

where $\dot{A}^s_{ij}$ is the cofactor of the (i, j)th entry of $\dot{A}^s$.

Now, each A^s has only finitely many (say, k_s) nonsingular submatrices which define probability vectors through (11) and (12). Let us number these $A^s_1, A^s_2, \ldots, A^s_{k_s}$ for each s. Consider all permutations of the form

$$\kappa = (k(1), k(2), \ldots, k(S)),$$

where

$$k(s) \in \{1, 2, \ldots, k_s\}.$$

There are

$$\mu = \prod_{s=1}^{S} k_s$$

such permutations, and they can be labeled $\kappa_1, \kappa_2, \ldots, \kappa_\mu$, according to some ordering. Thus, we have μ vectors

$$\theta_l = (\theta_{l(1)}(1), \theta_{l(2)}(2), \ldots, \theta_{l(S)}(S)),$$

where l corresponds to κ_l, which in turn corresponds to the selection

$$(A^1_{l(1)}, A^2_{l(2)}, \ldots, A^S_{l(S)})$$

of submatrices of A^1 through A^s. Such a selection, of course, determines player I's strategy in states $1, 2, \ldots, S_1$ and player II's strategy in states $S_1+1, \ldots, S$ via (11) and (12). Now, define a stationary strategy f^l for player I, corresponding to

$$\kappa_l = (l(1), \ldots, l(S)),$$

by

$$f^l(s) = \begin{cases} \theta_{l(s)}(s)\underline{1}, (A^s_{l(s)})^{-1}, & \text{if } s \le S_1, \\ \text{arbitrary and fixed}, & \text{if } s > S_1; \end{cases}$$

and similarly, for player II, let

$$g^l(s) = \begin{cases} \text{arbitrary and fixed}, & \text{if } s \le S_1, \\ \theta_{l(s)}(s)(A^s_{l(s)})^{-1}\underline{1}, & \text{if } s > S_1. \end{cases}$$

Thus, we have formed μ stationary strategy pairs (f^l, g^l). Let

$$Q_l = Q_l(f^l, g^l)$$

be the probability transition matrix determined by the pair (f^l, g^l). Then, for $\beta \in (0, 1)$, the matrix $I - \beta Q_l$ is nonsingular; so, we may define the following S-vector rational functions of β:

$$u_\beta = (I - \beta Q_l)^{-1}\theta_l, \qquad l = 1, 2, \ldots, \mu. \tag{13}$$

We are now in a position to prove the following theorem for β-discounted games.

Theorem 4.1. Let Γ_β be constrained by Hypothesis (H1), for $\beta \in (0, 1)$. Then, the following results hold:

(i) $v_\beta(s)$ is a rational function of β for all s, if β is sufficiently near 1;

(ii) if $s \leq S_1$, player I has a *uniformly discounted optimal strategy* $f^\circ(s)$ (i.e., optimal for all β sufficiently near 1), while player II has a uniformly discounted optimal strategy $g^\circ(s)$, for $s > S_1$.

Proof. By the Snow–Shapley theorem and (8), to every $\beta \in (0, 1)$ there corresponds a permutation κ_l of submatrices $A^s_{l(s)}$ of A^s for $s = 1, 2, \ldots, S$ and a pair of optimal strategies (f^l, g^l) satisfying (8)–(12) for every s. In particular, (9) and (8) can be combined to express $\theta_{l(s)}(s)$ as

$$\begin{aligned}\theta_{l(s)}(s) &= \sum_i f^l_i(s) a^s_{ij} = v_\beta(s) - \beta \sum_{s'} \left[\sum_i f^l_i(s) q(s'/s, i)\right] v_\beta(s') \\ &= v_\beta(s) - \beta \sum_{s'} q(s'/s, f^l) v_\beta(s'), \qquad \text{if } s \leq S_1;\end{aligned}$$

and similarly, if $s > S_1$, we have

$$\theta_{l(s)}(s) = v_\beta(s) - \beta \sum_{s'} q(s'/s, g^l) v_\beta(s').$$

Thus, in matrix form, we can rewrite the above equations as

$$[I - \beta Q(f^l, g^l)]^{-1}\theta_l = v_\beta, \tag{14}$$

where

$$v_\beta = (v_\beta(1), v_\beta(2), \ldots, v_\beta(S)).$$

Hence, v_β which is continuous in β, coincides with one of the μ rational functions u^l_β for each $\beta \in (0, 1)$ [see (13)]. Thus, by Lemma 3.3, there exists some $\beta_0 \in (0, 1)$ and some fixed l_0, such that

$$v_\beta \equiv u^{l_0}_\beta, \qquad \text{for all } \beta \in (\beta_0, 1).$$

This proves (i). Further,

$$f^{\beta}(s)=f^{l_0}(s), \qquad \text{for all } \beta\in(\beta_0,1), \text{ if } s\le S_1,$$

and

$$g^{\beta}(s)=g^{l_0}(s), \qquad \text{for all } \beta\in(\beta_0,1), \text{ if } s> S_1,$$

by (11)–(12). Now, (ii) follows from (i).

Corollary 4.1. If, in addition to Hypothesis (H1), we restrict Γ_β by Hypothesis (H2) and take any rational $\beta\in(0,1)$, then $v_\beta(s)$ is a rational number, and there is at least one optimal strategy pair (f^β, g^β) which has only rational entries.

Proof. Equation (14) still applies for some l, which corresponds to some selection $A^1_{l(1)},\ldots,A^S_{l(S)}$ of the submatrices of $A^1,\ldots,A^S$. Thus, $v_\beta(s)$ is rational for each s. The rationality of f^l and g^l follows immediately from (8).

Assume now that β_0 and l_0 are as in Theorem 4.1. For simplicity, we shall write A^s_0, $f^\circ(s)$, $g^\circ(s)$ in place of $A^s_{l_0}$, $f^{l_0}(s)$, $g^{l_0}(s)$. Further, for $\beta>\beta_0$ and for the pair (f^β, g^β) chosen by the Snow–Shapley theorem, we have from (8)

$$f^{\beta}(s)=v_\beta(s)\underline{1}[A^s_0(\beta)]^{-1}, \qquad \text{if } s>S_1. \tag{15}$$

From Theorem 4.1, it follows that $f^\beta(s)$ for $s>S_1$ is a rational function of β. Since it is a probability vector,

$$\lim_{\beta\to1^-} f^\beta(s)=\tilde f(s)$$

exists and is itself a probability vector. Similarly, let

$$\tilde g(s)=\lim_{\beta\to1^-} g^\beta(s)=\lim_{\beta\to1^-}\{v_\beta(s)[A^s_0(\beta)^{-1}\underline{1}\}, \qquad \text{for } s\le S_1.$$

We can now form a stationary strategy pair $(\tilde f, \tilde g)$, defined by

$$\begin{aligned}\tilde f&=(f^\circ(1),\ldots,f^\circ(S_1),\tilde f(S_1+1),\ldots,\tilde f(S)),\\ \tilde g&=(\tilde g(1),\ldots,\tilde g(S_1),g^\circ(S_1+1),\ldots,g^\circ(S)).\end{aligned} \tag{16}$$

In the next section, we shall show that $(\tilde f, \tilde g)$ is an optimal strategy pair in the undiscounted game Γ.

5. Undiscounted Game

Let (f^β, g^β), $(\tilde{f}, \tilde{g})$ be defined as in Section 4. Then, if $\beta > \beta_0$, we have

$$Q(f^\beta, g^\beta) = Q(\tilde{f}, \tilde{g}). \tag{17}$$

The above equation holds, because, for $\beta > \beta_0$ and $s \le S_1$,

$$q(s'/s, f^\beta) = \sum_{i=1}^{m_s} q(s'/s, i) f_i^\beta(s) = \sum_{i=1}^{m_s} q(s'/s, i) f_i^\circ(s)$$
$$= q(s'/s, f^\circ),$$

by Theorem 4.1(ii); and similarly,

$$q(s'/s, g^\beta) = q(s'/s, g^\circ), \qquad \text{if } s \ge S_1 + 1.$$

Likewise, for $\beta > \beta_0$ and any fixed stationary strategy f for player I, we have

$$Q(f, g^\beta) = Q(f, \tilde{g}).$$

Thus, for all $\beta > \beta_0$,

$$Q^*(f^\beta, g^\beta) = Q^*(\tilde{f}, \tilde{g}) \qquad \text{and} \qquad Q^*(f, g^\beta) = Q^*(f, \tilde{g}). \tag{18}$$

Similarly, note that, for $\beta > \beta_0$,

$$r(f^\beta, g^\beta)(s) = f^\beta(s) A^s g^\beta(s) = \theta_{l_0}(s), \qquad \text{for every } s. \tag{19}$$

For instance, if $s \le S_1$,

$$f^\beta(s) A^s g^\beta(s) = f^\circ(s) A^s_{l_0} g^\beta(s) = \theta_{l_0}(s) \underline{1} g^\beta(s) = \theta_{l_0}(s).$$

Theorem 5.1. In the undiscounted stochastic game restricted by Hypothesis (H1), the values $v(s)$, $s = 1, 2, \ldots, S$, and a pair of optimal stationary strategies $(\tilde{f}, \tilde{g})$ exist. Further,

$$v(s) = \lim_{\beta \to 1^-} (1-\beta) v_\beta(s), \qquad \text{for each } s.$$

Proof. Consider only $\beta > \beta_0$. Then, we have

$$(1-\beta) v_\beta = (1-\beta)\Phi_\beta(f^\beta, g^\beta) = (1-\beta) \sum_{n=0}^{\infty} \beta^n Q^n(f^\beta, g^\beta) r(f^\beta, g^\beta)$$
$$= (1-\beta) \sum_{n=0}^{\infty} \beta^n Q^n(\tilde{f}, \tilde{g}) \theta_{l_0},$$

by (18) and (19). However,

$$\lim_{\beta \to 1^-} (1-\beta) \sum_{n=0}^{\infty} \beta^n Q^n(\tilde{f}, \tilde{g}) \theta_{l_0} = Q^*(\tilde{f}, \tilde{g}) \theta_{l_0},$$

by an argument similar to Blackwell's (Ref. 4, pp. 722–723). Note also that

$$r(\tilde{f}, \tilde{g}) = \theta_{l_0},$$

as in (19). Thus, it follows that

$$\lim_{\beta \to 1} (1-\beta) v_\beta = \Phi(\tilde{f}, \tilde{g}). \tag{20}$$

Choose any stationary strategy f for player I. It can be checked that

$$\lim_{\beta \to 1^-} (1-\beta) \sum_{n=0}^{\infty} \beta^n Q^n(f, \tilde{g})[r(f, g^\beta) - r(f, \tilde{g})] = 0. \tag{21}$$

This follows from the fact that (see Blackwell, Ref. 4, p. 722)

$$\lim_{\beta \to 1^-} \sum_{n=0}^{\infty} \beta^n [Q^n(f, \tilde{g}) - Q^*(f, \tilde{g})] = [I - Q(f, \tilde{g}) - Q^*(f, \tilde{g})]^{-1} - Q^*(f, \tilde{g})$$

and the fact that

$$r(f, g^\beta) - r(f, \tilde{g}) \to 0, \qquad \text{as } \beta \to 1^-.$$

Now, by (18), we have

$$\begin{aligned}(1-\beta)\Phi_\beta(f, g^\beta) &= \sum_{n=0}^{\infty} (1-\beta)\beta^n Q^n(f, \tilde{g}) r(f, g^\beta) \\ &= (1-\beta) \sum_{n=0}^{\infty} \beta^n Q^n(f, \tilde{g}) r(f, \tilde{g}) \\ &\quad + (1-\beta) \sum_{n=0}^{\infty} \beta^n Q^n(f, \tilde{g})[r(f, g^\beta) - r(f, \tilde{g})].\end{aligned}$$

So,

$$\lim_{\beta \to 1^-} (1-\beta)\Phi_\beta(f, g^\beta) = \Phi(f, \tilde{g}), \tag{22}$$

by (21) and the above equation. However, for any $\beta > \beta_0$,

$$(1-\beta) v_\beta = (1-\beta)\Phi_\beta(f^\beta, g^\beta) \geq (1-\beta)\Phi_\beta(f, g^\beta).$$

So, (20) and (22) imply that

$$\Phi(\tilde{f}, \tilde{g}) \geq \Phi(f, \tilde{g}), \tag{23}$$

for any stationary f. Similarly, it can be shown that

$$\Phi(\tilde{f}, \tilde{g}) \leq \Phi(\tilde{f}, g),$$

for any stationary g. This shows that

$$\Phi(\tilde{f}, \tilde{g})(s) = v(s), \qquad \text{for every } s.$$

Corollary 5.1. If the game is restricted by Hypothesis (H2) in addition to the hypotheses of Theorem 5.1, then $v(s)$, $s = 1, 2, \ldots, S$, and the components of $(\tilde{f}, \tilde{g})$ are all rational numbers.

Proof. The rationality of $f^\circ(s)$ for $s \le S_1$ and the rationality of $g^\circ(s)$ for $S_1 + 1 \le S$ follow immediately from (11) and (12). To see that $\tilde{f}(s)$ is rational for $s \ge S_1 + 1$, recall that, in view of Theorem 4.1, (15) defines a rational function of β which takes rational values whenever $\beta \in (\beta_0, 1)$ is rational. Thus, by Lemma 3.1, we can assume that $f_i^\beta(s)$ is a ratio of polynomials in β with rational coefficients for each i; and, by Lemma 3.2,

$$\lim_{\beta \to 1^-} f_i^\beta(s) = \tilde{f}_i(s)$$

is a rational number. The rationality of $\tilde{g}(s)$ for $s \le S_1$ follows similarly. Further,

$$\lim_{\beta \to 1^-} (1 - \beta) v_\beta(s) = v(s),$$

by Theorem 5.1; and, by (14), $(1 - \beta) v_\beta$ is rational whenever $\beta \in (\beta_0, 1)$ is rational; thus, $v(s)$ is a rational number by Lemma 3.2, for each s.

References

1. PARTHASARATHY, T., and RAGHAVAN, T. E. S., *An Order Field Property for Stochastic Games when One Player Controls Transitions*, Journal of Optimization Theory and Applications, Vol. 33, No. 3, 1981.
2. FILAR, J. A., and RAGHAVAN, T. E. S., *An Algorithm for Solving an Undiscounted Stochastic Game in Which One Player Controls Transitions*, Research Memorandum, University of Illinois, Chicago, Illinois, 1979.
3. BEWLEY, T., and KOHLBERG, E., *On Stochastic Games with Stationary Strategies*, Mathematics of Operations Research, Vol. 3, pp. 104–125, 1978.
4. BLACKWELL, D., *Discrete Dynamic Programming*, Annals of Mathematical Statistics, Vol. 33, pp. 719–726, 1962.
5. SHAPLEY, L. S., *Stochastic Games*, Proceedings of the National Academy of Science, Vol. 39, pp. 1095–1100, 1953.
6. SHAPLEY, L. S., and SNOW, R. N., *Basic Solutions of Discrete Games*, Annals of Mathematics Studies, Princeton University Press, Princeton, New Jersey, Vol. 24, pp. 27–37, 1950.