Int J Game Theory (1998) 27:61–81

International Journal of Game Theory

On the existence of good stationary strategies for nonleavable stochastic games

Piercesare Secchi

Dipartimento di Economia Politica e Metodi Quantitativi, Università di Pavia, Via San Felice, 5, 27100 Pavia, Italy

Received December 1995/Final version 1997

Abstract. This paper discusses the problem regarding the existence of optimal or nearly optimal stationary strategies for a player engaged in a nonleavable stochastic game. It is known that, for these games, player I need not have an ε-optimal stationary strategy even when the state space of the game is finite. On the contrary, we show that uniformly ε-optimal stationary strategies are available to player II for nonleavable stochastic games with finite state space. Our methods will also yield sufficient conditions for the existence of optimal and ε-optimal stationary strategies for player II for games with countably infinite state space. With the purpose of introducing and explaining the main results of the paper, special consideration is given to a particular class of nonleavable games whose utility is equal to the indicator of a subset of the state space of the game.

Key words: Stochastic games, leavable games, gambling theory

1. Introduction

Let S be a countable nonempty set of states, A and B two finite nonempty sets of actions for players I and II respectively and q a law of motion which assigns to each triple $(x, a, b) \in S \times A \times B$ a probability distribution defined on the subsets of S. On S define a bounded, real valued function u called the utility function. Consonant with Maitra and Sudderth [1992] S, A, B, q and u define a zero-sum, nonleavable stochastic game whose dynamics we can describe as follows. Given an initial state $x \in S$, player I chooses, possibly at random, an action $a_1 \in A$, player II chooses, possibly at random, an action $b_1 \in B$ and the game moves to its next state X_1 according to the probability distribution $q(\cdot|x, a_1, b_1)$. The new state X_1 is announced to both players along with their chosen actions. By iterating this procedure a random sequence of states $x, X_1, X_2, \ldots$ is produced and the payoff from player II to player I is fixed to be the expected value of

$$u^* = \limsup_{n\to\infty} u(X_n).$$

Maitra and Sudderth [1992] proved that any nonleavable game has a value, but the question about the existence of good stationary strategies for one or both players is still open in its complete generality. Our main result asserts that in any nonleavable game with finite state space there is a uniformly ε-optimal stationary strategy available to player II. An example of Nowak and Ragahvan [1991] implies that this result cannot be extended to nonleavable games with countably infinite state space. However our methods will also yield sufficient conditions for the existence of good stationary strategies for player II: for example, we will show that a uniformly optimal stationary strategy is always available to player II when the value function V of the nonleavable game is greater than or equal to its utility function u. If $V(x) < u(x)$ for some $x \in S$ there need not be an optimal strategy for player II even when the state space of the game is finite.

The same results do not hold for player I since Maitra and Sudderth [1996] showed with an example that, even when S is finite, there need not be an ε-optimal stationary strategy available to player I for a given initial state of the game.

The techniques used in the paper are those of Maitra and Sudderth [1992, 1996] and, in general, of the theory of gambling. In the next section we will set notation and terminology, which will follow that of Maitra and Sudderth [1996], and we will recall a few general results about leavable and nonleavable stochastic games. Section 3 examines the question as to whether stationary strategies exist for one or both players with respect to a special class of nonleavable games, namely those with utility function equal to a subset of the state space S: the section should be considered as an introduction to the main results of the paper which are presented in Section 4. A concluding remark will close the paper.

2. Preliminaries

Let $Z = S \times A \times B$ and define $H = Z \times Z \times \cdots$ to be the space of histories or sequences $h = (z_1, z_2, \ldots)$ of elements of Z. For $n = 1, 2, \ldots$ a partial history of length n is a sequence $p = (z_1, \ldots, z_n)$ of n elements of Z. A strategy α for player I is a sequence $\alpha_0, \alpha_1, \ldots$ such that α_0 is an element of $\mathscr{P}(A)$, the collection of all probability distributions defined on the subsets of A, and, for $n = 1, 2, \ldots, \alpha_n$ is a mapping which assigns to each partial history p of length n an element of $\mathscr{P}(A)$. Strategies β for player II are defined in the same way with B in place of A. A family of strategies $\bar{\alpha}$ for player I is a mapping from S to the collection $\mathscr{A}$ of all available strategies for player I; that is, for every $x \in S, \bar{\alpha}(x)$ is a strategy for player I. A family of strategies $\bar{\alpha}$ for player I is called stationary if there exists a function μ from S to $\mathscr{P}(A)$ such that

$$\bar{\alpha}(x)_0 = \mu(x),$$

$$\bar{\alpha}(x)_n(z_1, \ldots, z_n) = \mu(x_n)$$

for any integer $n, x \in S$ and $z_1 = (x_1, a_1, b_1), \ldots, z_n = (x_n, a_n, b_n)$. We write $\bar{\alpha} = \mu^\infty$. Analogous definitions and notations hold for player II.

Given an initial state $x \in S$, the law of motion q along with a strategy α for player I and a strategy β for player II determines a probability distribution $P_{x,\alpha,\beta}$ on the sigma-field of subsets of H generated by the coordinate functions

$$Z_n(z_1, z_2, \ldots, z_n, \ldots) = z_n$$

for $n = 1, 2, \ldots$. The expected value of a bounded, Borel measurable function g from H to the reals will be indicated with $\int g dP_{x,\alpha,\beta}$ or $E_{x,\alpha,\beta} g$.

If α is a strategy and $p = (z_1, \ldots, z_n)$ is a partial history of length n, the conditional strategy $\alpha[p]$ is defined by

$$\alpha[p]_0 = \alpha_n(p),$$

$$\alpha[p]_m(z'_1, \ldots, z'_m) = \alpha_{n+m}(z_1, \ldots, z_n, z'_1, \ldots, z'_m)$$

for all $m \geq 1$ and $(z'_1, \ldots, z'_m)$.

A stopping time t is a mapping from H to $\{0, 1, \ldots\} \cup \{\infty\}$ such that if h' agrees with h in the first $t(h)$ coordinates, then $t(h') = t(h)$. A stopping time which is everywhere finite is called a stop rule. Let t be a stop rule and $p = (z_1, \ldots, z_n)$ a partial history of length n; if $t(z_1, \ldots, z_n, z'_1, z'_2, \ldots) \geq n$ for any history $h' = (z'_1, z'_2, \ldots)$, we define the conditional stop rule $t[p]$ by setting

$$t[p](z'_1, z'_2, \ldots) = t(z_1, \ldots, z_n, z'_1, z'_2, \ldots) - n.$$

When t is a non-zero stop rule, define on H the function p_t by setting, for every history $h = (z_1, z_2, \ldots), p_t(h) = (z_1, \ldots, z_{t(h)})$. If g is a bounded, Borel measurable function from H to the reals, we will make frequent use of the following conditioning formula,

$$E_{x,\alpha,\beta} g = \int \{E_{X_t, \alpha[p_t], \beta[p_t]}(gp_t)\} \, dP_{x,\alpha,\beta}, \tag{2.1}$$

where, for every $h \in H, gp_t(h)$ is defined by

$$gp_t(h)(z'_1, z'_2, \ldots) = g(z_1, \ldots, z_{t(h)}, z'_1, z'_2, \ldots).$$

The same formula holds even when t is a non-zero stopping time such that $P_{x,\alpha,\beta}[t < \infty] = 1$.

Finally we introduce the one-day operator G defined, for every bounded, real-valued function ϕ on S and for every $x \in S$, by

$$(G\phi)(x) = \inf_{\nu} \sup_{\mu} E_{x,\nu,\mu} \phi, \tag{2.2}$$

where μ and ν range over $\mathscr{P}(A)$ and $\mathscr{P}(B)$ respectively and

$$E_{x,\mu,\nu}\phi = \sum_{a\in A}\sum_{b\in B}\sum_{x_1\in S}\phi(x_1)q(x_1|x,a,b)\mu\{a\}\nu\{b\}.$$

By von Neumann's Theorem [von Neumann and Morgenstern, 1947], $(G\phi)(x)$ is the value of the one-day game $\mathscr{A}(\phi)(x)$ where x is the initial state, players I and II choose, possibly at random, actions $a \in A$ and $b \in B$ respectively, and the game moves according to the law of motion q to the new state X_1; finally II pays I the expected value of $\phi(X_1)$. The same result by von Neumann proves also the existence of optimal randomized actions $\mu \in \mathscr{P}(A)$ and $\nu \in \mathscr{P}(B)$ for players I and II respectively.

We are now in the position to define the notions of leavable and nonleavable stochastic game. Given S, A, B, u and q the leavable game $\mathscr{L}(u)(x)$ with initial position $x \in S$ is a game where player I chooses a strategy $\alpha \in \mathscr{A}$ and a stop rule t, player II chooses a strategy $\beta \in \mathscr{B}$, and II pays to I the quantity $E_{x,\alpha,\beta}u(X_t)$. When I is not allowed to stop, the game $\mathscr{N}(u)(x)$ where the payoff from II to I is $E_{x,\alpha,\beta}u^*$ is called nonleavable.

Maitra and Sudderth [1992] proved that, for every $x \in S$,

$$\inf_{\beta}\sup_{\alpha,t} E_{x,\alpha,\beta}u(X_t) = \sup_{\alpha,t}\inf_{\beta} E_{x,\alpha,\beta}u(X_t) = U(x)$$

and

$$\inf_{\beta}\sup_{\alpha} E_{x,\alpha,\beta}u^* = \sup_{\alpha}\inf_{\beta} E_{x,\alpha,\beta}u^* = V(x).$$

The functions U and V are the values of the games $\mathscr{L}(u)$ and $\mathscr{N}(u)$ respectively. Given an initial state $x \in S$, a strategy $\alpha \in \mathscr{A}$ is optimal (ε-optimal) for player I for $\mathscr{N}(u)(x)$ if

$$E_{x,\alpha,\beta}u^* \geq V(x) \quad (E_{x,\alpha,\beta}u^* \geq V(x) - \varepsilon),$$

for any strategy β of player II. By reversing the inequalities we obtain analogous definitions for optimal (ε-optimal) strategies for player II.

In Section 3 we will be concerned with a special class of nonleavable games, namely those where u is the indicator function of a subset W of S. In this case the expressions above have a particular meaning. It is in fact easy to check that if, for every $x \in S, u(x) = I[x \in W]$, then

$$U(x) = \inf_{\beta}\sup_{\alpha} P_{x,\alpha,\beta}[\text{reach } W] = \sup_{\alpha}\inf_{\beta} P_{x,\alpha,\beta}[\text{reach } W]$$

where [reach W] is the event that is true if $X_n \in W$ for some $n \geq 0$. The expression of $V(x)$ becomes instead

$$\begin{aligned} V(x) &= \inf_{\beta}\sup_{\alpha} P_{x,\alpha,\beta}[X_n \in W \text{ infinitely often}] \\ &= \sup_{\alpha}\inf_{\beta} P_{x,\alpha,\beta}[X_n \in W \text{ infinitely often}] \end{aligned}$$

since, for every history $h = ((x_1, a_1, b_1), (x_2, a_2, b_2), \ldots) \in H$,

$$u^*(h) = \limsup_{n\to\infty} u(x_n) = \begin{cases} 1 & \text{if } x_n \in W \text{ infinitely often,} \\ 0 & \text{otherwise.} \end{cases}$$

Considerable use will be made of the following lemma, due to Maitra and Sudderth [1992], which characterizes the value function of a leavable game.

2.3 Lemma. *The value function U for the leavable game $\mathscr{L}(u)$ solves the optimality equation*

$$U = u \vee GU,$$

and is the least, bounded, real valued function ϕ defined on S such that

$$(a)\ \ \phi \geq u \quad \text{and} \quad (b)\ \ G\phi \leq \phi.$$

In order to introduce a similar characterization for the value V of a nonleavable game $\mathscr{N}(u)$ we need to define an operator T which maps every bounded, real valued function u defined on S to the bounded, real valued function Tu defined, for every $x \in S$, by

$$Tu(x) = (GU)(x)$$

where U is the value of the leavable game $\mathscr{L}(u)$. The operator T has many properties for which we refer to Maitra and Sudderth [1996, Chapter 7]. In this paper we will need the following two properties of T.

2.4 Lemma. *Let u, u_1 and u_2 be bounded, real valued functions defined on S and let c be a nonnegative real number. Then*

(i) *If $u_1 \leq u_2, Tu_1 \leq Tu_2$.*
(ii) *$T(cu) = cTu$.*

The proof of (i) and (ii) above is trivial and follows immediately from the definitions of the operators T and G. By means of the operator T Maitra and Sudderth [1992] characterized the function V as it is shown in the next lemma.

2.5 Lemma. *The value function V for the nonleavable game $\mathscr{N}(u)$ is the largest, bounded, real valued function ϕ defined on S such that*

$$T(u \wedge \phi) = \phi.$$

Here is an immediate consequence of Lemma 2.5.

2.6 Lemma. *The value function V for the nonleavable game $\mathscr{N}(u)$ is also the value function for the nonleavable game $\mathscr{N}(u \wedge V)$.*

Proof: Let V' be the value of the nonleavable game $\mathcal{N}(u \wedge V)$. Clearly $V' \leq V$ since $u \wedge V \leq u$. On the other hand,

$$V = T(u \wedge V) = T((u \wedge V) \wedge V)$$

so that $V \leq V'$ by Lemma 2.5. $\diamond$

Before concluding the section we want to recall a formula which is due to Sudderth [1971]. Let $x \in S$ and α and β be strategies for player I and II respectively. Then for every bounded, real-valued function u defined on S

$$E_{x,\alpha,\beta} u^* = \inf_{s} \sup_{t \geq s} E_{x,\alpha,\beta} u(X_t) \tag{2.7}$$

where s and t range over the set of stop rules.

3. Reaching a set infinitely often

In this section we will consider a special class of nonleavable stochastic games with utility function equal to the indicator of a subset of the state space.

Let S be a countable nonempty set of states and assume that W is a nonempty subset of S. For every $x \in S$, set $u(x) = I[x \in W]$. Define, as before, U to be the value function of the leavable game $\mathcal{L}(u)$ and V to be the value function of $\mathcal{N}(u)$. In the game $\mathcal{N}(u)$ the aim of player I is to choose a strategy which maximizes her probability of reaching the set W infinitely often, while player II seeks a strategy which minimizes the same quantity. When W has only one element, we will prove that player II always has an optimal stationary family of strategies, whereas player I need not have a stationary strategy which is ε-optimal for a given initial state. However if W has more than one element there need not be an optimal strategy, and hence not even an optimal stationary strategy, for player II even when the state space of the game is finite.

We will begin with a zero-one result which states that, if u is the indicator of a subset W, then either $\sup_{x \in W} V(x) = 1$ or $V \equiv 0$.

3.1 Theorem. *Let W be a nonempty subset of S and, for every $x \in S$, set $u(x) = I[x \in S]$. If $\sup_{x \in W} V(x) < 1$, then $V \equiv 0$.*

Proof: Let $p = \sup_{x \in W} V(x) < 1$ and consider the nonleavable game $\mathcal{N}(pu)$ which has the same state space, action sets and law of motion as $\mathcal{N}(u)$, but utility pu. Write V' for the value of $\mathcal{N}(pu)$.

We claim that

$$V' = pV. \tag{3.2}$$

In fact note that, for every history $h = ((x_1, a_1, b_1), (x_2, a_2, b_2), \ldots) \in H$,

$$(pu)^*(h) = \lim_{n \to \infty} \sup (pu)(x_n) = p \lim_{n \to \infty} \sup u(x_n) = (pu^*)(h).$$

For every $\varepsilon > 0$ and $x \in S$, let $\alpha_\varepsilon(x)$ be an ε-optimal strategy for player I in the game $\mathscr{N}(u)(x)$. Then

$$E_{x,\alpha_\varepsilon(x),\beta}(pu)^* = pE_{x,\alpha_\varepsilon(x),\beta}u^* \geq p(V(x) - \varepsilon)$$

for all $x \in S$ and all strategies β of player II; therefore $V'(x) \geq pV(x)$. Analogously, by considering the game from player II's point of view, one can show that $V'(x) \leq pV(x)$ for all $x \in S$ and this proves that $V' = pV$.

However it is also true that

$$V \leq V' \tag{3.3}$$

since, by Lemma 2.6, V is the value of $\mathscr{N}(u \wedge V)$ and $u \wedge V \leq pu$ because $V(x) \leq p$ for all $x \in W$.

Equations (3.2) and (3.3) imply that $V = pV$ and thus $V \equiv 0$ because $p < 1$. ◇

When the set W is finite, the previous theorem asserts that either there is an $x \in W$ such that $V(x) = 1$ or $V \equiv 0$. In particular if g is an element of S, which we may call the goal, and $W = \{g\}$, then either $V(g) = 1$ or $V \equiv 0$. In both cases there is an optimal stationary family of strategies available to player II: this is the content of Theorem 3.6 below. Before proving it we want however to show a result which conveys the idea that, in order to reach a goal g infinitely often, one must first reach g and then return to g infinitely often. The same result was proved by Sudderth [1969] for nonleavable gambling problems with a goal.

3.4 Theorem. *Let $g \in S$ and set, for every $x \in S, u(x) = I[x = g]$. Then, for every $x \in S$,*

$$V(x) = U(x)V(g). \tag{3.5}$$

Proof: Since $W = \{g\}$, Theorem 3.1 implies that either $V \equiv 0$ or $V(g) = 1$.

If $V \equiv 0$, then (3.5) is obviously true.

If $V(g) = 1$, then proving (3.5) reduces to showing that $V = U$. Since it is always true that $V \leq U$, we need to prove that in the present situation it is also $V \geq U$. Notice first that $V(g) = u(g) = 1$ implies that $V \geq u$. Therefore

$$V = T(u \wedge V) = Tu = GU$$

and thus

$$GV = G(GU) \leq GU = V$$

where the inequality is true because $GU \leq U$, by Lemma 2.3, and because it is trivial to show that $Gu_1 \leq Gu_2$ if $u_1 \leq u_2$ are two bounded, real valued functions on S. But then, by Lemma 2.3 again, $V \geq U$. ◇

Let ν be a function which maps every $x \in S$ to a probability measure belonging to $\mathscr{P}(B)$ which is optimal for player II in the one-day game $\mathscr{A}(U)(x)$

and consider the stationary family of strategies ν^∞. Maitra and Sudderth [1992] proved that ν^∞ is always optimal for player II for the leavable game $\mathscr{L}(u)$. The next theorem shows that ν^∞ is also optimal for player II for $\mathscr{N}(u)$ when u is the indicator function of a point.

3.6 Theorem. *Let $g \in S$ and, for every $x \in S$, set $u(x) = I[x = g]$. Then ν^∞ is an optimal stationary family of strategies for player II for $\mathscr{N}(u)$.*

Proof: Once again, because of Theorem 3.1, there are only two mutually exclusive cases to be considered.

Case 1: $V(g) = 1$. Fix $x \in S$ and a strategy α for player I. Then the sequence $\{U(X_n)\}$ is a bounded supermartingale with respect to $P_{x,\alpha,\nu^\infty(x)}$. In fact U is bounded between 0 and 1, and, for any $n \geq 0$,

$$\begin{aligned} E_{x,\alpha,\nu^\infty(x)}[U(X_{n+1})|X_1,\ldots,X_n] &= E_{X_n,\alpha[p_n]_0,\nu(X_n)}U \\ &\leq (GU)(X_n) \\ &\leq U(X_n) \end{aligned}$$

where the first inequality is true because $\nu(X_n)$ is optimal for player II in the one-day game $\mathscr{A}(U)(X_n)$ whose value is $(GU)(X_n)$, whereas the second inequality follows from Lemma 2.3. By Doob's Optional Sampling Theorem [Doob, 1953] we thus have that, for any stop rule t,

$$E_{x,\alpha,\nu^\infty(x)}u(X_t) \leq E_{x,\alpha,\nu^\infty(x)}U(X_t) \leq U(x) = V(x)$$

where the first inequality holds because $u \leq U$ while the last equality is true because of (3.5). Applying formula (2.7), we get

$$\begin{aligned} E_{x,\alpha,\nu^\infty(x)}u^* &= \inf_s \sup_{t\geq s} E_{x,\alpha,\nu^\infty(x)}u(X_t) \\ &\leq \sup_{t\geq 0} E_{x,\alpha,\nu^\infty(x)}u(X_t) \\ &\leq V(x). \end{aligned}$$

Since this is true for every $x \in S$ and for all strategies α of player I, we have proved that ν^∞ is optimal for player II in this case.

Case 2: $V \equiv 0$. In order to prove that ν^∞ is optimal for player II we will show that, for every $x \in S$ and for all strategies α of player I,

$$P_{x,\alpha,\nu^\infty(x)}[X_n = g \text{ infinitely often}] = 0.$$

First notice that, since $V(g) = 0$, then $(GU)(g) < 1$. In fact, set $V_0 = Tu = GU$. By way of contradiction, assume that $V_0(g) = 1$. Then $V_0 \geq u$, so that

$$T(u \wedge V_0) = Tu = V_0$$

and thus $V_0 \le V$ by Lemma 2.5. But this implies that $V(g) = 1$ contradicting the assumption that $V \equiv 0$.

Set

$$\delta = \frac{1 - (GU)(g)}{2} > 0.$$

Then, for all $\lambda \in \mathscr{P}(A)$,

$$\begin{aligned}(GU)(g) &\ge E_{g,\lambda,\nu(g)} U \\ &= \int_{\{U(X_1) < (GU)(g)+\delta\}} U(X_1)\, dP_{g,\lambda,\nu(g)} \\ &\quad + \int_{\{U(X_1) \ge (GU)(g)+\delta\}} U(X_1)\, dP_{g,\lambda,\nu(g)} \\ &\ge ((GU)(g) + \delta) P_{g,\lambda,\nu(g)}[U(X_1) \ge (GU)(g) + \delta].\end{aligned}$$

Therefore

$$\begin{aligned}P_{g,\lambda,\nu(g)}[U(X_1) < 1 - \delta] &= P_{g,\lambda,\nu(g)}[U(X_1) < (GU)(g) + \delta] \\ &\ge 1 - \frac{(GU)(g)}{(GU)(g) + \delta} \\ &\ge \frac{\delta}{1+\delta} > 0 \end{aligned} \tag{3.7}$$

for all $\lambda \in \mathscr{P}(A)$. This says that, if player II uses ν^∞, whenever the process $\{X_n\}$ reaches the goal g there is a strictly positive probability that the following state of the game will belong to set $\{x \in S : U(x) < 1 - \delta\}$.

Now fix $x \in S$ and a strategy α for player I; for the sake of simplicity we will write P for the probability $P_{x,\alpha,\nu^\infty(x)}$ and E for the expected value computed according to P. Notice that, as in Case 1, the sequence $\{U(X_n)\}$ is a bounded supermartingale with respect to P and therefore converges almost surely with respect to P.

Equation (3.7) and the almost sure convergence of the sequence $\{U(X_n)\}$ imply that $P[X_n = g \text{ infinitely often}] = 0$ since, if this was not the case, the supermartingale $\{U(X_n)\}$ would downcross the interval $[1-\delta, 1]$ infinitely often with positive probability and therefore would not converge on a set of histories of probability greater than zero.

To make the argument more precise let

$$D = \{h \in H : X_n(h) = g \text{ infinitely often}\}$$

and define a sequence of stopping times $\{t_n\}$ by setting

$$t_0 = \inf\{n \ge 0 : X_n = g\}$$

and, for any $n \geq 1$,

$$t_n = \inf\{n > t_{n-1} : X_n = g\}.$$

Then $D = \bigcap_{i=0}^{\infty}\{t_i < \infty\}$. For any $n \geq 1$, define the set

$$A_n = \{h \in H : U(X_n(h)) < U(X_{n-1}(h)) - \delta\}$$

If $h = (h_1, h_2, \ldots) \in D$,

$$\sum_{n=1}^{\infty} P[A_{n+1}|X_1, \ldots, X_n](h) = \infty$$

since, for every $j \geq 1$, $X_{t_j(h)}(h) = g$, $U(X_{t_j(h)}(h)) = 1$ and thus, by (3.7),

$$P[A_{t_j+1}|X_1, \ldots, X_{t_j}](h) = P_{g,\alpha_{t_j(h)}(h_1,\ldots,h_{t_j(h)}),\nu(g)}[U(X_1) < 1 - \delta] \geq \frac{\delta}{1+\delta} > 0.$$

Therefore

$$D \subseteq \{h \in H : \sum_{n=1}^{\infty} P[A_{n+1}|X_1, \ldots, X_n](h) = \infty\}$$

and

$$P[D] \leq P[\sum_{n=1}^{\infty} P[A_{n+1}|X_1, \ldots, X_n] = \infty] = P[A_n \text{ infinitely often}]$$

where the last equality holds because of the Extended Borel-Cantelli Lemma [Breiman, 1968]. However, if $h \in \{A_n \text{ infinitely often}\}$, then $\{U(X_n(h))\}$ does not converge. Hence

$$P[A_n \text{ infinitely often}] = 0$$

and $P[D] = 0$. $\diamond$

3.8 Remark: When u is the indicator of a finite subset W of the state space S, the result of the previous theorem is still true in the sense that an optimal stationary family of strategies is available to player II if $V(x) = 1$ for all $x \in W$ or if $V \equiv 0$. When the former hypothesis is true, $V \geq u$ so that the result becomes a special case of Theorem 4.1 which will be proved in the next section. If $V \equiv 0$, then $W = \{x \in S : V(x) < u(x)\}$ and thus, for games with finite state space, the result follows immediately from part (i) of Lemma 4.9 of the next section. If S is countably infinite, W is finite and $V \equiv 0$, it is still true that player II has an optimal stationary family of strategies. To prove this claim one may follow the argument which will be sketched in Remark 4.14 $\diamond$

The previous theorem settles affirmatively the question about the existence of an optimal stationary family of strategies for player II when $W = \{g\}$. The situation is different for player I. Obviously, if $V \equiv 0$, any strategy is optimal for player I. However Example 7.13.4 of Maitra and Sudderth [1996], a variation of an example of Kumar and Shiau [1981], shows that, even when S is finite, player I need not have an ε-optimal stationary strategy at a fixed initial state of the game if $V(g) = 1$.

When W has more than one element, there need not be an optimal strategy for player II even if the state space S is finite. This follows from Example 7.13.5 of Maitra and Sudderth [1996] which is another variation of the example of Kumar and Shiau. In the next section we will prove however that in nonleavable games with finite state space, for any $\varepsilon > 0$, player II has a stationary family of strategies which is ε-optimal. This result cannot be extended to games with countably infinite state space since, by modifying an example of Nowak and Raghavan [1991], Maitra and Sudderth [1996, Example 7.13.6] got a nonleavable game with countably infinite state space in which no ε-optimal stationary family of strategies exist for player II.

4. Stationary strategies for player II

Let us now return to the general nonleavable stochastic game $\mathcal{N}(u)$ where u is a real valued bounded function defined on a state space S, player I chooses a strategy α, player II chooses a strategy β and the payoff from II to I is $E_{x,\alpha,\beta}u^*$ when x is the initial state of the game.

The main property of the stationary family v^∞, which in the previous section we showed to be optimal for II when u is the indicator of an element $g \in S$, is to keep the value V of $\mathcal{N}(u)$ from increasing along the way. To make this idea more precise given the initial state $x \in S$ of $\mathcal{N}(u)$ and a real valued, bounded function ϕ defined on S, say that a strategy β of player II conserves ϕ at x if, for any strategy α of I, $E_{x,\alpha_0,\beta_0}\phi \le \phi(x)$ and, for any $n \ge 1$ and any partial history $p = ((x_1, a_1, b_1), (x_2, a_2, b_2), \ldots)$ of length n, $E_{X_n,\alpha[p],\beta[p]}\phi \le \phi(X_n)$. When u is the indicator of a $g \in S$, the stationary family v^∞ conserves the value V of $\mathcal{N}(u)$ at every $x \in S$. However this property alone was not sufficient to prove optimality of v^∞. In fact if $V \equiv 0$ any strategy of II conserves V; optimality of v^∞ followed in this case from the fact that, no matter what the strategy of player I is, the set $\{g\}$ is reached infinitely often on a set of histories of probability zero, as we proved in Case 2 of Theorem 3.6. Note that, when $V \equiv 0, \{g\}$ coincides with the set of states where $V < u$ while this latter set is empty if $V(g) = 1$.

For a general nonleavable game $\mathcal{N}(u)$ define a partition of the state space S by setting

$$W^0 = \{x \in S : V(x) < u(x)\} \text{ and } W^1 = \{x \in S : V(x) \ge u(x)\}.$$

When W^0 is empty, we will prove with Theorem 4.1 that v^∞ is still optimal for II. When W^0 is nonempty and the state space of $\mathcal{N}(u)$ is finite, for any $\varepsilon > 0$ we will construct a stationary family $\bar{v}$ for II such that, regardless of the strategy of I, the probability of reaching W_0 infinitely often is zero and $\bar{v}$ conserves at every $x \in S$ a function ϕ whose distance from V is less than $\varepsilon/2$. It will follow from Theorem 4.16 that $\bar{v}$ is an ε-optimal stationary family for II.

4.1 Theorem. *If $W^0 = \emptyset$, v^∞ is optimal for player II.*

Proof: When $W^0 = \emptyset$, $V(x) \geq u(x)$ for every $x \in S$. Since $GU = V_0 \geq V$, this implies that $GU \geq u$. But $G(GU) \leq GU$ since $GU \leq U$ by Lemma 2.3 and it is trivial to show that $Gu_1 \leq Gu_2$ if $u_1 \leq u_2$ are two bounded, real valued functions defined on S. Hence, by Lemma 2.3 again, $U \leq GU$ and thus $U = GU$. Notice that

$$T(u \wedge U) = Tu = GU = U$$

and thus, by Lemma 2.5, $U \leq V$. Since it is always true that $V \leq U$, this proves that $V = U$ when $W^0 = \emptyset$.

Now note that, for every $x \in S$ and for every $\gamma \in \mathscr{P}(A)$,

$$E_{x,\gamma,v(x)} U \leq (GU)(x) \leq U(x)$$

where the first inequality holds because $v(x)$ is optimal in the one-day game $\mathscr{A}(U)(x)$ whose value is $(GU)(x)$ and the second inequality is true because of Lemma 2.3. This is enough to prove that $\{U(X_n)\}$ is a bounded supermartingale with respect to $P_{x,\alpha,v^\infty(x)}$ for every $x \in S$ and for all strategies α of player I.

Now fix $x \in S$ and a strategy α for player I. By Doob's Optional Sampling Theorem [Doob, 1953] we have that, for any stop rule t,

$$E_{x,\alpha,v^\infty(x)} u(X_t) \leq E_{x,\alpha,v^\infty(x)} U(X_t) \leq U(x) = V(x)$$

where the first inequality follows from the fact that $u \leq U$. Therefore, applying formula (2.7), we get

$$\begin{aligned} E_{x,\alpha,v^\infty(x)} u^* &= \inf_s \sup_{t \geq s} E_{x,\alpha,v^\infty(x)} u(X_t) \\ &\leq \sup_{t \geq 0} E_{x,\alpha,v^\infty(x)} u(X_t) \\ &\leq V(x). \end{aligned}$$

Being this true for every $x \in S$ and for all strategies α of player I, we have proved that v^∞ is optimal for player II. $\diamond$

4.2 Remark: Assume that u is the indicator of a subset W of S. Then $W^0 = \emptyset$ if and only if $V(x) = 1$ for every $x \in W$; therefore the theorem proves that, when this last condition is satisfied, an optimal stationary family of strategies is available to player II. This covers, for example, Case 1 of the proof of Theorem 3.6. $\diamond$

Define a sequence of functions $\{V_n\}$ by setting $V_0 = Tu$ and, for any $n \geq 1$, $V_n = T(u \wedge V_{n-1})$. By the definition of the operator T, $V_0 = GU$ where U is the value of the leavable game $\mathscr{L}(u)$. For any $n \geq 1$, let U^n be the value of the leavable game $\mathscr{L}(u \wedge V_{n-1})$. Then, for any $n \geq 0$, $V_n = GU^n$ if we set $U^0 = U$.

For any given $x \in S$, $\{V_n(x)\}$ is a decreasing sequence of numbers bounded between $V(x)$ and $U(x)$. The same is true for the sequence $\{U^n(x)\}$ as the next lemma implies.

4.3 Lemma. *For all* $n \geq 1$,

$$u \wedge V_{n-1} \leq U^n \leq V_{n-1} = GU^{n-1} \leq U^{n-1}.$$

Therefore

$$U^n(x) = V_{n-1}(x) = (GU^{n-1})(x)$$

if $V_{n-1}(x) \leq u(x)$.

Proof: For all $n \geq 1$, $V_{n-1} = GU^{n-1} \leq U^{n-1}$, where the first equality is true by definition whereas the second follows from Lemma 2.3. Since $V_{n-1} \geq u \wedge V_{n-1}$ and $GV_{n-1} = G(GU^{n-1}) \leq GU^{n-1} = V_{n-1}$, from Lemma 2.3 again it follows that $U^n \leq V_{n-1}$. But U^n is the value of $\mathscr{L}(u \wedge V_{n-1})$ and thus $U^n \geq u \wedge V_{n-1}$.

Note that, if $x \in S$ and $V_{n-1}(x) \leq u(x)$, then $(u \wedge V_{n-1})(x) = V_{n-1}(x)$ and thus $U^n(x) = V_{n-1}(x)$. $\diamond$

In general the sequence $\{V_n\}$ does not converge to the value function V of $\mathscr{N}(u)$. However, when S is finite, Maitra and Sudderth [1996] proved that

$$V = \lim_{n \to \infty} V_n. \tag{4.4}$$

In the rest of the section we will assume that S is finite and that the set W^0 is nonempty.

For every $x \in W^0$, define $l = l(x)$ to be the first integer greater than or equal to zero such that $V_l(x) < u(x)$.

4.5 Lemma. *For every* $x \in W^0$,

$$(GU^l)(x) < U^l(x).$$

Proof: Let $x \in W^0$. If $l(x) = 0$, then $(GU)(x) = V_0(x) < u(x) \leq U(x)$. If $l(x) \geq 1$, then

$$V_l(x) < u(x) \leq V_{l-1}(x)$$

and thus

$$(GU^l)(x) = V_l(x) < (u \wedge V_{l-1})(x) \leq U^l(x)$$

where the last inequality follows form the fact that U^l is the value of $\mathscr{L}(u \wedge V_{l-1})$. $\diamond$

4.6 Remark: If u is the indicator of a set W, then $W^0 \subseteq W$. In this case, let us define

$$W_0 = \{x \in W^0 : (GU)(x) < U(x)\}$$

and, for any $n \geq 1$,

$$W_n = \{x \in W_0 - \bigcup_{j=1}^{n-1} W_j : (GU^n)(x) < U^n(x)\}.$$

Then it is sufficient to assume that W is finite for proving that, when W^0 is nonempty, there is a $k \geq 0$ such that $W_0, \dots, W_k$ are all nonempty and

$$W^0 = \bigcup_{j=1}^{k} W_j$$

even when S is countably infinite. $\diamond$

We are now ready to define the stationary family $\bar{\nu}$ which we will prove to be ε-optimal for player II when the state space S of the game is finite and W^0 is nonempty.

Let

$$k = \max\{l(x) : x \in W^0\}. \tag{4.7}$$

Fix $\varepsilon > 0$ and let m be an integer greater than k. Define $m+1$ quantities $\eta_0, \dots, \eta_m$ by setting

$$\eta_0 = \frac{1}{2}[\min\{U(x) - (GU^j)(x) : x \in W^0, j \in \{0, \dots, m\}, \\ (GU^j)(x) < U^j(x)\} \wedge \varepsilon]$$

and, for any $i \in \{1, \dots, m\}$.

$$\eta_i = \frac{1}{2}[\min\{U^i(x) - (GU^j)(x) : x \in W^0, j \in \{i, \dots, m\}, \\ (GU^j)(x) < U^j(x)\} \wedge \eta_{i-1}]$$

where we use the convention that $\min\{\emptyset\} = \infty$. Note that

$$0 < \eta_m < \eta_{m-1} < \dots < \eta_0 \leq \frac{\varepsilon}{2}.$$

Now define a real valued, bounded function ϕ by setting, for every $x \in S$,

$$\phi(x) = \min\{U^i(x) - \eta_i : i \in \{0, \dots, m\}\}.$$

If $x \in W^1$ let $r = r(x)$ be such that $\phi(x) = U^r(x) - \eta_r$, but if $x \in W^0$ set

$$r(x) = \max\{0 \le j \le m : (GU^j)(x) < U^j(x)\}.$$

Note that, for every $x \in W^0, r(x)$ is well defined and is greater than or equal to $l(x)$ because of Lemma 4.5. Note also that, if $x \in W^0$ and $r(x) < m$, then

$$\begin{aligned}(GU^m)(x) &= U^m(x) = (GU^{m-1})(x) = U^{m-1}(x)\\ &= \cdots = (GU^r)(x) < U^r(x) \le \cdots \le U(x). \qquad (4.8)\end{aligned}$$

In fact since $l \le r$, for every $i \in \{r+1, \dots, m\}$,

$$V_{i-1}(x) \le V_l(x) < u(x)$$

and thus

$$U^i(x) = V_{i-1}(x) = (GU^{i-1})(x) = U^{i-1}(x)$$

where the first equality follows from Lemma 4.3, the second one is true by the definition of V_{i-1} and the last equality holds because of the way we defined $r(x)$.

Finally, for every $x \in S$, let

$$\lambda(x) = v_r(x)$$

where, for any $0 \le j \le m, v_j$ is a function which maps every $x \in S$ to an element of $\mathscr{P}(B)$ which is optimal for player II in the one-day game $\mathscr{A}(U^j)(x)$. Note that the function λ depends on the quantity ε and on the integer $m \ge k$ which were both chosen before. Define $\bar{v} = \bar{v}_{(m,\varepsilon)} = \lambda^\infty$.

4.9 Lemma. *For all $x \in S$ and for all strategies α of player I,*

(i)

$$P_{x,\alpha,\bar{v}(x)}[X_n \in W^0 \text{ infinitely often}] = 0$$

and

(ii)

$$E_{x,\alpha,\bar{v}(x)}u^* \le U^m(x) + \frac{\varepsilon}{2}.$$

Proof: Let $\bar{v} = \bar{v}_{(m,\varepsilon)}$. In order to prove part (i) of the lemma we will show that, for all $x \in S$ and for all strategies α of player I, the sequence $\{\phi(X_n)\}$ is a bounded supermartingale with respect to $P_{x,\alpha,\bar{v}(x)}$ which does not converge on a set of probability greater than zero if the process $\{X_n\}$ visits the set W^0 infinitely often with positive probability.

Let $M = \sup_{x\in S}|u(x)|$.

Proving that $\{\phi(X_n)\}$ is a bounded supermartingale with respect to $P_{x,\alpha,\bar{v}(x)}$, for all $x \in S$ and for al strategies α of player I, is tantamount to showing that

$$E_{x,\gamma,\nu_r(x)}\phi \le \phi(x)$$

for all $x \in S$ and for all $\gamma \in \mathscr{P}(A)$.

If $x \in W^1$ then

$$E_{x,\gamma,\nu_r(x)}\phi \le E_{x,\gamma,\nu_r(x)} U^r - \eta_r \le (GU^r)(x) - \eta_r \le U^r(x) - \eta_r = \phi(x)$$

for all $\gamma \in \mathscr{P}(A)$, where the last equality follows by the way we defined $r = r(x)$.

Let now $x \in W^0$. Then

$$E_{x,\gamma,\nu_r(x)}\phi \le E_{x,\gamma,\nu_r(x)} U^r - \eta_r \le (GU^r)(x) - \eta_r \le \phi(x)$$

for all $\gamma \in \mathscr{P}(A)$ where the last inequality is true because, for every $i \in \{0, \dots, m\}$,

$$(GU^r)(x) - \eta_r \le U^i(x) - \eta_i. \tag{4.10}$$

In fact $(GU^r)(x) < U^r(x)$ so that, for $i \in \{0, \dots, r\}$, we obtain (4.10) by noticing that

$$\begin{aligned}\eta_i - \eta_r &\le \eta_i \\ &\le \min\{U^i(x) - (GU^j)(x) : x \in W^0, j \in \{i, \dots, m\}, (GU^j)(x) < U^j(x)\} \\ &\le U^i(x) - (GU^r)(x).\end{aligned}$$

On the other hand, if $r < m$ and $i \in \{r+1, \dots, m\}$, then equation (4.8) implies that $U^i(x) = (GU^r)(x)$ and thus (4.10) holds in this case because $\eta_r - \eta_i > 0$.

Hence $\{\phi(X_n)\}$ is a bounded supermartingale with respect to $P_{x,\alpha,\bar{\nu}(x)}$ and thus converges with $P_{x,\alpha,\bar{\nu}(x)}$ probability one, for all $x \in S$ and for all strategies α of player I.

Set now $\eta_{m+1} = 0$ and let

$$\delta = \frac{1}{2}\min\{\eta_j - \eta_{j+1} : j \in \{0, \dots, m\}\} > 0.$$

We want to show that, for all $x \in W^0$,

$$P_{x,\gamma,\nu_r(x)}[\phi(X_1) \le \phi(x) - \delta] \ge \frac{\delta}{2M + \delta} > 0 \tag{4.11}$$

for all $\gamma \in \mathscr{P}(A)$.

Let $x \in W^0$. Then

$$(GU^r)(x) - \eta_r + \delta \le \phi(x) - \delta \tag{4.12}$$

since, for every $i \in \{0, \dots, m\}$,

$$(GU^r)(x) - \eta_r + \delta \le U^i(x) - \eta_i - \delta. \tag{4.13}$$

In fact $(GU^r)(x) < U^r(x)$ so that, if $i \in \{0, \ldots, r\}$, by the definitions of δ and η_i,

$$\delta \le \frac{\eta_m}{2} \le \frac{\eta_i}{2} \le \frac{[U^i(x) - (GU^r)(x)] - [\eta_i - \eta_r]}{2}.$$

On the other hand, if $r < m$ and $i \in \{r+1, \ldots, m\}$, we know from equations (4.8) that $U^i(x) = (GU^r)(x)$ and thus (4.13) follows in this case from the fact that

$$\delta \le \frac{\eta_{i-1} - \eta_i}{2} \le \frac{\eta_r - \eta_i}{2}.$$

Now note that, since $-M \le u \le M$, then $-M \le V \le U \le M$ and thus

$$0 \le M + V \le M + V_r = M + (GU^r) \le M + U^r \le M + U \le 2M.$$

Therefore

$$\begin{aligned}(GU^r)(x) + M &\ge E_{x,\gamma,v_r(x)} U^r + M \\ &= \int_{\{U^r(X_1) \le (GU^r)(x)+\delta\}} (U^r(X_1) + M)\, dP_{x,\gamma,v_r(x)} \\ &\quad + \int_{\{U^r(X_1) > (GU^r)(x)+\delta\}} (U^r(X_1) + M)\, dP_{x,\gamma,v_r(x)} \\ &\ge ((GU^r)(x) + M + \delta) P_{x,\gamma,v_r(x)}[U^r(X_1) > (GU^r)(x) + \delta]\end{aligned}$$

for all $\gamma \in \mathscr{P}(A)$. Hence, by (4.12) and the definition of ϕ.

$$\begin{aligned}P_{x,\gamma,v_r(x)}[\phi(X_1) \le \phi(x) - \delta] &\ge P_{x,\gamma,v_r(x)}[\phi(X_1) \le (GU^r)(x) - \eta_r + \delta] \\ &\ge P_{x,\gamma,v_r(x)}[U^r(X_1) \le (GU^r)(x) + \delta] \\ &\ge 1 - \frac{(GU^r)(x) + M}{(GU^r)(x) + M + \delta} \\ &> \frac{\delta}{2M + \delta}\end{aligned}$$

for all $\gamma \in \mathscr{P}(A)$.

Having proved that (4.11) holds for all $x \in W^0$ and for all $\gamma \in \mathscr{P}(A)$ we will now show that, for every $x \in S$,

$$P_{x,\alpha,\bar{v}_r(x)}[X_n \in W^0 \text{ infinitely often}] = 0$$

for any strategy α of player I.

Fix an initial state $x \in S$ and a strategy α for player I and write P for $P_{x,\alpha,\bar{v}(x)}$.

Let

$$D = \{h \in H : X_n(h) \in W^0 \text{ infinitely often}\}$$

and define a sequence of stopping times $\{t_n\}$ by setting

$$t_0 = \inf\{n \geq 0 : X_n \in W^0\}$$

and, for any $n \geq 1$,

$$t_n = \inf\{n > t_{n-1} : X_n \in W^0\}.$$

Then $D = \bigcap_{i=0}^{\infty}\{t_i < \infty\}$. For every $n \geq 1$, define the set

$$A_n = \{h \in H : \phi(X_n(h)) \leq \phi(X_{n-1}(h)) - \delta\}.$$

If $h = (h_1, h_2, \ldots) \in D$.

$$\sum_{n=1}^{\infty} P[A_{n+1}|X_1, \ldots, X_n](h) = \infty$$

since, for every $j \geq l, y = X_{t_j(h)}(h) \in W^0$ and thus, by (4.11),

$$P[A_{t_j+1}|X_1, \ldots X_{t_j}](h) = P_{y,\alpha_{t_j(h)}(h_1\ldots h_{t_j(h)}),v_r(y)}[\phi(X_1) \leq \phi(y) - \delta]$$
$$\geq \frac{\delta}{2M + \delta} > 0.$$

Therefore

$$D \subseteq \left\{h \in H : \sum_{n=1}^{\infty} P[A_{n+1}|X_1, \ldots X_n](h) = \infty\right\}$$

and

$$P[D] \leq P\left[\sum_{n=1}^{\infty} P[A_{n+1}|X_1, \ldots, X_n] = \infty\right] = P[A_n \text{ infinitely often}]$$

where the last equality holds because of the Extended Borel-Cantelli Lemma [Breiman, 1968]. However, if $h \in \{A_n \text{ infinitely often}\}$, then $\{\phi(X_n(h))\}$ does not converge. Hence

$$P[A_n \text{ infinitely often}] = 0.$$

This proves that $P[D] = 0$ and concludes the proof of part (i) of the lemma.

Now define a function u_1 by setting, for every $x \in S$,

$$u_1(x) = u(x)I[x \in W^1] - MI[x \in W^0].$$

Note that $u_1 \leq \phi + \eta_0$. In fact, if $x \in W^1$,

$$u(x) \leq V(x) \leq U^m(x) \leq U^{m-1}(x) \leq \cdots \leq U(x)$$

and thus, for every $i \in \{0, \ldots, m\}$,

$$u_1(x) - \eta_0 = u(x) - \eta_0 \leq U^i(x) - \eta_0 \leq U^i(x) - \eta_i$$

which shows that $\phi(x) + \eta_0 \geq u_1(x)$. On the other hand, if $x \in W^0$, $\phi(x) \geq -M - \eta_0 = u_1(x) - \eta_0$.

Let $u_1^* = \limsup_{n\to\infty} u_1(X_n)$. Because of part (i), which has already been proved,

$$P_{x,\alpha,\bar{v}(x)}[u^* = u_1^*] = 1$$

for all $x \in S$ and for all strategies α of player I. Therefore, for all $x \in S$ and for all strategies α of player I, by applying formula (2.7), we get

$$\begin{aligned}
E_{x,\alpha,\bar{v}(x)}u^* &= E_{x,\alpha,\bar{v}(x)}u_1^* \\
&= \inf_s \sup_{t \geq s} E_{x,\alpha,\bar{v}(x)}u_1(X_t) \\
&\leq \sup_{t \geq 0} E_{x,\alpha,\bar{v}(x)}u_1(X_t) \\
&\leq \sup_{t \geq 0} E_{x,\alpha,\bar{v}(x)}\phi(X_t) + \eta_0 \\
&\leq \phi(x) + \eta_0 \\
&\leq U^m(x) + \eta_0 - \eta_m \\
&\leq U^m(x) + \frac{\varepsilon}{2}
\end{aligned}$$

where the third inequality holds because of Doob's Optional Sampling Theorem [Doob, 1953] and the last inequality is true since $\eta_0 < \varepsilon/2$. This proves part (ii) and concludes the proof of the lemma. ◇

4.14 Remark: When u is the indicator of a finite set W, and $V \equiv 0$ there is an optimal stationary family of strategies available to player II. In fact, in this case $W^0 = W$ so that the result follows directly from part (i) of the lemma above when S is finite. When S is countably infinite, one can partition the set W^0 in the way indicated in Remark 4.6 and then proceed with the same arguments we used above to prove that (i) still holds. ◇

4.15 Remark: If player II is a dummy with only one action, the game $\mathcal{N}(u)$ is a nonleavable gambling problem for player I and part (i) of the previous Lemma implies that, whatever the strategy I plays, the probability of visiting the set where $V < u$ infinitely often is zero. This corresponds to Theorem 3.7.1 of Dubins and Savage [1976] ◇

We are finally ready to prove the main result of the section.

4.16 Theorem. *Let S be finite. Then for every $\varepsilon > 0$ there is a stationary family of strategies which is ε-optimal for player II in the nonleavable game $\mathcal{N}(u)$.*

Proof. Let V be the value of $\mathcal{N}(u)$ and $S = W^0 \cup W^1$ where, as before,

$$W^0 = \{x \in S : V(x) < u(x)\} \quad \text{and} \quad W^1 = \{x \in W : V(x) \geq u(x)\}.$$

If $W^0 = \emptyset$, Theorem 4.1 states that there exists an optimal stationary family for player II. So let us assume that W^0 is nonempty. Fix k as in (4.7). Since S is finite, equation (4.4) implies that there is an $m \geq k$ such that

$$V_m \leq V + \frac{\varepsilon}{2}.$$

Let player II use the stationary family of strategies $\bar{v} = \bar{v}_{(m+1,\varepsilon)}$. Then by part (ii) of Lemma 4.9, for every $x \in S$,

$$E_{x,\alpha,\bar{v}(x)} u^* \leq U^{m+1}(x) + \frac{\varepsilon}{2} \leq V_m(x) + \frac{\varepsilon}{2} \leq V(x) + \varepsilon$$

for all strategies α of player II, where the next to the last inequality holds because of Lemma 4.3. This proves that $\bar{v}$ is ε-optimal for player II. ◇

4.17 Remark: The theorem implies the results derived in Everett [1957] and in Thuijsman and Vrieze [1992] for recursive games. ◇

5. A concluding remark

Example 7.13.6 of Maitra and Sudderth [1996] implies that ε-optimal stationary strategies need not exist for player II when the state space of the game is countably infinite. However one can show that in the nonleavable game considered in that example, for any given initial state, there is an ε-optimal stationary strategy available to player II. If this is true in general for all nonleavable games is still unknown.

Acknowledgement. This article is part of my Ph. D. thesis at the University of Minnesota. I want to thank my advisor William Sudderth for his precious guidance and for the stimulating discussions we had during the course of this research.

References

Breiman L (1968) Probability. Addison-Wesley, Reading, Massachusetts

Doob JL (1953) Stochastic processes. Wiley, New York

Dubins L, Savage L (1976) Inequalities for stochastic processes: How to gamble if you must. Dover, New York

Everett H (1957) Recursive games. In: Dresher M et al. (ed.) Contributions to the theory of games. III, volume 39 of Annals of Mathematical Studies, Princeton University Press, New Jersey, pp. 47–78

Kumar PR, Shiau TH (1981) Existence of value and randomized strategies in zero-sum discrete time stochastic dynamic games. Siam J. Control and Optimization 19(5):617–634

Maitra A, Sudderth W (1992) An operator solution of stochastic games. Israel Journal of Mathematics 78:33–49

——— (1996) Discrete gambling and stochastic games. Springer-Verlag, New York

Nowak AS, Raghavan TES (1991) Positive stochastic games and a theorem of Ornstein. In: Raghavan TES et al. (ed.) Stochastic games and related topics, Kluwer Academic Press. The Netherlands, pp. 127–134

Sudderth WD (1969) On measurable, nonleavable gambling houses with a goal. The Annals of Mathematical Statistics 40(1):66–70

——— (1971) On measurable gambling problems. Annals of Mathematical Statistics 42:260–269

Thuijsman F, Vrieze K (1992) Note on recursive games. In: Dutta B et al. (ed.) Game theory and economic applications, volume 389 of Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, pp. 133–145

von Neumann J, Morgenstern O (1947) Theory of games and economic behavior. Princeton University Press, New Jersey