International Institute for Applied Systems Analysis
IIASA
www.iiasa.ac.at

Persistently Good Strategies for Nonleavable Stochastic Games With Finite State Space

Secchi, P. and Sudderth, W.D.

IIASA Interim Report
November 1997

IIASA

International Institute for Applied Systems Analysis • A-2361 Laxenburg • Austria
Tel: +43 2236 807 • Fax: +43 2236 71313 • E-mail: info@iiasa.ac.at • Web: www.iiasa.ac.at

INTERIM REPORT IR-97-085 / November

Persistently Good Strategies for Nonleavable Stochastic Games With Finite State Space

Piercesare Secchi (secchi@univpv.it)
William D. Sudderth (bill@stat.umn.edu)

Approved by
Giovanni Dosi (dosi@iiasa.ac.at)
Leader, *TED Project*

About the Authors

Piercesare Secchi is from the Dipartimento di Economia Politica e Metodi Quantitativi, Università di Pavia, Via San Felice 5, 27100 Pavia, Italy. William D. Sudderth is from the School of Statistics, University of Minnesota, Minneapolis, MN 55455-0488, USA.

Acknowledgement

This paper was completed while the first author was visiting the International Institute for Applied System Analysis in Laxenburg, Austria.

The authors wish to thank Paolo Bertoletti for some stimulating discussions on subgame-perfect equilibria.

This research was supported by CNR (Italy) and the National Science Foundation Grant DMS–9703285.

Abstract

The notion of persistently optimal strategy in gambling theory is analogous to that of subgame-perfect equilibria in game theory. We prove the existence of persistently ϵ-optimal strategies for the players engaged in a nonleavable stochastic game with finite state space.

Contents

Persistently Good Strategies for Nonleavable Stochastic Games With Finite State Space

Piercesare Secchi
William D. Sudderth

1 Introduction and Preliminaries

The purpose of this paper is to prove the existence of persistently ϵ-optimal strategies for the players engaged in a nonleavable stochastic game. This is a two-person, multi-stage, zero-sum competition which Maitra and Sudderth (1992) defined by means of five elements: a state space S, two finite sets of actions A and B, for player I and II respectively, a law of motion q which maps $S \times A \times B$ into the set of probability distributions defined on S and a bounded utility function $u : S \to \Re$. The dynamics of a nonleavable stochastic game $\mathcal{N}(u)(x)$ with initial state $x \in S$ are the following; player I chooses an action $a_1 \in A$ while simultaneosly player II chooses an action $b_1 \in B$ and the game moves to the state X_1 according to the probability distribution $q(\cdot|x, a_1, b_1)$. This ends the first stage of the game; in the second stage the players choose actions $a_2 \in A$ and $b_2 \in B$ respectively and the game moves to the state X_2 according to the probability distribution $q(\cdot|X_1, a_2, b_2)$. By iterating this procedure for an inifinite number of stages an infinite sequence $\{X_n\}$ of states is produced; the payoff of the game from player II to player I is defined to be the expected value of

$$u^* = \limsup u(X_n).$$

Maitra and Sudderth (1992) proved the existence of ϵ-optimal strategies for the players engaged in a nonleavable game $\mathcal{N}(u)$ by showing that the game always has a value V in the sense that, for every initial state $x \in S$,

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

where α and β range over the set of strategies available to player I and II respectively. For a complete introduction to nonleavable stochastic games we refer to the book of Maitra and Sudderth (1996); in what follows we will use the same notation and terminology as in the book and in (Secchi,1997b) without any further comment.

Given $\epsilon > 0$ and the initial state $x \in S$ of the game $\mathcal{N}(u)$, a strategy α is said to be persistently ϵ-optimal for I at x if it is ϵ-optimal and, for every $n \geq 1$ and every partial history $p = ((x_1, a_1, b_1), \ldots, (x_n, a_n, b_n))$ of length n, the conditional strategy $\alpha[p]$ is ϵ-optimal for I for the game $\mathcal{N}(u)(x_n)$. A family of strategies $\bar{\alpha}$ is persistently ϵ-optimal if $\bar{\alpha}(x)$ is persistently ϵ-optimal for all $x \in S$. Analogous definitions hold for II. We note that, for all $x \in S, n \geq 1$ and all partial histories $p = ((x_1, a_1, b_1), \ldots, (x_n, a_n, b_n))$, the game $\mathcal{N}(u)(x_n)$ is a proper subgame of $\mathcal{N}(u)(x)$; therefore when a player uses a persistently ϵ-optimal strategy α he cannot gain more than ϵ by deviating from α in any subgame of $\mathcal{N}(u)(x)$. In this sense the definition of persistently ϵ-optimal strategy coincides with that of subgame-ϵ-perfect equilibria of Selten (1965). Our terminology stems from gambling theory and from the analogous definitions due in that setting to Dubins and Sudderth (1977).

The notion of persistently optimal strategy is related to that of stationarity. In fact, if $\bar{\alpha}$ is an ϵ-optimal stationary family of strategies, then $\bar{\alpha}$ is persistently ϵ-optimal since, for every $x \in S, n \geq 1$ and partial history $p = ((x_1, a_1, b_1), \ldots, (x_n, a_n, b_n))$, $\bar{\alpha}(x)[p] = \bar{\alpha}(x_n)$. In (Secchi, 1997b) it is shown that an ϵ-optimal stationary family is always available to player II when the state space S of the nonleavable game $\mathcal{N}(u)$ is finite. However Example 7.13.4 of Maitra and Sudderth (1996), a variation of an example of Kumar and Shiau (1981), shows that player I need not have an ϵ-optimal stationary strategy at a fixed initial state of the game, even if S is finite. Therefore the following theorem takes on interest; together with the results in (Secchi, 1997b) it settles affirmatively the question regarding the existence of persistently ϵ-optimal strategies for the players engaged in a nonleavable game with finite state space.

1.1 **Theorem**. *Let $\epsilon > 0$. For any nonleavable game $\mathcal{N}(u)$ defined on a finite state space S, a persistently 4ϵ-optimal family of strategies $\bar{\mu}$ is available to player I.*

When the state space of the game S is infinite and the utility u is unbounded, player I need not have a persistently ϵ-optimal strategy. This follows from the next example which appears in van der Wal (1981) and is a modification of an example given by Blackwell (1967). Incidentally the example shows that subgame-ϵ-perfect equilibria need not exist for all zero-sum games with a value.

1.2 **Example**. Let $S = \{\ldots, -2, -1, 0, 1, 2, \ldots\}, A = \{0, 1\}$, and $B = \{0\}$; player II is therefore a dummy and the game is really a one-player one. If $x < 0$, set $u(x) = 2^{-x}$ and $q(x|x, 0, 0) = q(x|x, 1, 0) = 1$; if $x > 0$, set $u(x) = 0$, $q(x + 1|x, 1, 0) = 1 - q(0|x, 1, 0) = b(x)/(2b(x + 1))$, with

$$b(x) = 1 + \frac{1}{x},$$

and $q(-x|x, 0, 0) = 1$. Finally define $u(0) = 0$ and $q(0|0, 0, 0) = q(0|0, 1, 0) = 1$. Therefore all nonpositive states are absorbing and the utility u is non null only on the strictly negative states.

For all strategies σ available to player I, let $\pi_0(\sigma)$ be the probability that action 1 is chosen at the first stage of the game. Given an initial state $x \in S$, we also define the partial histories $p_n = ((x+1, 1, 0), (x+2, 1, 0), \ldots, (x+n, 1, 0)), n \geq 1$. Let $\underline{0}$ be the only strategy available to player II. Then, for all $x > 0$ and all strategies σ of player I we may compute

$$\begin{aligned} E_{x,\sigma,\underline{0}} u^* \;=\; & (1 - \pi_0(\sigma))2^x + \pi_0(\sigma)(1 - \pi_0(\sigma[p_1]))\frac{b(x)}{2b(x+1)}2^{x+1} + \ldots \\ & \ldots + \pi_0(\sigma)[\Pi_{i=1}^{n-1}\pi_0(\sigma[p_i])](1 - \pi_0(\sigma[p_n]))\frac{b(x)}{2b(x+1)} \cdots \frac{b(x+n-1)}{2b(x+n)}2^{x+n} + \ldots. \end{aligned} \tag{1.3}$$

Hence

$$\sup_{\sigma} E_{x,\sigma,\underline{0}} u^* \leq \sup\{\frac{b(x)}{b(x+n)}2^x : n \geq 0\} = b(x)2^x.$$

Let σ^* be the strategy selecting action 1 with probability one for the first $n + 1$ stages of the game, $n \geq 1$, and then action 0 forever; then

$$E_{x,\sigma^*,\underline{0}}u^* = \frac{b(x)}{b(x+n)}2^x.$$

This shows that, for all $x > 0$,

$$\sup_\sigma E_{x,\sigma,\underline{0}} \geq \sup\{\frac{b(x)}{b(x+n)}2^x : n \geq 0\} = b(x)2^x.$$

Therefore the game has value $V(x) = b(x)2^x$. Obviously, $V(0) = 0$ and $V(x) = 2^{-x}$ for all $x < 0$.

Before proving that no persistently ϵ-optimal strategy is available to player I if the initial state of the game is big enough, we find a necessary condition for a strategy σ to be ϵ-optimal for I when the initial state of the game is strictly positive. Let $0 < \epsilon \leq 1$ and $x > 0$. If σ is ϵ-optimal for I at x, then

$$\begin{aligned} b(x)2^x - \epsilon &= V(x) - \epsilon \\ &\leq (1-\pi_0(\sigma))2^x + \pi_0(\sigma)\frac{b(x)}{2b(x+1)}E_{x,\sigma[p_1],\underline{0}}u^* \\ &\leq (1-\pi_0(\sigma))2^x + \pi_0(\sigma)\frac{b(x)}{2b(x+1)}V(x+1) \\ &= (1-\pi_0(\sigma))2^x + \pi_0(\sigma)b(x)2^x. \end{aligned}$$

Hence,

$$1 - \pi_0(\sigma) \leq \frac{2^{-x}}{b(x)-1}\epsilon \leq x2^{-x}.$$

Now let $\epsilon = 1, x \geq 5$ and assume there is a strategy μ persistently 1-optimal for I at x. Then for all $n \geq 1$ and all partial histories $p = ((x_1, a_1, 0), \ldots, (x_n, a_n, 0))$ with $x_i > 0$, and $a_i \in A, i = 1, \ldots, n$,

$$1 - \pi_0(\mu[p]) \leq x_n 2^{-x_n}.$$

This inequality and equation (1.3) imply that

$$\begin{aligned} E_{x,\mu,\underline{0}}u^* &\leq (1-\pi_0(\mu))2^x + \frac{b(x)}{2b(x+1)}(1-\pi_0(\mu[p_1])2^{x+1} + \ldots \\ &\quad \ldots + \frac{b(x)}{2b(x+1)} \cdots \frac{b(x+n-1)}{2b(x+n)}(1-\pi_0(\mu[p_n])2^{x+n} + \ldots \\ &\leq x + \frac{2}{3}(x+1) + \ldots + (\frac{2}{3})^n(x+n) + \ldots \\ &= 3x + 6 \end{aligned}$$

where we used the fact that $b(x+n)/(2b(x+n+1)) \leq 2/3$ for all $n \geq 0$, if $x \geq 1$. But $3x + 6 < V(x) - 1$, if $x \geq 5$, and thus μ is not persistently 1-optimal at x since it is not 1-optimal. ◇

When S is countably infinite and u is bounded, it is still an open question whether a player engaged in a nonleavable game $\mathcal{N}(u)$ has a persistently ϵ-optimal family of strategies.

2 Proof of Theorem 1.1

Before constructing a family of strategies persistently good for player I in the nonleavable game $\mathcal{N}(u)$, we begin by considering a leavable game $\mathcal{L}(u)$ defined on the finite state space S by means of the same law of motion q, utility u and sets of actions A and B for I and II respectively as $\mathcal{N}(u)$, but with the additional feature that player I is now allowed to stop the game according to a stop rule t of her choice and the payoff from II to I is the expected value of the utility $u(X_t)$ of the state reached by the game when it is stopped. Leavable games have a value and an optimal stationary family of strategies is always available to player II (Maitra and Sudderth, 1992); for all $\epsilon > 0$, the existence of an ϵ-optimal stationary family for player I was proved by Secchi (1997a).

Set $U_0 = u$ and, for all $n \geq 1$, define $U_n = u \vee GU_{n-1}$ where G is the one-day operator defined for any bounded, real-valued function ϕ on S and for every $x \in S$, by

$$(G\phi)(x) = \inf_{\nu} \sup_{\mu} E_{x,\nu,\mu}\phi, \tag{2.4}$$

with μ and ν ranging over $\mathcal{P}(A)$ and $\mathcal{P}(B)$, the sets of probability distributions defined on A and B respectively. For all $n \geq 1$, $U_n \geq U_{n-1}$ and the sequence $\{U_n\}$ converges to the value U of $\mathcal{L}(u)$ (Maitra and Sudderth, 1992); the convergence is uniform since S is finite.

Fix $\epsilon > 0$ sufficiently small and such that, for all $x \in S$, $|U(x) - u(x)| \leq \epsilon$ if and only if $U(x) = u(x)$. Let m be an integer such that $sup_{x \in S}|U_m(x) - U(x)| < \frac{\epsilon}{2}$, and let

$$0 = \epsilon_0 < \epsilon_1 < \ldots < \epsilon_m = \frac{\epsilon}{2}.$$

Define, for all $x \in S$,

$$\phi(x) = Max\{U_k(x) - \epsilon_k : 0 \leq k \leq m\}.$$

Note that $u \leq \phi \leq U$; when $\phi(x) = u(x)$

$$U(x) - \epsilon \leq U_m(x) - \frac{\epsilon}{2} \leq \phi(x) = u(x).$$

Therefore $\phi(x) = u(x)$ if and only if $u(x) = U(x)$.

We now introduce a function $\mu_\epsilon : S \to \mathcal{P}(A)$ generating a stationary family μ_ϵ^∞ which is 2ϵ-optimal for I in $\mathcal{L}(u)$. If $\phi(x) > u(x)$, let $k = k(x)$ be such that $\phi(x) = U_k(x) - \epsilon_k$ and set $\mu_\epsilon(x) \in \mathcal{P}(A)$ to be optimal for I in the one-day game $\mathcal{A}(U_{k-1})(x)$. If $\phi(x) = u(x)$ let $\mu_\epsilon(x)$ be optimal for I in the one-day-game $\mathcal{A}(U)(x)$.

For all $k \geq 0$, define

$$\tau(k) = \inf\{n \geq k : \phi(X_n) = u(X_n)\} = \inf\{n \geq k : U(X_n) = u(X_n)\}. \tag{2.5}$$

The next two inequalities (2.6) and (2.7) regarding the stopping time $\tau(0)$ were demonstrated in the course of the proof of Theorem 3.2 of Secchi (1997a) and will be very useful also in the present context. The first implies that, no matter what player II does, the game will reach the set of states where $U = u$ with probability one if player I uses the stationary family μ_ϵ^∞. In fact, set

$$\zeta = \frac{1}{2} \min\{\epsilon_k - \epsilon_{k-1} : 0 \leq k \leq m\}.$$

Then, for all $x \in S$ and $n \geq 1$,

$$P_{x,\mu_\epsilon^\infty(x),\beta}[\tau(0) \geq n] \leq \frac{2\,\|u\|}{n\zeta} \tag{2.6}$$

for all strategies β of II where $\|u\| = \max\{|u(x)| : x \in S\}$; therefore $\tau(0)$ is finite with probability one. Now choose n big enough so that $4\,\|u\|^2(n\zeta)^{-1} \leq \epsilon$. Then, for all $x \in S$,

$$E_{x,\mu_\epsilon^\infty(x),\beta}u(X_{\tau(0)\wedge n}) \geq U(x) - 2\epsilon \tag{2.7}$$

against any strategy β of player II; this immediately implies that μ_ϵ^∞ is 2ϵ-optimal for I in $\mathcal{L}(u)$.

When player I is not allowed to stop the game $\mathcal{L}(u)$ immediately, but must play at least for one day, the following result is of use.

2.8 **Lemma**. *For all $x \in S$ and every strategy β of II,*

$$P_{x,\mu_\epsilon^\infty(x),\beta}[\tau(1) < \infty] = 1 \tag{2.9}$$

and

$$E_{x,\mu_\epsilon^\infty(x),\beta}u(X_{\tau(1)}) \geq GU(x) - 2\epsilon. \tag{2.10}$$

Proof. Equation (2.9) follows from (2.6) and the fact that the conditional stopping time $\tau(1)[X_1]$ is equal to $\tau(0)$. In fact, for all $x \in S$,

$$P_{x,\mu_\epsilon^\infty(x),\beta}[\tau(1) < \infty] = \int P_{X_1,\mu_\epsilon^\infty(X_1),\beta[X_1]}[\tau(0) < \infty]dP_{x,\mu_\epsilon^\infty(x),\beta} = 1$$

for all β of II.

In order to prove (2.10), assume first that $U(x) > u(x)$. Then $\tau(1) = \tau(0)$ and, if $4\,\|u\|^2(n\zeta)^{-1} \leq \epsilon$,

$$E_{x,\mu_\epsilon^\infty(x),\beta}u(X_{\tau(1)\wedge n}) = E_{x,\mu_\epsilon^\infty(x),\beta}u(X_{\tau(0)\wedge n}) \geq U(x) - 2\epsilon = GU(x) - 2\epsilon$$

against any β of II, where the last equality follows from the fact that $U = u \vee GU$ (Maitra and Sudderth, 1992). Since u is bounded and $\tau(1)$ is finite with probability one, we get (2.10) in this case by letting $n \to \infty$. If $U(x) = u(x)$,

$$\begin{aligned}
&E_{x,\mu_\epsilon^\infty(x),\beta}u(X_{\tau(1)}) \\
&\quad = \int E_{X_1,\mu_\epsilon^\infty(X_1),\beta[X_1]}u(X_{\tau(0)})dP_{x,\mu_\epsilon^\infty(x),\beta} \\
&\quad \geq \int (U(X_1) - 2\epsilon)dP_{x,\mu_\epsilon^\infty(x),\beta} \\
&\quad \geq GU(x) - 2\epsilon.
\end{aligned}$$

for any β of II; the next to the last inequality follows from (2.7) by letting $n \to \infty$, whereas the last inequality is a consequence of the definition of μ_ϵ. ◇

We are now ready to return to the nonleavable game $\mathcal{N}(u)$ defined on the finite state space S. Lemma 4.1 of Secchi (1997a) shows that the value V of $\mathcal{N}(u)$ is also the value of the leavable game $\mathcal{L}(u \wedge V)$ which is defined on S and has the same law of motion q and action spaces A and B for I and

II respectively as the nonleavable game $\mathcal{N}(u)$, but utility $u \wedge V$. Applying the previous arguments to the game $\mathcal{L}(u \wedge V)$, set $\epsilon > 0$ small enough and such that, for all $x \in S$, $V(x) - (u \wedge V)(x) < \epsilon$ if and only if $V(x) = (u \wedge V)(x)$. Then, for all $k \geq 0$, the stopping time $\tau(k)$ defined by (2.5) becomes

$$\tau(k) = \inf\{n \geq k : V(X_n) \leq u(X_n)\}.$$

Lemma 2.8 implies that there is a stationary family μ_ϵ^∞ for player I such that, for every $x \in S$ and every strategy β of II,

$$P_{x,\mu_\epsilon^\infty(x),\beta}[\tau(1) < \infty] = 1 \tag{2.11}$$

and

$$E_{x,\mu_\epsilon^\infty(x),\beta}(u \wedge V)(X_{\tau(1)}) \geq GV(x) - 2\epsilon = V(x) - 2\epsilon, \tag{2.12}$$

where the last equality follows from the fact that $GV = V$ (Maitra and Sudderth, 1996).

In order to introduce the family of strategies $\bar{\mu}_\epsilon$ that will be proved to be persistently 4ϵ-optimal for I in $\mathcal{N}(u)$, define a sequence of stopping times τ_n by setting $\tau_0 = \underline{0}$ and, for all $n \geq 0$, $\tau_{n+1} = \tau(\tau_n + 1)$. The family $\bar{\mu}_\epsilon$ is constructed by sequential composition of the strategies $\{\mu^\infty_{\frac{\epsilon}{2^{n-1}}}\}$ along the sequence $\{\tau_n\}, n = 1, 2, \ldots$. That is, for all $x \in S$, $\bar{\mu}_\epsilon(x)$ follows $\mu_\epsilon^\infty(x)$ up to time τ_1, then uses $\mu^\infty_{\frac{\epsilon}{2}}(X_{\tau_1})$ up to time τ_2 and so on. We will begin by proving, with the next two lemmas, that $\bar{\mu}_\epsilon$ is 4ϵ-optimal for player I.

2.13 **Lemma**. *For every $x \in S$ and β of II,*

$$P_{x,\bar{\mu}_\epsilon(x),\beta}[\bigcap_{n=1}^{\infty}[\tau_n < \infty]] = 1.$$

Proof. Note that, for every $n \geq 0, x \in S$ and β of player II,

$$P_{x,\bar{\mu}_\epsilon(x),\beta}[\tau_{n+1} < \infty] = \int P_{X_{\tau_n},\mu^\infty_{\frac{\epsilon}{2^n}}(X_{\tau_n}),\beta[p_{\tau_n}]}[\tau(1) < \infty]\, dP_{x,\bar{\mu}_\epsilon(x),\beta}.$$

if $P_{x,\bar{\mu}_\epsilon(x),\beta}[\tau_n < \infty] = 1$. Now use (2.11) and induction on n. ◇

2.14 **Lemma**. *$\bar{\mu}_\epsilon$ is 4ϵ-optimal for I in $\mathcal{N}(u)$.*

Proof. Fix $x \in S$ and a strategy β for II. For ease of notation, in the following we use P for the probability $P_{x,\bar{\mu}_\epsilon(x),\beta}$ and E for the expected values computed with respect to P.

Define $Y_0 = V(x)$ and, for every $n \geq 1$, set

$$Y_n = V(X_{\tau_n}) + 2\epsilon \sum_{j=0}^{n-1} 2^{-j}.$$

The sequence $\{Y_n\}$ is a bounded submartingale with respect to P. In fact, for every $n \geq 0$, $|Y_n| < \|u\| + 4\epsilon$ and

$$
\begin{aligned}
&E[Y_{n+1}|X_1,\ldots,X_{\tau_n}]\\
&\quad = E[V(X_{\tau_{n+1}})|X_1,\ldots,X_{\tau_n}] + 2\epsilon\sum_{j=0}^{n} 2^{-j}\\
&\quad = E_{X_{\tau_n},\mu^{\infty}_{\frac{\epsilon}{2^n}}(X_{\tau_n}),\beta[p_{\tau_n}]} V(X_{\tau_{n+1}[p_{\tau_n}]}) + 2\epsilon\sum_{j=0}^{n} 2^{-j}\\
&\quad \geq E_{X_{\tau_n},\mu^{\infty}_{\frac{\epsilon}{2^n}}(X_{\tau_n}),\beta[p_{\tau_n}]} u\wedge V(X_{\tau(1)}) + 2\epsilon\sum_{j=0}^{n} 2^{-j}\\
&\quad \geq V(X_{\tau_n}) - 2\frac{\epsilon}{2^n} + 2\epsilon\sum_{j=0}^{n} 2^{-j}\\
&\quad = Y_n
\end{aligned}
$$

where the first inequality holds because $\tau_{n+1}[p_{\tau_n}] = \tau(1)$ and $V \geq u\wedge V$, whereas the second inequality follows from (2.12).

Being a bounded submartingale, $\{Y_n\}$ converges with P probability one and

$$E[\lim_{n\to\infty} Y_n] \geq Y_0 = V(x).$$

Therefore

$$E[\lim_{n\to\infty} V(X_{\tau_n})] \geq V(x) - 4\epsilon$$

This proves the lemma since

$$Eu^* = E[\limsup_{n\to\infty} u(X_n)] \geq E[\limsup_{n\to\infty} u(X_{\tau_n})] \geq E[\lim_{n\to\infty} V(X_{\tau_n})]$$

where the first inequality follows from Lemma 2.13 and the second from the definition of τ_n. $\diamond$

Final steps of the proof of Theorem 1.1. In order to prove that $\bar{\mu}_\epsilon$ is persistently 4ϵ-optimal, it is enough to show that, for every $x\in S$, $j\geq 1$ and partial history $p = ((x_1,a_1,b_1),\ldots,(x_j,a_j,b_j))$ of length j,

$$E_{x_j,\bar{\mu}_\epsilon(x)[p],\beta} u^* \geq V(x_j) - 4\epsilon$$

against any strategy β of II.

Let $k\geq 0$ be such that, for every history $h' = ((x_1',a_1',b_1'),(x_2',a_2',b_2'),\ldots)$,

$$\tau_k(ph') \leq j < \tau_{k+1}(ph')$$

with $ph' = ((x_1,a_1,b_1),\ldots,(x_j,a_j,b_j),(x_1',a_1',b_1'),(x_2',a_2',b_2'),\ldots)$. Now $\bar{\mu}_\epsilon(x)[p]$ is the sequential composition of the strategies $\{\mu^{\infty}_{\frac{\epsilon}{2^{k+n-1}}}\}$ along the sequence $\{\tau_{k+n}\}$, $n = 1,2,\ldots$, and, hence, equals $\bar{\mu}_{\frac{\epsilon}{2^k}}(x_j)$. The desired result now follows from Lemma 2.14. $\diamond$

References

Blackwell, D. (1967). Positive dynamic programming. In *Proceedings of the 5th Berkeley symposium on Mathematical Statistics and Probability*, Vol.1, 415–418.

Dubins, L.E. and Sudderth, W.D. (1977). Persistently ϵ-optimal strategies. *Mathematics of Operation Research*, 2, 125–134.

Kumar, P.R. and Shiau, T.H. (1981). Existence of value and randomized strategies in zero-sum discrete time stochastic dynamic games. *SIAM Journal of Control and Optimization*, 19, 617–634.

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

Maitra, A. and Sudderth, W.D. (1996). *Discrete gambling and stochastic games.* Springer-Verlag, New York.

Secchi, P. (1997a). Stationary strategies for recursive games. *Mathematics of Operation Research*, 22, 494–512.

Secchi, P. (1997b). On the existence of good stationary strategies for nonleavable stochastic games. Forthcoming in the *International Journal of Game Theory.*

Selten, R. (1965). Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragetragheit. *Zeitscrift fur die gesante Staatswissenschaft*, 12, 301–324.

van der Wal, J. (1981). *Stochastic dynamic programming.* The Mathematical Centre, Amsterdam.