JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 131, No. 1, pp. 115–134, October 2006 (© 2006)
DOI: 10.1007/s10957-006-9128-2

Approximation of Noncooperative Semi-Markov Games[1]

A. JAŚKIEWICZ[2] AND A. S. NOWAK[3]

Communicated by D. A. Carlson

Published Online: 29 November 2006

Abstract. An approximation of a general V-ergodic semi-Markov game with Borel state space by discrete-state space strongly-ergodic games is studied. The standard expected ratio-average criterion as well as the expected time-average criterion are considered. New theorems on the existence of ϵ-equilibria are given.

Key Words. Noncooperative semi-Markov games, Nash equilibria, approximation of dynamic games.

1. Introduction

This paper deals with nonzero-sum semi-Markov games with Borel state space satisfying a natural V-ergodicity assumption. Our model is a generalization of the discrete-time Markov game, studied by Altman and Nowak (Ref. 1), to continuous time, when the time between successive jumps from state to state of the underlying stochastic process is a random variable. Zero-sum games of this type were recently studied by Jaśkiewicz (Ref. 2) and Vega-Amaya (Ref. 3).

The existence of Nash equilibria in stochastic games with uncountable state space is not easy to prove, even in the discrete-time case. Some partial results are given mainly for discounted Markov games as well as Markov games with additive reward and transition structure (see Ref. 4). The most general case with transition probabilities satisfying some additivity condition was obtained by Nowak in Ref. 5. Much more complete theory was developed for correlated equilibria involving i.i.d. public signals. The first result in this area was given by Nowak and Raghavan (Ref. 6) for discounted Markov games. A considerable extension to semi-Markov

---

[1] The authors thank an anonymous referee for constructive comments. This work is supported by MEiN Grant 1P03A 01030.

[2] Assistant Professor, Institute of Mathematics and Computer Sciences, Wroclaw University of Technology, Wroclaw, Poland.

[3] Professor, Faculty of Mathematics, Computer Sciences, and Economics, University of Zielona Góra, Zielona Góra, Poland.

models of games with the expected average criteria was reported in Ref. 7. For a broad survey of the literature on nonzero-sum stochastic games, the reader is referred to Refs. 1,7,8 and the references therein. We point out that the V-ergodicity condition and related ones were used frequently in the theory of Markov control processes and discrete-time games (Refs. 9–12).

Nonzero-sum stochastic games with countable state spaces satisfying some stochastic stability or ergodicity conditions are much easier to study. Therefore, the theory of such games is more complete. The most general result in this area for Markov games was established by Altman, Hordijk, and Spieksma in Ref. 12. Nonzero-sum uniformly ergodic semi-Markov games were first studied in Ref. 13. The reader can find further information on research done in this direction in Ref. 8. The idea of approximating sequential (stochastic) games with Borel state space goes back to Whitt (Ref. 14), who considered uniformly continuous discounted models without any ergodicity assumptions. A related result on the existence of an ϵ-equilibrium point was given by Nowak (Ref. 15) under much weaker conditions. Recently, Altman and Nowak (Ref. 1) constructed an approximation of nonzero-sum V-ergodic and discounted Markov games with unbounded payoff functions. Applying the approximation method and the main result from Ref. 12, they were able to obtain ϵ-equilibria in a quite general framework.

In this paper, we generalize the main result from Ref. 1 to the class of semi-Markov games with Borel state space satisfying the V-ergodicity assumption. Our result on ϵ-equilibria applies to the two basic expected payoff criteria: the ratio-average and the time-average payoffs. On the other hand, our approximation technique is different from that of Ref. 1. The idea is to use an approximation of a general V-ergodic game with possibly unbounded payoffs by strongly ergodic games with bounded payoff functions that resemble stochastic games with countable state space. At the same time, we show that the distance between the original game and the approximating strongly-ergodic bounded games can be calculated using some primitive data and main results from Refs. 16–17. Such a result is not included in Ref. 1.

2. Model

Let Y be a Borel (subset of a complete separable metric) space. We denote by $\mathcal{B}(Y)$ the Borel σ-algebra on Y.

We consider an m-person nonzero-sum stochastic game model $\mathcal{M} := \{S, X, Q, r\}$ under the following scenario:

(i) S is a Borel state space.
(ii) X_i is a compact metric space of actions for player i. We put $X = X_1 \times X_2 \times \cdots \times X_m$.
(iii) $Q(\cdot|s, x)$ is a regular transition measure from $S \times X$ into $R_+ \times S$, where $R_+ = [0, \infty)$. It is assumed that $Q(B|s, x)$ is a Borel function on $S \times X$

for any Borel subset $B \subset R_+ \times S$ and that $Q(\cdot|s, x)$ is a probability measure on $R_+ \times S$ for any $s \in S$ and $x \in X$.

Let

$$Q(t, \tilde{S}|s, x) := Q([0, t] \times \tilde{S}|s, x),$$

for any $\tilde{S} \in \mathcal{B}(S)$. Next, define

$$q(\cdot|s, x) := Q(R_+, \cdot|s, x), \quad s \in S, x \in X.$$

If an $x \in X$ is selected in state s, then $Q(t, \tilde{S}|s, x)$ is the joint probability that the sojourn time is not greater than $t \in R_+$ and the next state s' belongs to $\tilde{S}$. Let $\tau(s, x)$ be the mean holding time in state s, i.e.,

$$\tau(s, x) = \int_0^{+\infty} t H(dt|s, x),$$

where $H(t|s, x) = Q(t, S|s, x)$ is the distribution function of the sojourn time of the process in state s when $x \in X$ is selected by the players.

(iv) $r_i(s, x)$ is a Borel measurable payoff function for player i. We assume that

$$r_i(s, x) = r_i^1(s, x) + r_i^2(s, x)\tau(s, x), \tag{1}$$

where $r_i^1(s, x)$ is the immediate reward at the transition time and $r_i^2(s, x)$ is the reward rate in the time interval interval between succesive transitions.

Strategies for the players are defined in the usual way. A strategy for a player is a Borel measurable mapping which associates with each given history a probability distribution on the set of actions available to him. A stationary strategy for player i is a mapping which associates with each state $s \in S$ a probability distribution on the available set of actions, independently of the history that led to the state s. Thus, a stationary strategy for player i can be identified with a Borel measurable transition probability f_i from S to X_i. For a detailed discussion of transition probability functions consult Ref. 18. Let F_i be the set of all stationary strategies of player i. Put

$$F = F_1 \times F_2 \times \cdots \times F_m.$$

Put $T_0 = 0$ and let $\{T_n\}$ denote a sequence of random decision epochs in the game, which proceeds as follows. If the initial state is s_0 and a vector of actions $x_{(0)} = (x_1^0, \ldots, x_m^0)$ is selected by the players, then the immediate payoff $r_i^1(s_0, x_{(0)})$ is incurred for player i and the game remains in state s_0 until the time $T = T_1 - T_0 = T_1$. The payoff $r_i^2(s_0, x_{(0)})$ to player i is incurred until the next transition occurs. Afterward, the system jumps to the next state s_1 according to the probability measure (transition law) $q(\cdot|s_0, x_{(0)}) = Q(R_+, \cdot|s_0, x_{(0)})$. The players choose some $x_{(1)} \in X$ and the game remains in state s_1 for a random time $T_2 - T_1$. The player i receives the payoff $r_i(s_1, x_{(1)})$ and the new state s_2

is generated according to the distribution $q(\cdot|s_1, x_{(1)})$. The situation repeats itself yielding a trajectory $(s_0, x_{(0)}, t_1, s_1, x_{(1)}, t_2, \ldots)$ of some stochastic process, where $s_n \in S$, $x_{(n)} \in X$, and t_{n+1} describe the state, the actions chosen by the players, and the decision epoch, respectively, on the nth stage of the game. Clearly, t_{n+1} is a realization of the random variable T_{n+1}. The distribution function of the random holding time $T_{n+1} - T_n$ is $H(\cdot|s_n, x_{(n)})$. Let $\mathcal{H}^\infty := (S \times X \times R_+)^\infty$ be the space of all infinite histories of the game, endowed with the product σ-algebra. Then, for any profile of strategies $\pi = (\pi_1, \ldots, \pi_m)$ of the players and every initial state $s_0 = s \in S$, a probability measure P_s^π is defined uniquely on $\mathcal{H}^\infty$ according to the Ionescu-Tulcea theorem (see Chapter 7 in Ref. 18 or Proposition V.1.1 in Ref. 19).

Let $N(t)$ be the number of jumps that have occured prior to time t, i.e.,

$$N(t) := \max\{n : T_n \leq t\}.$$

Under our assumptions, for each intitial state $s \in S$, any strategy profile π, and $t \geq 0$, we have $P_s^\pi(N(t) < \infty) = 1$ (see Ref. 20).

For each profile of strategies $\pi = (\pi_1, \ldots, \pi_m)$ and every initial state $s \in S$, there are two basic ways for defining the expected average payoffs:

(i) the time-average payoff to player i,

$$j_i(s, \pi) = \liminf_{t \to \infty} E_s^\pi \left(\sum_{n=0}^{N(t)} r_i\left(s_n, x_{(n)}\right) \right) / t; \text{ or}$$

(ii) the ratio-average payoff to player i:

$$J_i(s, \pi) = \liminf_{n \to \infty} E_s^\pi \left(\sum_{k=0}^{n-1} r_i\left(s_k, x_{(k)}\right) \right) \Big/ E_s^\pi \left(\sum_{k=0}^{n-1} \tau\left(s_k, x_{(k)}\right) \right).$$

Here, E_s^π means the expectation operator with respect to the probability measure P_s^π.

Remark 2.1. Some authors studying semi-Markov decision or game models (Refs. 7, 21) assume that the payoff functions are of the form

$$\bar{r}_i\left(s_n, x_{(n)}, t\right), \tag{2}$$

where $s_n \in S$, $x_{(n)} \in X$, and t is a realization of the random holding time $T_{n+1} - T_n$ in the state s_n. From the construction of the probability measure P_s^π and the properties of the conditional expectation, it follows that such a game model is equivalent (in terms of the expected average payoffs) to the semi-Markov game, where the payoff of any player i (corresponding to any $(s_n, x_{(n)})$) is

$$\hat{r}_i\left(s_n, x_{(n)}\right) = \int_0^\infty \bar{r}_i\left(s_n, x_{(n)}, t\right) H\left(dt|s_n, x_{(n)}\right). \tag{3}$$

Observe that (3) is of the form (1) with $r_i^1 = \bar{r}_i$ and $r_i^2 \equiv 0$. Therefore, the case (2) is seemingly more general.

Let $\pi^* = (\pi_1^*, \dots, \pi_m^*)$ be a fixed profile of strategies of the players. For any strategy π_i of player i, we write (π_{-i}^*, π_i) to denote the strategy profile obtained from π^* by replacing π_i^* with π_i.

Definition 2.1. Let $\epsilon \geq 0$. A strategy profile $\pi^* = (\pi_1^*, \dots, \pi_m^*)$ is called ϵ-equilibrium for the time-average (ratio-average) payoff semi-Markov game iff, for every player i and any policy π_i,

$$j_i(s, \pi^*) \geq j_i(s, (\pi_{-i}^*, \pi_i)) - \epsilon,$$
$$[J_i(s, \pi^*) \geq J_i(s, (\pi_{-i}^*, \pi_i)) - \epsilon],$$

for each $s \in S$. A 0-equilibrium is called a Nash equilibrium.

3. Assumptions

We make the following assumptions:

(A1) V-Geometric Ergodicity.

(A1) (i) There exist a Borel measurable function $V : S \mapsto [1, +\infty)$ and a Borel set $C \subset S$ such that, for some $\lambda \in (0, 1)$ and $\eta > 0$, we have

$$\int_S V(s')q(ds'|s, x) \leq \lambda V(s) + \eta 1_C(s),$$

for each $s \in S$ and $x \in X$.

(A1) (ii) The function V is bounded on C.

(A1) (iii) There exist $\xi \in (0, 1)$ and a probability measure ν concentrated on the Borel set C with the property that

$$q(D|s, x) \geq \xi \nu(D),$$

for each Borel set $D \subset C$, $s \in C$, and $x \in X$.

For any Borel measurable function $u : S \mapsto R$, we define the weighted norm as

$$\|u\|_V := \sup_{s \in S} |u(s)|/V(s).$$

We denote by L_V^∞ the Banach space of all Borel measurable functions u for which $\|u\|_V$ is finite.

Let $f_i \in F_i$. For any $s \in S$, the probability measure $f_i(s)$ is denoted also by $f_i(\cdot|s)$. For any Borel measurable function $w : S \times X \mapsto R$ and any $f = (f_1, \dots, f_m)$,

$$w(s, f) := w(s, f(s)) = \int_X \cdots \int_X w(s, x_1, \ldots, x_m) f_1(dx_1|s) \ldots f_m(dx_m|s).$$

Assumption (A1) is basic for this paper. It was used to study Markov control processes and discrete-time Markov games in many papers (see Refs. 1,9,22 and their references). Investigations of semi-Markov control processes and games based on this assumptions are contained in Refs. 2, 7, 16, 20. Inequality (A1) (i) is called the drift inequality and the set C satisfying (A1) (iii) is called the small set (see Refs. 17, 23). They imply that the state process $\{s_n\}$ governed by any $f \in F$ is a positive recurrent aperiodic Markov chain with the unique invariant probability measure, denoted by π_f; see Theorem 11.3.4 and page 116 in Ref. 23. Moreover, $\{s_n\}$ is V-uniformly ergodic [Theorem 2.3 in (Ref. 17)], i.e., there exist $\theta > 0$ and $\kappa \in (0, 1)$ such that

$$\left| \int_S u(s') q^n(ds'|s, f) - \int_S u(s') \pi_f(ds') \right| \leq V(s) \|u\|_V \theta \kappa^n, \tag{4}$$

for every $u \in L_V^\infty$ and $s \in S$, $n \geq 1$. Here, $q^n(\cdot|s, f)$ denotes the n-stage transition probability induced by q and f.

(A2) Basic Continuity Assumptions.

(A2) (i) For each $s \in S$ and $i = 1, \ldots, m$, $r_i(s, \cdot)$ is continuous on X; moreover, there exists a constant $L > 0$ such that $\|r_i\|_V \leq L$.

(A2) (ii) For each $s \in S$, $\tau(s, \cdot)$ is continuous on X and there exist positive constants b and B such that

$$b \leq \tau(s, x) \leq B, \quad s \in S, \quad x \in X.$$

(A2) (iii) There exist a probability measure μ on S and a density function ρ such that

$$q(D|s, x) = \int_D \rho(s', s, x) \mu(ds')$$

for each Borel set $D \subset S$, $x \in X$, and $s \in S$; moreover, for any sequence of joint action tuples $\{x^n\}$ converging to some x^0, it holds that

$$\lim_{n \to \infty} \int_S |\rho(s', s, x^n) - \rho(s', s, x^0)| V(s') \mu(ds') = 0.$$

In order to study the time-average payoff criterion, we shall need two additional assumptions.

(A3) Regularity Condition. There exist $\epsilon^* > 0$ and $\beta^* < 1$ such that

$$H(\epsilon^*|s, x) \leq \beta^*,$$

for all $s \in C$ and $x \in X$.

(A4) Uniform Integrability Condition.

$$\lim_{t\to\infty} \sup_{s\in C} \sup_{x\in X}[1 - H(t|s, x)] = 0.$$

It follows from (4) that, under (A1), for any $f \in F$, we have

$$J_i(s, f) =: J_i(f) = \int_S r_i(s', f)\pi_f(ds') \Big/ \int_S \tau(s', f)\pi_f(ds'). \tag{5}$$

We point out that, under Assumptions (A1)–(A4), for any $f \in F$, it holds that

$$j_i(s, f) = J_i(s, f). \tag{6}$$

For the details, consult Ref. 20. By (5) and (6), we have also

$$j_i(s, f) =: j_i(f) = \int_S r_i(s', f)\pi_f(ds') \Big/ \int_S \tau(s', f)\pi_f(ds'), \quad f \in F.$$

Definition 3.1. Any semi-Markov game satisfying Assumption (A1) is called V-crgodic. If (A1) (iii) holds with $C = S$, then the game is called strongly ergodic. A semi-Markov game is called bounded if all the payoff functions are bounded.

It should be noted that, if (A1) (iii) holds with $C = S$, then the remaining conditions in (A1) hold trivially with $V(s) = 1$ for all $s \in S$ and η sufficiently large. In such a very special case, the n-step transition probabilities $q^n(\cdot|s, f)$ induced by any $f \in F$ and q converge (geometrically fast) to π_f in the total variation norm; see Lemma 3.3 in Ref. 24. Assumption (A1), with C essentially smaller than the whole state space S, has much more applications, especially when the function V is unbounded on S; see Refs. 9, 16, 17, 23 and references therein. Bounded strongly ergodic semi-Markov games seem to be easier to solve. In Section 4, our aim is to show that they can be used to approximate a general V-ergodic game.

Remark 3.1. (a) One may think that the upper bound for the average time in Assumption (A2) (ii) is a limitation. The upper bound of $\tau(s, x)$ plays an essential role in Theorem 1 in Ref. 16. However, in that case, we may allow for an unbounded average sojourn time; that is, we may replace B by $BV(x)$ and Theorem 1 in Ref. 16 will be valid with appropriately modified constants. The next place where we shall use the upper bound is the proof of Lemma 5.1. Note, however, that we wish to approximate the unbounded game by games with bounded costs. Since Assumption (A2) (ii) holds, we do not have to approximate unbounded $\tau(s, x)$ by bounded ones. We believe that Assumption (A2) (ii) may be weakened in the sense given above, but the proofs will require additional calculations. On the other hand, we emphasize that, in many papers (e.g. Refs. 2, 3, 20), an upper bound on $\tau(s, x)$ is imposed.

(b) Our results crucially depend on the continuity assumption of the transition probability [i.e., (A2) (iii)]. Obviously, the norm topology is stronger than the topology of weak convergence of probability measures. However, this assumption allows a fairly wide class of transition probabilities. For example, it includes the important case in which the transition probability is induced by the normal distribution with different parameters. A version of this is the case in which the transition probability determines tomorrow's state as a function of the current state and actions almost deterministically, except for some white noise. In fact, one can work with a transition probability that induces distributions having density functions and where the convergence of the transition probability become synonymous with the uniform convergence of their density functions. On the other hand, we realize that our case does not handle a transition probability that induces, for instance, the degenerate distributions on the space of probability measures. The question of how to approximate stochastic games without assuming (A2) (iii) remains open.

(c) Assumption (A2) (iii) may be regarded as a restrictive one. To the best of our knowledge, all related results on Markov games (Refs. 1, 15) make use of the fact that the transition probability has a density function. In addition, we point out that the existence of stationary correlated equilibria in Markov games is not known without assuming that q is dominated by some measure; for a broader discussion, see Ref. 8.

4. Approximation by Bounded Strongly Erogodic Semi-Markov Games

Let us fix $s^* \in S$. We assume without loss of generality that $\mu(\{s^*\}) = 0$. Let

$$\varepsilon^* := (1 - \lambda)/\{2[V(s^*) - \lambda]\}. \tag{7}$$

For any $\varepsilon \in (0, \varepsilon^*)$, $s \in S$, and $x \in X$, we put

$$q_\varepsilon(\cdot|s, x) := (1 - \varepsilon)q(\cdot|s, x) + \varepsilon\delta_{s^*}(\cdot), \tag{8}$$

where δ_{s^*} is the probability measure concentrated at the fixed state $s^* \in S$. We define new payoff functions in the following way:

$$r_i^\varepsilon(s, x) := \begin{cases} r_i(s, x), & \text{if } |r_i(s, x)| \le 1/\varepsilon, \\ 1/\varepsilon, & \text{if } r_i(s, x) > 1/\varepsilon, \\ -1/\varepsilon, & \text{if } r_i(s, x) < -1/\varepsilon. \end{cases}$$

We denote by $\mathcal{M}^\varepsilon := \{S, X, q_\varepsilon, H, r^\varepsilon\}$ a new semi-Markov game model with the transition law q_ε, new payoff functions r_i^ε, and the same state space S, action space X, and time distribution H.

Note that, if the model $\mathcal{M}$ satisfies (A1), then $\mathcal{M}^\varepsilon$ also satisfies such a condition, but with ξ replaced by $(1 - \varepsilon)\xi$ and $(\lambda + 1)/2 < 1$ instead of λ. The set C and $\eta > 0$ remain unchanged. Again, from Refs. 17, 23, it follows that the state process $\{s_n\}$ governed by any $f \in F$ is a positive recurrent aperiodic Markov chain with the unique invariant probability measure denoted by π_f^ε. Moreover, there exist constants $\theta_1 > 0$ and $\kappa_1 \in (0, 1)$ independent of ε such that

$$\left| \int_S y(s') q_\varepsilon^n(ds'|s, f) - \int_S u(s') \pi_f^\varepsilon(ds') \right| \leq V(s) \|u\|_V \theta_1 \kappa_1^n, \tag{9}$$

for every $u \in L_V^\infty$ and $s \in S, n \geq 1$. Obviously, $q_\varepsilon^n(\cdot|s, f)$ denotes the n-stage transition probability induced by q_ε and $f \in F$. Consequently, for any initial state s and any strategy profile $\pi = (\pi_1, \ldots, \pi_m)$, the expected ratio-average payoff $J_i^\varepsilon(s, \pi)$ to player i in the game model $\mathcal{M}^\varepsilon$ is also well defined. Moreover, from (9), it follows that, for any profile of stationary strategies $f = (f_1, \ldots, f_m)$, the expected ratio-average payoff $J_i^\varepsilon(s, f)$ is independent of the initial state $s \in S$. From our discussion above, we can conclude also that

$$J_i^\varepsilon(s, f) = \int_S r_i^\varepsilon(s, f) \pi_f^\varepsilon(ds) \Big/ \int_S \tau(s, f) \pi_f^\varepsilon(ds).$$

Definition 4.1. The game with the transition probability of the type (8) is called strongly ergodic.

Note that, for a strongly ergodic game, we have $q_\varepsilon(\{s^*\}|s, x) \geq \varepsilon$, when $s \in S$ and $x \in X$.

The following result is a strengthened version of Theorem 1 in Ref. 16, where approximations by strongly ergodic but unbounded semi-Markov control models are given.

Theorem 4.1. Let (A1)–(A2) hold. Then, for each player i, we have

$$\lim_{\varepsilon \to 0} \sup_{f \in F} |J_i^\varepsilon(s, f) - J_i(s, f)| = 0.$$

Proof. Note that, by the triangle inequality, we have

$$\begin{aligned} \left| J_i^\varepsilon(s, f) - J_i(s, f) \right| \leq & \left| \int_S r_i^\varepsilon(s, f) \pi_f^\varepsilon(ds) \Big/ \int_S \tau(s, f) \pi_f^\varepsilon(ds) \right. \\ & \left. - \int_S r_i(s, f) \pi_f^\varepsilon(ds) \Big/ \int_S \tau(s, f) \pi_f^\varepsilon(ds) \right| \\ & + \left| \int_S r_i(s, f) \pi_f^\varepsilon(ds) \Big/ \int_S \tau(s, f) \pi_f^\varepsilon(ds) \right. \\ & \left. - \int_S r_i(s, f) \pi_f(ds) \Big/ \int_S \tau(s, f) \pi_f(ds) \right|. \end{aligned} \tag{10}$$

By Theorem 1 in Ref. 16, the second term tends to zero uniformly in f when $\varepsilon \to 0$. The first term can be treated in the following way. Define two sets:

$$A_1 = \{(s,x) : r_i(s,x) > 1/\varepsilon\},$$
$$A_2 = \{(s,x) : r_i(s,x) < -1/\varepsilon\}.$$

Making use of (A2) (ii), we get

$$\begin{aligned} c_\varepsilon :=& \left| \int_S r_i^\varepsilon(s,f)\pi_f^\varepsilon(ds) \Big/ \int_S \tau(s,f)\pi_f^\varepsilon(ds) \right. \\ &\left. - \int_S r_i(s,f)\pi_f^\varepsilon(ds) \Big/ \int_S \tau(s,f)\pi_f^\varepsilon(ds) \right| \\ \le& (1/b) \int_S |r_i^\varepsilon(s,f) - r_i(s,f)| \pi_f^\varepsilon(ds) \\ \le& (1/b) \Bigg[\underbrace{\int\int}_{A_1} |1/\varepsilon - r_i(s,x)| f(dx|s)\pi_f^\varepsilon(ds) \\ &+ \underbrace{\int\int}_{A_2} |-1/\varepsilon - r_i(s,x)| f(dx|s)\pi_f^\varepsilon(ds) \Bigg]. \end{aligned} \tag{11}$$

Furthermore, note that

$$|1/\varepsilon - r_i(s,x)| \le LV(s), \quad \forall (s,x) \in A_1,$$
$$|1/\varepsilon + r_i(s,x)| \le LV(s), \quad \forall (s,x) \in A_2,$$

and

$$A_1 \cup A_2 \subset \{s : LV(s) \ge 1/\varepsilon\} \times X. \tag{12}$$

We recall that the constant L is taken from (A2)(i). By (11) and (12), for any $f \in F$, we get

$$c_\varepsilon \le \int_S \int_X LV(s) f(dx|s)\pi_f^\varepsilon(ds) = \int_S R^\varepsilon(s)\pi_f^\varepsilon(ds),$$

where

$$R^\varepsilon(s) := \begin{cases} LV(s), & \text{if } LV(s) \ge 1/\varepsilon, \\ 0, & \text{if } LV(s) < 1/\varepsilon, \end{cases}$$

for each $i = 1, \ldots, m$. Therefore, to show that the first term in (10) tends to 0, as $\varepsilon \to 0$, it is enough to prove that

$$\limsup_{\varepsilon \to 0} \sup_{f \in F} \int_S R^\varepsilon(s) \pi_f^\varepsilon(ds) = 0. \tag{13}$$

Clearly, $R^\varepsilon \in L_V^\infty$. Consider an auxiliary Markov decision process with the reward function R^ε, transition law q_ε, state space S, and X as the set of all actions available to the single decision maker in any state $s \in S$. Under our assumptions, it follows from Theorem 10.3.6 in Ref. 9 that there exist a constant g^ε (which is the optimal expected average payoff) and a function $z^\varepsilon \in L_V^\infty$ (unique up to an additive constant) that satisfy the following optimality equation:

$$z^\varepsilon(s) + g^\varepsilon = R^\varepsilon(s) + \max_{x \in X} \int_S z^\varepsilon(s') q_\varepsilon(ds' | s, x). \tag{14}$$

From a standard measurable selection theorem (Refs. 9, 18), it follows that there exists a Borel measurable function $f^0 = (f_1^0, \ldots, f_m^0)$, $f_i^0 \in F_i$ that attains the maximum on the right-hand side of (14). Moreover, by the methods of dynamic programming (Ref. 9), we infer that

$$g^\varepsilon = \sup_{f \in F} \int_S R^\varepsilon(s') \pi_f^\varepsilon(ds') = \int_S R^\varepsilon(s') \pi_{f^0}^\varepsilon(ds'). \tag{15}$$

If we put

$$\overline{z}^\varepsilon(s) : E_s^{f^0} \left[\sum_{k=0}^{\infty} (R^\varepsilon(s_k) - g^\varepsilon) \right], \tag{16}$$

then using (15) and (16) one can obtain the following Poisson equation:

$$\overline{z}^\varepsilon(s) + g^\varepsilon = R^\varepsilon(s) + \int_S \overline{z}^\varepsilon(s') q_\varepsilon(ds' | s, f^0).$$

See Proposition 10.2.3 in Ref. 9. Note that, by (9) and (16), it holds that

$$\|\overline{z}^\varepsilon\|_V \le L\theta_1/(1 - \kappa_1). \tag{17}$$

Proceeding analogously as in Theorem 2 in Ref. 16 and subtracting $\overline{z}^\varepsilon(s)$ from $z^\varepsilon(s)$, we get

$$z^\varepsilon(s) - \overline{z}^\varepsilon(s) = \int_S (z^\varepsilon(s') - \overline{z}^\varepsilon(s'))\, q(ds' | s, f^0).$$

Iterating this equation n times and letting $n \to \infty$, we infer that

$$z^\varepsilon(s) - \overline{z}^\varepsilon(s) = \int_S (z^\varepsilon(s') - \overline{z}^\varepsilon(s'))\, \pi_{f^0}^\varepsilon\, (ds').$$

Denote the right-side of this equality by ω. By putting

$$z^{\varepsilon}(s) = \bar{z}^{\varepsilon}(s) + \omega, \quad \text{for all } s \in S,$$

in (14) and subtracting ω from both sides of this equation, we get

$$\bar{z}^{\varepsilon}(s) + q^{\varepsilon} = R^{\varepsilon}(s) + \max_{x \in X} \int_S \bar{z}^{\varepsilon}(s') q_{\varepsilon}(ds'|s, x). \tag{18}$$

Let $\{\varepsilon_n\}$ be a sequence converging to zero. Define

$$\bar{z}^n := \bar{z}^{\varepsilon_n}, \quad R^n := R^{\varepsilon_n}, \quad g^n := g^{\varepsilon_n}, \quad q_n := q_{\varepsilon_n}.$$

By (17), $\{\bar{z}^n\}$ is a uniformaly bounded sequence in L_V^{∞}. Now, (18) can be written as

$$\bar{z}^n(s) + g^n = R^n(s) + \max_{x \in X} \int_S \bar{z}^n(s') q_n(ds'|s, x). \tag{19}$$

Since $R^n(s) \to 0$ for $s \in S$, we conclude that $g^n \to g$ for $n \to \infty$. Let W be the space of all μ-equivalence classes of Borel measurable functions

$$z : S \mapsto R \quad \text{such that} \quad \|z\|_V \leq L\theta_1/(1 - \kappa_1), \quad \mu - \text{a.e.}$$

It is well-known that W is a compact and metrizable subset of $L^{\infty} = L^{\infty}(S, u)$ equipped with the relative weak star-topology $\sigma(L^{\infty}, L^1)$; see the Alaglu theorem in Ref. 25. Set

$$u^n(s) := \bar{z}^n(s)/V(s).$$

There is no loss of generality in assuming that $\{u^n\}$ converages in the weak-star topology to some $u \in W$. By Lemma 7.2.7 in Ref. 18, we may say that u is represented by a Borel measurable function, denoted for convenience by the same symbol. Define $z(s) = V(s)u(s)$. Proceeding as in the proof of the theorem in Ref. 22, we get

$$\begin{aligned} &\lim_{u \to \infty} \max_{x \in X} \left[-g^n + R^n(s) + \int_S \bar{z}^n(s') q_n(ds'|s, x) \right] \\ &= \max_{x \in X} \left[-g + \int_S z(s') q(ds'|s, x) \right], \end{aligned} \tag{20}$$

for every $s \in S$. From (19) and (20), we conclude that

$$z'(s) := \lim_{n \to \infty} \bar{z}^n(s)$$

exists for every $s \in S$. Since both z and z' are Borel measurable, the set

$$S_0 := \{s \in S : z(s) \neq z'(s)\}$$

is Borel. Moreover, we have $\mu(S_0) = 0$. Finally, we obtain some $f' \in F$ such that

$$z(s) + g = \max_{x \in X} \left[\int_S z(s')q(ds'|s, x) \right] = \int_S z(s')q(ds'|s, f'), \quad \forall s \notin S_0.$$

By integrating both sides with respect to $\pi_{f'}$ and noting that $\pi_{f'}(S_0) = 0$, we obtain $g = 0$, which implies (13). □

Remark 4.1. The game $\mathcal{M}^\varepsilon$ is obviously bounded. Since $q_\varepsilon(B|s, x) \geq \varepsilon\delta_{s^*}(B)$ for every $B \in \mathcal{B}(S)$, $s \in S$, $x \in X$, it is also strongly ergodic.

5. ϵ-Equlibrium Points for V-Ergodic Semi-Markov Games

The main objective in this paper is to prove the following results.

Theorem 5.1. Let (A1)–(A2) hold. Then, for any $\epsilon > 0$, the V-ergodic semi-Markov game $\mathcal{M}$ with the ratio-average criterion has a stationary ϵ-equlibrium.

Theorem 5.2. Let (A1)–(A4) hold. Then, for any $\epsilon > 0$, the V-ergodic semi-Markov game $\mathcal{M}$ with the time-average crierion has a stationary ϵ-equlibrium.

Before we give the proof, we describe preliminary material and an approximation technique for strongly ergodic games with Borel state space by games can be solved by the methods use in studying models with denumerable state space.

Let P_1 and P_2 be probability measures on $(S, \mathcal{B}(S))$. Then, the total uariation norm of $P_1 - P_2$ is defined as

$$\|P_1 - P_2\| := 2 \sup_{B \in \mathcal{B}s} |P_1(B) - P_2(B)|.$$

If P_1 and P_2 have densities p_1 and p_2 with respect to some σ-finite measure ν on S, then by the Scheffé theorem,

$$\|P_1 - P_2\| := \int_S |p_1(s) - p_2(s)|\nu(ds).$$

Let Δ be the set of all density functions in $L_1(S, \mu)$ and let $\mathcal{D}$ be the space of all continuous mappings from X into Δ with the metric d defined as

$$d(\phi_1, \phi_2) := \max_{x \in X} \int_S |\phi_1(x)(s) - \phi_2(x)(s)|\mu(ds).$$

Obviously, Δ is a nonempty and closed subset of $L_1(S, \mu)$. Therefore, by Theorem I.5.1 in Ref. 26, D is a complete separable metric space. Let N be the set of positive integers. We denote by $C(X)$ the Banach space of all continuous functions on X, endowed with the supremum norm $\|\cdot\|_{\sup}$.

Let $\varepsilon \in (0, \varepsilon^*)$, with ε^* defined by (7). Consider the game $\mathcal{M}^\varepsilon$. To simplify our notation, we put

$$p(\cdot|s, x) := q_\varepsilon(\cdot|s, x) \quad \text{and} \quad R_i(s, x) : r_i^\varepsilon(s, x),$$

for every player i, $s \in S$ and $x \in X$.

For each $s \in S$, the transition probability density ρ of the original game induces an element $\phi(s, \cdot)$ of $\mathcal{D}$ defined as

$$\phi(s, x) = \rho(s, \cdot, x).$$

From the product measurability of ρ on $S \times S \times X$, it follows that $s \mapsto \phi(s, \cdot)$ is a measurable mapping from S into D.

Let $\{\phi^k\}$ be a countable dense set in $\mathcal{D}$ and let $\{R^k\}$, $\{\tau^k\}$ be countable dense sets in $C(X)$. For any positive integers $k_1, k_2, \ldots, k_m, l_1, l_2$, put

$$A(k_1, \ldots, k_m, l_1, l_2) := \left\{ s \in S \backslash \{s^*\} : \sum_{i=1}^{m} \|R_i(s, \cdot) - R^{k_i}\|_{\sup} + d(\phi(s, \cdot), \phi^{l_1}) \right.$$
$$\left. + \|\tau(s, \cdot) - \tau^{l_2}\|_{\sup} < \delta \right\}.$$

Let α be a fixed one-to-one correspondence between N and N^{m+2}. Define

$$T^n := A(\alpha(n)), \quad n \in N.$$

Next, put

$$\tilde{Y}^1 := T^1 \quad \text{and} \quad \tilde{Y}^l := T^l - \bigcup_{j<l} \tilde{Y}^j,$$

for $l \geq 2$. Let $\{Y^n\}$ be the enumeration of all nonempty sets $\{\tilde{Y}^l\}$ where n ranges over some subset $\widetilde{N}$ of N. Put

$$Y^0 = \{s^*\} \quad \text{and} \quad N^0 = \widetilde{N} \cup \{0\}.$$

Clearly, $\{Y^n\}$, where $n \in N^0$, is a measurable partition of the state space S. Note that each set Y^n, $n \in \widetilde{N}$, corresponds to some $\rho^n \in \mathcal{D}$, $R_i^n \in C(X)$, and $\tau^n \in C(X)$. Define

$$R_i^\delta(s, x) = R_i^n(x), \quad \tau^\delta(s, x) = \tau^n(x),$$
$$p^\delta(B|s, x) = (1 - \varepsilon) \int_B \rho^n(s', x)\mu(ds') + \varepsilon\delta_{s^*}(B),$$

for all $s \in Y^n$, $x \in X$, and $n \in \widetilde{N}$. Further, put

$$R_i^\delta(s^*, x) = R_i(s^*, x), \quad \tau^\delta(s^*, x) = \tau(s^*, x), \quad p^\delta(B|s^*, x) = p(B|s^*, x),$$

for each $x \in X$. The game obtained above for any fixed δ will be denoted by $\mathcal{G}^\delta$. Observe that this is also a strongly ergodic game. From our construction, it follows that, for all $s \in Y^n$, $n \in N^0$, we have

$$d(\phi(s,\cdot),\rho^n) < \delta, \quad \|R_i(s,\cdot) - R_i^n\|_{\sup} < \delta, \quad \|\tau(s,\cdot) - \tau^n\|_{\sup} < \delta.$$

Hence, for $s \in Y^n$, $n \in N^0$, $x \in X$, we obtain

$$\sup_{s\in S}\sup_{x\in X}|R_i^\delta(s,x) - R_i(s,x)| \leq \delta \text{ and } \sup_{s\in S}\sup_{x\in X}|\tau^\delta(s,x) - \tau(s,x)| \leq \delta. \quad (21)$$

Moreover, it holds that

$$\sup_{s\in S}\sup_{x\in X}\|p(\cdot|s,x) - p^\delta(\cdot|s,x)\| \leq (1-\varepsilon)\delta. \quad (22)$$

Now we describe formally the meaning of the strong ergodicity of the games $\mathcal{M}^\varepsilon$ and $\mathcal{G}^\delta$. Since $p(\cdot|s,x) \geq \varepsilon\delta_{s^*}(\cdot)$ for all $s \in S$ and $x \in X$, by Lemma 3.3 in Ref. 24 we get

$$\sup_{s,s'\in S}\sup_{x,x'\in X}\|p(\cdot|s,x) - p(\cdot|s',x'),\| \leq 2\beta, \quad (23)$$

where $\beta = (2-\varepsilon)/2$. Using Lemma 3.3 in Ref. 24 again, we infer that the Markov chain induced by the transition probabity $p(\cdot|s,f)$, where $f \in Fs \in S$ is uniformly ergodic (not only V-ergodic as stated in Section 4); i.e., there exists an invariant probability measure π_f^ε such that

$$\sup_{s\in S}\sup_{f\in F}\|p^n(\cdot|s,f) - \pi_f^\varepsilon(\cdot)\| \leq 2\beta^n.$$

Note that the same reasoning also applies to $p^\delta(\cdot|s,f)$ and its invariant probability measure π_f^δ. By (22), (23), and Corollary 2 in Ref. 27, we obtain the following important inequality:

$$\|\pi_f^\delta - \pi_f^\varepsilon\| \leq \delta(1-\varepsilon)/(1-\beta) = 2\delta(1-\varepsilon)/\varepsilon. \quad (24)$$

Let $J_i^\delta(f)$ be the ratio-average criterion for the semi-Markov game $\mathcal{G}^\delta$.

Lemma 5.1. Let $\varepsilon \in (0,\varepsilon^*)$. For any $\gamma > 0$, there exist some $\delta > 0$ and a game $\mathcal{G}^\delta$ such that

$$|J_i^\varepsilon(f) - J_i^\delta(f)| \leq \gamma,$$

for each player i and for all $f \in F$.

Proof. Obviously, we have

$$|J_i^\delta(f) - J_i^\varepsilon(f)|$$
$$\leq \left| \int_S R_i^\delta(s,f)\pi_f^\delta(ds) \Big/ \int_S \tau^\delta(s,f)\pi_f^\delta(ds) \right.$$

$$
\begin{aligned}
&- \int_S R_i(s, f)\pi_f^\varepsilon(ds) \Big/ \int_S \tau(s, f)\pi_f^\varepsilon(ds)\Big| \\
&= \Big|\Big[\int_S R_i^\delta(s, f)\pi_f^\delta(ds) \int_S \tau(s, f)\pi_f^\varepsilon(ds) \\
&- \int_S R_i(s, f)\pi_f^\varepsilon(ds) \int_S \tau^\delta(s, f)\pi_f^\delta(ds)\Big] \\
&\Big/\Big[\int_S \tau(s, f)\pi_f^\varepsilon(ds) \int_S \tau^\delta(s, f)\pi_f^\delta(ds)\Big]\Big|. \qquad (25)
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
&\Big| \int_S R_i^\delta(s, f)\pi_f^\delta(ds) \int_S \tau(s, f)\pi_f^\varepsilon(ds) - \int_S R_i(s, f)\pi_f^\varepsilon(ds) \int_S \tau^\delta(s, f)\pi_f^\delta(ds)\Big| \\
&\leq \Big| \int_S R_i^\delta(s, f)\pi_f^\delta(ds) \Big[\int_S \tau(s, f)\pi_f^\varepsilon(ds) - \int_S \tau^\delta(s, f)\pi_f^\delta(ds)\Big]\Big| \\
&\quad + \Big| \int_S \tau^\delta(s, f)\pi_f^\delta(ds) \Big[\int_S R_i^\delta(s, f)\pi_f^\delta(ds) - \int_S R_i(s, f)\pi_f^\varepsilon(ds)\Big]\Big|.
\end{aligned}
$$

Next, using (21) and (24), we obtain

$$
\begin{aligned}
&\Big| \int_S \tau(s, f)\pi_f^\varepsilon(ds) - \int_S \tau^\delta(s, f)\pi_f^\delta(ds)\Big| \\
&\leq \Big| \int_S \tau(s, f)\pi_f^\varepsilon(ds) - \int_S \tau^\delta(s, f)\pi_f^\varepsilon(ds)\Big| | \\
&\quad + \Big| \int_S \tau^\delta(s, f)\pi_f^\varepsilon(ds) - \int_S \tau^\delta(s, f)\pi_f^\delta(ds)\Big| \\
&\leq \delta + (B + \delta)\|\pi_f^\varepsilon - \pi_f^\delta\| = \delta + (B + \delta)2\delta(1 - \varepsilon)/\varepsilon.
\end{aligned}
$$

Similarly, taking into account that $|R(s, f)| \leq 1/\varepsilon$, we infer that

$$
\Big| \int_S R(s, f)\pi_f^\varepsilon(ds) - \int_S R^\delta(s, f)\pi_f^\delta(ds)\Big| \leq \delta + (1/\varepsilon + \delta)2\delta(1 - \varepsilon)/\varepsilon.
$$

Therefore, if $\delta < b$, from the above two upper bounds and (25), we obtain the following inequality:

$$
\begin{aligned}
&|J_i^\delta(f) - J_i^\varepsilon(f)| \\
&\leq [\delta(1/\varepsilon + B + 2\delta) + (1/\varepsilon + \delta)(B + \delta)4\delta(1 - \varepsilon)/\varepsilon]/[b(b - \delta)],
\end{aligned}
$$

which implies the assertion. □

Let F_i^0 be the class of strategies of player i that are piecewise constant: f_i belongs to F_i^0 iff $s \mapsto f_i(\cdot|s)$ is constant on each set Y^n of the partition $\{Y^n\}$ of $S, n \in N^0$. Put $F^0 := F_1^0 \times \ldots \times F_m^0$.

Tha game $\mathcal{G}^\delta$ with piecewise constant stationary strategies for the players resembles a semi-Markov game with countable state space N^0. The transition probabilities $P_{ij}(x)$ (of moving from a state i to a state j, when a vector of actions $x = (x_1, \ldots, x_m) \in X$ is selected) are defined as follows:

$$P_{ij}(x) := (1-\varepsilon)\int_{Y^j} \rho^i(s', x)\mu(ds') + \varepsilon\delta_{s^*}(Y^j), \quad i \in \widetilde{N}, j \in N^\circ,$$

$$P_{0j}(x) := (1-\varepsilon)\int_{Y^j} \rho^i(s^*, s', x)\mu(ds') + \varepsilon\delta_{s^*}(Y^j), \quad j \in N^0.$$

From Theorem 1 in Ref. 13, it follows immediately that the countable state space game defined above has a stationary Nash equilibrium point. Clearly, this implies that the game $\mathcal{G}^\delta$ has a Nash equilibrium $f^0 = (f_1^0, \ldots, f_m^0)$ in the class of F^0 of piecewise constant stationary strategies. The arguments are based on the average cost optimality equations and are precisely given for Markov games in the Appendix of the Nowak and Altman paper (Ref. 1). An extension to semi-Markov games is obvious. One can use the optimality equation derived in Ref. 16.

Lemma 5.2. The game $\mathcal{G}^\delta$ described above has a stationary Nash equilibrium $f^0 = (f_1^0, \ldots, f_m^0) \in F^0$ within the class of all strategies of the players.

We are now ready to prove our main results.

Proof of Theorem 5.1. Fix $\epsilon > 0$. By Theorem 4.1, there exists some $\varepsilon \in (0, \varepsilon^*)$ such that

$$\sup_{f \in F} |J_i(f) - J_i^\varepsilon(f)| < \epsilon/4, \tag{26}$$

for each player i. By Lemma 5.1, there exists some δ and a game $\mathcal{G}^\delta$ such that, we have

$$\sup_{f \in F} |J_i^\varepsilon(f) - J_i^\varepsilon(f)| < \epsilon/4, \tag{27}$$

for each player i. From Lemma 5.2, we know that $\mathcal{G}^\delta$ has a stationary Nash equilibrium $f^0 \in F^0 \subset F$. This, (26), and (27) imply that f^0 is an ϵ-equilibrium for the game model $\mathcal{M}$. □

Proof of Theorem 5.2. Let $(f_1^0, \ldots, f_m^0) \in F^0$ be a Nash equilibrium in the game $\mathcal{M}$ with the ratio-average criterion. From the main result in in (Ref. 20), we know that

$$J_i(f^0) = j_i(f^0),$$

for each player i. Let Π_i be the set of all strategies of player i. From Ref. 20, we know also that

$$\begin{aligned} j_i(f^0) &= \sup_{\pi_i \in \Pi_i} j_i(s, (f^0_{-i}, \pi_i)) \\ &= \sup_{\pi_i \in \Pi_i} J_i(s, (f^0_{-i}, \pi_i)) \\ &= J_i(f^0), \quad s \in S. \end{aligned}$$

Hence, we conclude Theorem 5.2 from Theorem 5.1. □

6. Concluding Remarks

In this paper, we show how to approximate a general unbounded V-ergodic game by a sequence of bounded strongly ergodic games which are easier to solve. The reason is that, from the proofs of Theorem 4.1 (Refs. 16–17) and Lemma 5.1, one can calculate the distance between the original game and the approximating one (expressed in terms of the average payoffs corresponding to stationary strategies). It should be noted that the evaluation of the distance between the invariant probability distributions given in (24) is of a very simple form. There is no such result in Ref. 1. Lemma 3.4 in Ref. 1 and its proof are much more complicated and the evaluation of similar difference of the invariant probabilities is rather theoretical. We point out that, in Ref. 1, the authors approximate the original game (which is unbounded) by unbounded games with countably many states. Next, they apply (in a quite involved manner) the main result from Ref. 12 (see pp. 1830–1831 in Ref. 1). In this paper, we omit such difficulties because our approximating games have a much simpler structure.

References

1. NOWAK, A. S., and ALTMAN, *ε-Equilibria for Stochastic Games with Uncountable State Space and Unbounded Cost*, SIAM Journal on Control and Optimization, Vol. 40, pp. 1821-1839, 2002.
2. JAŚKIEWICZ, A., *Zero-Sum Semi-Markov Games*, SIAM Journal on Control and Optimization, Vol. 41, pp. 723-739, 2002.
3. VEGA-AMAYA, O., *Zero-Sum Average Semi-Markov Games: Fixed-Point Solutions of the Shapley Equation*, SIAM Journal on Control and Optimization, Vol. 42, pp. 1876-1894, 2003.
4. KÜENLE, H. U., *Equilibrium Strategies in Stochastic Games with Additive Cost and Transition Structure*, International Game Theory Review, Vol. 1, pp. 131–147, 1999.
5. NOWAK, A. S., *On a New Class of Nonzero-Sum Discounted Stochastic Games Having Stationary Nash Equilibrium Points*, International Journal of Game Theory, Vol. 32, pp. 121–132, 2003.

6. NOWAK, A. S., and RAGHAVAN, T. E. S., *Existence of Stationary Correlated Equilibria with Symmetric Information for Discounted Stochastic Games*, Mathematics of Operations Research, Vol. 17, pp. 519–526, 1992.
7. NOWAK, A. S., and JAŚKIEWICZ, A., *Nonzero-Sum Semi-Markov Games with the Expected Average Payoffs*, Mathematical Methods of Operations Research. Vol. 62, 2005.
8. NOWAK, A. S., *N-Person Stochastic Games: Extensions of the Finite State Space Case and Correlation*, Stochastic Games and Applications, Edited by A. Neyman and S. Sorin, NATO Science Series, Series C: Mathematical and Physical Sciences, Vol. 570, pp. 93–106, 2003.
9. HERNÁNDEZ-LERMA, O., and LASSERRE, J. B., *Further Topics on Discrete-Time Markov Control Processes*, Springer Verlag, New York, NY, 1999.
10. HERNÁNDEZ-LERMA, O., and LASSERRE, J. B., *Zero-Sum Stochastic Games in Borel Spaces: Average Payoff Criteria*. SIAM Journal on Control and Optimization, Vol. 39, pp. 1520–1539, 2001.
11. KÜENLE, H. U., and SCHURATH, R., *The Optimality Equation and ε-Optimal Strategies in Markov Games with Average Reward Criterion*, Mathematical Methods of Operations Research, Vol. 56, pp. 439–449, 2003.
12. ALTMAN, E., HORDIJK, A., and SPIEKSMA F. M., *Contraction Conditions for Average and α-Discount Optimality in Countable State Markov Games with Unbounded Rewards*, Mathematics of Operations Research, Vol. 22, pp. 588–618, 1997.
13. POŁOWCZUK, W., *Nonzero-Sum Semi-Markov Games with Countable State Spaces*, Applicationes Mathematicae, Vol. 27, pp. 395–402, 2000.
14. WHITT, W., *Representation and Approximation of Noncooperative Sequential Games*, SIAM Journal on Control and Optimization, Vol. 18, pp. 33–48, 1980.
15. NOWAK, A. S., *Existence of Equilibrium Stationary Strategies in Discounted Noncooperative Stochastic Games with Uncountable State Space*, Journal of Optimization Theory and Applications. Vol. 45, pp. 591–602, 1985.
16. JAŚKIEWICZ, A., *An Approximation Approach to Ergodic Semi-Markov Control Processes*, Mathematical Methods of Operations Research, Vol. 54, pp. 1–19, 2001.
17. MEYN, S. P., and TWEEDIE, R. L., *Computable Bounds for Geometric Convergence Rates of Markov Chains*, Annals of Applied Probability, Vol. 4, pp. 981–1011, 1994.
18. BERTSEKAS, D. P., and SHREVE, S. E., *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, NY, 1978.
19. NEVEU, J., *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco, California, 1965.
20. JAŚKIEWICZ, A., *On the Equivalence of Two Expected Average Cost Criteria for Semi-Markov Control Processes*, Mathematics of Operations Research, Vol. 29, pp. 326–338, 2004.
21. FEINBERG, E. A., *Constrained Semi-Markov Decision Processes with Average Rewards*, Mathematical Methods of Operations Research, Vol. 39, pp. 257–288, 1994.
22. JAŚKIEWICZ, A., *On Strong 1-Optimal Policies in Markov Control Processes with Borel State Spaces*, Bulletin of the Polish Academy of Sciences, Mathematics Series, Vol. 48, pp. 439–450, 2000.

23. MEYN, S. P., and TWEEDIE, R. L., *Markov Chains and Stochastic Stability*, Springer Verlag, New York, NY, 1993.
24. HERNÁNDEZ-LERMA, O., *Adaptive Markov Control Processes*, Springer Verlag, New York, NY, 1989.
25. DUNFORD, N., and SCHWARTZ, J. T., *Linear Operators, Part 1: General Theory*, Wiley Interscience Publishers, New York, NY, 1958.
26. WARGA, J., *Optimal Control of Differential and Functional Equations*, Academic Press, New York, NY, 1977.
27. NOWAK, A. S., *A Generalization of Ueno's Inequality for n-Step Transition Probabilities*, Applicationes Mathematice, Vol. 25, pp. 295–299, 1998.