11. R. Cavazos-Cadena and **E. Fernández-Gaucherand**, "Value Iteration in a Class of Average Controlled Markov Chains with Unbounded Costs: Necessary and Sufficient Conditions for Pointwise Convergence," *Journal of Applied Probability* **33** (1996) 986-1002.

J. Appl. Prob. 33, 986–1002 (1996)
Printed in Israel

VALUE ITERATION IN A CLASS OF AVERAGE CONTROLLED MARKOV CHAINS WITH UNBOUNDED COSTS: NECESSARY AND SUFFICIENT CONDITIONS FOR POINTWISE CONVERGENCE

ROLANDO CAVAZOS-CADENA,* *Universidad Autónoma Agraria Antonio Narro*
EMMANUEL FERNÁNDEZ-GAUCHERAND,** *The University of Arizona*

Abstract

This work concerns controlled Markov chains with denumerable state space, (possibly) *unbounded* cost function, and an expected *average* cost criterion. Under a *Lyapunov function condition*, together with mild continuity-compactness assumptions, a simple *necessary and sufficient* criterion is given so that the relative value functions and differential costs produced by the *value iteration* scheme converge pointwise to the solution of the optimality equation; this criterion is applied to obtain convergence results when the cost function is bounded below or bounded above.

CONTROLLED MARKOV CHAINS; AVERAGE COST CRITERION; LYAPUNOV FUNCTION CONDITION; VALUE ITERATION SCHEME; POINTWISE CONVERGENCE; NECESSARY AND SUFFICIENT CONDITIONS

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 93E20; 90C40

1. Introduction

This work considers controlled Markov chains (CMC) with denumerable state space and an average cost criterion. The cost function is (possibly) *unbounded* and, besides standard continuity-compactness conditions, the main assumption on the structure of the model is that the *Lyapunov function condition* (LFC) — introduced by Hordijk in [14] — holds true. The LFC implies the existence of a (generally *unbounded*) solution of the *average cost optimality equation* (ACOE), yielding optimal stationary policies. In this context, the main contribution of this paper is to formulate a simple criterion so

Received 23 January 1995; revision received 11 July 1995.

* Postal address: Departamento de Estadistica y Cálculo, Universidad Autónoma Agraria Antonio Narro, Buenavista, Saltillo COAH 25315, México.

** Postal address: Systems and Industrial Engineering Department, The University of Arizona, Tucson AZ 85721, USA.

* Research supported by a U.S.-México Collaborative Research Program funded by the National Science Foundation under Grant NSF-INT 9201430, and by the Consejo Nacional de Ciencia y Tecnologia (CONACyT), México.

** The work of Cavazos-Cadena was partially supported by the PSF Organization under Grant No. 200-150/94-01, and by the MAXTOR Foundation for Applied Probability and Statistics (MAXFAPS) under Grant No. 01-01-56/01-94.

Performance index. The (lim-inf expected) *average cost* at state $x \in S$ under policy π is defined by

$$J(x, \pi) := \liminf_{k \to \infty} \frac{1}{k+1} E_x^\pi \left[\sum_{t=0}^{k} C(X_t, A_t) \right], \tag{2.1}$$

and

$$J^*(x) := \inf_\pi J(x, \pi) \tag{2.2}$$

is the optimal average cost at state x. A policy π^* is *average optimal* (AO) if $J(x, \pi^*) = J^*(x)$ for all $x \in S$.

The use of the limit inferior in (2.1) implicitly assumes an *optimistic* viewpoint from the decision-maker, since what is being optimized is the *best performance* attained by a policy. On the other hand, an opposite *pessimistic* viewpoint could be adopted, by using the limit superior instead in (2.1). However, under Assumption 2.2 below, both criteria are *equivalent*; see [3].

Optimality equation. To establish the existence of optimal stationary policies it is necessary to complement Assumption 2.1 with additional conditions [1], [19], [17]. One such condition is introduced in Assumption 2.2 below. First, let $z \in S$ be an arbitrarily selected state, *fixed* throughout the remainder of the paper, and define the first passage time T as follows:

$$T := \min\{n > 0 \mid X_n = z\}, \tag{2.3}$$

where, by the usual convention, the minimum of the empty set is ∞. The next condition was introduced by Hordijk in [14]; see also [1], [4], [5], [6], [7], [9], [10].

Assumption 2.2. Lyapunov function condition. There exists a (Lyapunov) function $l : S \to [0, \infty)$ satisfying the following conditions (i)–(iii):

(i) $1 + |C(x, a)| + \Sigma_{y \neq z}\ p_{xy}(a) l(y) \leq l(x)$ for all $x \in S$ and $a \in A$.

(ii) For each $x \in S$, the mapping $f \mapsto \Sigma_y\ p_{xy}(f(x)) l(y) = E_x^f[l(X_1)]$ is continuous in $f \in \mathbb{F}$.

(iii) For each $f \in \mathbb{F}$ and $x \in S$, $E_x^f[l(X_n) I[T > n]] \to 0$ as $n \to \infty$.

Notice that the above condition imposes a growth restriction on the one-stage cost function. Moreover, under continuity and boundedness conditions for the one-stage cost function and a communicating assumption (under every stationary policy), it has been shown in [7] that the Lyapunov function condition is *equivalent* to several stability/ergodicity conditions on the controlled transition law of the system. Under Assumptions 2.1 and 2.2 the *average cost optimality equation* (ACOE) given by (2.4) below has a solution yielding an optimal stationary policy.

Lemma 2.1. Suppose that Assumptions 2.1 and 2.2 hold true. Then there exist $h : S \to \mathbb{R}$ and $g \in \mathbb{R}$ such that (i)–(iv) below hold true.

(i) $g = J^*(x)$ for each $x \in S$.

(ii) $h(z) = 0$ and for some constant $c > 0$, $|h(x)| \leq c \cdot l(x)$ for all $x \in S$.

(iii) The ACOE is satisfied by g and $h(\cdot)$, i.e.

$$g + h(x) = \min_a \left[C(x, a) + \sum_y p_{xy}(a) h(y) \right], \qquad x \in S. \tag{2.4}$$

(iv) An optimal stationary policy exists: for each $x \in S$ the right-hand side of (2.4) considered as a function of $a \in A$ has a minimizer $f^*(x)$, and the corresponding policy $f^* \in \mathbb{F}$ is optimal.

A proof of this result can be found in [14, ch. 5]; see also [6] for a proof of (ii). In addition to this lemma other (somewhat technical) consequences of Assumptions 2.1 and 2.2 will be used; they are stated in Lemmas A.1–A.3 in the appendix. Notice that g in Lemma 2.1 is *uniquely determined*, since it is the optimal average cost at every state. The function h is also unique, as established in Lemma A.2(iv).

As already mentioned, the main objective of this paper is to study the VI procedure to obtain approximations to the solution $(g, h(\cdot))$ of the ACOE. Theorems 3.1 and 3.2 in the next section point in this direction and require the following additional condition.

Assumption 2.3. For each $a \in A$, $p_{zz}(a) > 0$.

Remark 2.1. Assumption 2.3 is readily verifiable. Furthermore, it does *not* imply any loss of generality, since it can be obtained by making an appropriate transformation on the transition law. In fact, suppose that $M = \langle S, A, C, P \rangle$ satisfies Assumptions 2.1 and 2.2 and define the transformed transition law $P^* = [p_{xy}^*(\cdot)]$ as follows:

$$p_{xy}^*(a) := (1 - \alpha)\delta_{xy} + \alpha \cdot p_{xy}(a), \qquad (x, a) \in S \times A, \quad y \in S,$$

where $\alpha \in [0, 1)$ is a given number and $\delta_{xy} := 1$ (resp. 0) if $x = y$ (resp. $x \neq y$). Now set $M^* := \langle S, A, C, P^* \rangle$, which clearly satisfies Assumptions 2.1 and 2.3. Moreover, it is not difficult to see that $l^*(\cdot) = l(\cdot)/\alpha$ is a Lyapunov function for M^*, so that Assumption 2.2 is also satisfied by M^*. On the other hand, M and M^* are *equivalent* CMC in the following sense. Let the pair (g, h) be the solution to the ACOE for model M and let (g^*, h^*) be the corresponding pair for model M^*. Then (a) $g^* = g$, (b) $h^* = h/\alpha$, and (c) a policy $f \in \mathbb{F}$ is optimal for model M if and only if f is optimal for M^*. Furthermore, (d) a policy $f \in \mathbb{F}$ is such that, for all states x, $f(x)$ minimizes the mapping $a \mapsto C(x, a) + \Sigma_y\ p_{xy}(a) h(y)$, if and only if it minimizes $a \mapsto C(x, a) + \Sigma_y\ p_{xy}^*(a) h^*(y)$. The transformation $p \mapsto p^*$ was introduced by Schweitzer in [20]; see also [17, pp. 371–373].

3. Main results

This section contains the main results in the paper. To begin with, the necessary notions are introduced.

Definition 3.1. The VI scheme. (i) The sequence $\{V_n : S \to \mathbb{R} \mid n = -1, 0, 1, \cdots\}$ of *value iteration functions* is recursively defined as follows: $V_{-1} \equiv 0$ and, for $n \geq 0$,

$$V_n(x) = \min_a \left[C(x, a) + \sum_y p_{xy}(a) V_{n-1}(y) \right], \qquad x \in S. \tag{3.1}$$

(ii) The *relative value functions* $R_n : S \to \mathbb{R}$ are defined by $R_n(x) := V_n(x) - V_n(z)$, for $x \in S$, $n = -1, 0, 1, 2, \cdots$.

(iii) For each $x \in S$ and $n \in \mathbb{N}$ define the *nth differential cost* at x by $g_n(x) := V_n(x) - V_{n-1}(x)$.

Remark 3.1. From Assumption 2.1, a standard induction argument yields that for each $x \in S$ and $n \in \mathbb{N}$ ([1], [11]), *the mapping*

$$a \mapsto C(x, a) + \sum_y p_{xy}(a) C(x, a), \qquad a \in A$$

is continuous. Therefore, the minimum in (3.1) is indeed attained.

The following are well known results; see [1], [11], [17].

Lemma 3.1. (i) The value iteration functions satisfy

$$V_n(x) = \inf_\pi E_x^\pi \left[\sum_{t=0}^n C(X_t, A_t) \right], \qquad x \in S. \tag{3.2}$$

(ii) Furthermore, there exists a Markovian policy π^n which attains the infimum in (3.2).

The differential costs and relative value functions are natural candidates to approximate the solution $(g, h(\cdot))$ of the ACOE. Consider the following conditions.

C1. For all $x \in S$, $g_n(x) \to g$ and $R_n(x) \to h(x)$ as $n \to \infty$.

C2. The sequence $\{g_n(z)\}$ is *bounded*, i.e. there exists $b \in [0, \infty)$ such that $|g_n(z)| \leq b$ for all $n \in \mathbb{N}$.

It is clear that C1 implies C2. The first main result of the paper is as follows.

Theorem 3.1. Suppose that Assumptions 2.1–2.3 hold true. Then (i) and (ii) below hold true.

(i) Conditions C1 and C2 are *equivalent*.

(ii) Suppose that C2 holds, and that for each $n \in \mathbb{N}$, policy $f_n \in \mathbb{F}$ is such that, for each $x \in S$, $f_n(x)$ is a minimizer of the mapping

$$a \mapsto C(x, a) + \sum_y p_{xy}(a) R_n(y), \qquad a \in A. \tag{3.3}$$

Then

(3.4) every limit point of $\{f_n\} \subset \mathbb{F}$ is AO.

A proof of this result is contained in Section 5. By part (i), establishing convergence in C1 is equivalent to verifying the — apparently weaker — condition of boundedness of the sequence $\{g_n(z)\}$. This criterion is now used to obtain the following.

Theorem 3.2. Suppose that Assumptions 2.1–2.3 hold true and that for some constant $B \in [0, \infty)$, either $C(\cdot,\cdot) \geq -B$ or $C(\cdot,\cdot) \leq B$. Then, C1 holds.

Proof. By Theorem 3.1, it is sufficient to verify C2. With this in mind, suppose first that $C(\cdot,\cdot) \geq -B$ and observe that (see (3.2))

$$V_n(z) = \min_\pi E_z^\pi \left[\sum_{t=0}^n C(X_t, A_t) \right] \geq \min_\pi E_z^\pi \left[-B + \sum_{t=0}^{n-1} C(X_t, A_t) \right]$$
$$\geq -B + \min_\pi E_z^\pi \left[\sum_{t=0}^{n-1} C(X_t, A_t) \right] = -B + V_{n-1}(z)$$

so that $V_n(z) - V_{n-1}(z) \geq -B$, $n \in \mathbb{N}$. Since $V_n(z) - V_{n-k}(z) = \sum_{r=0}^{k-1} V_{n-r}(z) - V_{n-r-1}(z)$ for $k \leq n+1$, it follows that

$$V_n(z) - V_{n-k}(z) \geq -k \cdot B, \qquad n, k \in \mathbb{N}, \quad n+1 \geq k. \tag{3.5}$$

In particular, setting $k = n+1$ and recalling that $V_{-1} \equiv 0$,

$$V_n(z) \geq -(n+1)B. \tag{3.6}$$

On the other hand,

$$V_n(z) = E_z^{\pi^n} \left[\sum_{t=0}^n C(X_t, A_t) \right]$$
$$= E_z^{\pi^n} \left[\sum_{t=0}^n C(X_t, A_t) I[T \leq n] \right] + E_z^{\pi^n} \left[\sum_{t=0}^n C(X_t, A_t) I[T > n] \right]$$
$$= E_z^{\pi^n} \left[\left\{ \sum_{t=0}^{T-1} C(X_t, A_t) + V_{n-T}(z) \right\} I[T \leq n] \right] + E_z^{\pi^n} \left[\sum_{t=0}^n C(X_t, A_t) I[T > n] \right],$$

where the third equality follows from Bellman's optimality principle. Hence,

$$-B \leq V_n(z) - V_{n-1}(z) = E_z^{\pi^n} \left[\left\{ \sum_{t=0}^{T-1} C(X_t, A_t) - (V_{n-1}(z) - V_{n-T}(z)) \right\} I[T \leq n] \right]$$
$$+ E_z^{\pi^n} \left[\sum_{t=0}^n C(X_t, A_t) I[T > n] \right] - V_{n-1}(z) P_z^{\pi^n}[T > n].$$

Since $T \geq 1$, (3.5) yields that $-(V_{n-1}(z) - V_{n-T}(z)) I[T \leq n] \leq (T-1) B I[T \leq n] \leq T B I[T \leq n]$, and from (3.6), $-V_{n-1}(z) \leq nB \leq (n+1)B$. Therefore,

$$-B \leq V_n(z) - V_{n-1}(z) \leq E_z^{\pi^n} \left[\left\{ \sum_{t=0}^{T-1} C(X_t, A_t) + T \cdot B \right\} I[T \leq n] \right]$$
$$+ E_z^{\pi^n} \left[\sum_{t=0}^n C(X_t, A_t) I[T > n] \right] + (n+1) B P_z^{\pi^n}[T > n]$$

$$\leq E_z^{\pi^n}\left[\sum_{t=0}^{T-1}(C(X_t, A_t)+B)I[T\leq n]\right]$$

$$+E_z^{\pi^n}\left[\sum_{t=0}^{n}(C(X_t, A_t)+B)I[T>n]\right]$$

$$\leq E_z^{\pi^n}\left[\sum_{t=0}^{T-1}(|C(X_t, A_t)|+B)\right]$$

$$\leq (B\vee 1)E_z^{\pi^n}\left[\sum_{t=0}^{T-1}(|C(X_t, A_t)|+1)\right]$$

$$\leq (B\vee 1)l(z),$$

where the last inequality follows from Lemma A.1(i). Hence C2 holds true. Similarly, it can be proved that $C(\cdot,\cdot)\leq B$ implies that the sequence of differential costs at z is bounded, and as already mentioned this completes the proof.

4. Preliminaries

This section contains the main technical results (Lemmas 4.1 and 4.2) that will be used in the proof of Theorem 3.1. Throughout the remainder of the section, Assumptions 2.1–2.3 as well as Condition C2 are supposed to hold true. Before going any further, the reader may find it convenient to glance at Lemmas A.1–A.3 in the appendix.

Lemma 4.1. Let $b>0$ be as in C2. Then, for all $x\in \boldsymbol{S}$ and $n\in\mathbb{N}$, (i) $|R_n(x)|\leq (b\vee 1)l(x)$, and (ii) $|g_n(x)|\leq 3(b\vee 1)l(x)$.

Proof. First notice that part (ii) is a consequence of part (i). In fact, from Definition 3.1 it follows easily that $|g_n(x)|=|R_n(x)+g_n(z)-R_{n-1}(x)|$ so that (i) implies, via the triangle inequality, that $|g_n(x)|\leq (b\vee 1)l(x)+b+(b\vee 1)l(x)$ and the inequality in part (ii) follows, since $l(x)\geq 1$; see Assumption 2.2(i). Therefore, it is a sufficient to provide a proof of part (i), and the following arguments are along the lines used to establish Theorem 3.2. Notice that $|g_n(z)|=|V_n(z)-V_{n-1}(z)|\leq b$ for all $n\in\mathbb{N}$ implies that

$$|V_n(z)-V_{n-k}(z)|=\left|\sum_{t=0}^{k-1}V_{n-t}(z)-V_{n-t-1}(z)\right|\leq k\cdot b \tag{4.1}$$

for all non-negative integers n and k with $k\leq n+1$; in particular,

$$|V_n(z)|=|V_n(z)-V_{-1}(z)|\leq (n+1)\cdot b. \tag{4.2}$$

Now let π^n be as in Lemma 3.1(ii) and observe that Bellman's optimality principle yields that

$$E_x^{\pi}\left[\sum_{t=0}^{n}C(X_t, A_t)I[T\leq n]\right]=E_x^{\pi^n}\left[\left\{V_{n-T}(z)+\sum_{t=0}^{T-1}C(X_t, A_t)\right\}I[T\leq n]\right]$$

so that

$$V_n(x)=E_x^{\pi^n}\left[\sum_{t=0}^{n}C(X_t, A_t)\right]$$

$$=E_x^{\pi^n}\left[\left\{V_{n-T}(z)+\sum_{t=0}^{T-1}C(X_t, A_t)\right\}I[T\leq n]\right]+E_x^{\pi^n}\left[\sum_{t=0}^{n}C(X_t, A_t)I[T>n]\right]$$

and then

$$|V_n(x)-V_n(z)|=\left|E_x^{\pi^n}\left[\left\{V_{n-T}(z)-V_n(z)+\sum_{t=0}^{T-1}C(X_t, A_t)\right\}I[T\leq n]\right]\right.$$

$$\left.+E_x^{\pi^n}\left[-V_n(z)+\sum_{t=0}^{n}C(X_t, A_t)I[T>n]\right]\right|$$

which combined with (4.1) and (4.2) yields

$$|V_n(x)-V_n(z)|\leq E_x^{\pi^n}\left[\left\{b\cdot T+\sum_{t=0}^{T-1}|C(X_t, A_t)|\right\}I[T\leq n]\right]$$

$$+E_x^{\pi^n}\left[b\cdot(n+1)+\sum_{t=0}^{n}|C(X_t, A_t)|I[T>n]\right]$$

$$\leq (b\vee 1)E_x^{\pi^n}\left[\sum_{t=0}^{T-1}(|C(X_t, A_t)|+1)\right],$$

and then Lemma A.1(i) yields that $|R_n(x)|=|V_n(x)-V_n(z)|\leq (b\vee 1)l(x)$.

To continue it is convenient to introduce additional notation.

Definition 4.1. Define the functions $U, L:\boldsymbol{S}\to\mathbb{R}$ as follows. For each $x\in\boldsymbol{S}$,

$$U(x):=\limsup_{n\to\infty}g_n(x)\quad\text{and}\quad L(x):=\liminf_{n\to\infty}g_n(x).$$

By Lemma 4.1, $|U(\cdot)|, |L(\cdot)|\leq 3(b\vee 1)l(\cdot)$. The following two lemmas establish useful properties of these functions.

Lemma 4.2. For each $x\in\boldsymbol{S}$,

(i) $L(z)\leq L(x)\leq U(x)\leq U(z)$.

(ii) Let $\{n(k)\}\subset\mathbb{N}$ be such that $n(k)\nearrow\infty$ as $k\nearrow\infty$. If $\lim_{k\to\infty}g_{n(k)}(z)=L(z)$, then $\lim_{k\to\infty}g_{n(k)-s}(z)=L(z)$ for all $s\in\mathbb{N}$.

Similarly,

(iii) If $\lim_{k\to\infty}g_{n(k)}(z)=U(z)$ as $k\to\infty$, then $\lim_{k\to\infty}g_{n(k)-s}(z)=U(z)$ for all $s\in\mathbb{N}$.

Proof. (i) Let $x\in\boldsymbol{S}$ and the positive integer n be arbitrary and select $a_n\in\boldsymbol{A}$ (which depends on $x\in\boldsymbol{S}$) such that

$$V_n(x) = C(x, a_n) + \sum_y p_{xy}(a_n) V_{n-1}(y);$$

see (3.1) and Remark 3.1. In this case

$$V_{n-1}(x) \leqq C(x, a_n) + \sum_y p_{xy}(a_n) V_{n-2}(y)$$

and then

$$g_n(x) = V_n(x) - V_{n-1}(x) \geqq \sum_y p_{xy}(a_n)(V_{n-1}(y) - V_{n-2}(y)) = \sum_y p_{xy}(a_n) g_{n-1}(y). \tag{4.3}$$

Now select a sequence $\{n(k)\}$ of positive integers increasing to ∞ such that $\lim_k g_{n(k)}(x) = L(x)$. Since $\{a_{n(k)}\} \subset A$ and the metric space A is compact, it can be assumed, taking a subsequence if necessary, that

$$a_{n(k)} \to a(x) \in A \quad \text{as} \quad k \to \infty. \tag{4.4}$$

Then, replacing n by $n(k)$ in (4.3) and taking the limit inferior as $k \to \infty$ in the resulting inequality, it follows that

$$\begin{aligned} L(x) = \liminf_k g_{n(k)}(z) &\geqq \liminf_k \sum_y p_{xy}(a_{n(k)}(x)) g_{n(k)-1}(y) \\ &\geqq \sum_y p_{xy}(a(x)) \liminf_k g_{n(k)-1}(y) \\ &\geqq \sum_y p_{xy}(a(x)) L(y) \end{aligned} \tag{4.5}$$

where Lemma 4.1(ii) and Lemma A.3(ii) were used to obtain the second inequality, whereas the third one follows from the definition of $L(y)$ as the limit inferior of the *whole* sequence $\{g_n(y)\}$. Setting $f(x) := a(x)$, $x \in S$, it follows that $L(x) \geqq E_x^f[L(X_1)]$ for all $x \in S$, and then Lemma A.2(iii) yields that $L(x) \geqq L(z)$ for all $x \in S$. Similarly, it can be established that $U(\cdot) \leqq U(z)$, whereas the inequality $L(\cdot) \leqq U(\cdot)$ is clear; see Definition 4.1.

(ii) By induction, we need only consider the $s=1$ case, which we do next. Suppose that $\lim_k g_{n(k)}(z) = L(z)$ and let $L'(z)$ be an arbitrary limit point of $\{g_{n(k)-1}(z)\}$. It is clear that the desired conclusion will be reached if it is proved that $L'(z) = L(z)$. To verify this equality observe that without loss of generality it can be assumed that

$$g_{n(k)-1}(z) \to L'(z) \quad \text{as} \quad k \to \infty. \tag{4.6}$$

In the arguments used to establish part (i) set $x = z$ and take an additional subsequence (if necessary) such that (4.4) holds with $x = z$. In this case (4.5) becomes

$$\begin{aligned} L(z) &\geqq \liminf_k \sum_y p_{zy}(a_n(z)) g_{n(k)-1}(y) \\ &\geqq \sum_y p_{zy}(a(z)) \liminf_k g_{n(k)-1}(y) \quad \text{:by Lemmas 4.1(ii) and A.3(ii)} \\ &\geqq \sum_y p_{zy}(a(z)) L(y) \geqq L(z) \end{aligned} \tag{4.7}$$

where the last inequality follows from part (i). Since $\liminf_k g_{n(k)-1}(\cdot) \geqq L(\cdot)$, (4.7) yields that

$$\liminf_k g_{n(k)-1}(y) = L(y) \qquad \text{if } p_{zy}(a(z)) > 0. \tag{4.8}$$

Then Assumption 2.3, (4.6) and (4.8) together yield that $L'(z) = L(z)$ and, as already mentioned, this completes the proof of part (ii), whereas (iii) can be established in a similar way.

Remark 4.1. The key point in parts (ii) and (iii) in Lemma 4.2 is that $\{n(k)\}$ and $\{n(k)-s\}$, s a positive integer, are *totally different* sequences, in general. The latter is not a delayed or shifted version of the former.

Lemma 4.3. (i) There exist functions $\hat{R}_s : S \to \mathbb{R}$, $s \in \mathbb{N}$, such that
(a) $|\hat{R}_s(\cdot)| \leqq (b \vee 1) l(\cdot)$, and
(b) $U(z) + \hat{R}_s(x) = \min_a [C(x, a) + \Sigma_y p_{xy}(a) \hat{R}_{s+1}(y)]$, for all $x \in S$ and $s \in \mathbb{N}$.
Similarly,
(ii) There exist functions $\tilde{R}_s : S \to \mathbb{R}$, $s \in \mathbb{N}$, such that
(a) $|\tilde{R}_s(\cdot)| \leqq (b \vee 1) l(\cdot)$, and
(b) $L(z) + \tilde{R}_s(x) = \min_a [C(x, a) + \Sigma_y p_{xy}(a) \tilde{R}_{s+1}(y)]$ for all $x \in S$ and $s \in \mathbb{N}$.

Proof. (i) Pick an increasing sequence $\{n(k)\} \subset \mathbb{N}$ such that $g_{n(k)}(z) \to U(z)$ as $k \to \infty$, and notice that Lemma 4.2(ii) yields that for all $s \in \mathbb{N}$

$$\lim_{k \to \infty} g_{n(k)-s}(z) = U(z). \tag{4.9}$$

For k large enough, $n(k) - s \geqq 0$, and the recursive relation (3.1) with $n(k) - s$ instead of n can be written (see Definition 3.1) as

$$g_{n(k)-s}(z) + R_{n(k)-s}(x) = \min_a \left[C(x, a) + \sum_y p_{xy}(a) R_{n(k)-s-1}(y) \right]. \tag{4.10}$$

Now set $R_t(\cdot) := 0$ for $t < 0$. By Lemma 4.1(i), $\{R_{n(k)-s}(\cdot)\}$ is a sequence in the compact metric space

$$\mathbb{D} := \prod_{x \in S} [-(b \vee 1) l(x), (b \vee 1) l(x)], \tag{4.11}$$

so that taking a subsequence, if necessary, it can be assumed that

$$\lim_{k \to \infty} R_{n(k)-s}(x) := \hat{R}_s(x) \in [-(b \vee 1) l(x), (b \vee 1) l(x)] \tag{4.12}$$

exists for all $x \in S$ and $s \in \mathbb{N}$. Taking limits as $k \to \infty$ in both sides of (4.10) it follows, via (4.9), (4.12) and Lemma A.3(iv), that

$$U(z)+\hat{R}_s(x)=\min_a\left[C(x,a)+\sum_y p_{xy}(a)\hat{R}_{s+1}(y)\right],\qquad x\in S,\ s\in\mathbb{N},$$

and, by (4.12), this completes the proof of part (i). Part (ii) can be proved along the same lines.

5. Proof of Theorem 3.1

The preliminaries in the previous section will now be used to prove Theorem 3.1.

Proof of Theorem 3.1. (i) As already mentioned it suffices to prove that C2 implies C1. First it will be established that

(5.1) $$\lim_{n\to\infty} g_n(x)=g,\qquad x\in S.$$

From Definition 3.1 this is equivalent to $U(\cdot)=L(\cdot)\equiv g$, which, by Lemma 4.2(i), reduces to

(5.2) $$L(z)=U(z)=g,$$

a relation that is verified as follows. Let $\{\hat{R}_s\}$ and $\{\tilde{R}_s\}$ be as in Lemma 4.3. A simple induction argument using the equations in Lemma 4.3(i) yields that for each policy π,

$$U(z)+\frac{1}{n+1}\hat{R}_0(x)\leqq\frac{1}{n+1}E_x^\pi\left[\sum_{t=0}^{n}C(X_t,A_t)\right]+\frac{1}{n+1}E_x^\pi[\hat{R}_{n+1}(X_{n+1})],\quad x\in S,\ n\in\mathbb{N}.$$

Recall now that $|\hat{R}_s(\cdot)|\leqq(b\vee 1)l(\cdot)$ for all $s\in\mathbb{N}$, by Lemma 4.3(i), so that taking the limit inferior as $n\to\infty$ in both sides of the above inequality it follows, via Lemma A.1(ii), that $U(z)\leqq J(x,\pi)$, and since policy π was arbitrary this yields that $U(z)\leqq J(x)$, i.e.

(5.3) $$U(z)\leqq g;$$

see Lemma 2.1(i). Now, for each $s\in\mathbb{N}$ select a stationary policy $f_s\in\mathbb{F}$ such that

$$L(z)+\tilde{R}_s(x)=C(x,f_s(x))+\sum_y p_{xy}(f_s(x))\tilde{R}_{s+1}(y),\qquad x\in S,$$

and form the Markov policy $\pi:=\{f_s\}$; since A is compact such a policy exists by Lemma A.2(i). In this case a simple induction argument yields that for all $x\in S$ and $n\in\mathbb{N}$,

$$L(z)+\frac{1}{n+1}\tilde{R}_0(x)=\frac{1}{n+1}E_x^\pi\left[\sum_{t=0}^{n}C(X_t,A_t)\right]+\frac{1}{n+1}E_x^\pi[\tilde{R}_{n+1}(X_{n+1})];$$

taking the limit inferior as n increases to ∞ in both sides of this inequality it follows that

$$L(z)=J(x,\pi)\geqq J(x)=g;$$

see (2.1), (2.2) and Lemma 2.1(i). Since $L(x)\leqq U(z)$, the last relation and (5.3) together imply (5.2) which is equivalent to (5.1). To conclude, it will be established that

(5.4) $$\lim_{n\to\infty}R_n(x)=h(x),\qquad x\in S.$$

To verify this observe that, by Lemma 4.1, $\{R_n\}\subset\mathbb{D}$ and that $\mathbb{D}$ is a compact metric space; see (4.11). Therefore, it is sufficient to see that any limit point of $\{R_n\}$ *in* $\mathbb{D}$ coincides with h. With this in mind, let $Q\in\mathbb{D}$ be such that

(5.5) $$\lim_{k\to\infty}R_{n(k)}(x)=Q(x),\qquad x\in S,$$

for some increasing sequence $\{n(k)\}$. From Definition 3.1 and (5.1) it follows that $R_n(x)-R_{n-1}(x)=g_n(x)-g_n(z)\to 0$ as $n\to\infty$, so that

(5.6) $$\lim_{k\to\infty}R_{n(k)-1}(x)=Q(x),\qquad x\in S$$

also holds. Next observe that (3.1) implies, via straightforward calculations using Definition 3.1, that

$$g_{n(k)}(z)+R_{n(k)}(x)=\min_a\left[C(x,a)+\sum_y p_{xy}R_{n(k)-1}(y)\right],$$

and taking the limit as $k\to\infty$ in both sides of this equality, it follows, via (5.1), (5.5), (5.6) and Lemma A.3(iv), that

$$g+Q(x)=\min_a\left[C(x,a)+\sum_y p_{xy}(a)Q(y)\right],\qquad x\in S,$$

i.e. the pair $(g, Q(\cdot))$ satisfies the ACOE. Observing that $Q(z)=\lim_{k\to\infty}R_{n(k)}(z)=0$ (see Definition 3.1), Lemma A.2(iv) yields that $Q(\cdot)=h(\cdot)$. In short, any limit point of $\{R_n(\cdot)\}\subset\mathbb{D}$ has been shown to coincide with $h(\cdot)$, and as already mentioned, this establishes (5.4) and the proof of part (i) is complete.

(ii) Let $\{f_n\}\subset\mathbb{F}$ such that for all $x\in S$, $n\in\mathbb{N}$ and $a\in A$,

(5.7) $$C(x,f_n(x))+\sum_y p_{xy}(f_n(x))R_n(y)\leqq C(x,a)+\sum_y p_{xy}(a)R_n(y).$$

Now let $f\in\mathbb{F}$ be a limit point of $\{f_n\}$ and select a sequence $\{n(k)\}\subset\mathbb{N}$ increasing to ∞ such that $\lim_{k\to\infty}f_{n(k)}=f$. In this case replace n by $n(k)$ in (5.7). Taking the limit as $k\to\infty$ in both sides of the resulting inequality, Lemma 4.1(i), part (i) and Lemma A.3(iii) together yield that

$$C(x,f(x))+\sum_y p_{xy}(f(x))h(y)\leqq C(x,a)+\sum_y p_{xy}(a)h(y),\quad \text{for all } x\in S,\ a\in A,$$

and then f is AO, by Lemma 2.1(iv).

6. Concluding remarks

In this paper, we studied the value iteration scheme in the context of denumerable CMC endowed with the average cost criterion and satisfying Assumptions 2.1–2.3. The main result — namely, Theorem 3.1 — shows the *equivalence* of C1 and C2, and this was applied in Theorem 3.2 to obtain that the differential costs and relative value functions converge pointwise to the solution of the ACOE when the cost function is bounded above or below. Although Assumption 2.3 does *not* imply any loss of generality, it played a subtle (but *essential*) role in the argumentation to establish Theorem 3.1; see the proof of Lemma 4.2(ii). Thus, Assumption 2.3 must be checked before starting the VI scheme. If that condition is not satisfied, Schweitzer's transformation should be performed and the VI scheme must be applied with the transformed model. Fortunately, this will not create any additional problem since both models are equivalent, as described in Remark 2.1.

On the other hand, the literature on the VI scheme is extensive and it is important to point out the main differences between the results in this paper and those already available; for comments on results obtained under stability conditions different to Assumption 2.2 — as the (simultaneous) Doeblin or Scrambling conditions — or applications of the VI method see, for instance, [11], [12], [16], [5], [9], [10], [17], [21] and the references therein. The following remarks refer to the relation between Theorem 3.2 and other results previously obtained under LFC.

(i) A fundamental paper on the VI scheme under the Lyapunov function condition is [15], where C1 was established under Assumptions 2.1–2.3 *and* the following additional boundedness condition:

$$\sup_{x\in S} |e(x)| < \infty, \tag{6.1}$$

where the *first error function* is defined as $e(\cdot) := V_0 - [g + h(\cdot)]$. Such a condition is quite restrictive and is *not* satisfied in common applications; in fact, it is closely related to the simultaneous Doeblin condition — see Example 3.1 and Remark 3.1 in [5]. Theorem 3.2 avoids (6.1) at the expense of assuming that the cost function is bounded above or below; however, this latter assumption is usually satisfied since, frequently, costs are non-negative, or at least bounded below. On the other hand, it would be very interesting to verify whether C1 holds under Assumptions 2.1–2.3 and nothing more. This problem reduces, by Theorem 3.1(i), to establishing that Assumptions 2.1–2.3 imply that the differential costs at state z form a bounded sequence. We have not been able to prove or disprove such a result up to now.

(ii) In [5], C1 was obtained under Assumptions 2.1–2.3 for an arbitrary *bounded* cost function. On the other hand, in [5], (6.1) was avoided but an additional communicating condition was required, under which $\{g_n(z), R_n(\cdot)\}$ was shown to converge to (g, h) in the Cesàro sense.

In short, Theorem 3.2 can be seen as a significant extension of the results in [15], [4] and [5]. As already mentioned, an interesting open problem is to investigate if C1 can be obtained within the context of Section 2 without requiring a lower or upper bound for the cost function.

Appendix

Throughout the remainder Assumptions 2.1 and 2.2 are supposed to hold true. These conditions imply the results in Lemmas A.1–A3 below, which are used extensively in the arguments contained in Sections 4 and 5.

Lemma A.1. For each policy π we have that:
(i) $E_x^\pi[T] \leqq E_x^\pi[\Sigma_{t=0}^{T-1}(|C(X_t, A_t)| + 1)] \leqq l(x)$, $x \in S$.
(ii) $[1/(n+1)]E_x^\pi[l(X_n)] \to 0$ as $n \to \infty$, $x \in S$.

Proof. (i) The first inequality is clear. To establish the other inequality observe that Assumption 2.2(i) and an induction argument together yield, for every policy π, that

$$E_x^\pi\left[\sum_{t=0}^{n}(1+|C(X_t, A_t)|)I[T>t] + l(X_{n+1})I[T>n+1]\right] \leqq l(x), \quad x\in S,\ n\in\mathbb{N};$$

since $l(\cdot) \geqq 0$, it follows that

$$E_x^\pi\left[\sum_{t=0}^{T-1}(1+|C(X_t, A_t)|)\right] = \lim_{n\to\infty} E_x^\pi\left[\sum_{t=0}^{n}(1+|C(X_t, A_t)|)I[T>t]\right] \leqq l(x).$$

(ii) A proof of this part — using only conditions (ii) and (iii) in Assumption 2.2 — can be seen, for instance, in [4], [5] or [6].

Lemma A.2. Let $W: S \to \mathbb{R}$ be such that $|W(\cdot)| \leqq Bl(\cdot)$ for some constant $B > 0$.
(i) For each $x \in S$, the mapping $a \mapsto \Sigma_y p_{xy}(a)W(y)$ is continuous in $a \in A$.
(ii) If $f \in \mathbb{F}$ is such that $W(x) \leqq E_x^f[W(X_1)]$ for all $x \in S$, then $W(x) \leqq W(z)$, $x \in S$. Similarly,
(iii) If for some policy $f \in \mathbb{F}$, $W(x) \geqq E_x^f[W(X_1)]$ for all $x \in S$, then $W(x) \geqq W(z)$, $x \in S$.
(iv) (Uniqueness of the solution of the ACOE.) Suppose that W_1 and W_2 satisfy
(a) $|W_i(\cdot)| \leqq Bl(\cdot)$, $i = 1, 2$ for some constant $B > 0$;
(b) $W_1(z) = W_2(z) = 0$, and
(c) W_1 and W_2 satisfy the ACOE, i.e.

$$g + W_i(x) = \min_a\left[C(x,a) + \sum_y p_{xy}(a)W_i(y)\right], \quad i = 1, 2, \quad x \in S. \tag{A.1}$$

Then $W_1 = W_2$.

Proof. (i) Let $\{S_k\}$ be a sequence of *finite* subsets of S with $S_n \nearrow S$ as $n \nearrow \infty$. Let $x \in S$ be fixed. By Assumption 2.1, for each $n \in \mathbb{N}$ the mappings

$$a \mapsto \sum_{y\in S_n} p_{xy}(a)l(y) \quad \text{and} \quad a \mapsto \sum_{y\in S_n} p_{xy}(a)W(y) \quad \text{are continuous in } a \in A.$$

On the other hand, as $n \nearrow \infty$, $\Sigma_{y\in S_n} p_{xy}(a)l(y)$ increases to $\Sigma_y p_{xy}(a)l(y)$, which depends continuously on $a \in A$, by Assumption 2.2(ii). Since A is a compact space, by Dini's theorem (see [8], [18]), for every $\varepsilon > 0$ there exists an integer $n(\varepsilon) > 0$ such that

$$\sum_{y\notin S_k} p_{xy}(a)l(y) \leqq \varepsilon \qquad \text{for } k \geqq n(\varepsilon),\ a \in A. \tag{A.2}$$

Then, for all $k \geqq n(\varepsilon)$ and $a \in A$,

$$\left| \sum_{y \in S_k} p_{xy}(a)W(y) - \sum_{y \in S} p_{xy}(a)W(y) \right| \leqq \sum_{y \notin S_k} p_{xy}(a)|W(y)|$$

$$\leqq B \sum_{y \notin S_k} p_{xy}(a)l(y) \leqq B\varepsilon.$$

Thus, as a uniform limit of continuous functions, $a \mapsto \Sigma_{y \in S} p_{xy}(a)W(y)$ is itself a continuous mapping.

(ii) Notice that $W(x) \leqq E_x^f[W(X_1)] = W(z)P_x^f[T=1] + E_x^f[W(X_1)I[T>1]]$. Using this, a simple induction argument yields that, for all $x \in S$ and $n \in \mathbb{N}$,

$$W(x) \leqq W(z)P_x^f[T \leqq n] + E_x^f[W(X_{n+1})I[T>n+1]]$$

$$\leqq W(z)P_x^f[T \leqq n] + BE_x^f[l(X_{n+1})I[T>n+1]].$$

Observe now that $P_x^f[T<\infty]=1$ (a consequence of Lemma A.1(i)), so that taking the limit as $n \to \infty$ in the above inequality and using Assumption 2.2(iii), it follows that $W(\cdot) \leqq W(z)$. The proof of part (iii) follows along the same lines.

(iv) Lemma 3.3 in [13] and (A.1) together imply that, for all $x \in S$,

$$|W_1(x) - W_2(x)| \leqq \sup_a \sum_y p_{xy}(a)|W_1(y) - W_2(y)|.$$

Observing that $|W_1(\cdot) - W_2(\cdot)| \leqq 2Bl(\cdot)$, the compactness of A and part (i) together imply that there exists a policy $f \in \mathbb{F}$ such that

$$|W_1(z) - W_2(x)| \leqq \sum_y p_{xy}(f(x))|W_1(y) - W_2(y)| = E_x^f[|W_1(X_1) - W_2(X_1)|], \qquad x \in S,$$

and part (ii) yields that $|W_1(\cdot) - W_2(\cdot)| \leqq |W_1(z) - W_2(z)|$. Then $W_1(\cdot) = W_2(\cdot)$ follows since $W_1(z) = W_2(z) = 0$.

Lemma A.3. Let the functions $W: S \to \mathbb{R}$, and $W_n: S \to \mathbb{R}$, $n \in \mathbb{N}$ be such that for some constant $B > 0$,

(A.3) $$|W(\cdot)| \leqq Bl(\cdot) \quad \text{and} \quad |W_n(\cdot)| \leqq Bl(\cdot), \qquad n \in \mathbb{N}.$$

Also, let $\{a_n\} \subset A$ be such that $a_n \to a \in A$ as $n \to \infty$. Then, for each $x \in S$, (i)–(iv) hold true.

(i) $\limsup_{n\to\infty} \Sigma_y p_{xy}(a_n)W_n(y) \leqq \Sigma_y p_{xy}(a) \limsup_{n\to\infty} W_n(y)$.

Similarly,

(ii) $\liminf_{n\to\infty} \Sigma_y p_{xy}(a_n)W_n(y) \geqq \Sigma_y p_{xy}(a) \liminf_{n\to\infty} W_n(y)$.

Suppose, additionally, that $W_n(x) \to W(x)$ as $n \to \infty$, for each $x \in S$. In this case, for all $x \in S$,

(iii) $\lim_{n\to\infty} \Sigma_y p_{xy}(a_n)W_n(y) = \Sigma_y p_{xy}(a)W(y)$

and

(iv) as $n \to \infty$,

$$M_n^x := \min_a \left[C(x,a) + \sum_y p_{xy}(a)W_n(y) \right] \to \min_a \left[C(x,a) + \sum_y p_{xy}(a)W(y) \right] =: M^x.$$

Proof. Let the sequences $\{S_n\}$ be as in the proof of Lemma A.2, pick $\varepsilon > 0$ and select an integer $n(\varepsilon) > 0$ such that (A.2) holds. For a fixed $k \geqq n(\varepsilon)$, (A.2) and (A.3) together yield that

(A.4) $$\sum_y p_{xy}(a_n)W_n(y) \leqq \sum_{y \in S_k} p_{xy}(a_n)W_n(y) + B\varepsilon,$$

so that

$$\limsup_{n\to\infty} \sum_y p_{xy}(a_n)W_n(y) \leqq \limsup_{n\to\infty} \sum_{y \in S_k} p_{xy}(a_n)W_n(y) + B\varepsilon$$

$$\leqq \sum_{y \in S_k} \limsup_{n\to\infty} p_{xy}(a_n)W_n(y) + B\varepsilon$$

$$= \sum_{y \in S_k} p_{xy}(a) \limsup_{n\to\infty} W_n(y) + B\varepsilon,$$

where the second inequality follows from the finiteness of S_k and the equality is a consequence of Assumption 2.1. Since $|\limsup_{n\to\infty} W_n(\cdot)| \leqq Bl(\cdot)$ (by (A.3)), it follows that

$$\limsup_{n\to\infty} \sum_y p_{xy}(a_n)W_n(y) \leqq \sum_y p_{xy}(a) \limsup_{n\to\infty} W_n(y) + 2B\varepsilon$$

(see (A.2)), and the conclusion is reached, since $\varepsilon > 0$ was arbitrary. Part (iii) can be established along the same lines.

Now assume that $\{W_n\}$ converges pointwise to W. In this case (iii) is obtained combining (i) and (ii). To verify (iv), pick $\varepsilon > 0$ and a fixed $k \geqq n(\varepsilon)$. Next observe that Lemma 3.3 in [13] yields that

$$|M_n - M| \leqq \sup_a \sum_y p_{xy}(a)|W_n(y) - W(y)|$$

$$\leqq \sup_a \sum_{y \in S_k} p_{xy}(a)|W_n(y) - W(y)| + 2B\varepsilon \quad \text{: see (A.2) and (A.3)}$$

$$\leqq \sum_{y \in S_k} |W_n(y) - W(y)| + 2B\varepsilon.$$

Since S_k is finite and $\{W_n\}$ converges pointwise to W, it follows that $\limsup_{n\to\infty} |M_n - M| \leqq 2B\varepsilon$ which yields the conclusion since $\varepsilon > 0$ was arbitrary.

References

[1] ARAPOSTATHIS, A., BORKAR, V. S., FERNÁNDEZ-GAUCHERAND, E., GOSH, M. K. AND MARCUS, S. I. (1993) Discrete-time controlled Markov processes with an average cost criterion: a survey. *SIAM J. Control Optim.* **31**, 282–344.

[2] BORKAR, V. S. (1991) *Topics in Controlled Markov Chains* (Pitman Research Notes in Mathematics Series **240**). Longman, London.

[3] CAVAZOS-CADENA, R. (1991) Recent results on conditions for the existence of average optimal stationary policies. *Ann. Operat. Res.* **28**, 3–28.

[4] CAVAZOS-CADENA, R. (1994) Cesàro convergence of the undiscounted value iteration method in Markov decision processes under the Lyapunov stability condition. *Boletín de la Sociedad Mathemática Mexicana*. To appear.

[5] CAVAZOS-CADENA, R. (1995) Undiscounted value iteration in stable Markov decision processes with bounded rewards. *J. Math. Systems, Estimation and Control.* To appear.

[6] CAVAZOS-CADENA, R. AND FERNÁNDEZ-GAUCHERAND, E. (1995) Denumerable controlled Markov chains with average reward criterion: sample path optimality. *ZOR-Math. Meth. Operat. Res.* **41**, 89–108.

[7] CAVAZOS-CADENA, R. AND HERNÁNDEZ-LERMA, O. (1992) Equivalence of Lyapunov stability criteria in a class of Markov decision processes. *Appl. Math. Optim.* **26**, 113–137.

[8] DUGUNDJI, J. (1966) *Topology.* Allyn and Bacon, Boston.

[9] FEDERGRUEN, A. AND SCHWEITZER, P. J. (1978) Discounted and undiscounted value iteration in Markov decision problems: A survey. In *Dynamic Programming and its Applications.* ed. M. L. Puterman. Academic Press, New York. pp. 23–52.

[10] FEDERGRUEN, A. AND SCHWEITZER, P. J. (1980) A survey of asymptotic value-iteration for undiscounted Markovian decision processes. In *Recent Developments in Markov Decision Processes.* ed. R. Hartley, L. C. Thomas and D. J. White. Academic Press, New York. pp. 73–109.

[11] HERNÁNDEZ-LERMA, O. (1989) *Adaptive Markov Control Processes.* Springer, New York.

[12] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (1993) Value iteration and rolling plans for Markov control processes with unbounded rewards. *J. Math. Anal. Appl.* **177**, 38–55.

[13] HINDERER, K. (1970) *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter.* Springer, New York.

[14] HORDIJK, A. (1974) *Dynamic Programming and Markov Potential Theory.* (Mathematical Centre Tract **51**.) Mathematisch Centrum, Amsterdam.

[15] HORDIJK, A, SCHWEITZER, P. J. AND TIJMS, H. C. (1975) The asymptotic behavior of the minimal total expected cost for the denumerable state Markov decision model. *J. Appl. Prob.* **12**, 298–305.

[16] MONTES-DE-OCA, R. AND HERNÁNDEZ-LERMA, O. (1993) Value iteration in average cost Markov control processes on Borel spaces. *Submitted.*

[17] PUTERMAN, M. (1994) *Markov Decision Processes.* Wiley, New York.

[18] ROYDEN, H. L. (1968) *Real Analysis.* 2nd edn. Macmillan, New York.

[19] ROSS, S. M. (1970) *Applied Probability Models with Optimization Applications.* Holden-Day, San Francisco.

[20] SCHWEITZER, P. J. (1971) Iterative solution of the functional equations for undiscounted Markov renewal programming. *J. Math. Anal. Appl.* **34**, 495–501.

[21] SENNOTT, L. I. (1991) Value iteration in countable state average cost Markov decision processes with unbounded costs. *Ann. Operat. Res.* **28**, 261–272.

[22] THOMAS, L. C. (1965) Connectedness conditions for denumerable state Markov decision processes. In *Recent Developments in Markov Decision Processes.* ed. R. Hartley, L. C. Thomas and D. J. White. Academic Press, New York. pp. 181–204.