MATHEMATICS OF OPERATIONS RESEARCH
Vol. 4, No. 2, May 1979
Printed in U.S.A.

APPROXIMATIONS OF DYNAMIC PROGRAMS, II*†

WARD WHITT‡

Yale University and Bell Laboratories

This paper extends a procedure for approximating dynamic programs due to Fox (1971). Here, the monotone contraction operator model of Denardo (1967) is approximated by replacing the state space with a subset and defining two approximate local income functions so that the two associated approximate optimal return functions serve as lower and upper bounds for the optimal return function in the original model. Conditions are also given implying convergence of a sequence of approximate optimal return functions to the optimal return function in the original model.

1. Introduction and summary. In [8] we introduced a general framework for constructing and analyzing approximations of dynamic programming models. The purpose of this paper is to discuss a different approach, which is an extension of Fox (1971). As in [8], the model is the monotone contraction operator model of Denardo (1967) which includes, for example, Markov decision models with a criterion of discounted present value over an infinite horizon. The approximation is constructed by making the new state space a subset of the original state space. Associated with the smaller state space, we define two approximating local income functions in such a way that the two approximate optimal return functions f^- and f^+ are lower and upper bounds for the original optimal return function f, i.e., $f^- \leqslant f \leqslant f^+$. Moreover, this can be done in such a way that if the approximating state space is enlarged, the new bounds are always at least as good as the old ones, if not better. Finally, we present conditions under which the sequences of optimal return functions associated with a sequence of approximating models converge to the optimal return function in the original model.

Our approximation scheme uses a fixed function to characterize future returns outside the designated subset. The goal is to obtain good decisions for many states inside the subset without examining the behavior outside the subset in detail. We actually discuss three approximation schemes which differ only in the function or functions we use outside the designated subset. Theorems 8–10 apply to an arbitrary function; Theorems 1 and 2 apply to two functions e and g with $e \leqslant f \leqslant g$; and Theorems 3–7 apply to two functions e and g satisfying an extra monotonicity condition, namely, (3.3).

It turns out that the lower approximation is better behaved than the upper in two important ways. First, with our construction, any policy δ^*_- for the original problem which is an extension of a policy attaining the supremum f^- over the designated subset in the lower approximate model has a return function $v_{\delta^*_-}$ satisfying $f^- \leqslant v_{\delta^*_-} \leqslant f \leqslant f^+$; for the upper approximate model we can say only that $v_{\delta^*_+} \leqslant f \leqslant f^+$, cf. Corollary to Theorem 3. Second, it is easier to conclude that a sequence of lower bounds $\{f_k^-, k \geqslant 1\}$ converges to f than a sequence of upper bounds $\{f_k^+, k \geqslant 1\}$, cf.

* Received May 23, 1977; revised June 6, 1978.
AMS 1970 subject classification. Primary 90C40.
IAOR 1973 subject classification. Main: Markov decision programming. Cross reference: Dynamic programming.
Key words. Dynamic programming, Markov decision processes, bounds, approximations.
† This research was initiated in the School of Organization and Management at Yale University and completed at Bell Laboratories.
‡ The author now is associated with Bell Laboratories, Holmdel, New Jersey.

0364-765X/79/0402/0179$01.25

§4. However, both functions are important because together they provide a measure of the error.

Just as with [8], the results here extend to N-stage contraction models, so that corresponding results exist for finite-stage models with nonstationary strategies.

The paper is organized as follows. First, the monotone contraction operator model is defined in §2. Two-sided bounds are established in §3; limit theorems are proved in §4; and a few concluding remarks are made in §5.

Finally, we mention that related research has recently been reported by White (1977, 1978). He establishes a priori bounds for both Fox's (1971) approximation and a successive approximation scheme similar to the one we suggest in Remark 5.1.

2. Monotone contraction operators. We now introduce the monotone contraction operator model of Denardo (1967), modified to allow for unbounded rewards. For further discussion of the treatment of unbounded rewards, see Van Nunen and Wessels (1977) and references there. Let the *state space* S be a nonempty set. For each $s \in S$, let the *action space* be a nonempty set A_s. Let the *policy space* Δ be the Cartesian product of the action spaces. Let $\alpha : S \to (0, \infty)$ and $\beta : S \to R$ be given functions. Let the space V of potential return functions be

$$V = \left\{ v : S \to R \,|\, \sup_{s \in S} |\alpha(s)[v(s) - \beta(s)]| < \infty \right\}. \tag{2.1}$$

Let $\|v\| = \sup\{|v(s)| : s \in S\}$ for any $v \in R^S$, where R^S is the set of all functions mapping S into R. Let

$$d(v_1, v_2) = \|\alpha(v_1 - v_2)\| = \sup\{|\alpha(s)(v_1(s) - v_2(s))| : s \in S\} \tag{2.2}$$

for $v_1, v_2 \in V$. This definition makes (V, d) a complete metric space. Let a *local income function* $h(s, a, v)$ be defined by assigning a real number to each triple (s, a, v) with $s \in S$, $a \in A_s$ and $v \in V$. For each $\delta \in \Delta$, let H_δ be a mapping of V into R^S defined by $(H_\delta v)(s) = h(s, \delta(s), v)$. Assume that the functions H_δ satisfy the basic three properties:

(B) There exist constants K_1 and K_2 such that

$$\|\alpha(H_\delta v - \beta)\| \leqslant K_1 + K_2\|\alpha(v - \beta)\|$$

for all $\delta \in \Delta$ and $v \in V$.

(M) If $v_1 \leqslant v_2$ in V, i.e., if $v_1(s) \leqslant v_2(s)$ for all $s \in S$, $H_\delta v_1 \leqslant H_\delta v_2$ for all $\delta \in \Delta$.

(C) For a fixed c, $0 \leqslant c < 1$,

$$d(H_\delta v_1, H_\delta v_2) \leqslant c d(v_1, v_2)$$

for all $\delta \in \Delta$ and $v_1, v_2 \in V$.

Property (B) implies that the range of H_δ and the ranges of associated extremal operators are contained in V. Property (C) implies that H_δ has a unique fixed point, say v_δ, in V for each $\delta \in \Delta$. The function v_δ is called the *return function* for policy δ. Let $f(s) = \sup\{v_\delta(s) : \delta \in \Delta\}$. The function f is called the *maximal return function*. Let the maximal operator F be defined as $(Fv)(s) = \sup\{(H_\delta v)(s) : \delta \in \Delta\}$. As in [1], the operator F inherits properties B, M, C, and has f as its unique fixed point. Paralleling Theorem 1 of [1], we have

$$d(v_\delta, v) \leqslant (1 - c)^{-1} d(H_\delta v, v) \quad \text{for all } \delta \in \Delta. \tag{2.3}$$

3. Two-sided bounds. Assume that we have found functions e and g in V such that $e \leqslant f \leqslant g$. We shall use the functions e and g together with subsets S_k of the state space S in order to generate approximate models. Let $\gamma(v_1, v_2, S_k)$ be a function in V defined by

$$\gamma(v_1, v_2, S_k)(s) = \begin{cases} v_1(s), & s \in S_k, \\ v_2(s), & s \in S - S_k. \end{cases} \tag{3.1}$$

for any $v_1, v_2 \in V$. For any $v \in V$, let $v_k^+ = \gamma(v, g, S_k)$ and $v_k^- = \gamma(v, e, S_k)$. For each k and $\delta \in \Delta$, define new return operators $H_{k\delta}^+$ and $H_{k\delta}^-$ on V by

$$H_{k\delta}^+ v = (H_\delta v_k^+)_k^+ \quad \text{and} \quad H_{k\delta}^- v = (H_\delta v_k^-)_k^-. \tag{3.2}$$

Also define associated maximal operators $F_k^\pm$, defined by $F_k^\pm v = (Fv_k^\pm)_k^\pm$ or, equivalently, by $F_k^\pm v = \sup\{H_{k\delta}^\pm v, \delta \in \Delta\}$. This is a minor modification of the scheme on page 666 of Fox (1971). While $H_{k\delta}^\pm$ is defined on V with state space S, it is obviously equivalent to a monotone contraction operator model with state space S_k. It is easy to verify that $H_{k\delta}^+$ and $H_{k\delta}^-$ are operators on V satisfying B, M, C for each k and $\delta \in \Delta$. Let $v_{k\delta}^+$, $v_{k\delta}^-$, f_k^+ and f_k^- denote the fixed points of $H_{k\delta}^+$, $H_{k\delta}^-$, F_k^+ and F_k^-, respectively. Just as with f, $f_k^\pm = \sup\{v_{k\delta}^\pm, \delta \in \Delta\}$. To avoid confusion, let $(f)_k^+ = \gamma(f, g_1 S_k)$.

THEOREM 1. $f_k^- \leqslant f \leqslant f_k^+$.

PROOF. Consider only $+$. By property (M), $F_k^+ f = (Ff_k^+)_k^+ \geqslant (Ff)_k^+ = (f)_k^+ \geqslant f$. By property (M) plus induction, $F_k^{+n} f \geqslant f$ for all $n \geqslant 1$. Since $d(F_k^{+n} f, f_k^+) \to 0$ as $n \to \infty$, $f_k^+ \geqslant f$. ∎

The quality of the bounds obviously depends on the functions e and g. Bounds for Theorem 1 are always available using K_1 from condition (B):

THEOREM 2. *For all* $\delta \in \Delta$, $\beta - (1-c)^{-1} K_1 \alpha^{-1} \leqslant v_\delta \leqslant \beta + (1-c)^{-1} K_1 \alpha^{-1}$.

PROOF. Using the triangle inequality, (C) and (B) with $v = \beta$, we obtain

$$\|\alpha(H_\delta^n \beta - \beta)\| \leqslant \sum_{k=1}^{n} \|\alpha(H_\delta^k \beta - H_\delta^{k-1} \beta)\|$$

$$\leqslant \sum_{k=1}^{n} c^{k-1} \|\alpha(H_\delta \beta - \beta)\| \leqslant (1-c)^{-1} K_1,$$

so that $\|\alpha(v_\delta - \beta)\| \leqslant (1-c)^{-1} K_1$. ∎

REMARK. The standard Markov decision model with discounting in which all one-step rewards are bounded by K_1 is covered by Theorem 2 with $\alpha(s) = 1$ and $\beta(s) = 0$ for all $s \in S$.

To get monotone approximations, we now assume the functions e and g satisfy

$$H_{\delta_0} e \geqslant e \text{ for } \textit{some } \delta_0 \in \Delta \quad \text{and} \quad H_\delta g \leqslant g \text{ for } \textit{all } \delta \in \Delta. \tag{3.3}$$

Such functions e and g exist because v_{δ_0} and f are possible assignments, but we are not providing a procedure for finding them. Let $\Delta_e = \{\delta \in \Delta : H_\delta e \geqslant e\}$. Let S_1 and S_2 be subsets of S with $S_1 \subseteq S_2$.

LEMMA 1. *Let e and g be given satisfying* (3.3).
(a) *For any* $v \in V$ *with* $v \leqslant g$ *and* $\delta \in \Delta$, $H_{2\delta}^+ v \leqslant H_{1\delta}^+ v \leqslant g$.
(b) *For any* $v \in V$ *with* $v \geqslant e$ *and* $\delta \in \Delta_e$, $e \leqslant H_{1\delta}^- v \leqslant H_{2\delta}^- v$.

PROOF. (a) Since $v \leqslant g$, $v_2^+ \leqslant v_1^+ \leqslant g$. By (M), $H_\delta v_2^+ \leqslant H_\delta v_1^+ \leqslant H_\delta g \leqslant g$. Therefore,

$$H_{2\delta}^+ v = (H_\delta v_2^+)_2^+ \leqslant (H_\delta v_1^+)_1^+ = H_{1\delta}^+ v$$

and

$$H_{1\delta}^+ v = (H_\delta v_1^+)_1^+ \leqslant (H_\delta g_1^+)_1^+ = (H_\delta g)_1^+ \leqslant g_1^+ = g.$$

(b) The reasoning is similar. ∎

Let $(f)_k^+ = \gamma(\mathrm{f}, g, S_k)$, $f_{ke}^+ = \sup\{v_{k\delta}^-, \delta \in \Delta_e\}$ and $f_e = \sup\{v_\delta, \delta \in \Delta_e\}$.

THEOREM 3. *Let e and g satisfy* (3.3).
(a) *For each* $\delta \in \Delta$, $v_\delta \leqslant v_{2\delta}^+ \leqslant v_{1\delta}^+ \leqslant g$.
(b) *For each* $\delta \in \Delta_e$, $e \leqslant v_{1\delta}^- \leqslant v_{2\delta}^- \leqslant v_\delta$.
(c) $e \leqslant f_{1e}^- \leqslant f_{2e}^- \leqslant f_e \leqslant f \leqslant f_2^+ \leqslant f_1^+ \leqslant g$.

PROOF. (a) By Lemma 1, $H_{2\delta}^+ v_{1\delta}^+ \leqslant H_{1\delta}^+ v_{1\delta}^+ = v_{1\delta}^+$. By Property (M) and induction, $H_{2\delta}^{+n} v_{1\delta}^+ \leqslant v_{1\delta}^+$ for all n. Since $d(H_{2\delta}^{+n} v, v_{2\delta}^+) \to 0$ as $n \to \infty$, $v_{2\delta}^+ \leqslant v_{1\delta}^+$. Similarly, $H_{2\delta}^+ v_\delta = (H_\delta(v_\delta)_2^+)_2^+ \geqslant (H_\delta v_\delta)_2^+ = (v_\delta)_2^+ \geqslant v_\delta$, so that $v_{2\delta}^+ \geqslant v_\delta$. Finally, $H_{1\delta}^+ g \leqslant g$ by Lemma 1, so that $v_{1\delta}^+ \leqslant g$ by the same reasoning. ∎

REMARK. If $e = -(1-c)^{-1}K_1$ and $g = (1-c)^{-1}K_1$ in the standard Markovian decision model with discounting in which all one-step rewards are bounded by K_1 ($\alpha(s) = 1$ and $\beta(s) = 0$ for all s), then $H_\delta e \geqslant e$ and $H_\delta g \leqslant g$ for all $\delta \in \Delta$, so that the functions e and g satisfy (3.3) with $\Delta_e = \Delta$.

Suppose δ_k^- and δ_k^+ are policies in Δ which are extensions of ϵ-optimal policies in the lower and upper approximate models with respect to S_k, which exist by Corollary 1 of [1] extended to the case of unbounded rewards. Obviously $v_{\delta_k^\pm} \geqslant e$ if $\delta_k^\pm \in \Delta_e$, by (3.3). However, note that $v_{\delta_k^+}(s) \geqslant f_k^+(s) - \epsilon/\alpha(s)$ for all $s \in S_k$ need not hold, but

COROLLARY. *Assume* (3.3). *If* $\delta_k^- \in \Delta_e$, *then*

$$f_k^-(s) - \epsilon/\alpha(s) \leqslant v_{\delta_k^-}(s) \leqslant f(s) \leqslant f_k^+(s), \qquad s \in S_k.$$

4. Convergence. Suppose that we have functions e and g satisfying (3.3) and a sequence $\{S_k, k \geqslant 1\}$ of subsets of S such that $S_k \subseteq S_{k+1}$ for all k and $\bigcup_{k=1}^\infty S_k = S$. From Theorem 3, it follows immediately that the sequences $\{v_{k\delta}^-, k \geqslant 1\}$ (for $\delta \in \Delta_e$) and $\{v_{k\delta}^+, k \geqslant 1\}$ converge pointwise on S monotonically to limits v_δ^- and v_δ^+ such that $e \leqslant v_\delta^- \leqslant v_\delta \leqslant v_\delta^+ \leqslant g$. However, we need not have $v_\delta^- = v_\delta = v_\delta^+$; see the first example on page 669 of Fox [2].

THEOREM 4. *If $h(s, \delta(s), v_n) \to h(s, \delta(s), v)$ whenever $v_n \to v$ pointwise monotonically in V, then $v_\delta^+ = v_\delta$ and, if $\delta \in \Delta_e$, $v_\delta^- = v_\delta$.*

PROOF. Consider only $+$. It suffices to show that the unique fixed point of H_δ is v_δ^+. Since $v_{k\delta}^+ = H_{k\delta}^+ v_{k\delta}^+ = (H_\delta v_{k\delta}^+)_k^+$, $v_{k\delta}^+(s) = h(s, \delta(s), v_{k\delta}^+)$ for $s \in S_k$. For any fixed s, $v_{k\delta}^+(s) \to v_\delta^+(s)$ monotonically. By the continuity condition, $h(s, \delta(s), v_{k\delta}^+) \to h(s, \delta(s), v_\delta^+)$. Hence, $H_\delta v_\delta^+ = v_\delta^+$ as desired. ∎

REMARKS. The first example on page 669 of [2] illustrates how the continuity condition in Theorem 4 can fail to hold. Notice that this condition is always satisfied in the affine case (Markov decision model), see Lemma 3 of [2]. The monotonicity is important for treating unbounded rewards.

Let f^- and f^+ be the pointwise-convergent limits of $\{f_k^-\}$ and $\{f_k^+\}$.

THEOREM 5. *Under the conditions of Theorem 4, $f^- = f_e$.*

PROOF. For any $s \in S$ and $\epsilon > 0$, there exists a $\delta^* \in \Delta_e$ such that $v_{\delta^*}(s) \geqslant f_e(s) - \epsilon$. By Theorem 3,

$$\begin{aligned} f_e(s) \geqslant f_{ke}^-(s) \geqslant v_{k\delta^*}^-(s) &\geqslant v_{\delta^*}(s) - |v_{k\delta^*}^-(s) - v_{\delta^*}(s)| \\ &\geqslant f_e(s) - \epsilon - |v_{k\delta^*}^-(s) - v_{\delta^*}(s)|. \end{aligned}$$

By Theorem 4, $|v_{k\delta^*}^-(s) - v_{\delta^*}(s)| \to 0$ as $k \to \infty$, so that $\limsup_{k\to\infty} |f_e(s) - f_{ke}^-(s)| \leqslant \epsilon$. Since ϵ and s were arbitrary, the proof is complete. ∎

REMARK. Obviously Theorem 5 is most useful when $f_e = f$, which certainly occurs when $\Delta_e = \Delta$. See the remark following Theorem 3.

The second example on page 669 of [2] shows that $f^+ = f$ need not hold without extra conditions. Extra conditions are contained in the easily verified

THEOREM 6. *If* $v_{k\delta}^{+}(s) \to v_{\delta}(s)$ *as* $k \to \infty$ *uniformly in* δ *for* $\delta \in \Delta$, *then* $f^{+}(s) = f(s)$.

THEOREM 7. *If, in addition to the condition of Theorem* 4, A_s *is a compact topological space and* $h(s, a, v)$ *is a continuous function of* a, *then* $v_{k\delta}^{+}(s) \to v_{\delta}(s)$ *uniformly in* δ, $\delta \in \Delta$, *and* $v_{k\delta}^{-}(s) \to v_{\delta}(s)$ *uniformly in* δ, $\delta \in \Delta_e$.

PROOF. By Theorem 4 and its proof, for all k sufficiently large, $v_{k\delta}^{+}(s) - v_{\delta}(s) = h(s, \delta(s), v_{k\delta}^{+}) - h(s, \delta(s), v_{\delta}^{+}) \to 0$ monotonically. By Tychonoff's theorem, page 166 of Royden (1968), Δ is compact. By Dini's theorem, page 162 of Royden (1968), and the new conditions, the convergence is uniform in δ. ∎

A stronger mode of convergence than pointwise convergence follows from stronger conditions. What is more important, it is no longer necessary for the functions e and g to satisfy (3.3) or even $e \leqslant f \leqslant g$. Any function w in V can be used outside of S_k. We use $v_{k\delta}$, $(v)_k$, etc. without + or − to indicate that w is used outside S_k. Let

$$\omega(v, \delta, k) = \sup_{s \in S_k} \{\alpha(s) | h(s, \delta(s), (v)_k) - h(s, \delta(s), v)|\}. \tag{4.1}$$

THEOREM 8. *If* $\omega(v, \delta, k) \to 0$ *as* $k \to \infty$ *for each* $v \in V$, *then* $d(v_{k\delta}, (v_{\delta})_k) \to 0$.

PROOF. Consider only +. By (2.3),

$$d(v_{k\delta}, (v_{\delta})_k) \leqslant (1-c)^{-1} d(H_{k\delta}(v_{\delta})_k, (v_{\delta})_k),$$
$$\leqslant (1-c)^{-1} \omega(v_{\delta}, \delta, k),$$

since $H_{k\delta}(v_{\delta})_k = (H_{\delta}(v_{\delta})_k)_k$. ∎

REMARK. Obviously $d(v_{k\delta}, v_{\delta}) \to 0$ is *not* possible.

To obtain corresponding results for the optimal return function, let

$$\omega(v, k) = \sup_{\delta \in \Delta} \omega(v, \delta, k). \tag{4.2}$$

THEOREM 9. *If* $\omega(v, k) \to 0$ *as* $k \to \infty$, *then* $d(f_k, (f)_k) \to 0$.

PROOF. As in the proof of Theorem 8,

$$d(f_k, (f)_k) = d\Big(\sup_{\delta \in \Delta} v_{k\delta}, \sup_{\delta \in \Delta} (v_{\delta})_k\Big)$$
$$\leqslant \sup_{\delta \in \Delta} d(v_{k\delta}, (v_{\delta})_k) \leqslant \omega(v, k). \quad ∎$$

With some extra conditions, there is pointwise convergence of $v_{k\delta}$ to v_{δ} for each δ and pointwise convergence of f_k to f, using the arbitrary fixed function w outside S_k for each k.

THEOREM 10. (a) *If*

(i) S *is countable*,

(ii) H_{δ} *maps* V_0 *into itself, where*

$$V_0 = \{v \in V : \gamma_1(s) \leqslant v(s) \leqslant \gamma_2(s)\}$$

and $\gamma_1(s)$, $\gamma_2(s)$ *are real-valued functions, and*

(iii) $h(s, \delta(s), v_n) \to h(s, \delta(s), v)$ *as* $n \to \infty$ *for each* $s \in S$, $v \in V_0$ *and sequence* $\{v_n, n \geqslant 1\}$ *in* V_0 *converging pointwise to* v,

then $v_{k\delta}$ *converges pointwise to* v_{δ} *as* $k \to \infty$.

(b) *If, in addition,* (ii) *holds for all* $\delta \in \Delta$ *and the convergence in* (iii) *is uniform in* δ, *then* f_k *converges pointwise to* f *as* $k \to \infty$.

PROOF. (a) By (i) and (ii), V_0 with the product topology is a compact metric space. By (ii), $v_{k\delta} \in V_0$ for all k. Hence, every subsequence of $\{v_{k\delta}, k \geqslant 1\}$ has a convergent subsequence. Let v_0 be the limit of some convergent subsequence. By the continuity condition (iii), for k sufficiently large (so that $s \in S_k$), $v_{k\delta}(s) = h(s, \delta(s), v_{k\delta}) \to h(s, \delta(s), v_0)$ as $k \to \infty$, while $v_{k\delta}(s) \to v_0(s)$ as $k \to \infty$, so that $H_\delta v_0 = v_0$. Since v_δ is the unique fixed point of H_δ in V_0, $v_0 = v_\delta$. Since all convergent subsequences of $\{v_{k\delta}\}$ have the same limit, the sequence $\{v_{k\delta}\}$ itself converges to this limit.

(b) Note that

$$\begin{aligned} |f_k(s) - f(s)| &= |\sup_{\delta \in \Delta} v_{k\delta}(s) - \sup_{\delta \in \Delta} v_\delta(s)| \\ &\leqslant \sup_{\delta \in \Delta} |v_{k\delta}(s) - v_\delta(s)| \\ &= \sup_{\delta \in \Delta} |h(s, \delta(s), v_{k\delta}) - h(s, \delta(s), v_\delta)| \to 0. \quad \blacksquare \end{aligned}$$

REMARKS. (1) Condition (ii) of Theorem 10 holds for the discounted Markov decision problem with

$$V_0 = \{v \in V : \|\alpha(v - \beta)\| \leqslant (1 - c)^{-1}(M_1 + M_2)\}$$

if

$$\|\alpha(r_\delta - (1 - c)\beta)\| \leqslant M_1, \qquad \|\alpha(q_\delta \beta - c\beta)\| \leqslant M_2$$

and

$$\|\alpha q_\delta \alpha^{-1}\| = \sup_{s \in S} \left| \alpha(s) \sum_{j=1}^{\infty} (1/\alpha(j)) q_\delta(j \mid s) \right| \leqslant c < 1,$$

where r_δ is the one-step reward function and q_δ is the Markov transition kernel, cf. Lemma 3.2.2 of [4] or Theorem 6.1 of [9].

(2) For the discounted Markov decision problem above with $\alpha(s) = 1$ and $\beta(s) = 0$ for all s, condition (iii) of Theorem 10 always holds. Convergence uniformly in δ for Theorem 10(b) obviously holds if

$$\lim_{m \to \infty} \sup_{\delta \in \Delta} \sum_{j=m}^{\infty} q_\delta(j \mid s) = 0. \tag{4.3}$$

In this case, pointwise convergence of f_k to f has also been established under stronger conditions by D. J. White (1977, 1978) by different methods.

5. Closing remarks. (1) A promising approach is to combine the technique here with successive approximations. For example, we could calculate $F_k^{+\tau_k} F_{k-1}^{+\tau_{k-1}} \cdots F_1^{+\tau_1} g$, where $F_j^+ v = \sup\{H_{j\delta}^+ v : \delta \in \Delta\}$ and τ_j is a positive integer or a positive integer-valued stopping time, as described in [4]. Obviously, $f_1^+ \leqslant F_1^{+\tau_1} g \leqslant g$ and $f_k^+ \leqslant F_k^{+\tau_k} \cdots F_1^{+\tau_1} g \leqslant F_{k-1}^{+\tau_{k-1}} \cdots F_1^{+\tau_1} g$.

(2) We could also work with subsets of the action spaces, but then only the inequalities for the lower approximation are valid.

(3) This approximation scheme applies to two-person zero-sum stochastic games, just as indicated in §V of Fox (1971). Let the local income function $h(s, \delta(s), \xi(s), v) = \int\int h(s, a_1, a_2, v)\delta(s)(da_1)\xi(s)(da_2)$, where $\delta(s)$ and $\xi(s)$, are probability measures on action spaces $A_I(s)$ and $A_{II}(s)$. The associated return operator is $[H_{\delta\xi} v](s) = h(s, \delta(s), \xi(s), v)$. If the initial bounds e and g satisfy $H_{\delta\xi} g \leqslant g$ and $H_{\delta\xi} e \geqslant e$ for all δ and ξ, then $e \leqslant f_k^- \leqslant f \leqslant f_k^+ \leqslant g$ as in Theorem 3, where f is the value. In order to get $f^- = f = f^+$, it suffices to have $A_I(s)$ and $A_{II}(s)$ be compact metric spaces and $h(s, a_1, a_2, v)$ continuous in (a_1, a_2). Then, by basic weak convergence theory, the spaces of probability measures on $A_I(s)$ and $A_{II}(s)$ with the topology of weak

convergence are metrizable as compact metric spaces and $h(s, \delta_n(s), \xi_n(s), v) \to h(s, \delta(s), \xi(s), v)$ whenever $\delta_n(s) \to \delta(s)$ and $\xi_n(s) \to \xi(s)$.

Acknowledgments. I am grateful to the referees for their helpful comments and D. J. White for sending me his unpublished papers. Theorem 10 here was added after seeing these papers.

References

[1] Denardo, E. V. (1967). Contraction Mappings in the Theory Underlying Dynamic Programming. *SIAM Rev.* **9** 165–177.

[2] Fox, B. L. (1971). Finite-state Approximations to Denumerable-state Dynamic Programs. *J. Math. Anal. Appl.* **34** 665–670.

[3] Royden, H. L. (1968). *Real Analysis*, second edition. MacMillan, London.

[4] Van Nunen, J. A. E. E. (1976). *Contracting Markov Decision Processes*. Mathematical Centre Tract No. 71, Amsterdam.

[5] ——— and Wessels, J. (1977). Markov Decision Processes with Unbounded Rewards. *Markov Decision Theory*, Tijms, H. and Wessels, J. (eds.), Mathematical Centre Tract No. 93, Amsterdam, 1–24.

[6] White, D. J. (1977). Finite State Approximations for Denumerable State Infinite Horizon Discounted Markov Decision Processes. Note Number 43, Department of Decision Theory, University of Manchester.

[7] ———. (1978). Finite State Approximations for Denumerable State Infinite Horizon Discounted Markov Decision Processes: The Method of Successive Approximations. Note Number 46, Department of Decision Theory, University of Manchester.

[8] Whitt, W. (1978). Approximations of Dynamic Programs, I. *Math. Oper. Res.* **3** 231–243.

[9] ———. (1978a). Representation and Approximation of Noncooperative Sequential Games. Bell Laboratories, Holmdel, New Jersey.

BELL LABORATORIES, HOLMDEL, NEW JERSEY 07733