MATHEMATICS OF OPERATIONS RESEARCH
Vol. 3, No. 3, August 1978
Printed in U.S.A.

APPROXIMATIONS OF DYNAMIC PROGRAMS, I*†

WARD WHITT

Yale University and Bell Laboratories

A general procedure is presented for constructing and analyzing approximations of dynamic programming models. The models considered are the monotone contraction operator models of Denardo (1967), which include Markov decision processes and stochastic games with a criterion of discounted present value over an infinite horizon plus many finite-stage dynamic programs. The approximations are typically achieved by replacing the original state and action spaces by subsets. Tight bounds are obtained for the distances between the optimal return function in the original model and (1) the extension of the optimal return function in the approximate model and (2) the return function associated with the extension of an optimal policy in the approximate model. Conditions are also given under which the sequence of bounds associated with a sequence of approximating models converges to zero.

1. Introduction and summary. If the state and action spaces in a dynamic programming model are large (infinite, for example), it is often convenient to use an approximate model in order to apply a dynamic programming algorithm to obtain an approximate solution. A natural way to construct an approximate model is to let the new state and action spaces be subsets of the original state and action spaces; then define the new transition and reward structure using the transition and reward structure of the original model. Having defined the smaller model, calculate the optimal return function and optimal policies for the smaller model and use them to define approximately optimal return functions and approximately optimal policies for the original model by a straightforward extension. An interesting question in this setting is: what desirable properties do these extensions have for the original model? It is the purpose of this paper to partially answer this question.

We begin in §2 with a definition of the model to be studied, which is the monotone contraction operator model of Denardo (1967). We indicate how two such models can be compared in §3 and give tight bounds on the difference between the optimal return function in one model and the extensions from the other model. These comparisons can be made when the state and action spaces of one model are subsets of the corresponding state and action spaces of the other model, but also in other circumstances. The special case in which the state and action spaces of one model are in fact subsets of the state and action spaces in the other model is discussed in §4. Several different methods for defining the transition and reward structure in the smaller model are considered. In §5 we prove limit theorems. Under appropriate conditions, a sequence of approximately optimal return functions generated from a sequence of approximate models converges uniformly to the optimal return function in the original model. In §6 we consider a special case of the monotone contraction operator model—the standard stochastic sequential decision model. Finally, extensions are discussed in §7. For example, corresponding results exist for finite-stage dynamic programs, stochastic games and models with unbounded rewards.

* Received June 27, 1975; revised January 12, 1978.
AMS 1970 subject classification. Primary 90C40.
IAOR 1973 subject classification. Main: Markov decision programming. Cross references: Dynamic programming.
Key words. Approximation, aggregation, dynamic programming, monotone contraction operators, fixed points, bounds.
† Partially supported by National Science Foundation Grant GK-38149 in the School of Organization and Management, Yale University.

An account of related work in dynamic programming appears in Hinderer (1978) and Morin (1978). Our work was originally motivated by the discovery of an error in the proof of the theorem in Fox (1973); a minor modification of the methods here provides a new proof. Thomas (1977) has applied the results here to study approximations of capacity expansion models. The results here have been extended by Hinderer (1978), who also treats finite-stage dynamic programs. For related investigations in linear programming, see Zipkin (1977) and references there.

2. Monotone contraction operators. Consider the dynamic programming model introduced by Denardo (1967) with the following notation. Let the *state space* be a nonempty set S. 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. Each element δ in the set Δ is thought of as a *stationary policy*, specifying action $\delta(s)$ to be taken in state s. Let V be the set of all bounded real-valued functions on S with the supremum norm: $\|v\| = \sup\{|v(s)| : s \in S\}$. The essential ingredient in the model specification is the *local income function* h, which assigns a real number to each triple (s, a, v) with $s \in S$, $a \in A_s$ and $v \in V$. The local income function h generates a *return operator* H_δ on V for each $\delta \in \Delta$, i.e., $[H_\delta(v)](s) = h(s, \delta(s), v)$. We make three basic assumptions about the return operators H_δ:

(B) *Boundedness. There exist numbers K_1 and K_2 such that $\|H_\delta v\| \leqslant K_1 + K_2\|v\|$ for all $v \in V$ and $\delta \in \Delta$.*

(M) *Monotonicity. If $v \geqslant u$ in V, i.e., if $v(s) \geqslant u(s)$ for all $s \in S$, then $H_\delta v \geqslant H_\delta u$ in V for all $\delta \in \Delta$.*

(C) *Contraction. For some fixed c, $0 \leqslant c < 1$,*

$$\|H_\delta u - H_\delta v\| \leqslant c\|u - v\|$$

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

The contraction assumption implies that H_δ has a unique fixed point in V for each $\delta \in \Delta$. The unique fixed point of H_δ, denoted by v_δ, is called the *return function* associated with policy δ. Let f denote the *optimal return function*, defined by $f(s) = \sup\{v_\delta(s) : \delta \in \Delta\}$. Let F be the *maximization operator* on V, defined by $[F(v)](s) = \sup\{[H_\delta(v)](s) : \delta \in \Delta\}$. Perhaps the key structural property of this model is that the operator F inherits properties (B, M, C) and has f as its unique fixed point. Call a policy δ *optimal* if $v_\delta = f$ and *ϵ-optional* if $v_\delta(s) \geqslant f(s) - \epsilon$ for all $s \in S$. By Corollary 1 of Denardo (1967), there exists an ϵ-optimal policy for each $\epsilon > 0$. We frequently apply the following basic result, which is Theorem 1 of Denardo (1967).

THEOREM 2.1. *For all $\delta \in \Delta$ and $v \in V$,*

$$\|v_\delta - v\| \leqslant (1-c)^{-1}\|H_\delta v - v\| \quad \textit{and} \quad \|f - v\| \leqslant (1-c)^{-1}\|Fv - v\|.$$

PROOF.

$$\|v_\delta - v\| = \lim_{n\to\infty} \|H_\delta^n v - v\|$$
$$\leqslant \sum_{k=1}^{\infty} \|H_\delta^k v - H_\delta^{k-1} v\| \leqslant \sum_{k=1}^{\infty} c^{k-1}\|H_\delta v - v\|. \quad ∎$$

3. Comparing dynamic programs. Let $(S, \{A_s, s \in S\}, h, c)$ and $(\tilde{S}, \{\tilde{A}_{\tilde{s}}, \tilde{s} \in \tilde{S}\}, \tilde{h}, \tilde{c})$ be two dynamic programming models as defined in §2. We say that these models are *comparable* and that the second is the *image* of the first if the following mappings are defined: (1) a mapping p of S *onto* $\tilde{S}$, (2) a mapping p of A_s *onto* $\tilde{A}_{p(s)}$ for each $s \in S$, (3) a mapping e of $\tilde{S}$ *into* S such that $p(e[\tilde{s}]) = \tilde{s}$ for each $\tilde{s} \in \tilde{S}$ and (4) a mapping e_s of $\tilde{A}_{p(s)}$ *into* A_s such that $p(e_s[\tilde{a}]) = \tilde{a}$ for each $\tilde{a} \in \tilde{A}_{p(s)}$ and $s \in S$. Given two comparable dynamic programs, define additional mappings: (1) $e : \tilde{V} \to V$ with $e(\tilde{v})(s) = \tilde{v}(p(s))$ for each $s \in S$, (2) $p : V \to \tilde{V}$ with $p(v)(\tilde{s}) = v(e[\tilde{s}])$ for each $\tilde{s} \in \tilde{S}$, (3) $e : \tilde{\Delta} \to \Delta$ with $e(\tilde{\delta})(s) = e_s(\tilde{\delta}[p(s)])$ for each $s \in S$ and (4) $p : \Delta \to \tilde{\Delta}$ with $p(\delta)(\tilde{s}) = p[\delta(e[\tilde{s}])]$ for each $\tilde{s} \in \tilde{S}$. Note that $e(\tilde{\delta}) \in \Delta$ for every $\tilde{\delta} \in \tilde{\Delta}$. Note that the composition $p \circ e$ is the identity map on $\tilde{V}$ and $\tilde{\Delta}$, while $e \circ p$ on V and Δ is typically not. The models can be said to be in one-to-one correspondence if $e \circ p$ is in fact the identity map; we shall not dwell on this case. The "distance" between these two models can be expressed in terms of

$$K(v) = \sup_{\substack{a_s \in A_s \\ s \in S}} |h(s, a_s, v) - \tilde{h}(p(s), p(a_s), p(v))|, \quad v \in V, \tag{3.1}$$

with $K(\tilde{v})$ understood to mean $K(e[\tilde{v}])$, $\tilde{v} \in \tilde{V}$. It turns out that $K(\tilde{v})$ is much more useful than $K(v)$.

THEOREM 3.1. $\|p(f) - \tilde{f}\| \leqslant \|f - e(\tilde{f})\| \leqslant (1 - c)^{-1}K(\tilde{f})$.

PROOF. The first inequality is obvious; we consider the second. Since $\tilde{H}_{\tilde{\delta}}\tilde{f} \leqslant \tilde{F}\tilde{f} = \tilde{f}$ for each $\tilde{\delta} \in \tilde{\Delta}$, we can substitute $e(\tilde{f})$ for v in (3.1) to obtain

$$\begin{aligned} [H_\delta e(\tilde{f})](s) = h(s, \delta(s), e(\tilde{f})) &\leqslant \tilde{h}(p(s), p[\delta(s)], \tilde{f}) + K(\tilde{f}) \\ &\leqslant e(\tilde{f})(s) + K(\tilde{f}) \end{aligned}$$

for each $s \in S$ and $\delta \in \Delta$. Then, as a consequence of properties M and C plus induction,

$$[H_\delta^n e(\tilde{f})](s) \leqslant e(\tilde{f})(s) + (1 + c + \cdots + c^{n-1})K(\tilde{f})$$

for all n and δ. Since $\|H_\delta^n v - v_\delta\| \to 0$ as $n \to \infty$,

$$v_\delta(s) \leqslant e(\tilde{f})(s) + (1 - c)^{-1}K(\tilde{f})$$

for all $\delta \in \Delta$, so that

$$f(s) \leqslant e(\tilde{f})(s) + (1 - c)^{-1}K(\tilde{f}).$$

Similarly, for any $\epsilon > 0$, there exists a $\tilde{\delta} \in \tilde{\Delta}$ such that

$$\begin{aligned} [H_{e(\tilde{\delta})} e(\tilde{f})](s) = h(s, e(\tilde{\delta})(s), e(\tilde{f})) &\geqslant \tilde{h}(p(s), \tilde{\delta}[p(s)], \tilde{f}) - K(\tilde{f}) \\ &\geqslant e(\tilde{f})(s) - \epsilon - K(\tilde{f}) \end{aligned}$$

for all $s \in S$. Then, reasoning as above,

$$v_{e(\tilde{\delta})}(s) \geqslant e(\tilde{f})(s) - (1 - c)^{-1}(\epsilon + K(\tilde{f})),$$

so that

$$f(s) \geqslant e(\tilde{f})(s) - (1 - c)^{-1}K(\tilde{f}). \quad \blacksquare$$

THEOREM 3.2. *For any* $\tilde{\delta} \in \tilde{\Delta}$,

$$\|e(\tilde{v}_{\tilde{\delta}}) - v_{e(\tilde{\delta})}\| \leqslant (1 - c)^{-1}K(\tilde{v}_{\tilde{\delta}}).$$

PROOF. Substituting $e(\tilde{\delta})$ for δ and $\tilde{v}_{\tilde{\delta}}$ for $\tilde{v}$ in (3.1) yields

$$\|H_{e(\tilde{\delta})}e(\tilde{v}_{\tilde{\delta}}) - e(\tilde{v}_{\tilde{\delta}})\| \leqslant K(\tilde{v}_{\tilde{\delta}}),$$

which implies the desired conclusion by virtue of Theorem 2.1. $\blacksquare$

LEMMA 3.1. *For all* $\tilde{u}, \tilde{v} \in \tilde{V}$, $|K(\tilde{u}) - K(\tilde{v})| \leq (c + \tilde{c})\|\tilde{u} - \tilde{v}\|$.

PROOF. By the triangle inequality,

$$K(\tilde{u}) \leq \sup_{\substack{\delta \in \Delta \\ s \in S}} \{|h(s, \delta(s), e(\tilde{u})) - h(s, \delta(s), e(\tilde{v}))| + |h(s, \delta(s), e(\tilde{v})) - \tilde{h}(p(s), p[\delta(s)], \tilde{v})| + |\tilde{h}(p(s), p[\delta(s)], \tilde{v}) - \tilde{h}(p(s), p[\delta(s)], \tilde{u})|\},$$

so that $K(\tilde{u}) \leq c\|e(\tilde{u}) - e(\tilde{v})\| + K(\tilde{v}) + \tilde{c}\|\tilde{u} - \tilde{v}\|$. ∎

COROLLARY. *If* $\tilde{\delta}^*$ *is an* ϵ*-optimal policy in* $\tilde{\Delta}$, *then*

$$v_{e(\tilde{\delta}^*)}(s) \leq f(s) \leq v_{e(\tilde{\delta}^*)}(s) + 2(1 - c)^{-1}K(\tilde{f}) + (1 + \tilde{c})(1 - c)^{-1}\epsilon,$$

for all $s \in S$.

PROOF. Since $e(\tilde{\delta}^*) \in \Delta$, $v_{e(\tilde{\delta}^*)} \leq f$. The triangle inequality plus Theorems 3.1 and 3.2 imply that

$$\begin{aligned} \|f - v_{e(\tilde{\delta}^*)}\| &\leq \|f - e(\tilde{f})\| + \|e(\tilde{f}) - e(\tilde{v}_{\tilde{\delta}^*})\| + \|e(\tilde{v}_{\tilde{\delta}^*}) - v_{e(\tilde{\delta}^*)}\| \\ &\leq (1 - c)^{-1}K(\tilde{f}) + \epsilon + (1 - c)^{-1}K(\tilde{v}_{\tilde{\delta}^*}) \\ &\leq 2(1 - c)^{-1}K(\tilde{f}) + [1 + (c + \tilde{c})(1 - c)^{-1}]\epsilon, \end{aligned}$$

with Lemma 3.1 being used in the last step. ∎

REMARKS. (1) It is easy to construct examples in which the inequalities here are equalities.

(2) Note that the bounds here involve $K(\tilde{v})$ rather than $K(v)$. The first part of the proof of Theorem 3.1 can be imitated to obtain $e(\tilde{f})(s) \leq f(s) + (1 - \tilde{c})^{-1}K(f)$, but the second part does not yield corresponding results. More generally, it is easy to construct examples which show that it is *not possible* to bound $\|v_\delta - e(\tilde{v}_{p(\delta)})\|$ or $\|v_\delta - v_{e \circ p(\delta)}\|$ using $K(v)$ in (3.1). Bounds on $\|f - e(\tilde{f})\|$ in terms of f and different distance measure appear in §5 of Hinderer (1978).

(3) It is possible to examine the effect of replacing a good policy δ^* by the "cruder" policy $e \circ p(\delta^*)$. If δ^* is ϵ-optimal, then $v_{e \circ p(\delta^*)}(s) \geq f(s) - (1 - c)^{-1}(3K(f) + 6\epsilon)$.

4. Smaller models. We now consider the special case of two comparable models in which $\tilde{S} \subseteq S$ and $\tilde{A}_{p(s)} \subseteq A_{p(s)}$ for each $s \in S$. We can obtain the maps $p : S \to \tilde{S}$, $p : A_s \to \tilde{A}_{p(s)}$, $e : \tilde{S} \to S$ and $e_s : \tilde{A}_{p(s)} \to A_s$ by constructing partitions of S and A_s for each $s \in S$, and then selecting one point from each partition subset. In particular, let $\{S_i, i \in I\}$ be a partition of subsets of S, i.e., $S = \cup_{i \in I} S_i$ and $S_{i_1} \cap S_{i_2} = \emptyset$ if $i_1 \neq i_2$, with no restriction on the cardinality of I. Let $\tilde{S} = \{s_i, i \in I\}$ be obtained by selecting one point from each subset in the partition of S. For each $i \in I$ and each $s \in S_i$, let $\{A_{sj}, j \in J_i\}$ be partitions of nonempty subsets of A_s, where again there is no restriction on the cardinality of the index sets J_i. We require that the cardinality of the partitions of A_s be the same for all $s \in S_i$, but it may vary with i. Moreover, the jth subset $A_{s_1 j}$ of A_{s_1} is matched with the jth subset $A_{s_2 j}$ of A_{s_2} for all $s_1, s_2 \in S_i$. For $s \in S_i$ and each $i \in I$, let $\tilde{A}_s = \{a_{sj}, j \in J_i\}$ be new action spaces obtained by selecting one point a_{sj} from each subset A_{sj}. Let $\tilde{A}_i$, A_{ij} and a_{ij} represent the quantities $\tilde{A}_{s_i}$, $A_{s_i j}$ and $a_{s_i j}$ associated with states $s_i \in S$. The mappings p and e here can be thought of as *projections* and *extensions*: (1) $p : S \to \tilde{S}$ defined by $p(s) = s_i$ for $s \in S_i$, (2) $p : A_s$

$\to \tilde{A}_{p(s)}$ defined by $p(a_s) = a_{ij}$ if $s \in S_i$ and $a_s \in A_{sj}$, (3) $e : \tilde{S} \to S$ defined by $e(s_i) = s_i$ and (4) $e_s : \tilde{A}_{p(s)} \to A_s$ defined by $e_s(a_{ij}) = a_{sj}$ if $p(s) = s_i$. We have constructed the defining mappings from partitions and identified points in each partition subset, but we can go the other way. The mappings determine partitions and identified points in each partition subset, e.g., $S_i = p^{-1}(s_i) = \{s \in S : p(s) = s_i\}$ for $s_i \in \tilde{S}$ and $a_{sj} = e_s(a_{ij})$ for $p(s) = s_i$.

To complete the definition of the smaller model, we need to define a local income function $\tilde{h}$ such that the induced operators $\tilde{H}_{\tilde{\delta}}$ on $\tilde{V}$ satisfy properties B, M, C. Motivated by the desire to have the small model be a simple approximation for the original model, we are led to the definition

$$\tilde{h}(s_i, a_{ij}, \tilde{v}) = h(s_i, a_{ij}, e(\tilde{v})) \tag{4.1}$$

for all $a_{ij} \in \tilde{A}_i$, $s_i \in \tilde{S}$, $\tilde{v} \in \tilde{V}$. Obviously, the associated return operators $\tilde{H}_{\tilde{\delta}}$ on $\tilde{V}$ satisfy B, M, C with a contraction modulus less than or equal to c.

There are, of course, many other possible definitions for the approximate local income function $\tilde{h}$. For example,

$$\tilde{h}_1(s_i, a_{ij}, \tilde{v}) = 2^{-1}\left\{ \sup_{\substack{a_s \in A_{sj} \\ s \in S_i}} h(s, a_s, e(\tilde{v})) + \inf_{\substack{a_s \in A_{sj} \\ s \in S_i}} h(s, a_s, e(\tilde{v})) \right\} \tag{4.2}$$

and

$$\tilde{h}_2(s_i, a_{ij}, \tilde{v}) = \int_{B_{ij}} h(s, a_s, e(\tilde{v}))\, d\mu_{ij}(s, a_s) \tag{4.3}$$

where $B_{ij} = \{(s, a_s) : s \in S_i,\ a_s \in A_{sj}\}$, μ_{ij} is a probability measure on B_{ij} such as the uniform distribution (when applicable) and the appropriate measurability assumptions are made so that the integral makes sense. Obviously (4.1) is a special case of (4.3). Obviously $K(\tilde{v})$ in (3.1) as a function of $\tilde{h}$ is minimized for every fixed $\tilde{v}$ by (4.2). In both (4.2) and (4.3) properties B, M, C hold with a contraction modulus less than or equal to c.

Recall that the theorems in §3 apply to all such smaller models, but both $\tilde{f}$ and $K(\tilde{v})$ change from model to model, so that the bound $K(\tilde{f})$ varies in an unpredictable way. To obtain a bound applicable to a variety of models, let

$$L(v) = \sup_{\substack{i \in I \\ j \in J_i}} \sup_{\substack{s', s'' \in S_i \\ a_{s'} \in A_{s'j} \\ a_{s''} \in A_{s''j}}} |h(s', a_{s'}, v) - h(s'', a_{s''}, v)| \tag{4.4}$$

for any $v \in V$. As with $K(v)$ in (3.1), we write $L(\tilde{v})$ for $L(e(\tilde{v}))$. As an immediate consequence of (3.1) and (4.4), we obtain

LEMMA 4.1. *If*

$$\inf_{\substack{s \in S_i \\ a_s \in A_{sj}}} h(s, a_s, v) \leqslant \tilde{h}(s_i, a_{ij}, p(v)) \leqslant \sup_{\substack{s \in S_i \\ a_s \in A_{sj}}} h(s, a_s, v) \tag{4.5}$$

for all i and j, then $K(v) \leqslant L(v)$.

Note that (4.5) is satisfied for $v = e(\tilde{v})$ and the choices of $\tilde{h}$ in (4.1)–(4.3).

REMARK. It is interesting that any of these smaller models $(\tilde{S}, \{\tilde{A}_s, s \in \tilde{S}\}, \tilde{h}, \tilde{c})$ can be embedded in the setting of the larger model by setting $\tilde{\delta}'(s) = \tilde{\delta}[p(s)]$ and

$$\tilde{h}'(s, \tilde{\delta}'(s), v) = \tilde{h}(p(s), \tilde{\delta}[p(s)], p(v))$$

for any $s \in S$, $\tilde{\delta} \in \tilde{\Delta}$ and $v \in V$. (Note that $\tilde{\delta}'$ might not be in Δ.) Operators $\tilde{H}'_{\tilde{\delta}'}$ can then be defined on V in the usual way by $[\tilde{H}'_{\tilde{\delta}'}(v)](s) = \tilde{h}'(s, \tilde{\delta}'(s), v)$. This embedding is the symmetric extension introduced in §7 of Denardo [2]. The primed model here is specified completely by $S' = S$, $A'_s = \tilde{A}_{p(s)}$ for $s \in S$, and $\tilde{h}'$ above. It is elementary that $v_{\tilde{\delta}'} = e(v_{\tilde{\delta}})$ for all $\tilde{\delta} \in \tilde{\Delta}$, cf. Theorem 5 of [2]. Hence, the new model $(S', \{A'_s, s \in S'\}, \tilde{h}')$ is a representation of the smaller model $(\tilde{S}, \{\tilde{A}_i, i \in I\}, \tilde{h})$ with the same state space as the larger model. Moreover, the theorems in §3 comparing the models $(S, \{A_s, s \in S\}, h)$ and $(\tilde{S}, \{\tilde{A}_i, i \in I\}, \tilde{h})$ are natural generalizations of the symmetry theorem in §7 of [2], comparing the models $(S', \{A'_s, s \in S'\}, \tilde{h}')$ and $(\tilde{S}, \{\tilde{A}_i, i \in I\}, \tilde{h})$.

5. Limit theorems. Given a dynamic programming model as defined in §2, we should expect that it is possible to construct a sequence of approximating smaller models as defined in §4 such that the smaller models become better and better approximations for the original models as the partitions associated with the smaller models become finer and finer. In particular, we should hope that the sequences of return functions $\{e_n(\tilde{f}_n)\}$ and $\{v_{e_n(\tilde{\delta}_n^*)}\}$ generated from the sequence of smaller models converge to the optimal return function f. The results in this section apply to all the local income functions in (4.1)–(4.3); we only assume $\tilde{h}$ has modulus c and (4.5) is satisfied for all $\tilde{v} \in \tilde{V}$. In fact, it is only used for $\tilde{v} = p(f)$.

THEOREM 5.1. *If* (4.5) *holds for* $v = e_n \circ p_n(f)$ *for all* n, *then there exists a sequence of finite partitions of* S *and* A_s, $s \in S$, *such that* $\lim_{n\to\infty} L_n(\tilde{f}_n) = 0$.

REMARKS. (1) The proof is constructive, but the construction is of limited practical value because the optimal return function f is used. This does suggest a heuristic method: first make a rough estimate of f and then apply the construction using it. For example, use $e_{n-1}(\tilde{f}_{n-1})$ to define the partitions in the nth approximating model with $\tilde{f}_n$.

(2) By Theorem 3.1 and Lemma 4.1, Theorem 5.1 implies that $\|f - e_n(\tilde{f}_n)\| \to 0$, but this is shown directly in the proof.

The proof uses the following lemma, which applies to a single approximate model. For any $v \in V$, let

$$\omega(v) = \sup\{|v(s') - v(s'')| : s', s'' \in S_i, i \in I\}. \tag{5.1}$$

LEMMA 5.1. *Assume* (4.5). *If* $\omega(f) \leq \epsilon$ *and* $\tilde{\delta}'$ *is any policy in* $\tilde{\Delta}$ *such that*

$$\sup\{|h(s, a_s, f) - f(s_i)| : i \in I, s \in S_i, a_s \in A_{sj}, \tilde{\delta}'(s_i) = a_{ij}\} \leq \epsilon,$$

then

(a) $\|\tilde{v}_{\tilde{\delta}'} - p(f)\| \leq (1 + c)(1 - c)^{-1}\epsilon$,

(b) $\|\tilde{f} - p(f)\| \leq (1 + c)(1 - c)^{-1}\epsilon$, *and*

(c) $\|e(\tilde{f}) - f\| \leq 2(1 - c)^{-1}\epsilon$.

PROOF OF LEMMA 5.1. (a) To simplify notation, let Sup represent the supremum over $i \in I$, $s \in S_i$ and $a_s \in A_{sj}$, where $\tilde{\delta}'(s_i) = a_{ij}$. By (4.5) for $v = e \circ p(f)$, the triangle inequality and the two conditions,

$$\begin{aligned}\|\tilde{H}_{\tilde{\delta}'}p(f) - p(f)\| &\leq \text{Sup}|h(s, a_s, e \circ p(f)) - f(s_i)| \\ &\leq \text{Sup}|h(s, a_s, e \circ p(f)) - h(s, a_s, f)| + \text{Sup}|h(s, a_s, f) - f(s_i)| \\ &\leq c\|e \circ p(f) - f\| + \epsilon \leq c\omega(f) + \epsilon \leq (1 + c)\epsilon,\end{aligned}$$

which, with Theorem 2.1, implies that

$$\|\tilde{v}_{\tilde{\delta}'} - p(f)\| \leq (1 - c)^{-1}\|\tilde{H}_{\tilde{\delta}'}p(f) - p(f)\| \leq (1 + c)(1 - c)^{-1}\epsilon.$$

(b) By (4.5) and the first condition, for any $\tilde{\delta} \in \tilde{\Delta}$ with $\tilde{\delta}(s_i) = a_{ij}$,

$$[\tilde{H}_{\tilde{\delta}} p(f)](s_i) \leqslant \sup_{\substack{s \in S_i \\ a_s \in A_{sj}}} h(s, a_s, e \circ p(f)) \leqslant \sup_{\substack{s \in S_i \\ a_s \in A_{sj}}} h(s, a_s, f) + c\epsilon$$

$$\leqslant \sup_{s \in S_i} f(s) + c\epsilon \leqslant p(f)(s_i) + (1 + c)\epsilon,$$

so that, by the argument used in Theorem 3.1,

$$\tilde{v}_{\tilde{\delta}}(s_i) \leqslant p(f)(s_i) + (1 + c)(1 - c)^{-1}\epsilon.$$

Since $\tilde{\delta}$ was arbitrary,

$$\tilde{f}(s_i) \leqslant p(f)(s_i) + (1 + c)(1 - c)^{-1}\epsilon.$$

On the other hand, by part (a),

$$\tilde{f}(s_i) \geqslant \tilde{v}_{\tilde{\delta}'}(s_i) \geqslant p(f)(s_i) - (1 + c)(1 - c)^{-1}\epsilon.$$

(c) By the triangle inequality,

$$\|e(\tilde{f}) - f\| \leqslant \|e(\tilde{f}) - e \circ p(f)\| + \|e \circ p(f) - f\| \leqslant \|\tilde{f} - p(f)\| + \omega(f). \quad ∎$$

PROOF OF THEOREM 5.1. Since $f \in V$, property B implies there exists a constant K such that $|h(s, \delta(s), f)| \leqslant K$ for all $s \in S$ and $\delta \in \Delta$. Hence, for each $n \geqslant 1$, a finite partition of nonempty subsets of the set of all state-action pairs can be constructed by letting

$$B_k = \{(s, a) : s \in S, a \in A_s, k/n \leqslant h(s, a, f) < (k + 1)/n\}, \qquad -\infty < k < \infty,$$

where we suppress the dependence on n. For each $s \in S$ and subset B_k, define the s-section of B_k in the usual way as

$$(B_k)_s = \{a \in A_s : (s, a) \in B_k\}.$$

Associate with each state s the necessarily finite set I_s of indices for which the s-sections are nonempty, i.e.,

$$I_s = \{k : (B_k)_s \neq \emptyset\}.$$

Define an equivalence relation on S by saying that s_1 is equivalent to s_2 if $I_{s_1} = I_{s_2}$. Let $\{S_1, \ldots, S_m\}$ be the finite partition of equivalence classes in S. Form the state space $\tilde{S}$ of the smaller model by selecting one point s_i from each subset S_i of this partition of S. For each i and $s \in S_i$, use the subcollection of B_k subsets for $k \in I_{s_i}$ to form the partitions of the action space A_s. For example, suppose one such subcollection has been relabeled as $\{B_j, 1 \leqslant j \leqslant k\}$. Then let

$$A_{sj} = \{a \in A_s : (s, a) \in B_j\}, \qquad 1 \leqslant j \leqslant k, s \in S_i.$$

Then select one point a_{sj} from each subset A_{sj} associated with the subcollection $\{B_j, 1 \leqslant j \leqslant k\}$. Let the projection be defined as $p(s) = s_i$ and $p(a_s) = a_{s_i j}$ if $s \in S_i$ and $a_s \in A_{sj}$. After reintroducing n, this construction yields $L_n(f) \leqslant n^{-1}$, cf. (4.4), $\omega_n(f) \leqslant n^{-1}$, cf. (5.1), and

$$\sup\{|h(s, a_s, f) - f(s_i)| : i \in I, s \in S_i, a_s \in A_{sj}, \tilde{\delta}'(s_i) = a_{ij}\} \leqslant n^{-1},$$

where $\tilde{\delta}'(s_i)$ attains the maximum over j of $h(s_i, a_{ij}, f)$, so that both conditions in Lemma 5.1 hold with $\epsilon \leqslant n^{-1}$. Finally,

$$L_n(\tilde{f}_n) \leqslant L_n(f) + 2c\|f - e_n(\tilde{f}_n)\| \to 0 \quad \text{as } n \to \infty. \quad ∎$$

In order to guarantee convergence of sequences of approximately optimal return functions associated with a sequence of approximating finite models via Theorem 5.1, great care must be taken in the choice of partitions. With appropriate continuity and compactness, the approximation scheme is relatively insensitive to the specific choice of partitions. In the following, there is only one action space A and the same partition of A is used for all $s \in S$.

THEOREM 5.2. *If*

(i) *S and A are compact metric spaces,*

(ii) *$h(\cdot, \cdot, v)$ is continuous for each continuous v in V, and*

(iii) *for each $n \geq 1$, $\{S_{n1}, \ldots, S_{nk_n}\}$ and $\{A_{n1}, \ldots, A_{nm_n}\}$ are finite partitions of S and A such that the subsets S_{ni} and A_{nj} are all contained in ϵ_n-balls, where $\epsilon_n \to 0$ as $n \to \infty$, then $\lim_{n\to\infty} L_n(\tilde{f}_n) = 0$.*

PROOF. Since $h(\cdot, \cdot, v)$ is continuous for each continuous v, $F(v)(\cdot) = \sup_{\delta \in \Delta} H_\delta v(\cdot)$ is continuous for each continuous v. Hence, F maps the closed subset of continuous functions in V into itself, so that the unique fixed point f of F is continuous. Hence, f is uniformly continuous on S and $h(\cdot, \cdot, f)$ is uniformly continuous on $S \times A$, i.e., for any $\epsilon > 0$, there exists an $\epsilon' > 0$ such that if $\{S_1, \ldots, S_m\}$ and $\{A_1, \ldots, A_n\}$ are finite partitions of S and A such that the diameter of each partition subset is less than ϵ', then

$$\sup\{h(s, a, f) - h(s', a', f)| : s, s' \in S_i, a, a' \in A_j, i \in I, j \in J\} \leq \epsilon.$$

The rest of the proof follows the proof of Theorem 5.1. ∎

REMARK. By virtue of Theorem 3.1, the Corollary to Theorem 3.2 and Lemma 4.1, the conclusion in Theorems 5.1 and 5.2 implies that $\|f - e_n(\tilde{f}_n)\| \to 0$ and $\|f - v_{e_n(\tilde{\delta}_n^*)}\| \to 0$ and $n \to \infty$.

6. A stochastic sequential decision model. We now consider a special case. Let the sets S and A_s, $s \in S$, be endowed with σ-fields; let $r(s, a)$ be a real number bounded in absolute value by M for each $s \in S$ and $a \in A_s$; let $q(\cdot \mid s, a)$ be a subprobability measure on S for each $s \in S$ and $a \in A_s$; and let the local income function be

$$h(s, a, v) = r(s, a) + c \int_S v(x) q\,(dx \mid s, a),$$

where the integral is an abstract Lebesgue integral if v is measurable and an upper integral otherwise, cf. Example 3 of §8 in Denardo (1967). Given such a stochastic sequential decision model, consider finite or countably infinite partitions of measurable subsets of S and A_s for each $s \in S$, according to the scheme of §4. (We could also express everything in the more general setting of §3.) Let $\tilde{r}(s_i, a_{ij})$ be a real number bounded in absolute value by M for each i, j; let $\tilde{q}(\cdot \mid s_i, a_{ij})$ be a subprobability measure on $\tilde{S}$ for each i, j; and let the local income function $\tilde{h}$ be defined by

$$\tilde{h}(s_i, a_{ij}, \tilde{v}) = \tilde{r}(s_i, a_{ij}) + c \sum_n \tilde{v}(s_n) \tilde{q}(\{s_n\} \mid s_i, a_{ij}).$$

In order to compare the models, let

$$K_r = \sup_{\substack{s \in S \\ a_s \in A_s}} |r(s, a_s) - \tilde{r}(p(s), p(a_s))|,$$

$$K_q = \sup_{\substack{s \in S \\ a_s \in A_s}} \sum_{n=1}^{\infty} |q(S_n \mid s, a_s) - \tilde{q}(\{s_n\} \mid p(s), p(a_s))|, \tag{6.1}$$

$$\gamma(v) = \sup_{s \in S} v(s) - \inf_{s \in S} v(s), \qquad v \in V, \quad \text{and}$$

$$\gamma(\tilde{v}) = \gamma(e(\tilde{v})), \qquad \tilde{v} \in \tilde{V}.$$

We now obtain a bound for $K(\tilde{v})$ in (3.1) using K_r, K_q and γ in (6.1).

THEOREM 6.1. (a) *For any* $\tilde{v} \in \tilde{V}$, $K(\tilde{v}) \leqslant K_r + c\|\tilde{v}\|K_q$.
(b) *If* $q(S \mid s, a_s) = \tilde{q}(\tilde{S} \mid p(s), p(a_s)) = 1$ *for all* $a_s \in A_s$ *and* $s \in S$, *then*

$$K(\tilde{v}) \leqslant K_r + c\gamma(\tilde{v})K_q/2.$$

Theorem 6.1 yields an obvious bound on $K(\tilde{f})$ which does not require that we solve either problem. For this a priori bound, let

$$\gamma_{\tilde{r}} = \sup_{\substack{s \in S \\ a_s \in A_s}} \tilde{r}(p(s), p(a_s)) - \inf_{\substack{s \in S \\ a_s \in A_s}} \tilde{r}(p(s), p(a_s)). \tag{6.2}$$

COROLLARY. (a) $K(\tilde{f}) \leqslant K_r + c(1-c)^{-1}MK_q$.
(b) *If* $q(S \mid s, a_s) = \tilde{q}(\tilde{S} \mid p(s), p(a_s)) = 1$ *for all* $a_s \in A_s$ *and* $s \in S$, *then*

$$K(\tilde{f}) \leqslant K_r + c(1-c)^{-1}\gamma_{\tilde{r}}K_q/2.$$

PROOF OF THE COROLLARY. (a) Note that $\|\tilde{v}_{\tilde{\delta}}\| \leqslant (1-c)^{-1}\|\tilde{r}(\cdot, \tilde{\delta}(\cdot))\| \leqslant (1-c)^{-1}M$. (b) Note that $\gamma(\tilde{f}) \leqslant (1-c)^{-1}\gamma_{\tilde{r}}$. ∎

In the proofs of Theorems 6.1 and 6.2 we use the fact that the upper integral $\int_B v\,d\mu$ is countably additive as a function of B. We first state a lemma which does not exploit partitions.

LEMMA 6.1. *If* μ_1 *and* μ_2 *are two finite measures on* S, *then*

$$\left|\int_S v\,d\mu_1 - \int_S v\,d\mu_2\right| \leqslant \gamma(v)\min\{\mu_1(S), \mu_2(S)\} + \|v\|(|\mu_1(S) - \mu_2(S)|).$$

PROOF. Observe that the upper integrals satisfy

$$\begin{aligned}
\int_S v\,d\mu_1 - \int_S v\,d\mu_2 &\leqslant \mu_1(S)\sup_{s\in S} v(s) - \mu_2(S)\inf_{s\in S} v(s) \\
&\leqslant \min\{\mu_1(S), \mu_2(S)\}\Big(\sup_{s\in S} v(s) - \inf_{s\in S} v(s)\Big) \\
&\quad + (\mu_1(S) - \min\{\mu_1(S), \mu_2(S)\})\sup_{s\in S} v(s) \\
&\quad - (\mu_2(S) - \min\{\mu_1(S), \mu_2(S)\})\inf_{s\in S} v(s) \\
&\leqslant \gamma(v)\min\{\mu_1(S), \mu_2(S)\} + \|v\|(|\mu_1(S) - \mu_2(S)|).
\end{aligned}$$

To obtain the inequality in the other direction, change the subscripts of μ_1 and μ_2. ∎

We now exploit the partitions. For this purpose, let

$$\begin{aligned}
\gamma_n(v) &= \sup_{s\in S_n} v(s) - \inf_{s\in S_n} v(s) \quad \text{and} \\
\|v\|_n &= \sup_{s\in S_n} |v(s)|.
\end{aligned} \tag{6.3}$$

Note that $\omega(v) = \sup_n \gamma_n(v)$ for ω in (5.1). Let

$$K_\mu = \sum_{n=1}^{\infty} |\mu_1(S_n) - \mu_2(S_n)|,$$

where μ_1 and μ_2 are finite measures on S.

LEMMA 6.2. (a) *If* μ_1 *and* μ_2 *are two finite measures on* S, *then*

$$\begin{aligned}
\left|\int_S v\,d\mu_1 - \int_S v\,d\mu_2\right| &\leqslant \sum_n \big[\gamma_n(v)\min\{\mu_1(S_n), \mu_2(S_n)\} \\
&\qquad + \|v\|_n(|\mu_1(S_n) - \mu_2(S_n)|)\big] \\
&\leqslant \omega(v)\big[(\mu_1(S) + \mu_2(S) - K_\mu)/2\big] + \|v\|K_\mu.
\end{aligned}$$

(b) *If also* $\mu_1(S) = \mu_2(S)$, *then*

$$\left|\int_S v\,d\mu_1 - \int_S v\,d\mu_2\right| \leqslant \sum_n \gamma_n(v)\min\{\mu_1(S_n), \mu_2(S_n)\} + \gamma(v)K_\mu/2$$
$$\leqslant \omega(v)(\mu_1(S) - [K_\mu/2]) + \gamma(v)K_\mu/2.$$

PROOF. (a) Apply the triangle inequality with Lemma 6.1, using the fact that $\min\{x, y\} = (x + y - |x - y|)/2$ in the last step.

(b) Apply the proof of Lemma 6.1 on the partition subsets to obtain

$$\int_S v\,d\mu_1 - \int_S v\,d\mu_2 = \sum_n \left(\int_{S_n} v\,d\mu_1 - \int_{S_n} v\,d\mu_2\right)$$
$$\leqslant \sum_n \min\{\mu_1(S_n), \mu_2(S_n)\}\left(\sup_{s\in S_n} v(s) - \inf_{s\in S_n} v(s)\right)$$
$$+ \sum_n (\mu_1(S_n) - \min\{\mu_1(S_n), \mu_2(S_n)\}) \sup_{s\in S_n} v(s)$$
$$- \sum_n (\mu_2(S_n) - \min\{\mu_1(S_n), \mu_2(S_n)\}) \inf_{s\in S_n} v(s)$$
$$\leqslant \sum_n \gamma_n(v)\min\{\mu_1(S_n), \mu_2(S_n)\} + \gamma(v)K_\mu/2.$$

Since $\gamma_n(v) \leqslant \omega(v)$ and $\min\{x, y\} = (x + y - |x - y|)/2$, the second inequality in (b) holds too. ∎

PROOF OF THEOREM 6.1. (a) Note that

$$|h(s, \delta(s), e(\tilde{v})) - \tilde{h}(p(s), p[\delta(s)], \tilde{v})|$$
$$\leqslant K_r + c\left|\int_S e(\tilde{v})(x)q(dx \mid s, \delta(s)) - \int_S e(\tilde{v})(x)\tilde{q}(dx \mid p(s), p[\delta(s)])\right|$$
$$\leqslant K_r + c\|\tilde{c}\|K_q,$$

by Lemma 6.2(a), since $\omega(e(\tilde{v})) = 0$.

(b) Apply Lemma 6.2(b) above. ∎

It is easy to construct examples in which all the inequalities here are equalities. Even though the bounds in Theorems 3.2 and 6.1 are tight, combining them does not yield a tight bound for $\|v_{e(\tilde{\delta})} - e(\tilde{v}_{\tilde{\delta}})\|$. The extra structure of the stochastic sequential decision model yields a better bound. Suppose $\gamma_r = M_r - m_r$, where

$$m_r \leqslant r(s, a_s) \leqslant M_r, \qquad \text{for all } s \in S,\ a_s \in A_s,$$
$$m_r \leqslant \tilde{r}(p(s), p(a_s)) \leqslant M_r, \quad \text{for all } s \in S,\ a_s \in A_s. \tag{6.4}$$

THEOREM 6.2. *If* $q(S \mid s, \delta(s)) = \tilde{q}(\tilde{S} \mid p(s), p[\delta(s)]) = 1$ *for all* $s \in S$ *and* $\delta \in \Delta$, *then*

(a) $\|v_{e(\tilde{\delta})} - e(\tilde{v}_{\tilde{\delta}})\| \leqslant \alpha \equiv [1 - c(1 - K_q/2)]^{-1}(K_r + c(1 - c)^{-1}\gamma_r K_q/2)$ *for all* $\tilde{\delta} \in \tilde{\Delta}$, *and*

(b) $\tilde{f}(p(s)) \leqslant f(s) + \alpha$ *for all* $s \in S$.

PROOF. (a) We use the following extension of Lemma 6.1 for measures μ_1 and μ_2 and bounded functions v_1 and v_2:

$$\int_S v_1\,d\mu_1 - \int_S v_2\,d\mu_2 \leqslant \left(\sup_{s\in S} v_1(s) - \inf_{s\in S} v_2(s)\right)\min\{\mu_1(S), \mu_2(S)\}$$
$$+ \sup_{s\in S} v_1(s)(\mu_1(S) - \min\{\mu_1(S), \mu_2(S)\})$$
$$- \inf_{s\in S} v_2(s)(\mu_2(S) - \min\{\mu_1(S), \mu_2(S)\}).$$

Moreover, if v_2 is constant, then

$$\sup_{s\in S} v_1(s) - \inf_{s\in S} v_2(s) \leqslant \|v_1 - v_2\|.$$

Applying these inequalities on the partition subsets, we obtain

$$\begin{aligned} v_{e(\tilde\delta)}(s) - e(\tilde v_{\tilde\delta})(s) &= r(s, e(\tilde\delta)(s)) + c\int_S v_{e(\tilde\delta)}(x) q(dx \mid s, e(\tilde\delta)(s)) \\ &\quad - \tilde r(p(s), \tilde\delta[p(s)]) - c\sum_{n=1}^{\infty} \tilde v_{\tilde\delta}(s_n)\tilde q(\{s_n\} \mid p(s), \tilde\delta[p(s)]) \\ &\leqslant K_r + c\|v_{e(\tilde\delta)} - e(\tilde v_{\tilde\delta})\| Q(s,\tilde\delta) + c\sup_{s\in S} v_{e(\tilde\delta)}(s)(1 - Q(s,\tilde\delta)) \\ &\quad - c\inf_{s\in S} e(\tilde v_{\tilde\delta})(s)(1 - Q(s,\tilde\delta)), \end{aligned}$$

where

$$\begin{aligned} Q(s,\tilde\delta) &= \sum_{n=1}^{\infty} \min\{q(S_n \mid s, e(\tilde\delta)(s)), \tilde q(\{s_n\} \mid p(s), \tilde\delta[p(s)])\} \\ &= 1 - 2^{-1}\sum_{n=1}^{\infty}\left[q(S_n \mid s, e(\tilde\delta)(s)) - \tilde q(\{s_n\} \mid p(s), \tilde\delta[p(s)])\right]. \end{aligned}$$

Since $Q(s,\tilde\delta) \geqslant 1 - K_q/2$,

$$(1-c)^{-1}m_r \leqslant v_{e(\tilde\delta)}(s) \qquad \text{and} \qquad e(\tilde v_{\tilde\delta})(s) \leqslant (1-c)^{-1}M_r,$$

$$\begin{aligned} v_{e(\tilde\delta)}(s) - e(\tilde v_{\tilde\delta})(s) &\leqslant K_r + c(1-c)^{-1}\gamma_r \\ &\quad - c\big((1-c)^{-1}\gamma_r - \|v_{e(\tilde\delta)} - e(\tilde v_{\tilde\delta})\|\big)Q(s,\tilde\delta) \\ &\leqslant K_r + c(1-c)^{-1}\gamma_r - c\big((1-c)^{-1}\gamma_r - \|v_{e(\tilde\delta)} - e(\tilde v_{\tilde\delta})\|\big)(1 - K_q/2). \end{aligned}$$

Similarly, $e(\tilde v_{\tilde\delta})(s) - v_{e(\tilde\delta)}(s)$ has the same upper bound, so that

$$\big(1 - c[1 - K_q/2]\big)\|v_{e(\tilde\delta)} - e(\tilde v_{\tilde\delta})\| \leqslant K_r + c(1-c)^{-1}\gamma_r K_q/2.$$

(b) For any $\epsilon > 0$, there is a $\tilde\delta \in \tilde\Delta$ such that $\tilde v_{\tilde\delta}(p(s)) \geqslant \tilde f(p(s)) - \epsilon$. Then

$$\tilde f(p(s)) \leqslant \tilde v_{\tilde\delta}(p(s)) + \epsilon \leqslant v_{e(\tilde\delta)}(s) + \alpha + \epsilon \leqslant f(s) + \alpha + \epsilon. \quad \blacksquare$$

REMARKS. It is easy to construct an example showing that the bound in Theorem 6.2 is tight. We have not yet determined a tight upper bound for $f(s) - \tilde f(p(s))$.

7. Extensions. (1) The contraction assumption (C) can be replaced by the N-stage contraction assumption, cf. §5 of Denardo (1967). For example, suppose $\|H_\delta u - H_\delta v\| \leqslant m\|u - v\|$ and $\|H_\delta^N u - H_\delta^N v\| \leqslant c\|u - v\|$ for all $\delta \in \Delta$, where $c < 1$ but not necessarily $m < 1$; then the bound in Theorem 3.1 becomes $(1 + m + \cdots + m^{N-1})(1-c)^{-1}K(\tilde f)$. This extension includes many N-stage dynamic programs (usually with $m = 1$ and $c = 0$). The representation of finite-stage dynamic programs as monotone contraction operator models is facilitated by including the stage as part of the state description. For example, after this modification, there is no loss of generality in considering only stationary policies. Approximations of finite-stage dynamic programs are studied directly by Hinderer (1978).

(2) Corresponding theorems hold for the case of unbounded rewards, cf. Wessels (1977). For example, suppose $b : S \to R$, $\mu : S \to (0, \infty)$,

$$V = \left\{ v : S \to R \mid \sup_{s\in S} |[v(s) - b(s)]/\mu(s)| < \infty \right\}$$

and

$$d(u, v) = \sup_{s \in S} |[u(s) - v(s)]/\mu(s)|, \qquad u, v \in V.$$

Suppose the operators H_δ satisfy properties (B, M, C) on the complete metric space (V, d). Let $\tilde{V}$ and $\tilde{d}$ be defined just as V, d with $\tilde{S}$ instead of S and suppose that $e(\tilde{v}) \in V$ for all $\tilde{v} \in \tilde{V}$. Let the operators $\tilde{H}_{\tilde{\delta}}$ satisfy properties (B, M, C) on $(\tilde{V}, \tilde{d})$. Then Theorem 3.1 extends: $d(f, e(\tilde{f})) \leqslant (1 - c)^{-1} K(\tilde{f})$, where

$$K(v) = \sup_{\substack{s \in S \\ a_s \in A_s}} |(h(s, a_s, v) - \tilde{h}(p(s), p(a_s), p(v)))/\mu(s)|.$$

(3) The results here extend to zero-sum stochastic games, cf. Example 2 of Denardo (1967), which is very important because with randomized strategies the policy spaces are invariably large. Theorem 3.1 remains valid if f and $\tilde{f}$ represent the value of the stochastic games and

$$K(\tilde{v}) = \sup_{\substack{s \in S \\ \delta \in \Delta \\ \gamma \in \Gamma}} |h(s, \delta(s), \gamma(s), e(\tilde{v})) - \tilde{h}(p(s), p[\delta(s)], p[\gamma(s)], \tilde{v})|,$$

where $\delta(s)$ and $\gamma(s)$ ($\tilde{\delta}(\tilde{s})$ and $\tilde{\gamma}(\tilde{s})$) are the randomized policies of players I and II in the original (small) model,

$$h(s, \delta(s), \gamma(s), v) = \int_{A_s} \int_{B_s} h(s, a, b, v) \delta(s)(da) \gamma(s)(db),$$

A_s and B_s are the action spaces for players I and II in state s and

$$h(s, a, b, v) = r(s, a, b) + c \int_S v(x) q(dx \mid s, a, b).$$

To see how the analogue of Theorem 3.1 can be proved, suppose $\tilde{\delta}^*$ and $\tilde{\gamma}^*$ are optimal policies for players I and II in the small model. Then, for any $\delta \in \Delta$,

$$\begin{aligned}[H_{\delta, e(\tilde{\gamma}^*)} e(\tilde{f})](s) &= h(s, \delta(s), e(\tilde{\gamma}^*), e(\tilde{f})) \\ &\leqslant \tilde{h}(p(s), p[\delta(s)], \tilde{\gamma}^*(p(s)), \tilde{f}) + K(\tilde{f}) \leqslant e(\tilde{f})(s) + K(\tilde{f}),\end{aligned}$$

so that

$$v_{\delta, e(\tilde{\gamma}^*)}(s) \leqslant e(\tilde{f})(s) + (1 - c)^{-1} K(\tilde{f}).$$

Similarly, for any $\gamma \in \Gamma$,

$$v_{e(\tilde{\delta}^*), \gamma}(s) \geqslant e(\tilde{f})(s) - (1 - c)^{-1} K(\tilde{f}).$$

The convergence results can be used to prove that a stochastic game with large state and action spaces has a value. Approximations of noncooperative sequential games are discussed in Whitt (1977).

(4) Our purpose was to define and analyze deliberate approximations, but the second model could arise in other ways, for example, because of lack of information. When the system is in state s, the decision maker may only know that the system is in some subset $p(s)$ of S or the decision maker may only have some probability distribution $p(s)$ on the set S of possible states. The distance $\|f - e(\tilde{f})\|$ may be considered the value of information in this context.

(5) The approach in the proof of Theorem 5.1 yields a new proof of a Dubins-Savage measurable selection theorem, cf. §9 of Wagner (1977) and Whitt (1976).

Acknowledgment. I am grateful to Eric Denardo, Karl Hinderer, Arie Hordijk, Uriel Rothblum and Frank Van der Duyn Schouten for helpful comments and corrections. Theorems 3.1 and 3.2 in their present form are due to Eric Denardo.

References

[1] Bertsekas, D. P. (1975). Convergence of Discretization Procedures in Dynamic Programming. *IEEE Trans. Automatic Control.* **20** 415–419.

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

[3] Fox, B. L. (1973). Discretizing Dynamic Programs. *J. Optimization Theory Appl.* **11** 228–234.

[4] Hinderer, K. (1978). On Approximate Solutions of Finite-Stage Dynamic Programs. *Proceedings of the International Conference on Dynamic Programming, University of British Columbia, Vancouver.* M. Puterman (ed.). To appear.

[5] Morin, T. L. (1978). Computational Advances and Reduction of Dimensionality in Dynamic Programming: A Survey. *Proceedings of the International Conference on Dynamic Programming, University of British Columbia, Vancouver.* M. Puterman (ed.). To appear.

[6] Thomas, A. (1977). *Models for Optimal Capacity Expansion.* Ph.D. Dissertation, School of Organization and Management, Yale University.

[7] Wagner, D. H. (1977). Survey of Measurable Selection Theorems. *SIAM J. Control Optimization* **15** 859–903.

[8] Wessels, J. (1977). Markov Programming by Successive Approximations with Respect to Weighted Supremum Norms. *J. Math. Anal. Appl.* **58** 326–335.

[9] Whitt, W. (1976). Baire Classification of Measurable Selections of Extrema.

[10] ———. (1977). Representation and Approximation of Noncooperative Sequential Games.

[11] Zipkin, P. H. (1977). *Aggregation in Linear Programming.* Ph.D. Dissertation, School of Organization and Management, Yale University.

BELL LABORATORIES, HOLMDEL, NEW JERSEY 07733