MONOTONE MAPPINGS IN DYNAMIC PROGRAMMING*

Dimitri P. Bertsekas

Department of Electrical Engineering and
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

Abstract

In this paper we consider a class of monotone mappings underlying many sequential optimization problems over a finite or infinite horizon which are of interest in applications. This class of problems includes deterministic and stochastic optimal control problems, minimax control problems, Semi-Markov Decision problems and others. We prove some fixed point properties of the optimal value function and we analyze the convergence properties of a related generalized Dynamic Programming algorithm. We also give a sufficient condition for convergence, which is widely applicable and considerably strengthens known related results.

1. Introduction

It is well known that Dynamic Programming (D.P. for short) is the principal method for analysis of sequential optimization problems. It is also known that it is possible to describe each iteration of a D.P. algorithm by means of a certain mapping which maps a space of functions defined on the state space into itself. In problems with a finite, say k, number of stages, after k successive applications of this mapping (i.e., after k steps of the D.P. algorithm) one obtains the optimal value function of the problem. In problems with an infinite number of stages one hopes that the sequence of functions generated by successive application of the D.P. iteration converges in some sense to the optimal value function of the problem.

To illustrate this viewpoint let us consider the deterministic optimal control problem of finding a finite sequence of control functions $\pi = \{\mu_0,\mu_1,\ldots,\mu_{k-1}\}$ which minimize

(1) $\quad f_{k,\pi}(x_0) = \sum_{i=0}^{k-1} h[x_i,\mu_i(x_i)]$

subject to the system equation constraint

(2) $\quad x_{i+1} = g[x_i,\mu_i(x_i)] \quad i = 0,1,\ldots,k-1.$

The states x_i belong to a state space X and the controls $\mu_i(x_i)$ are elements of a control space U. The initial state x_0 is given, and h,g are given functions. The D.P. algorithm for this problem is given by

(3) $\quad f_0(x) = 0$

(4) $\quad f_{i+1}(x) = \inf_u\{h(x,u) + f_i[g(x,u)]\} \quad i = 0,1,\ldots,k-1$

and the optimal value of the problem is obtained at the kth step of the algorithm

$$\inf_\pi f_{k,\pi}(x_0) = f_k(x_0).$$

Notice also that one may obtain the value $f_{k,\pi}(x_0)$ of (1) corresponding to any $\pi = \{\mu_0,\mu_1,\ldots,\mu_{k-1}\}$ from the kth step of the D.P. algorithm

(5) $\quad f_{0,\pi}(x) = 0$

(6) $\quad f_{i+1,\pi}(x) = h[x,\mu_{k-1-i}(x)] + f_{i,\pi}[g(x,\mu_{k-1-i}(x))]$

Now it is possible to formulate the problem above as well as to describe the D.P. algorithm (3),(4) by means of the mapping H given by

(7) $\quad H(x,u,f) = h(x,u) + f[g(x,u)].$

Let us define

(8) $\quad T(f)(x) = \inf_u H(x,u,f)$

and for any function μ: $X \to U$

(9) $\quad T_\mu(f)(x) = H[x,\mu(x),f]$

Then, in view of (5),(6), we may write the cost functional $f_{k,\pi}$ of (1) as

(10) $\quad f_{k,\pi}(x_0) = (T_{\mu_0} T_{\mu_1} \ldots T_{\mu_{k-1}})(f_0)(x_0)$

where f_0 is the zero function of (3), ($f_0(x) = 0$, $\forall x \in X$) and $(T_{\mu_0} T_{\mu_1} \ldots T_{\mu_{k-1}})$ is the composition of the mappings $T_{\mu_0},T_{\mu_1},\ldots,T_{\mu_{k-1}}$.

Similarly the D.P. algorithm (3),(4) may be described by

(11) $\quad f_{i+1}(x) = T(f_i)(x) \quad i = 0,1,\ldots,k-1$

and we have

(12) $\quad \inf_\pi f_{k,\pi}(x_0) = T^k(f_0)(x_0)$

where T^k is the composition of T with itself k times.

One may consider also an infinite horizon version of the deterministic problem above where-

*Work supported by the Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DAAB-07-72-C-0259 and by NSF under Grant ENG 74-19332.

by we seek a sequence $\pi = \{\mu_0, \mu_1, \ldots\}$ of control functions which minimize

$$(13)\quad f_\pi(x_0) = \lim_{k\to\infty} \sum_{i=0}^{k-1} h[x_i, \mu_i(x_i)]$$

subject to the system equation constraint

$$(14)\quad x_{i+1} = g[x_i, \mu_i(x_i)] \quad i = 0,1,\ldots$$

We assume that the function h above is such that the limit in (13) is well defined for every π. A question of considerable computational and analytical interest concerns the validity of the equation

$$(15)\quad f^*(x) = \inf_\pi f_\pi(x) = \lim_{k\to\infty} T^k(f_o)(x)$$

where T is the mapping of (8). When (15) holds, the D.P. algorithm yields in the limit the optimal value of the problem. Another important question is whether the optimal value function f^* satisfies Bellman's functional equation

$$(16)\quad f^*(x) = \inf_u\{h(x,u) + f^*[g(x,u)]\}$$

or equivalently whether

$$(17)\quad f^*(x) = T(f^*)(x)$$

and f^* is a fixed point of the mapping T. Considerable amount of research has been directed towards resolving these and other related questions. It has been proved under fairly general assumptions that (16) holds true. However the equality (15) has been proved only under restrictive assumptions. In fact (15) may fail even for very simple problems as the following example shows:

Example: Let $X = [0,\infty)$, $U = (0,\infty)$ be the state and control spaces respectively. Let the system equation be

$$x_{i+1} = 2x_i + u_i \quad i = 0,1,\ldots$$

and let the cost per stage be defined by

$$h(x,u) = x + u.$$

Then it can easily be verified that

$$f^*(x) = \inf_\pi f_\pi(x) = +\infty \qquad \forall\ x \in X$$

while

$$T^k(f_0)(0) = 0 \qquad \forall\ k = 1,2,\ldots$$

The deterministic optimization problem described above is representative of a plethora of sequential optimization problems of practical interest which may be formulated in terms of mappings similar to the mapping H of (7). Furthermore D.P. algorithms corresponding to such problems may also be described in terms of mappings similar to the mapping T of (8). A general class of such sequential optimization problems will be described in the next section together with examples of specific classes of problems of interest. Similarly as in deterministic problems, questions arise as to whether equations such as (15) and (17) hold. The purpose of this paper is to provide an analysis of these questions in a general setting. To this end we take as our starting point a class of mappings and subsequently construct a class of abstract optimization problems which contains as special cases several classes of problems of interest. The validity of several important properties of these problems is traced directly to properties of their associated mappings. Since the framework adopted is abstract and general in nature, the analysis provides a measure of the extent to which results obtained earlier for specific cases hold true in a more general setting. In addition existing results for important special cases, such as deterministic and stochastic optimal control problems, concerning the convergence of the D.P. algorithm in infinite horizon problems are substantially strengthened. In view of the space limitation, no detailed proofs are given in this paper. They may be found in a report [15] available from the author on request.

2. A Class of Monotone Mappings

Let X,U be two sets, let b be some extended real number and let F be the set of all functions $f:X \to [b,\infty]$. For any two functions $f,f'\epsilon F$ we write

$$f = f' \quad \text{if} \quad f(x) = f'(x) \quad \forall\ x\epsilon X$$

$$f \le f' \quad \text{if} \quad f(x) \le f'(x) \quad \forall\ x\epsilon X.$$

For any collection $\{f_\alpha | \alpha\epsilon A\}$ of functions in F we denote by $\inf_{\alpha\epsilon A} f_\alpha$ ($\sup_{\alpha\epsilon A} f_\alpha$) the pointwise infimum (supremum) of f_α over A. For any sequence $\{f^k\}$ with $f^k \epsilon F$ we denote by $\lim_{k\to\infty} f^k$ the pointwise limit of $\{f^k\}$ (assuming it is well defined as an extended real valued function), and by $\liminf_{k\to\infty} f^k$ and $\limsup_{k\to\infty} f^k$ the pointwise limit inferior or limit superior of $\{f^k\}$. Throughout this paper the convergence analysis is carried out within the set of extended real numbers, i.e. $+\infty$ or $-\infty$ are allowed as limits of sequences of extended real numbers.

Consider now a mapping $H:X \times U \times F \to [b,\infty]$ which has the following monotonicity property

$$(18)\quad f \le f' \Rightarrow H(x,u,f) \le H(x,u,f') \qquad \forall (x,u)\epsilon X \times U, f,f'\epsilon F.$$

For any function $\mu:X \to U$ define the mapping $T_\mu:F \to F$ by means of

$$(19)\quad T_\mu(f)(x) = H[x,\mu(x),f] \quad \forall\ x\epsilon X.$$

Define also the mappings $T:F \to F$ and $\tilde{T}:F \to F$ by means of

$$(20)\quad T(f)(x) = \inf_u H(x,u,f) \quad \forall\ x\epsilon X.$$

$$(21)\quad \tilde{T}(f)(x) = \sup_u H(x,u,f) \quad \forall\ x\epsilon X.$$

Relation (18) implies the following monotonicity relations

$$(22)\quad f \le f' \Rightarrow T_\mu(f) \le T_\mu(f') \quad \forall\ f,f'\epsilon F,\ \mu: X \to U$$

$$(23)\quad f \le f' \Rightarrow T(f) \le T(f'),\ \tilde{T}(f) \le \tilde{T}(f') \qquad \forall\ f,f'\epsilon F$$

(24) $f \le f' \Rightarrow (T_{\mu_0} T_{\mu_1} \dots T_{\mu_k})(f)$

$$\le (T_{\mu_0} T_{\mu_1} \dots T_{\mu_k})(f') \quad \forall\, f,f' \epsilon F$$

$$\mu_i : X \to U, \; i=0,\dots,k.$$

In relation (24) above $(T_{\mu_0} \dots T_{\mu_k})$ represents the composition of the mappings $T_{\mu_0}, T_{\mu_1}, \dots, T_{\mu_k}$. We shall denote by T_μ^k, T^k, $\tilde{T}^k$ the respective composition of T_μ, T, $\tilde{T}$ with itself k times.

Mappings of the type considered are of interest in several classes of dynamic optimization problems under certainty or under uncertainty involving stationary or nonstationary dynamic systems and an infinite horizon. Some examples are provided below. Further examples may be found in the author's forthcoming textbook [16] and in the paper by Denardo [11] who considered similar mappings under additional boundedness and contraction assumptions.

Deterministic Optimal Control with Additive Cost Functional (see e.g. [1])

(25) $H(x,u,f) = h(x,u) + \alpha(x,u) f[g(x,u)]$

If $h(x,u) \ge 0$, $\alpha(x,u) \ge 0$, $\forall (x,u) \epsilon X \times U$, $b=0$, then H above fall within the described framework. Here X is the state space, U is the control space, $h(x,u)$ represents cost per stage, $g: X \times U \to X$ represents the system function, and $\alpha(x,u)$ may be viewed as a discount factor whenever $\alpha(x,u) < 1$. This discount factor may depend on x and u.

Stochastic Optimal Control with Additive Cost Functional (see e.g. [2,4])

(26) $H(x,u,f) = \underset{w}{E}\{h(x,u,w) + \alpha(x,u,w) f[g(x,u,w)] \,|\, x,u\}$

Here w is an uncertain parameter-element of a countable set W with given probability distribution for every $(x,u) \epsilon X \times U$, and $E\{\cdot\}$ denotes expectation. If $h(x,u,w) \ge 0$, $\alpha(x,u,w) \ge 0$ $\forall (x,u,w) \epsilon X \times U \times W$, $b=0$, then H above falls within our framework.

Minimax Control Problems with Additive Cost Functionals (see e.g. [5-7])

(27) $H(x,u,f) = \sup\limits_{w \epsilon W(x,u)} \{h(x,u,w)$

$$+ \alpha(x,u,w) f[g(x,u,w)]\}$$

Here again w is an uncertain parameter-element of a set W, and W(x,u) is a given subset of W for every $(x,u) \epsilon X \times U$. If $h(x,u,w) \ge 0$, $\alpha(x,u,w) \ge 0$ $\forall (x,u,w) \epsilon X \times U \times W$, $b=0$ then H above falls within our framework. The same is true if "sup" in (27) is replaced by "inf" in which case (27) is of interest in Max-Min problems.

Stochastic Control Problems with Exponential Cost Functionals (see e.g. [8,9])

(28) $H(x,u,f) = \underset{w}{E}\{e^{h(x,u,w)} f[g(x,u,w)] \,|\, x,u\}$

Here everything is as in (26).

In the next section we formulate two optimization problems corresponding to the mapping H and we consider related generalized Dynamic Programming Algorithms. Various existence, characterization and convergence results are given subsequently in Section 4.

3. A Class of Optimization Problems

Let Π denote the set of all sequences $\Pi = \{\mu_0, \mu_1, \dots\}$ of functions $\mu_k : X \to U$ (also referred to as policies). Suppose that $f_0 \epsilon F$ is a function such that

(29) $f_0(x) \le H(x,u,f_0) \quad \forall (x,u) \epsilon X \times U$

and define for every $x \epsilon X$, $\pi \epsilon \Pi$

(30) $f_\pi(x) = \lim\limits_{k\to\infty} (T_{\mu_0} T_{\mu_1} \dots T_{\mu_k})(f_0)(x).$

By (29) we have $f_0 \le T_{\mu_k}(f_0)$, $k=0,1,\dots$, and using (24) we obtain

(31) $f_0 \le T_{\mu_0}(f_0) \le \dots \le (T_{\mu_0} \dots T_{\mu_k})(f_0) \le \dots$

(32) $f_0 \le T(f_0) \le T^2(f_0) \le \dots \le T^k(f_0) \le \dots$

(33) $f_0 \le \tilde{T}(f_0) \le \tilde{T}^2(f_0) \le \dots \le \tilde{T}^k(f_0) \le \dots .$

Hence the limit in (30) is well defined as an extended real number.

Consider now the optimization problems of finding

(34) $f^*(x) = \inf\limits_{\pi \epsilon \Pi} f_\pi(x)$

(35) $\tilde{f}^*(x) = \sup\limits_{\pi \epsilon \Pi} f_\pi(x)$

as well as policies attaining the infimum or the supremum above (if any exist). These problems contain as special cases dynamic optimization problems over an infinite horizon which have been the subject of much attention in recent years. One example is the deterministic optimization problem of the past section under the assumption that the function h in (1) satisfied $h(x,u) \ge 0$ for all x,u. In this case f_0 is taken to be the function which is identically zero on X ($f_0(x)=0, \forall\, x \epsilon X$). As another example take the mapping (26) and assume that $0 \le h(x,u,w)$, $\alpha(x,u,w) = 1$, $\forall (x,u,w) \epsilon X \times U \times W$ $f_0(x) = 0$, $\forall\, x \epsilon X$. Then the problems associated with (34) and (35) fall within the framework of the negative and positive dynamic programming problems examined by Blackwell [14], Strauch [2] and subsequently by Hinderer [3]. Similarly for the mapping (27) under the same assumptions on h and α, problem (34) represents an infinite horizon version of minimax control problems considered by Witsenhausen [5] and contains as a special case a reachability problem considered by the author [10].

It is to be noted that, since the mapping H is extended real valued, one may handle constraints on x and u which may be present in the minimization problem (34) simply by letting H take the value $+\infty$ whenever x or u violate these constraints. For the maximization problem (35) one may handle constraints on x or u when $b = -\infty$ by letting H take the value $-\infty$ whenever the constraints are violated. When $b \ne -\infty$ a reformulation of the problem is necessary and is possible for many problems of interest by using reachability methods [10].

Now consider the functions f_k and $\tilde{f}_k$ given by

(36) $f_k = \inf_{\pi\in\Pi} (T_{\mu_0} \cdots T_{\mu_{k-1}})(f_0) \quad k = 1,2,\ldots$

(37) $\tilde{f}_k = \sup_{\pi\in\Pi} (T_{\mu_0} \cdots T_{\mu_{k-1}})(f_0) \quad k = 1,2,\ldots$

The functions f_k, $\tilde{f}_k$ represent the optimal value functions of "truncated" (finite horizon) versions of problems (34) and (35) respectively. By (31), $\{f_k\}$, $\{\tilde{f}_k\}$ are increasing sequences and their pointwise limits f_∞, $\tilde{f}_\infty$

(38) $f_\infty(x) = \lim_{k\to\infty} f_k(x) \quad \forall x\in X$

(39) $\tilde{f}_\infty(x) = \lim_{k\to\infty} \tilde{f}_k(x) \quad \forall x\in X$

are well defined as functions in F.

The objective of this paper is to provide conditions under which the following equalities hold

(40) $f^* = T(f^*) \quad , \quad \tilde{f}^* = \tilde{T}(\tilde{f}^*)$

(41) $f_\infty = f^* \quad , \quad \tilde{f}_\infty = \tilde{f}^*$.

Equations (40) may be viewed as generalized versions of Bellman's equation. They are known to hold for discounted, negative and positive D.P. models [14],[2],[3]. They do not necessarily hold in our more general setting in the absence of additional assumptions which are specified in the next section. The validity of equations (41) is of considerable analytical and computational interest since the functions f_∞ and $\tilde{f}_\infty$ may usually (but not always within our setting) be obtained in the limit by the D.P. Algorithm

(42) $f_\infty = \lim_{k\to\infty} T^k(f_0) \quad , \quad \tilde{f}_\infty = \lim_{k\to\infty} \tilde{T}^k(f_0)$.

In special cases such as those considered by Strauch [2] and Hinderer [3] there holds $\tilde{f}_\infty = \tilde{f}^*$ but the equality $f_\infty = f^*$ can be guaranteed to hold only under restrictive finiteness assumptions on the space U (see [2], Th.9.1) in negative D.P. models. An example where $f_\infty \neq f^*$ was given in Section 1 and other examples where $f_\infty \neq f^*$ have been given in [2] (p.880) and [10] (p.608). Within our more general setting we prove in the next section that it is always true that $\tilde{f}_\infty = \tilde{f}^*$ and we provide conditions under which the equality $f_\infty = f^*$ also holds. These conditions (Proposition 10) strengthen substantially existing results.

Several other subsidiary questions are examined in the next section. Conditions under which the equalities (42) hold are provided. In addition the question of existence and characterization of optimal stationary policies is examined.

It is to be noted that the equations (40), (41), and (42) hold under boundedness and contraction assumptions similar to those introduced by Denardo in [11]. Since Denardo's analysis is complete and satisfying there is no reason of duplicating any portion of it here. For this reason we shall restrict our attention to situations where contraction assumptions such as those of [11] need not be satisfied.

4. Main Results

The first relations between f_∞, $\tilde{f}_\infty$ and f^*, $\tilde{f}^*$ are given in the following two propositions. The results for problem (34) differ from those for problem (35).

Proposition 1: There holds

(43) $\tilde{f}_\infty = \tilde{f}^*$.

Proposition 2: Assume that $f_\infty = T(f_\infty)$ and furthermore the infimum in the relation

(44) $f_\infty(x) = \inf_u H(x,u,f_\infty)$

is attained for every $x\in X$. Then

(45) $f^* = T(f^*) = f_\infty = T(f_\infty)$

and if $\mu^*(x)$ attains the infimum in (44) for each $x\in X$ then the stationary policy $\pi^* = \{\mu^*,\mu^*,\ldots\}$ minimizes $f_\pi(x)$ for every $x\in X$.

Further properties and relations of f_∞, $\tilde{f}_\infty$, f^*, $\tilde{f}^*$ follow upon the introduction of the following two continuity properties of H.

P1. If $\{g_k\}$ is any sequence with $g_k\in F$, and $g_k \leq g_{k+1}$ for all k, then

(46) $\lim_{k\to\infty} H(x,u,g_k) = H(x,u,\lim_{k\to\infty} g_k) \quad \forall (x,u)\in X \times U$.

P2. There exists a scalar $\alpha > 0$ such that for all scalars $r > 0$ and functions $f\in F$ there holds

(47) $H(x,u,f) \leq H(x,u,f+re) \leq H(x,u,f) + \alpha r \quad \forall (x,u)\in X \times U$

where e denotes the unit function on X ($e(x) = 1$, $\forall x\in X$).

The following two propositions provide conditions under which the functions f_k and $\tilde{f}_k$ of (36),(37) can be computed iteratively by means of generalized Dynamic Programming algorithms.

Proposition 3: Let P2 hold and assume that $f_0(x) > -\infty$ for all $x\in X$. Then $f_k = T^k(f_0)$ for all $k = 1,2,\ldots$ and hence

$$f_\infty = \lim_{k\to\infty} f_k = \lim_{k\to\infty} T^k(f_0).$$

Proposition 3 may fail to hold if either of its two assumptions is not in effect.

Counterexample 1: Take $X = \{0\}$, $U = (0,1)$, $b = 0$, $f_0(0) = 0$, $H(0,u,f) = 1$ if $f(0) > 0$, $H(0,u,f) = u$ if $f(0) = 0$. Then $(T_{\mu_0} \cdots T_{\mu_{k-1}})(f_0)(0) = 1$ for every $\pi\in\Pi$ and $k \geq 2$. Hence $f_k(0) = f_\infty(0) = f^*(0) = 1$. But we have $T^k(f_0)(0) = 0$ for all k. Here P2 is violated.

Counterexample 2: Take $X = \{0,1\}$, $U = (-\infty,0]$, $b = -\infty$, $f_0(0) = f_0(1) = -\infty$, $H(0,u,f) = u$ if $f(1) = -\infty$, $H(0,u,f) = 0$ if $f(1) > -\infty$, and $H(1,u,f) = u$. Then $(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{k-1}})(f_0)(0) = 0$, $(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{k-1}})(f_0)(1) = \mu_0(1)$ for every $\pi\in\Pi$ and $k \geq 2$. Hence we have $f_k(0) = f_\infty(0) = 0$,

$f_k(1) = f_\infty(1) = -\infty$ for every $k \geq 2$. On the other hand we have $T^k(f_0)(0) = T^k(f_0)(1) = -\infty$ for all $k \geq 1$. Notice also that for this problem we have $f^* = f_\infty$ and $T(f^*)(0) = T(f_\infty)(0) = T(f^*)(1) = T(f_\infty)(1) = -\infty$. Hence $f^* \neq T(f^*)$ and $f_\infty \neq T(f_\infty)$. Here P2 is satisfied but the assumption $f_0(x) > -\infty$, $\forall\ x\epsilon X$ is violated.

Proposition 4: Let P2 hold and assume that $\tilde{T}^k(f_0)(x) < +\infty$ for all $x\epsilon X$. Then $\tilde{f}_k = \tilde{T}^k(f_0)$ for all $k = 1,2,\ldots$ and hence

$$\tilde{f}_\infty = \lim_{k\to\infty} f_k = \lim_{k\to\infty} T^k(f_0).$$

Proposition 4 may fail to hold if either of its two assumptions is not in effect.

Counterexample 3: Take $X = \{0\}$, $U = (0,1)$, $b = 0$, $f_0(0) = 0$, $H(0,u,f) = u$ if $0 \leq f(0) < 1$, $H(0,u,f) = f(0) + u$ if $1 \leq f(0)$. Then $\tilde{f}_k(0) = \sup_{\pi\epsilon\Pi}(T_{\mu_0}\ldots T_{\mu_{k-1}})(\tilde{f}_0)(0) = 1$ but $\tilde{T}^k(f_0)(0) = k$ for every $k \geq 1$. Here P2 is violated. Notice also that here $\tilde{f}^*(0) = 1$ but $\tilde{T}(\tilde{f}^*)(0) = 2$ and $f^* \neq \tilde{T}(\tilde{f}^*)$.

Counterexample 4: Take $X = \{0,1\}$, $U = [0,\infty)$, $b = 0$, $f_0(0) = f_0(1) = 0$, $H(0,u,f) = u$ if $f(1) = \infty$, $H(0,u,f) = 0$ if $f(1) < \infty$, and $H(1,u,f) = u$. Then $(T_{\mu_0}T_{\mu_1}\ldots T_{\mu_{k-1}})(f_0)(0) = 0$,

$(T_{\mu_0}T_{\mu_1}\ldots T_{\mu_{k-1}})(f_0)(1) = \mu_0(1)$ for every $k \geq 1$.

Hence we have $\tilde{f}_k(0) = \tilde{f}_\infty(0) = 0$, $\tilde{f}_k(1) = \tilde{f}_\infty(1) = \infty$ for all $k \geq 1$. On the other hand we have $\tilde{T}^k(f_0)(0) = \tilde{T}^k(f_0)(1) = \infty$ for all $k \geq 2$. Notice also that for this problem we have $\tilde{f}^* = \tilde{f}_\infty$ (as Proposition 1 predicts) and $\tilde{T}(\tilde{f}^*)(0) = \tilde{T}(\tilde{f}^*)(1) = \infty$. Hence $\tilde{f}^* \neq \tilde{T}(\tilde{f}^*)$ and $\tilde{f}_\infty \neq \tilde{T}(\tilde{f}_\infty)$. Here P2 is satisfied but the assumption $\tilde{T}^k(f_0)(x) < \infty$, $\forall\ x\epsilon X$ is violated.

The following propositions provide conditions under which the optimal value functions f^* and $\tilde{f}^*$ are fixed points of T and $\tilde{T}$ respectively. Furthermore under appropriate conditions f^* and $\tilde{f}^*$ are characterized as the "smallest" fixed points of T and $\tilde{T}$ respectively.

Proposition 5: Let P2 hold and assume that $f'\epsilon F$ is a function such that $f' \geq f_0$, $f'(x) > -\infty$ for all $x\epsilon X$, and $f' = T(f')$. Then $f^* \leq f'$.

In Counterexample 1 we have that $f'(0) = 0$ satisfies $f' = T(f')$. Yet $f^*(0) = 1 > f'(0)$. Here P2 is violated.

Proposition 6: Let P2 hold and assume that $f'\epsilon F$ is a function such that $f' \geq f_0$, $f' \geq \tilde{T}(f')$ and either $f'(x) < \infty$ or else $\sup_\pi(T_{\mu_0}\ldots T_{\mu_{k-1}})(f')(x) = \tilde{T}^k(f')(x)$ for all $x\epsilon X$. Then $\tilde{f}^* \leq f'$.

It is to be noted that Proposition 6 may form the basis for computation of the function $\tilde{f}^*$ when the set X is a finite set, $X = \{x_1, x_2, \ldots, x_n\}$, and $\tilde{f}^*$ satisfies $\tilde{f}^* = \tilde{T}(\tilde{f}^*)$. Under these conditions it follows from Proposition 6 that $\{\tilde{f}^*(x_1), \ldots, \tilde{f}^*(x_n)\}$ solve the problem

$$\min \sum_{i=1}^{n} \lambda_i$$

subject to

$$\lambda_i \geq \sup_u H(x_i,u,f_\lambda) \qquad i = 1,\ldots,n$$

$$\lambda_i \geq f_0(x_i) \qquad i = 1,\ldots,n$$

where f_λ is the function taking values $f_\lambda(x_i) = \lambda_i$, $i = 1,\ldots,n$.

Proposition 7: Assume that P1 holds, $f_0(x) > -\infty$ for all $x\epsilon X$, and furthermore the mapping H has the property that for every $(x,u)\epsilon X \times U$ there exists a sequence of policies $\{\pi_k\}$ such that

(48) $H(x,u,f_{\pi_k}) \to H(x,u,f^*)$.

Then

$$f^* = T(f^*).$$

Proposition 8: Assume that P1 holds and furthermore the mapping H has the property that for every $(x,u)\epsilon X \times U$ there exists a sequence of policies $\{\pi_k\}$ such that

(49) $H(x,u,f_{\pi_k}) \to H(x,u,\tilde{f}^*)$.

Then

$$\tilde{f}^* = \tilde{T}(\tilde{f}^*).$$

Counterexamples 2, 3 and 4 show the manner in which the relations $f^* = T(f^*)$ and $\tilde{f}^* = \tilde{T}(\tilde{f}^*)$ may fail to hold.

It is to be noted that the conditions (48) and (49) may be quite difficult to verify in specific problems. For example while they do hold in positive and negative dynamic programming problems [14], [2], [3], their verification is rather complicated. The following proposition provides alternative, and in some cases more easily verifiable conditions, which guarantee that $\tilde{f}^* = \tilde{T}(\tilde{f}^*)$.

Proposition 9: Let P1, P2 hold and assume further that $\tilde{T}^k(f_0)(x) < \infty$ for all $x\epsilon X$ or otherwise $\tilde{f}_k = \tilde{T}^k(f_0)$ for all $k \geq 1$. Then

$$\tilde{f}^* = \tilde{T}(\tilde{f}^*) = \tilde{f}_\infty = \tilde{T}(\tilde{f}_\infty).$$

The following proposition provides a compactness assumption under which the equality $f_\infty = f^* = T(f^*)$ is satisfied and furthermore an optimal stationary policy exists. It is readily verifiable in many problems of interest and constitutes a substantial improvement over available sufficient cconditions for the D.P. algorithm to yield in the limit the optimal value function f^*.

Proposition 10: Let U be a Hausdorff topological space and assume that P1, P2 hold, we have $f_0(x) > -\infty$ for all $x\epsilon X$, and there exists a nonnegative integer $\bar{k}$ such that for each $x\epsilon X$, $\lambda\epsilon(-\infty,\infty)$ and $k \geq \bar{k}$ the set

(50) $U_k(x,\lambda) = \{u\epsilon U | H(x,u,f_k) \leq \lambda\}$

is compact. Then

$$f^* = T(f^*) = f_\infty = T(f_\infty).$$

Furthermore there exists a stationary policy $\pi^* = \{\mu^*,\mu^*,\ldots\}$ which minimizes $f_\pi(x)$ for all $x\epsilon X$.

The compactness of the sets $U_k(x,\lambda)$ of (50) may be verified in a number of important special cases. One such case is when $U_k(x,\lambda)$ is a finite set for all k, x, λ. Simply consider the discrete topology on U [12], i.e. the topology consisting of all subsets of U. In this topology a set is compact if and only if it is finite. For this case the relation $f_\infty = f^*$ for the negative model of Strauch has been shown earlier [2]. There are many other important cases where the compactness of $U_k(x,\lambda)$ can be verified. It is not our intention to provide an extensive list of such cases. Instead we state as an illustration two sets of assumptions which guarantee compactness of $U_k(x,\lambda)$ in the case of the mapping

(51) $H(x,u,f) = h(x,u) + \alpha(x,u)\, f[g(x,u)]$

corresponding to a deterministic optimal control problem.

Assume in (51) that $h(x,u) \geq 0$, $\alpha(x,u) \geq 0$ for all $(x,u)\epsilon X \times U$ and take F to be the set of functions $f:X \to [0,\infty]$ and $f_0(x) = 0$, $\forall x\epsilon X$. Then compactness of $U_k(x,\lambda)$ is guaranteed if:

a) $X = R^n$ (n-dimensional Euclidean space), $U = R^m$, h, g, α are continuous in (x,u) and h satisfies $\lim_{n\to\infty} h(x_n,u_n) = \infty$ for every bounded sequence $\{x_n\}$ and every sequence $\{u_n\}$ for which $|u_n| \to \infty$ ($|\cdot|$ is a norm on R^m).

b) $X = R^n$, U is a nonempty and compact subset of R^m, g and α are continuous, h is continuous on $R^n \times U$.

Aside from the result of Strauch mentioned earlier, other general sufficient conditions which guarantee that an optimal stationary policy exists for special cases of problem (17) are those of Maitra for discounted problems (see [3] Th.5.11), and Kushner for free end time problems [13]. In these cases the basic mapping has contraction properties which guarantee that $f_\infty = f^*$. In both cases the sufficient conditions for existence of an optimal stationary policy follow from Proposition 10.

Finally we point out that our results are limited in two ways. First we have allowed as admissible policies all sequences of arbitrary functions μ_k: $X \to U$. In many problems of interest there may be restrictions on the class of functions μ_k under consideration. For example μ_k may be required to be measurable, when X and U are measurable spaces. Our results, unless the problem is reformulated, are not applicable to such a case. A second limitation is due to the assumption (29) on the initial function f_0 which rules out the applicability of our results to certain classes of problems which do not fit the framework of Positive and Negative D.P. models. Current research is aimed at relaxation, to the extent possible, of these restrictive assumptions.

References

[1] Bellman, R., Dynamic Programming, Princeton Univ. Press, Princeton, N.J., 1957.

[2] Strauch, R.E., "Negative Dynamic Programming," Annals of Math. Stat., Vol. 37, 1966, pp.871-890.

[3] Hinderer, K., Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter, Springer-Verlag, N.Y., 1970.

[4] Ross, S.M., Applied Probability Models with Optimization Applications, Holden Day, San Francisco, 1970.

[5] Witsenhausen, H.S., "Minimax Control of Uncertain Systems," Ph.D. Dissertation, Dept. of Electrical Engineering, MIT, Cambridge, Mass., May 1966.

[6] Witsenhausen, H.S., "A Minimax Control Problem for Sampled Linear Systems," IEEE Trans. on Automatic Control, Vol AC-13, 1968, p.5-21.

[7] Bertsekas, D.P. and Rhodes, I.B., "Sufficiently Informative Functions and the Minimax Feedback Control of Uncertain Systems," IEEE Trans. on Automatic Control, Vol. AC-18 1973, pp.117-124.

[8] Jacobson, D.H., "Optimal Stochastic Linear Systems with Exponential Performance Criteria and their Relation to Deterministic Differential Games," IEEE Trans. on Automatic Control, Vol. AC-18, 1973, pp.124-131.

[9] Howard, R. and Matheson, J., "Risk Sensitive Markov Decision Processes," Management Science, Vol.18, 1972, pp.356-369.

[10] Bertsekas, D.P., "Infinite Time Reachability of State Space Regions by Using Feedback Control," IEEE Trans. on Automatic Control, Vol.AC-17, 1972, pp.604-613.

[11] Denardo, R.V., "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review, Vol.9, 1967, pp.165-177.

[12] Dugundji, J., Topology, Allyn and Bacon, Boston, Mass., 1968.

[13] Kushner, H.J., Introduction to Stochastic Control, Holt, Rinehart and Winston, Inc., New York, 1971.

[14] Blackwell, D., "Positive Dynamic Programming," Proc. Fifth Berkeley Symp. on Math. Stat. and Probability, Vol.1, 1965, pp.415-418.

[15] Bertsekas, D.P., "On Fixed Points of Some Monotone Mappings with Application in Dynamic Programming," Coordinated Science Lab. Working Paper, U. of Ill., Nov. 1974.

[16] Bertsekas, D.P., Dynamic Programming and Stochastic Control, Academic Press, N.Y., to appear in 1976.