The Annals of Probability
1974, Vol. 2, No. 5, 942–949

THE OPTIMAL REWARD OPERATOR IN SPECIAL CLASSES OF DYNAMIC PROGRAMMING PROBLEMS[1]

BY DAVID A. FREEDMAN

University of California, Berkeley

Consider a dynamic programming problem with separable metric state space S, constraint set A, and reward function $r(x, P, y)$ for $(x, P) \in A$ and $y \in S$. Let Tf be the optimal reward in one move, for the reward function $r(x, P, y) + f(y)$. Three results are proved. First, suppose S is compact, A closed, and r upper semi-continuous; then $T^n 0$ is upper semi-continuous, and there is an optimal Borel strategy for the n-move game. Second, suppose S is compact, A is an F_σ, and $\{r > a\}$ is an F_σ for all a; then $\{T^n 0 > a\}$ is an F_σ for all a, and there is an ε-optimal Borel strategy for the n-move game. Third, suppose A is open and r is lower semi-continuous; then $T^n 0$ is lower semi-continuous, and there is an ε-optimal Borel measurable strategy for the n-move game.

1. Introduction. A dynamic programming problem can be specified in terms of three objects: the state space S, the constraint set A, and the reward function r. Let S be a separable metric set, endowed with the Borel σ-field $\sigma(S)$ generated by the topology. Let $\pi(S)$ be the set of probabilities on $\sigma(S)$, endowed with the weak* topology and the Borel σ-field generated by this topology. So $\pi(S)$ is also separable metric. Suppose A is a Borel subset of $S \times \pi(S)$, whose x-section A_x is nonempty for all $x \in S$. Suppose r is a nonnegative extended real-valued function on $A \times S$. Informally, when you are at $x \in S$, you can select any $P \in A_x$, move to y chosen at random from S according to P, and receive the reward $r(x, P, y)$.

The optimal reward operator T was defined in [1] as follows:

$$(1) \qquad (Tf)(x) = \sup_{P \in A_x} \int_S [r(x, P, y) + f(y)]P(dy) ,$$

provided the integral makes sense. Write 0 for the function which vanishes identically. If, for instance, S is a Borel subset of a complete separable metric space, then [1] identifies $T^n 0$ as the optimal reward in n moves, and demonstrates the existence of universally measurable ε-optimal strategies. In that degree of generality, $T^n 0$ need not be Borel, although it has to be universally measurable; and there need not be any Borel strategies whatsoever, optimal or otherwise.

This note will apply the argument of [1] to three special cases, where the measurability issues are much easier to resolve. Before stating the conditions

Received December 27, 1972; revised February 1974.

[1] Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-71-2100. The United States government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

AMS 1970 *subject classifications*. 49C99, 60K99, 90C99, 28A05.

Key words and phrases. Dynamic programming, optimal reward, optimal strategy, gambling.

and the results, there is a review of some properties of strategies. In checking them, remember that the Borel σ-field in $\pi(S)$ is generated by the sets $\{\mu: \mu(B) > a\}$, for $0 < a < 1$ and $B \in \sigma(S)$. A function t_u into $\pi(S)$ is measurable, then, if $u \to t_u(B)$ is measurable for an algebra of B which generate $\sigma(S)$.

A *Borel strategy* s of length n is by definition a sequence of Borel measurable functions

$$x_1 \to s_{x_1}, (x_1, x_2) \to s_{x_1 x_2}, \cdots, (x_1, \cdots, x_n) \to s_{x_1, \cdots, x_n}$$

from $S, S^2, \cdots, S^n$ into $\pi(S)$, subject to the constraint

$$s_{x_1, \cdots, x_i} \in A_{x_i} .$$

The strategy s and starting state x determine a probability $\mathbf{s}_x$ on S^n, by the requirement that

$$\int_{S^n} \phi(x_2, \cdots, x_{n+1}) \mathbf{s}_x(dx_2, \cdots, dx_{n+1})$$

be the n-fold iterated integral of $\phi(x_2, \cdots, x_{n+1})$ relative to

$$s_{x x_2, \cdots, x_n}(dx_{n+1}) \cdots s_x(dx_2) ,$$

for all nonnegative Borel ϕ on S^n. So $x \to \mathbf{s}_x$ is Borel. The *reward* of s at $(x_1, x_2, \cdots, x_{n+1}) \in S^{n+1}$ is

$$r_n(s, x_1, x_2, \cdots, x_{n+1}) = r(x_1, s_{x_1}, x_2) + \cdots + r(x_n, s_{x_1, \cdots, x_n}, x_{n+1}) ,$$

a Borel function. The *expected reward* of s at $x \in S$ is

$$\rho_n(s, x) = \int_{S^n} r_n(s, x, x_2, \cdots, x_{n+1}) s_x(dx_2, \cdots, dx_{n+1}) ,$$

a Borel function of x. For $x \in S$, the *x-section* of s is this Borel strategy of length $n - 1$:

$$x_1 \to s_{x x_1}, (x_1, x_2) \to s_{x x_1 x_2}, \cdots, (x_1, \cdots, x_{n-1}) \to s_{x x_1, \cdots, x_{n-1}} .$$

So $(x, y) \to \mathbf{s}_x{}^y$ is Borel, as is $(x, y) \to \rho_{n-1}(s^x, y)$. As usual,

$$\rho_n(s, x) = \int_S [r(x, s_x, y) + \rho_{n-1}(s^x, y)] s_x(dy) . \tag{2}$$

A Borel strategy s of length n is *optimal* if $\rho_n(s, x) = (T^n 0)(x)$ for all $x \in S$. It is *ε-optimal* if

$$\begin{aligned} \rho_n(s, x) &> (T^n 0)(x) - \varepsilon \quad && \text{for all } x \text{ with } (T^n 0)(x) < \infty \\ &> 1/\varepsilon && \text{for all } x \text{ with } (T^n 0)(x) = \infty . \end{aligned}$$

This definition is proper because $T^n 0$ is an upper bound to $\rho_n(s, x)$, as shown in [1].

Here are the three sets of conditions on S, A, and r. The conventions of the first paragraph are to be understood to apply in all cases.

(3) Suppose S is compact metric, so $\pi(S)$ is too. Suppose A is a closed subset of $S \times \pi(S)$. Suppose r is a nonnegative, real-valued *upper semi-continuous* function on $A \times S$, that is, $\{r \geqq a\}$ is closed for all $a > 0$.

(4) Suppose S is compact metric, so $\pi(S)$ is too. Suppose A is an F_σ-subset of $S \times \pi(S)$, that is, a countable union of closed sets. Suppose r is a nonnegative, extended real-valued function on $A \times S$ of *type* F_σ, that is, $\{r > a\}$ is an F_σ for all $a > 0$.

(5) Suppose S is separable metric, but not necessarily Borel or even analytic. Suppose A is an open subset of $S \times \pi(S)$. Suppose r is a nonnegative, extended real-valued *lower semi-continuous* function on $A \times S$, that is, $\{r > a\}$ is open for all $a > 0$.

Here are the results.

(6) THEOREM. *Suppose* (3). *If f is a nonnegative, finite, upper semi-continuous function on S, so is Tf. In particular, so is $T^n 0$. For each n, there is an optimal Borel measurable strategy of length n.*

(7) THEOREM. *Suppose* (4). *If f is a nonnegative, extended real-valued function on S of type F_σ, so is Tf. In particular, so is $T^n 0$. For each n and positive ε, there is an ε-optimal Borel measurable strategy of length n.*

(8) THEOREM. *Suppose* (5). *If f is a nonnegative, extended real-valued lower semi-continuous function on S, so is Tf. In particular, so is $T^n 0$. For each n and positive ε, there is an ε-optimal Borel measurable strategy of length n.*

REMARKS. (a) Suppose (3), (4), or (5) holds. Then $T^n 0$ is non-decreasing with n. Call the limit u_∞. Then u is the optimal reward for the infinite game. If (4) holds, then u_∞ is of type F_σ. If (5) holds, then u_∞ is lower semi-continuous. In all three cases, the infinite game admits an ε-optimal Borel strategy which stops everywhere.

(b) There are four classes of sets considered in [1] and here: the analytic sets, the compact sets, the σ-compact sets, and the open sets. Each class is closed under finite unions and intersections, and under projections. If B is in a class, so is $\{\mu(B) > a\}$—or $\{\mu(B) \geqq a\}$ for the compacts. If a set in one of these classes is embedded in a product space, it admits a reasonable selector. These properties make the proofs go; but I do not see any other interesting class of sets with these properties.

(c) Theorem (6) is a variation on Dubins and Savage (1965, Theorem 1, page 36).

2. Compact metric selectors. Let S be a compact metric set. Let 2^S be the set of all nonempty compact subsets of S, in the usual compact metric topology (Hausdorff (1957, Section 28) or Kuratowski (1968, Section 42–43)). Endow 2^S with the Borel σ-field generated by the topology. Each open subset U of S generates two of the sub-basic open sets of 2^S:

$$\{K\colon K \in 2^S \text{ and } K \subset U\} \quad \text{and} \quad \{K\colon K \in 2^S \text{ and } K \cap U \neq \phi\}.$$

Open sets in S are F_σ's, and closed sets are G_δ's, so either class of sub-basic open sets generates the full Borel σ-field in 2^S. So

(9) Let Ω be a set endowed with a σ-field $\mathscr{F}$. Let f_n and f be functions from Ω to 2^S.

(a) f is $\mathscr{F}$-measurable iff $\{\omega : f(\omega) \subset U\} \in \mathscr{F}$ for all open subsets U of S.

(b) f is $\mathscr{F}$-measurable iff $\{\omega : f(\omega) \cap K \neq \phi\} \in \mathscr{F}$ for all compact subsets K of S.

(c) Suppose each f_n is $\mathscr{F}$-measurable, and $f_1(\omega) \supset f_2(\omega) \supset \cdots$ for all ω. Then $\omega \to \bigcap_n f_n(\omega)$ is $\mathscr{F}$-measurable.

(10) The map $x \to \{x\}$ is continuous and one-to-one from S into 2^S. This map has a compact range, and a continuous inverse.

The next result is mentioned in Kuratowski (1968, Section 43):

(11) LEMMA. *Let S be compact metric. There is a measurable function σ from 2^S into S, such that $\sigma(K) \in K$ for all $K \in 2^S$.*

PROOF. For each n, construct a finite collection $\mathscr{C}_n = \{K_{n1}, K_{n2}, \cdots\}$ of compact subsets of S, with $\mathscr{C}_1 = \{S\}$, and $d_n \to 0$ as $n \to \infty$, where

$$d_n = \max_j \quad \text{diameter} \quad K_{nj}\ ,$$

and this nesting property: for each n, there are positive integers $j_1 = 1 < j_2 < j_3 < \cdots$ such that

$$K_{ni} = K_{(n+1)j_i} \cup \cdots \cup K_{(n+1)(j_{i+1}-1)}\ .$$

Let $f_n(K) = K \cap K_{nj}$, where $j = j(K)$ is the least index with $K \cap K_{nj} \neq \phi$. Then f_n is measurable by (9b), and $f_n \supset f_{n+1}$, so $\bigcap_n f_n$ is measurable by (9c). But the diameter of $f_n(K)$ is at most d_n, so $\bigcap_n f_n(K)$ consists of a single point, $\sigma(K)$. and $K \to \sigma(K)$ is measurable by (10). Finally, $f_n(K) \subset K$, so $\sigma(K) \in K$. □

(12) COROLLARY. *Let S and T be compact metric sets. Let A be a compact subset of $S \times T$, and let $B = \mathrm{proj}_S\, A$. So B is compact. There is a Borel selector t for A, that is, a Borel function from B to T, with $(x, t(x)) \in A$ for all $x \in B$.*

PROOF. Suppose A is nonempty. Let A_x be the x-section of A, for $x \in B$. So $A_x \in 2^S$, and $x \to A_x$ is measurable by (9b). Compose this function with the σ of (11): that is, let $t(x) = \sigma(A_x)$. □

(13) COROLLARY. *Let S and T be compact metric sets. Let A be an F_σ-subset of $S \times T$, and let $B = \mathrm{proj}_S\, A$. So B is an F_σ. There is a Borel selector t for A.*

PROOF. Let $A = \bigcup_n A_n$, where $A_1 \subset A_2 \subset \cdots$ are closed. Let $B_n = \mathrm{proj}_S\, A_n$, so $B_1 \subset B_2 \subset \cdots$ are closed, and $B = \bigcup_n B_n$. Let t_n be a Borel selector for A_n, as constructed in (12). Let $t = t_1$ on B_1, and $t = t_n$ on $B_{n+1} \backslash B_n$ for $n = 1, 2, \cdots$. □

(14) COROLLARY. *Let S and T be compact metric sets. Let h be a nonnegative, real-valued, upper semi-continuous function defined on a closed subset A of $S \times T$, with* $\mathrm{proj}_S\, A = S$. *Let*

$$h^*(x) = \sup_{y \in A_x} h(x, y)\ .$$

(a) *For each* x, *the set of* $y \in A_x$ *with* $h(x, y) = h^*(x)$ *is nonempty and closed.*

(b) h^* *is a nonnegative, real-valued upper semi-continuous function on* S.

(c) *There is a Borel function* t *from* S *to* T, *with* $(x, t(x)) \in A$ *and* $h(x, t(x)) = h^*(x)$ *for all* x.

PROOF. Claim (a). Let $h(x, y_n) \to h^*(x)$. By passing to a subsequence, suppose $y_n \to y$. Then $h^*(x) \geqq h(x, y) \geqq \lim h(x, y_n) = h^*(x)$.

Claim (b). Using (a),

$$\{x: x \in S \text{ and } h^*(x) \geqq a\} = \text{proj}_S \{(x, y): (x, y) \in A \text{ and } h(x, y) \geqq a\}.$$

Claim (c). Let $C(x) = \{y: y \in A_x \text{ and } h(x, y) = h(x)\}$. So $C(x) \in 2^T$ by (a). To show $x \to C(x)$ is measurable, use (9b): fix $K \in 2^T$, and let $\phi(x) = \sup_{y \in K \cap A_x} h(x, y)$. Then ϕ is upper semi-continuous by (b). And

$$\{x: C(x) \cap K \neq \phi\} = \{x: \phi(x) = h^*(x)\}$$

is Borel in S. □

(15) COROLLARY. *Let* S *and* T *be compact metric sets. Let* h *be a nonnegative, extended real-valued function of type* F_σ *defined on an* F_σ*-subset* A *of* $S \times T$. *Suppose* $\text{proj}_S A = S$. *Let*

$$h^*(x) = \sup_{y \in A_x} h(x, y).$$

(a) h^* *is a nonnegative, extended real-valued function of type* F_σ *on* S.

(b) *If* $\varepsilon > 0$, *there is a Borel function* t *from* S *to* T, *with* $(x, t(x)) \in A$ *for all* x, *and*

$$\begin{aligned} h(x, t(x)) &> h^*(x) - \varepsilon \quad &\textit{when} \quad h^*(x) < \infty \\ &> 1/\varepsilon \quad &\textit{when} \quad h^*(x) = \infty. \end{aligned}$$

PROOF. Claim (a). Clearly, $\{h^* > a\} = \text{proj}_S \{A \text{ and } h > a\}$.

Claim (b). Fix a positive integer $k > 1/\varepsilon$. Let $A_j = \{A \text{ and } h > j/k\}$, an F_σ in $S \times T$. Let $B_j = \text{proj}_S A_j$, an F_σ in S. Clearly, $\{h > 0\} = A_0 \supset A_1 \supset \cdots$. So $B_0 \supset B_1 \supset \cdots$. Let $B^* = \bigcap_n B_n$, a Borel subset of S. Let $\bar{t}$ be a Borel selector for A, as constructed in (13). Let t_j be a Borel selector for A_j, as constructed in (13). Let

$$\begin{aligned} t &= \bar{t} \quad &\text{on} \quad &S \backslash B_0 \\ &= t_j \quad &\text{on} \quad &B_{j+1} \backslash B_j, \qquad \text{for } j = 0, 1, \cdots \\ &= t_k \quad &\text{on} \quad &B^*. \end{aligned}$$

Clearly, t is a Borel selector for A. If $x \in S \backslash B_0$, then $h^*(x) = 0$, and $h(x, t(x)) > h^*(x) - \varepsilon$. If $x \in B_{j+1} \backslash B_j$, then $h^*(x) \leqq (j + 1)/k$, and $h(x, t(x)) > j/k > h^*(x) - \varepsilon$. Finally, if $x \in B^*$, then $h^*(x) = \infty$ and $h(x, t(x)) > k > 1/\varepsilon$. □

3. Open selectors.

(16) LEMMA. *Let* S *and* T *be separable metric sets. Let* A *be an open subset of*

$S \times T$. Then $B = \text{proj}_S A$ is an open subset of $S \times T$. There is a Borel selector t for A.

PROOF. Suppose A is nonempty. Let $\{y_1, y_2, \cdots\}$ be a dense subset of T. For $x \in B$, let $t(x)$ be the y_n with least n such that $(x, y_n) \in A$. □

(17) COROLLARY. *Let S and T be separable metric sets. Let h be a nonnegative, extended real-valued, lower semi-continuous function defined on an open subset A of $S \times T$. Suppose $\text{proj}_S A = S$. Let*

$$h^*(x) = \sup_{y \in A_x} h(x, y) .$$

(a) *h^* is nonnegative, extended real-valued, lower semi-continuous function on S.*

(b) *If $\varepsilon > 0$, there is a Borel function t from S to T, with $(x, t(x)) \in A$ for all x, and*

$$\begin{aligned} h(x, t(x)) &> h^*(x) - \varepsilon \quad \text{when} \quad h^*(x) < \infty \\ &> 1/\varepsilon \quad\quad\quad\ \text{when} \quad h^*(x) = \infty . \end{aligned}$$

PROOF. As in (15). □

4. The weak* topology. There is a quick review of the weak* topology in [1], and a detailed discussion in [6]. The proofs of the next three results are omitted, being routine.

(18) LEMMA. *Let X be a compact metric set. Let $\pi(X)$ be the set of probabilities on X, endowed with the weak* topology. Let r be a nonnegative, real-valued upper semi-continuous function on X. Then $\mu \to \int_X r\, d\mu$ is upper semi-continuous on $\pi(X)$.*

(19) LEMMA. *Let X be a compact metric set. Let $\pi(X)$ be the set of probabilities on X, endowed with the weak* topology. Let r be a nonnegative, extended real-valued function on X, of type F_σ. Then $\mu \to \int_X f\, d\mu$ is of type F_σ on $\pi(X)$.*

(20) LEMMA. *Let X be a separable metric set. Let $\pi(X)$ be the set of probabilities on X, endowed with weak* topology. Let r be a nonnegative, extended real-valued function on X, which is lower semi-continuous. Then $\mu \to \int_X r\, d\mu$ is lower semi-continuous on $\pi(X)$.*

5. Proving the theorems. The proofs are very similar to one another and to the argument in [1]. So I will sketch the proof of (6), and omit the other arguments. Suppose condition (3).

(21) LEMMA. *Let $h(x, P) = \int_S r(x, P, y)P(dy)$, for $(x, P) \in A$. Then h is upper semi-continuous on A.*

PROOF. Let $\varphi = \varphi_{(x,P)}$ map A into $\pi(A \times S)$ by sending (x, P) into P installed on the (x, P) slice of $A \times S$. More formally, if $\phi = \phi(x, P, y)$ is a continuous function on $A \times S$,

$$\int_{A \times S} \phi\, d\varphi_{(x,P)} = \int_S \phi(x, P, y)P(dy) .$$

Then φ is continuous, for ϕ can be unifomly approximated by sums of functions of the form $\phi_1(x)\phi_2(P)\phi_3(y)$, each ϕ_i being continuous. Now h is the composition of the function $\mu \to \int_{A\times S} r\,d\mu$, upper semi-continuous on $\pi(A \times S)$ by (18), with the continuous mapping φ from A to $\pi(A \times S)$. □

(22) COROLLARY. *If f is a nonnegative, real-valued, upper semi-continuous function on S, so is Tf.*

PROOF. The modified reward function $r(x, P, y) + f(y)$ is still upper semi-continuous on $A \times S$. Use (21) on this modified r:

$$h(x, P) = \int_S [r(x, P, y) + f(y)]P(dy)$$

is upper semi-continuous on A. Now use (14b). □

(23) COROLLARY. *Suppose* (3). *Then T^n0 is upper semi-continuous on S.*

(24) LEMMA. *Suppose* (3). *For each n, there is an optimal Borel strategy.*

PROOF. The case $n = 1$. This is immediate from (21) and (14c).

The induction. Suppose the lemma holds for $n = k$. Let

$$r_k(x, P, y) = r(x, P, y) + (T^k0)(y),$$

which is upper semi-continuous on $A \times S$ by (23). If you integrate out y against P, and sup out $P \in A_x$, you get $(T^{k+1}0)(x)$, by (1). Use the case $n = 1$ on the reward function r_k to get a Borel strategy t of length one, such that

$$\int_S [r(x, t_x, y) + (T^k0)(y)]t_x(dy) = (T^{k+1}0)(x).$$

Use the induction hypothesis to generate a Borel strategy t^* of length k, with $\rho_k(t^*, y) = (T^k0)(y)$. There is a unique Borel strategy s of length $k + 1$, such that

$$s_x = t_x \quad \text{and} \quad s^x = t^* \qquad \text{for all } x \in S.$$

And s is optimal by (2). □

Note. The optimal Borel strategy s constructed by this induction is Markovian: $s_{x_1,\cdots,x_i}$ depends only on i and x_i.

Note on proving (8). If you follow the pattern set by (21), you have to show that φ is continuous. This is less obvious. However, as explained in [1], (10–12), you can embed S into a compact metric set S^*; this embeds $\pi(S)$ into the set of $\mu \in \pi(S^*)$ which assign outer measure 1 to S. The continuity in the general case now follows from the continuity in the compact case.

REFERENCES

[1] BLACKWELL, D., FREEDMAN, D., and ORKIN, M. (1974). The optimal reward operator in dynamic programming. *Ann. Probability* **2** 926–941.

[2] DUBINS, L. E. and SAVAGE L. J. (1965). *How to Gamble if You Must.* McGraw-Hill, New York.

[3] HAUSDORFF, F, (1957). *Set Theory.* Chelsea, New York.

[4] KURATOWSKI, K. (1966). *Topology*, **1** Academic Press, New York.
[5] KURATOWSKI, K. (1968). *Topology*, **2** Academic Press, New York.
[6] PARTHASARATHY, K. (1967). *Probability Measures on Metric Spaces*. Academic Press, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720