The Annals of Mathematical Statistics
1971, Vol. 42, No. 1, 297–303

INFINITESIMAL LOOK-AHEAD STOPPING RULES[1]

BY SHELDON M. ROSS

University of California, Berkeley

1. Introduction. Let $X = (X_t, t \geqq 0)$ be a strong Markov Process having stationary transition distributions, and sample paths which are almost surely right continuous and have only jump discontinuities. The state space S of the process is assumed to be a Borel subset of a complete separable metric space and we consider the problem of selecting a stopping time τ maximizing

$$E^x[e^{-\lambda\tau} f(X_\tau) - \int_0^\tau e^{-\lambda s} c(X_s)\, ds], \tag{1}$$

where f and c are continuous real-valued functions on S, $\lambda \geqq 0$, and E^x denotes expectation conditional on $X_0 = x$.

In the second section of this paper, we show that under certain conditions an infinitesimal look-ahead procedure is optimal. This result generalizes certain discrete time results given by Derman–Sacks (1960) in [5] and independently by Chow–Robbins (1961) in [4]. In the third section, a related approach is described and the resultant procedure is shown to be optimal under slightly more general situations. The fourth section considers a class of continuous time Markovian Decision Processes for which the criterion function is closely related to (1).

2. Infinitesimal look-ahead stopping rule. A stopping time τ is defined to be any nonnegative extended real-valued random variable such that for all $t > 0$, $\{\tau \leqq t\}$ is contained in the sigma field generated by $\{X_s, 0 \leqq s \leqq t\}$. A stopping time τ^* is said to be optimal at $x \in S$ if

$$E^x[e^{-\lambda\tau^*} f(X_{\tau^*}) - \int_0^{\tau^*} e^{-\lambda s} c(X_s)\, ds] = \max_\tau E^x[e^{-\lambda\tau} f(X_\tau) - \int_0^\tau e^{-\lambda s} c(X_s)\, ds].$$

If τ^* is optimal at x for every $x \in S$, then it is said to be optimal.

Define the infinitesimal operator $\alpha(x)$ by

$$\alpha(x) = \lim_{h\to 0^+} E^x\left[\frac{f(X_h) - f(x)}{h}\right], \qquad x \in S. \tag{2}$$

We assume that f and X are such that the limit in (2) exists.

We first state the following well-known result. For a proof, the reader should consult Breiman [3], page 376.

LEMMA 2.1. *Suppose that both f and α are bounded and continuous.*

(a) *For any stopping time τ and $\lambda > 0$,*

$$E^x[e^{-\lambda\tau} f(X_\tau)] - f(x) = E^x[\int_0^\tau e^{-\lambda s}(\alpha(X_s) - \lambda f(X_s))\, ds].$$

Received December 12, 1968.

[1] This research has been partially supported by the Office of Naval Research under Contract Nonr-222(83), the National Science Foundation under Grant GP-8695 and the U.S. Army Research Office-Durham under Contract DA-31-124-ARO-D-331 with the University of California.

(b) *For any stopping time* τ *such that* $E^x\tau < \infty$,

$$E^x[f(X_\tau)] - f(x) = E^x[\int_0^\tau \alpha(X_s)\, ds].$$

Now define the set $B_\lambda \subset S$ *as follows*

$$B_\lambda = \{x : \alpha(x) - \lambda f(x) - c(x) \leqq 0\}$$

and let

$$\tau_\lambda^* = \inf\{t \geqq 0 : X_t \in B_\lambda\}.$$

Let P^x *denote probability conditional on* $X_0 = x$.

THEOREM 2.2. *Suppose that f and* α *are bounded and continuous.*

(a) *For* $\lambda > 0$, *if* B_λ *is closed in the sense that*

$$P^x\{\exists\, t \geqq 0 : X_t \notin B_\lambda\} = 0 \quad \textit{for all} \quad x \in B_\lambda,$$

and if τ_λ^* *is finite with probability one for all starting points* $x \in S$, *then* τ_λ^* *is optimal.*

(b) *For* $\lambda = 0$, *if* B_0 *is closed in the sense of Part* (a), *if* $\inf_x c(x) > 0$, *and if* τ_0^* *is finite with probability one for all starting points* $x \in S$, *then* τ_0^* *is optimal.*

PROOF.

(a) By Lemma 2.1, we have reduced the problem to one in which there is no reward given for stopping, and there is a cost $\alpha(x) - \lambda f(x) - c(x)$ per unit time for being in state x. The result follows from this.

(b) If $E^x\tau = \infty$ and $\inf_x c(x) > 0$, then $E^x[f(X_\tau) - \int_0^\tau c(X_s)\, ds] = -\infty$, and hence we need only consider rules such that $E^x\tau < \infty$. Now, by Lemma 2.1, we have that

$$E^x[f(X_\tau) - \int_0^\tau c(X_s)\, ds] = f(x) + E^x[\int_0^\tau (\alpha(X_s) - c(X_s))\, ds] \tag{3}$$

whenever $E^x\tau < \infty$. The argument now follows as in Part (a). □

What we have done can perhaps best be described as follows: We define τ_λ^*, the infinitesimal look-ahead (ILA) rule, to be the one which stops at state x if and only if the infinitesimal look-ahead gain is no greater than the stopping gain. Theorem 2.2 then says that if the set of stopping states is closed then τ_λ^* is optimal. This result is clearly the continuous time analogue of the Derman–Sacks, Chow–Robbins result of optimality of the one-stage look-ahead rule in the monotonic case (see [4] or [5]).

EXAMPLE 1. Let $Y_1, Y_2, \cdots$ be a sequence of i.i.d. bounded random variables with cdf F, and let $(N_t, t \geqq 0)$ be a nonhomogeneous Poisson Process, independent of the Y_i's, and with a continuous non-increasing rate function $\mu(t)$. Let $M_t = \max(Y_1, \cdots, Y_{N_t})$, and consider the Markov Process $\{X_t = (t, M_t), t \geqq 0\}$. We take $f(t, m) = m$ and assume that $c(t, m)$ is non-decreasing in both t and m. This is, of course, the continuous time analogue of the famous house-selling problem (though for the sake of generality we have not required that $F(0) = 0$, see [4], [5] and [10]).

$$\alpha(t, m) = \lim_{h\to 0} E\left[\frac{M_{t+h}-M_t}{h}\,\middle|\, M_t = m\right]$$
$$= \mu(t)E[\max(Y_i, m) - m]$$
$$= \mu(t)\int y\, dF(y+m).$$

Since $\int_y y\, dF(y+m)$ is non-increasing in m, it follows from Theorem 2.2 that

$$\tau_\lambda^* = \inf\{t \geqq 0: \mu(t)\int y\, dF(y+M_t) \leqq c(t, M_t) + \lambda M_t\}$$

is optimal.

Example 2. Now, consider Example 1 with the exception that once an offer is rejected it is no longer available. Clearly, the optimal return for this problem is no greater than the optimal return for the original problem. Thus, since the optimal policy τ^* is a legitimate policy for this new problem (as it never accepts an old offer) it follows that it is also optimal for this problem. This is related to certain results given by Elfving [8] and Siegmund [13].

Example 3. Let $(N_t, t \geqq 0)$ be a Poisson Process with rate μ and consider the Markov Process $(X_t = (t, N_t), t \geqq 0)$. Let $c(t, N_t) = N_t$ and $f(t, N_t) = -\mu(T-t)^2/2$, where T is some fixed constant. Then $\alpha(t, N_t) = \mu(T-t)$ and from Theorem 2.2 it follows that $\tau_\lambda^* = \inf\{t \geqq 0: N_t \geqq \mu(T-t)(1+\lambda(T-t)/2)\}$ is optimal. This problem arises in determining the optimal intermediate time to dispatch a Poisson Process (see Ross [11]).

Example 4. Let $(N_t, t \geqq 0)$ be a nonhomogeneous Poisson Process with rate $\mu(t)$. Suppose the reward for stopping when $N_t = x$ is x and the continuation rate at $N_t = x$ is $c(x)$. Suppose further that $\mu(t)$ is continuous non-increasing, and $c(x)$ is continuous non-decreasing, and let λ be the discount factor.

The state space is thus $X_t = (t, N_t)$ and

$$\alpha(t, n) = \lim_{h\to 0} E^{t,n}\left[\frac{N_h - N_0}{h}\right]$$
$$= \mu(t).$$

Thus, from Theorem 2.2, we have that

$$\tau_\lambda^* = \inf\left\{t \geqq 0: N_t \geqq \frac{\mu(t) - c(N_t)}{\lambda}\right\}$$

is optimal. This example with $c(x) \equiv c$, and $\mu(t) \equiv \mu$ was treated in Taylor [14] by a different method.[2]

[2] Taylor's answer differs somewhat from ours as he supposed that $[\log_e(1+\lambda/\mu)]^{-1} - c/\lambda$ was an integer.

3. A related approach. Let $Z_t = e^{-\lambda t} f(X_t) - \int_0^t e^{-\lambda s} c(X_s)\, ds$, and let $\bar{B}_\lambda = \{x \in S : E^x Z_t \leqq f(x) \text{ for all } t \geqq 0\}$. Thus, $\bar{B}_\lambda$ is the set of states at which stopping is better than continuing for any *fixed* amount of time.

LEMMA 3.1. *If* $P^x\{\exists\, t \geqq 0 : X_t \notin \bar{B}_\lambda\} = 0$ *for all* $x \in \bar{B}_\lambda$, *then*

$$E^x[Z_t \mid X_u, 0 \leqq u \leqq s] \leqq Z_s \quad \text{a.s.} \qquad \forall\, x \in \bar{B}_\lambda, \quad \forall\, s < t.$$

PROOF.

$$\begin{aligned} E^x[Z_t \mid X_u, 0 \leqq u \leqq s] &= Z_s + e^{-\lambda s} E^x[e^{-\lambda(t-s)} f(X_t) - \textstyle\int_s^t e^{-\lambda(u-s)} c(X_u)\, du - f(X_s) \mid X_s] \\ &= Z_s + e^{-\lambda s} E^{X_s}[e^{-\lambda(t-s)} f(X_{t-s}) - \textstyle\int_0^{t-s} e^{-\lambda u} c(X_u)\, du - f(X_0) \mid X_s]. \end{aligned}$$

Since $x \in \bar{B}_\lambda$ implies by hypothesis that $X_s \in \bar{B}_\lambda$ a.s., the result follows from the definition of $\bar{B}_\lambda$. □

LEMMA 3.2. *If* $E^x|Z_\tau| < \infty$ *and* $\liminf_t \int_{\tau > t} |Z(t)|\, dP^x = 0$ *for all* x *and all* τ, *and if*

$$P^x\{\exists\, t \geqq 0 : X_t \notin \bar{B}_\lambda\} = 0 \qquad \forall\, x \in \bar{B}_\lambda,$$

then

$$E^x Z_\tau \leqq f(x) \qquad \forall\, x \in \bar{B}_\lambda, \quad \forall\, \tau.$$

PROOF. From Lemma 3.1, we have that $(Z_t, t \geqq 0)$ is a supermartingale whenever the initial state is in $\bar{B}_\lambda$. Thus, the result follows from a standard supermartingale systems theorem (see Breiman [3], page 302).

THEOREM 3.3. *Under the conditions of Lemma* 3.2,

$$\begin{aligned} \sup_\tau E^x(Z_\tau) &= f(x) \qquad x \in \bar{B}_\lambda \\ &= > f(x) \qquad x \notin \bar{B}_\lambda. \end{aligned}$$

PROOF. When $x \in \bar{B}_\lambda$, sup $= f(x)$ from Lemma 3.2. When $x \notin \bar{B}_\lambda$, the result follows from the definition of $\bar{B}_\lambda$. □

Let us define the stopping time $\bar{\tau}_\lambda$ by

$$\bar{\tau}_\lambda = \inf\{t \geqq 0 : X_t \in \bar{B}_\lambda\}.$$

Now, suppose that there exists an optimal stopping rule τ for x. Let $\tau_1 = \min(\tau, \bar{\tau}_\lambda)$, then from Lemma 3.2 it follows that τ_1 is also optimal for x.[3] But it is easily seen that τ_1 must be a.s. equal to $\bar{\tau}_\lambda$, for if $\tau_1 < \bar{\tau}_\lambda$ with positive probability then with positive probability τ_1 will stop in some final state $y \notin \bar{B}_\lambda$. However, for each $y \notin \bar{B}_\lambda$, there is a fixed time t_y such that $E^x Z_{t_y} > f(y)$. Hence, if we define the stopping time τ_2 by

$$\begin{aligned} \tau_2 &= \tau_1 \qquad \text{if } \tau_1 \text{ stops in a state } y \in \bar{B}_\lambda, \\ &= \tau_1 + t_y \qquad \text{if } \tau_1 \text{ stops in state } y \notin \bar{B}_\lambda, \end{aligned}$$

[3] All of this is assuming, of course, the conditions of Lemma 3.2.

then, using the strong Markov property, it follows that τ_2 will be strictly better than τ_1, which would contradict the optimality of τ_1. Therefore, if there exists an optimal rule for each x, then it follows that $\bar{\tau}_\lambda$ is optimal. It should also be noted that $\bar{\tau}_\lambda$ is just the continuous time analogue of the functional equation rule (see Bellman [1]).

Some sufficient conditions for the existence of an optimal stopping rule are given in Dynkin [7] and Taylor [14]. To determine the connection between $\bar{\tau}_\lambda$ and the ILA $\tau_\lambda{}^*$, we first note that $\bar{B}_\lambda \subset B_\lambda$ and so $\bar{\tau}_\lambda \geqq \tau_\lambda{}^*$. To go the other way, we need the following:

COROLLARY 3.4. *If the conditions of Theorem* 2.2 *hold, then*

$$\bar{B}_\lambda = B_\lambda \qquad \textit{and hence}$$

$$\bar{\tau}_\lambda = \tau_\lambda{}^*.$$

PROOF. If $x \in B_\lambda$, then since B_λ is closed, it follows that $E^x[\int_0^t e^{-\lambda s}(\alpha(X_s) - \lambda f(X_s) - c(X_s))\,ds] \leqq 0$ for all t, and so the result follows from Lemma 2.1 and Lemma 3.1. □

Aside from its own interest, the reason we have considered this approach is that one may easily construct examples in which $\bar{B}_\lambda$ is closed but B_λ is not. The idea is also illuminating, and we paraphrase it as follows: Call a state bad if stopping at that state is better than continuing from that state for any *fixed* amount of time. Then if this set of states is closed and if an optimal rule exists, then the rule which stops the first time it enters a bad state is optimal. Since a discrete time Markov Process may be regarded as a continuous time Markov Process (with $X_t' = (t, X_t)$), it follows that this result also holds for the discrete time problem.

4. Some related criteria. In this section, we consider the problem of choosing a stopping time τ maximizing either

$$\psi_\tau = \frac{E^x[e^{-\lambda t} f(X_\tau) - \int_0^\tau e^{-\lambda s} c(X_s)\,ds]}{E^x[1 - e^{-\lambda\tau}]}, \qquad \text{where } \lambda > 0 \tag{4}$$

or

$$\phi_\tau = \frac{E^x[f(X_\tau) - \int_0^\tau c(X_s)\,ds]}{E^x\tau}, \qquad \text{where } 0 < E^x\tau < \infty. \tag{5}$$

Criterion (4) represents expected total discounted return, and (5) the long-run average return per unit time, when a sequence of independent stopping games are played, each starting at x. These criterions also arise in connection with a 2-action, continuous time Markovian Decision Process (see [12]) in which the "stop" action resets the process to a fixed initial state x. (In this connotation, $-f(y)$ is usually thought of as the cost of resetting from state y.) For any constant b, let

$$\psi_\tau(b) = (\psi_\tau - b)E^x[1 - e^{-\lambda\tau}] \tag{6}$$

$$= E^x[e^{-\lambda\tau} f(X_\tau) - \int_0^\tau e^{-\lambda s}(c(X_s) + \lambda b)\,ds], \tag{7}$$

and let

(8) $$\phi_\tau(b) = (\phi_\tau - b)E^x\tau$$

(9) $$= E^x[f(X_\tau) - \int_0^\tau (c(X_s) + b)\, ds].$$

LEMMA 4.1.

(i) *If for some* b, $0 = \psi_{\tau^*}(b) = \max_\tau \psi_\tau(b)$, *then* $b = \psi_{\tau^*} = \max_\tau \psi_\tau$, *and conversely*;

(ii) *If for some* b, $0 = \phi_{\tau^*}(b) = \max_{\tau\in\Gamma} \phi_\tau(b)$ *where* $\Gamma = \{\tau : 0 < E^x\tau < \infty\}$, *then* $b = \phi_{\tau^*} = \max_{\tau\in\Gamma}\phi_\tau$, *and conversely.*

PROOF. Follows directly from (6) and (8).

REMARK. Part (i) of the above lemma seems to be new, as criterion (4) does not seem to have been previously considered in stopping rule literature. Part (ii) is not new and may be found in either Breiman [2] or Taylor [14].

We shall suppose for the remainder that optimal rules exist for criterions (4) and (5) and we let $V = \max_\tau \psi_\tau$, and $g = \max_{\tau\in\Gamma}\phi_\tau$.

THEOREM 4.2. *Under the conditions of Theorem* 2.2,

(i) *If* $B_1 = \{x : \alpha(x) - \lambda f(x) - c(x) \leqq \lambda V\}$ *is closed, and if* $\tau^* = \inf\{t \geqq 0 : X_t \in B_1\}$ *is* a.s. *finite for all starting states, then* τ^* *is optimal for* (4).

(ii) *If* $B_2 = \{x : \alpha(x) - c(x) \leqq g\}$ *is closed, and if* $\tau^* = \inf\{t \geqq 0 : X_t \in B_2\}$ *is* a.s. *finite for all starting states, then* τ^* *is optimal for* (5) *whenever* $0 < E^x\tau^* < \infty$.

PROOF. Follows directly from (7), (9) and Theorem 2.2. □

REMARK. Even though V and g are in general unknown, Theorem 4.2 is quite useful as it often enables us to determine the structure of the optimal rule. Also when bounds are known on V or g, then Theorem 4.2 may be applied to determine the optimal action at least in some states. For example, if $L \leqq V \leqq U$ and the conditions of Theorem 4.2 are satisfied, then it follows that the optimal rule stops in state x whenever $\alpha(x) - \lambda f(x) - c(x) \leqq \lambda L$ and continues whenever $\alpha(x) - \lambda f(x) - c(x) > \lambda U$.

EXAMPLE 5. Let $(N_t, t \geqq 0)$ be any right-continuous counting process with left limits. Let $c(n)$ be the cost rate when there are n in the system, and suppose that $c(n)$ is nonnegative and non-decreasing. Let $f(x) \equiv -R$ (i.e., R is the reset cost). Then the related Markov Process is

$$X_t = (N_s, s \leqq t), \qquad \text{and}$$

$$\alpha(X_t) = 0.$$

Hence, $\tau^* = \inf\{t \geqq 0 : c(N_t) \geqq \lambda(R - V)\}$ is optimal for (4), and $\tau^* = \inf\{t \geqq 0 : c(N_t) \geqq -g\}$ is optimal for (5). Thus, for the average cost case, it is optimal to reset the process whenever the present cost rate is at least as large as the optimal average cost per unit time.

Acknowledgment. I thank the referee for his helpful comments.

REFERENCES

[1] BELLMAN, R. (1957). *Dynamic Programming.* Princeton Univ. Press.

[2] BREIMAN, L. (1964). Stopping rule problems. *Applied Combinatorial Mathematics* (E. Beckenbach, Ed.). Wiley, New York.

[3] BREIMAN, L. (1968). *Probability.* Addison-Wesley, Reading.

[4] CHOW, Y. S. and ROBBINS, H. (1961). A martingale systems theorem and applications. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.

[5] DERMAN, C. and SACKS, J. (1960). Replacement of periodically inspected equipment, (An optimal optional stopping rule). *Naval Res. Logist. Quart.* **7** 597–607.

[6] DOOB, J. L. (1957). *Stochastic Processes.* Wiley, New York. (1957).

[7] DYNKIN, E. B. (1963). Optimal selection of stopping time for a Markov process. *Dokl. Adak. Nauk. USSR* **150** 238–240. (English translation in *Soviet Math.* **4** 627–629).

[8] ELFVING, G. (1967). A persistency problem connected with a point process. *J. Appl. Probability* **4** 77–89.

[9] GRIGELIONIS, B. I. and SHIRYAEV, A. N. (1966). On Stetans problem and optimal stopping rules for Markov processes. *Theor. Probability Appl.* **11** 541–555.

[10] MACQUEEN, J. and MILLER, R. G. Optimal persistency policies. *Operations Res.* **8** 362–380.

[11] ROSS, S. M. (1969). Optimal dispatching of a Poisson process. *J. Appl. Probability* **6** 692–699.

[12] ROSS, S. M. (1968). Arbitrary state Markovian decision processes. *Ann. Math. Statist.* **39** 2118–2123.

[13] SIEGMUND, D. O. (1967). Some problems in the theory of optimal stopping rules. *Ann. Math. Statist.* **38** 1627–1640.

[14] TAYLOR, H. (1968). Optimal stopping in a Markov process. *Ann. Math. Statist.* **39** 1333–1344.