Journal of the Operations Research
Society of Japan
Vol. 28, No. 1, March 1985

SEMI-MARKOV DECISION PROCESSES AND THEIR APPLICATIONS IN REPLACEMENT MODELS

Masami Kurano
Chiba University

(Received January 13, 1984: Final November 8, 1984)

Abstract We consider the problem of minimizing the long-run average expected cost per unit time in a semi-Markov decision process with arbitrary state and action space. Using the idea of successive approximations, sufficient conditions for the existence of an optimal stationary policy are given. These results are applied to solve the replacement problem with a semi-Markov shock model.

1. Introduction

We consider the problem of minimizing the long-run average expected cost per unit time in a semi-Markov decision process(semi-MDP) with arbitrary state and action space. Lippman[9] has given the sufficient conditions under which, for each $\varepsilon > 0$, there is a stationary policy which is ε-optimal. We, using the idea of successive approximations(for example, see[16]), give another conditions for the existence of an ε-optimal policy that is stationary. These results are used to analyse the replacement model with a semi-Markov shock process under the average cost criterion. The related replacement model is investigated by Kao[6], Taylor[13], Feldman[4] and Zuckerman[18]. Taylor[13] studied the case in which the damage process is a compound Poisson process. Zuckerman[18] derived an optimal replacement policy in an extended model by using the infinitesimal operator of the Markov process. We analyse Zuckerman's model by an approach from a semi-MDP and derive an optimal replacement policy which is contained in the class of state-age replacement rules of Kao[6].

2. Semi-MDP's with Arbitrary State

In this section we formulate a semi-MDP with arbitrary state and action space refering to Federgruen and Tijms[3] and Kolonko[7]. By a Borel set we mean a Borel subset of some complete separable metric space. For a Borel set X, F_X denote the Borel subsets of X. If X is a non-empty Borel set, $\mathsf{B}_m(X)$ [$\mathsf{B}_a(X)$] denotes the set of all bounded Baire[lower semianalytic] functions on X. The product space of the set $D_1, D_2, \cdots$ will be denoted by $D_1 D_2 \cdots$. For any non-empty Borel sets X and Y, a transition probability measure on Y given X is a function $p(\cdot|\cdot)$ on $\mathsf{F}_Y X$ such that for each $x \in X$ $p(\cdot|x)$ is a probability measure on F_Y and for each Borel set $B \in \mathsf{F}_Y$ $p(B|\cdot)$ is a Baire function on X. The set of all transition probability measures on Y given X is denoted by $\mathsf{P}(Y|X)$. Also, we denote the set of all [analytically] measurable functions from X to Y by $\mathsf{B}_m(X \to Y)$ [$\mathsf{B}_a(X \to Y)$].

We consider a semi-Markov decision model specified by six objects (S, A, Z, Q, c, τ), where S, A and Z are any Borel sets and denote the state space, the action space and the space of additional observations respectively, $Q \in \mathsf{P}(ZS|SA)$ is the law of motion, $c \in \mathsf{B}_m(SAZS)$ is one step cost function and $\tau \in \mathsf{B}_m(SAZS \to (0,\infty))$ is a weighting function for defining the average cost.

Let Π denote the set of all policies, i.e., for $\pi = (\pi_0, \pi_1, \cdots) \in \Pi$ let $\pi_t \in \mathsf{P}(A|S(AZS)^t)$.

A policy $\pi = (\pi_0, \pi_1, \cdots)$ is called [analytically measurable] Markov policy if for each t there is a $f_t \in \mathsf{B}_m(S \to A)$ [$\mathsf{B}_a(S \to A)$] such that $\pi_t(\{f_t(x_t)\}|x_0, a_0, z_1, x_1, \cdots, z_t, x_t) = 1$ for all $(x_0, a_0, z_1, x_1, \cdots, z_t, x_t) \in S(AZS)^t$. Such policy will be denoted by $(f_0, f_1, \cdots)$. A Markov policy $(f_0, f_1, \cdots)$ with $f = f_t$ for each $t \geq 0$ is called stationary and denoted by f^∞.

The sample space is the product space $\Omega = S(AZS)^\infty$. Let X_t, Z_t and Δ_t be random quantities defined on Ω by $X_t(\omega) = x_t$, $Z_t(\omega) = z_t$ and $\Delta_t(\omega) = a_t$ for all $\omega = (x_0, a_0, z_1, x_1, a_1, \cdots) \in \Omega$.

For each policy $\pi = (\pi_0, \pi_1, \cdots)$ and $x \in S$, we assume that

$$\text{Prob}(\Delta_t \in D_1 \mid X_0 = x, \Delta_0, \cdots, Z_t, X_t) = \pi_t(D_1 | X_0 = x, \Delta_0, \cdots, Z_t, X_t)$$

and

$$\text{Prob}(Z_{t+1} \in D_2, X_{t+1} \in D_3 | X_0 = x, \cdots, Z_t, X_t, \Delta_t) = Q(D_2 D_3 | X_t, \Delta_t)$$

for each $t \geq 0$, $D_1 \in \mathsf{F}_A$, $D_2 \in \mathsf{F}_Z$ and $D_3 \in \mathsf{F}_S$.

Then, for each $\pi \in \Pi$ and starting state $x \in S$ we can define the probability measure P_π^x on Ω in an obvious way. We shall consider the following average cost criterion:

$$g[\pi](x) = \limsup_{T\to\infty} \frac{\sum_{t=0}^{T-1} E_\pi^x[c(X_t,\Delta_t,Z_{t+1},X_{t+1})]}{\sum_{t=0}^{T-1} E_\pi^x[\tau(X_t,\Delta_t,Z_{t+1},X_{t+1})]}$$

for $\pi\in\Pi$ and $x\in S$ if this expression exists, where E_π^x is the expectation operator with respect to P_π^x.

For $\varepsilon \geqq 0$, we say $\pi^*\in\Pi$ is ε-optimal if $g[\pi^*](x) \leqq g[\pi](x) + \varepsilon$ for all $\pi\in\Pi$ and $x\in S$. A 0-optimal policy is simply called optimal. We shall discuss the existence of an ε-optimal stationary policy by introducing the following condition.

Condition(*). There are a $v\in B_a(S)$ and a constant g^* satisfying that

$$(2.1)\qquad v(x) = \inf_{a\in A}\{ c(x,a) - g^*\tau(x,a) + \int v(x')Q^S(dx'|x,a) \}$$

where

$c(x,a) = \int c(x,a,z,x')Q(d(z,x')|x,a)$,

$\tau(x,a) = \int \tau(x,a,z,x')Q(d(z,x')|x,a)$ and $Q^S\in P(S|SA)$ is the state transition probability defined by $Q^S(D|x,a) = \int_{ZD} Q(d(z,x')|x,a)$ for each $D\in F_S$.

An examination of the proof of Theorem 7.6 in [12] and the selection theorem of [14] give the following theorem.

Theorem 2.1. Let Condition(*) be satisfied. Then

(a) $g[\pi](x) \geqq g^*$ for any $\pi\in\Pi$ and $x\in S$,

(b) for each $\varepsilon > 0$, there exists a $f\in B_a(S\to A)$ such that

$$(2.2)\qquad v(x) + \varepsilon \geqq c(x,f(x)) - g^*\tau(x,f(x)) + \int v(x')Q^S(dx'|x,f(x))$$

for any $x\in S$,

and the stationary policy f^∞ is ε-optimal.

Remark 2.1. In Theorem 2.1, if there exists a $f\in B_a(S\to A)$ satisfying the equation (2.2) for $\varepsilon = 0$, f^∞ is optimal.

Condition 1. There are positive constants m, M and M' such that $m < \tau(x,a) < M$ and $\int \tau^2(x,a,z,x')Q(d(z,x')|x,a) < M'$ for all $x\in S$ and $a\in A$.

For each $T \geqq 0$, let $n_T = \max\{ N \mid \sum_{t=0}^{N-1} \tau(X_t,\Delta_t,Z_{t+1},X_{t+1}) \leqq T \}$. Then, we have the following.

Corollary 2.1. Let Condition(*) and Condition 1 be satisfied. Then

(a) for any $\pi\in\Pi$

$$(2.3)\qquad \liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{n_T} c(X_t,\Delta_t,Z_{t+1},X_{t+1}) \geqq g^* \quad P_\pi^x\text{-a.s.},$$

and

(b) there is an analytically measurable Markov policy $\pi^* = (f_0, f_1, \cdots)$ which is optimal and

(2.4) $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{n_T} c(X_t, \Delta_t, Z_{t+1}, X_{t+1}) = g^* \quad P_{\pi^*}^{x}$ -a.s. .

Proof: For g^* and $v \in B_a(S)$ as in Theorem 2.1., we define

$$\phi(x,a,z,x') = v(x) + g^*\tau(x,a,z,x') - c(x,a,z,x') - v(x')$$

and

$$\phi(x,a) = \int \phi(x,a,z,x')Q(d(z,x')|x,a).$$

By the stability theorem of Loéve[10], it holds that

(2.5) $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \{ \phi(X_t, \Delta_t, Z_{t+1}, X_{t+1}) - \phi(X_t, \Delta_t)\} = 0 \quad P_{\pi}^{x}$ -a.s. .

For (a), since $\frac{1}{T} \sum_{t=0}^{T-1} \phi(X_t, \Delta_t) \leqq 0$, it follows from (2.5) that

(2.6) $\limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \phi(X_t, \Delta_t, Z_{t+1}, X_{t+1}) \leqq 0 \quad P_{\pi}^{x}$ -a.s. .

We observe that

(2.7) $$\sum_{t=0}^{T-1} \phi(X_t, \Delta_t, Z_{t+1}, X_{t+1}) = g^* \sum_{t=0}^{T-1} \tau(X_t, \Delta_t, Z_{t+1}, X_{t+1}) - \sum_{t=0}^{T-1} c(X_t, \Delta_t, Z_{t+1}, X_{t+1}) + v(X_0) - v(X_T).$$

Since Condition 1 implies that $\frac{1}{T} \sum_{t=0}^{n_T} \tau(X_t, \Delta_t, Z_{t+1}, X_{t+1}) \to 1 \quad P_{\pi}^{x}$ -a.s. as $T \to \infty$, we obtain (2.3) from (2.6) and (2.7).

For (b), from Theorem 2.1. for each $\varepsilon > 0$ there exists a $f^{\varepsilon} \in B_a(S \to A)$ such that

$$\phi(x, f^{\varepsilon}(x)) \geqq -\varepsilon \text{ for all } x \in S.$$

For a sequence $\{\varepsilon(t)\}_{t=0}^{\infty}$ such that $\varepsilon(t) > 0$ and $\frac{1}{T} \sum_{t=0}^{T-1} \varepsilon(t) \to 0$ as $T \to \infty$, define a Markov policy $\pi^* = (f_0, f_1, \cdots)$ by $f_t = f^{\varepsilon(t)}$ for each $t \geqq 0$. Since $0 \geqq \frac{1}{T} \sum_{t=0}^{T-1} \phi(X_t, f_t(X_t)) \geqq \frac{1}{T} \sum_{t=0}^{T-1} \{-\varepsilon(t)\} \to 0$ as $t \to \infty$, it holds $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \phi(X_t, f_t(X_t)) = 0$. Therefore, by repeating the above discussion we obtain (2.4). Also, from (2.7) and the definition of π^*, we have

$$g^* - E_{\pi^*}^{x}[\tilde{W}_T]^{-1} \sum_{t=0}^{T-1} E_{\pi^*}^{x}[c(X_t, \Delta_t, Z_{t+1}, X_{t+1})]$$
$$\geqq E_{\pi^*}^{x}[\tilde{W}_T]^{-1}(v(X_T) - v(X_0)) - E_{\pi^*}^{x}[\tilde{W}_T]^{-1} \sum_{t=0}^{T-1} \varepsilon(t),$$

where $\tilde{W}_T = \sum_{t=0}^{T-1} \tau(X_t, \Delta_t, Z_{t+1}, X_{t+1})$, so that , as $T \to \infty$, it holds from Condition 1 that $g^* \geqq g[\pi^*](x)$. So, by theorem 2.1 π^* is optimal. Q.E.D.

3. Sufficient Conditions for Condition(*)

In this section, using the idea of the successive approximations we shall give the sufficient conditions for which Condition(*) holds.

Condition(**). There exist a measure γ on S and $0 < \beta < 1$ such that

(a) $Q^S(D|x,a) \geqq \tau(x,a)\gamma(D)$ for any $D \in \mathrm{F}_S$

and

(b) $\gamma(S) > (1 - \beta)/\tau(x,a)$ for any $(x,a) \in SA$.

In non-discounted Markov Decision Processes($\tau \equiv 1$), Condition(**) insures the ergodicity of the process and corresponds to the conditions for the existence of an optimal stationary policy (see [11], [15] and [17]). In the next section we shall give the replacement problem which satisfies Condition(**).

We define the map U on $\mathrm{B}_a(S)$ by

(3.1) $$Uu(x) = \inf_{a \in A} U(x,a,u),$$

where

(3.2) $$U(x,a,u) = c(x,a) + \int u(x')Q^S(dx'|x,a) - \tau(x,a)\int u(x')\gamma(dx')$$

for each $u \in \mathrm{B}_a(S)$, $x \in S$ and $a \in A$.

Lemma 3.1. Under Condition(**), we have

(a) $u \in \mathrm{B}_a(S)$ implies $Uu \in \mathrm{B}_a(S)$

(b) U is monotone, i.e., if $u \leqq u'$, $Uu \leqq Uu'$

(c) $u \geqq -H/(1-\beta)$ implies that $Uu \geqq -H/(1-\beta)$,
where $H = \sup_{x \in S, a \in A} |c(x,a)|$.

Proof: From the results of [2], (a) follows. Since $Q^S(S|x,a) - \tau(x,a)\gamma(S) < \beta$, (b) and (c) hold.

Theorem 3.1. Condition(**) implies Condition(*).

Proof: Let $u_0 = H/(1-\beta)$ and $u_{n+1} = Uu_n$ for $n \geqq 0$. Then, by Lemma 3.1, $u_n \in \mathrm{B}_a(S)$ for all n. We observe that $u_1 = Uu_0 \leqq H + \beta H/(1-\beta) = H/(1-\beta) = u_0$. Thus, by Lemma 3.1, $u_{n+1} \leqq u_n$ and $u_n \geqq -H/(1-\beta)$ for all n. Let $v = \lim_{n\to\infty} u_n$. Then, $v \in \mathrm{B}_a(S)$. By nonincreasing convergence, $v \leqq Uv$. On the other hand, for any $\varepsilon > 0$, we have $v(x) \geqq Uu_n(x) - \varepsilon \geqq U(x,a^*,u_n) - 2\varepsilon \geqq Uv(x) - 2\varepsilon$ for some n and $a^* \in A$. As $\varepsilon \to 0$, $v \geqq Uv$. Therefore, we have

(3.3) $$v(x) = Uv(x) = \inf_{a \in A} U(x,a,v).$$

If we put $g^* = \int v(x')\gamma(dx')$, (3.3) implies (2.1). Q.E.D.

4. Generalized Age Replacement

Consider a failure model for a system being subject to a sequence of randomly occurring shocks. Each shock causes a random amount of damage and these damages accumulate additively. The damage process is a semi-Markov process. Given that the process is in a state x, it remains there for a random length of time, called the sojourn time, which is denoted by $\alpha(x)$. And the magnitude of shock associated with a state x will be denoted by $\beta(x)$. Let $F_x(\cdot)$ be the distribution function of $\alpha(x)$ and $G_x(\cdot|z)$ the conditional distribution of $\beta(x)$ given $\alpha(x) = z$. Assume that $F_x(z)$ and $G_x(y|z)$ are measurable in x, y and z and $F_x(\cdot)$ has a continuous density with respect to a Lebesgue measure.

A failure of the system can only happen at the time of a shock. If at time t the accumulated damage is x and a shock magnitude y occurs, then the system fails with known probability $1 - r(x + y)$. The function $r(\cdot)$ is called the survival function. The informations available to a decision maker are the time when a shock has occurred, the shock magnitude and whether or not the system has failed at the time of a shock. Under a replacement policy the system is always replaced at failure or at the end of a planned time, whichever occurs first, with respectively costs K and C ($K > C > 0$).

Here, we note that if $r(x) = 0$ for all $x > 0$ and $G_x(0|z) = 0$ for all x and z, the replacement model given above is reduced to the age replacement problem (for example, see [1] and [5]). Thus, our model is called a generalized age replacement. Our objective is to find the replacement policy which minimizes the long-run average cost.

If we define the state of the system by the magnitude of the accumulated damage at the time of a shock, the generalized age replacement is equivalently transformed in a semi-MDP treated in Sections 2 and 3.

Let $R^+ = (0,\infty)$, $\bar{R}^+ = [0,\infty)$ and $\bar{A} = R^+ \cup \{\infty\}$. Define $S = \bar{R}^+$, $A = \bar{A} \cup \{(0,a) | a \in \bar{A}\}$ and $Z = R^+$. A stage is the period starting just after a shock and ending just after the next shock. The length of each stage is represented as the element of Z and $a \in R^+$ and ∞ correspond to the action of the planned replacement time a and non-planned replacement respectively, and $(0,a)$ with $a \in \bar{A}$ means that we replace the system instantly and take an action a. Further define $Q^S(D|x,a)$ $= I_0(D) \{\bar{F}_x(a) + \int_0^a F_x(dz)\int_0^\infty (1-r(x+y))G_x(dy|z)\} + \int_0^a F_x(dz)\int_{\{x+y\in D\}} r(x\ y)\cdot$ $G_x(dy|z)$ and $Q^S(D|x,(0,a)) = Q^S(D|0,a)$ for all $x \in S$, $a \in \bar{A}$ and $D \in \mathsf{F}_S$, where $I_0(D) = 1$ (0) if $0 \in D$ (otherwise) and $\bar{F}_x(z) = 1 - F_x(z)$ for all $z \in R^+$.

Finally define $c(x,a,z,x') = K$ if $z \leq a$ and $x' = 0$; $= C$ if $z > a$ and $x' = 0$; $= 0$ otherwise, and $\tau(x,a,z,x') = z$.

To analyse the generalized age replacement by applying the results of Sections 2 and 3, we introduce the following assumptions A1-A5.

Let $h_x(z)$ be the hazard rate associated with the distribution function $F_x(\cdot)$ at the point z.

A1. $h_{x'}(z) \geq h_x(z)$ if $x' \geq x$ and $h_x(z') \geq h_x(z)$ if $z' \geq z$.

A2. $F_x(\cdot)$ is stochastically decreasing, i.e., $F_x(z) \leq F_{x'}(z)$ for $x \leq x'$.

A3. For any non-negative real valued function $u \in B_a(R^+)$ which is nonincreasing,

$$\int u(y)r(x'+y)G_{x'}(dy|z) \leq \int u(y)r(x+y)G_x(dy|z) \text{ if } x' \geq x \text{ and}$$

$$\int u(y)r(x+y)G_x(dy|z') \leq \int u(y)r(x+y)G_x(dy|z) \text{ if } z' \geq z.$$

A4. $r^* = \int F_0(dz)\int r(y)G_0(dy|z) < 1$, i.e., the probability that the system in state 0 does not fail at the next shock is less than 1.

A5. There exist two positive numbers m and M such that $m < \int zF_x(dz) < M$ for all $x \in \bar{R}^+$.

Notice that if $r(x)$ is nonincreasing and $G_x(\cdot|z)$ is stochastically increasing in $(x,z) \in \bar{R}^+R^+$, i.e., $G_{x'}(y|z) \leq G_x(y|z)$ if $x' \geq x$ and $G_x(y|z') \leq G_x(y|z)$ for $z' \geq z$, A3 holds (see, Lehmann[8] p.73). Also, notice that A1 implies A2.

If we take an action $a \in \bar{A}$ in a state x, the mean sojourn time is

$$\tau(x,a) = \int_0^a zF_x(dz) + a\bar{F}_x(a). \tag{4.1}$$

By (4.1), we observe that $\tau(x,a) \to 0$ as $a \in \bar{A} \to 0$, so that Condition(**) in Section 3 does not hold. For any fixed $\varepsilon > 0$ let $A_\varepsilon = \bar{A}_\varepsilon \cup \{(0,a)|a \in \bar{A}_\varepsilon\}$, where $\bar{A}_\varepsilon = [\varepsilon,\infty) \cup \{\infty\}$. Putting $\delta = (1-r^*)/M$, define a measure γ on $\bar{R}^+$ by, for each $D \in F_{\bar{R}^+}$, $\gamma(D) = \delta$ (0) if $0 \in D$ (elsewhere). Then, since $Q^S(\{0\}|x,a) \geq 1 - r^*$ under A3 and $0 < \inf_{x \in S,\, a \in A_\varepsilon} \tau(x,a) < \sup_{x \in S, a \in A_\varepsilon} \tau(x,a) < M$ under A2 and A5, it is easily verified that Condition(**) holds for the restricted action space A_ε. The function $U(x,a,u)$ defined by (3.2) is, for each $u \in B_a(\bar{R}^+)$,

$$U(x,a,u) = KF_x(a) + C\bar{F}_x(a) + \int_0^a F_x(dz)\int(u(x+y) - K - u(0))r(x+y)G_x(dy|z)$$
$$+ u(0)(1 - \delta\tau(x,a)),$$

and

$$U(x,(0,a),u) = C + U(0,a,u) \quad \text{for } a \in \bar{A}.$$

Lemma 4.1. Let A2 and A3 be satisfied. Then, for any non-negative valued and nondeacreasing $u \in B_a(\bar{R}^+)$ with $u(x) \leq K + u(0)$ for all $x \in S$, it holds that $U(x,a,u) \leq U(x',a,u)$ if $x \leq x'$.

Proof: From A3, we observe that if $x \leqq x'$, $\int(u(x+y) - K - u(0))r(x+y)\cdot G_x(dy|z) \leqq \int(u(x'+y) - K - u(0))r(x+y)G_x(dy|z) \leqq \int(u(x'+y) - K - u(0))r(x'+y)\cdot G_{x'}(dy|z)$. Also, it holds from A2 that $\tau(x,a) \geqq \tau(x',a)$ if $x \leqq x'$. Q.E.D.

Proposition 4.1. Let A2-A5 be satisfied. Then there exist a nondecreasing function $v \in B_a(\bar{R}^+)$ and a constant g^* satisfying that

$$(4.2) \qquad v(x) = \inf_{a\in \bar{R}^+ \cup \{\infty\}} \{ KF_x(a) + C\bar{F}_x(a) + \int_0^a F_x(dz)\int(v(x+y) - K - v(0))r(x+y)G_x(dy|z) + v(0) - g^*\tau(x,a) \},$$

which is corresponding to (2.1).

Proof: For the restricted action space A_ε, define the operator U^ε by

$$(4.3) \qquad U^\varepsilon u(x) = \min\{ C + \inf_{a\in \bar{A}_\varepsilon} U(0,a,u), \inf_{a\in \bar{A}_\varepsilon} U(x,a,u) \}$$

$$= \min\{ C + U^\varepsilon u(0), \inf_{a\in \bar{A}_\varepsilon} U(x,a,u) \} \quad \text{for } u\in B_a(\bar{R}^+).$$

Then, from Theorem 3.1 there exists a non-negative real valued function $v_\varepsilon \in B_a(\bar{R}^+)$ such that

$$(4.4) \qquad v_\varepsilon(x) = U^\varepsilon v_\varepsilon(x) \quad \text{for } x\in \bar{R}^+.$$

First we shall show that $v_\varepsilon(x)$ is nondecreasing in x. Let $u_0 = 0$ and $u_n = U^\varepsilon u_{n-1}$ for $n \geqq 1$. By induction on n it holds from Lemma 4.1 that $u_n(x)$ is nondecreasing for each n.

Since $u_n \to v_\varepsilon$ as $n \to \infty$, v_ε is nondecreasing. Obviously, $v_\varepsilon(x) \geqq v_{\varepsilon'}(x)$ if $\varepsilon \geqq \varepsilon'$. Now, let $v(x) = \lim_{\varepsilon\to 0} v_\varepsilon(x)$ for each $x\in \bar{R}^+$. Then, $v(x)$ is nondecreasing in x. From the monotonicity of $U(x,a,\cdot)$, we have

$$(4.5) \qquad \lim_{\varepsilon\to 0}\inf_{a\in \bar{A}_\varepsilon} U(x,a,v_\varepsilon) \geqq \inf_{a\in \bar{A}} U(x,a,v).$$

On the other hand, for any $\eta > 0$, there exists $a' \in \bar{A}$ with

$$\inf_{a\in \bar{A}} U(x,a,v) > U(x,a',v) - \eta.$$

Also, we have, by the monotone convergence theorem, $\lim_{\varepsilon\to 0} U(x,a',v_\varepsilon) = U(x,a',v)$, so that, since there exists $\varepsilon' > 0$ such that $a' \in \bar{A}_\varepsilon$ for any $\varepsilon < \varepsilon'$, we get

$$\lim_{\varepsilon\to 0}\inf_{a\in \bar{A}_\varepsilon} U(x,a,v_\varepsilon) \leqq U(x,a',v).$$

Therefore, we have

$$\inf_{a\in \bar{A}} U(x,a,v) \geqq \lim_{\varepsilon\to 0}\inf_{a\in \bar{A}_\varepsilon} U(x,a,v_\varepsilon) - \eta.$$

As $\eta \to 0$, it holds that

$$\inf_{a\in \bar{A}} U(x,a,v) \geqq \lim_{\varepsilon\to 0}\inf_{a\in \bar{A}_\varepsilon} U(x,a,v_\varepsilon). \tag{4.6}$$

From (4.5) and (4.6), we have

$$\lim_{\varepsilon\to 0}\inf_{a\in \bar{A}_\varepsilon} U(x,a,v_\varepsilon) = \inf_{a\in \bar{A}} U(x,a,v). \tag{4.7}$$

As $\varepsilon \to 0$ in (4.4), we obtain from (4.3) and (4.7) that

$$v(x) = \min\{ C + v(0),\ \inf_{a\in \bar{A}} U(x,a,v) \} = \inf_{a\in \bar{R}^+\cup\{\infty\}} U(x,a,v). \tag{4.8}$$

In (4.8), putting $g^* = v(0)\delta$ we obtain (4.2). Q.E.D.

For $v\in B_a(\bar{R}^+)$ and g^* as in Proposition 4.1, consider the following equation.

$$J(x,a) = g^*,\quad a\in[0,\infty), \tag{4.9}$$

where $J(x,a) = \{ K - C + \int(v(x+y) - K - v(0))r(x+y)G_x(dy|a) \}h_x(a)$

Proposition 4.2. Suppose that A1 and A3-A5 hold. Then

(a) there is an extended real valued function $a^* = a^*(x)$ such that for each $x\in\bar{R}^+$, $a^*(x)$ is a solution of (4.9) and $a^*(x)$ is nonincreasing in $x\in\bar{R}^+$, where if no solution exists on $[0,\infty)$, let $a^*(x) = \infty$.

(b) the optimal replacement level for each $x\in\bar{R}^+$ is represented by $a^*(x)$, that is, under the optimal replacement policy we replace the system if the sojourn time in state x is greater than $a^*(x)$.

Proof: By taking the derivative of $U(x,a,v)$ with respect to a, setting it equal to zero and simplifying, we obtain the equation (4.9). From Proposition 4.1, $v(x)$ is nondecreasing in x and $v(x) \leqq v(0) + K$ for all $x\in\bar{R}^+$. Thus, from A1 and A3, $J(x',a) \geqq J(x,a)$ if $x' \geqq x$ and $J(x,a') \geqq J(x,a)$ if $a' \geqq a$, so that (a) follows. Also, by Theorem 2.1 and Remark 2.1, (b) holds clearly.

Q.E.D.

Remark 4.1. Putting $I = \{x\in\bar{R}^+ \mid 0 < a^*(x) < \infty\}$, let $a_1^* = \inf I$ and $a_2^* = \sup I$. Then, from Proposition 4.2 we observe that under an optimal replacement policy we don't replace if the accumulated damage x is smaller than a_1^* and replace instantly if x is greater than a_2^*. The optimal replacement policy given in Proposition 4.2 is contained in the class of state-age replacement rules of Kao[6].

To obtain the further results, we shall introduce the lexicographic order $\succ$ in R^2. We say that $(x_1,y_1)\in R^2$ is lexicographical positive, written $(x_1,y_1) \succ 0$ if $(x_1,y_1) \neq 0$ and the first non-vanishing coordinate is positive.

If $(x_1-x_2,y_1-y_2)\succ 0$, we write $(x_1,y_1)\succ(x_2,y_2)$.

A1' $h_x(z)$ is nondecreasing in $(x,z)\in \bar{R}^+R^+$ w.r.t. the lexicographic order $\succ$.

A3' For any non-negative real valued function $u\in B_a(R^+)$ which is nonincreasing, $\int u(y)r(x+y)G_x(dy|z)$ is nonincreaing in $(x,z)\in \bar{R}^+R^+$ w.r.t. the lexicographic order $\succ$.

Note that A1' and A3' imply A1 and A3 respectively. So, we can obtain the following:

Proposition 4.3. Suppose that A1', A3', A4 and A5 hold. Then, there exists $x^*\in \bar{R}^+$ and $a^*\in \bar{A}$ for which the optimal replacement level, $a^*(x)$, is defined by

$$(4.10)\qquad a^*(x) = \begin{cases} \infty & \text{if } x < x^* \\ a^* & \text{if } x = x^* \\ 0 & \text{if } x > x^* . \end{cases}$$

Proof: From A1' and A3', $J(x,a)$ is nondecreasing in $(x,a)\in \bar{R}^+R^+$ w.r.t. the lexicographic order $\succ$, so that there exists $x^*\in \bar{R}^+$ and $a^*\in \bar{A}$ such that

$$J(x,a)\begin{Bmatrix} \leqq \\ = \\ > \end{Bmatrix} g^* , \quad \text{if } \begin{cases} x \leqq x^* \\ x = x^* \text{ and } a = a^* \\ x > x^* . \end{cases}$$

This means that $a^*(x)$ defined by (4.10) is the optimal replacement level at state x. Q.E.D.

Remark 4.2. The optimal replacement level, $a^*(x)$, given in Proposition 4.3 has the same properties as that in Theorem 2 in Zuckerman[18], who has investigated the structure of an optimal replacement policy using the infinitesimal operator of the Markov processes. Our conditions assumed in Proposition 4.3 are slightly stronger than those in Zuckerman[18], but our proofs are more constructive.

Remark 4.3. In the case which the system always fails when a shock occurs, i.e., $r(x) = 0$ and $G_x(0|z) = 0$ for all $x,z > 0$, the optimal equation (4.5) is as follows:

$$(4.11)\qquad v(0) = \inf_{a\in R^+\cup\{\infty\}} \{ KF_0(a) + C\bar{F}_0(a) + v(0)(1 - \delta\tau(0,a) \}.$$

By putting $g^* = \delta v(0)$, (4.11) becomes

$$(4.12)\qquad g^* = \inf_{a\in R^+\cup\{\infty\}} \{ C + (K - C)F_0(a) \}/\tau(0,a).$$

Also, the equation (4.9) is $(K-C)h_0(a) = g^*$, whose solution is represented by a^*. Then, since $\tau(0,a) = \int_0^a \bar{F}_0(y)dy$, from (4.12) a^* satisfies that

(4.13) $h_0(a^*)\int_0^{a^*} \bar{F}_0(y)dy - F_0(a^*) = C/(K-C),$

which agrees with the well-known results(see, e.g., Barlow and Proschan[1], p.87).

Acknowledgement

The author wishes to express his thanks to referees for their very helpful comments which let to an improvement of the presentation of the material.

References

[1] Barlow, R. E. and Proschan, F. : *Mathematical Theory of Reliability*. Wiley, New york, 1965.

[2] Blackwell, D., Freedman, D. A. and Orkin, M.: The optimal reward operator in dynamic programming. *Ann. Probability*, Vol.2(1974), 926-941.

[3] Federgruen, A. and Tijms, H. C.: The optimality equation in average cost denumerable state semi-Markov decision problems, recurrency conditions and algorithms. *Journal of Applied Probability*, Vol.15(1978), 356-373.

[4] Feldman, R. M.: Optimal replacement with semi-Markov shock models. *Journal of Applied Probability*, Vol.13(1976), 108-117.

[5] Fox, B. L.: Age replacement with discounting. *Operetions Research*, Vol. 14(1966), 533-537.

[6] Kao, E. P. C.: Optimal replacement rules when changes of state are semi-Markovian. *Operations Research*, Vol.21 (1973), 1231-1249.

[7] Kolonko, M.: The average-optimal adaptive control of a Markov renewal model in presence of an unknown parameter. *Optimization*, Vol.13 (1982), 567-591.

[8] Lehmann, E. L.: *Testing statistical hypotheses*. Wiley, New York, 1959.

[9] Lippman, S. A.: Maximal average-reward policies for semi-Markov decision processes with arbitrary state and action space. *Ann. Math. Statist.*, Vol.42 (1971), 1717-1726.

[10] Loéve, M.: *Probability Theory, Second edition*, D. Van Nostrand. Co. Inc. New Jersey, 1960.

[11] Ross, S. M.: Arbitrary state Markovian decision processes. *Ann. Math. Statist.*, Vol.39 (1968), 2118-2122.

[12] Ross, S. M.: *Applied probability models with optimization applications*. Holden-Day, San Francisco, 1970.

[13] Taylor, H. M.: Optimal replacement under additive damage and other failure models. *Naval Res. Logist. Quart.*, Vol.22 (1975), 1-8.

[14] Shreve, S. E. and Bertsekas, D. P.: Alternative theoretical frameworks for finite horizon discrete-time stochastic optimal control. *Siam J. Control and Optimization*, Vol.16 (1978), 953-978.

[15] Tijms, H. C.: On dynamic programming with arbitrary state space, compact action space and the average return criterion. *Report BW 55/75 Mathematical Centre*, Amsterdam. 1975.

[16] White, D. J.: dynamic programming, Markov chains and the method of successive approximations. *Journal of Mathematical Analysis and Applications*, Vol.6 (1963), 373-376.

[17] Yamada, K.: Duality theorem in Markovian decision processes. *Journal of Mathematical Analysis and Applications*, Vol.50 (1975), 579-595.

[18] Zuckerman, D.: Optimal stopping in a semi-Markov shock model. *Journal of Applied Probability*, Vol.15 (1978), 629-634.

Masami KURANO: Department of Mathematics,
Faculty of Education, Chiba University,
Yayoi-cho, Chiba, 260, Japan.