Journal of the Operations Research Society of Japan
Vol. 26, No. 2, June 1983

OPTIMAL SERVICE-RATE CONTROL OF EXPONENTIAL QUEUEING SYSTEMS

Kyung Y. Jo
George Mason University

(Received May 21, 1982; Final February 8, 1983)

Abstract We analyze M/M/1 queueing systems with variable service rates under weak cost and probabilistic assumptions by using semi-Markov decision-process formulations. The methods used in studying the characteristics of optimal policies are induction arguments involving supermodular (or submodular) properties of the objective function. First, we establish monotonicity properties of the optimal policies and sufficient conditions for extremal-policy optimality for the finite-horizon problem. The analyses are then extended to show that the monotone optimal policies carry over to the discounted infinite-horizon and long-run average-return problems.

1. Introduction

The model in consideration is a generalized version of the models considered by Crabill [3], Sabeti [14], Lippman [11], and Bengtsson [2]. Further detailed discussion on the service-rate control models can be seen in [5], [19], and [21]. In contrast to the previous literature, we use weaker cost assumptions and obtain more extensive characteristics of the (exponential) service-rate control models. Another feature of this paper is to have a unified treatise on the subject, since the previous literature reveals diverse, and often repetitive, arguments on the characteristics of the control model.

The control problem is modeled as a semi-Markov decision process with an infinite denumerable state space S and a closed action space A. The state variable denotes the number of customers in the system. We use Lippman's device (see [11]), so that the time between observations is exponentially distributed with a fixed parameter. The objective is to maximize the expected total return of operating the system, both with and without discounting, starting from each given state. The decision for a given state is to select a service rate for the system. This service rate will be in effect until the next observation point (decision epoch).

In section 2 we formulate a semi-Markov decision model of the $M/M/1$ queueing system with variable service rates. We can eliminate some nonoptimal actions before actual maximization steps, given the structure of the objective function of dynamic programming.

For the finite-horizon problem, we prove, in Section 3, some monotonicity properties of optimal policies. It is shown that the optimal service rate is nondecreasing in the number of customers in the system, nonincreasing in the discount rate, and nondecreasing in the number of remaining periods. The method used in proofs is induction on the horizon length. At each step we establish the supermodularity or submodularity of the objective function (cf. Topkis [24]) of the dynamic-programming formulation.

In sections 4 and 5 the infinite-horizon problems, with and without discounting, are considered, respectively. The monotonicity results for the finite-horizon problem are shown to carry over to the infinite-horizon problem with discounting. It is also shown that the n-horizon optimal-value functions as well as optimal policies converge to those of the infinite-horizon problem. We next derive sufficient conditions for extremal policies to be optimal for the infinite-horizon problems by using shift transformations. In section 5 we study the long-run average-return problem. To prove the existence of strongly optimal policies with the property of monotonicity in the number of customers, we use an extremal-policy shift transformation in the proof.

2. Formulation as Semi-Markov Decision Process

We construct a semi-Markov decision model using the "new device" of Lippman [11]. The arrival rate λ is fixed. The decision variable μ is the exponential service rate to be selected from the closed interval $[0, \bar{\mu}]$, $\bar{\mu} > \lambda$, so as to maximize the total expected return starting from each given state. We observe the system at points of arrivals, which occur at rate λ; service completions, which occur at rate μ; and null events which occur at rate $\bar{\mu} - \mu$. The time between observation points is therefore exponentially distributed with parameter $\Lambda = \lambda + \bar{\mu}$, not depending on state and action. The system is said to be in state $i \in S$ whenever there are i customers in the system. The action to be taken at an observation point is the selection of a service rate μ for the system so as to maximize the total return of operating the system over the remaining periods. The service rate μ is in effect until the next observation point. Thus, the action space for each state i is $A = [0, \bar{\mu}]$.

We need the following assumptions for the cost structure of our model.

Assumption 2.1. The service cost rate, $C(\mu)$, is a left-continuous nondecreasing function of μ, with $C(0) = 0$.

Assumption 2.2. The holding cost rate, $h(i)$, is a convex nondecreasing function of i, with $h(1) > h(0) \geq 0$.

The reason for assuming $C(\mu)$ to be a left-continuous function is to assure an α-optimal n-period policy for the closed action space. The left continuity of $C(\mu)$ implies that the objective function is upper semicontinuous (cf. [18]).

The continuously discounted one-period return $r(\cdot, \mu)$ is given by

$$r(i, \mu) = (\alpha+\Lambda)^{-1}\{-C(\mu) - h(i)\}$$

and the transition probability from state i to state j, for given μ, is

$$\begin{aligned} P_{ij}(\mu) &= Pr\{X_{t+1} = j \mid X_t = i,\ A_t = \mu\} \\ &= [\lambda \underset{\sim}{1}(j = i+1) + \mu \underset{\sim}{1}(j = i-1) + (\bar{\mu}-\mu) \underset{\sim}{1}(j = i)]/\Lambda \end{aligned}$$

where $\underset{\sim}{1}(E)$ is the indicator function of the event E.

As a function of state, define $V_{n,\alpha}$ as the maximal discounted total return of operating the system over n remaining periods, $n \geq 1$ ($V_{0,\alpha} \equiv 0$). For our cost and probability structure, $V_{n,\alpha}$ is well defined for all states $i \in S$ and $n \geq 0$ (cf. [9], [15], and [17]). We have the following dynamic programming recursive equation for $n \geq 1$:

$$\begin{aligned} V_{n,\alpha}(i) &= \max_{0\leq\mu\leq\bar{\mu}}[-C(\mu) - h(i) + \lambda V_{n-1,\alpha}(i+1) \\ &\quad + \mu V_{n-1,\alpha}(i-1) + (\bar{\mu}-\mu) V_{n-1,\alpha}(i)]/(\alpha+\Lambda),\quad i \geq 1, \end{aligned} \tag{2.1}$$

where $V_{n,\alpha}(0) = [-h(0) + \lambda V_{n-1,\alpha}(1) + \bar{\mu} V_{n-1,\alpha}(0)]/(\alpha+\Lambda)$.

Disregarding the constant terms in the right hand of (2.1), we are only concerned with the following function $g(\cdot,\mu)$ of μ to be maximized

$$g_{n,\alpha}(i,\mu) = -C(\mu) + \mu\{V_{n-1,\alpha}(i-1) - V_{n-1,\alpha}(i)\}. \tag{2.2}$$

In exploiting the special structure of (2.2), we can use the criterion for exclusion of nonoptimal actions of Crabill [4]. In this context, we state the next theorem and corollary.

Theorem 2.1. A sufficient condition for nonoptimality of decision μ for any state is that there exist two other decisions μ' and μ'' such that

$$\mu = \gamma\mu' + (1-\gamma)\mu'' \text{ and } C(\mu) > \gamma C(\mu') + (1-\gamma)\, C(\mu'')$$

where $\gamma \in (0,1)$.

Proof: Since our problem has a form $-C(\mu) + D\mu$, where D is a constant, we directly apply Theorem 1 of Crabill [4].

Corollary 2.1. If $C(\bar{\mu})/\bar{\mu} < C(\mu)/\mu$ for all $\mu \in (0,\bar{\mu})$, the optimal policy is extremal: the optimal action in each state is either $\mu = \bar{\mu}$ (full-service) or $\mu = 0$ (passive).

Proof: The condition $C(\bar{\mu})/\bar{\mu} < C(\mu)/\mu$ excludes all μ except 0 and $\bar{\mu}$ with $\gamma = \mu/\bar{\mu}$ by Theorem 2.1.

Corollary 2.1 states the extremal-policy optimality condition for our $M/M/1$ control model. The choice between the two extremal actions (full service or no service) is decided by the value of $V_{n,\alpha}(i-1) - V_{n,\alpha}(i)$. The full-service optimality condition for other models was derived by Sobel [20] and Schleef [16].

3. Characteristics of Optimal Service Policies for Finite Horizon

In this section we study the characteristics of optimal policies for the finite-horizon problem. Of all characteristics, the monotonicity of optimal policies is perhaps most appealing. We prove that the optimal service rate is nondecreasing in the number of customers, nondecreasing in the horizon length, and nonincreasing in the discount rate. The advantage of using the monotonicity properties of optimal policies is the reduction of the action space to be searched, so that the savings in maximization effort can be considerable.

The most general approach in the literature for establishing monotonicity involves showing that the objective function is a supermodular (submodular) function of state and action (cf. [24]), which in turn implies that the optimal action is a monotonically increasing (decreasing) function of the state. A function $g(s,a)$ is said to be *supermodular* (*submodular*) if, for all $a_1 < a_2$, $g(s,a_2) - g(s,a_1)$ is monotonically increasing (decreasing) in s. For each state $s \in S$, let $a^*(s)$ be an action that maximizes $g(s,\cdot)$ over all $a \in A$. Ties are resolved, in all cases, by choosing the smallest maximizer. The existence of such an $a^*(s)$ is ensured if $g(s,\cdot)$ is an upper semicontinuous function (cf. [18]).

Lemma 3.1. If $g(s,a)$ is supermodular (submodular), then $a^*(s)$ is a monotonically increasing (decreasing) in s.

Proof: Let $s' > s$. By definition, we have $g(s,a^*(s)) > g(s,a)$ for all $a < a^*(s)$. Hence, by the supermodularity of $g(s,a)$, $g(s',a^*(s)) - g(s',a) > g(s,a^*(s)) - g(s,a) > 0$, for all $a < a^*(s)$. Therefore, $a^*(s') \geq a^*(s)$. We use a similar method to prove the monotonicity of the maximizer of a submodular function.

Using the structure of $g(s,a)$ for all our control models, we can infer sufficient (and necessary) conditions for $g(s,a)$ to be supermodular or submodular, so that we can establish the monotonicity properties of the optimal policies. In this connection, we derive a necessary and sufficient condition for $g_{n,\alpha}(i,\mu)$ of (2.2) to be supermodular in i and μ in the next lemma.

Lemma 3.2. A necessary and sufficient condition that $g_{n,\alpha}(i,\mu)$ is supermodular in i and μ, for a fixed n, α, is that $V_{n-1,\alpha}(i-1) - V_{n-1,\alpha}(i)$ is a nondecreasing function of i ($V_{n,\alpha}(i)$ is concave in i).

Proof: We suppress the subscript α in the expressions of $g_{n,\alpha}$ and $V_{n,\alpha}$. We must show for $\mu_2 > \mu_1$ that

$$g_n(i,\mu_2) - g_n(i,\mu_1) \le g_n(i+1,\mu_2) - g_n(i+1,\mu_1). \tag{3.1}$$

Direct subsititution from (2.2) yields

$$\begin{aligned} &g_n(i+1,\mu_2) - g_n(i+1,\mu_1) - g_n(i,\mu_2) + g_n(i,\mu_1) \\ &\quad = (\mu_2-\mu_1)[\{V_{n-1}(i) - V_{n-1}(i+1)\} - \{V_{n-1}(i-1) - V_{n-1}(i)\}]. \end{aligned}$$

Since $\mu_2-\mu_1 > 0$, we have the desired result.

Let $\mu_{n,\alpha}{}^*(i)$ be the optimal service rate when there are i customers in the system, n periods remain, and the discount rate is α. Then the optimal service rate maximizes $g_{n,\alpha}(i,\mu)$ among all actions. The ties are resolved by choosing the smallest maximizer. The next theorem establishes the monotonicity of optimal policies in the number of customers in the system.

Theorem 3.1. For $\alpha \ge 0$, $n \ge 0$, and $i \in S$, $V_{n,\alpha}(i-1) - V_{n,\alpha}(i)$ is a nonnegative nondecreasing function of i ($V_{n,\alpha}$ is concave), so that $\mu_{n,\alpha}{}^*(i)$ is a nondecreasing function of i.

Proof: We drop the subscript α of V. We need to prove

$$V_n(i-1) - V_n(i) \le V_n(i) - V_n(i+1). \tag{3.2}$$

We use an induction on n to prove this. For $n = 1$, (3.2) is trivial since $V_1(i) = -h(i)/(\alpha+\Lambda)$. Assume (3.2) holds for $n - 1$ for each $i \in S$. Let μ' be the optimal action in state $i+1$ and μ'', in $i-1$, with n periods remaining. Then we have

$$\begin{aligned} &\{V_n(i) - V_n(i+1)\}(\alpha+\Lambda) \\ &\ge h(i+1) - h(i) + \lambda\{V_{n-1}(i+1) - V_{n-1}(i+2)\} \\ &+ (\bar{\mu}-\mu')\{V_{n-1}(i) - V_{n-1}(i+1)\} + \mu'\{V_{n-1}(i-1) - V_{n-1}(i)\} \ge 0. \end{aligned}$$

and

$$\{v_n(i-1) - v_n(i)\}(\alpha+\Lambda)$$
$$\le h(i) - h(i-1) + \lambda\{v_{n-1}(i-1) - v_{n-1}(i)\}$$
$$+ (\bar{\mu}-\mu'')\{v_{n-1}(i-1) - v_{n-1}(i)\} + \mu''\{v_{n-1}(i-2) - v_{n-1}(i-1)\}.$$

Combining the above inequalities gives

$$[\{v_n(i) - v_n(i+1)\} - \{v_n(i-1) - v_n(i)\}](\alpha+\Lambda)$$
$$\ge \{h(i+1) - h(i)\} - \{h(i) - h(i-1)\}$$
$$+ \lambda[\{v_{n-1}(i+1) - v_{n-1}(i+2)\} - \{v_{n-1}(i) - v_{n-1}(i+1)\}]$$
$$+ (\bar{\mu}-\mu')[\{v_{n-1}(i) - v_{n-1}(i+1)\} - \{v_{n-1}(i-1) - v_{n-1}(i)\}]$$
$$+ \mu''[\{v_{n-1}(i-1) - v_{n-1}(i)\} - \{v_{n-1}(i-2) - v_{n-1}(i-1)\}]$$
$$\ge 0,$$

where the last inequality results from the convexity of $h(i)$ and the induction hypothesis. By Lemma 3.2 $g_{n,\alpha}(i,\mu)$ is supermodular in i and μ. The existence of an action maximizing $g_{n,\alpha}(i,\mu)$ is justified since it is upper semicontinuous and A is closed. By Lemma 3.1, $\mu_{n,\alpha}^*(i)$ is a nondecreasing function of i. This completes the proof.

Theorem 3.1 establishes the monotonicity of the optimal policy in the number of customers in the system. Such an optimal policy is also called a switch-over or connected policy (cf. [3] and [11]). We are now interested in seeing how the optimal service rate will be affected if the number of periods remaining increases. Since we are assuming zero terminal rewards, we can verify monotonicity of optimal policies in the number of periods remaining.

Theorem 3.2 For each $i \in S$ and $\alpha \ge 0$, $v_{n,\alpha}(i-1) - v_{n,\alpha}(i)$ is a nondecreasing function of n, so that $\mu_{n,\alpha}^*$ is a nondecreasing function of n.

Proof: By analogy with Lemma 3.2 we need to prove $g_{n,\alpha}(i)$ is supermodular in n and μ. So it suffices to show for all $i \in S$

(3.3) $$v_{n,\alpha}(i-1) - v_{n,\alpha}(i) \ge v_{n-1,\alpha}(i-1) - v_{n-1,\alpha}(i).$$

We drop the subscript α on v. For n=1, this is trivially true. Assume (3.3) is true for $n-1$. Let μ' be an optimal policy associated with $v_{n-1,\alpha}(i-1)$ and μ'', with $v_{n,\alpha}(i)$. Then we have

$$\{v_n(i-1) - v_n(i)\}(\alpha+\Lambda)$$
$$\ge h(i) - h(i-1) - c(\mu') + c(\mu'')$$
$$+ \lambda\{v_{n-1}(i) - v_{n-1}(i+1)\} + (\mu-\bar{\mu}'')\{v_{n-1}(i-1) - v_{n-1}(i)\}$$

$$+ \mu'\{v_{n-1}(i-2) - v_{n-1}(i-1)\} \geq 0$$

and

$$\{v_{n-1}(i-1) - v_{n-1}(i)\}(\alpha+\Lambda)$$
$$\geq h(i) - h(i-1) - c(\mu') + c(\mu'')$$
$$+ \lambda\{v_{n-2}(i) - v_{n-2}(i+1)\} + (\bar{\mu}-\mu'')\ v_{n-2}(i-1) - v_{n-2}(i)\}$$
$$+ \mu'\{v_{n-2}(i-2) - v_{n-2}(i-1)\}.$$

Combining these two inequalities gives

$$[\{v_n(i-1) - v_n(i)\} - \{v_{n-1}(i-1) - v_{n-1}(i)\}](\alpha+\Lambda)$$
$$\geq \lambda[\{v_{n-1}(i) - v_{n-1}(i+1)\} - \{v_{n-2}(i) - v_{n-2}(i+1)\}]$$
$$+ (\bar{\mu}-\mu'')[\{v_{n-1}(i-1) - v_{n-1}(i)\} - \{v_{n-2}(i-1) - v_{n-2}(i)\}]$$
$$+ \mu'[\{v_{n-1}(i-2) - v_{n-1}(i-1)\} - \{v_{n-2}(i-2) - v_{n-2}(i-1)\}]$$
$$\geq 0,$$

where the last inequality results from the induction hypothesis. Therefore, $g_{n,\alpha}(i,\mu)$ is supermodular in n and μ and, accordingly, $\mu_{n,\alpha}*(i)$ is a nondecreasing function of n.

We now see how the optimal action behaves when we are allowed to change the parameter α, the discount rate. An examination of the dynamic-programming optimality equation reveals that the optimal service rate depends only on the function $g_{n,\alpha}(i,\mu)$. So we shall be concerned with seeing how $g_{n,\alpha}(i,\mu)$ behaves. To this end, we present the next lemma.

Lemma 3.3. A necessary and sufficient condition that $g_{n,\alpha}(i,\mu)$ is submodular in α and μ is that $v_{n-1,\alpha}(i-1) - v_{n-1,\alpha}(i)$ is a nonincreasing function of α for each $i \in S$.

Proof: It suffices to show for $\alpha_1 < \alpha_2$ and $\mu_1 < \mu_2$,

(3.4) $$g_{n,\alpha_1}(i,\mu_2) - g_{n,\alpha_1}(i,\mu_1) \geq g_{n,\alpha_2}(i,\mu_2) - g_{n,\alpha_2}(i,\mu_1).$$

By a direct substitution of (2.2), we have

$$v_{n-1,\alpha_1}(i-1) - v_{n-1,\alpha_1}(i) \geq v_{n-1,\alpha_2}(i-1) - v_{n-1,\alpha_2}(i).$$

The converse is also true.

As we vary the discount rate, the corresponding optimal service rate will change monotonically as suggested in the next theorem.

Theorem 3.3. For each $i \in S$ and $n \geq 1$, $v_{n,\alpha}(i-1) - v_{n,\alpha}(i)$ is a nonincreasing function of $\alpha \geq 0$, so that $\mu_{n,\alpha}*(i)$ is a nonincreasing function of α, Proof: By Lemma 3.3, it is (necessary and) sufficient to show, for $\alpha_2 \geq \alpha_1$,

$$(3.5)\qquad v_{n,\alpha_1}(i-1) - v_{n,\alpha_1}(i) \geq v_{n,\alpha_2}(i-1) - v_{n,\alpha_2}(i).$$

We use an induction argument as a proof. For $n = 1$ (3.5) is trivially true, since $v_{1,\alpha}(i) = -h(i)/(\alpha+\Lambda)$. Assume (3.5) holds for $n-1$ for each i. Let μ' be the optimal solution associated with $v_{n,\alpha_2}(i-1)$ and μ'' with $v_{n,\alpha_1}(i)$. Then we have

$$\begin{aligned}&\{v_{n,\alpha_1}(i-1) - v_{n,\alpha_1}(i)\}(\alpha_1+\Lambda)\\ &\geq h(i) - h(i-1) + c(\mu'') - c(\mu')\\ &+ \lambda\{v_{n-1,\alpha_1}(i) - v_{n-1,\alpha_1}(i+1)\} + (\bar{\mu}-\mu'')\{v_{n-1,\alpha_1}(i-1)\\ &- v_{n-1,\alpha_1}(i)\} + \mu'\{v_{n-1,\alpha_1}(i-2) - v_{n-1,\alpha_1}(i-1)\}\end{aligned}$$

and

$$\begin{aligned}&\{v_{n,\alpha_2}(i-1) - v_{n,\alpha_2}(i)\}(\alpha+\Lambda)\\ &\leq h(i) - h(i-1) + c(\mu'') - c(\mu')\\ &+ \lambda\{v_{n-1,\alpha_2}(i) - v_{n-1,\alpha_2}(i+1)\} + (\bar{\mu}-\mu'')\{v_{n-1,\alpha_2}(i-1)\\ &- v_{n-1,\alpha_2}(i)\} + \mu'\{v_{n-1,\alpha_2}(i-2) - v_{n-1,\alpha_2}(i-1)\}.\end{aligned}$$

Combining these two inequalities gives

$$\begin{aligned}&(\alpha_1+\Lambda)\{v_{n,\alpha_1}(i-1) - v_{n,\alpha_1}(i)\} - (\alpha_2+\Lambda)\{v_{n,\alpha_2}(i-1) - v_{n,\alpha_1}(i)\}\\ &\geq \lambda[\{v_{n-1,\alpha_1}(i) - v_{n-1,\alpha_1}(i+1)\} - \{v_{n-1,\alpha_2}(i) - v_{n-1,\alpha_2}(i+1)\}]\\ &+ (\bar{\mu}-\mu'')[\{v_{n-1,\alpha_1}(i-1) - v_{n-1,\alpha_1}(i)\}\\ &- \{v_{n-1,\alpha_2}(i-1) - v_{n-1,\alpha_2}(i)\}]\\ &+ \mu'[\{v_{n-1,\alpha_1}(i-2) - v_{n-1,\alpha_1}(i-1)\}\\ &- \{v_{n-1,\alpha_2}(i-2) - v_{n-1,\alpha_2}(i-1)\}]\\ &\geq 0,\end{aligned}$$

where the last inequality results form the induction hypothesis. Therefore, we have

$$v_{n,\alpha_1}(i-1) - v_{n,\alpha_1}(i) \geq \{(\alpha_2+\Lambda)/(\alpha_1+\Lambda)\}\{v_{n,\alpha_2}(i-1) - v_{n,\alpha_2}(i)\}$$

$$\geq V_{n,\alpha_2}(i-1) - V_{n,\alpha_2}(i),$$

the desired result (3.5). By Lemma 3.1 we conclude that $\mu_{n,\alpha}{}^*(i)$ is a non-increasing function of $\alpha \geq 0$. This completes the proof.

We have established the monotonicity properties of optimal service rate for the finite-horizon problem of $M/M/1$ control models. In the next sections, we study the infinite-horizon problems with and without discounting to see if some monotonicity results still hold.

4. Infinite-Horizon Problem with Discounting

In this section, we consider infinite-horizon problems with discounting and show that the monotonicity properties of optimal policies established for the finite-horizon problem also hold under the same cost and probability structure. We also derive a sufficient condition for full-service policy optimality or passive policy optimality using a shift transformation.

Suppose we have an infinite-horizon control model where there is no limit to the operating time and future costs are discounted continuously at rate $\alpha > 0$. Since the expected length of time between observation points is $1/\Lambda$, we have a discount factor $\beta = e^{-\alpha/\Lambda}$ for one period. Define $V_\alpha(i)$ as the maximal expected discounted total return over an infinite horizon starting in state i. For our problem, the expected discounted positive part of the total return at the n^{th} decision epoch is bounded above by zero. Hence, by standard arguments from the theory of dynamic programming (see, e.g., [9], [15], and [17]), V_α is well defined and satisfies the optimality equation,

$$(4.1)\qquad \begin{aligned} V_\alpha(i) = \max_{0\leq\mu\leq\bar{\mu}} & [-C(\mu) - h(i) + \lambda V_\alpha(i+1) + \mu V_\alpha(i-1) \\ & + (\bar{\mu}-\mu)\, V_\alpha(i)]/(\alpha+\Lambda) \end{aligned}$$

for $i \geq 1$,

where $V_\alpha(0) = [-h(0) + \lambda V_\alpha(1) + \bar{\mu} V_\alpha(0)]/(\alpha+\Lambda)$.

Moreover, the stationary policy $\mu_\alpha{}^*(i)$ that maximizes the right-hand side of (4.1) is optimal among all policies.

The next theorem and corollary show that the optimal-value functions and the optimal policies for the n-period problem monotonically approach those for the infinite-horizon problem as n goes to infinity. Hence the characteristics of optimal policies, especially the monotonicities in α and in i, carry over to the infinite-horizon problem.

Theorem 4.1. For all $i \in S$, $V_\alpha(i) = \lim_{n\to\infty} V_{n,\alpha}(i)$. In addition, $\mu_\alpha^*(i) = \lim_{n\to\infty} \mu_{n,\alpha}^*(i)$, so that the limit of a sequence of n-period optimal policies is optimal for the infinite-horizon problem.

Proof: By Schäl [15] and Serfozo [17], we conclude that the n-period optimal-value function approaches a well-defined limit as n goes to infinity and that the infinite-horizon optimal policy exists and $\mu_\alpha^*(i) = \lim_{n\to\infty} \mu_{n,\alpha}^*(i)$.

Corollary 4.1. For each $i \in S$, $V_\alpha(i-1) - V_\alpha(i)$ is nonnegative nondecreasing in i and nonincreasing in α, so that $\mu_\alpha^*(i)$ is nondecreasing in i and nonincreasing in α.

Proof: Immediate from Theorem 3.1, 3.3, and 3.4.

For some cost structure, the optimal policy turns out to be an extreme policy, either a full-service policy or a passive policy. In this case, we can solve the problem with almost no computational effort. We shall derive the sufficient conditions for the extremal-policy optimality using a shift transformation. Further discussion of shift-transformations is found in [7], [25], and [22].

Let g be a stationary policy which selects the service rate $\mu'(i)$ whenever the system is in state i. Under policy g, we have the corresponding value function V_α^g for the infinite horizon problem with discount rate α. We define a new return function as

$$\tilde{r}(f) = r(f) + P(f)V_\alpha^g - V_\alpha^g.$$

Using $V_\alpha^g = r(g) + P(g)V_\alpha^g$, we have the component-wise expression of $\tilde{r}$ as

$$\tilde{r}(i,\mu) = [C(\mu'(i)) - C(\mu) - (\mu'(i)-\mu)(V_\alpha^g(i-1) - V_\alpha^g(i))]/(\alpha+\Lambda)^{-1},$$

(4.2)

$$\tilde{r}(0,\cdot) = 0.$$

This is the reward function for a new Markov decision process induced by the shift transformation. The shift transformation is one of the methods to convert a problem with unbounded returns to an equivalent problem with bounded returns. The value function V_α^g has the following properties.

Theorem 4.2 When g is a full-service policy or a passive policy, i.e., $\mu(i) = \bar{\mu}$ for all $i \geq 1$ or $\mu(i) = 0$ for all $i \geq 0$, $V_{n,\alpha}^g(i-1) - V_{n,\alpha}^g(i)$ is nondecreasing in i, nondecreasing in n, and nonincreasing in α.

Proof: We follow the similar induction arguments as in Theorems 3.1, 3.2 and 3.3.

Next, we establish sufficient conditions for extremal-policy optimality using shift transformation. Let $g(i)$ be the action of the stationary policy g when the system is in state i.

Theorem 4.3. (i) A sufficient condition that a full-service policy is optimal for all $i \geq 1$ is

$$C(\bar{\mu}) - C(\mu) < (\bar{\mu} - \mu)(V_\alpha^g(i-1) - V_\alpha^g(i)), \tag{4.3}$$

for all $\mu \in [0,\bar{\mu})$, where $g(i) = \bar{\mu}$; (ii) a sufficient condition that a passive policy is optimal for all $i \geq 0$ is

$$C(\mu)/\mu \geq V_\alpha^g(i-1) - V_\alpha^g(i), \tag{4.4}$$

for all $\mu \in (0,\bar{\mu}]$, where $g(i) = 0$.

Proof: The return function $\tilde{r}(i,\mu)$ can be interpreted as the net savings from taking action μ rather than $g(i)$ in the current period, when policy g is taken in all future periods. Hence, a sufficient condition for a stationary policy g to be optimal for all $i \in S$ is that $\tilde{r}(i,\mu)$ is nonpositive for all $\mu \in [0,\bar{\mu})$. With direct application of (4.2), $\tilde{r}(i,\mu) \leq 0$ gives (4.3) and (4.4) for the full-service optimality and the zero-service optimality respectively. Since we are choosing the smallest maximizer to resolve the ties, we have the strict inequality in (4.3). The implication of Theorem 4.3 is different from that of Corollary 2.1 where we exploited the property of the function $C(\mu)$ in that the extremal-policy sufficient conditions for the former apply even when the condition of $C(\mu)$ for the latter is violated.

If we have a prior knowledge about $V_\alpha^g(i-1) - V_\alpha^g(i)$ for an extremal policy g, we can deduce more easily verified sufficient conditions. For a control model in which we have a linear holding cost, the rigorous expression of $V_\alpha^g(i)$, $i \geq 0$, is possible using a busy-period regenerative process. In this connection, we consider a linear holding cost rate with $h(i) = h\cdot i$. Let $B^*(s)$ and $G^*(s)$ be the Laplace-Stieltjes transforms of the service-time distribution and the busy-period distribution, respectively. Then we have

$$G^*(s) = B^*[s + \lambda - \lambda G^*(s)].$$

For an $M/M/1$ queueing model under a full-service policy, we have

$$G^*(s) = \frac{\bar{\mu} + \lambda + s - [(\bar{\mu} + \lambda + s)^2 - 4\mu\lambda]^{1/2}}{2\lambda}. \tag{4.5}$$

From the Appendix, we obtain the expression for $V_\alpha^g(i)$ for the full-service policy case with $g(i) = \bar{\mu}$,

$$\begin{aligned} V_\alpha^g(i) = &- \frac{(1 - G^*(\alpha))h\lambda}{\alpha^2} + \frac{hG^*(\alpha)(1 - [G^*(\alpha)]^i)}{\alpha(1 - G^*(\alpha))} - \frac{hi}{\alpha} \\ &- \frac{C(\bar{\mu})}{\alpha} + \frac{[G^*(\alpha)]^i C(\bar{\mu})}{\alpha+\lambda-\lambda G^*(\alpha)}. \end{aligned} \tag{4.6}$$

For the passive-service case with $g(i) = 0$, we have

$$(4.7) \qquad V_\alpha^g(i) = -\frac{h\lambda}{\alpha^2} - \frac{hi}{\alpha}.$$

Given the expression for $V_\alpha^g(i)$ with extremal policies for the control model with a linear holding cost, we present the sufficient condition for either extremal policy in the next thorem.

Theorem 4.4. For an $M/M/1$ control model with a linear holding cost, (i) a sufficient condition that a full-service policy is optimal for all $i \varepsilon S$ is

$$(4.8) \qquad \frac{C(\bar{\mu}) - C(\mu)}{\bar{\mu} - \mu} < \frac{h(1 - G^*(\alpha))}{\alpha} + \frac{(1 - G^*(\alpha))}{\alpha+\lambda-\lambda G^*(\alpha)} C(\bar{\mu})$$

for all $\mu \varepsilon [0,\mu)$;

(ii) a sufficient condition that a passive policy is optimal for all $i \varepsilon S$ is

$$(4.9) \qquad \frac{C(\mu)}{\mu} \geq \frac{h}{\alpha}$$

for all $\mu \varepsilon [0,\bar{\mu}]$.

Proof: Since $V^g(i-1) - V^g(i)$ is nondecreasing in i, (4.3) is satisfied for all i if

$$C(\bar{\mu}) - C(\mu) < (\bar{\mu}-\mu)(V_\alpha^g(0) - V_\alpha^g(1)).$$

Similarly, (4.4) is satisfied for all i if

$$C(\mu) \geq \mu \lim_{i\to\infty} (V_\alpha^g(i-1) - V_\alpha^g(i)).$$

From (4.6) and (4.7) we have after some algebra (4.8) and (4.9).

Corollary 4.2. For an $M/M/1$ control model with a linear holding cost and with $C(\bar{\mu})/\bar{\mu} < C(\mu)/\mu$ for all $\mu \varepsilon [0,\bar{\mu}]$ (i) a sufficient condition that a full-service policy is optimal for all $i \varepsilon S$ is

$$(4.10) \qquad C(\bar{\mu})/\bar{\mu} < \left[\frac{(1 - G^*(\alpha))(\alpha+\lambda - \lambda G^*(\alpha))}{(\alpha+\lambda-\lambda G^*(\alpha)-\bar{\mu}+\bar{\mu}G^*(\alpha))}\right][h/\alpha],$$

(ii) a sufficient condition that a passive policy is optimal for all $i \varepsilon S$ is

$$(4.11) \qquad C(\bar{\mu})/\bar{\mu} \geq h/\alpha.$$

Proof: By Corollary 2.1, the optimal policy is either $\bar{\mu}$ or 0 for all states. From (4.8) with $\mu = 0$, we have (4.10) after rearrangement. We also obtain (4.11) from (4.9) with $\mu = \bar{\mu}$.

When the value $C(\bar{\mu})/\bar{\mu}$ satisfies neither (4.10) nor (4.11) for the control model considered in Corollary 4.2, there exists a critical state above which the full-service option is optimal.

5. Long-run Average-Return Problem

Now we consider a long-run average return problem of our control model. To prove the existence of strongly optimal policies with monotonicity properties for our control model with infinite denumerable state space and an unbounded return, we use an extremal-policy shift transformation in the proof.

For each policy π, let $V^{\pi}(i,t)$ denote the total expected reward earned by time t when starting from state i. For each initial state i, we define $\phi_{\pi}(i)$, the expected average return per unit time, as

$$\phi_{\pi}(i) = \liminf_{t\to\infty} V^{\pi}(i,t)/t.$$

A policy π^* is optimal if

$$\phi_{\pi^*}(i) = \sup_{\pi} \phi_{\pi}(i) \text{ for all } i \in S.$$

Theorem 5.1. For the long-run average-return problem, if $\{c(\bar{\mu}) - c(\mu)\}/(\bar{\mu}-\mu) \le M < \infty$ for all $\mu \in [0,\mu)$, the function $v(i) = \lim_{\alpha\to 0^+} \{V_{\alpha}(i) - V_{\alpha}(0)\}$ exists and satisfies the functional equation, $i \ge 1$,

$$v(i) = \max_{0\le\mu\le\bar{\mu}} \{-c(\mu) - h(i) + \lambda v(i+1) + \mu v(i-1) + (\bar{\mu}-\mu)v(i) - \bar{g}\}/\Lambda \tag{5.1}$$

where $v(0) = \{-h(0) + \lambda v(1) + \bar{\mu}v(0) - \bar{g}\}/\Lambda$, and $\bar{g}$ is the optimal long-run average return and $\bar{g} = \lim_{\alpha\to 0^+} \alpha V_{\alpha}(0)$. Moreover, the stationary policy, $\mu^*(i)$, that attains the maximum in (5.1) is optimal and a nondecreasing function of i.

Proof: Let g be a full-service policy. First, we prove that $V_{\alpha}^{g}(i-1) - V_{\alpha}^{g}(i)$ goes to infinity as $i\to\infty$ and $\alpha\to 0^+$. Let $V_{n,\alpha}^{g}$ be the n-period value function under policy g. By Theorem 4.2, we have that $V_{n,\alpha}^{g}(i-1) - V_{n,\alpha}^{g}(i)$ is nondecreasing in i and nondecreasing in n. Let $\delta = h(1) - h(0)$. Then

$$(\alpha+\Lambda)\{V_{n,\alpha}^{g}(i-1) - V_{n,\alpha}^{g}(i)\} \ge \delta + \lambda\{V_{n-1,\alpha}^{g}(i) - V_{n-1,\alpha}^{g}(i+1)\} + \bar{\mu}\{V_{n-1,\alpha}^{g}(i-2) - V_{n-1,\alpha}^{g}(i-1) \ge \delta + \Lambda\{V_{n-1,\alpha}^{g}(i-2) - V_{n-1,\alpha}(i-1)$$

Iterating this relation gives

$$V_{n,\alpha}^{g}(i-1) - V_{n,\alpha}^{g}(i) \ge \delta \cdot \sum_{j=0}^{\min[n-1,i-1]} [\Lambda/(\alpha+\Lambda)]^{j}/(\alpha+\Lambda).$$

Then

$$\lim_{n\to\infty}\lim_{i\to\infty} \{V_{n,\alpha}^{g}(i-1) - V_{n,\alpha}^{g}(i)\} \ge \frac{\delta}{\alpha+\Lambda}\left(1 + \frac{\Lambda}{\alpha+\Lambda} + \left(\frac{\Lambda}{\alpha+\Lambda}\right)^2 + \ldots\right) = \delta/\alpha.$$

Since $V^g_{n,\alpha}(i-1) - V^g_{n,\alpha}(i)$ is monotonically increasing in n and i, we have by interchanging limits

$$\lim_{i\to\infty}\left[V^g_\alpha(i-1) - V^g_\alpha(i)\right]$$
$$= \lim_{i\to\infty}\lim_{n\to\infty}\left[V^g_{n,\alpha}(i-1) - V^g_{n,\alpha}(i)\right]$$
$$= \lim_{n\to\infty}\lim_{i\to\infty}\left[V^g_{n,\alpha}(i-1) - V^g_{n,\alpha}(i)\right] \geq \delta/\alpha.$$

Therefore, $V^g_\alpha(i-1) - V^g_\alpha(i)$ goes to infinity as $i\to\infty$ and $\alpha\to 0^+$. Since $(C(\bar{\mu}) - C(\mu))/(\bar{\mu} - \mu)$ is bounded, the sufficient condition (4.3) for full-service optimal policy is satisfied as $i\to\infty$ for the long-run average-return problem. We assumed that $\bar{\mu} > \lambda$. Hence, the assumptions of Theorem 4 of Lippman [10] are readily satisfied. Therefore, $V(\cdot)$ exists and satisfies (5.1) and any stationary policy that chooses a service rate that maximizes the right-hand side of (5.1) is average optimal. $\bar{g} = \lim_{\alpha\to\infty^+} \alpha V_\alpha(0)$ follows from the Tauberian theorem (see [6] and Ross [13]). Left continuity of $C(\mu)$ and the monotonicity of $\mu_\alpha{*}(i)$ in α give

$$\max_{0\leq\mu\leq\bar{\mu}} [-C(\mu) + \mu\{v(i-1) - v(i)\}]$$
$$= \lim_{\alpha\to 0^+} \max_{0\leq\mu\leq\bar{\mu}} [-C(\mu) + \mu\{V_\alpha(i-1) - V_\alpha(i)\}]$$
$$= \lim_{\alpha\to 0^+} [-C(\mu_\alpha{*}(i)) + \mu_\alpha{*}(i)\{V_\alpha(i-1) - V_\alpha(i)\}]$$
$$= -C(\mu{*}(i)) + \mu{*}(i)\{v(i-1) - v(i)\},$$

so that $\mu{*}(i)$ is average optimal. From Theorem 4.1 and Corollary 4.1, we also see that $v(i-1)-v(i)$ is nonnegative nondecreasing in i, so that $\mu{*}(i)$ is a nondecreasing function of i. This completes the proof.

We have established the monotonicity properties of optimal policies for the long-run average-return problem. What distinguishes our arguments from other induction arguments in the monotonicity proofs is that we use the extremal-policy shift transformation.

Our control model is a special case of the right-skip-free Markov decision process (see [23]). By exploiting the special structure inherent to our model, we can obtain an equivalent functional equation for the long-run average-return, which is computationally more tractable. Low [12] used a similar approach to an arrival control model. From Theorem 5.1, we know that if $\bar{g}$ and the function $v(\cdot)$ are given, then an optimal stationary policy can be obtained by solving (5.1). A computationally more tractable functional equation is introduced by the next theorem.

Theorem 5.2. Let $w(i) = v(i) - v(i+1)$. Suppose $\phi_{\pi*} = \overline{g}$ for some stationary policy $\pi* = (\mu*(0), *(1),\ldots)$. Then $\mu*(i)$ maximizes

(5.2) $$w(i) = \max_{0\le\mu\le\overline{\mu}} \{-C(\mu) - h(i) + \mu w(i-1) - \overline{g}\}/\lambda, \quad j \ge 1$$

where: $w(0) = \{-h(0) - \overline{g}\}/\lambda$.

Proof: From (5.1) we have for all $\mu \varepsilon [i,\overline{\mu}]$

$$v(i) \ge \{-C(\mu) - h(i) + \lambda v(i+1) + \mu v(i-1) + (\overline{\mu}-\mu)v(i) - \overline{g}\}/\Lambda.$$

The equality holds for some μ (actually $\mu*(i)$). After simple algebraic manipulation, we have

$$w(i) \ge \{-C(\mu) - h(i) + \mu w(i-1) - \overline{g}\}/\lambda.$$

Hence, we have a new functional equation (5.2). This completes the proof.

The equivalent functional equation (5.2) can be efficiently used when we are to solve a long-run average-return problem numerically, since the value function can be calculated recursively from (5.2). For details, see [23].

6. Appendix

We compute the functional values under a given policy which uses the same service rate whenever the system is not idle, using regeneration processes (cf. Heyman [8], Bell [1], Schleef [16]). Since a fixed service rate is used whenever the server is turned on, the service structure can be regarded as a special case of Bell [1]. The regeneration points for our system are the epochs at which the server begins to be idle.

We consider the holding costs and service costs separately. First, we consider the holding cost sector. Let $b(i)$ be the expected α-discounted holding cost until the next regeneration epoch starting with i customers in the system. When a unit cost rate is continuously charged for a random time T with a distribution function $F(\cdot)$, the expected cost is $(1 - F*(\alpha))/\alpha$, where $F*(s)$ is the Laplace-Stieltjes transform of $F(\cdot)$. Let $\eta = G*(\alpha)$, where $G(\cdot)$ is the distribution function for the busy period. Using these arguments we have

$$b(i+1) = hi(1-\eta)/\alpha + b(1) + \eta b(i), \quad i \ge 1.$$

For our probability structure $b(1)$ is expressed as

$$b(1) = h(1 - \frac{\Lambda}{\alpha + \Lambda})/\alpha + \frac{\lambda}{\alpha + \Lambda} b(2) + \frac{\overline{\mu} - \mu}{\alpha + \Lambda} b(1).$$

Using these relations and $G*(s) = B*[s + \lambda - \lambda G*(s)]$ gives, after some algebra,

$$(6.1)\qquad b(1) = h(1-\eta)/\alpha + h\lambda(1-\eta)^2/\alpha^2.$$

This yields the new recursive relation of $b(\cdot)$ as follows

$$(6.2)\qquad b(i) = hi(1-\eta)/\alpha + h\lambda(1-\eta)^2/\alpha^2 + \eta b(i-1).$$

Let $k(i) = hi(1-\eta)/\alpha + h\lambda(1-\eta)^2/\alpha^2$, then

$$b(i) = k(i) + b(i-1)$$

$$= \frac{(1-\eta)h\lambda(1-\eta^i)}{\alpha^2} + \frac{h(1-\eta)}{\alpha}\eta^i[\sum_{j=0}^{i} j(1/\eta)^j].$$

We use the relation $\sum_{j=0}^{n} jx^j = x\frac{(1-x^n)}{(1-x)^2} - \frac{nx^{n+1}}{1-x}$ to obtain

$$(6.3)\qquad b(i) = \frac{(1-\eta)(1-\eta^i)h\lambda}{\alpha^2} - \frac{h\eta(1-\eta^i)}{\alpha(1-\eta)} + hi/\alpha,\ i \ge 1.$$

Let $H(i)$ be the expected discounted total holding cost over an infinite horizon starting with i customers in the system and let $A^*(s)$ be the Laplace-Stieltjes transform of the interarrival distribution. Then

$$(6.4)\qquad H(0) = b(0)[1 + A^*(\alpha)\eta + (A^*(\alpha)\eta)^2 + \dots\dots\dots],$$

where $b(0) = A^*(\alpha)b(1)$.
From (6.1) and (6.4), we have

$$(6.5)\qquad H(0) = (1-\eta)\lambda h/\alpha^2.$$

After the next i busy periods, $H(i)$ becomes

$$H(i) = b(i) + \eta^i H(0).$$

Hence (6.3) and (6.5) yield

$$(6.6)\qquad H(i) = \frac{(1-\eta)h\lambda}{\alpha^2} - \frac{h\eta(1-\eta^i)}{\alpha(1-\eta)} + hi/\alpha.$$

Let $w(i)$ be the expected discounted service cost until the next regeneration epoch starting with i customers in the system and let $S(i)$ be the expected discounted total service cost for the infinite horizon starting with i customers in the system. By a similar approach applied for the holding cost sector, we have

$$(6.7)\qquad \begin{aligned} w(i) &= c(\mu)(1-\eta^i)/\alpha,\ i \ge 1,\\ w(0) &= A^*(\alpha)w(1). \end{aligned}$$

Then $S(0)$ is expressed as

$$(6.8)\quad \begin{aligned} S(0) &= w(0)[1 + A^*(\alpha)\eta + (A^*(\alpha)\eta)^2 + \dots\dots\dots] \\ &= w(0)/(1 - A^*(\alpha)\eta). \end{aligned}$$

From (6.7) and (6.8) we have

$$S(0) = \frac{\lambda c(\mu)(1 - \eta)}{\alpha(\alpha + \lambda - \lambda\eta)} .$$

After the ith busy period, $S(i)$ becomes

$$S(i) = w(i) + \eta^i S(0).$$

Hence (10) and (12) yield

$$(6.9)\quad S(i) = c(\mu)/\alpha - \frac{\eta^i c(\mu)}{(\alpha + \lambda - \lambda\eta)} .$$

Summing up the two cost sectors, we have

$$V^g(i) = -[H(i) + S(i)],\ i \geq 0.$$

If g is the full service policy, $\mu = \bar{\mu}$,

$$(6.10)\quad V^g(i) = -\frac{(1 - \eta)h}{\alpha^2} + \frac{h\eta(1 - \eta^i)}{\alpha(1 - \eta)} - \frac{hi}{\alpha} - \frac{c(\bar{\mu})}{\alpha} + \frac{\eta^i c(\bar{\mu})}{\alpha + \lambda - \lambda\eta} .$$

If g is the passive policy, $\mu = 0$,

$$(6.11)\quad V^g(i) = -hi/\alpha - h\lambda/\alpha^2.$$

References

[1] Bell, C.E.: Characterization and Computation of Optimal Policies for Operating an M/G/1 Queueing System with Removable Server. *Opns. Res.* 19, 208-217, 1971.

[2] Bengtsson, B.: Optimal Control of Queues, Part II: The Infinite-Capacity Case. *Report 415*, Dept. of Electrical Engineering, Linköping Univ., Sweden, 1980.

[3] Crabill, T. B.: Optimal Control of Service Facility with Variable exponential Service Times and Constant Arrival Rate. *Management Sci.* 18, 560-566, 1972.

[4] Crabill, T. B.: Optimal Control of a Maintenance System with Variable Service Rates. *Opns. Res.* 22, 736-745, 1974.

[5] Crabill, T. B. Gross, D. and Magazine, M. J.: A Classified Bibliography of Research on Optimal Design and Control of Queues. *Opns. Res.* 25, 219-232, 1977.

[6] Feller, W.: *An Introduction to Probability Theory and Its Applications* Vol. II, Wiley, New York, 1971.

[7] Harrison, J. M.: Discrete Dynamic Programming with Unbounded Rewards. *Ann. Math. Statist.* 43, 636-644, 1972.

[8] Heyman, D. P.: Optimal Operating Policies for M/G/1 Queueing Systems. *Opns. Res.* 16, 362-382, 1968.

[9] Hinderer, K.: *Foundation of Non-Stationary Programming with Discrete Time Parameter.* Lecture Notes in Operations Research and Mathematical Systems 33, New York, Springer-Verlag, 1970.

[10] Lippman, S. A.: Semi-Markov Decision Processes with Unbounded Rewards. *Management Sci.* 19, 717-731, 1973.

[11] Lippman, S. A.: Applying a New Device in the Optimization of Exponential Queueing System. *Opns. Res.* 23, 687-710, 1975.

[12] Low, D. W.: Optimal Dynamic Pricing Policies for an M/M/1 Queue. *Opns. Res.* 22, 545-561, 1974.

[13] Ross, S. M.: *Applied Probability with Optimization Applications.* Holden-Day, San Francisco, 1970.

[14] Sabeti, H.: Optimal Selection of Service Rates in Queueing with Different Cost. *J. Opns, Res. Soc. of Japan* 16, 15-35.

[15] Schäl, M.: Conditions for Optimality in Dynamic Programming and for the Limit of n-Stage Optimal Policies to be Optimal. *Wahrscheinlichkeitstheorie Verw. Geb.* 32, 179-196, 1975.

[16] Schleef, H.: Optimal Control Models for Multi-Server Exponential Queueing Systems. Unpublished PhD Dissertation, University of Chicago, 1977.

[17] Serfozo, R. F.: Monotone Optimal Policies for Markov Decision Processes. *Math. Programming Study* 6, 202-215, 1976.

[18] Serfozo, R. F.: Optimal Control of Random Walks, Birth and Death Processes, and Queues. *Adv. Appl. Prob.* 13, 61-83, 1981.

[19] Sobel, M. J.: Optimal Operations of Queues. *Mathematical Methods in Queueing Theory.* Lecture Notes in Economics and Mathematical Systems 98, New York, Springer-Verlag, 231-261, 1974.

[20] Sobel, M. J.: The Optimality of Full-service Policies. *Opns. Res.* 30, 636-649, 1982.

[21] Stidham, S., Jr. and Prabhu, N. U.: Optimal Control of Queues. *Mathematical Methods in Queueing Theory.* Lecture Notes in Economics and Mathematical Systems 98, New York, Springer-Verlag, 263-293, 1974.

[22] Stidham, S., Jr. and van Nunen, J. A. A. E.: The Shift-Function Approach for Markov Decision Processes with Unbounded Returns. *O. R. Report 173.* North Carolina State University, 1981.

[23] Stidham, S., Jr. and Wijngaard, J.: Average-Reward Markovian Decision Processes with Skip-free-to-the-right Transitions. Paper presented at the Joint Meeting of ORSA/TIMS, Colorado Springs, November 10-12, 1980.

[24] Topkis, D.: Minimizing a Submodular Function on a Lattice. *Opns. Res.* 26, 305-321, 1978.

[25] van Nunen, J. A. A. E.: *Contracting Markov Decision Processes.* Mathematical Centre Tracts 71, the Netherlands, 1976.

Kyung Y. Jo: Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030, U.S.A.