Adv. Appl. Prob. **19**, 202–218 (1987)
Printed in N. Ireland

OPTIMAL CONTROL OF SERVICE RATES IN NETWORKS OF QUEUES

RICHARD R. WEBER,* *University of Cambridge*
SHALER STIDHAM JR, ** *University of North Carolina*

Abstract

We prove a monotonicity result for the problem of optimal service rate control in certain queueing networks. Consider, as an illustrative example, a number of $\cdot/M/1$ queues which are arranged in a cycle with some number of customers moving around the cycle. A holding cost $h_i(x_i)$ is charged for each unit of time that queue i contains x_i customers, with h_i being convex. As a function of the queue lengths the service rate at each queue i is to be chosen in the interval $[0, \bar{\mu}]$, where cost $c_i(\mu)$ is charged for each unit of time that the service rate μ is in effect at queue i. It is shown that the policy which minimizes the expected total discounted cost has a monotone structure: namely, that by moving one customer from queue i to the following queue, the optimal service rate in queue i is not increased and the optimal service rates elsewhere are not decreased. We prove a similar result for problems of optimal arrival rate and service rate control in general queueing networks. The results are extended to an average-cost measure, and an example is included to show that in general the assumption of convex holding costs may not be relaxed. A further example shows that the optimal policy may not be monotone unless the choice of possible service rates at each queue includes 0.

CONTROL OF QUEUES; CYCLES OF QUEUES; DYNAMIC PROGRAMMING; MONOTONE POLICIES; SERIES OF QUEUES

1. Control of queues

The control of arrival and service rates in a single queue is well studied in the literature of queueing and communications theory (see Sobel (1974), Stidham and Prabhu (1974), Crabill et al. (1977), Johansen and Stidham (1980), Serfozo (1981) and Stidham (1985) for surveys and examples). Studies have described various circumstances under which an optimal control policy has monotone structure. It is usually assumed that a holding cost is charged

Received 24 June 1983; revision received 19 November 1985.

* Postal address: Queens' College, Cambridge, CB3 9ET, UK.

** Postal address: Department of Operations Research and Systems Analysis, University of North Carolina, Chapel Hill, NC 27514, USA.

The research of this author was partially supported by the U.S. Army Research Office, Contract DAAG29-82-k-0152, at North Carolina State University, Raleigh, NC, USA, and by the Science and Engineering Research Council, at the University of Cambridge, UK.

according to the number of customers in the queue, rewards are gained when customers enter the queue and costs are paid while customers are served. Expected total-discounted or time-average cost is to be minimized by controlling the arrival and/or service rate. In many circumstances, it turns out that the optimal arrival rate and service rate are respectively non-increasing and non-decreasing in the number of customers (or amount of work) in the queue.

Recent research has considered the optimal control of systems of more than one queue. Winston (1977) and Weber (1978) showed that for a system of identical servers in parallel, where an arriving customer can be directed to any queue, sending the customer to the shortest queue minimizes the expected average waiting time. Ephremides et al. (1980) considered this problem for two parallel servers, as did Davis (1977) with the additional possibility that arrivals may be rejected. In these papers and most others the models considered are memoryless. Throughout the rest of this paper we are concerned only with models of this type. That is, all arrival processes are Poisson and all service times are exponentially distributed (though possibly with controllable parameters). A consequence of the memoryless assumption is that the service and arrival rates chosen by an optimal policy need only change at times of customer arrivals or service completions.

A number of authors have considered the control of the series system $M/M/1 \rightarrow /M/1$ in which there are Poisson arrivals at rate λ and two memoryless servers, serving at rates μ_1 and μ_2 in the first and second queues respectively. Ghoneim (1980) showed that, for fixed μ_1 and μ_2, and controllable λ, the optimal λ is non-increasing in the number of customers in either queue. Moreover, he showed that the same is true for a controllable process of arrivals at the second queue. Rosberg et al. (1982) considered a model in which λ and μ_2 are fixed, and μ_1 is controllable: $\mu_1 \in \{0, a\}$, and in which no cost is charged for use of the positive rate a. They showed that the optimal μ_1 is non-decreasing in the number of customers in the first queue and non-increasing in the number of customers in the second queue. Their proof makes use of a linear programming formulation and is therefore only applicable to a model with linear holding costs. A more general model for two queues has been considered by Hajek (1984). In this model queues 1 and 2 receive Poisson arrivals at rates λ_1 and λ_2 respectively. A third stream of arrivals at rate λ can be directed to either queue. The stations have servers of rates μ_1 and μ_2 and a third server of rate μ which can be assigned to either queue; customers whose service is completed by these servers leave the system. Finally, there are two servers, of rates γ_{12} and γ_{21}, the first of which serves queue 1 and sends customers to queue 2, the second of which serves queue 2 and sends customers to queue 1. Hajek restricts attention to linear holding costs. The rates γ_{12} and γ_{21} are to be controlled between 0 and fixed levels, the customers arriving at rate λ are to be routed to one or the other of the queues, and the server of rate μ is to be allocated to a queue; all these decisions are to be made dynamically, as a function of the number of customers in the queues. Hajek shows that all these controls have a monotone structure by demonstrating that the benefit obtained by removing a customer from queue 1 or queue 2 is an increasing function of the number of customers in either queue, and the benefit obtained by moving a customer from queue 1 to queue 2 is a increasing function of the number of customers in queue 1, and a decreasing function of the number of customers in the queue 2.

2. Optimal control of a network of queues

To illustrate the type of general queueing network with which this paper is concerned and to fix ideas, we begin by considering a cycle of m queues in which a customer who completes service in queue i joins queue $i+1$ (where we identify queue $m+1$ as queue 1). At queue i customers arrive from outside the cycle in a Poisson stream of rate λ_i. Customers in queue i are served by a memoryless server. The rate μ_i of this server is subject to control and may be varied continuously within the interval $[0, \bar{\mu}]$, where $\bar{\mu}$ is finite. A cost is charged at the rate of $c_i(\mu_i)$ per unit time while the service rate μ_i is in effect at queue i. We assume that each function c_i is continuous and convex. If there is a reward rather than a cost associated with serving queue i at a rate μ then $c_i(\mu)$ is negative. The number of customers in queue i will be denoted by x_i and a state of the system by the vector $x = (x_1, \dots, x_m)$, with $x_i \geqq 0$, $i = 1, \cdots, m$. The holding cost per unit time is given by $h(x) = \sum h_i(x_i)$, where each function h_i is non-negative, convex (but not necessarily monotone) in its argument.

Remark. In fact, the restriction of the service rates to an interval $[0, \bar{\mu}]$ and the requirement that each c_i be convex is without loss of generality. For suppose that the set of feasible service rates for queue i is closed and contains 0 and a maximal element $\bar{\mu}$, but is otherwise arbitrary, and that the service cost function c_i is continuous, but otherwise arbitrary. Then it can be shown (see Crabill (1972), Jo and Stidham (1983)) that an equivalent optimization problem results if the set of feasible service rates is extended to its convex hull, $[0, \bar{\mu}]$, and c_i is replaced by its lower convex envelope on $[0, \bar{\mu}]$.

With simultaneous control of all the service rates and with continuous discounting at rate $\alpha > 0$, our aim is to minimize the expected total discounted cost. (Notice that although we allow the service rates to be varied at any time, the memoryless assumptions imply that an optimal policy need only do so at times of customer arrivals or service completions.) We begin by considering the minimization of expected discounted costs over a horizon of n random stages (or observation points), in which we observe the system at every customer arrival or 'potential' service completion. In doing this, we adopt the convention that the server at queue i is always running, producing potential service completions at the maximum rate $\bar{\mu}$. When a potential service completion occurs, it is 'accepted' with probability $\mu_i/\bar{\mu}$ and 'rejected' with probability $1-(\mu_i/\bar{\mu})$. This 'uniformization' procedure has the desirable property that the time between observation points is rendered independent of the state of the control variables (being exponentially distributed with parameter $\Lambda = m\bar{\mu} + \sum \lambda_i$). Change of service rates need only be considered at observation points and therefore the dynamic programming equation takes a simple form in (1) below. The procedure is fairly standard in the literature on stochastic dynamic optimization (see Lippman (1975) and Serfozo (1979) and can be shown to result in a decision problem over the infinite horizon that is equivalent to the problem in which the system is observed continuously in time or at each arrival and actual service completion.

With this convention the expected one-stage total discounted cost when starting from state x and employing service rates μ_i, $i=1,\cdots,m$, is in effect $[h(x)+\sum c_i(\mu_i)]/(\alpha+\Lambda)$ and the expected one-stage discount factor is $\Lambda/(\alpha+\Lambda)$. Let e_i denote the m-component vector which has a 1 in component i and 0 in other components; let $d_i=(e_{i+1}-e_i)$. With probability λ_i/Λ the next observation point is an arrival to queue i and the system moves to state $x+e_i$. If the service rate in effect at queue i is μ_i then with probability μ_i/Λ the next observation point is an actual service completion at queue i and the system moves to state $x+d_i$. With probability $(m\bar{\mu}-\sum \mu_i)/\Lambda$ the next observation point is a 'null event', corresponding to the rejection of a potential service completion; in this case the system remains in state x.

Define $V_n(x)$ as the minimal n-state expected discounted cost starting from state x, with $V_0(x)=0$ for all x. Since $\alpha>0$ and the one-stage costs are uniformly bounded below (by $-\sum \max_{\mu_i}|c_i(\mu_i)|/(\alpha+\Lambda)$) it follows by standard results (Schal (1977), Bertsekas (1976), Whittle (1983)) that V_n is well defined for each n and satisfies the dynamic programming equation

$$V_{n+1}(x)=\sum_{i=1}^{m}\min_{\mu_i}\,[h_i(x_i)+c_i(\mu_i)+\lambda_i V_n(x+e_i)$$

$$+\mu_i V_n(x+d_i)+(\bar{\mu}-\mu_i)V_n(x)]/(\alpha+\Lambda)$$

where the minimization is over μ_i in $[0,\bar{\mu}]$, $i=1,\cdots,m$, and it is understood that $\mu_i=0$ is selected if $x_i=0$. Without loss of generality suppose $\alpha+\Lambda=1$.

Then we can rewrite these equations in the equivalent form

$$
\begin{aligned}
V_{n+1}(x) = {} & \sum_{i=1}^{m} [h_i(x_i) + \lambda_i V_n(x + e_i) + \bar{\mu} V_n(x)] \\
& + \sum_{i=1}^{m} \min_{\mu_i} \{c_i(\mu_i) - \mu_i[V_n(x) - V_n(x + d_i)]\}.
\end{aligned} \tag{1}
$$

A decision rule for selecting the service rates can be described by a function $\mu(x) = (\mu_1, \cdots, \mu_m)$, in which $\mu_i(x)$ is the service rate chosen for queue i when the state is x. An optimal policy for the finite-horizon problem at stage $n + 1$ is specified by a decision rule $\mu(x)$ such that, for each $i = 1, \cdots, m,$ the expression in { } on the right-hand side of (1) is minimized by $\mu_i = \mu_i(x)$. We shall resolve ties by selecting the smallest minimizer.

It does not seem possible to prove a general result concerning the way in which the optimal service rates might change as a single customer is added to queue i. However, we can say what happens as a customer moves from queue i to queue $i + 1$. We call decision rule $\mu(x)$ *transition-monotone* if we have

$$\mu_j(x) \leqq \mu_j(x + d_i), \quad \text{for each} \quad i \neq j.$$

Our main theorem states that there is an optimal policy with a transition-monotone decision rule. In other words, when a customer moves from queue i to queue $i + 1$, the service rate at each queues j, $j \neq i$, does not decrease. An immediate corollary of this is that the service rate at queue i does not increase. To prove the corollary, one imagines taking a customer around the cycle: from queue $i+1$ to queue $i+2$ to queue $i+3$ to . . . to queue i. The theorem implies that at each movement from queue to queue, the optimal service rate at queue i does not decrease. The state reached after the final movement is just that which would have been obtained if the customer had been moved directly from queue $i + 1$ to queue i (that is, $-d_i = d_{i+1} + \cdots + d_{i-1}$). Therefore, when a customer moves from queue i to queue $i + 1$, the service rate at queue i does not increase. The following theorem gives the main result of this paper.

Theorem. There exist transition-monotone optimal policies for both the n-stage and the infinite-horizon problems.

Proof. An examination of the expression { } in (1) shows that it suffices to prove

$$V_n(x) - V_n(x + d_i) - V_n(x + d_j) + V_n(x + d_i + d_j) \leqq 0, \tag{2}$$

for all $j \neq i$ and x such that x, $x + d_i$, $x + d_j$ and $x + d_i + d_j \geqq 0$. We use induction on n with (1) to show that if (2) holds, then it also holds when n is replaced by $n+1$. The induction begins trivially, with $V_0 = 0$. Notice that if we write the left side of (2) with $n+1$ replacing n, and then substitute for V_{n+1} from (1), letting the summation be over index k, then the terms multiplied by λ_k give

$$\lambda_k[V_n(x+e_k) - V_n(x+e_k+d_i) - V_n(x+e_k+d_j) + V_n(x+e_k+d_i+d_j)].$$

This is non-positive trivially by the inductive hypothesis for n. Similarly, the inductive step is easy for the terms multiplied by $\bar{\mu}$. By the convexity of h_k the sum of terms in h_k also results in a non-negative quantity. To check the inductive step for the terms coming from $\{\ \}$ in (1) we shall prove a lemma whose corollary is the following proposition (for this cycle-of-queues problem). The application of this result with $g = V_n$ establishes the inductive step for the terms in $\{\ \}$ and completes the proof of the theorem for the cycle-of-queues problem.

Proposition. Suppose g is a function such that

$$(3) \qquad g(x) - g(x+d_i) - g(x+d_j) + g(x+d_i+d_j) \leqq 0,$$

for all $i \neq j$ and x such $x,\ x+d_i,\ x+d_j,\ x+d_i+d_j \geqq 0$. Define for some k,

$$f_k(x) = \min\{c_k(\mu) + \mu g(x+d_k) + (\bar{\mu}-\mu)g(x)\},$$

where the minimization is over μ in the interval $[0, \bar{\mu}]$ and it is understood that $\mu = 0$ is selected if $x_k = 0$. Then (3) also holds when g is replaced by f_k.

Rather than prove the proposition as it stands, we shall reformulate these ideas in a more general setting. The reader may find it helpful to hold in mind the example of the cycle-of-queues model as we proceed. Suppose for the moment that all states $x \in \mathbb{Z}^m$ are permissible. Let the q vectors $d_1, \cdots, d_q \in \mathbb{Z}^m$ represent possible transformations $d_i : \mathbb{Z}^m \to \mathbb{Z}^m$ defined by $x \to x + d_i$. Here d_i can be any vector in $\mathbb{Z}^m$: the action of $x \to x + d_i$ is to add customers to some queues and subtract them from others. If in state x the rate of server i is chosen to be $\mu_i \in [0, \bar{\mu}]$, then a cost $c_i(\mu_i)$ is incurred per unit time and the system moves to state $x + d_i$ at rate μ_i. (Note the difference between the general model and the cycle-of-queues model: in the general model, server i is associated with a transformation of the system $x \to x + d_i$, which may be much more complex than just the movement of one customer from queue i to $i+1$. We have also allowed the count of customers x_i to be negative.) Let $\boldsymbol{D}$ be the set of all subsets of $\{d_1, \cdots, d_q\}$ (including the empty set). Consider the partial order on $\boldsymbol{D}$ given by $D_i \leqq D_j$ if and only if $D_i \subseteq D_j$. For this partial order, $\boldsymbol{D}$ is a lattice. The greatest lower bound and least upper bound of D_i and D_j are denoted $D_i \wedge D_j$ and $D_i \vee D_j$. These are in fact $D_i \cap D_j$ and $D_i \cup D_j$ respectively. The function $f : \boldsymbol{D} \to R$ is said to be *submodular* on $\boldsymbol{D}$ if for all D_i,

$D_j \in \boldsymbol{D}$ we have

$$f(D_i \cap D_j) - f(D_i) - f(D_j) + f(D_i \cup D_j) \leqq 0.$$

For $D = \{d_{i_1}, \cdots, d_{i_k}\} \in \boldsymbol{D}$, let $x + D$ denote $x + d_{i_1} + \cdots + d_{i_k}$. We now suppose that not all $x \in \mathbb{Z}^m$ are actually permissible states, but that only those $x \in X$ are actually permissible, where X is a subset of $\mathbb{Z}^m$. We say that $\boldsymbol{D}$ is compatible with X if for all $x \in \mathbb{Z}^m$ and D_i, $D_j \in \boldsymbol{D}$ such that $x + D_i$ and $x + D_j$ are in X, then $x + D_i \cap D_j$ and $x + D_i \cup D_j$ are also in X. For each $x \in X$ let $\boldsymbol{D}(x) = \{D \in \boldsymbol{D} : x + D \in X\}$. It is easy to check that if $\boldsymbol{D}$ is compatible with X then $\boldsymbol{D}(x)$ is a partially ordered subset of $\boldsymbol{D}$ which is closed under operations of intersection and union of its members. Note that it is X and the set $\{d_1, \cdots, d_q\}$ that define the geometry of the queueing network. We now state the key lemma which generalizes the proposition above.

Lemma. Suppose that we are given $x \in X$, $d_k \in \{d_1, \cdots, d_q\}$ and a function $g: X \to R$. Suppose $\boldsymbol{D}$ is compatible with X. Define $g(x, \cdot)$ on $\boldsymbol{D}$ by

$$g(x, D) = \begin{cases} g(x + D) \\ \infty \end{cases} \quad \text{as} \quad x + D \begin{cases} \in X \\ \notin X. \end{cases}$$

Suppose that for all $x \in X$ the function $g(x, \cdot)$ is submodular on $\boldsymbol{D}(x)$. Define $f_k(x, \cdot)$ on $\boldsymbol{D}$ by

$$f_k(x, D) = \min \{c_k(\mu) + \mu g(x + d_k, D) + (\bar{\mu} - \mu) g(x, D)\},$$

where the minimization is over $\mu \in [0, \bar{\mu}]$. Then $f_k(x, D)$ is also submodular on $\boldsymbol{D}(x)$.

Note the importance of the assumption that μ can be controlled to 0. It ensures that if $x + d_k + D$ is not in X then we can take $\mu = 0$, thereby avoiding $f_k(x, D) = \infty$. Thus $f(x, \cdot)$ is finite on $\boldsymbol{D}(x)$. The above lemma puts our work in the context of the theory of submodular functions on a lattice which has been developed by Topkis (1978) and others. For example, in another paper dealing with general models, Serfozo (1981) uses submodularity to prove monotonicity results for simple random walk models. However, our lemma does not appear to be implied by previously known results; it differs in the essential prominence given to the condition that controllable service rates be controllable to 0. Before proving the lemma, we remark that for the cycle-of-queues problem, $X = \{(x_1, \ldots, x_m) : x_i \geqq 0,\ i = 1, \cdots, m\}$. It is easy to check that taking $d_i = (e_{i+1} - e_i)$, $i = 1, \cdots, m - 1$, and $d_m = (e_1 - e_m)$, results in a $\boldsymbol{D}$ which is compatible with X. With only a few other alterations, the reader will obtain a proof of the specialized proposition for the cycle-of-queues problem by reading the following proof for the general setting and everywhere mentally removing

$\{\ \}$'s, replacing 'D' with 'd' and '$\in D$' with '$=d$', and within all $g(\cdot,\cdot)$'s replacing ',' with '+'.

Proof. Consider D_i, $D_j \in \boldsymbol{D}(x)$. It is required to prove

$$\Delta = f_k(x, D_i \cap D_j) - f_k(x, D_i) - f_k(x, D_j) + f_k(x, D_i \cup D_j) \leqq 0.$$

Without loss of generality, we can assume that $D_i \cap D_j = \varnothing$. For if not, we can let $x' = x + D_i \cap D_j$, $D_i' = D_i - D_i \cap D_j$ and $D_j' = D_j - D_i \cap D_j$, and note that Δ is unchanged if x, D_i and D_j are replaced by x', D_i' and D_j'. So assuming $D_i \cap D_j = \varnothing$ we write $D_i \cup D_j = D_i + D_j$.

Suppose that for $i \neq j$, the minimizing controls are $0 \leqq \mu_i \leqq \bar{\mu}$, $0 \leqq \mu_j \leqq \bar{\mu}$, such that

$$f_k(x + D_i) = c_k(\mu_i) + \mu_i g(x + d_k, D_i) + (\bar{\mu} - \mu_i) g(x, D_i),$$
$$f_k(x + D_j) = c_k(\mu_j) + \mu_j g(x + d_k, D_j) + (\bar{\mu} - \mu_j) g(x, D_j).$$

Assume $\mu_i \geqq \mu_j$ (the case $\mu_i \leqq \mu_j$ being symmetric). We consider two cases, the first being $d_k \notin D_j$. Then

$$\begin{aligned}
\Delta \leqq\ & c_k(\mu_j) + \mu_j g(x + d_k) + (\bar{\mu} - \mu_j) g(x) \\
& - c_k(\mu_i) - \mu_i g(x + d_k, D_i) - (\bar{\mu} - \mu_i) g(x, D_i) \\
& - c_k(\mu_j) - \mu_j g(x + d_k, D_j) - (\bar{\mu} - \mu_j) g(x, D_j) \\
& + c_k(\mu_i) + \mu_i g(x + d_k + D_i + D_j) + (\bar{\mu} - \mu_i) g(x, D_i + D_j) \\
=\ & (\bar{\mu} - \mu_j)[g(x) - g(x, D_i) - g(x, D_j) + g(x, D_i + D_j)] \\
& + (\mu_i - \mu_j)[g(x, D_i) - g(x + d_k, D_i) - g(x, D_i + D_j) + g(x + d_k, D_i + D_j)] \\
& + \mu_j[g(x + d_k) - g(x + d_k, D_i) - g(x + d_k, D_j) + g(x + d_k, D_i + D_j)].
\end{aligned}$$

On the right-hand side of the above, the first term is non-positive by the submodularity of $g(x, \cdot)$. If $\mu_i > 0$, then this implies $x + d_k + D_i \in X$. Also, since $\boldsymbol{D}$ is compatible with X, we can take the union of $\{d_k\}$ and D_j, which are both in $\boldsymbol{D}(x + D_i)$, to deduce $x + d_k + D_i + D_j \in X$. Thus the second term is equal to

$$(\mu_i - \mu_j)[g(x + D_i, 0) - g(x + D_i, \{d_k\}) - g(x + D_i, D_j) + g(x + D_i, \{d_k\} + D_j)],$$

which is non-positive because $g(x + D_i, \cdot)$ is submodular and $\{d_k\} \cap D_j = 0$. If also $\mu_j > 0$, then this implies $x + d_k + D_j$, and by taking the intersection and union of $\{d_k\} + D_i$ and D_j, which are both in $\boldsymbol{D}(x)$, we can deduce $x + d_k$, $x + d_k + D_i + D_j \in X$. Thus the third term is non-positive by the submodularity of $g(x + d_k, \cdot)$.

Now suppose $d_k \in D_j$, $d_k \notin D_i$. Then

$$\begin{aligned}\Delta \leqq\; & c_j(\mu_i) + \mu_i g(x + d_k) + (\bar{\mu} - \mu_i) g(x) \\ & - c_j(\mu_i) - \mu_i g(x + d_k, D_i) - (\bar{\mu} - \mu_i) g(x, D_i) \\ & - c_j(\mu_j) - \mu_j g(x + d_k, D_j) - (\bar{\mu} - \mu_j) g(x, D_j) \\ & + c_j(\mu_j) + \mu_j g(x + d_k, D_i + D_j) + (\bar{\mu} - \mu_j) g(x, D_i + D_j) \\ =\; & (\bar{\mu} - \mu_i)[g(x) - g(x, D_i) - g(x, D_j) + g(x, D_i + D_j)] \\ & + (\mu_i - \mu_j)[g(x + d_k) - g(x + d_k, D_i) - g(x, D_j) + g(x, D_i + D_j)] \\ & + \mu_j[g(x + d_k) - g(x + d_k, D_i) - g(x + d_k, D_j) + g(x + d_k, D_i + D_j)].\end{aligned}$$

On the right-hand side of the above, the first term is non-positive by the submodularity of $g(x, \cdot)$. If $\mu_i > 0$, then this implies $x + d_k + D_i \in X$. By applying the compatibility condition to the intersection of D_j and $\{d_k\} + D_i$, which are both in $\boldsymbol{D}(x)$, we can deduce $x + d_k \in X$. So noting that $d_k \notin D_i$, we write the second term as

$$(\mu_i - \mu_j)[g(x + d_k) - g(x + d_k, D_i) - g(x + d_k, D_j') + g(x + d_k, D_i + D_j')],$$

where $D_j' = D_j - \{d_k\}$. This is non-positive by the submodularity of $g(x + d_k, \cdot)$. If also $\mu_j > 0$, the third term is non-positive as in the first case. This completes the proof of the lemma.

By a similar induction to that used for the cycle-of-queues result, and application of the lemma with $D_i = \{d_i\}$ and $D_j = \{d_j\}$, $j \neq i$, we conclude that the theorem of this section is true for the general model. We need only the additional assumption that $h(x, D)$ is submodular on $\boldsymbol{D}(x)$ for all $x \in X$. The theorem is that when the state moves from x to $x + d_i$ the optimal service rates μ_j, $j \neq i$, do not decrease.

Recall that, for the cycle-of-queues model, there were uncontrolled arrivals at rate λ_k to queue k. In the general model, the inductive proof also goes through when the model includes some uncontrolled transformations. Suppose $a_k \in \mathbb{Z}^m$ represents a transformation of the state, $x \to x + a_k \in X$. This occurs at rate λ_k whenever $x + a_k \in X$ and at rate 0 otherwise. We say that this uncontrolled transformation a_k is $\boldsymbol{D}$*-independent* if $x + a_k \in X$ implies $x + a_k + D \in X$ for all $D \in \boldsymbol{D}$. The inductive proof of the theorem uses a similar argument to that which we used in the proof of the theorem for the cycle-of-queues model when we considered the terms multiplied by λ_k. Thus the theorem is true for a general queueing network if all the conditions described thus far hold, and any uncontrolled transformations are $\boldsymbol{D}$-independent. The arrival transformation, e_i, is always $\boldsymbol{D}$-independent if $X = \{x : x \geqq 0\}$, but there can be other interesting $\boldsymbol{D}$-independent transformations. Note that the theorem makes a statement about how the optimal $\mu_i(x)$

changes, as $x \to x + d_j$, $j \neq i$. It does not say anything about how $\mu_i(x)$ changes after some uncontrolled transformation, $x \to x + a_k$. For example, in the cycle-of-queues model, it is an open problem as to how the $\mu_i(x)$'s change as $x \to x + e_i$.

We shall now give some applications of the theorem to other systems. For example, a model of a series of queues can be constructed from the model for a cycle of queues by redefining $d_m = -e_m$ and adding a new transformation $d_0 = e_1$ with a corresponding c_0. (Alternatively, a series of m queues can be constructed from a cycle of $m+1$ queues, $0, 1, \ldots, m$, by starting queue 0 with an infinite number of customers and taking $h_0 = 0$.) The theorem implies that upon the movement of a customer from queue i to $i+1$ the service rate at queue i is not increased and the service rates elsewhere are not decreased. A second result describes the effect of adding a customer to some queue: that upon the addition of a customer to queue i the service rate at queue i and the service rates downstream of queue i are not decreased and the arrival rate to queue 1 and service rates upstream of queue i are not increased. This conclusion comes from noting that as $x \to x + d_0 \to x + d_0 + d_1 \to \cdots \to x + d_0 + d_1 + \cdots + d_{i-1} = x + e_i$ the service rate at queue j, $j \geqq i$, does not decrease, and as $x \to x + d_i \to x + d_i + d_{i+1} \to \cdots \to x + d_i + d_{i+1} + \cdots + d_m = x - e_i$ the service rate at queue j, $j < i$, does not decrease.

In another example, suppose queue $m+1$ is a queue with no holding costs, which is fed by a parallel system of queues: $1, \ldots, m$. Define $d_i = (e_{m+1} - e_i)$, for $i = 1, 2, \cdots, m$. When a customer leaves queue i, $i = 1, \cdots, m$, it enters queue $m+1$. For this model $\boldsymbol{D}$ is compatible with $X = \{x : x \geqq 0\}$. Take $h(x) = \sum h_{ij}(|x_i - x_j|) + h_i(x_i)$, the sum running over $i < j$, $i, j = 1, \cdots, m$, and h_{ij}, h_i increasing and convex. Then it is easy to check that $h(x, D) = h(x + D)$ is submodular on $\boldsymbol{D}(x)$, and the problem which is posed is one of minimizing the sum of service costs, ordinary holding costs and holding costs which penalise imbalances amongst the sizes of the queues. The theorem gives the intuitive plausible result that upon completion of a service in queue i the optimal service rate should not decrease at any other queue.

Notice finally, that if deficits of customers are allowed, so that x_i may be negative and $Z = \mathbb{Z}^m$, then any $\boldsymbol{D}$ is compatible with X. We only require that $h(x, D)$ be submodular on $\boldsymbol{D}$ for each x. If this is the case transition-monotonicity holds for all possible transitions.

3. Transition-monotonicity for average-cost optimal policies

For the cycle of queues in Section 2 the time-average cost is finite only if all $\lambda_i = 0$. If this is the case, the system is closed and a finite number of customers unendingly move around the cycle. Since the state space is finite we can use well-known arguments to let α tend to 0 and deduce the transition-monotone structure of the control policy for a time-average-cost criterion (see Ross (1970)).

When the state space is infinite, results concerning average-cost optimal policies are more difficult to establish since a direct argument based on letting α tend to 0 is no longer valid (depending as it does upon costs being uniformly bounded). Rosberg et al. (1982) have commented that the discount optimal policy (for two queues in series) may induce a transient Markov chain, in which the number of customers in first queue tends to ∞, even though an average-cost optimal policy induces an ergodic chain. However, this behaviour does not occur in the limit as α tends to 0, and we shall see that average-cost versions of the theorem can be deduced via those for discounted cost as α tends to 0. Rosberg et al., and also Hajek (1984), have analysed the average-cost problem via a limit of finite-horizon problems, but we have been unable to adapt their methods to the more general models of this paper. Other authors have dealt with the average-cost problem through the discounted-cost problem. Serfozo (1981) and Lu and Serfozo (1984) show that the limit of the discount optimal policies can be taken as $\alpha \to 0$ if the total number of possible policies is finite. This happens, for example, if each c_i is linear (in which case $\mu_i(x)$ is just 0 or $\bar{\mu}$) and one can argue by some means that $\mu_i(x) = \bar{\mu}$ if x_i lies outside a finite set where $h_i(x_i)$ is relatively low. However, this condition, like others given in the literature, can be difficult to verify.

Our treatment of the average-cost problem proceeds via a new argument for taking the limit of discounted-cost problems, when the state space is countably infinite and one-stage costs are unbounded. The conditions under which the validity of our limiting scheme applies are easy to check for the models of §2. For a general Markov decision problem they are as follows:

(a) The state space X is countable.

(b) The set of actions $A(x)$ which is available in state x is a compact metric space.

(c) The probability $P_a(x, y)$, of transition to state y when action a is taken in state x, is continuous in $a \in A(x)$,

(d) The one-stage cost $c_a(x)$, of taking action a in state x, is non-negative and continuous in $a \in A(x)$.

(e) It is possible to go from any state x to any other state y with finite expected cost.

(f) For each x there are only finitely many y for which $P_a(x, y) > 0$ for some $a \in A(x)$.

(g) If there is some policy which achieves a finite average cost, say γ^*, then the number of states in which the one-stage cost can be no more than γ^* is finite.

A similar result to ours, showing that an average-cost optimal policy can be found by taking the limit of discounted-cost optimal policies, has been proved by Schal (1977) under assumptions (a)–(d) and an additional assumption (h): that for each policy π, which takes action $\pi(x)$ in state x, the transition probability matrix P_π induces an irreducible and positive recurrent Markov chain whose stationary probabilities are continuous in π. Hordijk and Van Der Duyn Schouten (1983) have also proved this result under a condition similar to (h), but by using Abelian limit theorems, rather than the optimality equation. We are grateful to the last-named author for drawing attention to the relation between our work and that in these papers. For our models, (e)–(g) appear easier to verify than (h). Conditions (a)–(c) and (e)–(f) are obviously valid. Since each $c_i(\mu)$ is bounded below on $\mu \in [0, \bar{\mu}]$ we can simply increase each c_i by some constant amount to make $c_i(\mu) \geqq 0$. If we also assume that $h(x)$ is bounded below, we can without loss of generality assume $h(x) \geqq 0$, and so (d) holds. In order that (g) hold, we assume each $h_i(x_i)$ is convex increasing in $|x_i|$ for $|x_i|$ large enough (which is necessary anyway if the problem is to be interesting).

We describe the limiting scheme which applies under assumptions (a)–(g). Let $V_\alpha(x)$ denote the minimum expected total discounted cost over the infinite horizon when starting in state x and discounting at rate $\alpha > 0$. Suppose x_α is a most favourable starting state: that is, $V_\alpha(x_\alpha) \leqq V_\alpha(x)$ for all x. Then for all x

$$V_\alpha(x_\alpha) \leqq V_\alpha(x) \leqq M(x, x_\alpha) + V_\alpha(x_\alpha),$$

giving

$$0 \leqq V_\alpha(x) - V_\alpha(x_\alpha) \leqq M(x, x_\alpha), \tag{4}$$

where $M(x, x_\alpha)$ is the expected undiscounted cost of first passage from x to x_α under some policy which makes this cost finite (here using assumption (e)). We can write

$$\begin{aligned} &V_\alpha(x) - V_\alpha(x_\alpha) \\ &\qquad = -\alpha V_\alpha(x_\alpha) + \inf_a \Big[c_a(x) + (1-\alpha) \sum_y P_a(x, y)\{V_\alpha(y) - V_\alpha(x_\alpha)\} \Big]. \end{aligned} \tag{5}$$

We shall shortly let α tend to 0 in (5) after establishing that $V_\alpha(x) - V_\alpha(x_\alpha)$ converges to a limit as α tends to 0 in some sequence of discount rates. The convergence of $V_\alpha(x) - V_\alpha(x_\alpha)$ in a sequence of discount rates which tend to 0 is assured if the upper bound, $M(x, x_\alpha)$, in (4) can be bounded above uniformly in α. We shall see that this is possible because x_α can take only finitely many values, $0 < \alpha < \infty$. Invoke assumption (f): suppose that there is some stationary policy π^* such that the average cost is finite. (For example, in the series of queues, π^* could be the full service policy which sets the controllable arrival rate at queue 1 to 0 and all service rates at queues to the maximum possible value, $\bar{\mu}$. Note that finite average cost under some π^* places certain restrictions on the parameters of the problem.) Let x_α^* be a most favourable starting state when the discount rate is α and policy π^* is to be followed over the infinite horizon. Suppose that the undiscounted time average cost under policy π^* has the finite value γ^*. Then

$$V_\alpha^*(x) = c^*(x) + (1-\alpha)\sum_y P^*(x, y) V_\alpha^*(y), \tag{6}$$

where starting in state x and using policy π^*, $V_\alpha^*(x)$, $c^*(x)$ and $P^*(x, y)$ are respectively the expected total discounted cost, one-stage cost and transition matrix. Observe that $V_\alpha^*(x_\alpha^*)$ is no more than γ^*/α. This follows from the fact that $\alpha V_\alpha^*(x_\alpha^*)$ can be interpreted as the minimal average cost in an undiscounted setting when opportunities to restart the system occur randomly according to a Poisson process with rate α. It is not hard to see that if opportunities to restart the system are so introduced than an optimal average-cost policy will avail itself of these opportunities and on each occasion restart the system in state x_α^*. The resulting average cost is then, by a renewal-reward calculation, $\alpha V_\alpha^*(x_\alpha^*)$; moreover, it is clear that $\alpha V_\alpha^*(x_\alpha^*)$ is non-decreasing as α tends to 0 and that it is no more than the average cost γ^* achieved by the policy π^* which does not take advantage of opportunities to restart. So we find

$$\begin{aligned} V_\alpha(x_\alpha) &= \inf_a \left\{ c_a(x_\alpha) + (1-\alpha)\sum_y P_a(x, y) V_\alpha(y) \right\} \\ &\geqq \min_a \{c_a(x_\alpha)\} + (1-\alpha) V_\alpha(x_\alpha). \end{aligned}$$

$$\gamma^*/\alpha \geqq V_\alpha^*(x_\alpha^*) \geqq V_\alpha(x_\alpha^*) \geqq V_\alpha(x_\alpha) \geqq \min_a \{c_a(x_\alpha)\}/\alpha. \tag{7}$$

From (7) we have $\min \{c_a(x_\alpha)\} \leqq \gamma^*$ for all α and thus x_α must be a member of the finite set of states, S, in which one-stage costs can be no more than γ^* (invoking assumption (f)). This implies that $M(x, x_\alpha)$ can be bounded above uniformly in α by the maximum of $M(x, y)$ over $y \in S$. This is sufficient for $V_\alpha(x) - V_\alpha(x_\alpha)$ to have a limit $\phi(x)$ in some sequence of discount rates which tend to 0. By a diagonalization argument we can ensure that $V_\alpha(x) - V_\alpha(x_\alpha)$ tends to a limit for all x in a sequence of discount rates which tend to 0. Notice that for each $\alpha > 0$, $V_\alpha(x, D) - V_\alpha(x_\alpha, D)$ inherits submodularity from $V_\alpha(x, D)$ and will therefore in the limit be submodular on $\boldsymbol{D}(x)$. Also (4) implies $\phi(x) \geqq 0$ for all x. Taking limits in (5) we have

$$\phi(x) = -\gamma + \inf_a \left\{ c_a(x) + \sum_y P_a(x, y)\phi(y) \right\}, \tag{8}$$

where γ is the limit of $\alpha V_\alpha(x_\alpha)$ in the sequence of discount rates tending to 0. In deriving the right-hand side of (8), we must bring the limiting operation inside the infimum. It is easy to check that conditions (c), (d) and (g) are sufficient for this step to be valid. The infimum is achieved by a policy $\bar{\pi}$ for which $\bar{\pi}(x)$ is an action achieving the infimum in (8). In our models, the rate μ_i is chosen by $\bar{\pi}$ to minimise $c_i(\mu) + \mu\phi(x + d_i) + (\bar{\mu} - \mu)\phi(x)$.

It remains to show that γ is the minimum average cost. Observe first that $\alpha V_\alpha(x_\alpha)$ is not greater than the minimal average cost. This follows by the same argument used above: $\alpha V_\alpha(x_\alpha)$ can be interpreted as the minimal average cost in an undiscounted setting when opportunities to restart the system occur randomly according to a Poisson process with rate α. Moreover, $\alpha V_\alpha(x_\alpha)$ tends monotonically to γ as α tends to 0, and therefore γ is no greater than the minimal average cost.

Let X_t, $t = 1, 2, 3, \cdots$, denote the state of the Markov chain induced by $\bar{\pi}$ after t stages when starting from state x. From (8) we can deduce

$$\phi(x) = E_{\bar{\pi}}\left[\sum_{t=0}^{n-1} c_{\bar{\pi}}(X_t)\right] - n\gamma + E_{\bar{\pi}}[\phi(X_n)].$$

So

$$\begin{aligned} E_{\bar{\pi}}\left[\sum_{t=0}^{n-1} c_{\bar{\pi}}(X_t)\right]\Big/ n &= \gamma + \phi(x)/n - E_{\bar{\pi}}[\phi(X_n)]/n \\ &\leqq \gamma + \phi(x)/n. \end{aligned}$$

The last inequality above follows from the fact that $\phi(x) \geqq 0$ for all x. Thus taking the lim sup in the above as n tends to ∞ we deduce that policy $\bar{\pi}$ has average cost no more than γ and it is therefore an average-cost-optimal policy. For the models of Section 2, the submodularity of $\phi(x)$ implies that $\bar{\pi}$ has transition-monotone structure.

The proof of the theorem required the assumption that holding cost in each queue be convex. In another paper, Stidham and Weber (1987), we consider control of the arrival and/or service rate at a single queue when the holding cost is a non-decreasing, but not necessarily convex, function of the number of customers in the queue. It turns out that the transition-monotone structure of the optimal policy is lost, in general, if the criterion is expected total discounted cost, but retained if the criterion is average cost. This raises the question as to whether the same might be true regarding the control of more complicated systems.

The following example shows that in the average-cost case the assumption of convex holding costs may not in general be relaxed if we are to guarantee that the optimal policy be transition-monotone. The example concerns a series of two queues for which the non-zero parameters are $\lambda_1 = 0{\cdot}001$, $\mu_1 \in [0, 1]$,

$\mu_2 \in [0, 0{\cdot}1]$, $c_1(\mu) = \mu$ and $c_2(\mu) = 0$. This means that the arrival rate is fixed at 0·001. The service rate at queue 1 can lie within $[0, 1]$ and the service rate at queue 2 will always be chosen to be 0·1. Let the non-decreasing, but non-convex holding costs be as follows:

$$h_1(x_1) = x_1, \quad x_1 = 0, 1, 2, \cdots, \quad h_2(x_2) = \begin{cases} 0, \\ 3, \end{cases} \quad x = \begin{cases} 0, 1 \\ 2, 3, 4, \cdots. \end{cases}$$

Computation shows that the minimal average cost is 0·0021 and it is optimal to serve queue 1 at rate 1 in states $(1, 0)$, $(1, 3)$, $(1, 4)$, $\cdots$ but at rate 0 in state $(1, 1)$ and $(1, 2)$. Thus the optimal control of the service rate at queue 1 is not monotone in the number of customers in queue 2.

4. Importance of the zero service rate option

A particular feature of our model has been the ability to control all service rates to 0. If for either the expected discounted or time average-cost cases this option is not available then there may not be a transition-monotone optimal policy. This fact may seem surprising, but an example shows that it is the case. Consider the series system $M/M/1 \rightarrow /M/1$ and the average-cost criterion. The transformations are $d_0 = e_1$, $d_1 = e_2 - e_1$ and $d_2 = -e_2$. Suppose the service rate at queue 1 is not controllable to 0. The rate at queue 1 is fixed at 1 and the rate at queue 2 is controllable in $[0, 2]$, but with $c_1(1) = c_2(\mu) = 0$. Suppose the arrival rate μ_0 can be controlled between 0 and 0.01, and that the reward for admitting arrivals is reflected in a cost $c_0(\mu_0) = -0{\cdot}944\mu_0$. Let

$$h_1(x_1) = \begin{cases} 0, \\ \infty, \end{cases} \quad x_1 = \begin{cases} 0, 1, 2 \\ 3, 4, 5, \cdots \end{cases}, \quad h_2(x_2) = \begin{cases} x_2, \\ \infty, \end{cases} \quad x_2 = \begin{cases} 0, 1, 2, 3, 4 \\ 5, 6, 7, \cdots. \end{cases}$$

The infinite holding costs are assumed to make calculations easy (since there can then be at most 12 recurrent states if the cost is to be finite when the system starts in state 0), but these could be replaced by any large values with the h_i's convex increasing. With the linear reward $0{\cdot}944\mu_0$ for admitting arrivals, the optimal value of μ_0 will be at endpoints of the interval $[0, 0{\cdot}01]$. At queue 2 customers will always be served with $\mu_2 = 2$. Computation shows that the minimal average cost is $-0{\cdot}00441470$ and the uniquely optimal policy admits arrivals only in the states $(0, 0)$, $(1, 0)$, $(0, 1)$, $(1, 1)$ and $(1, 2)$. Since the optimal arrival rate is set at 0 in state $(0, 2)$ and at 0.01 in state $(1, 2)$ the optimal arrival rate is not non-increasing in x_1 and the optimal policy is not transition-monotone. The reason for this surprising behaviour is that the service rate at queue 1 may not be set to 0. It turns out that $\phi(x, D)$ is not submodular. We find

$$\phi((2, 2)) - \phi((1, 2)) = 0{\cdot}8625 < 0{\cdot}944 < 0{\cdot}9470 = \phi((1, 2)) - \phi((0, 2)).$$

Whereas, taking $x=(1,2)$, $D_i=\{d_0\}$ and $D_j=\{d_1, d_2\}$, submodularity would require the inequality in the opposite direction. Note, that if $\boldsymbol{D}$ is restricted to subsets of $\{d_0, d_2\}$, then d_1 is a $\boldsymbol{D}$-independent uncontrollable transformation. The theorem now does apply and we can deduce that μ_0 is non-increasing in x_2, as the calculations confirm.

Fixed rates can be consistent with a transition-monotone structured optimal policy provided the holding costs are sensibly arranged. For example, in the series-of-queues model, the rate in queue i can be fixed at μ_i and monotone structure retained provided it is always advantageous to move a customer from queue i to $i+1$. This is the case if, for example, $h_i(x_i)=w_ix_i$, $h_{i+1}(x_{i+1})=w_{i+1}x_{i+1}$ with $w_i \geqq w_{i+1}$. We just pretend that the service rate at queue i is controllable within $[0, \mu_i]$, with $c_i(\mu)=0$ for all μ in this interval, and observe that it must be worthwhile to set the service rate to the highest rate μ_i since service completions move a customer to a less expensive position at no cost. Thus while the availability of 0 service rates is sufficient for there to be a transition-monotone optimal policy, it is not necessary.

References

BERTSEKAS, D. (1976) *Dynamic Programming and Stochastic Control.* Academic Press, New York.

CRABILL, T. (1972) Optimal control of a service facility with variable exponential service time and constant animal rate. *Management Sci.* **18,** 560–566.

→ CRABILL, T., GROSS, D., AND MAGAZINE, M. (1977) A classified bibliography of research on optimal design and control of queues. *Operat. Res.* **25,** 219–232.

DAVIS, E. (1977) Optimal Control of Arrivals to a Two-Server Queueing System with Separate Queues. Ph.D. Dissertation, Program in Operations Research, North Carolina State University, Raleigh.

EPHREMIDES, A., VARAIYA, P., AND WALRAND, J. (1980) A simple dynamic routing problem. *IEEE Trans. Autom. Control* **25,** 690–693.

GHONEIM, H. (1980) Optimal Control of Arrivals to a Network of Two Queues in Series. Ph.D dissertation, Program in Operations Research, North Carolina State University, Raleigh.

HAJEK, B. (1984) Optimal control of two interacting service stations. *IEEE Trans. Autom. Control* **29,** 491–499.

HORDIJK, A. AND VAN DER DUYN SCHOUTEN, F. A. (1983) Average optimal policies in Markov decision drift processes with applications to a queueing and replacement model. *Adv. Appl. Prob.* **15,** 274–303.

JOHANSEN, S. AND STIDHAM, S. (1980) Control of arrivals to a stochastic input output system. *Adv. Appl. Prob.* **12,** 972–999.

JO, K. Y. AND STIDHAM, S. (1983) Optimal service-rate control of M/G/1 queueing systems using phase methods. *Adv. Appl. Prob.* **15,** 616–637.

→ LIPPMAN, S. (1975) Applying a new device in the optimization of exponential queueing systems. *Operat. Res.* **23,** 687–710.

→ LIPPMAN, S. AND STIDHAM, S. (1977) Individual versus social optimization in exponential congestion systems. *Operat. Res.* **25,** 233–247.

→ LU, F. V. AND SEFOZO, R. F. (1984) $M/M/1$ queueing decision processes with monotone hysteritic optimal policies. *Operat. Res.* **32,** 1116–1132.

ROSBERG, Z., VARAIYA, P., AND WALRAND, J. (1982) Optimal control of service in tandem queues. *IEEE Trans. Autom. Control* **27,** 600–610.

ROSS, S. (1970) *Applied Probability Models with Optimization Applications.* Holden-Day, San Francisco.

SCHAL, M. (1977) On negative dynamic programming with irreducible Markov chains and the average cost criterion. *Bonner Math. Schr.* **98,** 93–97.

SERFOZO, R. (1979) An equivalence between discrete and continuous time Markov decision processes. *Operat. Res.* **27,** 616–620.

SERFOZO, R. F. (1981) Optimal control of random walks, birth and death processes, and queues. *Adv. Appl. Prob.* **13,** 61–83.

SOBEL, M. (1974) Optimal operation of queues. In *Mathematical Methods in Queueing Theory,* ed. A. B. Clarke, Springer-Verlag, Berlin, 231–261.

STIDHAM, S. (1985) Optimal control of admission to a queueing system. *IEEE Trans. Autom. Control.* To appear.

STIDHAM, S. AND PRABHU, N. (1974) Optimal control of queueing systems. In *Mathematical Methods in Queueing Theory,* ed. A. B. Clarke, Springer-Verlag, Berlin, 263–294.

STIDHAM, S. AND WEBER, R. (1987) Monotonic and insensitive optimal policies for control of arrival and service rates in queues. *Operat. Res.* To appear.

TOPKIS, D. (1978) Minimizing a submodular function on a lattice. *Operat. Res.* **26,** 305–321.

WEBER, R. (1978) On the optimal assignment of customers to parallel servers. *J. Appl. Prob.* **15,** 406–413.

WHITTLE, P. (1983) *Optimization over Time.* Wiley, Chichester.

WINSTON, W. (1977) Optimality of the shortest line discipline. *J. Appl. Prob.* **14,** 181–189.