A survey of Markov decision models for control of networks of queues

Shaler Stidham Jr.[1]

Department of Operations Research, CB 3180, Smith Building, University of North Carolina, Chapel Hill, NC 27599-3180, USA

Richard Weber

University of Cambridge, Judge Institute of Management Studies, Mill Lane, Cambridge, CB2 1RX, England

Received 17 January 1992; revised 23 November 1992

We review models for the optimal control of networks of queues. Our main emphasis is on models based on Markov decision theory and the characterization of the structure of optimal control policies.

Keywords: Networks of queues, optimal control, Markov decision process.

1. Introduction

A network of queues can serve as a useful model for systems arising in many contexts, including manufacturing, distributed computer systems, voice and data communications, and vehicular traffic flow. Descriptive models can help to evaluate and predict the performance of existing and proposed systems, and thus improve the design of a system. Control models take into account the possibility of dynamically varying some of the parameters of the system, such as the arrival or service rates, or the rules used to determine the routing of jobs or the order in which they are processed at various nodes. In principle, dynamic (that is, state-dependent) rules for setting such parameters offer the possibility of significantly improved performance, as compared to static (state-independent) rules. The performance

[1] This research was partially supported by the National Science Foundation under Grant No. DDM-8719825. The Government has certain rights in this material. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The research was also partially supported by the C.I.E.S. (France), while the author was on leave at INRIA, Sophia-Antipolis, 1991–92.

improvement can take the form of reduced congestion, as measured by the numbers of jobs in the buffers at the various nodes or by the time spent by jobs waiting to be served, for a given load (throughput). Alternatively, it can be reflected in increased throughput for a given level of congestion.

The importance of dynamic control has received increasing recognition in recent years in a variety of applications involving large-scale systems of interacting components. At the same time, the introduction of automation has made sophisticated control rules feasible to implement for the first time. Computer-controlled manufacturing facilities offer an example, as do communication systems with automated, digital switches.

As an illustration of how optimal control models can be used in applications of networks of queues, consider a data communication network. Whether operating under packet switching or virtual-circuit switching, such a network can often be modelled as a network of queues. The messages or packets generated by a session are the customers. Each channel in a link between an adjacent pair of nodes of the communication network is a server. The buffer at a node containing the packets waiting for transmission on a particular link (consisting of one or more channels) constitutes a queue. For any given routing and flow-control policy, the resulting network of queues can be analyzed with the help of analytical, numerical, and/or simulation models. Moreover, the problem of determining a routing and flow control policy that is optimal (with respect to performance measures such as throughput and/or congestion) can sometimes be formulated as a Markov decision process. The techniques of dynamic programming can then (in principle) be used to characterize the structure of an optimal policy and/or numerically solve for its parameters. For communication networks of realistic size, solving for optimal control policies in this way is typically computationally intractable. Examining the structure of optimal policies for smaller networks, however, can lead to insights that are useful for the development of heuristic rules for larger networks. (For a review of both static and dynamic models specifically relevant to control of routing in data communication networks, see Hariharan et al. [32].)

This paper summarises a large number of models and results for the control of networks of queues. We shall focus our attention on the use of Markov decision models to examine the structure of optimal control rules. We shall not discuss computational issues, nor shall we consider, except in passing, the use of other techniques, such as sample-path analysis, for determining the structure of optimal policies. We assume that the reader is familiar with the basic theory of Markov decision processes, as contained for example in Bertsekas [6], Whittle [72,73], or Walrand [65]. The last-named reference is an excellent introduction to the general topic of networks of queues, including descriptive as well as control models.

We shall concentrate on discounted-cost problems. Often the structural properties of optimal control policies for an average-cost problem are the same as those for the corresponding discounted-cost problem. Techniques for deriving the former from the latter are now well developed. Recent results specifically motivated by control of queues may be found in Borkar [8–10], Weber and Stidham [67], Cavazos-Cadena [12,13], Sennott [54,55]. For a survey, see Arapostathis et al. [2].

The paper starts in section 2, with a description of a general model for control of transition rates in Markov processes and a proof of a characterization of the optimal control. In the remainder of the paper we review of a large number of models and results. Our review is organised by categories of control: namely, service rates (section 3), admission (section 4), routing (section 5), and scheduling (section 6). Although most results are reviewed without proofs, there is in each section at least one result for which the proof is given: in sections 3, 4, and 5 this is by application of the general ideas introduced in section 2; in section 6 there is a solution for a scheduling problem in a series of queues.

2. Control of transitions in a Markov process: A general model

In this section we present a general model for control of the transition rates in a continuous-time Markov chain. This model is essentially the same as that developed by Weber and Stidham [67] and applied there to control of service rates in cycles and series of queues. We also mention recent extensions by Veatch and Wein [62] and Glasserman and Yao [28].

We consider a continuous-time Markov decision process with countable state space $X \subseteq Z^m$. That is, the state is an m-vector of non-negative integers, $x = (x_1, \ldots, x_m)$. In the context of networks of queues, x_l typically denotes the number of customers at node (queue) l. There are two varieties of transition, *controlled* and *uncontrolled*. Uncontrolled transitions of type r $(r = 1, \ldots, p)$ are defined in terms of an operator A_r: $X \to X$. They occur at fixed rate λ_r and move the system from state $x \in X$ to state $A_r x \in X$. Controlled transitions of type i $(i = 1, \ldots, q)$ are defined in terms of an operator T_i: $X \to Z^m$. They move the system from state x to $T_i x$ and occur at rate μ_i, which is subject to control, with $\mu_i \in [0, \bar{\mu}_i]$, $0 < \bar{\mu}_i < \infty$. It is understood that the system controller must select rate $\mu_i = 0$ whenever the system is in a state x such that $T_i x \notin X$. (Infeasible controlled transitions are forbidden.)

An application that we shall come to in the next section is a network consisting of m infinite-capacity queues arranged in a cycle, with a single exponential server at each node i, performing services at controllable rate μ_i (cf. Weber and Stidham [67]). Each node i receives external input of customers according to an uncontrolled Poisson arrival process with rate λ_i, $i = 1, \ldots, m$. With state variable $x = (x_1, \ldots, x_m)$, $A_i x := x + e_i$, and $T_i x := x - e_i + e_{i+1}$, $i = 1, \ldots, m$, (and identifying $m + 1$ as 1), this is an example within the general framework above.

The cost rate is a combination of transition-rate costs and holding costs. While the rate for type-i controllable transitions is μ_i, a cost is incurred at rate $c_i(\mu_i)$ per unit time, where c_i is continuous $(i = 1, \ldots, q)$. While in state x the system incurs a (holding) cost at rate $h(x)$, where h is non-negative. Future costs are continuously discounted at rate $\alpha \geqslant 0$ and the objective of the system controller is to minimize the expected total α-discounted cost over an infinite horizon.

Let $V(x)$ denote the minimum total expected α-discounted cost over an infinite horizon, starting from state x. Then V satisfies the dynamic-programming optimality equation

$$V(x) = h(x) + \sum_{r=1}^{p} \lambda_r V(A_r x) + \sum_{i=1}^{q} \min_{\mu_i \in [0,\bar{\mu}_i]} \{c_i(\mu_i) + \mu_i V(T_i x) + (\bar{\mu}_i - \mu_i) V(x)\}, \tag{1}$$

where we have assumed that $\alpha + \sum_{r=1}^{p} \lambda_r + \sum_{i=1}^{q} \bar{\mu}_i = 1$, without loss of generality. Moreover, an optimal control rule chooses, whenever the system is in state x, a rate $\mu_i(x)$ for type-i transitions that achieves the minimum in the ith term of the second summation in (1), where again it is understood that the minimization operator selects $\mu_i(x) = 0$ if $T_i x \notin X$. The arguments for the validity of this optimality equation and the optimality of the (stationary, deterministic) control rule, $\mu(x) := (\mu_1(x), \ldots, \mu_q(x))$, are based on the principle of optimality and special properties of the exponential distribution. For detailed discussion of this optimality equation and those for other problems in the control of queues, see Weber and Stidham [67], Stidham [57,58].

We can rewrite eq. (1) in the equivalent form

$$V(x) = h(x) + \sum_{r=1}^{p} \lambda_r V(A_r x) + \sum_{i=1}^{q} \bar{\mu}_i V(x) + \sum_{i=1}^{q} \min_{\mu_i \in [0,\bar{\mu}_i]} \{c_i(\mu_i) - \mu_i[V(x) - V(T_i x)]\}. \tag{2}$$

The quantity $V(x) - V(T_i x)$ can be interpreted as the benefit of a transition of type i that occurs in state x. Thus we should be willing to choose a faster rate for this type of transition as this benefit increases.

The model of Weber and Stidham [67] provides a general framework in which an optimal control rule for this problem is *transition monotone*, that is,

$$\mu_i(x) \leqslant \mu_i(T_j x) \text{ for all } x \in X \text{ such that } T_j x \in X, \quad \text{for all } j \neq i.$$

In words, a control rule exhibits transition monotonicity if

the rate for transitions of type i does not decrease as a transition type j occurs for all $j \neq i$.

One can establish this property by showing that the benefit, $V(x) - V(T_i x)$, of a type-i transition in state x is no greater than the corresponding benefit, $V(T_j x) - V(T_i T_j x)$, in state $T_j x$, that is, V satisfies the functional inequality

$$f(x) - f(T_i x) - f(T_j x) + f(T_i T_j x) \leqslant 0, \tag{3}$$

for all $j \neq i$ and all $x \in X$ such that $T_i x \in X$, $T_j x \in X$, and $T_i T_j x \in X$.

We say that a function f satisfying (3) is (X, T)-*multimodular*, following Hajek [30] who introduced a special case of the concept in a different context. To prove that V is (X, T)-multimodular requires additional conditions on the holding cost $h(x)$ and on the relationship between the state space X and the sets of transition operators $A := \{A_r, r = 1, \ldots, p\}$ and $T := \{T_i, i = 1, \ldots, q\}$. In particular, we shall need the following definition.

DEFINITION

The set T of controlled transition operators is said to be *compatible* with the state space X if for all $i \neq j$ and all $x \in Z^m$ such that $T_i x \in X$ and $T_j x \in X$, the states x and $T_i T_j x$ are also in X.

THEOREM 1

Suppose the set T of controlled transition operators is commutative and compatible with X. Suppose also that A_r commutes with T_i for all $r = 1, \ldots, p$ and $i = 1, \ldots, q$. If $h(x)$ is (X, T)-multimodular, then the optimal value function V is also (X, T)-multimodular and an optimal control policy is transition monotone.

Remark

Weber and Stidham [67] actually reformulate and prove the transition-monotonicity property in terms of submodularity (Topkis [60]) of the optimal value function V with respect to a partial ordering based on set inclusion, but the proof sketched below is essentially equivalent and perhaps somewhat easier to follow.

The proof of theorem 1 uses induction on a sequence of finite-horizon problems. The key step is to show that (X, T)-multimodularity is preserved by transformations of the form

$$f_k(x) = \min_{\mu_k \in [0, \bar{\mu}_k]} \{c_k(\mu_k) + \mu_k g(T_k x) + (\bar{\mu}_k - \mu_k) g(x)\}, \tag{4}$$

when T is commutative and compatible with X. To prove this result, assume that $g: X \to R$ is (X, T)-multimodular, let $x \in X$, $T_i x \in X$, $T_j x \in X$, $T_i T_j x \in X$, $j \neq i$, and let

$$\Delta := f_k(x) - f_k(T_i x) - f_k(T_j x) + f(T_i T_j x).$$

Our goal is to show that $\Delta \leqslant 0$. To this end let μ_k and μ_k' achieve the minimum in (4) with x replaced by $T_i x$ and $T_j x$, respectively. Note that Δ is symmetric in i and j, since the operators T_i and T_j commute. Hence, without loss of generality we can assume that $\mu_k' \geqslant \mu_k$. (Otherwise reverse the roles of i and j.) We distinguish two cases $k \neq i$ and $k = i$.

Case 1: $k \neq i$. First we observe that μ_k is a feasible control in state x. If $\mu_k = 0$, this is trivially true. On the other hand, if $\mu_k > 0$ then $\mu_k' > 0$ and hence both $T_k T_i x$ and $T_k T_j x$ are in X, from which it follows by commutativity and compatibility that

$T_k x \in X$. Hence μ_k is feasible for state x. Similarly, μ'_k is feasible in state $T_i T_j x$. Again, if $\mu'_k = 0$, this is trivially true. Suppose $\mu'_k > 0$. Since g is (X, T)-multimodular and $k \neq i$, the minimizing control in state $T_i T_j x$ is at least as great as that in state $T_j x$, namely μ'_k. Hence the minimizing control in state $T_i T_j x$ is positive, which implies that $T_k T_i T_j x \in X$, and hence μ'_k is feasible for $T_i T_j x$.

Thus we can use control μ_k in state x and control μ'_k in state $T_i T_j x$ to get the following upper bound on Δ.

$$\begin{aligned}
\Delta \leqslant\; & c_k(\mu_k) + \mu_k g(T_k x) + (\bar{\mu}_k - \mu_k) g(x) \\
& - c_k(\mu_k) - \mu_k g(T_k T_i x) - (\bar{\mu}_k - \mu_k) g(T_i x) \\
& - c_k(\mu'_k) - \mu'_k g(T_k T_j x) - (\bar{\mu}_k - \mu'_k) g(T_j x) \\
& + c_k(\mu'_k) + \mu'_k g(T_k T_i T_j x) + (\bar{\mu}_k - \mu'_k) g(T_i T_j x) \\
=\; & (\bar{\mu}_k - \mu_k)[g(x) - g(T_i x) - g(T_j x) + g(T_i T_j x)] \\
& + (\mu'_k - \mu_k)[g(T_j x) - g(T_i T_j x) - g(T_k T_j x) + g(T_k T_i T_j x)] \\
& + \mu_k [g(T_k x) - g(T_k T_i x) - g(T_k T_j x) + g(T_k T_i T_j x)] \\
=\; & (\bar{\mu}_k - \mu_k)[g(x) - g(T_i x) - g(T_j x) + g(T_i T_j x)] \\
& + (\mu'_k - \mu_k)[g(T_j x) - g(T_i T_j x) - g(T_k T_j x) + g(T_k T_i T_j x)] \\
& + \mu_k [g(T_k x) - g(T_i T_k x) - g(T_j T_k x) + g(T_i T_j T_k x)]\,,
\end{aligned}$$

where we have used commutativity to derive the second equality. Since $k \neq i \neq j$, each of the last three terms is non-positive, by the assumption that g is (X, T)-multimodular. Thus $\Delta \leqslant 0$.

Case 2: $k = i$. First note that μ'_k is a feasible control for state x, since $T_k x = T_i x \in X$ by assumption. Also, μ_k is feasible in state $T_i T_j x$. If $\mu_k = 0$, this is trivially true. If $\mu_k > 0$, then $T_k T_k x = T_k T_i x \in X$. Moreover, $\mu'_k \geqslant \mu_k > 0$, so that $T_j T_k x = T_k T_j x \in X$ (using commutativity). Then, since $k = i \neq j$, it follows from commutativity and compatibility that $T_k T_i T_j x = T_k T_k T_j x = T_k T_j T_k x \in X$, and hence μ_k is feasible in state $T_i T_j x$.

Thus we can use control μ'_k in state x and control μ_k in state $T_i T_j x$ to get the following upper bound on Δ.

$$\begin{aligned}
\Delta \leqslant\; & c_k(\mu'_k) + \mu'_k g(T_k x) + (\bar{\mu}_k - \mu'_k) g(x) \\
& - c_k(\mu_k) - \mu_k g(T_k T_i x) - (\bar{\mu}_k - \mu_k) g(T_i x) \\
& - c_k(\mu'_k) - \mu'_k g(T_k T_j x) - (\bar{\mu}_k - \mu'_k) g(T_j x) \\
& + c_k(\mu_k) + \mu_k g(T_k T_i T_j x) + (\bar{\mu}_k - \mu_k) g(T_i T_j x) \\
=\; & (\bar{\mu}_k - \mu'_k)[g(x) - g(T_k x) - g(T_j x) + g(T_k T_j x)] \\
& + \mu_k [g(T_k x) - g(T_k T_k x) - g(T_k T_j x) + g(T_k T_k T_j x)] \\
=\; & (\bar{\mu}_k - \mu'_k)[g(x) - g(T_k x) - g(T_j x) + g(T_k T_j x)] \\
& + \mu_k [g(T_k x) - g(T_k T_k x) - g(T_j T_k x) + g(T_i T_k T_k x)]\,,
\end{aligned}$$

where we have used the fact that $k = i$ to derive the first equality and commutativity to derive the second equality. Since $k = i \neq j$, each of the last two terms is non-positive, by the assumption that g is (X, T)-multimodular. Thus $\Delta \leqslant 0$.

This completes the proof that (X, T)-multimodularity is preserved by transformations of the form (4).

Remark

Veatch and Wein [62] have extended these results to prove transition monotonicity with respect to uncontrolled as well as controlled transitions and to the case where a controlled transition of type i may include a routing decision – specifically, a choice between moving from state x to state $T_{i1}x$ at rate μ_i and to state $T_{i0}x$ at rate $\bar{\mu}_i - \mu_i$. Glasserman and Yao [28] have developed a general model for controlled discrete-state Markov processes in the framework of a generalized semi-Markov process (GSMP). They show that an optimal control policy is monotonic, under structural conditions of *non-interruption* and *strong permutability* of events and submodularity and supermodularity conditions on the one-step cost function. These conditions are in the same spirit as those in Weber and Stidham [67], as is the method of analysis, which uses ideas from Topkis [60] and induction on a sequence of finite-horizon problems.

3. Control of service rates in a network of queues

The model of the previous section can be applied to control of the service rates at the nodes of a network of queues, for networks with special structure. In each potential application, the crucial question is whether or not the compatibility and commutability conditions are satisfied. We shall illustrate these issues in the context of the two applications given prominence in Weber and Stidham [67]. Throughout, we assume that there is cost rate $h_j(x_j)$ per unit time while x_j jobs are at queue j, where h_j is convex (but not necessarily non-decreasing).

CYCLE OF QUEUES

Consider the cycle of m queues, labeled $i = 1, 2, \dots, m$, that was introduced in the previous section. A job that completes service at node (queue) i goes to node $i + 1$. (We identify node $m + 1$ as node 1.) Jobs from outside the system enter node i at mean rate λ_i according to a Poisson process which is not subject to control. There is a single exponential server at node i who performs potential services at mean rate μ_i, where μ_i is to be chosen from the feasible set $[0, \bar{\mu}_i]$. The number of jobs in node i is denoted by x_i and the state of the system by the vector $x = (x_1, \dots, x_m)$. While in state x, the system incurs a holding cost per unit time, $h(x) = \sum_{i=1}^{m} h_i(x_i)$. Future costs are continuously discounted at rate $\alpha > 0$ and the objective is to minimize the expected total discounted cost over an infinite horizon.

The two types of state transitions will be denoted $x \to A_i x := x + e_i$, corresponding to an arrival at node i, and $x \to T_i x := x - e_i + e_{i+1}$, corresponding to a service completion at node i and resulting transfer to node $i+1$. (Here e_i denotes the unit m-vector with a one in the ith component.)

It is easily verified that the conditions of commutativity and compatibility are satisfied for this problem with these operators. Transition monotonicity in this case implies that

the optimal service rate at node i does not decrease as a job is moved from node j to node $j+1$, $j \neq i$.

By moving a job all the way around the cycle, from node $i+1$ to $i+2$ to ... to $i-1$ to i, it follows that the optimal service rate at node i does not decrease as a customer is moved from node $i+1$ to node i, or, equivalently,

the optimal service rate at node i does not increase as a job is moved from node i to node $i+1$.

SERIES OF QUEUES

The results for a cycle of queues can be applied to a series of queues ($i = 1, 2, \ldots, m-1$), with a Poisson arrival process with controllable arrival rate $\lambda \in [0, \bar{\lambda}]$ at the first node and a reward earned at rate $r(\lambda)$ while the arrival rate is λ. This is done by adding to the series a dummy node m, with no holding cost and an infinite supply of jobs, and letting that node receive all output from node $m-1$ and generate all input to node 1. With $\mu_m := \lambda$, serving at rate μ_m at the dummy node corresponds to selecting arrival rate λ at the first node in the series. The service-cost rate $c_m(\mu_m)$ at the dummy node is the negative of the reward rate $r(\lambda)$.

The monotonicity results referred to above apply to the series of queues, so that the optimal service rate at node i does not decrease as a job is transferred from any node j to $j+1$, $j \neq i$, and does not increase as a job is transferred from node i to $i+1$. In addition, by moving a job from node $j > i$ to $j+1$ to $j+2$, and so forth, until it reaches the dummy node m representing the external world, one can deduce that the optimal service rate at node i does not decrease as a job is removed from a downstream node (or, equivalently, does not increase as a job is added to a downstream node). In the setting of a series-of-queues model for a production line, this corroborates the *pull* effect. Similarly, it corroborates the *push* effect, in that if a job is added to an upstream node, the optimal service rate at node i does not decrease.

OPTIMALITY OF BANG-BANG CONTROLS

Consider the special case in which the service-cost-rate function at node i, $c_i(\mu_i)$, has the following *economy-of-scale* property.

$$c(\mu_i)/\mu_i \geqslant c(\bar{\mu}_i)/\bar{\mu}_i\,, \quad \mu_i \in [0, \bar{\mu}_i]\,.$$

This property holds, for example, if $c(\cdot)$ is concave – in particular, linear. Then it can be shown (cf. Crabill [17]) that the optimal service rate at node i in any state x is either 0 or $\bar{\mu}_i$; that is, an extremal or *bang-bang* policy is optimal for node i.

KANBAN POLICIES

Now consider a control model in which the service rate μ_i at node i can assume only two values, 0 and $\bar{\mu}_i$, with associated cost rates, 0 and c_i, respectively. This corresponds to the case of a single server at node i with fixed rate $\bar{\mu}_i$, which can be turned on or off. We can apply the monotonicity results for our model to this situation by extending the range of feasible values of μ_i to the interval $[0, \bar{\mu}_i]$ and choosing as the service-cost-rate function, $c(\mu_i) := (\mu_i/\bar{\mu}_i)c_i$, that is, the linear interpolation between the points (0, 0) and $(\bar{\mu}_i, c_i)$. Then, since a bang-bang policy is optimal, the only values of μ_i that will be used in an optimal policy for our model are 0 and $\bar{\mu}_i$, so that an optimal policy will be feasible, and hence optimal, for the problem in which the server at node i is controlled by turning it on and off.

The monotonicity results for our model now imply that it will be optimal to turn an off server at node i on as the numbers of jobs at the downstream nodes decrease, or as the numbers of jobs at upstream nodes increase. In the context of a model for a production line, we have therefore a justification for use of a *kanban* policy to avoid starving and blocking of nodes.

SYSTEMS WITH FINITE BUFFERS AND BLOCKING

A series system with finite buffers between the successive nodes can also be accommodated by the model of Weber and Stidham [67]. This is done by modifying the (convex) holding-cost function $h_i(x_i)$ by adding a sufficiently large increment whenever the number of jobs at node i exceeds the buffer capacity so that an optimal policy will always set the service rate at the immediately preceding node to zero when the buffer is full. Thus, one may derive a generalization of the monotonicity results in Chao and Chen [15], who study two queues in series with a finite intermediate buffer. They also derive conditions under which it is optimal to serve at the maximal possible rate when the buffer is not full.

4. Control of admission to a series of queues

CONTROL OF ADMISSION TO THE FIRST QUEUE

The model of Weber and Stidham [67] can also be used to examine the structure of an optimal policy for control of arrivals to the first node of a series of queues. Recall that, in the application of [67] to a series of queues, the external Poisson arrival process at node 1 is modelled as a service process at node m, which represents the external world, thus reducing the series system to an equivalent cycle system.

The monotonicity results for the cycle-of-queues model then apply.

In particular, suppose customers arrive to node 1 at rate $\bar{\lambda}$ and an arriving customer can be *accepted*, earning a reward r, or *rejected* with no reward earned. This model is equivalent to the series-of-queues model of the previous section, with controllable arrival rate $\lambda \in [0, \bar{\lambda}]$ and $r(\lambda) = r \cdot \lambda$, since a bang-bang control is optimal for the latter problem. We therefore deduce that the benefit of accepting an arriving job does not decrease as another job is transferred from any node j in the series to node $j+1$, or (by combining a sequence of such moves) as a job is removed from any node j in the series. Thus an optimal admission-control policy will be more likely to accept if either of these two types of state change is made.

Although the present model implicitly assumes that the service rate at each of the nodes in series is also controllable, the results apply to a system with fixed service rates, as long as the marginal holding costs do not increase from node i to node $i+1, i = 1,2,\ldots,m-1$. For in this case the problem with fixed service rates is equivalent to one in which service rates are controllable but the service-cost rate function is identically zero at each node; in the latter problem it will always be optimal to serve at the maximal rate and hence accelerate the movement of a job to a cheaper node at no cost. Note that this ordering of the marginal holding costs implies that each holding cost function $h_i(x_i)$ is non-decreasing, since the marginal holding cost at node m is identically zero. See Weber and Stidham [67] for further discussion of this point. As noted in section 3, by introducing a holding cost that is sufficiently large when more than a certain number of jobs are present at a particular node, one can model nodes in series with finite intermediate buffers. (See also Chao and Chen [15].)

CONTROL OF ARRIVALS TO EACH OF TWO QUEUES IN SERIES

Ghoneim [25] (see also Ghoneim and Stidham [26]) study two exponential servers in series (with mean service rates μ_1 and μ_2), each with an infinite-capacity queue. Arrivals to queue j are from a Poisson process with mean rate $\lambda_j, j = 1,2$. Jobs arriving to queue 1 must go on to queue 2 after finishing service at server 1. Jobs arriving to queue 2 leave the system after finishing service at server 2. The model thus describes, for example, a simple communication system consisting of two channels in series with a combination of local and long-distance traffic.

With the reward and cost structure described at the start of this section, an induction based on value iteration establishes that the optimal value function is concave in each argument, submodular, and satisfies a third condition. The three conditions taken together constitute the analogue of multimodularity for maximization problems in two dimensions, and they imply that it is optimal to accept a job to node 1 (node 2) in a particular state, then it is still optimal if a job is removed from node 1 or node 2 or moved from node 1 to node 2 (from node 2 to node 1). These properties also rule out certain increasing sets as candidates for the optimal rejection region, namely those whose boundaries have horizontal segments of length greater than one.

Note that this model is not a special case of the series-of-queues model of Weber and Stidham [67] since the latter allowed control of arrivals to the first node only.

5. Control of admission, routing and server allocation in parallel queues

Davis [18] considers two exponential servers (with mean rates μ_1 and μ_2) in parallel, each with its own queue, and a renewal arrival process – that is, i.i.d. interarrival times distributed as a random variable T. The system controller may reject an arriving job, admit it to queue 1, or admit it to queue 2, based on the state $x = (x_1, x_2)$ at the instant of arrival, where x_j is the number of jobs at queue j (including the one in service, if any), $j = 1, 2$. Davis [18] considers the symmetric case: $\mu_1 = \mu_2, h_1(\cdot) \equiv h_2(\cdot)$. Abdel-Gawad [1] considers the general case.

An inductive proof based on value iteration shows that the optimal value function satisfies the same three properties (equivalent to multimodularity) as in the series-of-queues model discussed in the previous subsection. These properties imply that an optimal policy is *admission monotonic*: if it is optimal to reject in state x, then it is also optimal to reject in states $x + e_j, j = 1, 2$; in other words, the rejection region $R := \{x : a(x) = 0\}$ is an *increasing set*. Moreover, an optimal policy is *routing monotonic*: if admitting to queue 2 (queue 1) is preferable to admitting to queue 1 (queue 2) in state x, then it will remain so in state $x + e_1(x + e_2)$. In other words, an optimal routing policy is characterized by a monotonic "switching curve". Finally, an additional property of the rejection region is demonstrated by the induction: if it is optimal to admit a job to queue 1 (queue 2) in state x, then it is also optimal to do so in state $x - e_1 + e_2(x - e_2 + e_1)$.

Attempts to generalize these structural results to more than two queues in parallel have not met with much success, except in the symmetric case (see below), in which the service rates and holding costs at the different queues are the same. Beutler and Teneketsis [7] have shown that the monotonicity properties of optimal routing policies extend to partially observable queues, but they have also been unsuccessful in handling systems with more than two parallel nodes.

Hariharan et al. [31] have extended the monotonicity properties derived by Davis [18] and Hajek [29] to control of admission and routing to two parallel queues, each with an infinite number of servers serving at the same rate μ. The holding costs at the two queues may be different, however, in contrast to the models discussed next.

THE SYMMETRIC CASE: JOIN-THE-SHORTEST-QUEUE RULE

Symmetric queues are ones that are not distinguished by number, so that if the contents of two queues are exchanged the departure process and holding costs are unaffected. Control of routing to symmetric parallel queues has been studied by many authors, including Winston [74], Foschini [23], Weber [66], Foschini and Salz [24], Ephremides et al. [19], Lehtonen [45], Whitt [70], Houck [40], Menich and Serfozo [48], Johri [41], Hordijk and Koole [39], and Farrar [21]. All these authors assume that all customers are admitted and the only decision is where to route an arriving customer. In this case, the optimal routing policy is very often (but not always) the JSQ ("join-the-shortest-queue") rule, in which an arriving customer is routed to queue i if $x_i = \min_j\{x_j\}$.

Most of the cited references use sample-path arguments involving stochastic ordering and/or coupling and very often show that the JSQ rule is optimal in stronger senses than simply minimization of expected total discounted cost.

Menich and Serfozo [48] use induction on value iteration in a Markov-decision-process model to show that the optimality of JSQ routing extends to a network of parallel nodes with possibly state-dependent Poisson arrival or compound Poisson arrival process, and a memoryless service mechanism at each node. The service rate at each node can depend on the number of customers at that node and at other nodes. Similarly, the holding-cost rate need not be separable. But both the service rates and the holding cost must satisfy certain interchangeability assumptions. A special case is where the arrival process is state-independent and compound Poisson and the service rate at each node is the same decreasing, concave, and bounded function of the number of customers at that node.

Menich and Serfozo's model does not cover the case in which stations are $\cdot/M/s$ – since then the service rate at each node is increasing rather than decreasing – but this case has been separately considered by Johri [41] who showed that JSQ routing minimizes (among other measures) the average waiting time of customers in the system.

THE SYMMETRIC CASE: SERVE-THE-LONGEST-QUEUE RULE

Many of the models above are special cases of a model considered by Farrar [21], in which the problem is to allocate dynamically both n streams of customers, arriving as non-homogeneous Poisson processes of differing rates, and n servers of differing speeds, so that exactly one arrival process and one server is allocated to each of n symmetric queues in parallel. This is a combined customer routing and server allocation problem. The natural generalization of JSQ is to allocate the jth fastest arrival stream to the jth shortest queue and the jth fastest server to the jth longest queue. This is the JSQ/SLQ ("join-the-shortest-queue/serve-the-longest-queue") rule. Menich and Serfozo also considered this rule for the case in which n identical arrival streams and n identical servers are already present, but there is one extra arrival stream to route and one extra server to allocate. Farrar's proofs are by a sample-path argument based on comparing states in terms of submajorization ordering of states. In this ordering, x is submajorized by y if their components have the same total, but the sum of the kth largest components of x are no more than the sum of the kth largest components of y, for all $k = 1, \ldots, n$. In other words, the customers are spread more evenly between queues in state x than they are in state y. The sample-path approach is more powerful than value iteration. Farrar shows that JSQ/SLQ stochastically minimizes the sum of the k largest queues, for all k, when the sets of instantaneous arrival rates and service rates are arbitrary functions of time. Moreover, as in Menich and Serfozo's work, these rates may also be functions of the state, provided that if x is better than y in the submajorization order (i.e. more spread out) the sets of arrival and service rates for state x are greater in submajorization order than those in y, i.e. less spread out.

Farrar also considers symmetric parallel queues with finite buffers. He proves that JSQ/SLQ stochastically minimizes departure times when buffers are different sizes and there is a single arrival stream to be routed, or if there are several streams to be allocated but buffers are the same size. The first of these results is also proved by Hordijk and Koole [39], who use a Markov decision model and show that JSQ is optimal for infinite buffers and batch arrivals that must be allocated to a single queue before their size is observed.

SUMMARY OF MONOTONICITY RESULTS FOR ADMISSION AND ROUTING

Each of the papers cited so far in this section and the previous section has something to say about the structure of optimal policies for admission and routing of customers in a network of queues, at least for small (e.g., two-node) networks or networks of special (e.g., cyclic or series) structure. We now summarize some of the implications of the results in these papers. For simplicity we only consider the case where the cost per unit time of holding customers is proportional to the total number of customers in the network. (That is, the holding-cost functions $h_i(x_i)$ at all nodes i are the same linear function.) If admission or rejection of arriving customers at a particular node is an option, then a (node-dependent) cost in incurred whenever a customer is rejected (or, equivalently, a reward is earned whenever a customer is accepted). The objective is to minimize the expected discounted total cost.

Consider a series of m nodes with a single exponential server at each node i and a controllable Poisson arrival process at node 1. An optimal policy for admission of arriving customers is monotonic in the following sense. If it is optimal to accept a customer in a particular state, then it is still optimal if a customer is removed from any node i or moved from node i to node $i+1$. (See Weber and Stidham [67].)

Consider a network of two nodes in series, each with an exponential server and a controllable Poisson arrival process. (See Ghoneim [25], Ghoneim and Stidham [26].) If it is optimal to accept a job at node 1 (node 2) in a particular state, then it is still optimal if a job is removed from node 1 or node 2 or moved from node 1 to node 2 (from node 2 to node 1).

Consider a network of two parallel nodes, each with an exponential server and a separate queue, and with a common controllable Poisson arrival process. An optimal policy for routing customers to the two queues is monotonic in the following senses. If it is preferable to route a customer to queue i rather than queue j in a particular state, then it is still preferable if a customer is removed from queue i or added to queue j. (That is, the optimal switching curve is monotonic. See Davis [18], Abdel-Gawad [1], Hajek [29].) If admission or rejection of arriving customers is also an option, then an optimal policy is also monotonic in the following senses. If it is optimal to admit a customer in a particular state, then it is also optimal to do so if a customer is removed from either queue. Moreover, if it is optimal to admit a job and route it to queue i then it is also optimal to do so if a customer is moved from node i to node j. If the service rates of the two servers are equal, then the JSQ ("join-the-shortest-queue") rule is optimal. This result extends to more that two parallel queues, to arbitrary arrival processes, and to service-time distributions with nondecreasing failure rates or state-dependent exponential servers, but not to arbitrary service-time distributions. If there are also servers to allocate then the SLQ ("serve-the-longest-queue") rule is optimal. (See Menich and Serfozo [48], Farrar [21].)

ROUTING TO PARALLEL SERVERS FROM A COMMON QUEUE

In the routing-control models discussed so far, each server has its own queue, a customer upon arrival is routed to one of the queues, and thereafter no jockeying between queues is permitted. Suppose instead that customers enter a common buffer when they arrive and the decision about which server should be assigned to the customer is postponed until the customer reaches the head of the queue. The trade-off is between immediately assigning the customer to a slow server or waiting for a faster server to become available.

Such a decision procedure, if feasible, is clearly preferable to the former approach according to any reasonable performance measure, since all decisions are made at later points in time when more information is available. In some systems, however, it may not be feasible or permissible to delay the routing decision beyond the time point at which the customer arrives. An example is the assignment of vehicles to alternative lanes at a tunnel or toll plaza. For communication systems, it can be argued that the separate-queue model is most appropriate for virtual-circuit routing, in which an arriving customer (session) is assigned a route (virtual circuit) at the instant of arrival (session generation).

Perhaps the first paper to examine the single-queue, parallel-server optimal routing problem was the unpublished Ph.D. dissertation of Farrell [22]. Later references include Sarachik [52], Lin and Kumar [46], Walrand [64], Viniotis and Ephremides [63], and Shenker and Weinrib [56]. Most references concentrate on proving the optimality of a threshold policy: always route the customer at the head of the queue to the faster server, if available, and route to the slower server only if the faster server is unavailable and the number of customers in the queue exceeds a threshold. The models generally assume Poisson arrivals, exponential servers, a linear holding-cost rate, and discounted or average cost criterion. Attempts to prove the optimality of a threshold policy for more than two servers have so far been unsuccessful. Shenker and Weinrib [56] compare the performance of threshold policies to policies based on various heuristics. They also analyze the performance of heuristic routing rules for the separate-queue, parallel-server model discussed previously.

OTHER ROUTING MODELS

Yum and Schwartz [76] proposed a JBSQ ("join-biased-shortest-queue") rule as an improvement on the JSQ rule when the servers have different rates. In this case an arriving customer is routed to the queue that minimizes the weighted number of customers in the queue. (The individually optimal rule, in which an arriving customer minimizes his expected waiting time, is an example of a JBSQ rule, but it is not optimal for the system as a whole when the servers serve at different rates.)

Bovopoulos and Lazar [11] use linear programming to examine numerically the optimal routing and flow-control policies for the separate-queue parallel-server problem. The objective is to maximize throughput, subject to a constraint on expected response time.

Krishnan [43] develops a model for optimal routing to parallel queues. His algorithm uses separable routing, based on an optimal static allocation of flow, as a base policy and then uses one step of policy improvement to yield a state-dependent policy with better performance. Extensive numerical studies suggest that this algorithm performs very well.

Chang [14] considers the separate-queue model where each queue feeds a series of identical exponential servers, with a Poisson arrival process. He shows that the optimal policy routes an arriving customer to the series that minimizes the sum of the queue sizes in the series.

TWO QUEUES WITH COMBINED CONTROL OF ROUTING AND SERVICE

Hajek [29] considers a general two-node model that incorporates many of the features of both the parallel and series queue models (but not the option of accepting or rejecting arriving jobs). In Hajek's model, queues 1 and 2 receive Poisson arrivals at rates λ_1 and λ_2, respectively. A third stream of Poisson arrivals at rate λ can be routed to either queue. The stations have fixed exponential servers with rates μ_1 and μ_2 and a third exponential server with rate μ that can be assigned to either queue; jobs whose service is completed by these servers leave the system. There are two additional exponential servers, with rates γ_{12} and γ_{21}, the first of which serves queue 1 and sends jobs to queue 2, the second of which serves queue 2 and sends jobs to queue 1. Service completions by these servers can be "accepted" or "rejected"; the jobs arriving at rate γ are to be routed to one or the other of the queues; and the server with rate μ is to be assigned to one or the other of the queues. All these decisions are to be made dynamically as a function of the number of jobs in the two queues. Hajek uses an inductive proof to establish the existence of a monotonic switching curve, on which all these decisions can be based. His analysis accommodates convex holding costs at each queue and costs associated with each switching decision.

A METHOD BASED ON PIECEWISE-LINEAR INTERPOLATION

Bartroli [5] presents a new mechanism for inductive proofs of monotonicity of optimal control policies for networks of queues, based upon piecewise-linear interpolation of the value function on an appropriate triangulation of Euclidean space. He applies this mechanism to the problem of optimal admission and routing to two parallel queues and obtains simpler proofs of the structural results mentioned above.

6. Scheduling in networks of queues

Scheduling problems arise in the control of networks of queues when one must decide which class of customers to process next at one or more nodes of a network. The class of a customer may determine its processing time, holding cost, and/or its route through the network.

The earliest results on scheduling a queueing system concern an isolated single-server facility fed by independent Poisson arrival processes, one for each customer class. For the case of a non-preemptive $M/GI/1$ queue in which customer class i has service rate μ_i and holding-cost rate c_i, various authors have shown that a policy known as the $c\mu$ rule minimizes the expected steady-state holding cost per unit time. The $c\mu$ rule always gives priority to a job of the class i with the largest value of $c_i\mu_i$ among the jobs present. It is an example of an *index* rule, in which an index (in this case $c_i\mu_i$) is computed for each job class i and the classes are served in decreasing order of the index value.

Harrison [34,35] considers a multi-class non-preemptive $M/GI/1$ queue, in which the objective is to maximize the expected discounted value of total service reward minus holding costs incurred over an infinite horizon. He uses Markov decision theory (specifically, a variant of the policy improvement technique) to show that an index rule is optimal for this problem. Whittle [71] shows that the optimality of this index rule can be derived from the general theory of Gittins indices (Gittins [27]). Other references on the optimality of index rules for preemptive as well as non-preemptive scheduling of a single-server facility include Whittle [72,73], Baras et al. [3], Baras et al. [4], Varaiya et al. [61], Walrand [65], Nain [50], Nain et al. [51], and Liu and Nain [47].

Klimov [42] considers a multi-class, non-preemptive $M/GI/1$ queue in which a customer of class i, upon completing service, re-enters the system as a class-j customer with probability p_{ij} and leaves the system with probability p_{i0}. He shows that the long-run average cost is minimized by an index rule. Tcha and Pliska [59] consider Klimov's problem and establish the optimality of an index rule for the infinite-horizon discounted-cost case.

Klimov's model can be regarded as the first model for scheduling in a network of queues, if one associates a node of the network with each customer class, with a single server who must select which node to serve next. Customers are routed from node to node independently of one another and of the state of the system according to the Markov routing matrix (p_{ij}). The optimality of an index policy depends crucially, however, on the property that only one node of the network may receive service at any point in time.

Problems in which service may take place simultaneously at more than one node are much harder to solve exactly and do not typically have index rules that are optimal. In section 4 we mentioned Farrar's work [21] on the optimality of the SLQ rule for allocating servers of different speeds to parallel queues. Farrar [20] also considers a system of two queues in tandem, in which each of the two stations has a fixed server, but there is also an additional server whose effort is to be dynamically allocated between the stations. He considers a system in which there are no arrivals (commonly called a clearing system) and shows that a policy that minimizes expected value of total linear holding cost until such time that system clears is transition monotone, in the sense that following a service completion at either station it cannot be optimal to move the extra server to that station if it was not already in use there. The conjecture that the optimal policy is transition-monotone when there are arrivals, or non-linear holding costs is supported by numerical work, but remains open.

Harrison [37] approximates a multi-class queueing network with Poisson input by a Brownian network. The optimal scheduling policies obtained using the Brownian network are nearly optimal for the corresponding multi-class network with heavy traffic. Other works on network scheduling that deal with Brownian approximations of a multi-class queueing network include Harrison [36], Harrison and Wein [38], Wein [68,69], and Laws and Louth [44].

Hariharan et al. [33] (see also Moustafa [49]) study a series of m queues (labeled $j = 1, 2, \dots, m$) with $m + 1$ classes of jobs (labeled $k = 0, 1, 2, \dots, m$). Class-0 jobs require service at node 1 only, whereas class-k jobs, $k = 1, 2, \dots, m$, require service at nodes k through m. The service time at each node is deterministic and equals one unit of time. The decision to be taken at the beginning of every service slot is whether to serve a class-0 or class-1 customer at node 1. The objective is to minimize the expected discounted holding cost over an infinite horizon. Hariharan et al. [33] formulate the problem as a Markov decision process and derive conditions under which it is optimal to schedule a class-0 job, showing in particular that it is optimal to do so in all but a finite polyhedral set of states in the case where all customer classes incur holding cost at the same linear rate.

The model where service times are deterministic finds applications in packet-switched data communication networks, where each message is divided at the source node into packets of fixed length. Since processing times at the nodes and links in a communication network are directly proportional to the length of the packets, the packets require the same service time at each node.

Recent work by Yang et al. [75] and Chen et al. [16] on a similar model with exponential service times is based on point processes with stochastic intensities. The model reduces to a discrete-parameter Markov decision model. They consider a two-node queueing network with two classes of customers. The problem includes the scheduling of the server at node 1 amongst different classes. Dividing the problem into several cases corresponding to certain parameter values, they derive optimal scheduling rules for some of these cases and propose heuristic rules for the other.

Our survey reaches a conclusion in the next section with a discussion of the model of Hariharan et al. [33], as an example of a Markov decision process model of a scheduling problem in a network of queues.

SCHEDULING A SERIES SYSTEM WITH CONSTANT SERVICE TIMES

Consider an m-node network of the type described above. Each node j has a single server ($j = 1, 2, \ldots, m$). The classes of jobs are defined according to their routes. In particular, we define $m + 1$ classes of jobs. Class-0 jobs require service at node 1 only. Class-k jobs ($k = 1, 2, \ldots, m$) require service at nodes $k, k+1, \ldots, m$ (in that order). The service times at each node are constant and equal. Without loss of generality we adopt this service time as our time unit, called a *slot*. Thus, in each time slot t ($t = 0, 1, \ldots$), exactly one customer is served at each node $j = 1, 2, \ldots, m$ (provided that the class to which the server is assigned has at least one customer in its queue at the beginning of the time slot). Let $A_t(k)$ denote the number of class-k arrivals during time slot t and let $A_t = (A_t(0), A_t(1), \ldots, A_t(m))$. We assume that the random vectors A_t, $t = 0, 1, \ldots$ are i.i.d. Customers of all classes incur holding cost at rate 1 while in the system. Future costs are discounted; the one-period discount factor is β. The objective is to minimize the expected total discounted holding cost over an infinite horizon.

Since all the jobs at node j (for $j = 2, 3, \ldots, m$) have the same remaining route and holding-cost rates, they can be regarded as equivalent to class-j jobs. Hence, for each node $j \geqslant 2$, it is sufficient to keep track of the total number of jobs at that node and the only scheduling decision that is to be made is at node 1, where the system controller has to choose between class-0 and class-1 jobs.

Let $x = (x_0, x_1, \ldots, x_m)$ denote the state of the system, where x_0 is the number of class-0 jobs at node 1, x_1 is the number of class-1 jobs at node 1, and x_k, $k = 2, 3, \ldots, m$, is the number of jobs at node k (which we shall call *level-k* jobs). To describe the effects of the scheduling decision, it will be convenient to introduce the following notation. Let

$$d_0(x) := -e_0 \mathbf{1}(x_0 > 0),$$

$$d_i(x) := (-e_i + e_{i+1})\mathbf{1}(x_i > 0), \quad i = 1, 2, \ldots, m-1, \tag{5}$$

$$d_m(x) := -e_m \mathbf{1}(x_m > 0),$$

where e_i is the unit m-vector with a one in the ith component and zeroes elsewhere. Thus, $d_i(x)$ measures the effect on the state vector x of the service of a level-i job, $i = 0, 1, \ldots, m$. Now define state-transformation operators $T_i x$ as follows:

$$T_0 x := x + d_0(x) + \sum_{i=2}^{m} d_i(x), \tag{6}$$

$$T_1 x := x + d_1(x) + \sum_{i=2}^{m} d_i(x). \tag{7}$$

Thus, given that the state of the system is x, if a class-0 customer is served at node 1, then the state of the system at the beginning of next service slot, without the addition of new arrivals, is $T_0 x$. Similarly, if a class-1 customer is served at node 1, then the state of the system at the beginning of next service slot, without the inclusion of the new arrivals, is $T_1 x$.

Let $V(x)$ be the minimal expected total β-discounted cost ($\beta < 1$) over an infinite horizon starting from state x. Then $V(x)$ satisfies the optimality equation

$$V(x) = \sum_{i=0}^{m} x_i + \min\{U(T_0 x), U(T_1 x)\}, \tag{8}$$

where $U(x) = \beta E[V(x + A)]$ and $A = (A(0), A(1), A(2), \ldots, A(m))$ has the common joint distribution of the random vectors $A_t = (A_t(0), A_t(1), A_t(2), \ldots, A_t(m))$, $t = 0, 1, \ldots$.

We now characterize certain sufficient conditions to be met by the state vector x under which it is optimal to serve customers of class 0 at node 1. To establish these properties for the infinite-horizon problem, Hariharan et al. [33] use successive approximations (value iteration) and in the process show that they are also valid for the finite-horizon case.

Let $V_n(x)$ be the minimal expected total β-discounted cost over n periods (time slots) starting in state $x, n \geqslant 1$ ($V_0 \equiv 0$). We can express the value functions V_n recursively as follows:

$$V_n(x) = \sum_{i=0}^{m} x_i + \min\{U_n(T_0 x), U_n(T_1 x)\}, \quad n \geqslant 1, \tag{9}$$

where $U_n(x) = \beta E[V_{n-1}(x + A)]$. It follows from the theory of Markov decision processes (see Schäl [53], Bertsekas [6]) that $V_n(x) \to V(x)$ and $U_n(x) \to U(x)$ as $n \to \infty$.

The following theorem contains the sought-for characterization. It shows that it is optimal to serve class-0 customers at node 1 in all but a finite set of states, which is the intersection of $m - 1$ partial-sum inequalities.

THEOREM 2

Serving class-0 customers at node 1 is optimal in state x if $x_0 \geqslant 1$ and $\sum_{i=2}^{k} x_i \geqslant k$ for some $k \in \{2, 3, \ldots, m\}$.

The proof of theorem 2 depends on the following two lemmas.

LEMMA 1

Given x such that $x_0 \geqslant 1$ and $x_1 \geqslant 1$, if

$$V_n(y - e_0) - V_n(y - e_1 + e_2) \leqslant 0, \quad \text{for all } y \geqslant x, \tag{10}$$

then

$$U_{n+1}(x - e_0) - U_{n+1}(x - e_1 + e_2) \leqslant 0. \tag{11}$$

LEMMA 2

For $j \geqslant i \geqslant 2$, let $X = \{x : \sum_{k=1}^{p} x_{j-k} \geqslant p$ for all $p \in \{1, 2, \ldots, j-i-1\}$ and $\sum_{k=0}^{l} x_{j+k} \geqslant l+1$ for some $l \in \{0, 1, 2, \ldots, m-j\}\}$. If V_0 satisfies the following inequality,

$$f(x+e_i) - f(x+e_j) \leqslant 0, \quad \text{for all } j \geqslant i \geqslant 2 \text{ and } x \in X, \tag{12}$$

then V_n and U_n also satisfy (12).

Lemma 1 follows directly from the definition U_n. Lemma 2 may be proved by induction on n (see Hariharan et al. [33]).

To prove theorem 2 it suffices to show that U_n satisfies the functional inequality

$$f(T_0 x) \leqslant f(T_1 x) \tag{13}$$

if $x_0 \geqslant 1$, $x_1 \geqslant 1$ and $\sum_{i=2}^{k} x_i \geqslant k$ for some $k \in \{2, 3, \ldots, m\}$. This can also be done by induction on n (see [33]), using lemmas 1 and 2.

Remark

Lemma 2 gives conditions on the state-vector x under which having an extra customer at node i instead of at node j ($j > i$), the number of customers at all other nodes being the same, does not result in an additional cost. This might seem counter-intuitive at first. However, the condition $x \in X$ implies that the extra job at node j in state $x + e_j$ is blocked by the jobs at nodes $j, j+1, \ldots, m$, and $i+1, \ldots, j-1$ such that the following statements hold.

(a) When this extra job reaches the terminal node (node m), there is at least one other job at the terminal node (refer to condition $\sum_{k=1}^{l} x_{j+k} \geqslant l$ for some $l \in \{0, 1, \ldots, m-j\}$), so that the extra job continues to stay at node m.

(b) The nodes in between i and j are sufficiently loaded such that, at least until the extra job at node i (state $x + e_i$) reaches the end node, there is a job (other than the extra job from node j, starting in state $x + e_j$) at node m (refer to condition $\sum_{k=1}^{p} x_{j-k} \geqslant p$ for all $p \in \{1, 2, \ldots, j-i-1\}$).

This ensures that the extra job (from state $x + e_j$) is not served at node m until the extra job at node i (from state $x + e_i$) reaches node m so that the extra jobs in the two states leave the system at the same time. In fact, the inequality (12) may not be valid without this condition. (See Hariharan et al. [33] for counterexamples.)

References

[1] E.F. Abdel-Gawad, Optimal control of arrivals and routing in a network of queues, Ph.D. dissertation, Program in Operations Research, N.C. State University, Raleigh (1984).

[2] A. Arapostathis, V.S. Borkar, E. Fernandez-Gaucherand, M.K. Ghosh and S.I. Marcus, Discrete-time controlled Markov processes with average cost criterion: a survey, to appear in SIAM J. Control Optim.

[3] J.S. Baras, A.J. Dorsey and A.M. Makowski, Two competing queues with linear costs and geometric service requirements: the μc rule is often optimal, Adv. Appl. Prob. 17 (1985) 186–209.

[4] J.S. Baras, D.-J. Ma and A.M. Makowski, K competing queues with geometric service requirements and linear costs: the μc rule is always optimal, Syst. Control Lett. 6 (1985) 173–180.

[5] M. Bartroli, On the structure of optimal control policies for networks of queues, Ph.D. dissertation, Department of Research, University of North Carolina at Chapel Hill (1989).

[6] D. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models* (Prentice–Hall, Englewood Cliffs, NJ, 1987).

[7] F.J. Beutler and D. Teneketzis, Routing in queueing networks under imperfect information, Stoch. Stoch. Reports (1989) 81–100.

[8] V.S. Borkar, Controlled Markov chains and stochastic networks, SIAM J. Control Optim. 21 (1983) 652–666.

[9] V.S. Borkar, Control of Markov chains with long-run average cost criterion, in: *Stochastic Differential Systems, Stochastic Control Theory and Applications*, volume 10, eds. W. Fleming and P.L. Lions, IMA Volumes in Mathematics and its Applications (Springer, 1988) pp. 57–77.

[10] V.S. Borkar, Control of Markov chains with long-run average cost criterion: the dynamic programming equations, SIAM J. Control Optim. 27 (1989) 642–657.

[11] A.D. Bovopoulos and A.A. Lazar, Optimal routing and flow control of a network of parallel processors with individual buffers, *Proc. 23rd Allerton Conf. on Communication, Control, and Computing* (1985) pp. 564–573.

[12] R. Cavazos-Cadena, Necessary conditions for the optimality equation in average reward Markov decision processes, Appl. Math. Optim. 19 (1989) 97–112.

[13] R. Cavazos-Cadena, Recent results on conditions for the existence of average optimal stationary policies, Ann. Oper. Res. 28 (1991) 3–26.

[14] M.H. Chang, Martingale dynamics and optimal routing in a network, Appl. Math. Comp. (1988) 309–327.

[15] X. Chao and H. Chen, Optimal intensity control for tandem queues with blocking, Faculty of Commerce and Business Administration, University of British Columbia (1990).

[16] H. Chen, P. Yang and D. Yao, Control and scheduling in a two-station queueing network: optimal policies, heuristics, and conjectures, Technical report, Department of Industrial Engineering and Operations Research, Columbia University (1991).
[17] T. Crabill, Optimal control of a service facility with variable exponential service time and constant arrival rate, Manag. Sci. 18 (1972) 560–566.
[18] E. Davis, Optimal control of arrivals to a two-server queueing system with separate queues, Ph.D. dissertation, Program in Operations Research, N.C. State University, Raleigh (1977).
[19] A. Ephremides, P. Varaiya and J. Walrand, A simple dynamic routing problem, IEEE Trans. Auto. Control AC-25 (1980) 690–693.
[20] T.M. Farrar, Optimal use of an extra server in a two station queueing network, IEEE Trans. Auto. Control AC-37 (1992).
[21] T.M. Farrar, Resource allocation in systems of queues, Ph.D. dissertation, University of Cambridge (1992).
[22] W. Farrell, Optimal switching policies in a non-homogeneous exponential queueing system, Ph.D. dissertation, Graduate School of Management, University of California at Los Angeles (1976).
[23] G. Foschini, On heavy traffic diffusion analysis and dynamic routing in packet-switched networks, *Computer Performance*, eds. K.M. Chandy and M. Reiser (North-Holland, 1977) pp. 499–513.
[24] G. Foschini and J. Salz, A basic dynamic routing problem and diffusion, IEEE Trans. Commun. COM-26 (1978) 320–327.
[25] H. Ghoneim, Optimal control of arrivals to a network of two queues in series, Ph.D. dissertation, Program in Operations Research, N.C. State University, Raleigh (1980).
[26] H. Ghoneim and S. Stidham, Optimal control of arrivals to two queues in series, Euro. J. Oper. Res. 21 (1985) 399–409.
[27] J.C. Gittins, Bandit processes and dynamic allocation indices, J. Roy. Statist. Soc. B41 (1979) 148–164.
[28] P. Glasserman and D. Yao, Monotone optimal control of permutable GSMPs, Technical Report, IE/OR Department, Columbia University, New York.
[29] B. Hajek, Optimal control of two interacting service stations, IEEE Trans. Auto. Control AC-29 (1984) 491–499.
[30] B. Hajek, Extremal splittings of point processes, Math. Oper. Res. 10 (1985) 543–556.
[31] R. Hariharan, V.G. Kulkarni and S. Stidham, Optimal control of admission and routing to two parallel infinite-server queues, *Proc. 29th IEEE Conf. on Decision and Control* (1990).
[32] R. Hariharan, V.G. Kulkarni and S. Stidham, A survey of research relevant to virtual-circuit routing in telecommunication networks, Preprint, Department of Operations Research, University of North Carolina at Chapel Hill (1990).
[33] R. Hariharan, M.S. Moustafa and S. Stidham, Scheduling in a multi-class queueing network, Department of Operations Research, University of North Carolina, Chapel Hill (1991).
[34] J.M. Harrison, Dynamic scheduling of a multi-class queue: discount optimality, Oper. Res. 23 (1975) 270–282.
[35] J.M. Harrison, A priority queue with discounted linear costs, Oper. Res. 23 (1975) 260–269.
[36] J.M. Harrison, *Brownian Motion and Stochastic Flow Systems* (Wiley, New York, 1985).
[37] J.M. Harrison, Brownian models of queueing networks with heterogeneous customers, presented at the *IMA Workshop on Stochastic Differential Systems*, Minneapolis, MI (1986).
[38] J.M. Harrison and L.M. Wein, Scheduling a two-station multiclass closed queueing network in heavy traffic, Oper. Res. (1987), to appear.
[39] A. Hordijk and G. Koole, Note on the optimality of the generalized shortest queue policy, Prob. Eng. Inf. Sci. (1992), to appear.
[40] D.J. Houck, Comparison of policies for routing customers of parallel queueing systems, Oper. Res. 35 (1987) 306–310.

[41] P.K. Johri, Optimality of the shortest line discipline with state-dependent service times, Euro. J. Oper. Res. 41 (1990) 157–161.
[42] G.P. Klimov, Time sharing systems, Theory Prob. Appl. 9 (1974) 532–551.
[43] K.R. Krishnan, Joining the right queue: a Markov decision rule, *Proc. 26th IEEE Conf. on Decision and Control* (1987) pp. 1863–1868.
[44] C.N. Laws and G.M. Louth, Dynamic scheduling of a four-station queueing network, Prob. Eng. Inf. Sci. 4 (1990) 131–156.
[45] T. Lehtonen, On the optimality of the shortest-line discipline, Ph.D. dissertation, Helsinki School of Economics (1981).
[46] W. Lin and P.R. Kumar, Optimal control of a queueing system with two heterogeneous servers, IEEE Trans. Auto. Control AC-29 (1984) 696–703.
[47] Z. Liu and P. Nain, Optimal scheduling in some multi-queue single-server systems, IEEE Trans. Auto. Control AC-37 (1992).
[48] R. Menich and R. Serfozo, Optimality of shortest-queue routing for dependent service stations, Queueing Systems 9 (1991) 403–418.
[49] M.S. Moustafa, Scheduling in a multi-class network, Ph.D. dissertation, Program in Operations Research, North Carolina State University, Raleigh (1987).
[50] P. Nain, Interchange arguments for classical scheduling problems in queues, Syst. Control Lett. 12 (1989) 177–184.
[51] P. Nain, P. Tsoucas and J. Walrand, Interchange arguments in stochastic scheduling, J. Appl. Prob. 27 (1989) 815–826.
[52] P.E. Sarachik and U. Ozgüner, On decentralized dynamic routing for congested traffic networks, IEEE Trans. Auto. Control AC-27 (1982) 1233–1238.
[53] M. Schäl, Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal, Z. Wahrscheinlichkeitstheorie verw. Geb. 32 (1975) 179–196.
[54] L. Sennott, Average cost optimal stationary policies in infinite state Markov decision processes, Oper. Res. 37 (1989) 626–633.
[55] L. Sennott, Average cost semi-Markov decision processes and the control of queueing systems, Prob. Eng. Inf. Sci. 3 (1989) 247–272.
[56] S. Shenker and A. Weinrib, The optimal control of heterogeneous queueing systems: a paradigm for load-sharing and routing, IEEE Trans. Comp. C-38 (1989) 1724–1735.
[57] S. Stidham, Optimal control of admission to a queueing system, IEEE Trans. Auto. Control 30 (1985) 705–713.
[58] S. Stidham, Scheduling, routing, and flow control in stochastic networks, *Stochastic Differential Systems, Stochastic Control Theory and Applications*, eds. W. Fleming and P.L. Lions, IMA vol. 10 (1988) pp. 529–561.
[59] D.W. Tcha and S.R. Pliska, Optimal control of single-server queueing networks and multi-class $M/G/1$ queues with feedback, Oper. Res. 25 (1977) 248–257.
[60] D. Topkis, Minimizing a submodular function on a lattice, Oper. Res. 26 (1978) 305–321.
[61] P. Varaiya, J. Walrand and C. Buyukkoc, Extensions of the multiarmed bandit problem: the discounted case, IEEE Trans. Auto. Control AC-30 (1985) 426–439.
[62] M. Veatch and L. Wein, Monotone control of queueing and production/inventory systems, Preprint, Operations Research Center, MIT (1991).
[63] I. Viniotis and A. Ephremides, An optimization problem in a multiserver queue, *Proc. 23rd Allerton Conf. on Communications, Control, and Computing* (1985) pp. 555–563.
[64] J. Walrand, A note on "Optimal control of a queueing system with heterogeneous servers", Syst. Control Lett. 4 (1984) 131–134.
[65] J. Walrand, *An Introduction to Queueing Networks* (Prentice–Hall, Englewood Cliffs, NJ, 1988).
[66] R. Weber, On the optimal assignment of customers to parallel servers, J. Appl. Prob. 15 (1978) 406–413.

[67] R. Weber and S. Stidham, Control of service rates in networks of queues, Adv. Appl. Prob. 24 (1987) 202–218.
[68] L.M. Wein, Asymptotically optimal scheduling of a two-station multi-class queueing network, Technical Report, Sloan School of Management, M.I.T. (1987).
[69] L.M. Wein, Reducing the variance of sojourn times in multiclass queueing systems, Technical Report, Sloan School of Management, M.I.T. (1989).
[70] W. Whitt, Deciding which queue to join; some counterexamples, Oper. Res. 34 (1986) 55–62.
[71] P. Whittle, Multi-armed bandits and the Gittins index, J. Roy. Statist. Soc. B42 (1980) 143–149.
[72] P. Whittle, *Optimization over Time: Dynamic Programming and Stochastic Control*, Vol. 1 (Wiley, New York, 1982).
[73] P. Whittle, *Optimization over Time: Dynamic Programming and Stochastic Control*, Vol. 2 (Wiley, New York, 1983).
[74] W. Winston, Optimality of the shortest-line discipline, J. Appl. Prob. 14 (1977) 181–189.
[75] P. Yang, H. Chen and D. Yao, Optimal control and scheduling in a multi-class queueing network: results and conjectures, SIAM Conf. Proc., New Orleans (1990).
[76] T.K. Yum and M. Schwartz, The join-biased-queue rule and its application to routing in computer communication networks, IEEE Trans. Commun. COM-29 (1981) 505–511.