ROUTING IN QUEUEING NETWORKS UNDER IMPERFECT INFORMATION: STOCHASTIC DOMINANCE AND THRESHOLDS

FREDERICK J. BEUTLER and DEMOSTHENIS TENEKETZIS

Department of Electrical Engineering and Computer Science, Communications and Signal Processing Laboratory, The University of Michigan, Ann Arbor, Michigan 48109

Optimal policies for routing between two servers under imperfect information is treated by discrete time dynamic programming. It is proved that certain inequalities involving stochastic ordering of information measures can be propagated inductively from one epoch to the next. Convexity conditions on the instantaneous costs insure proper initiation and inductive continuation of these properties. Consequently, an inductive procedure shows that threshold routing policies are optimal, and that the total cost is convex and monotone.

Two examples are provided. The first deals with a tandem queue having inputs to both work stations, and only inferential information available on the state of the second station. It is first shown that the optimal control is bang-bang, and then that the hypotheses dictating a threshold policy resulting in convex and monotone costs are satisfied. The second example considers optimal routing for customers arriving in a renewal stream, when the routing decision is between two parallel exponential servers, and the observations are both delayed and subject to random errors. It is again shown that the optimal policy is of the threshold type, and that the costs are monotone and convex.

KEY WORDS: Dynamic programming, stochastic ordering, stochastic dominance, optimal policy, partial information, queues, routing, optimal routing, bang-bang control, bang-bang policy, threshold policy, submodularity, convexity.

1. INTRODUCTION

Communication networks are often characterized by work stations that act individually, each possessing only a local knowledge of its immediate environment. Even when information can be exchanged among stations, there are propagation and processing delays that render such information partially obsolete; also, faults and transmission errors may render the data inaccurate. The optimization problem can then be generally expressed as: based on this partial knowledge of the system state, how should the individual work stations operate so as to maximize some measure of the overall utility for the entire system?

While we cannot expect to obtain universal results of that nature, it is possible to isolate certain network aspects for analysis and, in particular, to formulate relevant problems in terms of optimal decentralized control, or as problems of optimization based on incomplete knowledge of the state.

The work presented here is based on queueing models of communication

networks, and more specifically, the nature of optimal routing and flow control for such systems. These include considerations of whether and when a customer should be admitted for processing, to which work station an entering customer should be sent, how shall he be routed through the network, when is it advantageous to transfer a customer from one work station to the next, etc.

Papers in this area have, of course, appeared earlier; we specifically cite [1, 5–9, 15, 16, 18, 21, 22], all of which have treated the dynamic optimization of admission and customer transfer in simple queueing networks. In each case, the analysis is dependent on the state of the (Markov) network being completely known. The nature of these publications suggests that generalizations to partial information or multiple decentralized control will be extremely difficult. Nevertheless, we believe that we have been able to develop methodologies for elucidating qualitative properties of optimal policies.

Specifically, our approach treats queueing models that can be reduced to discrete time Markov systems through uniformization or embedding. This suggests dynamic programming as an appropriate tool for the analysis of optimal policies. The programming equation appears in terms of information states (see [2] or [10]), which reflects the assumption of partial knowledge.

In Section 2, we examine typical dynamic programming relations for routing applications. It is traditional to represent customer arrivals, routings, and departures by measure preserving operators on information measures. However, our work is phrased in the more general context of stochastic orderings (see [14]); this approach proved to be advantageous for two reasons: first, it circumvents the difficulties others (e.g., [6, 21]) have encountered with departures or transfers near the "boundary", and second, such orderings constitute a unifying principle for routing actions. Accordingly, inequalities on the value function are expressed in terms of stochastic dominance. These inequalities imply threshold properties for choice of routing between two servers, as well as monotonicity and convexity of the value function.

It is shown that if each of two functions is subject to these inequalities, so is their minimum. This suggests that the same properties are valid by induction at each step of the dynamic programming equation. To complete the necessary arguments, sufficient conditions are derived for the running (instantaneous) costs to initiate the inductive process, and to maintain it from each step to the next.

Section 3 offers two examples that generalize the existing literature. The first extends [13] by consideration of a tandem queue where customer entrance is effected to both work stations, and conditions at the second station are incompletely known. It is shown that convex instantaneous cost functions lead to an optimal policy of the threshold type, with monotone convex expected costs.

The second problem treats the optimal choice for routing customers arriving in accordance with a renewal process to a choice of one of two exponential servers. The routing decision is based on a delayed observation of the system state. Moreover, this observation is subject to random (independent) error, and the number of service completions between customer arrivals is unknown. Again, a convex cost function results in a threshold policy, and a convex monotone expected cost.

2. METHODOLOGY

This section treats certain analytic constructs used to deduce properties of optimal routing policies for Markov queueing systems under conditions of partial knowledge of the system state. Specifically, it is supposed that at epoch n, when the true state is $\mathbf{x}_n$, the observation at that epoch is specified by $z_{n+1} = h(\mathbf{x}_n, \mathbf{w}_n)$, where $\mathbf{w}_n$ is regarded as a random measurement error. It is well established (see [2] or [10]) that, under mild subsidiary conditions, the optimization problem can then be posed in terms of the dynamic programming equation, in which the argument of the future expected cost is an *information state*. Accordingly, we shall deal with information states that are probability measures on $\mathbf{Z}_2^+$ (two dimensional vectors composed of non-negative integer valued entries) for each set of observations over the past. The class of probability measures on $\mathbf{Z}_2^+$ is denoted by $\mathbf{M}$, and $\pi \in \mathbf{M}$ will be used to indicate a generic information state.

Specifically, we are concerned with Markov systems in which dynamic decisions on routing are applied to customers at their time of arrival and/or service completion, where the actions are chosen on the basis of incomplete knowledge. Thus, discrete time dynamic programming in terms of information state arguments become applicable.

While the resulting optimal policies are too complex to characterize explicitly, it is nevertheless possible to demonstrate qualitative properties. We shall consider in detail systems where the arriving customer is to be routed to either one of two work stations. Thus, the admissible (routing) actions move the process from $\mathbf{x}$ to $\mathbf{x} + e_i$, where e_i is the unit vector along component i. Then the optimal policy is said to possess the *threshold property* if, whenever the action i (which increases x_i by unity) is optimal for $\mathbf{x} + e_i$, the same action is also optimal for $\mathbf{x}$. Alternatively, if i is not optimal for $\mathbf{x}$, it cannot be optimal for $\mathbf{x} + e_i$.

Recognition that the optimal policy is of threshold type reduces the search over policies described by a threshold. We shall also demonstrate convexity and monotonicity of the expected total cost, which can be treated by similar techniques.

For routing problems, the dynamic programming equation often takes on the general form

$$V_n(\pi) = C_n(\pi) + \min\left[V_{n+1}(A_1\pi), V_{n+1}(A_2\pi)\right], \qquad (2.1)$$

where A_i is an operator on π generalizing the notion of an arrival by the relation

$$(A_i\pi)(\mathbf{x}) = \pi(\mathbf{x} - e_i). \qquad (2.2)$$

on $\mathbf{x} \in \mathbf{Z}_2^+$. Also, in (2.1), C_n represents the accumulated running cost in continuous time over the interval $[t_n, t_{n+1})$, in which t_n is the instant of the nth decision point.

The form of (2.1) suggests that induction constitutes a suitable technique for proving the threshold properties as well as other facts about the optimal policy and the resulting minimal costs. For the finite horizon problem with N epochs, one shows that $C_N(\cdot)$ possesses property $\mathbf{P}$, and then demonstrates that whenever

P is true for $V_{n+1}(A_i\pi)$, it also holds for the minimum of such functions. To begin the induction for desirable properties such as optimality of a threshold, and/or the convexity and monotonicity of costs, one generally supposes that $C_N(\cdot)$ is itself convex and monotone.

The inductive plan based on (2.3) below has been carried out by Davis [4] and by Hajek [6], who found conditions sufficient for propagation of the threshold property. These conditions intertwine convexity and monotonicity with a submodularity condition that directly describes the threshold property. The result is that if $f: \mathbf{Z}_2^+ \to \mathbf{R}^+$ meets (for $i, j = 1, 2$) the three conditions

$$f(\mathbf{x}) \leq f(A_i\mathbf{x}) \tag{2.3a}$$

$$f(A_i\mathbf{x}) - f(\mathbf{x}) \leq f(A_iA_j\mathbf{x}) - f(A_j\mathbf{x}) \tag{2.3b}$$

$$f(A_i\mathbf{x}) - f(A_j\mathbf{x}) \leq f(A_i^2\mathbf{x}) - f(A_jA_i\mathbf{x}) \tag{2.3c}$$

so does $a(\mathbf{x}) \triangleq \min [f(A_1\mathbf{x}), f(A_2\mathbf{x})]$. Functions (on whatever state space) satisfying the conditions (2.3) will be said to belong to **X**; hence, the desired result is that $f \in \mathbf{X}$ implies $a \in \mathbf{X}$.

Here (2.3c) is a submodularity relation that has a particular threshold optimality meaning; $f(A_i^2\mathbf{x}) - f(A_jA_i\mathbf{x})$ is the cost advantage of routing to i instead of to j for state $A_i\mathbf{x}$ and this advantage is enhanced if the i queue is shortened to state $\mathbf{x}$. A generalized form of (2.3c), appropriately rephrased in terms of information states, will appear in (2.7c). The Eqs. (2.3a) and (2.3b) relate respectively to the monotonicity and convexity of the expected cost rather than to the nature of the optimal policy.

Hajek's proof of (2.3a) to (2.3c) entails a division and subsequent subdivision of different cases, until there are eleven cases altogether. This suggests that his method is not readily amenable to generalization to information states, such as we have in mind. Specifically, we are interested in optimization under conditions of partial knowledge of the state, so that we shall have to work over measures on $\mathbf{Z}_2^+$ to apply the necessary inequalities to information states. The inequalities (2.3) will then be generalized to expressions involving **M** rather than vectors on $\mathbf{Z}_2^+$.

Thus, instead of showing that $a(\mathbf{x}) \triangleq \min [f(A_1\mathbf{x}), f(A_2\mathbf{x})]$, belongs to **X**, we should ask concerning the analogous property for

$$a(\pi) \triangleq \min [f(A_1\pi), f(A_2\pi)]. \tag{2.4}$$

Yet another generalization relates to departures and customer transfers, whose treatment (compare [6] and [21]) is inherently more difficult than that of arrivals, since the departure operator produces change only when customers are actually present. In systems with exponential service, one may then assume that departures occur in a Poisson stream, but are real when customers are present, and virtual if the queue is empty.

The generic departure operator will be denoted by D, and a departure from the

·coordinate i is represented by D_i. A departure from the first coordinate, for example, is described by its effect on a measure π on $\mathbf{Z}_2^+$ by

$$(D_1\pi)(\mathbf{x}) = \pi(\mathbf{x} + e_1)\delta(x_1 > 0) + [\pi((0, x_2)) + \pi((1, x_2))]\delta(x_1 = 0). \qquad (2.5)$$

The analytical description (2.5) of a departure corresponds in all respects to the intuitive behavior of departures. For instance, the operator D is a composition of commuting operators, and A_i commutes with D_j for $i \neq j$. Moreover, any D has range $\mathbf{M}$, while A_i is one to one from $\mathbf{M}$ into itself. Indeed, D_i is a left (but not a right) inverse for A_i. D_i is a left inverse for A_i, but D_i and A_i fail to commute. The latter reflects the fact that an arrival followed by a departure leaves the state unchanged, but that the same is not true for a departure preceding the arrival.

Since it is necessary to consider various combinations of arrivals and departures, the notation and arguments become quite cumbersome. Fortunately, we are able to avoid these complexities by introducing concepts from stochastic ordering; indeed, stochastic ordering acts as a unifying notion that enables us to treat models in which both departures and arrivals play a crucial role. We define a stochastic ordering (compare [14], Chapter 8) through

DEFINITION 2.1 μ is larger than v (in symbols: $v < \mu$) if

$$q^v((x_1, x_2)) \leq q^\mu((x_1, x_2)) \qquad (2.6)$$

for all elements of $\mathbf{Z}_2^+$.   ∎

In (2.6),

$$q^\pi(\mathbf{x}) \triangleq \sum_{\mathbf{y} \geq \mathbf{x}} \pi(\mathbf{y}),$$

where $\mathbf{x} \triangleq (x_1, x_2)$, and $\mathbf{y} \geq \mathbf{x}$ is taken componentwise for vectors in $\mathbf{Z}_2^+$. Examples of orderings of measures (Definition 2.1) include $\pi < A_i\pi$, $D\pi < \pi$, and $\pi < A_iD_i\pi$.

To analyze threshold policies, it is also desirable to describe probability shifts for routings to or departures from a particular work station. For notational convenience, we choose station #1, and define stochastic dominance in terms of that station.

DEFINITION 2.2 μ is larger than v along coordinate 1 (in symbols: $v <_1 \mu$) if

a) $v < \mu$

and

b) $q^v((0, x_2)) = q^\mu((0, x_2))$

for all $\mathbf{x} \in \mathbf{Z}_2^+$.   ∎

Thus, the first and third examples of the preceding paragraph illustrate ordering along coordinate i.

We can now rewrite (2.3) in terms of stochastic dominance. These inequalities now apply to $f: \mathbf{M} \to \mathbf{R}^+$ and take the form

$$f(v) \leqq f(\mu) \tag{2.7a}$$

$$f(A_j v) - f(v) \leqq f(A_j \mu) - f(\mu) \tag{2.7b}$$

whenever $v < \mu$, and for $v <_j \mu$

$$f(A_j v) - f(A_i v) \leqq f(A_j \mu) - f(A_i \mu). \tag{2.7c}$$

Of course, (2.7) specializes to (2.3) if the measures are taken to consist of a single atom, and μ is obtained from v by a single mass shift. As with (2.3), we shall use $\mathbf{X}$ to denote the class of functions satisfying (2.7). With suitable assumptions on $C_n(\cdot)$ in (2.1), proof of the inductive propagation of the inequalities (2.7) from V_n to V_{n+1} is then tantamount to showing that $f \in \mathbf{X}$ implies $a \in \mathbf{X}$, where $a(\cdot)$ is defined by (2.4).

Induction for the threshold property requires only submodularity (2.7c), as will be shown. The proof indicates that, if f satisfies (2.7c), the same will be true for a without reference to the other inequalities (2.7a) and (2.7b). Moreover, at each epoch (2.7c) implies the threshold property by the same argument as that following (2.3c).

As an example of the role of the departure operator relative to dominance along a coordinate, let us consider optimal policies for arrival routing in the presence of independent, unmeasurable departures over the interval between arrivals. This leads to terms that reflect the expectation E_D of the departure process. These are of the form

$$b(\pi) \triangleq \min E_D[f(A_1 D\pi), f(A_2 D\pi)] \tag{2.8a}$$

or

$$b(\pi) \triangleq E_D\{\min [f(A_1 D\pi), f(A_2 D\pi)]\}, \tag{2.8b}$$

in the dynamic programming equation, depending on the particular problem considered.

Since properties (2.7) are maintained under summation with non-negative coefficients, we can focus attention on $f(D\pi)$ only, rather than on the more complicated sums occurring in (2.8a). The same is true for (2.8b), since if $f(D \cdot) \in \mathbf{X}$ for each D the class $\mathbf{X}$ is closed under expectation with respect to D. A reduction can be accomplished by limiting consideration to $f(D_1 \pi)$. In fact, for $D = D_2^l D_1^k$ one uses commutativity to introduce the new measure $\bar{\pi} = D_2^l D_1^{k-1} \pi$. For this measure, we have $D_1 \bar{\pi} = D\pi$, so that the reduction is accomplished by writing the relevant functions in terms of the new argument $\bar{\pi}$. Consequently, the threshold property will hold for (2.8) it for each realization D, and each element of $\mathbf{M}$,

$$f(A_1 D_1 \pi) - f(A_2 D_1 \pi) \leqq f(A_1 \pi) - f(A_2 \pi). \tag{2.9}$$

This inequality falls within the purview of (2.7), since $D_1 \pi <_1 \pi$.

While inductive propagation of submodularity (2.7c) does not require the presence of either monotonicity (2.7a) or convexity (2.7b), the reverse is not true. Specifically, convexity need not extend to minima of convex functions, so convexity may not propagate in the absence of monotonicity or submodularity. We shall find that the convexity proof not only requires the other two properties, but turns out to be quite involved.

We formalize the threshold results by

THEOREM 2.3   Let (2.7c) hold for $f: \mathbf{M} \to \mathbf{R}^+$. Then the same inequality is true for $a(\cdot)$ as defined by (2.4).

Proof   As can be directly verified, $\min(a, b)$ is alternatively expressed as $a + (b-a)^-$ or $b - (b-a)^+$, where $(b-a)^- = \min(b-a, 0)$ and $(b-a)^+ = \max(b-a, 0)$; if these are applied respectively to $a(A_1 v)$ and $a(A_2 v)$, we get

$$a(A_1 v) - a(A_2 v) = [f(A_1(A_1 v)) - f(A_2(A_1 v))]^- + [f(A_1(A_2 v)) - f(A_2(A_2 v))]^+. \quad (2.10)$$

Now observe that $v <_1 \mu$ implies $A_1 v <_1 A_1 \mu$. If we majorize the right side of (2.10) by this inequality, we have the result (2.7c) for $a(\cdot)$ with $j = 1$ and $i = 2$. ∎

The proof that (2.7a) implies monotonicity for $a(\cdot)$ is immediate (even for the most general case), and will receive no further attention. Also (2.7c) has been taken care of in Theorem 2.3. Therefore, the remaining effort is directed at the inductive propagation of the convexity property (2.7b).

Suppose again that $v < \mu$, with the additional proviso that v and μ are chained. We say that v and μ are chained (i.e., connected by a chain) if there exists a finite sequence $\pi_1, \ldots, \pi_N$ in $\mathbf{M}$ such that

$$v <_{i_1} \pi_1 <_{i_2} \pi_2, \ldots, \pi_N <_s \mu. \quad (2.11)$$

For such chained measures, we have

THEOREM 2.4   Let the inequalities (2.7) hold for $f: \mathbf{M} \to \mathbf{R}^+$. Then (2.7b) applies to $a(\cdot)$ for all chained measures.

Remark   A proof that the more restricted inequality

$$f(A_j \pi) - f(\pi) \leqq f(A_j(A_i \pi)) - f(A_i \pi)$$

implies the same for $a(\cdot)$, can be accomplished by a single argument serving all circumstances. However, that proof cannot be extended to the more general situation treated here, so that we must consider three exhaustive cases as indicated below. Also note that the submodularity of f appears in the proof for Case III.

Proof   By the transitivity of the reals, we only need to show the asserted result for $v <_j \mu$. We wish to prove

$$a(A_2 v) - a(v) \leqq a(A_2 \mu) - a(\mu) \quad (2.12)$$

for each $v <_1 \mu$ and $v <_2 \mu$.

Now we prove (2.12) for $v <_1 \mu$. Three cases need to be considered separately. For the first case, assume

$$a(v) = f(A_2 v). \tag{2.13}$$

For $j = 1, 2$ we have

$$a(A_2 v) - a(v) \leqq f(A_j(A_2 v)) - f(A_2 v) \leqq f(A_j(A_2 \mu)) - f(A_2 \mu) \tag{2.14}$$

by the definition (2.4) of $a(\cdot)$, and by (2.7b). Now choose j such that $a(A_2 \mu) = f(A_j(A_2 \mu))$, and observe that $f(A_2 \mu) \geqq a(\mu)$ to show that the right side of (2.14) is dominated by $a(A_2 \mu) - a(\mu)$.

For Case II, assume the simultaneous equalities

$$a(v) = f(A_1 v) \quad \text{and} \quad a(A_2 \mu) = f(A_1(A_2 \mu)). \tag{2.15}$$

The string of inequalities

$$a(A_2 v) - a(v) \leqq f(A_1(A_2 v)) - f(A_1 v) \leqq f(A_2(A_1 \mu)) - f(A_1 \mu) \tag{2.16}$$

completes the argument, since the last expression in (2.16) is majorized by $a(A_2 \mu) - a(A_1 \mu)$.

If neither Case I nor Case II applies, we have the remaining possibility of Case III, which consists of the conditions

$$a(v) = f(A_1 v) \quad \text{and} \quad a(A_2 \mu) = f(A_2^2 \mu). \tag{2.17}$$

Using the first equality yields

$$a(A_2 v) - a(v) \leqq f(A_1(A_2 v)) - f(A_1 v). \tag{2.18}$$

But by (2.7c),

$$f(A_1(A_2 v)) - f(A_1 v) \leqq f(A_2(A_2 v)) - f(A_2 v) \tag{2.19}$$

because $v <_2 A_2 v$. Moreover the right side of (2.19) is majorized by the replacement of $A_2 v$ by $A_2 \mu$. The argument is completed by observing that

$$f(A_2(A_2 \mu)) - f(A_2 \mu) \leqq a(A_2 \mu) - a(\mu)$$

by use of the second equality of (2.17) and the relation $a(\mu) \leqq f(A_2 \mu)$.

Having finished the proof for $v <_1 \mu$, we now turn to the other possibility, namely $v <_2 \mu$. Here we find that

$$a(A_2 v) - a(v) = [a(A_2 v) - a(A_1 v)] + [a(A_1 v) - a(v)]. \tag{2.20}$$

The first term is majorized by $a(A_2\mu) - a(A_1\mu)$ by an application of (2.7c) to $a(\cdot)$. By what has been already proved, the second term of (2.20) can only be larger if v is replaced by μ. The result is then (2.12). ∎

For reasons too complicated to pursue here, the proofs of Theorems 2.3 and 2.4 do not extend to minima of more than two functions. Moreover, we have been unable to construct a suitable proof, even under alternate sets of assumptions. Nevertheless, we continue to conjecture that if f (now defined for measures over Z_m^+) belongs to X, so does

$$a(\pi) \triangleq \min_{1 \leq i \leq m} [f(A_i\pi)]. \tag{2.21}$$

If the conjecture is valid, we find that for $v <_j \mu$

$$f(A_j v) - a_j(v) \leq f(A_j\mu) - a_j(\mu). \tag{2.22}$$

in which $a_j(\cdot)$ is

$$a_j(\pi) \triangleq \min_{i \neq j} [f(A_i\pi)]. \tag{2.23}$$

The interpretation of the *threshold sufficiency criterion* (2.22) is as follows: if the optimal routing is to j for information state μ, the right side of (2.22) is nonpositive. Then the left side of (2.22) cannot be positive, implying an optimal routing to the same station j for state v. Conversely, if j is not optimal for v, it cannot be optimal for μ.

In the remainder of this Section, we investigate aspects of stochastic ordering relevant to properties of control with imperfect information, or decentralized control. Two general topics are considered. First, we introduce an approximation technique that extends the applicability of Theorem 2.4 to ordered measures that are not chained. Second, we treat in greater detail the contribution of the instantaneous (running) cost in (2.1) to the total cost V_n, with reference to maintaining the convexity of V_n and the threshold character of the total cost.

Let us again take X to be a random variable defined on Z_2^+, and distributed according to $\pi \in M$. We then use π_r to denote the approximation

$$q_r^\pi(x) = \sum_{k=0}^{2^r} k2^{-r} I(k2^{-r} \leq q^\pi(x) < (k+1)2^{-r}) \tag{2.24}$$

to π. In the above, $I(\cdot)$ is an indicator function, which is unity on the set inside the parentheses. It is easy to verify that π_r is a probability measure belonging to M. Further, $v < \mu$ implies $v_r < \mu_r$ for $r = 1, 2, \ldots$, and $v <_k \mu$ implies $v_r <_k \mu_r$.

To extend the sequence of approximations to f, we need a continuity hypothesis. For this purpose, we must have a topology on finite measures on Z_2^+. A

convenient (but by no means only) choice is obtained by defining open sets containing $\bar{\pi} \in \mathbf{M}$ by

$$N^{\pi}(n, \mathbf{x}) = \{\pi : |q^{\pi}(\mathbf{x}) - q^{\bar{\pi}}(\mathbf{x})| < n^{-1}\}.$$

We then require

HYPOTHESIS 2.5 $f: \mathbf{M} \to \mathbf{R}^{+}$ is continuous in the topology just mentioned.   ∎

A specific condition assuring the continuity of the V_n is that $C_n(\cdot)$ be continuous for all n. In particular, $V_0 = C_0$ is then continuous, and continuity of the subsequent V_n follows by induction from (2.1), the minimum of continuous functions again being continuous. Continuity and other properties of the C_n are discussed later.

The continuity hypothesis enables us to assert.

THEOREM 2.6   Let $f \in \mathbf{X}$ (i.e., satisfy the three inequalities (2.7)), and assume that f is continuous according to Hypothesis 2.5. Then $a \in \mathbf{X}$.

Proof   That $a(\cdot)$ satisfies (2.7a) and (2.7c) has already been demonstrated. We also know a satisfies (2.7b) for $v < \mu$ when these are chained. Therefore, the proof will utilize the approximation (2.24) to apply Theorem 2.4 to this inequality.

For any $\pi \in \mathbf{M}$ and any r, the total number of non-zero $q_r^{\pi}(\mathbf{x})$ is bounded, and each non-zero term can only assume one of a finite number of values $(k+1)2^{-r}$. Now suppose $v < \mu$, which entails $v_r < \mu_r$ for each r. Then v_r and μ_r differ in their probability mass distribution only in a finite number of multiples of 2^{-r}, each of which is moved in the direction of increasing coordinate values to go from v_r to μ_r. We observe that each such move of mass represents a step in a chain such as (2.11); hence v_r and μ_r are chained.

According to Theorem 2.4, (2.7b) is applicable to each

$$a_r(\pi) \triangleq \min [f(A_1\pi_r), f(A_2\pi_r)];$$

moreover, a is a continuous on $\mathbf{M}$, because the same is true for f. It is also trivial to note that $\pi_r \to \pi$ implies $A_i\pi_r \to A_i\pi$. In short, writing (2.7b) in terms of a_r, and taking limits on r produces the desired inequality for a.   ∎

In the applications envisaged by us, $f: \mathbf{M} \to R^{+}$ is invariably an expectation. Accordingly, for each such f there is a $K: \mathbf{Z}_2^{+} \to R^{+}$ such that

$$f(\pi) = E^{\pi}(K) = \sum_{\mathbf{x}} \pi(\mathbf{x})K(\mathbf{x}). \tag{2.25}$$

In this setting, stochastic ordering may be characterized in terms of expectations depending on q^{π} (cf. Definition 2.1). In fact, a simple computation shows that (2.25) is equivalent to

$$E^{\pi}(K) = \sum_{\mathbf{x}} q^{\pi}(\mathbf{x})[K(\mathbf{x}) - K(\mathbf{x} - e_1) - K(\mathbf{x} - e_2) + K(\mathbf{x} - e_1 - e_2)] \tag{2.26}$$

where we adopt the convention that $K(\mathbf{x}) = 0$ whenever any component of $\mathbf{x}$ is negative.

Equation (2.26) provides the connection between stochastic dominance and functions possessing properties (2.3a) and (2.3b). Indeed, from Definition 2.1 and (2.26) it follows that $v < \mu$ if and only if

$$E^v(K) \leq E^\mu(K) \tag{2.27}$$

for all $K: \mathbf{Z}_2^+ \to \mathbf{R}^+$ such that the right side of (2.27) is finite, with

$$K(\mathbf{x}) - K(\mathbf{x} - e_1) - K(\mathbf{x} - e_2) + K(\mathbf{x} - e_1 - e_2) \geq 0 \tag{2.28}$$

for all $\mathbf{x} \in \mathbf{Z}_2^+$. Similarly, $v <_1 \mu$ if and only if (2.27) holds for K monotone nondecreasing in the first coordinate, with equality for K depending only on the second coordinate.

Let us apply the above arguments to the running costs C_0 which also equals the initial total cost V_0. It is seen from the form (2.25), together with (2.27) and (2.28), that we can claim the validity of (2.7a) if K satisfies the monotonicity and convexity conditions (2.3a) and (2.3b). Unfortunately, however, we cannot assert (2.7b) or (2.7c) for V_0. Consequently, we are unable to start the inductive process that assures convexity and submodularity at each epoch.

As a practical matter, however, the running cost function (here denoted generically by $K(\cdot)$) is most often of the form

$$K(\mathbf{x}) = K_1(x_1) + K_2(x_2). \tag{2.29}$$

In that case, we have for the corresponding expected cost

$$E^\pi(K) = \sum_{i=1}^{2} \sum_{j=0}^{\infty} q_i^\pi(j) K_i(j) \tag{2.30}$$

in which $q_i^\pi(j)$ denotes the marginal probability in coordinate i, and for values of x_i equal to or greater than j.

The form (2.29) is of particular interest since, if $C_0 = V_0$ is of this type, it is easy to verify natural conditions under which all three of the inequalities (2.7) are met. In fact, we have

THEOREM 2.7   Let $K(\cdot)$ be as in (2.29), and let the K_i be non-negative, monotone, and convex. Then the functions f defined by

$$f(\pi) = E^\pi(K) \tag{2.31}$$

satisfy the inequalities (2.7).

Proof In the notation of (2.29) to (2.31),

$$f(A_i v) - f(v) = \sum_{k \geq 0} q_i^v(k)\{[K_i(k+1) - K_i(k)] - [K_i(k) - K_i(k-1)]\} + K_i(0). \quad (2.32)$$

By the hypothesized convexity, the expression in braces is non-negative. Since $v < \mu$ implies $q_i^v(k) \leq q_i^\mu(k)$, (2.7b) follows.

To verify (2.7c), one observes that for $v <_j \mu$ and $j \neq i$

$$f(A_j v) - f(v) \leq f(A_j \mu) - f(\mu) \quad (2.33a)$$

and

$$f(A_i v) - f(v) = f(A_i \mu) - f(\mu); \quad (2.33b)$$

this is a consequence of (2.32) and $q_i^v(k) = q_i^\mu(k)$. Subtracting (2.33b) from (2.33a) immediately yields (2.7c). ∎

According to the theorem, therefore, running costs taking on the form (2.29), and having non-negative, monotone and convex components permit the initiation of the inductive process. Not only does V_0 then have the properties (2.7), but also, the C_n will be consistent with the continuation of the induction.

3. APPLICATIONS

3.1 Routing Control for Tandem Queues

3.1.1 Problem description—features of the problem Consider the network of Figure 1 consisting of two stations in tandem. Arrivals at Station 1 are described by a discrete-time Bernoulli process with parameter λ_1, and service at Station 2 is also i.i.d., with probability μ_2 that a customer at that station departs the system. Station 2 of the network accepts outside customers whose arrivals are described by a discrete-time Bernoulli process with parameter λ_2. Finally, transfers from Station 1 to Station 2 are specified at each epoch by a controlled parameter $\mu \in [0, \bar{\mu}]$, where μ is the probability that a job at Station 1 (if any) leaves that station and moves to Station 2.
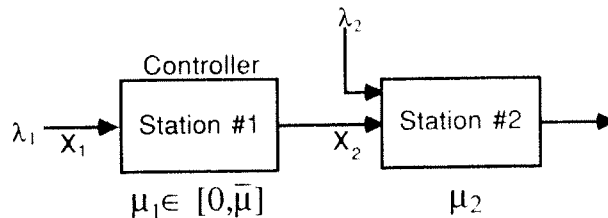

Figure 1 Controlled tandem queue.

At each epoch the controller observes arrivals to Station 1 as well as departures from that station. Thus, the controller knows precisely the number x_1 of customers

at Station 1, but has only inferential information regarding x_2, the number of customers at Station 2; the latter is expressed through a probability distribution $\pi_2(n)$. The objective is then to choose the controlled transfer parameter $\mu(n)$ to minimize an infinite horizon average discounted cost whose instantaneous components are $c_1(x_1(n)) + c_2(x_2(n))$, and $c_i(\cdot)$ $i = 1, 2$ are convex nondecreasing functions of their arguments. We allow simultaneous arrivals, services and transfers.

Simultaneous arrivals, services and transfers are permitted as a response to a peculiar phenomenon that arises when the usual uniformization technique is used to obtain the discrete-time equivalent of a continuous time system with imperfect observations. To illustrate this phenomenon consider the simple $M/M/1$ continuous tandem queues of Figure 2. This network is a Markov system whose state is
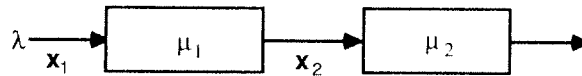

Figure 2 Uncontrolled tandem queue.

the tuple (x_1, x_2) (the number of customers at queues 1 and 2 respectively). If only x_1 can be observed and A is the infinitesimal operator for the system, the usual uniformization [3] can be accomplished by representing the system as a Markov chain with transition matrix

$$K = I + \gamma^{-1} A$$

where

$$\gamma = \lambda_1 + \mu_1 + \mu_2.$$

The diagonal elements of K are then zero for all states such that x_1 and x_2 are strictly positive. Suppose now that the system is in such a state $(x_1 > 0, x_2 > 0)$ and x_1 remains unchanged at some epoch n. Then, without observing x_2 directly one is able to conclude that x_2 has decreased by unity at this nth epoch. With less certainty one is also able to draw inferences on x_2 whenever x_1 fails to jump, no matter what the state of the process happens to be. In other words, the usual uniformization method renders it impossible to observe only x_1, since one gains information on x_2 not similarly obtained in continuous time.

Allowing simultaneous arrivals, services and transfers complicates the resulting computation somewhat, but without changing the fundamental nature of the problem and without affecting our results significantly.

A similar network consisting of two stations in tandem was studied earlier [13], except that the controller had perfect information of all the events occuring in the network. Because the complete state was known the selection of the service rate was based on the tuple $(x_1(n), x_2(n))$, consisting of the respective number of customers at the two service stations. In the proposed problem, the controller at Station 1 actually controls $\pi_2(n)$ as well as $x_1(n)$ by deciding whether (or not) to transfer a customer from Station 1 to Station 2. Consequently, our optimization problem can be viewed as follows: control $x_1(n)$ and $\pi_2(n)$ for all t to minimize a discounted cost of the form

$$J(g) = E^g \sum_{n=0}^{\infty} \beta^n [c_1(x_1(n)) + c_2(x_2(n))] \tag{3.1}$$

where $g \triangleq (g_0, g_1, \ldots)$.

We shall use $(x_1(n), \pi_2(n))$ as the information state for the optimization problem and will prove that the optimal control policy for a cost of the form (3.1) is not only bang-bang, but is even of the threshold type, that is

$$\mu^*(n) = \begin{cases} \bar{\mu} & \text{if } x(n) > l(\pi(n)) \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

3.1.2 Solution We first consider a finite horizon N, and then allow $N \to \infty$. The problem of minimizing (3.1) over a finite horizon under conditions of partial information represented by the information state $(x, \pi) \triangleq (x_1, \pi_2)$ is treated by a dynamic programming argument. Thus, we deal with the optimized value function

$$V_n^N(x, \pi) = \min E^g \left\{ \sum_{\tau=n}^{N} \beta^\tau [c_1(x_1(\tau)) + c_2(x_2(\tau))] \,\middle|\, x_1(n) = x, \pi_2(n) = \pi \right\} \tag{3.3}$$

where the minimum is taken over all admissible policies. These policies regulate the service rate $\mu(\tau)$, $n \le \tau \le N$ over the action space $[0, \bar{\mu}]$.

We write the dynamic programming equation (DPE) in two different forms. The first reveals that a bang-bang policy is optimal. The second leads to a monotone optimal policy under a plausible set of hypotheses discussed in Section 2.

Equation (3.3) and the assumptions imply a DPE of the form

$$V_{N+1}^N(x, \pi) = 0,$$

$$V_n^N(x, \pi) = c_1(x) + E^\pi [c_2(x_2(n))] + \beta \{ \Gamma_{n+1}^N(x, \pi) + \min_{\mu \in [0, \bar{\mu}]} \mu \Lambda_{n+1}^N(x, \pi) \} \tag{3.4}$$

for $x > 0$, where

$$\Gamma_{n+1}^N(x, \pi) = \lambda_1 V_{n+1}^N(x+1, \hat{\pi}) + (1 - \lambda_1) V_{n+1}^N(x, \hat{\pi}), \tag{3.5}$$

$$\hat{\pi} = \lambda_2(1 - \mu_2)\pi_+ + \lambda_2\mu_2\pi + (1 - \lambda_2)\mu_2\pi_- + (1 - \lambda_2)(1 - \mu_2)\pi, \tag{3.6}$$

$$\Lambda_{n+1}^N(x, \pi) = \lambda_1 [V_{n+1}^N(A_2(x, \hat{\pi})) - V_{n+1}^N(A_1(x, \hat{\pi}))]$$

$$+ (1 - \lambda_1)[V_{n+1}^N(A_2(x-1, \hat{\pi})) - V_{n+1}^N(A_1(x-1, \hat{\pi}))], \tag{3.7}$$

$$A_1(x, \pi) = (x+1, \pi), \quad A_2(x, \pi) = (x, \pi_+), \tag{3.8}$$

and

$$\pi_+(y) = \pi(y-1)\delta(y > 0),$$

$$(3.9)$$

$$\pi_-(y) = \pi(y+1)\delta(y>0) + [\pi(0)+\pi(1)]\delta(y=0).$$

When $x=0$, $\Lambda_{n+1}^N(0,\pi)=0$ by a direct computation so that the value of μ does not matter. This is as expected, since $x=0$ corresponds precisely to an empty Station 1; no transfer of a job from Station 1 to Station 2 is possible.

Equation (3.4) indicates that the last term of its right hand side (RHS) is minimized by $\mu=0$ if $\Lambda_{n+1}^N(x,\pi)>0$ and by $\mu=\bar\mu$ if $\Lambda_{n+1}^N(x,\pi)<0$. Therefore, there exists an optimal bang-bang solution for the subject problem.

To prove the monotonicity (threshold) property of the optimal policy we find it convenient to rewrite the second of Eqs. (3.4) according to

$$V_n^N(x,\pi) = c_1(x) + E^\pi(c_2(x_2(n))) + \beta\{\Gamma_{n+1}^N(x,\pi) + \bar\mu(\Lambda_{n+1}^N(x,\pi))^-\} \tag{3.10}$$

with the definition $\alpha^- = \min(\alpha,0)$. Applying the equality

$$(\beta-\alpha)^- = \min(\alpha,\beta) - \alpha \tag{3.11}$$

to $(\Lambda_{n+1}^N(x,\pi))^-$, letting α in (3.11) be comprised of the negative terms of $\Lambda_{n+1}^N(x,\pi)$, and observing the form of $\Gamma_{n+1}^N(x,\pi)$ and $\Lambda_{n+1}^N(x,\pi)$, we obtain $V_n^N(x,\pi)$ as the alternative expression

$$V_n^N(x,\pi) = c_1(x) + E^\pi(c_2(x_2(n))) + \beta\{(1-\bar\mu)\Gamma_{n+1}^N(x,\pi)$$

$$+ \bar\mu \min_{i=1,2} \{\lambda_1 V_{n+1}^N(A_i(x,\hat\pi)) + (1-\lambda_1)V_{n+1}^N(A_i(x-1,\hat\pi))\}\} \tag{3.12}$$

for $x>0$. As we have already explained, (3.12) is simpler for $x=0$, since the minimization argument terms do not appear (or what is equivalent, become zero) reflecting the fact that no transfer is possible.

The expression to be minimized in (3.12) suggests that for the inductive process to propagate the submodularity property discussed in Section 2, it is crucial that the minimum of two submodular functions again be submodular. Theorems 2.3 and 2.4 insure that this is indeed the case and provide the basis for the proof of the threshold property of the optimal policy.

Indeed, $V_{N+1}^N(x,\pi)$ defined by the first of Eqs. (3.4) has the property (2.7c), (as well as (2.7a) and (2.7b)), that is, $V_{N+1}^N(x,\pi)\in X$. Assuming that $V_{n+1}^N(x,\pi)\in X$, it remains to prove that $V_n^N(x,\pi)$, defined by (3.12), has the same properties as $V_{n+1}^N(x,\pi)$ for all (x,π). For this it is sufficient to prove that each of the terms involved in the sum defining $V_n^N(x,\pi)$ satisfies the properties (2.7a)–(2.7c). The instantaneous cost $c(x,\pi)=c_1(x)+E^\pi(c_2(x_2(n))$ has the abovementioned properties because $c_1(x)$ and $c_2(x)$ are convex monotone nondecreasing functions of their arguments. The term $\Gamma_{n+1}^N(x,\pi)$ is a finite sum of $V_{n+1}^N(\cdot)$ terms, with positive coefficients; consequently, $\Gamma_{n+1}^N(x,\pi)\in X$.

Finally, the last term on the RHS of (3.12) contains minima of finite sums of V_{n+1}^N, these being of the form

$$\alpha(x, \hat{\pi}) = \min \{\lambda_1 V^N_{n+1}(A_1(x, \hat{\pi})) + (1 - \lambda_1)V^N_{n+1}(A_1(x - 1, \hat{\pi})),$$

$$\lambda_1 V^N_{n+1}(A_2(x, \hat{\pi})) + (1 - \lambda_1)V^N_{n+1}(A_2(x - 1, \hat{\pi}))\}. \tag{3.13}$$

For $\alpha(x, \hat{\pi})$ the conclusions of Theorems 2.3 and 2.4 remain unchanged, and their proofs, while appearing more complex, are essentially the same and will not be detailed here. Therefore, $V^N_n(x, \pi) \in X$ for all (x, π), the proof of the induction step is complete, and the threshold property of the optimal policy is established for the finite horizon problem.

The optimal policy for the finite horizon control problem in tandem queues with imperfect information, formulated in this section, is described by

$$\mu^*(n) = \begin{cases} \bar{\mu} & \text{if } x(n) \geq l_n(\pi(n)) \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

To solve the infinite horizon problem, we must limit our consideration to convex instantaneous cost functions of polynomial order. Then, for $\beta < 1$, the infinite horizon cost $V^\infty(x, \pi)$ is finite, and, because of Theorem 1 of [12],

$$V^\infty(x, \pi) = \lim_{N \to \infty} V^N_n(x, \pi). \tag{3.15}$$

When $\beta < 1$, $V^\infty(x, \pi) \in X$ because of (3.15) and the fact that X is closed under pointwise limits. By the same argument as in the case of the finite horizon problem it is possible to show that the optimal policy for the infinite horizon problem is described by a time invariant threshold for each (x, π).

The results of the above analysis can be summarized by the following theorem:

THEOREM 3.1   *The optimal policy for the infinite horizon control problem in tandem queues with imperfect information, formulated in this section, is described by* (3.2). ∎

3.2 Routing Control for Parallel Queues

3.2.1 Problem description—features of the problem   Consider the network of Figure 3, consisting of two stations in parallel. Arrivals at the network are described by a process whose interarrival times τ_n are i.i.d. with $E(\tau_n) < \infty$. The
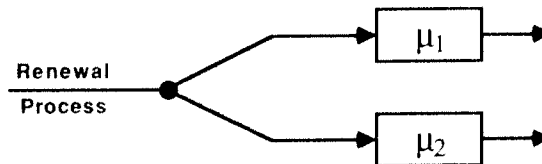

Figure 3   Input routing to two parallel queues.

services are exponential with respective rates μ_i. A controller observes perfectly the arrivals at the network and has to assign each arrival to one of the stations. Initially the controller has inferential information regarding $x_i(0)$, the number of customers at Station i at $n = 0$. This information is described by $\pi(0)$, which is the probability distribution of the number of customers in the network at $n = 0$. At each arrival epoch t_n, the controller receives noisy information about the number of customers at each station at t_{n-1}, i.e., the time of the last previous arrival. This information is described by

$$\mathbf{z}_{n+1} = h(\mathbf{x}_n, \mathbf{w}_n) \tag{3.16}$$

where $\mathbf{x}_n$ is the tuple whose kth component denotes the number of customers at the kth station at $t = t_n$, and $\mathbf{w}_n$ is the noise in the observations. It is assumed that $\{\mathbf{w}_n\}$ is independent of $\{\mathbf{x}_n\}$. Assuming that N arrivals are going to occur, the objective is to determine the controller's routing policy to minimize a cost of the form

$$J(g) = E^g \sum_{i=1}^{2} \int_0^{\infty} c_i(x_i(t))\, dt, \tag{3.17}$$

where each $c_i(\cdot)$ $(i = 1, 2)$ is a non-negative convex monotone nondecreasing function of its arguments, $g \triangleq (g_{t_1}, \ldots, g_{t_N})$, and $t_1, t_2, \ldots, t_N$ are the arrival epochs.

A network of two stations in parallel was studied earlier by [11, 19], except that the controller had perfect information of all the events occuring in the network. In the present problem the controller has only inferential information regarding $x_i(t_n)$, the number of customers at Station i at the nth arrival epoch $(n = 1, 2, \ldots, N)$. This information is expressed through the information state $\pi(t_n) \triangleq \pi(n)$, which is the probability distribution of the number of customers at the two stations, for the nth arrival epoch. The controller controls $\pi(n)$ by deciding where to route the arriving customers. Thus, our optimization problem can be viewed as: control $\pi(n)$ to minimize the cost (3.17). We shall prove that the resulting optimal control policy is of the threshold type.

3.2.2 Solution   The problem of minimizing (3.17) under conditions of imperfect information represented by the information state π is treated by a dynamic programming argument. By concentrating on the arrival epochs t_i we get the DPE

$$V_{t_N}(\pi) = \sum_{i=1}^{2} \sum_{\mathbf{x}, \mathbf{d}} \int_{t_N}^{\infty} c_i((\mathbf{x} - \mathbf{d})^+) P[D(t) = \mathbf{d}] \pi(\mathbf{x}_{t_N} = \mathbf{x} | \mathbf{y}^N)\, dt \tag{3.18a}$$

$$V_{t_{n-1}^+}(\pi) = \sum_{i=1}^{2} \sum_{\mathbf{x}, \mathbf{d}} \int_{t_{n-1}}^{t_n} c_i((\mathbf{x} - \mathbf{d})^+) P[D(t) = \mathbf{d}] \pi(\mathbf{x}_{t_{n-1}^+} = \mathbf{x} | \mathbf{y}^{n-1})\, dt$$

$$+ \min_{i=1, 2} \{ E_D[V_{t_n}(A_i D \pi)] \}, \quad n = 1, 2, \ldots, N-1 \tag{3.18b}$$

where D is a tuple whose components D_1 and D_2 denote the number of departures from the respective stations in (t_{n-1}, t_n), $\mathbf{x} = (x_1, x_2)$,

$$c_i((\mathbf{x} - \mathbf{d})^+) = c_i((x_i - d_i)^+), \tag{3.19}$$

$$\mathbf{y}^{n-1} = (z_{t_1}, \ldots, z_{t_{n-1}}, t_1, t_2, \ldots, t_{n-1}, u(t_1), u(t_2), \ldots, u(t_{n-1})), \tag{3.20}$$

and $u(t_1), \ldots, u(t_{n-1})$ denote the controller's decisions at $t_1, t_2, \ldots, t_{n-1}$ respectively.

To prove the threshold property of the optimal policy it is sufficient to show that the value function defined by (3.18a)–(3.18b) satisfies property (2.7c). We prove that this is true by induction.

To verify that (2.7c) is satisfied for $n = N$ it is sufficient to limit ourselves to an analysis of

$$g_i(\pi) = \sum_{x_i} c_i((x_i - d)^+) \pi(x_i) \tag{3.21}$$

for fixed $d \geq 0$. This is because if $g_i(\pi)$ satisfies (2.7c), then so does $V_{t_N^+}$. Note that $c_i(\cdot)$ non-negative, monotone nondecreasing and convex implies the same for

$$K_i(x_i) = c_i((x_i - d_i)^+). \tag{3.22}$$

Hence $g_i(\pi)$ satisfies (2.7a)–(2.7c) because of Theorem 2.7. To proceed with the induction we assume that $V_{t_n^+}(\pi)$ satisfies property (2.7c) and use Eq. (3.18b) to prove that $V_{t_{n-1}^+}(\pi)$ satisfies the same property. The first term on the right hand side (RHS) of (3.18b) is of the same form as the RHS of (3.18a); hence it satisfies (2.7c). The second term on the RHS of (3.18b) is of the form (2.8a); Theorem 2.3 insures that this term satisfies (2.7c). Consequently, $V_{t_{n-1}^+}(\pi)$ satisfies (2.7c) and the induction is complete.

The results of the above analysis are summarized by the following Theorem:

THEOREM 3.2 *The optimal policy for the network of parallel queues with imperfect information, formulated in this section, has the following threshold property: if for a certain $A_i\pi$, the incoming customer is assigned to Station i, the customer will also be assigned to this station when the information state is π rather than $A_i\pi$.* ∎

Submodularity of the value function $V_t(\pi)$ is a sufficient condition for the threshold property of the optimal policy to hold. Moreover, as pointed out in Section 2, inductive propagation of submodularity (2.7c) does not require the presence of either monotonicity (2.7a) or convexity (2.7b). Nevertheless, it is possible to prove that the value function $V_t(\pi)$ satisfies properties (2.7a)–(2.7b). The proof of monotonicity is immediate and will receive no further attention.

To prove convexity we start at $n = N$ and note that $g_i(\pi)$ satisfies property (2.7b). Thus, $V_{t_N}(\pi)$ is convex. To proceed with the induction we assume that $V_{t_n}(\pi)$ satisfies property (2.7b) and use Eq. (3.18b) to prove that $V_{t_{n-1}}(\pi)$ satisfies the same property. The first term on the RHS of (3.18b) is of the same form as the RHS of (3.18a). Hence it satisfies (2.7b). The second term on the RHS of (3.18b) is of the

form (2.8a). Theorem 2.6 insures that this term satisfies (2.7b). Therefore, $V_{t,n-1}^+$ satisfies (2.7b) and the induction is complete.

4. CONCLUSIONS

Optimal routing policies for queueing networks under conditions of imperfect information are of the threshold type under conditions specified in Section 2. The examples of Section 3 indicate that these conditions are realized in useful queueing models beyond those considered heretofore in the literature. Moreover, knowledge that the optimal policy is of threshold type can guide its computation, as it reduces the search over functions described by a threshold; nevertheless, computation of the optimal policy remains a very difficult and challenging problem.

It is our conjecture that the notion of information state in dynamic programming can be applied to a wider range of problems, including especially decentralized routing control of certain queueing networks. The use of inequalities such as those of Section 2 will then lead to an elucidation of monotonicity and convexity properties for the optimal policy and the resultant expected cost.

Acknowledgements

References

[1] C. W. Bell and S. Stidham Jr., Individual versus social optimization in the allocation of customers to alternative servers. *Management Science* **29** (1983), 831–839.

[2] D. P. Bertsekas, *Dynamic Programming and Stochastic Control*, Academic Press, New York, 1976.

[3] E. Cinlar, *Introduction to Stochastic Processes*, Prentice-Hall, Englewood Cliffs, NJ, 1975.

[4] E. Davis, *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, NC, 1977.

[5] A. Ephremides, P. Varaiya and J. C. Walrand, A simple dynamic routing problem, *IEEE Trans. Auto. Control* **AC-25** (1980), 690–693.

[6] B. Hajek, Optimal control of two interacting service stations, *IEEE Trans. Auto. Control.* **AC-29** (1984), 491–499.

[7] B. Hajek, Extremal splitting of point processes, *Math. of Operations Research* **10** (1985), 543–556.

[8] M. T. Hsiao and A. A. Lazar, *Optimal Decentralized Flow Control of Markovian Queueing Networks with Multiple Controllers*, School of Engineering and Applied Science, Columbia University, New York, 1986.

[9] M. T. Hsiao and A. A. Lazar, *Optimal Flow Control of Multi-Class Queueing Networks with Decentralized Information*, Department of Electrical Engineering, Columbia University, New York, 1986.

[10] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[11] 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.

[12] S. A. Lippman, Semi-Markov decision processes with unbounded rewards, *Management Science* **19** (1973), 717–731.

[13] Z. Rosberg, P. Varaiya and J. C. Walrand, Optimal control of service in tandem queues, *IEEE Trans. Auto. Control* **AC-27** (1982), 600–610.

[14] S. Ross, *Stochastic Processes*, Wiley, New York, NY, 1983.

[15] A. Segall and N. R. Sandell Jr., Dynamic file assignment in a computer network—Part II: decentralized control, *IEEE Trans. Auto. Control* **AC-24** (1979), 709–715.

[16] S. Stidham Jr., *Optimal Control of Admission, Routing and Service in Queues and Networks of Queues*, Department of Industrial Engineering, North Carolina State University, Raleigh, NC, 1984.

[17] S. Stidham Jr., Optimal control of admission to a queueing system, *IEEE Trans. Auto. Control* **AC-30** (1985), 705–713.

[18] S. Stidham Jr., *Scheduling, Routing and Flow Control in Stochastic Networks*, Tech. Report UNC/ORSA/TR-86-22, University of North Carolina at Chapel Hill, Chapel Hill, NC, 1986.

[19] J. Walrand, A note on optimal control of a queueing system with two heterogeneous servers, *Systems and Control Letters* **4** (1984), 131–134.

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

[21] R. Weber and S. Stidham Jr., Optimal control of service rates in networks of queues, *Adv. Appl. Prob.* **19** (1987), 202–218.

[22] W. Whitt, Deciding which queue to join: Some counterexamples, *Operations Research* **34** (1986), 55–62.