AN M/M/1 DYNAMIC PRIORITY QUEUE WITH OPTIONAL
PROMOTION

Ehud Kofman, Lecturer, Department of Business
Administration
Steven A. Lippman, University of California, Los
Angeles

#524

.

AN M/M/1 DYNAMIC PRIORITY QUEUE WITH OPTIONAL
PROMOTION

Ehud Kofman, Lecturer, Department of Business
Administration
Steven A. Lippman, University of California, Los
Angeles

#524

.

Summary:

We consider an M/M/1 queue with two types of customers: priority customers and
regular customers. They arrive at the service facility according to two independent
Poisson streams and form a single queue according to the order in which they arrive.
The two types of customers are distinguished by the holding costs charged per unit
time that each of them resides in the queue. The server can either serve customers
according to the order in which they arrive or pay a fixed fee R and promote a
priority customer, bypassing the customers ahead of him. The server selects the
customers to be served so as to minimize the expected average cost per unit of time
of operating the system. We show that whenever the number of regular customers
bypassed in a promotion times the expected holding costs per priority customer
per service period is greater than or equal to R, promotion is strictly optimal.
Moreover, for each state there exists a value of R, with R exceeding the number of
regular customers bypassed in a promotion times the expected holding costs per
priority customer per service period, for which promotion is optimal. This result
contradicts previous work in the literature. In addition we demonstrate that the
set of states from which promotion is optimal decreases in the sense of set inclu-
sion as R increases. This fact is the key to an efficient algorithm.

Acknowledgment:

AN M/M/1 DYNAMIC PRIORITY QUEUE WITH OPTIONAL PROMOTION

Most previous works in optimization of priority queueing systems
have been concerned primarily with the important issues of (i) assign-
ing priorities to different classes of customers, and (ii) evaluating
the performance of queueing systems under various regimes, using the
usual parameters of queueing systems like waiting time and delay time.
For example, Cobhan [4] treated the problem of assigning priorities to
each of k classes of customers (classified on the basis of their
respective holding costs and required service times) in an M/G/1
system. In Oliver and Pastalozzi's work [14] the basis for class-
ification of customers is the required service time, determined when
the customer arrives at the queue. The objective in those works was
to establish an optimal priority rule: to assign priorities to the
different classes of customers in a manner which will cause the system
to perform optimally. The main feature of this type of work is that the rule
being sought is static; once priorities are assigned they don't change.

Harrison [9] considered a problem of dynamic scheduling in an
M/G/1 queueing system so as to minimize the expected present value of
rewards received minus costs incurred over an infinite horizon, where
future costs and rewards are continuously discounted. In his model
customers from different classes incur different holding costs, receive
different rewards upon service completion, and possess different service
time distributions. Given the state of the system, the controller's

1

problem is to decide, at the completion of each service, which class to admit next into service. In Harrison's model there is, surprisingly, a static rule which is optimal, but assigning priorities to the different classes involves a very complex procedure. An excellent bibliography covering the vast literature on Priority queues can be found in Jaswal [10].

In contrast to previous works [5], Bell [2] treated an M/G/1 where priorities change dynamically. There are two types of customers: priority customers and regular customers. They arrive at the service facility according to two independent Poisson streams, and they form a single queue according to the order in which they arrive. The two types of customers are distinguished by the holding costs charged per unit time that each of them resides in the queue. The queue controller can either serve customers according to the order in which they arrived or pay a fixed fee R and promote a customer (with higher holding cost) thereby bypassing the customers ahead of him. Naturally, the optimal determination whether or not a priority customer will be promoted depends on the state of the system. It is in this sense that priorities change dynamically.

We treat Bell's problem under the assumption of exponential service and two cases: in the constrained case if the server decides to promote, he must select the priority customer who arrived first. In the unconstrained case, no such restriction exists. The criterion of optimality is average cost per unit of time.

In section 1 we formulate the problem as a standard Markov Decision Process (see [3,6]) and verify that the rule which from every state minimizes the expected cost of the remainder of the busy period is optimal. It is also demonstrated that the optimal rule is a limit (in the sense of Ross [15]) of a sequence of rules optimal for the problem with discounting.

In section 2 it is shown that the option to promote will be utilized to reduce the cost of operating the system. And, indeed, it will be proven that whenever the number of regular customers bypassed in a promotion times the expected holding cost per priority customer per service period exceeds the cost R of promotion, it is then optimal to promote.

It was previously thought that promotion is optimal only when the above condition is met. Even when R exceeds the number of regular customers bypassed times the expected cost of holding a priority customer for one service period, it may still be optimal to promote. This is discussed in section 3. Next a necessary condition for promotion to be optimal is provided. We follow this by presenting an example that shows this condition is not sufficient.

We conclude in section 4 by proposing a finite algorithm which produces the optimal rule when an upper bound is placed on the queue length. The algorithm depends upon the fact that the set of states from which promotion is optimal decreases, in the sense of set inclusion, as the promotion fee R is allowed to increase.

In this section we formulate our model as a Markov Decision process; examine two cases of our model, the constrained case and the unconstrained case; and provide results which demonstrate the optimal rule can be obtained as a limit from consideration of the discounted version of the problem.

In our M/M/1 queueing system two types of customers, labeled type "1" (priority) and type "2" (regular), arrive at the queue according to two independent Poisson processes with rates λ_1 and λ_2, respectively. Both types of customers require the same service whose exponential distribution has parameter μ. For the continuous time Markov Process generated by this system, the time between transitions when the server is busy is exponential with parameter $\lambda_1 + \lambda_2 + \mu$. Given that transition occurs, the probability of each type of event is as follows: the probability of arrival of a "1", arrival of a "2", or departure is, respectively, $\lambda_1/(\lambda_1 + \lambda_2 + \mu)$, $\lambda_2/(\lambda_1 + \lambda_2 + \mu)$, $\mu/(\lambda_1 + \lambda_2 + \mu)$. Two kinds of costs are incurred. There is a holding cost of h_1 (h_2) per unit time per "1" ("2") in t' queue. Utilizing an ingeneous but simple argument Bell [1] showed that we can, without loss of generality, assume that there is a holding cost of $h = h_1 - h_2 > 0$ per unit time per "1" customer in the queue and a zero cost for holding a "2". Second, each time a customer is served who is not at the head of the line, a fixed charge of $R > 0$ is incurred.

Immediately after service completion, the server must decide which customer to serve next. We denote his action by F (for first) if he decides to serve the customer at the head of the line next. If, on the other hand, he decides to serve a "1" who is not at the head of the line, thereby bypassing all of the customers ahead of him, we denote his action by P (for promotion). The action space is, therefore, given by $K = \{P, F\}$. Note that if there is only one type of customer present, then there is no decision to make.

We distinguish two cases of our model. The first case, which we call the constrained case, is characterized by the fact that if the server elects to promote a "1", he is constrained to promote the customer who, amongst the "1"'s presently in the queue, was the first to arrive. In the second case, referred to as the unconstrained case, the above constraint does not apply and the server has complete freedom to decide which "1" to promote. In the unconstrained model the action P is not yet well defined. But, as can easily be shown (see [11, pp. 13-15]), there is no loss of generality in assuming that when the server promotes he always promotes the "1" who came last. Thus, if there is to be a promotion from state s, in both cases the "1" who will be promoted is uniquely determined. We shall have need to refer to this customer and, accordingly, we label him 1_s.

The state space S is defined by

(1) $$S \equiv \{s : s = \langle s_i \rangle_{i=1}^{n} \ s_i \in \{1,2\}, \text{ for all } 1 \le i \le n, \ n = 0,1,2,\ldots\} .$$

That is to say, S is composed of elements each of which represents a queue, i.e., a number of customers, their type, and the order in which they arrived at the service facility. Specifically, $s_i = 1$ means that among the customers of s, the ith to arrive was a "1"; in particular, s_1 is the customer type at the head of the line.

If action P is taken at s, then the state of the system after one service period will be $s_P Y$, $Y \in S$. The random vector Y represents the number and order of arrivals which take place during one service period. The vector s_P represents the queue 1_s left behind when he was promoted. Similarly, if F is taken at s, then s_F is the queue left behind by the customer served, and at the end of his

service (i.e. one period later) the state of the system is $s_F Y$. Due
to the exponential nature of the service and the linearity of the hold-
ing cost, the expected holding cost associated with the service period
that commences when action P is chosen at s is $b(s_P)h/\mu +$
$E[b(Y)]h/\mu$, where $b(s)$ is the number of priority customers in s. On
the other hand if action F is taken at s, then the associated expected
holding cost is $b(s_F)h/\mu + E[b(Y)]h/\mu$.

The objective is to find a rule which will minimize the average
cost per unit time for the infinite horizon. We formulate this problem
as a discrete-time Markov Decision Process based on the embedded chain
which results from observing the system only at the times of service
completion. Henceforth, units of time will be service periods, and, as
before, the state of the system is the queue observed after service completion.

To determine the transition probabilities we notice that if at state s
action F(P) is taken the system moves next to $s_F Y(s_P Y)$, $Y \in S$. As before,
$Y = (X_1, \ldots, X_N)$ is a random vector, where N is a geometric random variable
with parameter $\mu/(\lambda_1 + \lambda_2 + \mu)$. For any i, $1 \le i \le N$, $X_i = j$ if the ith
interval during a service period is a "j", $j = 1,2$. Given that $N = n$, the
X_i are independent random variables with $P(X_i = 1) = \lambda_1/(\lambda_1 + \lambda_2)$ and
$P(X_i = 2) = \lambda_2/(\lambda_1 + \lambda_2)$. To ensure that the embedded Markov Chain is positive
recurrent we assume that $\mu > \lambda_1 + \lambda_2$. (There is no need to specify the probability
of transition from the empty state to a busy one because, as will later be shown,
the search for the optimal rule can be restricted to one busy period.)

Henceforth, we will assume, with no loss of generality that $h/\mu = 1$. Hence,
if promotion is taken at s and the next state reached is $s_P Y$, then $c(s,P)$,
the (immediate) one-period expected cost, is given by

(2a) $\qquad c(s,P) = R + b(s_P) + E(b(Y))$.

Similarly, the one-period expected cost of taking action F at s is

(2b) $\qquad c(s,F) = b(s_F) + E(b(Y))$.

In other words, the (one period) cost is the cost of promotion plus the expected number of "1"'s found in the queue at the end of the service period. Therefore, it is easy to see that $c(s,P) - c(s,F) = R - 1$. Moreover, as a result of promotion at s, the next state reached contains one less priority customer. In this sense s_P is a better state than s_F.

For any time i let H_i be the history of states and decisions up to and including time i. A rule π is a sequence of functions $\langle f_i(h_{i-1},s) \rangle_{i \geq 1}$ with values in the action space K, where h_{i-1} is a given history and s is the state at time i. Let $V_\pi(s)$ represent the expected cost of using π starting at s until the busy period ends, and denote by $V(s)$ the expected cost of the s-residual busy period using the optimal rule, where

$$V(s) = \inf_\pi V_\pi(s).$$

Let Ω be the underlying probability space whose points are sequences of arrivals and departures. For any fixed state s_0 let $T(s_0)$ be the time until the system empties after starting at s_0. (Of course, the distribution of $T(s_0)$ is independent of π but does depend upon the queue length of s_0.) Let $\langle S^\pi(i) \rangle_{i=0}^{T(s_0)}$ be the random sequence of states visited by the system under π when starting at s_0 and ending $T(s_0)$ units of time later $(S^\pi(0) = s_0)$. For any $\omega \in \Omega$, $\langle s^\pi(i,\omega) \rangle_{i=0}^{T(s_0,\omega)}$ is a realization of $\langle S^\pi(i) \rangle_{i=0}^{T(s_0,\omega)}$, and we denote such a

realization by $\omega(\pi)$. Let $C(\pi,\omega(\pi))$ be the cost of the s-residual busy period when using π and ω occurs. Formally, then,

$$V_{\pi}(s) = \int_{\Omega} C(\pi,\omega(\pi))P(d\omega).$$

Whenever ω is fixed it will be omitted from the notation.

Both cases of our model constitute a Markov Decision Process with unbounded rewards [8,12,13]. Application of Denardo's N-stage contraction [6] to these (and other) models in which the one-period cost function increases "moderately" in the state and the transition function is also well behaved enables us to conclude that there is a stationary (deterministic) rule which is optimal in the presence of discounting (see [13]). Then, using limit procedures reminiscent of those used by Taylor [17] and Ross [15], it can be shown [13] that there exists a stationary rule which is average-optimal. The assumptions on which this treatment is predicated hold in our case and can be verified easily [13, p. 1232]. Based on the existence of a stationary (deterministic) rule which is average-optimal, we begin by showing that π^*, the rule which from every state s minimizes the expected cost of the s-residual busy period, is average optimal.

Let B_i and I_i denote the length of the ith busy period and idle period respectively. Also, denote by W_i^{π} the cost of using the rule π during the ith busy period. For a stationary rule π, $\{B_i\}_{i \geq 1}$, $\{W_i^{\pi}\}_{i \geq 1}$ and $\{I_i\}_{i \geq 1}$ are three sequences of i.i.d. random variables. Furthermore, the distributions of B_i and I_i are rule-independent. The cost W_i^{π}, however, depends on π as indicated by the superscript. Because the above random variables have finite expectations, $\varphi(s,\pi)$, the average cost per unit of time when starting from s and using π, is well defined for every stationary rule π. Moreover, the requirement $\mu > \lambda_1 + \lambda_2$ ensures that the system

empties in a finite expected time so that ϕ has the representation (see Ross [16]),

$$\varphi(0,\pi) = \frac{\displaystyle\lim_{n\to\infty} \sum_{i=1}^{n} W_i^{\pi}}{\displaystyle\lim_{n\to\infty} \sum_{i=1}^{n} (B_i + I_i)} .$$

Furthermore, $\varphi(s,\pi) = \varphi(0,\pi)$ for every s. Now, by the strong law of large numbers, we have

$$\varphi(0,\pi) = \frac{\displaystyle\lim_{n\to\infty} \sum_{i=1}^{n} W_i^{\pi}/n}{\displaystyle\lim_{n\to\infty} \sum_{i=1}^{n} (B_i + I_i)/n} = \frac{EW_1^{\pi}}{E(B_1) + E(I_1)} .$$

Since $E(B_1) + E(I_1)$ does not depend on π, we observe that the rule which minimizes $E(W_1^{\pi})$ is the optimal rule. Hence it is clear that for the average cost criterion, the best stationary rule is the rule which minimizes the expected cost in a busy period.

Let π_s and π^* denote the stationary rule which minimizes the expected cost of the s-residual busy period and the rule which at s assigns the action $\pi_s(s)$. Because there exists a stationary rule which is optimal [13, p. 1232], it follows by the principle of optimality that the stationary rule π^* is an optimal rule.

THEOREM 1. The rule π^* is average optimal.

We now proceed to prove that π^* is a limit of a sequence of discount optimal rules. Consider the case where costs are discounted and all costs are incurred at the beginning of the appropriate period, and denote by $\psi_{\pi}(s,\beta)$ the expected

β-discounted cost of starting at s and using the rule π. Throughout the discussion we will assume (with no loss of optimality) that π_β, the β-optimal rule, is stationary and minimizes the β-discounted expected cost of the s-residual busy period, each s. Promotion is said to be optimal at a given state s if for any rule which does not promote at s there exists a rule which does and is at least as good. Promotion will be said to be β-optimal at s if for any rule which does not promote at s there is a rule which does promote at s and is as good.

Following Ross [15] we say that a sequence $\langle \pi_r \rangle_{r \geq 1}$ of stationary (deterministic) rules converges to a stationary (deterministic) rule π, if for each s there exists an integer r_s such that $\pi_r(s) = \pi(s)$ for all $r \geq r_s$. Because the state space is countable, diagonalization can be utilized to extract a convergent subsequence. Using this we now prove

THEOREM 2. The rule π* is a limit point of $\{\pi_\beta: 0 < \beta < 1\}$, the set of β-optimal rules.

Proof. Let $\langle \pi_{\beta_r} \rangle$ be a convenient subsequence, and denote by σ its limit. We claim that from any s, σ minimizes the expected cost of the s-residual busy period. If not, then $V_{\pi*}(s) < V_\sigma(s)$ for some state s. This implies that there exists an ε > 0 and r_0 large enough such that for $r \geq r_0$

$$\psi_{\pi*}(s,\beta_r) < \psi_\sigma(s,\beta_r) - \varepsilon \quad .$$

On the other hand, the optimality of π_β guarantees that for any β, 0 < β < 1, $\psi_{\pi*}(s,\beta) > \psi_{\pi_\beta}(s,\beta)$. In particular, we have

(3) $\psi_\sigma(s,\beta_r) - \varepsilon > \psi_{\pi*}(s,\beta_r) \geq \psi_{\pi_\beta}(s,\beta_r) ,$ for $r \geq r_0 .$

We claim

(4)
$$\lim_{r \to \infty} \left| \psi_\sigma(s, \beta_r) - \psi_{\pi_{\beta_r}}(s, \beta_r) \right| = 0 .$$

Since (4) contradicts (3), the proof is complete upon establishing (4).

In order to verify (4) we consider the set S_N of states with queue length less than or equal to N. For any finite N we can pick r_0 large enough so that for $r \geq r_0$, σ and π_{β_r} take the same actions in S_N. By the Gambler's Ruin Problem, it follows that from any state s with queue length $n(s)$ the probability p_N of leaving S_N before reaching the empty state is given by

$$p_N = \frac{1 - [\mu/(\lambda_1 + \lambda_2)]^{n(s)}}{1 - [\mu/(\lambda_1 + \lambda_2)]^N} ,$$

a quantity which converges geometrically to zero.

Because σ and π_{β_r} take the same actions while in S_N, the first state $\hat{\imath}$ S_N^c entered is the same for both policies. Label this state s^N. The number $n(s^N)$ of customers in the queue at s^N is bounded by $N + n(Y)$.

Recall that $T(s)$ represents the duration of the s-residual busy period when the state at time zero is s. Then the number of customers in the queue at time $t < T(s)$ is given by $n(s) + \sum_{i=1}^{t} n(Y_i) - t$ so that

$$V_\pi(s) \leq E \sum_{t=0}^{T(s)} [n(s) + \sum_{i=1}^{t} n(Y_i) - t + R]$$

$$= [n(s) + R] E(T(s)) + E[(T(s))(T(s)+1)/2] \left[\frac{\lambda_1 + \lambda_2}{\mu} - 1 \right] ,$$

as $n(Y_i)$ is a geometric random variable with parameter $\mu/(\lambda_1 + \lambda_2 + \mu)$. Furthermore, if $<Z_i>$ are i.i.d. random variables with the same distribution

as $T(s_0)$ when $n(s_0) = 1$, then $T(s)$ has the same distribution as $\sum_{i=1}^{n(s)} Z_i$.

Because $\mu > \lambda_1 + \lambda_2$, EZ_1 and EZ_1^2 are both finite (see [7, p. 198]).

Since $n(s^N) \leq N + n(Y_1)$, these facts enable us to conclude

$$|\psi_\sigma(s,\beta_r) - \psi_{\pi_{\beta_r}}(s,\beta_r)| \leq p_N[\psi_\sigma(s^N,\beta_r) + \psi_{\pi_{\beta_r}}(s^N,\beta_r)]$$

$$\leq 2p_N\Bigg\{ (N+EY_1+R)(N+EY_1)EZ_1 + \frac{1}{2}\left[\frac{\lambda_1+\lambda_2}{\mu} - 1\right]$$

$$[N^2+N(N+EY_1)EZ_1 + (E(Y_1^2)-EY_1)E(Z_1)^2 + EY_1E(Z_1^2) + (N+EY_1)EZ_1]\Bigg\} \to 0$$

as $N \to \infty$.

Q.E.D.

2. STRUCTURE OF THE OPTIMAL POLICY

Our first objective is to show that the option to promote can be utilized
to reduce the cost of operating the system. In other words, we will show that
FIFO, the rule which never uses P, is not optimal. Toward this end we define
the sets A^+ and A by

(5) $A^+ = \{s \in S: R < k(s)\}$ and $A = \{s \in S: R \leq k(s)\}$,

where $k(s)$ is the number of "2"'s bypassed when promotion is taken at s.
(Of course $k(s)$ differs according to whether or not we are in the constrained
or unconstrained cases, and $k(s) = 0$ when there are no priority customers in s,

THEOREM 3. If s is in A^+, then promotion is strictly optimal at s.

Proof. Let π be any rule which does not promote at s. Let $\tau(\omega)$ be the time
$\omega(\pi)$ first enters s, and let $\hat{\tau}$ be the first time after τ at which 1_s is
served under π.

Let π' be the non-stationary rule defined as follows:

(6) $\pi'(s,i) \equiv \begin{cases} \pi(s) & i < \tau \\ P & i = \tau \\ \pi(s^\pi(i)) & \tau < i \leq \hat{\tau} \\ F & i = \hat{\tau} \\ \pi(s) & i > \hat{\tau}, \end{cases}$

where $\pi'(s,i)$ is the action taken by the rule π' at time i, and
$s^\pi(i)$ is the state visited by $\omega(\pi)$ at time i.

We now compare $C(\pi,\omega(\pi))$ with $C(\pi',\omega(\pi'))$. If $\omega(\pi)$ contains no entry into s, then, obviously, $C(\pi',\omega(\pi')) = C(\pi,\omega(\pi))$. If an entry into s does occur, then there are two cases to consider: (i) l_s is eventually promoted under π, and (ii) l_s is served when it reaches the head of the line under π.

In case (i),

$$(7) \qquad C(\pi,\omega(\pi)) - C(\pi',\omega(\pi')) = \sum_{i=\tau}^{\hat{\tau}(\omega)} [c(s^\pi(i),\pi(s^\pi(i)))$$

$$- c(s^{\pi'}(i),\pi'(s^{\pi'}(i)))]$$

$$= -(R - 1) + \sum_{i=\tau+1}^{\hat{\tau}-1} 1 + R = \hat{\tau} - \tau > 0.$$

In case (ii) we have

$$(8) \qquad C(\pi,\omega(\pi)) - C(\pi',\omega(\pi')) = -R + 1 + \sum_{i=\tau+1}^{\hat{\tau}-1} 1$$

$$= -R + (\hat{\tau} - \tau) \geq k(s) - R > 0.$$

$$Q.E.D.$$

Coupling the fact that in A^+ promotion is optimal with the fact that during every busy period under every rule an entry into A^+ occurs with positive probability it follows from Theorem 3 that FIFO is not optimal.

It was previously thought [2] that A^+ includes all the states from which promotion is strictly optimal. Surprisingly, this is not true. We will now show that if $k(s) = R$ promotion is not only as good as F but rather it is strictly better.

THEOREM 4. If $k(s) = R$, then promotion is strictly optimal at s.

Proof. We will first prove the theorem for unconstrained cases.

Let s be any state such that $k(s) = R$ and let π be any rule which does not promote at s. As a result of Theorem 3, we can assume that π does promote in A^+. Let π' be the nonstationary rule defined in terms of π as in (6). An argument identical to that of Theorem 3 implies that

$$(9) \qquad C(\pi, \omega(\pi)) \geq C(\pi', \omega(\pi')).$$

We now show that there is a subset of Ω for which (9) holds with strict inequality.

The first case we consider is that π is such that the subset of Ω at which 1_s will eventually be promoted under π is of positive probability. In this case

$$C(\pi', \omega(\pi')) - C(\pi, \omega(\pi)) = (\hat{\tau} - \tau) > 0, \quad \text{as per (7).}$$

The second case we consider is that π is such that with probability one 1_s is served only when it reaches the head of the line. But, there is always a positive probability that $\omega(\pi)$ enters A^+ at time $\tau + 1$, whence, with positive probability, we have

$$(10) \qquad (\hat{\tau} - \tau) \geq k(s) + 1 > R.$$

For, as soon as $\omega(\pi)$ enters A^+, π will call for the promotion of a priority customer other than 1_s, hence, these promotions will cause 1_s to be delayed in the queue for more than

$k(s)$ periods. Since $(\hat{\tau} - \tau)$ is the number of periods after τ that 1_s is delayed in the queue, the proof for the unconstrained case follows from (10).

We now treat the constrained case. If $k(s) = R$, then s is of one of the forms: $s = [2 \overset{R}{\ldots} 2] [1 \overset{j}{\ldots} 1] 2 \ldots$, or $s = [2 \overset{R}{\ldots} 2] [1 \overset{j}{\ldots} 1]$ for some $j \geq 1$. For $m = 1, 2, \ldots, j$ we denote by 1_s^m the priority customer in s who was the mth to arrive. Let τ be the time $\omega(\pi)$ enters s, and let $\hat{\tau}_1$ be the time after τ that 1_s^j is served under π. We define the non-stationary rule $\pi'(s,i)$ as follows:

$$(11) \qquad \pi'(s,i) = \begin{cases} \pi(s) & i < \tau \\[2ex] P & \tau \leq i < \tau + j \\[2ex] F & \tau + j \leq i \leq \hat{\tau}_1 \\[2ex] \pi(s) & i > \hat{\tau}_1. \end{cases}$$

It is easy to see that $C(\pi, \omega(\pi)) \geq C(\pi', \omega(\pi'))$ and the inequality is strict if and only if any of the 1_s^m, $m = 1, 2, \ldots, j$ is eventually promoted under π. Since there are rules under which the probability of such an eventuality is zero we proceed as follows.

Let τ_1 be the first time after τ A^+ is entered under π', and if $\tau_1 \leq \tau(s)$ (which occurs with positive probability) let s^+ be the state where entry into A^+ occurs. Also, let $\hat{\tau}_2$ be the time 1_{s^+} is served under π.

Define τ by

$$\hat{\tau} \equiv \hat{\tau}_1 \, \chi_{[\tau_1 \neq \tau + j]} + \hat{\tau}_2 \, \chi_{[\tau_1 = \tau + j]},$$

where χ_E is the indicator function of E. Using $\pi'(s,i)$ we now define the non-stationary rule $\pi''(s,i)$ as follows:

$$(12) \qquad \pi''(s,i) = \begin{cases} \pi'(s,i) & \text{if} \quad \tau_1 \neq \tau + j \\[6pt] \pi'(s,i) & i < \tau + j, \qquad \tau_1 = \tau + j \\[6pt] P & i = \tau + j, \qquad \tau_1 = \tau + j \\[6pt] F & \tau + j + 1 \leq i \leq \hat{\tau}_2 \\[6pt] \pi'(s,i) & i > \hat{\tau}_2. \end{cases}$$

It is clear that if $\omega(\pi')$ enters A^+ at time $\tau + j$, then

$$(13) \qquad C(\omega(\pi),\pi) - C(\omega(\pi''),\pi')$$

$$\geq \min[k(s^+) - R), \, (\hat{\tau} - (\tau + j))] > 0.$$

But, the ω-set such that $\omega(\pi')$ enters A^+ at time $\tau + j$ is of positive probability, hence the theorem.

$$\text{Q.E.D.}$$

Of course, the proof of Theorem 4 suffices, mutatis mutandis, to assert that promotion is strictly β-optimal from any state in $A_\beta = \left\{ s : R \leq \sum_{i=1}^{k(s)} \beta^{i-1} \right\}$.

3. A NECESSARY CONDITION FOR THE OPTIMALITY OF PROMOTION

As revealed in Theorem 4 the set S_0 of states from which promotion is optimal contains A. We now show that this set is larger than A. Thus, membership in A is not a necessary condition for promotion to be optimal. We then present a necessary condition. This necessary condition is not, however, sufficient, as will be evidenced in a counterexample.

It is clear that A^+, A, and S_0 depend on R. To emphasize this dependence denote these sets by $A^+(R)$, $A(R)$, and $S_0(R)$ respectively. Similarly, we shall write $V_\pi(s,R)$ and $V(s,R)$ rather than $V_\pi(s)$ and $V(s)$. Whenever R is varying the optimal rule will be denoted by π_R rather than π^*. (When R is fixed the subscript R will be omitted.) We shall make extensive use of the nonstationary rule π_R^1 whose first action is P (regardless of the state of the system) followed by π_R, and the nonstationary rule π_R^2, whose first action is F followed by π_R.

THEOREM 5. For every s there exists an R strictly larger than k(s) such that π_R promotes at s.

Proof. If $R = k(s)$, then by Theorem 4 promotion is strictly optimal at s so that

$$(13) \qquad V_{\pi_{k(s)}^1}(s,k(s)) < V_{\pi_{k(s)}^2}(s,k(s)).$$

If we could establish the continuity of $V_{\pi_R^i}(s,R)$, $i = 1,2$, then (13) would enable us to assert that there exists an $R_1 > k(s)$ such that

$$(14) \qquad V_{\pi_{R_1}^1}(s,R_1) < V_{\pi_{k(s)}^2}(s,k(s)) < V_{\pi_{R_1}^2}(s,R_1),$$

where the last inequality is a direct consequence of the fact that $V(s,R)$ increases in R.

To show that $V_{\pi_R^i}(s,R)$, $i = 1,2$, are continuous in R observe that

$$V_{\pi_R^1}(s,R) = R + b(s_p) + E[b(Y_1)] + \sum_{y_1} V(s_p y_1,R)P(Y_1 = y_1).$$

For any vector y

$$V(s_p y,R) < V_{FIFO}(s_p y) < d(s_p y),$$

where $d(s_p y)$ is a positive constant which depends on $s_p y$ but not on R. Hence,

$$\sum_y V(s_p y),R)P(Y_1 = y) \leq \sum_y d(s_p y)P(Y_1 = y) < \infty.$$

But, since for each y, $V(s_p y,R)P(Y_1 = y) \geq 0$ it follows from the Weierstrass M-test Theorem that $\sum V(s_p y,R)P(Y_1 = y)$ converges uniformly in R. Coupling this fact with the continuity of $V(s_p y,R)$ for each y guarantees that the sum $\sum V(s_p y,R)P(Y_1 = y)$ converges to a continuous function.

Q.E.D.

In terms of our cost structure, Theorem 5 shows that it is optimal to promote even when there is a chance that the cost of promotion will never be recovered (in the form of reduced holding costs). In contrast, it had been claimed ([2], Theorem 2, page 782) that it is optimal to promote only when it is certain that the cost of promotion will be recovered.

As is clear, a necessary and sufficient condition for promotion to be strictly optimal at s is

$$(14) \qquad V_{\pi_R^1}(s,R) < V_{\pi_R^2}(s,R).$$

In analogy to this comparison consider the nonstationary rule $\pi^{\tau,1}$ which always promotes (if possible) for a random number τ of periods and then acts optimally and compare this with the nonstationary rule $\pi^{\tau,2}$ that uses FIFO for the same random number τ of periods and then acts optimally.

A state s is said to be __promotable__ if there exists an integer-valued stopping time $\tau \leq T(s)$, where τ is defined with respect to the embedded Markov chain, such that

$$(15) \qquad V_{\pi^{\tau,1}}(s) < V_{\pi^{\tau,2}}(s),$$

We will now show that promotability is a necessary condition for promotion to be optimal.

THEOREM 6. Promotion is strictly optimal at s only if s is promotable.

__Proof.__ Denote by π_0 the stationary rule which promotes whenever possible and suppose promotion at s is optimal. If $\tau \equiv 1$ then $\pi_R^{\tau,i} = \pi_R^i$, $i = 1,2$. Thus

$$V(s) = V_{\pi^{\tau,1}}(s) = V_{\pi^1}(s) < V_{\pi^2}(s) = V_{\pi^{\tau,2}}(s).$$

Q.E.D.

However, promotability is not a sufficient condition as will be shown next.

EXAMPLE. Consider the constrained case and let $s = 21[2 \ldots 2]1$ be the initial state with $R > 1$ and $j > 3(R - 1)$. Let τ be the smallest integer such that $\tau \geq j - R + 2$. For each ω, $s^{\pi_R^{\tau,1}}(0,\omega) = s$ and $s^{\pi_R^{\tau,1}}(1,\omega) \in A^+$, it follows from the definition that during the first τ periods $\pi_R^{\tau,1}$ and π_0 take the same actions, namely, promote whenever feasible. On the other hand $s^{\pi_R^{\tau,2}}(0,\omega) = s$ and $s^{\pi_R^{\tau,2}}(2,\omega) \in A^+$ for each ω. As was seen in Theorem 3, every priority customer who is promoted under $\pi_R^{\tau,1}$ and not under $\pi_R^{\tau,2}$ causes $V_{\pi_R^{\tau,1}}(s,R) - V_{\pi_R^{\tau,2}}(s,R)$ to decrease. Therefore, $V_{\pi_R^{\tau,1}}(s,R) - V_{\pi_R^{\tau,2}}(s,R)$ assumes its largest possible value when no arrivals of "1"'s occur during the first τ periods. Hence,

$$V_{\pi_R^{\tau,1}}(s,R) - V_{\pi_R^{\tau,2}}(s,R) \leq (R - 1) + (R - 1) - (j - R) < 0,$$

and s is thus promotable.

To show that promotion is not optimal at s we let σ be the first time there are no priority customers in the queue when π_0 is used. Fix R so large that $R - E(\sigma) > 0$. Next, fix j such that $j > 3(R - 1)$. Now, note that π_R^1 promotes for the first $(\sigma - 1)$ times, and π_R^2 promotes during the $\sigma - 2$ periods following the second period. Therefore, $s^{\pi_R^1}(i,\omega) = s^{\pi_R^2}(i,\omega)$ for $i \geq \sigma(\omega)$, all ω. Consequently for each ω we have

$$C(\pi_R^1, \omega(\pi_R^1)) - C(\pi_R^2, \omega(\pi_R^2)) = R - \sigma(\omega)$$

whence

(16)
$$V_{\pi_R^1}(s,R) - V_{\pi_R^2}(s,R) = R - E(\sigma) > 0,$$

so that promotion is not optimal.

<div align="right">Q.E D.</div>

In trying to interpret the above example   notice that regardless of the initial action, as soon as the first "1" is served the system enters A^+, at which time promotion is optimal. Thus, the difference in holding costs between promoting at .s and not promoting there (i.e., the difference in holding costs between π_R^1 and π_R^2) is 2. On the other hand, the difference in promotion costs is R. Hence, promotion is optimal only if R < 2.

4. COMPUTING THE OPTIMAL RULE IN A FINITE CAPACITY QUEUE

It was established that for any s there exists an $R > k(s)$ such that promotion is still strictly optimal at s. We now consider the unconstrained case and develop an algorithm which in the problem with finite queue capacity produces $R(s)$ for every s, and hence π_R, where $R(s)$ is the R-value for which promotion at s is optimal but not strictly optimal. That is, $R(s)$ is the unique solution to

$$(17) \qquad V_{\pi_R^1}(s,R) = V_{\pi_R^2}(s,R).$$

It is clear that $A^+(R_2) \subseteq A^+(R_1)$ and $A(R_2) \subseteq A(R_1)$ for any pair R_1, R_2 with $R_2 > R_1$. In other words, both $A^+(R)$ and $A(R)$ decrease (weakly) in the sense of set inclusion, as R increases. This suggests that $S_0(R)$ also decreases weakly as R increases. We believe that this is in fact true, but we are only able to prove the following.

THEOREM 7. If the queue capacity is finite, then $S_0(R)$ is nonincreasing in R, i.e., $S_0(R_1) \supseteq S_0(R_2)$ whenever $R_1 < R_2$.

Proof. To begin, observe that

$$(18) \qquad V_\pi(s,R) \text{ is linear in } R, \text{ for each } s \text{ and } \pi.$$

The veracity of (18) follows from the fact that the expected holding cost is independent of R and the expected cost associated with promotion is simply $R\eta(s,\pi)$, where $\eta(s,\pi)$ is the expected number of promotions when employing rule π and starting from state s.

Since the infimum of concave functions is itself concave, it
follows from (18) that

(19) $V(s,R)$ is concave in R for each s.

Moreover, because the state space is finite, there are only a finite
number of stationary rules. Consequently, it follows from (18) that

(20) $V(s,R)$ is piecewise linear for each s.

Now fix $R_2 > R_1$. If it is not true that $S_0(R_2) \subset S_0(R_1)$, then it
follows from (19) that there are states s_1 and s_2 such that $s_1 \in S_0(R_1)$,
$s_2 \in S_0(R_2), s_1 \notin S_0(R_2)$, and $s_2 \notin S_0(R_1)$. (However, we cannot assume
that $s_1 \in S_0(R)$ and $s_2 \notin S_0(R)$ for all $R_1 \leq R < R_2$ without additional
knowledge such as that provided by (20).)

In view of (20), we can assert that there is an $\epsilon > 0$ such
that

(21)   and $\begin{array}{l} \pi_1 \text{ is optimal on } [R_1, R_2] \\ \pi_2 \text{ is optimal on } [R_2, R_2 + \epsilon] , \end{array}$

where π_i promotes at $S_0(R_i)$ and $S_0(R_1) - \{s_1\} = S_0(R_2) - \{s_2\}$.

Now define $\widetilde{\pi}$ by

(22) $$\widetilde{\pi}(s) = \begin{cases} P & , \quad s = s_2 \\ \pi_1(s), & s \neq s_2 . \end{cases}$$

The continuity of $V_{\pi_i}(s,R)$ for each s, $i = 1,2$, ensures that $\widetilde{\pi}$
is optimal at R_2.

Fix s and note that the function $V_{\widetilde{\pi}}(s,R) - V_{\pi_1}(s,R)$ is
strictly increasing on $[R_1, R_2]$ and nonnegative at R_1 by (21)
and the optimality of π_1 at R_1, respectively. But this implies
$V_{\widetilde{\pi}}(s,R_2) - V_{\pi_1}(s,R_2) > 0$, contradicting the optimality of $\widetilde{\pi}$ at R_2.

Q.E.D.

Henceforth we assume that there is a bound N on the queue length. Let the state space S_N be the set of queues of length less than or equal to N. In order to establish an ordering of the state space vis-a-vis the optimality of promotion, we first denote by H the subset of S_N where promotion is at all possible. (Sometimes we write $H(0)$ rather than H.)

Theorem 7 revealed that the optimality set $S_0(R)$ shrinks as R increases. For any $s \in H$, $R(s)$ is the value such that s will no longer belong to $S_0(R)$ for any $R > R(s)$; i.e., promotion will no longer be optimal at s. Based on these facts it is legitimate to define a <u>hierarchy</u> among the M states in H as follows: a state s' is said to be as high as s or higher if $R(s') \geq R(s)$. Let s^1 be a state at which $R(s)$ is minimized over H. Let us order the states $s^1, s^2, \ldots, s^M$ in such a way that

(23) $$R(s^1) \leq R(s^2) \leq R(s^3) \leq \ldots \leq R(s^M),$$

where the minimum in (23) might not be unique.

Suppose we have produced the above hierarchy on H. Then for a given R either

    (a) $R \leq R(s^1)$ so that $\pi_R \equiv \pi_0$

    (b) $R > R(s^M)$ so that $\pi_R \equiv$ FIFO

    (c)   There exists j with $R(s^j) \leq R < R(s^{j+1})$ so that

$$\pi_R = \begin{cases} P \text{ at } s^i, & i > j \\ F \text{ at } s^i, & i \leq j. \end{cases}$$

Thus, the hierarchy explicitly produces an optimal rule.

Denote by $H(i)$ the set of states $\{s^j : j \geq i\}$. For s in $H(i)$, $1 \leq i < M$, define $r(s, i)$ to be the unique R-value which solves

(24) $$V_{\pi^1_{R(s^i)}}(s,R) = V_{\pi^2_{R(s^i)}}(s,R).$$

(Uniqueness follows from the fact that $V_{\pi^j_{R(s^i)}}(s,R)$ is a linear

function of R, $j = 1,2$.) For s in $H(0)$ define $r(s,0)$ to be the

unique solution of $V_{\pi^1_1}(s,R) = V_{\pi^2_1}(s,R)$. Next define $L(i)$ by

(25) $L(i) = \min\{r(s,i) : s \in H(i)\}$, $i = 0,1,2,\ldots,M-1$.

We now demonstrate that $L(i) = R(s^{i+1})$.

THEOREM 7.   For each i, $0 \le i < M$, we have

(26) $$L(i) = R(s^{i+1}).$$

Proof.   To begin, note that $V_{\pi^1_{R(s^i)}}(s,R) = V(s,R)$ for $s \in H(i)$

and $R \in [R(s^i), R(s^{i+1})]$ so that

(27) $V_{\pi^1_{R(s^i)}}(s,R) < V_{\pi^2_{R(s^i)}}(s,R)$, for $R \in [R(s^i), R(s^{i+1})]$ and $s \in H(i)$,

and, hence $r(s,i) \ge R(s^{i+1})$. (Thus $L(i) \ge R(s^{i+1})$.) Furthermore, it

follows from (17), (24), and the fact that $\pi_{R(s^i)}$ is optimal for

all $R \in [R(s^i), R(s^{i+1})]$ that $r(s^{i+1},i) = R(s^{i+1})$. Consequently,

$L(i) \le R(s^{i+1})$ and we are done.

                                                                Q.E.D.

Using Theorem 7 we now present an M-step algorithm for comput_ing $R(s^i)$, $i = 1, 2, \ldots, M$.

Step 1: (a) For every $s \in H(0)$ solve

$$V_{\pi_1^1}(s,R) = V_{\pi_1^2}(s,R), \quad \text{to arrive at } r(s,0).$$

(b) Set $L(0) = \min\{r(s,0), \; s \in H(0)\}$.

(c) Let s^1 be any state where the minimum in b is achieved and set $R(s^1) \equiv L(0)$.

(d) Set $H(1) \equiv H(0) - \{s^1\}$.

Step j+1: $(1 \leq j < M)$.

(a) For every $s \in H(j) \equiv H(0) - \{s^1, s^2, \ldots, s^j\}$ solve

$$V_{\pi^1_{R(s^j)}}(s,R) = V_{\pi^2_{R(s^j)}}(s,R), \quad \text{to arrive at } r(s,j).$$

(b) Let $L(j) = \min\{r(s,j), \; s \in H(j)\}$.

(c) Let s^{j+1} be any state at which the minimum in (b) is achieved and set $R(s^{j+1}) = L(j)$.

(d) Set $H(j + 1) \equiv H(j) - \{s^{j+1}\}$.

Step M:

(a) For s^M solve

$$V_{\pi^1_{R(s^{M-1})}}(s^M,R) = V_{\pi^2_{R(s^{M-1})}}(s^M,R) \quad \text{to arrive at } R(s^M).$$

Note that at each step we have to solve an equation of the type: $V_\pi(s,R) = V_{\pi'}(s,R)$, where π and π' are two given rules, furthermore, both functions are linear in R by (18), so we merely seek the R-value at which two straight lines intersect.

Two conjectures which, if true when the queue capacity is finite, can be used to expedite the process which produces $R(s)$ in the unconstrained case:

Conjecture 1. In the unconstrained case, if promotion is optimal
at s, then it is also optimal at ss', for any s' ∈ S.

Conjecture 2. In the unconstrained case, if promotion is optimal
at s, then it is also optimal at any s' which is obtained by re-
placing a "1" by a "2".

For example, if Conjecture (1) is true, and if promotion is opt-
imal at s = 21, then it is also optimal at ss' = 211. If Conjecture
(2) is true, and if promotion is optimal at s = 2112, then promotion
is also optimal at s' = 2122.

REFERENCES

[1]   Bell, C. E., "Characterization and Computation of Optimal Policies
       for Operating an M/G/1 Queueing System with Removable Server,"
       Operations Research, 19 (1971), pp. 208-218.

[2]   Bell, C. E.,"Efficient Operation of Optional-Priority Queueing
       Systems," Operations Research, 21 (1973), pp. 777-786.

[3]   Blackwell, D., "Discounted Dynamic Programming," Annals of
       Mathematical Statistics, 36 (1965), pp. 226-235.

[4]   Cobham, A., "Priority Assignment in Waiting Lines Problems,"
       Operations Research, 2 (1954), pp. 70-76.

[5]   Crabill, T. B., "Optimal Control of a Service Facility with
       Variable Exponential Service Times and Constant Arrival Rates,"
       Management Science, 18 (1972), pp. 560-566.

[6]   Denardo, E., "Contraction Mappings in the Theory Underlying
       Dynamic Programming," SIAM Reviews, 9 (1967), pp. 165-177.

[7]   Feller, W., An Introduction to Probability Theory and its Applica-
       tions, Vol. II, 1966, Wiley, New York.

[8]   Harrison, M. J., "Discrete Dynamic Programming with Unbounded
       Rewards," Annals of Mathematical Statistics, 43 (1972), pp. 636-644.

[9]   _____, "Dynamic Scheduling of a Multiclass Queue: Discount
       Optimality," Operations Research, 23 (1975), pp. 270-282.

[10]   Jaiswal, N. K., Priority Queues, 1968, Academic Press.

[11]   Kofman, E., An M/M/1 Dynamic Priority Queue with Optional Promo-
       tion," Ph.D. Dissertation, University of California, Los Angeles,
       1978.

[12] Lippman, S. A., "Semi-Markov Decision Processes with Unbounded Rewards," _Management Science_, 19 (1973), pp. 717-731.

[13] _____, "On Dynamic Programming with Unbounded Rewards," _Management Science_, 21 (1975), pp. 1225-1233.

[14] Oliver, R. M. and Pestalozzi, G., "On a Problem of Optimum Priority Classification," _J. Soc., Indust. and Appl. Math._, 13 (1965), pp. 890-901.

[15] Ross, S. M., "Non Discounted Denumerable Markovian Decision Model," _Annals of Mathematical Statistics_, 39 (1968), pp. 412-423.

[16] _____, _Applied Probability Models with Optimization Applications_, 1970, Holden-Day.

[17] Taylor, H., "Markovian Sequential Replacement Processes," _Annals of Mathematical Statistics_, 36 (1965), pp. 1677-1694.