MANAGEMENT SCIENCE
Vol. 29, No. 2, February 1983
Printed in U.S.A.

OPTIMAL ENTERING RULES FOR A CUSTOMER WITH WAIT OPTION AT AN M/G/1 QUEUE*

AVISHAI MANDELBAUM† AND URI YECHIALI†

A "smart" customer arrives at an M/G/1 queue. While every other arrival joins the system unconditionally, our customer is allowed to choose among three alternatives: (i) he may Enter the queue and stay there until his service is completed, (ii) he may Leave the system right away, or (iii) he may Wait outside the queue. If he Enters or Leaves the system, his decision is final and no further actions are taken. If he chooses to Wait, he makes a new decision at the next service completion where he may, again, select one of the three options: Enter, Leave, or Wait.

For a linear cost structure we show that for any n-period horizon ($0 \leqslant n \leqslant \infty$) the individual customer's optimal strategy is a 3-region (possibly degenerate) policy by which he Enters a small queue, Leaves a large one and Waits when the queue is of an intermediate size. We also give a necessary and sufficient condition for the Wait option to exist.

The special M/M/1 queue is further analyzed by allowing the Waiting customer to make decisions at instants of customers' arrivals as well as at instants of service completion. In this case, too, the optimality of the 3-region policy is derived and the "smart" customer dichotomy is recognized as the Gambler's Ruin problem.

Computational procedures are developed and numerical results are presented for commonly used service time distributions.

(QUEUES; OPTIMAL JOINING RULES)

1. Introduction

The subject of this paper is a single "smart" customer seeking self-optimization in an M/G/1 queueing system. Upon arrival our customer is made aware of the current state of the system and the monetary reward he will attain through completion of service. While all other customers join the queue unconditionally, our customer may choose between three alternatives: he may Enter the queue or Leave the system right away (which are the two standard alternatives used in previous models such as Yechiali [7], Adiri and Yechiali [1], Stidham [5], and others) or he may Wait outside the queue, monitor the system and defer his final decision—as to whether to Enter the queue or to Leave—to a later stage. While monitoring the development of the system, he reconsiders his status at each service completion, at which time he can again choose between the three alternatives—Enter, Leave, or Wait. Once he decides to Enter or to Leave, his decision is irrevocable and no further actions are taken.

There is a different cost for standing in line and waiting out of it. The decision of the customer is based on his concrete cost-benefit analysis and he reaches his conclusion on the basis of his narrow self-interest.

In contrast to exponential-service models, when dealing with general service time, we have to consider two phases of decisions. The first is the decision upon arrival. It depends on the number of customers queueing as well as on the outstanding service time of the customer being served (in [3] we derive an explicit formula for the conditional expectation of this outstanding service time given the number of customers present). If the customer chooses the "Waiting outside the queue" option, the second phase initiates and actions are then taken each time a service is completed. These actions depend on the queue size only.

*Accepted by Marcel F. Neuts, former Departmental Editor; received October 1, 1979. This paper has been with the authors 14 months for 3 revisions.

†Tel Aviv University, Israel.

0025-1909/83/2902/0174$01.25

The analysis is developed as follows: In §2 we present the model. Optimal strategies for Phase II are studied in §3 and shown to be of a 3-region type for both the finite and infinite horizons; that is, it is best for our customer to Enter small queues, Leave large ones and Wait if the queue is of an intermediate size. We obtain a necessary and sufficient condition under which it is best to exercise the Wait option, thus strictly improving upon policies with only Enter or Leave options. In addition, we derive a sufficient condition so that our customer can afford to be "polite" and still remain a "smart" one; that is, as long as he waits he can allow everyone else to be served before him and still maintain optimality.

The overall decision process is formulated in §4. In §5 the M/M/1 queue is studied where decisions are taken both at instants of arrival and departure. Here the problem converts to the Gambler's Ruin problem where the search for the parameters of the optimal policy reduces to a minimization problem of a function of two variables, unimodal in each variable. Computational procedures are decribed in §6, and numerical results are presented in §7.

2. The Model

We consider a FIFO regime, stationary M/GI/1 queue with Poisson arrivals at rate λ, i.i.d. service times having finite mean $1/\mu$, and traffic intensity $\rho = \lambda/\mu < 1$.

We are interested in a single individual customer, X, whose objective is to minimize his expected total cost. Upon arrival, customer X is facing Phase I of his decision process. At this stage he is willing to take one of three actions:

E: to Enter the queue (as all other customers do unconditionally), and incur queueing time losses at a rate of $A > 0$ monetary units per unit of time so as to gain a reward g upon his service completion,

L: to Leave the system and incur a penalty $l > 0$, or

W: to Wait outside the queue until the next service completion and then recalculate whether to Enter, Leave, or Wait.

While in the waiting stage, X incurs waiting time losses at a rate of $B > 0$.

If actions E or L are taken—Phase I and the entire decision process terminate. If option W is exercised, the decision process moves to Phase II. While in Phase II, actions are taken after each service completion. The same options—E, L or W—are available. When the decision is W—Phase II continues. As soon as decision E or L is taken—the phase and the entire process terminate.

3. Analysis of Phase II

We formulate the decision process in Phase II as a (semi) Markov Decision Process (MDP). We follow some of the notations of Ross [4, p. 119].

The state space is $\{0, 1, 2 \dots\}$, the number of customers observed by X at instants of service completion.

If the process is in state i and action a is taken then

$$P_{ij}(a) = \text{transition probability to state } j$$

and

$$C(i, a) = \text{expected one-step cost.}$$

We say that the MDP goes to the absorbing state ∞ when either E or L is chosen. Without loss of generality we ignore the waiting losses incurred by X while being served. We assume that $g > -l$ which is equivalent to saying that E is optimal at state $i = 0$.

The nonzero transition probabilities for $i \geqslant 1$ are

$$
\begin{aligned}
&P_{i,\infty}(E) = P_{i,\infty}(L) = 1 \qquad (1 \leqslant i \leqslant \infty), \\
&P_{ij}(W) = a_{j-i+1} \qquad (1 \leqslant i < \infty;\ i-1 \leqslant j < \infty),
\end{aligned} \tag{1}
$$

where $a_k = \int_0^\infty e^{-\lambda t}(\lambda t)^k / k!\, dG(t)$ $(0 \leqslant k < \infty)$ is the probability of k arrivals during a service period.

The expected one-step costs are:

$$
\begin{aligned}
&C(i, L) = l \qquad (1 \leqslant i < \infty), \\
&C(i, E) = (i/\mu)A - g \qquad (0 \leqslant i < \infty), \\
&C(i, W) = B/\mu \qquad (1 \leqslant i < \infty), \\
&C(\infty, a) = 0 \qquad (a = E, W, L).
\end{aligned} \tag{2}
$$

Let $U_n(i)$ be the value of an optimal policy for the n-period horizon starting from state i. That is to say, either E or L action must be taken within n steps.

We have

$$U_0(i) = \min\left\{ \frac{i}{\mu}A - g, l \right\}$$

and for $0 \leqslant n < \infty$,

$$
\begin{aligned}
&U_{n+1}(0) = -g, \\
&U_{n+1}(i) = \min\left\{ \frac{i}{\mu}A - g,\ \frac{B}{\mu} + \sum_{k=0}^{\infty} a_k U_n(i-1+k), l \right\} \qquad (1 \leqslant i < \infty).
\end{aligned} \tag{3}
$$

We now transform the MDP model into an equivalent Negative Dynamic Programming (NDP) model (Strauch [6], Ross [4, p. 135]). We need this transformation since under the MDP the costs are unbounded and take both positive and negative values.

We let

$$
\begin{aligned}
&V_n(i) = (U_n(i) + g)\mu/A, \\
&f = (l + g)\mu/A \quad \text{and} \quad c = B/A.
\end{aligned} \tag{4}
$$

This is the same as increasing the terminal cost by g (making the reward 0) and choosing as unit of cost the cost of lining up for one mean service time $(= A/\mu)$.

The NDP process has the same transition probabilities as in (1) and the following non-negative one-step expected costs

$$
\begin{aligned}
&C(i, L) = f \qquad (1 \leqslant i < \infty), \\
&C(i, E) = i \qquad (0 \leqslant i < \infty), \\
&C(i, W) = c \qquad (1 \leqslant i < \infty), \\
&C(\infty, a) = 0 \qquad (a = E, W, L).
\end{aligned} \tag{5}
$$

In terms of the $V_n(i)$, (3) takes the form

$$
\begin{aligned}
&V_0(i) = \min\{i, f\}, \\
&V_{n+1}\{0\} = 0 \qquad (0 \leqslant n < \infty), \\
&V_{n+1}(i) = \min\left\{ i, c + \sum_{k=0}^{\infty} a_k V_n(i-1+k), f \right\} \qquad (1 \leqslant i < \infty).
\end{aligned} \tag{6}
$$

Equations (6) can be extended to hold true for the infinite horizon case as a consequence of standard NDP results. Namely, letting $V_\infty(i) \equiv V(i)$ we have

THEOREM 1 (Infinite Horizon). (i) *An optimal policy exists and the optimal cost function $V(i)$ satisfies*

$$
\begin{aligned}
&V(0) = 0, \\
&V(i) = \min\left\{ i, c + \sum_{k=0}^{\infty} a_k V(i-1+k), f \right\} \qquad (1 \leqslant i < \infty).
\end{aligned} \tag{7}
$$

(ii) *The stationary policy which takes the action minimizing the right-hand side of* (7) *is an optimal policy.*

REMARK. (i) In case of ties we postulate the following transitive relation: L is preferable to E which is preferable to W.

Our aim now is to show that for both finite and infinite horizons the optimal policy of an individual customer is a 3-region policy as follows:

For any n-period horizon $(0 \leqslant n \leqslant \infty)$ there exist natural numbers s_n and t_n $(s_n < t_n < \infty)$ such that if the queue size i is not greater than s_n the optimal action is E; if $i \geqslant t_n$ the optimal action is L, and if $s_n < i < t_n$ the optimal action is W.

We establish these facts by the method of successive approximation using the following theorem:

THEOREM 2. *For $i \geqslant 0$, $\lim_{n\to\infty} V_n(i) = V(i)$. The convergence is monotone.*

PROOF. $V_n(i) \geqslant V_{n+1}(i) \geqslant V(i)$ $(n \geqslant 0)$ since what can be achieved in n steps could clearly be achieved in $n+1$ or more. Hence, $\exists \lim_{n\to\infty} V_n(i) \geqslant V(i)$. For $n = \infty$, let T_i be the random time at which an optimal policy stops starting from i. Then, following the same lines as in Ross [4, p. 136, Theorem 6.13], we derive

$$0 \leqslant V_n(i) - V(i) \leqslant P(T_i > n)(f - c).$$

Since T_i is bounded by the time it takes the queue to empty (i busy periods) then $E(T_i) \leqslant i/(1 - P) < \infty$. Thus, $\lim_{n\to\infty} P(T_i > n) = 0$, which completes the proof.

LEMMA 3. *For $i \geqslant 0$, $n \geqslant 0$*

(i) $0 \leqslant V_n(i) \leqslant V_n(i+1) \leqslant V_n(i) + 1$,

(ii) $0 \leqslant V(i) \leqslant V(i+1) \leqslant V(i) + \min\{1, c/(1-\rho)\}$.

PROOF. (i) The nonnegativity of $V_n(i)$ is obvious. The rest follows from (6) using induction on n.

(ii) Passing to limit in (i) results in $0 \leqslant V(i) \leqslant V(i+1) \leqslant V(i) + 1$. In addition, the expected number of service completions that it takes to get from $i+1$ to i is $1/(1-\rho)$. If at state $i+1$ X waits until the system reaches state i and then applies the optimal policy his expected loss is $c/(1-\rho) + V(i)$, which proves (ii).

We now characterize the structure of the optimal policies. For reasons of convenience we assume that f is an integer. All our qualitative results stay true for an aribtrary f except for somewhat more tedious calculations.

Denote by $\pi_n(i)$ the optimal action when in state i and there are (at most) n steps to go. The sequence of mappings $\{\pi_0, \pi_1, \ldots, \pi_n\}$ defines an optimal nonrandomized policy. An infinite-horizon optimal stationary policy where $\pi_0 = \pi_1 = \cdots = \pi_\infty$ will be abbreviated to π_∞.

We also introduce a special notation to denote 3-region policies. For any integers s, t satisfying $t - s \geqslant 1$ define a mapping $\pi(s,t):\{0,1,\ldots\} \to \{E, W, L\}$ by

$$\pi(s,t)(i) = \begin{cases} E, & 0 \leqslant i \leqslant s, \\ W, & s < i < t, \\ L, & i \geqslant t. \end{cases}$$

Note that in light of part (ii) of Theorem 1 and the remark following it, we have, for $0 \leqslant n \leqslant \infty$ and $i \geqslant 0$,

$$V_n(i) = f \quad \text{iff} \quad \pi_n(i) = L,$$

if

$$V_n(i) < f \quad \text{then} \quad V_n(i) = i \quad \text{iff} \quad \pi_n(i) = E,$$

$$V_n(i) < \min\{i, f\} \quad \text{iff} \quad \pi_n(i) = W.$$

LEMMA 4. *For $0 \leqslant n \leqslant \infty$,*

(a) *If $\pi_n(i) = L$ then $\pi_n(i+j) = L$ for $j \geqslant 1$.*

(b) *If $\pi_n(i) = E$ then $\pi_n(i-j) = E$ for $0 \leqslant j \leqslant i$.*

(c) *If $\pi_n(i) = W$ then $\pi_{n+j}(i) = W$ for $j \geqslant 1$.*

(d) *There always exists an L-state.*

In particular $\pi_n(f+n) = L$ for $n < \infty$.

(e) *There exist s_n and t_n, $s_n + 1 \leqslant t_n < \infty$, such that $\pi_n = \pi(s_n, t_n)$. That is to say, π_n is a 3-region policy.*

(f) $0 \leqslant t_{n+1} - t_n \leqslant 1$.

PROOF. (a) From Lemma 3(i) and from the optimality equations (6), $f \geqslant V_n(i+j) \geqslant V_n(i) = f$.

(b) Again, from (6) and Lemma 3, $i - 1 \geqslant V_n(i-1) \geqslant V_n(i) - 1 = i - 1$. Iterating results in (b).

(c) $\pi_n(i) = W$ iff $\min\{i, f\} > V_n(i)$, and $V_n(i) \geqslant V_{n+j}(i)$. Hence (c).

(d) First let $n = \infty$. Suppose there are no L-states. Since $\pi_\infty(i) \neq E$ for $i > f$, there exists a maximal E state, say s_∞, and $\pi_\infty(i) = W$ for $i \geqslant s_\infty + 1$. Starting with state $s_\infty + k$ there must be at least k service completions before X joins the queue. Thus $V(s_\infty + k) \geqslant c \cdot k + s_\infty$ with k as large as we please. This contradicts $V(s_\infty + k) \leqslant f$ and implies (d) for $n = \infty$.

Now let $n < \infty$. Suppose $V_n(i) = f$ for some i. Then $f \leqslant i < i + 1$. Using (a), $c + \sum_{k=0}^{\infty} a_k V_n(i+k) = c + \sum_{k=0}^{\infty} a_k \cdot f = c + f > f$. Hence $V_{n+1}(i+1) = f$. That is, $V_n(i) = f \Rightarrow V_{n+1}(i+1) = f$.

Starting with $V_o(f) = f$, we have, $\pi_n(f+n) = L$ for $n \geqslant 0$.

(e) The 3-region structure follows from (a) and (b). The existence of an E-state follows from $\pi_n(o) = E$. Finally, $t_n < \infty$ by (d).

(f) As in (d), $V_n(t_n) = f \Rightarrow V_{n+1}(t_n + 1) = f$. Hence, $t_{n+1} \leqslant t_n + 1$. $t_{n+1} \geqslant t_n$ since $\pi_n(t_n - 1) = E$ implies that $\pi_{n+1}(t_n - 1)$ is either E or W, and if $\pi_n(t_n - 1) = W$ then $\pi_{n+1}(t_n - 1) = W$ by (c).

Lemma 4 establishes that an optimal policy is a 3-region policy. However, the 3-region policy may degenerate to a simple 2-region policy where it is never optimal to exercise the W option. We now find the conditions under which a nondegenerate 3-region policy strictly improves the 2-region policy (with only the E and L options) assumed in all previous individual optimization studies.

For $0 \leqslant n \leqslant \infty$ denote the set of states for which the decision is E, W or L by E_n, W_n and L_n, respectively. That is,

$$E_n = \pi_n^{-1}(E); \; W_n = \pi_n^{-1}(W); \; L_n = \pi_n^{-1}(L).$$

Note that $W_1 = \emptyset$ implies $V_1 = V_0$, which implies that $\pi_0 \equiv \pi_1 \equiv \cdots \equiv \pi_\infty$, and hence $W_\infty = \emptyset$. On the other hand, $W_1 \neq \emptyset$ implies $\pi_1(i) = W$ for some i, which implies $\pi_\infty(i) = W$ (by Lemma 4(c)) and hence $W_\infty = \emptyset$. Therefore, $W_\infty = \emptyset$ iff $W_1 = \emptyset$.

It is reasonable to assume that when c is large enough there are no W-states (clearly this is the case when $c > f$), and if c is small enough there are no E-states except for 0. We claim

THEOREM 5. (a) *$W_\infty = \phi$ iff $c \geqslant a_0$.*
(b) *If $c < 1 - \rho$ then $E_\infty = \{0\}$.*

PROOF. (a) It suffices to establish the conditions for $W_1 \neq \emptyset$. From (d) and (a) of Lemma 4, $V_1(f+j) = f$ for $j \geqslant 1$. (Recall that f is an integer.) Since $W_1 \neq \emptyset$ implies that $f \in W_1$, it is enough to check state f. $V_1(f) = \min\{f,\ c + a_0(f-1) + (1-a_0)f,\ f\} = \min\{f, f + c - a_0, f\}$. Hence, $\pi_1(f) = W$ iff $c - a_0 < 0$.

(b) For $i \geqslant 1$ we have $c + \sum_{k=0}^{\infty} a_k V(i-1+k) \leqslant c + \sum_{k=0}^{\infty} a_k(i-1+k) = i - [(1-\rho) - c] < i$ since $\sum_{k=0}^{\infty} k a_k = \rho$. Hence, W is always preferable to E.

REMARK. Using the notation of the Markovian model one could have expected that $B < A$ would guarantee the existence of states where W is the optimal action. Nevertheless, part (a) of Theorem 5 implies that there are cases where $c < 1$ and still it is never optimal to wait. In the sequel we show that for the M/M/1 queue, where decisions are taken *both* at instants of arrival and service completion, the condition $c < 1$ will indeed guarantee the existence of optimal W-states.

Part (b) of Theorem 5 tells X that when $c < 1 - \rho$ he can afford being a 'polite' customer and still remain a "smart" one: as long as he waits he can allow everyone else to be served before him and still maintain optimality.

To summarize, for an optimal policy to have 3 nondegenerate regions (nondegenerate E-region must include states other than 0), it is necessary that c belongs to the interval $[1-\rho, a_0)$. Note that $1 - a_0 = \sum_{k=1}^{\infty} a_k \leqslant \sum_{k=1}^{\infty} k a_k = \rho$.

In Lemma 4(d) a sequence of upper bounds on the maximal W-state was established for the finite horizon problem. We now derive an upper bound on the number of W-states of an optimal policy π_∞. We accomplish this by deriving a lower bound on $V(i+1) - V(i), i \in W_\infty$.

Suppose $\pi_\infty = \pi(s,t)$ for some $s < t$. Let $m = t - s - 1$ and suppose $m \geqslant 1$. That is, $W_\infty \neq \emptyset$ which is equivalent to $c < a_0$ by Theorem 5.

The vector $(V(s+1), \ldots, V(s+m))$ solves the equations $V(s+j) = c + a_0 V(s+j-1) + \cdots + a_{m-j+1} V(s+m) + (1 - a_0 - \cdots - a_{m-j+1})f$, $1 \leqslant j \leqslant m$, where $V(s) = s$.

Let $x_j = f - V(s+j)$ for $0 \leqslant j \leqslant m$. Backward induction on j proves that $x_j - x_{j+1} \geqslant c/a_0$ for $0 \leqslant j \leqslant m-1$. Summing these m inequalities we get $x_0 - x_m \geqslant m \cdot c/a_0$, or $m \leqslant a_0/c[V(s+m) - s] < a_0/c[f - s]$. Since $c < a_0$ we have established the following:

THEOREM 6. *If $\pi_\infty = \pi(s_\infty, t_\infty)$, $m_\infty = t_\infty - s_\infty - 1$, $m_\infty \geqslant 1$, then*

(a)
$$m_\infty < \frac{a_0}{c}(f - s_\infty),$$

(b)
$$t_\infty < \frac{a_0}{c} f + 1.$$

Theorem 6 also gives rise to a computational procedure that identifies an optimal stationary 3-region policy when searching through only a finite set of candidates (see §6).

We emphasize that all our results are applicable to any model with transition probabilities (1) and cost structure (2) or (5), where the sequence $\{a_k\}_0^\infty$ need not be of the specific form of the M/GI/1 model. We summarize the results in a general framework which will be used later.

THEOREM 7. *Consider a Markov decision process with state space* $\{0, 1, \ldots, \infty\}$ *and an action space* $\{E, W, L\}$. *Suppose the transition probabilities are as in* (1) *and the cost structure as in* (2) *or* (5), *where, in addition,* $0 < a_0 < 1$ *and* $\rho = \sum_{k=0}^{\infty} k a_k < 1$.

Then there exist two sequences of integers $\{s_n\}_{n=0}^{\infty}$, $\{t_n\}_{n=0}^{\infty}$ *that satisfy*:

(a) $\pi_n = \pi(s_n, t_n)$ *for* $0 \leqslant n \leqslant \infty$. ($n = \infty$ *applies to the case of a stationary nonrandomized* 3-*region policy*.)

(b) $s_0 = f - 1$; $t_0 = f$.

(c) $0 \leqslant t_{n+1} - t_n \leqslant 1$ *for* $n \geqslant 0$.

(d) $0 \leqslant s_\infty \leqslant \cdots \leqslant s_1 \leqslant s_0 = f - 1 < f = t_0 \leqslant t_1 \leqslant \cdots \leqslant t_\infty < \infty$.

(e) *If* $c < 1 - \rho$ *then* $s_\infty = 0$ ("*smart*" *customer can afford being* "*polite*").

(f) $c \geqslant a_0$ *iff* $t_1 - s_1 = 1$ *iff* $t_\infty - s_\infty = 1$ (*i.e.*, $s_0 = s_1 = \cdots = s_\infty$, $t_0 = t_1 = \cdots = t_\infty$ *and there are no W-states*).

(g) *When* $c < a_0$, $t_\infty + s_\infty(a_0/c - 1) < (a_0/c)f + 1$.

4. Analysis of Phase I

The analysis of Phase I differs from that of Phase II as it depends on the outstanding service time, R, of the customer being served when X arrives and finds $i \geqslant 1$ customers in the system. We denote this state by $\hat{i}$.

We develop recursive formulae from which the optimal decision at instant of arrival can be obtained.

For $i \geqslant 1, j \geqslant 0$, let

$$R_i = E(R \mid \hat{i}) = \int_0^\infty x\,dP\{R \leqslant x \mid \hat{i}\}, \qquad D_i = R_i/(1/\mu),$$

$$b_{ij} = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!}\,dP\{R \leqslant x \mid \hat{i}\}.$$

R_i is the expected outstanding service time if there are i customers upon arrival, and D_i is R_i normalized with respect to the regular expected service time. b_{ij} is the probability that j customers will join the queue during the remaining service time, R, had X seen i upon arrival.

The optimal cost function for instants of arrival satisfies

$$V(\hat{o}) = 0,$$

$$V(\hat{i}) = \min\left\{ D_i + i - 1,\ D_i \cdot c + \sum_{j=0}^{\infty} b_{ij} V(i - 1 + j),\ f \right\},$$

where $V(i)$ is the optimal cost function for Phase II.

If $\{p_i\}_{i=0}^{\infty}$ denotes the stationary distribution of the M/GI/1 Markov chain ($p_0 = 1 - \rho$) it is shown in [3] that

$$D_i = \frac{1-\rho}{\rho}\,\frac{1 - (p_0 + \cdots + p_i)}{p_i}, \qquad i \geqslant 1, \tag{8}$$

and that

$$\begin{aligned}(a_0 p_i) b_{ij} + (a_1 p_{i-1}) b_{i-1,j} + \cdots + (a_{i-1} p_1) b_{1j} \\ = p_0 a_{j+i} + p_1 a_{j+i-1} + \cdots + p_{i-1} a_{j+1}.\end{aligned} \tag{9}$$

The following is a procedure to calculate the optimal policy at instants of arrival, assuming $\{a_j\}_{j=0}^{\infty}$ and ρ are given. (See [3] for recursive formulae of a_j in various M/GI/1 models.)

Step 1. Calculate $\{V(i)\}_{i=0}^{\infty}$ by successive approximation via $V_n(i)$ (Theorem 2). As shown in §3 the optimal policy at instants of service completion is a 3-region policy, say, $\pi(s_\infty, t_\infty)$.

Step 2. (a) Calculate $(p_0, p_1, \ldots, p_{t_\infty})$:

$$p_0 = 1 - \rho,$$

$$p_1 = \frac{1}{a_0}(p_0 - p_0 a_0),$$

$$p_i = \frac{1}{a_0}\left[p_{i-1} - p_0 a_{i-1} - \sum_{k=1}^{i-1} p_k a_{i-k}\right], \qquad 2 \leqslant i \leqslant t_\infty.$$

(b) Calculate $(D_0, D_1, \ldots, D_{t_\infty})$:

$$D_0 = 1,$$

$$D_i = \frac{P_{i-1}}{p_i} D_{i-1} - \frac{1-\rho}{\rho}, \qquad 1 \leqslant i \leqslant t_\infty.$$

(c) Starting with $b_{1j} = a_{j+1}/(1 - a_0)$, $0 \leqslant j \leqslant t_\infty - 1$, calculate row by row the half matrix (using (9)).

$$\begin{bmatrix} b_{10} & b_{11} \cdots & & b_{1,t_\infty - 1} \\ b_{20} & b_{21} \cdots & b_{2,t_\infty - 2} & \\ \vdots & \vdots & & \\ b_{t_\infty - 1,0} & b_{t_\infty - 1,1} & & \\ b_{t_\infty,0} & & & \end{bmatrix}$$

Step 3. The optimal cost function is

$$V(\hat{o}) = 0,$$

$$V(\hat{i}) = f, \qquad i \geqslant t_\infty + 1,$$

$$V(\hat{i}) = \min\left\{D_i + i - 1, cD_i + \sum_{k=0}^{t_\infty - i + 1} b_{ik} V(i - 1 + k) + \left(1 - \sum_{k=0}^{t_\infty - i + 1} b_{ik}\right) f, f\right\},$$

$$(1 \leqslant i \leqslant t_\infty).$$

Numerical results for various queueing models are presented in §7. We note that in all our calculations the optimal policies are found to be of the 3-region type not only for Phase II (as derived analytically) but for Phase I as well. This is probably due to the monotone failure rate property of the service distributions considered. The problem of deriving conditions under which a 3-region policy is optimal at Phase I remains open.

5. The M/M/1 Model—Full Control

In this section we study the M/M/1 model. Applying results (8) and (9) with $p_i = (1 - \rho)\rho^i$, $i \geqslant 1$, and $a_j = \rho^j/((1 + \rho)^{j+1})$, $j \geqslant 0$, we get

$$b_{ij} \equiv a_j; \qquad R_i \equiv \frac{1}{\mu}; \qquad D_i \equiv 1.$$

This is expected due to the memoryless property of the exponential service time. As a result we state:

THEOREM 8. *When service time is exponential, the optimal decision at instants of service completion is also optimal at the instant of arrival.*

The assumption of exponential service time facilitates the analysis considerably. Moreover, it enables us to analyze the M/M/1 queue when decisions may be taken both at instants of arrival and service completion. That is, X is allowed to make a decision any time the system changes its state. His first decision is made upon arrival. When taking the W option, X may reconsider his status any time a customer either joins or leaves the queue.

Formulating the queueing process as a semi-Markov process (Ross [4, p. 105]), we have:

$$P_{i,\infty}(E) = P_{i,\infty}(L) = 1,$$

$$P_{ij}(W) = \begin{cases} \lambda/(\lambda+\mu), & j = i+1, \\ \mu/(\lambda+\mu), & j = i-1, \qquad i \geqslant 1, \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

We concentrate in the sequel on the infinite horizon case. Using the original cost structure, the optimal cost function satisfies:

$$U(o) = -g,$$

$$U(i) = \min\left\{ i \cdot \frac{A}{\mu} - g, B \cdot \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} U(i-1) + \frac{\lambda}{\lambda+\mu} U(i+1), l \right\}, \qquad i \geqslant 1.$$

Transforming to the NDP notation gives, for $V(i) = (U(i) + g)\mu/A$,

$$\begin{cases} V(o) = 0, \\ V(i) = \min\left\{ i, \tilde{c} + \sum_{j=0}^{\infty} \tilde{a}_j V(i-1+j), \tilde{f} \right\}, \qquad i \geqslant 1, \end{cases} \tag{11}$$

where

$$\tilde{f} = (l+g)\frac{\mu}{A} = f,$$

$$\tilde{c} = \frac{B}{A}\frac{\mu}{\lambda+\mu} = c \cdot \frac{\mu}{\lambda+\mu}; \qquad c = \frac{B}{A},$$

$$\tilde{a}_0 = \frac{\mu}{\lambda+\mu} = \frac{1}{1+\rho}; \qquad \tilde{a}_2 = \frac{\lambda}{\lambda+\mu} = \frac{\rho}{1+\rho}; \qquad \tilde{a}_1 = \tilde{a}_3 = \tilde{a}_4 = \cdots = 0.$$

Calculating $\tilde{\rho} = \sum_{j=0}^{\infty} j\tilde{a}_j = 2 \cdot \lambda/(\lambda+\mu)$ we see that $\tilde{\rho} < 1$ iff $\rho < 1$, so, the conditions of Theorem 7 are met. Since $\tilde{c} < \tilde{a}_0$ iff $c < 1$ and $\tilde{c} < 1 - \tilde{\rho}$ iff $c < 1 - \rho$, we get:

THEOREM 9. *In a stationary* M/M/1 *model* ($\rho < 1$) *where decisions are made both at instants of customer arrival and service completion, the optimal policy* π (*infinite horizon*) *satisfies*:

(a) $\pi = \pi(s_\infty, t_\infty)$ *for some* s_∞, t_∞ ($t_\infty - s_\infty \geqslant 1$),
(b) *there are no* W-*states iff* $c \geqslant 1$ ($B \geqslant A$),
(c) *if* $c < 1 - \rho$, 0 *is the only* E-*state.*

We describe now a procedure to calculate the optimal policy in this particular M/M/1 model. We assume that $c < 1$. Since, otherwise, an optimal procedure has no W-states and the solution is trivial.

The procedure is based on the observation that any 3-region policy $\pi(s,t)$ with $t - s \geqslant 2$ can be regarded as a Gambler's Ruin Problem (Feller [2, pp. 344, 348]), with the following parameters:

$$q = \text{probability of a failure } (-1) = \tilde{a}_0$$

$$p = \text{probability of a success } (+1) = \tilde{a}_2. \qquad (p < q),$$

The game continues as long as X takes the W-option (the system remains between states s and t): "Winning" is defined as taking the L-option at t.

"Losing" is considered as taking the E-option at s.

Denote by $m = t - s - 1$ the number of W-states. We use the well-known results:

$$p_{i,L} = P\left\{\begin{matrix}\text{Leaving at} \\ \text{state } t\end{matrix}\middle|\begin{matrix}\text{Starting at} \\ \text{state } s+i\end{matrix}\right\} = \frac{1-(q/p)^i}{1-(q/p)^{m+1}}, \qquad 1 \leqslant i \leqslant m,$$

$$p_{i,E} = P\left\{\begin{matrix}\text{Entering at} \\ \text{state } s\end{matrix}\middle|\begin{matrix}\text{Starting at} \\ \text{state } s+i\end{matrix}\right\} = 1 - p_{i,L},$$

T_i = expected number of decisions, starting from $s + i$, until a final decision (E or L) is taken

$$T_i = \frac{i}{q-p} - \frac{m+1}{q-p} \cdot \frac{1-(q/p)^i}{1-(q/p)^{m+1}}.$$

If we denote by $V(i,s,m)$ the cost function of the policy $\pi(s, s+m+1)$ then

$$V(i,s,m) = s \cdot p_{i-s,E} + f \cdot p_{i-s,L} + \tilde{c} \cdot T_{i-s}, \qquad s+1 \leqslant i \leqslant t-1 = s+m$$

or

$$V(i,s,m) = \begin{cases} i, & 0 \leqslant i \leqslant s, \\ f, & i \geqslant s+m+1, \\ s\left(1 - \dfrac{c}{1-\rho}\right) + i \cdot \dfrac{c}{1-\rho} + \dfrac{c}{1-\rho} \cdot \dfrac{1-\rho^{i-s}}{\rho^{i-s}} & \\ \cdot \dfrac{\rho^{m+1}}{1-\rho^{m+1}}\left[(f-s)\dfrac{1-\rho}{c} - (m+1)\right], & s+1 \leqslant i \leqslant s+m. \end{cases}$$

We use the upper limit for m (Theorem 6) and the relation $(\tilde{a}_0/\tilde{c}) \cdot \tilde{f} = f/c$ and define a finite set

$$\tilde{\Gamma} = \left\{(s,m) \middle| 1 \leqslant m < \frac{f}{c}, \max[0, f-m] \leqslant s < \min\left[f, \frac{f}{c} - m\right]\right\}$$

to which the pair (s_∞, m_∞) that defines an optimal policy, must belong.

Thus $V(i) = \min_{(s,m) \in \tilde{\Gamma}} V(i,s,m)$, and the problem of finding the optimal cost function reduces to a minimization problem of a function of two variables (s,m).

We add a few observations to improve the minimization procedure:

(1) The procedure is applied only if $c < 1$, otherwise an optimal policy has only two regions: $W_\infty = \emptyset$.

(2) Fixing i and s ($i \geqslant s+1$, $s < f$) it is enough to find an m ($i - s \leqslant m < (f-s)/c$) that minimizes the function $\phi(m)$ defined by

$$\phi(m) = \frac{\rho^{m+1}}{1-\rho^{m+1}}\left[(f-s)\frac{1-\rho}{c} - (m+1)\right].$$

$\phi(m)$ is easily proved to be unimodal, so starting with $m = i - s$, we increase m until the first m satisfying $\phi(m+1) \geqslant \phi(m)$ which is the one we need.

(3) Our computational experience shows that an optimal policy is characterized by a unique pair (s, m) which minimizes $V(i, s, m)$ for $i = f$. (f must be a W-state if there are any.)

(4) If $c < 1 - \rho$ we know that state 0 is the only E-state. The problem reduces to minimization of a unimodal function with a single variable m, ($s = 0$).

6. A Search Procedure for the Calculation of $\{V_n(i)\}_{i=0}^{\infty}$

One way to calculate $\{V_n(i)\}_{i=0}^{\infty}$ is by successive approximation as indicated in Step 1 of §4. Another method is to extend the search procedure for the M/M/1 model, described in §5, and apply it to the general model as follows:

We regard any stationary 3-region policy as a random walk on $\{0, 1, 2, \ldots\}$ with absorbing regions $S = \{0, 1, \ldots s\}$ and $T = \{t, t+1, \ldots\}$ (Feller [2, p. 363]). The transition probabilities from any state z are given by

$$p_{z,z+k} = a_{k+1}, \qquad k = -1, 0, 1, 2, \ldots .$$

Suppose X starts in state i and pays c monetary units per step till absorption takes place. If absorbed at S he pays s units. If absorbed at T he pays f. The expected cost then is exactly the cost function $V(i, s, m)$ associated with $\pi(s, t)$ ($m = t - s - 1$).

The vector $(V(s+1, s, m), \ldots, V(t-1, s, m))$ is a solution to the following set of equations, with $(x_1, \ldots, x_m)$ being the unknowns:

$$\begin{cases} x_1 = c + a_0 s + a_1 x_1 + \cdots + a_m x_m + (1 - a_0 - \cdots - a_m) f \\ x_2 = c + a_0 x_1 + \cdots + a_{m-1} x_m + (1 - a_0 - \cdots - a_{m-1}) f \\ \vdots \\ x_m = c + a_0 x_{m-1} + a_1 x_m + (1 - a_0 - a_1) f. \end{cases} \tag{12}$$

Searching through the set

$$\Gamma = \left\{(s, m) | 1 \leqslant m < \frac{a_0}{c} f,\ \max[0, f - m] \leqslant s < \min\left[f, \frac{a_0}{c} f - m\right]\right\}$$

while solving (12) for any candidate policy, is guaranteed to result in identifying the values of an optimal policy.

We add a few comments concerning the solution of (12):

Let I_m be the unit matrix of order m.

Let M_m be the matrix of order m defined by

$$M_m = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \\ a_0 & a_1 & \cdots & a_{m-1} \\ 0 & a_0 & \cdots & a_{m-2} \\ \vdots & & & \\ 0 & \cdots & a_0 & a_1 \end{bmatrix}.$$

In order to solve (12) the matrix $[I_m - M_m]$ must be inverted.

Since $p_{z,z-1} = a_0 \geqslant 1 - \rho > 0$, the matrix $[I_m - M_m]$ has an inverse (Feller [2, p. 364]). Thus (12) has a unique solution. We also note that if in $M_m, a_j > 0$ for all $j \geqslant 0$ (which is usually the case in the M/GI/1 model) then a result for Minkowski-Leontieff matrices (see details in [3]) guarantees that the solution is positive.

Inversion of matrices is a costly operation. The following lemma, proved in [3], can be used for the calculation of $[I_m - M_m]^{-1}$, without actually doing any inversion. We emphasize that the search procedure requires the inversion of $[I_m - M_m]$ for $1 \leqslant m < (a_0/c)f$.

LEMMA 10. *Let*

$$z_m = (-a_{m+1}, -a_m, \ldots, -a_2),$$
$$y_m = (0, 0, \ldots, 0, -a_0), \qquad m \geqslant 1.$$

Define inductively the following sequence of matrices:

$$E_1 = \left[\frac{1}{1 - a_1} \right],$$

$$E_{m+1} = \begin{bmatrix} E_m & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{(1 - a_1) - y_m E_m z_m^t} \cdot \begin{bmatrix} E_m z_m^t y_m E_m & -E_m z_m^t \\ -y_m E_m & 1 \end{bmatrix},$$

then $E_m = [I_m - M_m]^{-1}$ *for* $m \geqslant 1$.

7. Numeric Results

We present some numerical results (Tables 1–5) for the following queueing models: M/M/1, M/E_k/1 ($k = 2$), M/D/1 and M/Gamma (0.5)/1. In all models $\rho = 0.8$, $c = 0.234$ and $f = 7.0$. In each table optimal values of the cost function $V_n(i)$ are given, where Theorem 2 is applied to obtain the values of $V_\infty(i)$. The last row in each table consists of the optimal values, $V_n(\hat{i})$, for Phase I, the calculation of which is described in §4 and [3]. As for Phase II, since $V_n(i) = f$ whenever $\pi_n(i) = L$, and $V_n(i) = i$ whenever $\pi_n(i) = E$ $(0 \leqslant n \leqslant \infty)$, we indicate the action (L or E) rather than the cost value $V_n(i)$ when appropriate, so that only those values of $V_n(i)$ correspond to the W-regions are presented. Observe that in each table the W-states form a pyramid-type region in the middle of the table—as was derived analytically in Lemma 4 and Theorem 7.

Tables 1–4 are for the case when decisions in Phase II are taken at instants of service completion (§§2–4). Table 5 corresponds to the M/M/1 queue where decisions are taken both at instants of arrival and departure (§5).

TABLE 1
The M/M/1 *Queue*

i	0	1	2	3	4	5	6	7	8	9	10
$V_0(i)$	E	.	.	.	.	.	E	L	.	.	L
$V_1(i)$	E	.	.	.	E	4.96	5.88	6.68	L	.	L
$V_2(i)$	E	.	.	E	3.99	4.93	5.79	6.53	L	.	L
$V_3(i)$	E	.	.	E	3.97	4.89	5.73	6.45	6.97	L	L
$V_5(i)$	E	.	E	E	3.95	4.84	5.65	6.35	6.87	L	L
$V_\infty(i)$	E	.	E	2.96	3.87	4.72	5.48	6.14	6.68	L	L
$V_\infty(\hat{i})$	0	1	2	2.96	3.87	4.72	5.48	6.14	6.68	7.00	7.00

The values of Phase I and Phase II are identical—see Theorem 8.

TABLE 2

The M/E$_2$/1 *Queue*

i	0	1	2	3	4	5	6	7	8	9	10
$V_0(i)$	E	.	.	.	.	.	E	L	.	.	L
$V_1(i)$	E	.	.	.	E	E	5.92	6.72	L	.	L
$V_4(i)$	E	.	.	E	E	4.94	5.79	6.50	7.00	L	L
$V_\infty(i)$	E	.	.	E	3.97	4.86	5.67	6.34	6.86	L	L
$V_\infty(\hat{i})$	0	0.79	1.75	2.74	3.71	4.63	5.46	6.17	6.73	7.00	7.00

TABLE 3

The M/D/1 *Queue*

i	0	1	2	3	4	5	6	7	8	9	10
$V_0(i)$	E	.	.	.	.	.	E	L	.	.	L
$V_1(i)$	E	.	.	.	.	E	5.98	6.78	L	.	L
$V_4(i)$	E	.	.	.	E	E	5.89	6.62	L	.	L
$V_\infty(i)$	E	.	.	.	E	4.97	5.84	6.56	L	.	L
$V_\infty(\hat{i})$	0	0.57	1.48	2.47	3.46	4.46	5.39	6.19	6.79	7.00	7.00

TABLE 4

The M/Gamma (0.5)/1 *Queue*

i	0	1	2	3	4	5	6	7	8	9	10
$V_0(i)$	E	.	.	.	.	.	E	L.	.		L
$V_1(i)$	E	.	E	2.99	3.96	4.90	5.80	6.61	L	.	L
$V_2(i)$	E	E	E	2.97	3.92	4.83	5.67	6.42	6.99	L	L
$V_5(i)$	E	E	1.97	2.92	3.82	4.67	5.45	6.14	6.70	L	L
$V_8(i)$	E	0.99	1.95	2.87	3.75	4.58	5.33	6.00	6.56	6.98	L
$V_\infty(i)$	E	0.96	1.88	2.76	3.60	4.38	5.10	5.74	6.30	6.75	L
$V_\infty(\hat{i})$	0	1.32	2.31	3.20	4.02	4.78	5.46	6.06	6.57	6.97	7.00

Note that here $V_\infty(\hat{i}) > V_\infty(i)$ as Gamma (0.5) is a DFR distribution.

TABLE 5

The M/M/1 *Queue—Decisions at Arrivals and Service Completions*

i	0	1	2	3	4	5	6	7	8	9	10
$V_0(i)$	E	.	.	.	.	.	E	L	.	.	L
$V_1(i)$	E	.	.	.	.	.	E	6.57	L	.	L
$V_2(i)$	E	.	.	.	.	E	5.83	6.57	6.89	L	L
$V_3(i)$	E	.	.	.	E	4.94	5.83	6.43	6.89	L	L
$V_4(i)$	E	.	.	E	3.99	4.94	5.74	6.43	6.81	L	L
$V_9(i)$	E	.	E	E	3.96	4.83	5.62	6.25	6.73	6.98	L
$V_{34}(i)$	E	E	E	2.96	3.86	4.69	5.44	6.07	6.57	6.89	L
$V_\infty(i)$	E	E	1.99	2.95	3.84	4.67	5.41	6.04	6.54	6.88	L
$V_\infty(\hat{i})$	0	1.00	1.99	2.95	3.84	4.67	5.41	6.04	6.54	6.88	7.00

As expected, there is a strict improvement in the cost function $V_\infty(\cdot)$ when compared with Table 1.

References

1. Adiri, I. and Yechiali, U., "Optimal Priority—Purchasing and Pricing Decisions in Nonmonopoly and Monopoly Queues," *Oper. Res.*, Vol. 22 (1974), pp. 1051-1066.
2. Feller, W., *An Introduction to Probability Theory and Its Applications*, Wiley International Edition, 3rd Edition, Volume 1 (1950).
3. Mandelbaum, A. and Yechiali, U., "Individual Optimization in the M/GI/1 Queue," Technical Report, Department of Statistics, Tel Aviv University (September 1969).
4. Ross, M.S., *Applied Probability Models with Optimization Applications*, Holden-Day, 1970.
5. Stidham, S., "Socially and Individually Optimal Control of Arrivals to a GI/M/1 Queue," *Management Sci.*, Vol. 24 (1978), pp. 1598-1610.
6. Strauch, R.E., "Negative Dynamic Programming," *Ann. Math. Statist.*, Vol. 37 (1966), pp. 871-890.
7. Yechiali, U., "On Optimal Balking Rules and Toll Charges in the GI/M/1 Queueing Process," *Oper. Res.*, Vol. 19 (1971), pp. 349-370.