Look-Ahead Policies for Admission to a Single Server Loss System
Author(s): Wim M. Nawijn
Source: *Operations Research*, Vol. 38, No. 5 (Sep. - Oct., 1990), pp. 854-862
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/171044
Accessed: 11-12-2015 12:40 UTC

LOOK-AHEAD POLICIES FOR ADMISSION TO A SINGLE SERVER LOSS SYSTEM

WIM M. NAWIJN

The University of Twente, Enschede, The Netherlands

(Received November 1987; revision received February 1989; accepted May 1989)

Consider a single server loss system in which the server, being idle, may reject or accept an arriving customer for service depending on the state at the arrival epoch. It is assumed that at every arrival epoch the server knows the service time of the arriving customer, the arrival time of the next customer and the service time. The server gets a fixed reward for every customer admitted to the system. The form of an optimal stationary policy is investigated for the discounted and average reward cases.

Many models exist for controlling the admission of customers to a service system. For a survey of these models we refer to Stidham (1985). In almost all cases, it is assumed that at a decision epoch there is no a priori information about future events. If, however, some information is available a server may reject a customer from entering the system in favor of a future customer, depending on the objective. We will investigate such a situation in the case of a single server loss system, that is, a system without waiting room, where customers that arrive while the server is busy are lost and never return. Customers arrive according to a renewal process, with interarrival time distribution $A(\cdot)$ and require a service time with probability distribution $B(\cdot)$. Service times are mutually independent and independent of the arrival process. It is assumed that at each arrival epoch the server knows the service time of the corresponding customer and the arrival epoch of the next customer together with the service time. For every customer entering the system the server gets a fixed reward. We will derive the form of a stationary optimal policy in the case where the server wants to maximize the total expected discounted reward as well as in the case where the server wants to maximize the average expected reward in the long run.

In the sequel, it will be assumed for convenience that $A(0+) = 0$ and $A(\cdot)$ is absolutely continuous, so that the corresponding renewal function is strictly increasing from a certain point onward.

In an earlier paper, Nawijn (1985) considered a related problem, concerning a single server model with finite, but positive, waiting room in which arrivals during service are lost. Decisions, taken at service completions, are based on information regarding the number of customers in the system and the time until the next arrival, which is exactly known. A decision is to either postpone service and admit the first arriving customer to the system or to immediately start servicing a waiting customer. The objective is to maximize the average number of serviced customers in the long run. In this model, however, the service time of the next customer is unknown. The present paper focuses on the form of an optimal stationary policy if this service time is also known; we restrict ourselves to the case of zero waiting room. Here it is more natural to take the arrival instants as decision epochs. The case of positive waiting room would have a state space involving all the known service times of waiting customers, which leads to an extremely complex decision problem.

1. THE DECISION MODEL

Let T_n, $n = 1, 2, \ldots, (T_1 = 0)$ denote the arrival epoch of the nth customer and P_n its service time. The decision epochs coincide with arrival epochs. We assume that at epoch T_n, the server knows P_n, T_{n+1}, P_{n+1}, and the residual service time Y_n, if a customer is being served. When the server is idle at time $T_n - 0$ we define the state of the system to be $(0; b; \tau, p)$, in which $b = P_n$, $\tau = T_{n+1} - T_n$ and $p = P_{n+1}$. When the server is busy at time $T_n - 0$ the state is defined as $(1; b; \tau, p)$, where $b = Y_n$ and τ and p are defined as before. Notice that P_n does not enter the state description because the customer is lost. The state-space $S = \{0, 1\} \times [0, \infty]^3$.

Subject classifications: Dynamic programming, semiMarkov: admission control to a loss system. Queues, optimization: input control to a single server loss system.

0030-364X/90/3805-0854 $01.25

Let $a_n(s)$, $s \in S$ denote the decision at epoch T_n. Obviously, we have

$$a_n(0; b; \tau, p) = 0 \text{ (reject)} \quad \text{or } 1 \text{ (accept)}$$
$$a_n(1; b; \tau, p) = 0 \tag{1}$$

for all b, τ and p and all admissible policies π.

Let $X_n = 1$ or 0 if the server is busy or idle at time $T_n - 0$, respectively. Moreover, let $Z_n = P_n$ if $X_n = 0$ and $Z_n = Y_n$ if $X_n = 1$. Denote the transition probabilities by

$$Q_{ij}^k(z, x, y \mid b, \tau, p) = \Pr\{X_{n+1} = j,\ Z_{n+1} \leqslant z,\ T_{n+2} - T_{n+1} \leqslant x,\ P_{n+2} \leqslant y \mid X_n = i,\ Z_n = b,\ T_{n+1} - T_n = \tau,\ P_{n+1} = p,\ a_n = k\}. \tag{2}$$

Then we have

$$Q_{0j}^0(z, x, y \mid b, \tau, p) = \delta_{0j} U(z - p) A(x) B(y)$$
$$Q_{ij}^1(z, x, y \mid b, \tau, p) = \delta_{1j} U(z - b + \tau) A(x) B(y), \quad b \geqslant \tau, \quad i = 0, 1 \tag{3}$$
$$Q_{ij}^1(z, x, y \mid b, \tau, p) = \delta_{0j} U(z - p) A(x) B(y), \quad b < \tau, \quad i = 0, 1$$

in which δ_{ij} is Kronecker's symbol and $U(\cdot)$ is the unit step function. Since the reward for admitting a customer into the system is fixed, we may simply define the reward $r(a_n, s)$ associated with decision a_n in state $s \in S$ as

$$r(a_n, s) = 1 \quad \text{if } a_n = 1$$
$$= 0 \quad \text{if } a_n = 0. \tag{4}$$

The objectives to be considered are: i) maximizing the expected total discounted reward, and ii) maximizing the expected average reward in the long run. In particular, we want to find an optimal stationary policy that maximizes for every initial state $s_1 \in S$

$$E_\pi\left\{\sum_{n=1}^{\infty} e^{-\rho T_n} r(a_n, s_n) \mid s_1\right\}, \quad \rho > 0 \tag{5}$$

and

$$\liminf_{N\to\infty} \frac{E_\pi\{\sum_{n=1}^{N} r(a_n, s_n) \mid s_1\}}{E_\pi\{T_n \mid s_1\}} = \lambda \liminf_{N\to\infty} E_\pi\left\{\sum_{n=1}^{N} r(a_n, s_n) \mid s_1\right\} \Big/ N \tag{6}$$

over all permissible policies π.

Remark 1

a. It will be assumed that $T_2 - T_1 < \infty$ and $Y_1 < \infty$.
b. The equality in (6) follows from the fact that the transition times are independent of the policy used.

2. THE DISCOUNTED REWARD CASE

The optimality equations for the discounted reward case are

$$V_0^\rho(b; \tau, p) = \max\begin{Bmatrix} 1 + e^{-\rho\tau}\bar{V}_0^\rho(p), & b < \tau \\ 1 + e^{-\rho\tau}\bar{V}_1^{\rho_1}(b - \tau), & b \geqslant \tau; \quad e^{-\rho\tau}\bar{V}_0^\rho(p) \end{Bmatrix} \tag{7}$$

$$V_1^\rho(b; \tau, p) = \begin{cases} e^{-\rho\tau}\bar{V}_0^\rho(p), & b < \tau \\ e^{-\rho\tau}\bar{V}_1^{\rho_1}(b - \tau), & b \geqslant \tau \end{cases} \tag{8}$$

in which

$$\bar{V}_i^\rho(b) = \int_0^\infty \int_0^\infty V_i^\rho(b; \tau, p)\, dA(\tau)\, dB(p), \quad i = 0, 1. \tag{9}$$

Since the reward is bounded, (7) and (8) have a unique solution and any policy that prescribes an action that maximizes the right side in (7) is optimal. It is immediately seen from (7) that if a customer arrives when the system is empty, and its service time is smaller than the time elapsed until the next arrival, i.e., if $b < \tau$, it is always optimal to admit this customer to the system. This is an obvious result since the next customer is never lost in doing so. Hence, the analysis concentrates on the optimal decisions in case $b \geqslant \tau$. So we exclude the trivial case $P(b < \tau) = 1$, in which the service times are always smaller than the interarrival times.

To start the analysis, let us first consider equation (8). It is readily verified by averaging (8) with respect to $A(\tau)$ and $B(p)$ that

$$\bar{V}_1^\rho(b) = v_0(\rho) \int_{\tau=b}^{\infty} e^{-\rho\tau}\, dA(\tau) + \int_{\tau=0}^{b} e^{-\rho\tau}\bar{V}_1^\rho(b - \tau)\, dA(\tau), \quad b \geqslant 0 \tag{10}$$

in which

$$v_0(\rho) = \int_0^\infty \bar{V}_0^\rho(b)\, dB(b). \tag{11}$$

Introducing the Laplace–Stieltjes transform of $A(\cdot)$

$$\alpha(\rho) = \int_0^\infty e^{-\rho t}\, dA(\tau), \quad \rho \geqslant 0 \tag{12}$$

equation (10) can be written as

$$e^{\rho b}[\bar{V}_1^\rho(b) - v_0(\rho)]$$
$$= \{\alpha(\rho) - 1\}v_0(\rho)e^{\rho b}$$
$$+ \int_0^b e^{\rho(b-\tau)}[\bar{V}_1^\rho(b-\tau) - v_0(\rho)]\, dA(\tau). \quad (13)$$

This convolution type of equation, which is related to the well known renewal equation, has the unique solution

$$\bar{V}_1^\rho(b) = \alpha(\rho)v_0(\rho) - (1 - \alpha(\rho))v_0(\rho) \cdot \int_0^b e^{-\rho t}\, dM(t), \quad b \geq 0 \quad (14)$$

in which $M(\cdot)$ denotes the renewal function.

Since

$$\int_0^\infty e^{-\rho t}\, dM(t) = \frac{\alpha(\rho)}{1-\alpha(\rho)}, \quad \rho > 0 \quad (15)$$

we can rewrite (14) as

$$\bar{V}_1^\rho(b) = (1 - \alpha(\rho))v_0(\rho) \int_b^\infty e^{-\rho t}\, dM(t), \quad b \geq 0. \quad (16)$$

Because, by assumption, $M(\cdot)$ is continuous and strictly increasing and, moreover, $\alpha(\rho) < 1$ if $\rho > 0$, it follows that $\bar{V}_1^\rho(\cdot)$ is continuous and strictly decreasing.

Letting $b \to \infty$ in (16) we obtain

$$\bar{V}_1^\rho(\infty) = 0, \quad \rho > 0 \quad (17)$$

an obvious result because no reward will be earned when starting in state $(1; \infty; \tau, p)$.

Observe that the Laplace–Stieltjes transform (15) of the renewal function can be viewed as the expected discounted reward offered to the system for $t > 0$. So, including the reward offered at time $t = 0$ ($T_1 = 0$), we conclude that $v_0(\rho) \leq 1/(1 - \alpha(\rho))$. For a decision problem in which the information comprises only the service time of an arriving customer, it is readily seen that if $\alpha(\rho) < ½$ (heavy discounting) it is optimal to accept every customer if possible. For the problem under consideration, equations (7) and (8) admit such a policy only in special cases. Suppose that this policy is optimal and thus it prescribes a maximizing action in (7). Then (19), which will be considered in a moment, holds with equality. Therefore, it follows from (16) and (7) that

$$1 - e^{-\rho\tau} \geq e^{-\rho\tau}(1 - \alpha(\rho))v_0(\rho) \int_p^{b-\tau} e^{-\rho t}\, dM(t)$$

must hold for all values of b, p and τ such that $p \geq 0$, $b > \tau \geq 0$ and which belong to the supports of $B(\cdot)$ and $A(\cdot)$. Sufficient conditions for this to be true are ρ being infinite or $\Pr\{\tau \leq b - p\} = 0$. The latter condition is satisfied for constant service times and when the interarrival times are always larger than the service times. If $B(\cdot)$ has finite support (s, S), it is also satisfied if $P(\tau > S - s) = 1$. That, in these cases, the above policy is optimal is intuitively clear. If $A(\cdot)$ has support $(0, \infty)$ it is easily seen that the above expression does not hold for τ small enough and, consequently, accepting is not a maximizing action.

We will need the following lower bounds for $\bar{V}_0^\rho(\cdot)$.

Lemma 1

$$\bar{V}_0^\rho(b) \geq 1 - A(b) + \alpha(\rho)v_0(\rho), \quad b \geq 0 \quad (18)$$

$$\bar{V}_0^\rho(b) \geq 1 + \bar{V}_1^\rho(b), \quad b \geq 0. \quad (19)$$

Proof. Since the discounted reward earned starting in state $(0; b; \tau, p)$ choosing a maximizing action and proceeding optimally in the future is at least as large as the discounted reward earned when rejecting the customer if $b \geq \tau$ and accepting if $b < \tau$ and proceeding optimally in the future we have from (7)

$$V_0^\rho(b; \tau, p) = 1 + e^{-\rho\tau}\bar{V}_0^\rho(p), \quad b < \tau$$

$$V_0^\rho(b; \tau, p) \geq e^{-\rho\tau}\bar{V}_0^\rho(p), \quad b \geq \tau.$$

Averaging with respect to (w.r.t.) the distribution functions $A(\tau)$ and $B(p)$ proves (18). Relation (19) follows from the observation that when accepting the arriving customer in state $(0; b; \tau, p)$, irrespective of b, τ and p, one immediately jumps to state $(1; b; \tau, p)$ and gets a direct reward 1. Hence, for an optimal policy $V_0^\rho(b; \tau, p) \geq 1 + V_1^\rho(b; \tau, p)$ from which (19) is a direct consequence.

Averaging (19) w.r.t. $B(b)$ and using relation (16) leads to the following lower bound for $v_0(\rho)$.

Corollary 1

$$v_0(\rho)\{1 - \alpha(\rho)\} \geq 1 \Big/ \left\{1 + \int_0^\infty \int_0^b e^{-\rho t}\, dM(t)\, dB(b)\right\}. \quad (20)$$

It can be shown that the right-hand side in (20) gives the expected discounted reward starting with an empty system and always accepting an arriving customer if possible.

To investigate the behavior of $V_0^\rho(b;\tau,p)$ as a function of b in the neighborhood of $b=\tau$, consider the difference

$$\Delta^\rho(\tau,p)=[1+e^{-\rho\tau}\bar{V}_0^\rho(p)] - \lim_{b\to\tau^+}[1+e^{-\rho\tau}\bar{V}_1^\rho(b-\tau)]. \quad (21)$$

From (14) we obtain

$$\Delta^\rho(\tau,p)=e^{-\rho\tau}[\bar{V}_0^\rho(p)-\alpha(\rho)v_0(\rho)]. \quad (22)$$

Therefore, by (18), we have

$$\Delta^\rho(\tau,p)\geqslant e^{-\rho\tau}(1-A(p)),\quad \tau\geqslant 0,\quad p\geqslant 0. \quad (23)$$

Consequently, since $\bar{V}_1^\rho(\cdot)$ is decreasing it follows that

$$V_0^\rho(b;\tau,p)\leqslant 1+e^{-\rho\tau}\bar{V}_0^\rho(p) \quad \text{for all } b\geqslant 0,\quad \tau\geqslant 0,\quad p\geqslant 0 \quad (24)$$

from which we obtain

$$\bar{V}_0^\rho(b)\leqslant 1+\alpha(\rho)v_0(\rho),\quad b\geqslant 0. \quad (25)$$

Combining (22), (23) and (25) leads to the following result.

Lemma 2. *For all $\tau\geqslant 0$ and $p\geqslant 0$*

$$e^{-\rho\tau}(1-A(p))\leqslant\Delta^\rho(\tau,p)\leqslant e^{-\rho\tau}. \quad (26)$$

So $V_0^\rho(b;\tau,p)$ has, as a function of b, a downward jump at $b=\tau$ of magnitude at most one with τ and p fixed. Moreover, since $1+e^{-\rho\tau}\bar{V}_1^\rho(b-\tau)$ is a continuous and decreasing function of b, we either have:

a. $V_0^\rho(b;\tau,p)=1+e^{-\rho\tau}\bar{V}_1^\rho(b-\tau)$ for all $b\geqslant\tau$, or b. There exists a number $c_\rho(\tau,p)>0$ such that $V_0^\rho(b;\tau,p)=1+e^{-\rho\tau}\bar{V}_1^\rho(b-\tau)$ for $\tau\leqslant b<\tau+c_\rho(\tau,p)$ and $V_0^\rho(b;\tau,p)=e^{-\rho\tau}\bar{V}_0^\rho(p)$ for $b\geqslant\tau+c_\rho(\tau,p)$. In view of the properties of $\bar{V}_1^\rho(\cdot)$, it follows from $\bar{V}_1^\rho(\infty)=0$ that a necessary and sufficient condition for the two cases (a and b) to occur are $e^{-\rho\tau}\bar{V}_0^\rho(p)\leqslant 1$ and $e^{-\rho\tau}\bar{V}_0^\rho(p)>1$, respectively.

Proposition 1. *There exists a function $k_\rho(p)$, $p\geqslant 0$, that satisfies*

$$\exp\{-\rho k_\rho(p)\}\bar{V}_0^\rho(p)=1. \quad (27)$$

The function $k_\rho(p)$ is nonnegative, bounded and decreasing. Moreover

$$\begin{aligned} &e^{-\rho\tau}\bar{V}_0^\rho(p)>1 \quad \text{if } \tau<k_\rho(p)\\ &e^{-\rho\tau}\bar{V}_0^\rho(p)\leqslant 1 \quad \text{if } \tau\geqslant k_\rho(p). \end{aligned} \quad (28)$$

Proof. Consider the equation in τ $e^{-\rho\tau}\bar{V}_0^\rho(p)=1$ for fixed $p\geqslant 0$. First we observe that in both the above cases (a and b) $V_0^\rho(b;\tau,p)$ is a decreasing function of b, for fixed τ and p, and, consequently, so is $\bar{V}_0^\rho(b)$. Secondly, it can be verified from (18) and (25) that

$$\bar{V}_0^\rho(0)=1+\alpha(\rho)v_0(\rho). \quad (29)$$

The interpretation is that when starting in a state $(0;0+;\tau,p)$ one certainly accepts the present customer, so the optimal expected discounted reward equals $V_0^\rho(0;\tau,p)=1+e^{-\rho\tau}\bar{V}_0^\rho(p)$, from which (29) is a direct consequence.

Since $\bar{V}_1^\rho(b)\geqslant 0$ for $b\geqslant 0$, it follows from (19) and (29) that

$$1\leqslant\bar{V}_0^\rho(p)\leqslant 1+\alpha(\rho)v_0(\rho),\quad p\geqslant 0. \quad (30)$$

Moreover, since $v_0(\rho)\alpha(\rho)>0$, cf. (20), we also have $\bar{V}_0^\rho(0)>1$. It is readily seen that the above equation implicitly defines a function $k_\rho(p)$ that satisfies the assertions.

Returning to the behavior of $V_0^\rho(b;\tau,p)$ as a function of b, we apparently have for $b\geqslant\tau$, with τ and p fixed, that case a occurs when $\tau\geqslant k_\rho(p)$ and case b occurs when $\tau<k_\rho(p)$. From the latter observation we obtain the following result.

Proposition 2. *There exists a function $c_\rho(\tau,p)\geqslant 0$ such that*

$$1+e^{-\rho\tau}\bar{V}_1^\rho(c_\rho(\tau,p))=e^{-\rho\tau}\bar{V}_0^\rho(p) \quad \text{if } \tau<k_\rho(p),\quad p\geqslant 0 \quad (31)$$

which is increasing in τ, for p fixed, and increasing in p with τ fixed. Moreover, for $\tau\leqslant k_\rho(p)$

a. $c_\rho(\tau,p)\to\infty$ *if* $\tau\uparrow k_\rho(p)$

b. $e^{\rho\tau}-1+A(p)$

$$\geqslant(1-\alpha(\rho))v_0(\rho)\int_0^{c_\rho(\tau,p)}e^{-\rho t}\,dM(t)$$

$$\geqslant e^{\rho\tau}-1$$

c. $c_\rho(0,0)=0$.

Proof. Since $\bar{V}_1^\rho(\cdot)$ is continuous and strictly decreasing there exists, for fixed τ and p, a number $b=\tau+c_\rho(\tau,p)$ such that

$$1+e^{-\rho\tau}\bar{V}_1^\rho(b-\tau)=e^{-\rho\tau}\bar{V}_0^\rho(p)$$

provided that $\tau<k_\rho(p)$, which proves (31). Writing (31) as

$$\bar{V}_1^\rho(c_\rho(\tau,p))=\bar{V}_0^\rho(p)-e^{\rho\tau}$$

and noting that $\bar{V}_0^\rho(\cdot)$ and $\bar{V}_1^\rho(\cdot)$ are both decreasing, it is readily seen that $c_\rho(\tau,p)$ is componentwise increasing.

From (27) and (31) it follows that $\bar{V}_1^\rho(c_\rho(\tau, p)) \to 0$ if $\tau \uparrow k_\rho(p)$. Hence, property a follows from (17) since $\bar{V}_1^\rho(\cdot)$ is strictly decreasing. Property b can be deduced from inequalities (18) and (25) by using (14) and (31). Part c is a direct consequence of b.

Taking $\tau = 0$ in (31), averaging the result with respect to $B(p)$, and using (16) we obtain the following relation between $v_0(\rho)$ and $c_\rho(0, p)$.

Corollary 2

$$v_0(\rho) = 1 \Big/ \left[\{1 - \alpha(\rho)\} \cdot \left\{ 1 + \int_0^\infty \int_0^{c_\rho(0,p)} e^{-\rho t}\, dM(t)\, dB(p) \right\} \right], \quad \rho > 0. \quad (32)$$

Comparing (32) and (20) gives the following result.

Corollary 3. *The critical function* $c_\rho(\tau, p)$ satisfies

$$c_\rho(0, p) \leqslant p, \quad p \geqslant 0. \quad (33)$$

This result is intuitively clear when considering the decisions in state $(0; b; 0, p)$ with $b > p$. Apparently, it is always better to reject the b-customer in favor of the p-customer, since the direct reward is equal for both decisions, while the expected discounted future rewards will be larger when rejecting the b-customer and accepting the p-customer since $p < b$. Hence, p must be an upper bound for $c_\rho(0, p)$.

We are now in a position to state the form of an optimal policy.

Theorem 1. *There exists a stationary optimal policy such that for all* $p \geqslant 0$

$$a_n(0; b; \tau, p) = \begin{cases} 0 & \text{if } \tau < k_\rho(p) \text{ and } b \geqslant \tau + c_\rho(\tau, p) \\ 1 & \text{otherwise.} \end{cases} \quad (34)$$

Moreover

$$\begin{aligned} V_0^\rho(b; \tau, p) &= 1 + e^{-\rho\tau}\bar{V}_0^\rho(p), \quad b < \tau \\ &= 1 + e^{-\rho\tau}\bar{V}_1^\rho(b - \tau), \quad \tau \leqslant b < \tau + c_\rho(\tau, p) \\ &\quad \text{and } \tau < k_\rho(p) \\ &\quad \text{or } \tau \leqslant b \text{ and } \tau \geqslant k_\rho(p) \\ &= e^{-\rho\tau}\bar{V}_0^\rho(p), \quad b \geqslant \tau + c_\rho(\tau, p) \\ &\quad \text{and } \tau < k_\rho(p) \end{aligned} \quad (35)$$

in which the functions $k_\rho(p)$ *and* $c_\rho(\tau, p)$ *satisfy the properties in Propositions* 1 *and* 2, *respectively.*

Remark 2

i. The optimality of the above policy follows from the fact that (34) prescribes a maximizing action in (7).
ii. Obviously, the functions $k_\rho(p)$ and $c_\rho(\tau, p)$ are still to be determined. Formal equations can be obtained by averaging (35) over $A(\tau)$ and $B(p)$ and using (27), (31) and (16). Since these are extremely complicated we will not pursue the analysis; see, however, the average reward case.

An upper bound for $k_\rho(p)$ can easily be obtained by the following argument. Suppose that a customer with service time b arrives when the server is idle, and τ and p are unknown. Then the optimal expected discounted reward equals $\bar{V}_0^\rho(b)$. When rejecting the customer, one can never earn more than $\alpha/(1 - \alpha)$ in view of the interpretation given earlier. On the other hand, when accepting the customer, with direct reward 1, the expected discounted reward offered in the future will be equal to $\int_b^\infty \exp(-\rho t)\, dM(t)$. Hence, we have

$$\bar{V}_0^\rho(b) \leqslant \max\left\{ \frac{\alpha(\rho)}{1 - \alpha(\rho)},\ 1 + \int_b^\infty e^{-\rho t}\, dM(t) \right\}$$

or, equivalently, cf. (27)

$$k_\rho(p) \leqslant \frac{1}{\rho} \ln\left[\max\left\{ \frac{\alpha(\rho)}{1 - \alpha(\rho)},\ 1 + \int_p^\infty e^{-\rho t}\, dM(t) \right\} \right], \quad p \geqslant 0. \quad (36)$$

Observe that when $\alpha(\rho) < ½$ (heavy discounting), (36) implies $k_\rho(\infty) = 0$ and, since (cf. (16), (27) and (31))

$$\bar{V}_0^\rho(p) = \exp[\rho k_\rho(p)] = 1 + (1 - \alpha(\rho)) v_0(\rho) \int_{c_\rho(0,p)}^\infty e^{-\rho t}\, dM(t) \quad (37)$$

we also have $\bar{V}_0^\rho(\infty) = 1$ and $c_\rho(0, \infty) = \infty$. This means that in state $(0; b; \tau, p)$, for very large p and heavy discounting, one always accepts the present customer irrespective of its service time b, as one would expect.

From the lower bounds for $\bar{V}_0^\rho(\cdot)$ in Lemma 1 and Corollary 1 we obtain

$$k_\rho(p) \geqslant \frac{1}{\rho} \ln\left[1 + \frac{\alpha(\rho)}{1 - \alpha(\rho)} \cdot \frac{1}{1 + I} - \min\left\{ A(p),\ \int_0^p e^{-\rho t}\, dM(t)/(1 + I) \right\} \right] \quad (38)$$

in which

$$I = \int_0^\infty \int_0^p e^{-\rho t}\, dM(t)\, dB(p).$$

Observe that $k_\rho(p)$ tends to infinity for all $p \geq 0$ if ρ tends to zero. Letting $p \to \infty$ in (38) yields

$$k_\rho(\infty) \geq \frac{1}{\rho} \ln\left[\max\left(1, \frac{\alpha}{1-\alpha} \cdot \frac{1}{1+I}\right)\right]. \tag{39}$$

For ρ small enough it follows that

$$k_\rho(\infty) \geq \frac{1}{\rho} \ln\left(\frac{\alpha}{1-\alpha} \cdot \frac{1}{1+I}\right) > 0$$

and, consequently, we conclude from (37) that $c_\rho(0, \infty) < \infty$, which in turn implies that $c_\rho(0, p)$ is a bounded function of p, and so is $c_\rho(\tau, p)$ for fixed $\tau < k_\rho(\infty)$. It may, therefore, be optimal (under weak discounting) in state $(0; b; \tau, \infty)$ to reject the arriving customer and, of course, the next one, contrary to the case of heavy discounting.

In Figure 1, typical control domains are depicted for two values of p. The diagram illustrates the fact that as p, the service time of the next customer, becomes larger, the server is more reluctant to reject the presently arrived customer with service time b, given the value of τ. Moreover, for a given value of p, the larger the value of τ, the larger the value of b for which rejection occurs.

If τ is too large, i.e., if $\tau \geq k_\rho(p)$, the server will always accept the presently arriving customer. The reason is that under discounting accepting an arriving customer in state $(0; b; \tau, p)$ gives an immediate reward 1, while rejecting the customer leads to an expected future discounted reward of $O(e^{-\rho\tau})$. Hence, if τ is large enough it will always be optimal to accept the present customer even if $b = \infty$, so that all future rewards will be missed. In the average reward case, which we will discuss next, this phenomenon does not

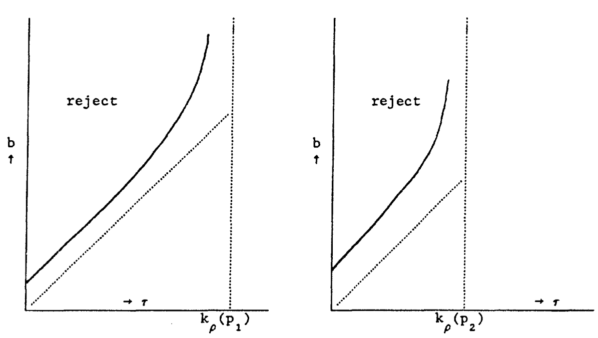

Figure 1. Control domain ($p_1 < p_2$).

occur because the potential contributions of all arriving customers are equal.

3. THE AVERAGE REWARD CASE

An optimal stationary policy will be obtained from the discounted reward case by appropriately letting $\rho \downarrow 0$, following the ideas of Ross (1983). First, observe from (20) and (30) that for $\rho > 0$

$$\left\{1 + \int_0^\infty \int_0^p e^{-\rho t}\, dM(t)\, dB(p)\right\}^{-1}$$

$$\leq \frac{[1-\alpha(\rho)]}{\rho} \cdot [\rho v_0(\rho)] \leq 1. \tag{40}$$

Since $[1 - \alpha(\rho)]/\rho \to 1/\lambda$ as $\rho \downarrow 0$, it follows from (40) that there exists some constant $\delta > 0$, such that $\rho v_0(\rho)$ is bounded for $\rho \in [0, \delta]$. Consequently, we have the following result.

Lemma 3. *There exists a sequence* $\{\rho_n\}$ *with* $\rho_n \downarrow 0$ *as* $n \to \infty$ *and a finite constant* g *such that* $\rho_n v_0(\rho_n) \to g$ *as* $n \to \infty$, *with*

$$\lambda\left\{1 + \int_0^\infty M(p)\, dB(p)\right\}^{-1} \leq g \leq \lambda. \tag{41}$$

Remark 3

The left side in (41) gives the expected number of service completions per unit of time if every customer arriving when the system is empty is admitted to the system.

Lemma 4

a. $\bar{V}_1^{\rho_n}(b) - v_0(\rho_n) \to -(1/\lambda)g[1 + M(b)]$ *as* $n \to \infty$.
b. $\bar{V}_0^\rho(p) - v_0(\rho)$ *is uniformly bounded for all* $\rho \in [0, \delta]$ *and all* $p \geq 0$.

Proof. Part a follows from (14) and Lemma 3. Combining inequalities (18) and (25) leads to

$$1 - A(p) - \left[\frac{1-\alpha(\rho)}{\rho}\right][\rho v_0(\rho)]$$

$$\leq \bar{V}_0^\rho(p) - v_0(p) \leq 1 - \left[\frac{1-\alpha(\rho)}{\rho}\right][\rho v_0(\rho)] \tag{42}$$

and assertion b readily follows from (40).

From the above lemmas it is standard to prove the following theorem (see Ross, p. 95), so the proof will be omitted.

Theorem 2

a. *There exist functions $f_0(b;\tau,p)$ and $f_1(b;\tau,p)$ and a constant g that satisfies*

$$g\tau + f_0(b;\tau,p) = \max\begin{bmatrix} 1+\bar{f}_0(p), & b<\tau & \\ & & ; \ \bar{f}_0(p) \\ 1+\bar{f}_1(b-\tau), & b\geq\tau & \end{bmatrix} \quad (43)$$

$$g\tau + f_1(b;\tau,p) = \begin{cases} \bar{f}_0(p), & b<\tau \\ \bar{f}_1(b-\tau), & b\geq\tau \end{cases} \quad (44)$$

in which

$$\bar{f}_i(b) = \int_0^\infty \int_0^\infty f_i(b;\tau,p)\,dA(\tau)\,dB(p), \quad i=0,1.$$

b. *For some sequence* $\{\rho'_n\}\downarrow 0$

$$\lim_{n\to\infty}[V_i^{\rho'_n}(b;\tau,p) - v_0(\rho'_n)] = f_i(b;\tau,p), \quad i=0,1.$$

c. $\lim_{\rho\downarrow 0}\rho v_0(\rho) = g.$

The form of the solution of the optimality Equations (43) and (44) can be deduced from the discounted case in view of part b of the theorem. There are two essential observations to be made in this limiting process. First, observe from Lemma 2 that we must have

$$0 \leq \lim_{b\downarrow\tau}[1+\bar{f}_1(b-\tau)] - \bar{f}_0(p) \leq A(p). \quad (45)$$

Hence, $1+\bar{f}_1(b-\tau)$, viewed as a function of b, starts off at $b=\tau$ in between $1+\bar{f}_0(p)$ and $\bar{f}_0(p)$. Secondly, considering the difference

$$1 + e^{-\rho\tau}[\bar{V}_1^\rho(b-\tau) - v_0(\rho)] - e^{-\rho\tau}[\bar{V}_0^\rho(p) - v_0(\rho)]$$

taking $\rho=\rho'_n$ and letting $n\to\infty$ we obtain

$$1+\bar{f}_1(b-\tau) - \bar{f}_0(p)$$
$$= 1 - \frac{g}{\lambda}[1+M(b-\tau)] - \bar{f}_0(p).$$

Hence, since $g>0$ and $M(\cdot)$ is a strictly increasing and continuous renewal function there always exists a number $b=\tau+c(p)$ at which the above difference becomes negative, taking into account the boundedness of $\bar{f}_0(p)$, cf. Lemma 4. In other words, $c_{\rho'_n}(\tau,p)$ tends to a function $c(p)$ that is independent of τ and $k_{\rho'_n}(p)$ tends to infinity for all $p\geq 0$, cf. (38).

Proposition 3. *There exists a function $c(p)$, $p\geq 0$ that satisfies*

a. $\bar{f}_0(p) = 1-(1/\lambda)g[1+M(c(p))], \quad p\geq 0.$

b. $c(p)\geq 0, \quad c(0)=0.$

c. $c(p)$ *is increasing and bounded.*

Proof. Parts b and c can be proved from Proposition 2, apart from the boundedness. Since $\bar{V}_0^\rho(0) = 1+\alpha(\rho)v_0(\rho)$, cf. (29), it is readily seen that $\bar{f}_0(0) = 1-(1/\lambda)g$. Hence from (45) and part a, it follows that

$$M(c(p)) \leq \frac{\lambda}{g}A(p), \quad p\geq 0. \quad (46)$$

So $c(p)$ must be bounded.

Corollary 4

$$g = \lambda \Big/ \left\{1 + \int_0^\infty M(c(p))\,dB(p)\right\}. \quad (47)$$

Proof. Since

$$\int_0^\infty [\bar{V}_0^\rho(p) - v_0(\rho)]\,dB(p) = 0$$

it follows that

$$\int_0^\infty \bar{f}_0(p)\,dB(p) = 0.$$

Integrating the expression in part a of Proposition 3 then leads to the desired result.

The solution to equation (43) is

$$g\tau + f_0(b;\tau,p)$$
$$= 2 - \frac{g}{\lambda}[1+M(c(p))], \quad b<\tau$$
$$= 1 - \frac{g}{\lambda}[1+M(b-\tau)], \quad \tau\leq b<\tau+c(p)$$
$$= 1 - \frac{g}{\lambda}[1+M(c(p))], \quad b\geq\tau+c(p) \quad (48)$$

in which $c(\cdot)$ is still to be determined. The form of the corresponding policy is stated in the next theorem.

Theorem 3. *There exists an optimal stationary policy such that*

$$a_n(0;b;\tau,p) = \begin{cases} 0, & b\geq\tau+c(p), \quad \tau\geq 0, \quad p\geq 0 \\ 1, & \textit{otherwise.} \end{cases}$$

Moreover

$$g=\lambda \max_{\pi} \lim_{N\to\infty} E_{\pi}\left\{\sum_{n=1}^{N} r(a_n, s_n) \mid s_1\right\} \Big/ N, \quad s_1 \in S.$$

Proof. The proof follows from Theorem 2 and a theorem of Ross (p. 93), once we have shown that

$$\lim_{n\to\infty} E_{\pi}\{f_{X_n}(Y_n;\ T_{n+1}-T_n,\ P_{n+1}) \mid X_1 = i,\ Y_n = b$$

$$T_2 = \tau,\ P_2 = p)\}/n = 0$$

for $0 \leqslant b < \infty$, $0 \leqslant \tau, p \geqslant 0$, and all admissible policies π. This condition holds since $c(\cdot)$ is bounded and $E_{\pi}\{T_{n+1} - T_n\} < \infty$ is independent of the policy used.

Finally, we give an equation from which $c(\cdot)$ can be determined, at least in principle. Observe that $c(\cdot)$ is implicitly determined by (48) and part a of Proposition 3. In particular, averaging relation (48) w.r.t. $A(\tau)$ and $B(p)$ and equating the result for $f_0(b)$ to the expression in part a of Proposition 3, gives an equation for $c(\cdot)$. We will omit the calculations but note the following.

When first averaging (48) w.r.t. $A(\tau)$, observe that there are two possibilities: either $b < c(p)$ or $b \geqslant c(p)$. If $b \geqslant c(p)$, one should distinguish between two cases when subsequently averaging over $B(p)$: either $b < c(\infty)$ or $b \geqslant c(\infty)$, in which $c(\infty) = \lim_{p\to\infty} c(p)$. In the latter case, $c(p) \leqslant b$ holds for all $p \geqslant 0$. In the former case, $c(p) \leqslant b$ induces the integration interval $0 \leqslant p \leqslant \gamma(b)$, in which $\gamma(\cdot)$ denotes the inverse function of $c(\cdot)$, which exists because $c(\cdot)$ is increasing. Ultimately, by using Corollary 3, we obtain

$$M(c(b)) = M(b) - \int_0^{\gamma(b)} \left\{ \int_0^{b-c(p)} [M(b-\tau) - M(c(p))]\, dA(\tau) \right\} dB(p), \qquad 0 \leqslant b < c(\infty) \quad (49)$$

$$M(c(b)) = M(b) - \int_0^{\infty} \left\{ \int_0^{b-c(p)} [M(b-\tau) - M(c(p))]\, dA(\tau) \right\} dB(p), \qquad b \geqslant c(\infty). \quad (50)$$

Using the renewal equation and integration by parts these relations can be rewritten, respectively, as

$$M(c(b)) = A(b) + \int_0^{\infty} \left\{ \int_0^{c(p)} A(b-u)\, dM(u) \right\} dB(p) - \int_{\gamma(b)}^{\infty} \left\{ \int_b^{c(p)} A(b-u)\, dM(u) \right\} dB(p) \quad (51)$$

and

$$M(c(b)) = A(b) + \int_0^{\infty} \left\{ \int_0^{c(p)} A(b-u)\, dM(u) \right\} dB(p), \qquad b \geqslant c(\infty). \quad (52)$$

Proposition 4

$$g = \lambda / M(c(\infty)) \quad (53)$$

$$M(c(b)) \geqslant A(b), \quad b \geqslant 0 \quad (54)$$

$$c(b) \leqslant b, \quad b \geqslant 0. \quad (55)$$

Proof. Letting $b \to \infty$ in (52), the first assertion follows from Corollary 4 and the uniform convergence of the integral. Relation (54) is obvious from (51) and (52). Relation (55) is easily verified from (49) and (50).

The inequality $c(p) \leqslant p$ is evident, since if $b > \tau + p$, the p-customer is completed before the b-customer and, consequently, the latter customer will certainly be rejected. Keep in mind that the objective is to maximize the average reward. Hence p is an upper bound for $c(p)$.

It follows from (41), (46) and (54) that

$$A(p) \leqslant M(c(p)) \leqslant A(p)\left\{1 + \int_0^{\infty} M(b)\, dB(b)\right\}. \quad (56)$$

For Poisson arrivals $M(t) = \lambda t$ so that

$$\frac{1}{\lambda} A(p) \leqslant c(p) \leqslant \left(\frac{1}{\lambda} + E(b)\right) A(p) \quad (57)$$

and

$$\frac{1}{\lambda} \leqslant c(\infty) \leqslant \frac{1}{\lambda} + E(b). \quad (58)$$

Apparently, for Poisson arrivals one will always reject a customer if $b > \tau + 1/\lambda + E(b)$. This result is intuitively clear noticing that $1/\lambda + E(b)$ is the expected service time plus the average waiting time for the next customer to arrive after service completion.

Since, in general, $M(t) \geqslant \lambda t - 1$ (see Barlow and Proschan 1965, p. 53), it follows that

$$c(p) \leqslant \frac{1}{\lambda} + \frac{1}{\lambda} A(p) \left\{ 1 + \int_0^\infty M(b)\, dB(b) \right\}. \qquad (59)$$

If $A(\cdot)$ belongs to the class NBUE (new better than used in expectation) (see Barlow and Proschan, p. 53), then $M(t) \leqslant \lambda t$, and we conclude from (56) and (59) that in this case

$$\frac{1}{\lambda} \leqslant c(\infty) \leqslant \frac{2}{\lambda} + E(b).$$

One may wonder whether the equation for $c(\cdot)$ possesses a closed-form solution for a particular choice of the underlying probability distributions. Even in the case of Poisson arrivals and negative exponentially distributed service times, the answer seems to be negative.

For Poisson arrivals, we obviously have $M(t) = \lambda t$, $t \geqslant 0$. It is readily seen that (52) yields

$$c(b) = \frac{1}{\lambda} + \int_0^\infty c(p)\, dB(p) - \frac{1}{\lambda} e^{-\lambda b} \int_0^\infty \exp\{\lambda c(p)\}\, dB(p), \quad b \geqslant c(\infty) \qquad (60)$$

while (51) can equivalently be replaced by the integro-differential equation

$$c'(c(x)) = 1 - B(x) + e^{-\lambda c(x)} \int_0^x \exp\{\lambda c(p)\}\, dB(p), \quad x \geqslant 0 \qquad (61)$$

with $c(0) = 0$.

Remark 4

If the service times are deterministic and equal β, then the equations can be solved to give $c(\beta) = \beta$. Consequently, all customers that arrive when the server is idle will be accepted because $\beta < \tau + \beta$, $\tau > 0$. This result is obvious since the decision problem makes sense only for variable service times.

REFERENCES

BARLOW, R. E., AND F. PROSCHAN. 1965. *Mathematical Theory of Reliability*. John Wiley, New York.

NAWIJN, W. M. 1985. The Optimal Look-Ahead Policy for Admission to a Single Server System. *Opns. Res.* **33,** 625–643.

ROSS, S. M. 1983. *Introduction to Stochastic Dynamic Programming*. Academic Press, New York.

STIDHAM, S. 1985. Optimal Control of Admission to a Queueing System. *IEEE Trans. Autom. Control* **AC-30,** 705–713.