stichting
mathematisch
centrum

MC

DEPARTMENT OF OPERATIONS RESEARCH BW 62/76 APRIL

G. DE LEVE, A. FEDERGRUEN & H.C. TIJMS

A GENERAL MARKOV DECISION METHOD, II: APPLICATIONS

Prepublication

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 90C45, 60K25

A General Markov Decision Method, II: Applications*)

by

G. De Leve, A. Federgruen & H.C. Tijms

ABSTRACT

In a preceding paper we have introduced a new approach for solving a wide class of Markov decision problems in which the state space may be general and the system may be continuously controlled. The criterion is the average cost. This paper discusses three applications of this approach. The first application considers an inventory-queueing system in which the workload can be controlled by choosing between two constant processing rates. The second application concerns a house-selling problem in which a constructor builds houses which may be sold at any stage of the construction and potential customers make offers depending on the stage of the construction. The third application considers an M/M/c queueing problem in which the number of operating servers can be controlled by turning servers on or off.

KEY WORDS & PHRASES: *Markov decision problems, average cost, general state space, continuous control, applications, inventory-queueing problem, house-selling problem, M/M/c queueing problem with variable number of servers.*

---

*) This paper is not for review; it is meant for publication elsewhere.

1. INTRODUCTION

In a preceding paper [4] we have introduced a new approach for solving a wide class of Markov decision problems with the *average cost* as criterion including problems in which the state space is general and the system can be continuously controlled. This paper discusses three applications of this approach. Each of these applications will be illustrated with numerical results.

The first application considers an inventory-queueing system in which the workload can be controlled by choosing between two constant processing rates. Using a formula developed in [4] for the average cost of a policy, we derive for the case of fixed switch-over costs an expression for the average cost of a control policy characterized by two switch-over levels.

The second application concerns a house-selling problem in which a constructor builds houses which may be sold at any stage of the construction and potential customers make offers depending on the stage of the construction. From the optimality equation given in [4], an integral-differential equation for the curve determining an optimal policy for accepting offers is derived.

The third application considers the well-known M/M/c queueing problem in which the number of servers turned on is variable. Using a general policy-iteration method developed in [4], we derive a special policy-iteration algorithm which exploits the structure of this problem and calculates an optimal policy within a certain class of structured policies for controlling the number of servers turned on.

In this paper we will follow the notation introduced in [4].

2. A CONTROL POLICY FOR AN INVENTORY-QUEUING SYSTEM WITH TWO CONSTANT PROCESSING RATES AND SWITCH-OVER COSTS

2.1. INTRODUCTION

We consider a single-server station where jobs arrive in accordance with a Poisson process with rate λ. Each job involves an amount of work. The amounts of work of the jobs are known upon arrival and are independently sampled from a general distribution having probability distribution function F with $F(0) = 0$ and finite first two moments μ and $\mu^{(2)}$. At any time the server may choose between the processing rates 1 and 2. When the server is in service and uses processing rate i an amount of work σ_i will be processed per unit time, $i = 1,2$. It is assumed that $\sigma_2 > \sigma_1 > 0$ and $\lambda\mu/\sigma_2 < 1$. Define the workload at time t as the total amount of work remaining to be processed in the system at time t, $t \geq 0$. The server provides service when the system is not empty and uses the following switch-over policy. The server switches from rate 1 to rate 2 only when the workload exceeds the level y_1 and switches from rate 2 to rate 1 only when the workload falls to the level y_2, where y_1 and y_2 are given numbers with $0 < y_2 \leq y_1$. It is assumed that it takes no time to switch from one processing rate to another.

The following costs are incurred. There is a holding cost of $h > 0$ per unit work in the system per unit time. When the server is busy and uses service rate i there is a service cost at rate $r_i \geq 0$, $i = 1,2$. There is a service cost at rate $r_0 \geq 0$ when the system is empty. The cost of switching from rate 1(2) to rate 2(1) is $K_1(K_2)$ where $K_1,K_2 \geq 0$ (we note that actually the analysis below permits also the holding cost to depend on the processing rate used and the switch-over cost to depend on the workload level at which the processing rate is changed).

Denote the above policy as the (y_1,y_2) policy. For the case where $K_1 = K_2 = 0$ and $\lambda\mu/\sigma_1 < 1$ the (y_1,y_2) policy with $y_1 = y_2$ was studied by THATCHER [15] who derived by busy-period analysis a formula for the average cost of this policy and proved that such a policy is average cost optimal among the class of stationary policies, cf. also DOSHI [5]. Related work was done by COHEN [2] who derived for the (y_1,y_2) policy with $y_1 = y_2$ several interesting quantities as the stationary distribution of the workload. In TIJMS [16] a formula for the average cost of the (y_1,y_2) policy was found for the M/M/1 queue with $\lambda\mu/\sigma_1 < 1$.

In this paper we use the approach in [4] in order to derive a formula for the average cost of the (y_1,y_2) policy. To do this, we consider a Markov decision problem with a single decision process associated with a fixed (y_1,y_2) policy. In section 2.2 we specify the elements 1-6 of section 2 of [4]. Next in section 2.3 we study for the (y_1,y_2) policy an embedded decision process and give the formula for the average cost. Finally, in section 2.4 we give some numerical results for the M/M/1 queue.

2.2. THE ELEMENTS

The state space, natural process and the feasible decisions will be of course specified to measure the (y_1,y_2) policy. Before doing this, we make the following observations. The natural process and the intervention must be chosen in such a way that the result of the natural process and the control by the interventions agrees with the process describing the workload when the (y_1,y_2) policy is used. However, these choices determine the set A_0. In its turn the set A_0 is determinative for the calculation of the k- and t-functions. It will be obvious that we shall try to choose the natural process and the interventions in such a way that the resulting set A_0 allows for a simple calculation of the k- and t-functions. Clearly, a convenient choice for the natural process is one where the server never switches from one processing rate to another. For this choice it would be pleasant when the state 0 (say) corresponding to the situation in which the system becomes empty while the server is adjusted to rate 1 belongs to A_0. However, in this state the (y_1,y_2) policy prescribes no change of the processing rate. Nevertheless, we can always achieve that state 0 is an intervention state for the (y_1,y_2) policy by choosing the natural process such that state 0 is an absorbing state for this process, e.g. imagine that in the natural process the service station is closed down in this state. This has as a consequence that we have to introduce both a fictitious intervention for state 0 (e.g. imagine that this intervention immediately reopens the station) and a fictitous state to which the system is instantaneously transferred by this intervention. All of this can be done provided that the result of the natural process and the control by the interventions agrees with the process describing the workload under the (y_1,y_2) policy, cf. section 2 of [4]. This observation will be used in the specification of the elements 1-6, of section 2 of [4].

We choose as state space

$$X = \{u \mid u \text{ real}, u \geq 0\} \cup \{u' \mid u \text{ real}, u \geq 0\} \cup \{\bar{0}\}.$$

State u(u') corresponds to the situation where the workload equals u and the server is adjusted to rate 1 (2). In addition, in state 0 the station is closed down. State $\bar{0}$ corresponds to the situation where the workload is zero, the station is open and the server is adjusted to rate 1.

The natural process is chosen such that in the natural process the server never switches from one processing rate to another. For any initial state u' we choose the natural process as the process describing the workload when always processing rate 2 is used. For initial state $u > 0$ the natural process is chosen as the process describing the workload under the use of processing rate 1 as long as the system is not empty. If the system becomes empty under rate 1, the natural process assumes state 0. This state is taken to be an absorbing state for the natural process. When the initial state is $\bar{0}$ the natural process stays in this state until the next job arrives. Then the natural process assumes state y when this job involves an amount of work y.

Since we consider a fixed (y_1,y_2) policy, the set of feasible decisions in each state consists of a single decision. We take both in state u with $0 < u \le y_1$, state u' with $u > y_2$ and in state $\bar{0}$ the null-decision is the only feasible decision. The null-decision does not disturb the natural process. In the other states the intervention $d = 1$ is the only possible decision. The intervention $d = 1$ in state u' with $0 \le u \le y_2$ prescribes to switch from rate 2 to rate 1 and causes an instantaneous transition to state u when $u > 0$ and to state $\bar{0}$ when $u = 0$. The intervention $d = 1$ in state u with $u > y_1$ prescribes to switch from rate 1 to rate 2 and causes an instantaneous transition to state u'. Finally the intervention $d = 1$ in state 0 prescribes to re-open the station and causes an instantaneous transition to state $\bar{0}$.

We take the following cost structure. In the natural process there is a holding cost at rate hu both in the states u and u', there is a service cost at rate $r_1(r_2)$ in state $u \neq 0(u' \neq 0')$ and a service cost at rate r_0 in each of the states 0, $\bar{0}$ and 0'. Further, there is an immediate decision cost of K_1 for taking intervention $d = 1$ in state u with $u > y_1$ and an immediate decision cost of K_2 for taking intervention $d = 1$ in state u' with $u \le y_2$.

Now, it will be clear that the result of this natural process and the control by the above decisions agrees with the process describing the workload under the (y_1,y_2) policy. Now, by the above choices, element 4 in [4] applies with

$$A_0 = \{0\} \cup \{u \mid u > y_1\} \cup \{u' \mid 0 \le u \le y_2\}.$$

To calculate the k-and t-functions introduced in element 5 of [4], we choose

$$A_{01} = A_{02} = \{0\} \cup \{u \mid u > y_1\} \cup \{0'\}.$$

Before we calculate the k-and t-functions, we first discuss the following "renewal-type" equation

$$(2.1) \qquad u(x) = a(x) + \int_0^{y_1-x} u(x+y)dH(y), \qquad 0 < x < y_1,$$

where a(x) is a given function, u(x) is unknown and H is defined by

$$H(x) = \frac{\lambda}{\sigma_1}\int_0^x \{1-F(y)\}dy \qquad \text{for } x \ge 0.$$

The solution of such an equation has been derived in COHEN [3] To give this solution, we define $\delta = 0$ when $\lambda\mu/\sigma_1 \le 1$ and define δ as the unique positive root to

$$\int_0^\infty e^{-xy}dH(y)-1 = 0$$

when $\lambda\mu/\sigma_1 > 1$. Further, we define the function G by $G(x) = 0$ for $x < 0$ and

$$G(x) = \int_0^x e^{-\delta y}dH(y) \qquad \text{for } x \ge 0.$$

Then G is a proper(defective) probability distribution function when $\lambda\mu/\sigma_1 \ge 1$ $(\lambda\mu/\sigma_1<1)$. Next we define the renewal function M by

$$M(x) = \sum_{n=1}^{\infty} G^n(x) \qquad \text{for } x \ge 0,$$

where G^n is the n-fold convolution of G with itself. Letting $\bar{u}(x) = e^{\delta x}u(x)$ and $\bar{a}(x) = e^{\delta x}a(x)$, we can write (2.1) in the equivalent form

$$\bar{u}(x) = \bar{a}(x) + \int_0^{y_1-x} \bar{u}(x+y)dG(y), \qquad 0 < x < y_1.$$

The solution of this renewal equation (cf. FELLER [6]) yields

$$u(x) = a(x) + \int_0^{y_1-x} e^{\delta y}a(x+y)dM(y), \qquad 0 < x < y_1 \tag{2.2}$$

From the definition of the k-and t-functions given in section 2 of [4],

$t(u;1) = t_0(u') - t_0(u)$, $k(u;1) = K_1 + k_0(u') - k_0(u)$ for $u > y_1$,
$t(u';1) = t_0(u) - t_0(u')$, $k(u';1) = K_2 + k_0(u) - k_0(u')$ for $0 < u \le y_2$,
$t(0;1) = t_0(\bar{0})$ and $k(0;1) = k_0(\bar{0})$.

We now determine the functions t_0 and k_0. By the choice of A_{02} and considering what may happen in a small time interval of length Δu, we get for $0 < u < y_1$,

$$t_0(u+\Delta u) = \frac{\Delta u}{\sigma_1} + \lambda\frac{\Delta u}{\sigma_1}\int_0^{y_1-u} t_0(u+y)dF(y) + (1-\lambda\frac{\Delta u}{\sigma_1})t_0(u) + 0(\Delta u)$$

so, for $0 < u < y_1$,

$$t_0'(u) = \frac{1}{\sigma_1} - \frac{\lambda}{\sigma_1}t_0(u) + \frac{\lambda}{\sigma_1}\int_0^{y_1-u} t_0(u+y)dF(y)$$

Using the relation

$$\frac{\partial}{\partial x}\int_0^{y_1-x} a(x+y)\ \{1-F(y)\}\ dy = -a(x) + \int_0^{y_1-x} a(x+y)dF(y), \qquad x \ge 0 \tag{2.3}$$

we get, for some constant a,

$$t_0(u) = \frac{u}{\sigma_1} + a + \int_0^{y_1-u} t_0(u+y)dH(y), \qquad 0 < u < y_1.$$

Together this relation, (2.1)-(2.2) and the fact that $\lim_{u\to 0} t_0(u) = t_0(0) = 0$ imply

$$t_0(u) = \frac{u}{\sigma_1} + a + \int_0^{y_1-u} e^{\delta y}\ (\frac{u+y}{\sigma_1}+a)dM(y), \qquad 0 \le u \le y_1$$

where

$$a = -\int_0^{y_1} ye^{\delta y}dM(y)/\sigma_1\{1+\int_0^{y_1} e^{\delta y}dM(y)\}.$$

Similarly, from

$$k_0(u+\Delta u) = hu\frac{\Delta u}{\sigma_1} + r_1\frac{\Delta u}{\sigma_1} + \lambda\frac{\Delta u}{\sigma_1}\int_0^{y_1-u} k_0(u+y)dF(y) + (1-\lambda\frac{\Delta u}{\sigma_1})k_0(u) + O(\Delta u), \quad 0 < u < y_1,$$

we derive

$$k_0(u) = \frac{hu^2}{2\sigma_1} + \frac{r_1 u}{\sigma_1} + b + \int_0^{y_1-u} e^{\delta y}\{\frac{h(u+y)^2}{2\sigma_1} + \frac{r_1(u+y)}{\sigma_1} + b\}\, dM(y),$$

$$0 \le u \le y_1$$

where

$$b = -\int_0^{y_1} (\frac{hy^2}{2} + r_1 y)e^{\delta y}dM(y)/\sigma_1\{1+\int_0^{y_1} e^{\delta y}dM(y)\}.$$

Next we find

$$t_0(\bar{0}) = \frac{1}{\lambda} + \int_0^{y_1} t_0(y)dF(y) \text{ and } k_0(\bar{0}) = \frac{r_0}{\lambda} + \int_0^{y_1} k_0(y)dF(y).$$

Finally, we have that $t_0(u')$ and $k_0(u')$ are equal to the expected time until the system is empty and the expected holding and service costs incurred until the system is empty when the initial workload is u and always processing rate 2 is used. Using standard arguments from busy-period analysis, it is routine to derive (e.g. Theorem 4 in THATCHER [15] and Theorem 1 in TIJMS [16])

$$t_0(u') = \frac{u}{\sigma_2-\lambda\mu} \text{ and } k_0(u') = \frac{hu^2}{2(\sigma_2-\lambda\mu)} + \frac{h\lambda\mu^{(2)}u}{2(\sigma_2-\lambda\mu)^2} + \frac{r_2 u}{\sigma_2-\lambda\mu}, \quad u \ge 0.$$

2.3. THE AVERAGE COST OF THE (y_1,y_2) POLICY.

In this section we determine a formula for the average cost of the (y_1,y_2) policy by using Theorem 1 of [4]. To do this, we have first to study the embedded Markov chain $\{I_n\}$ describing the state of the system at the epochs at which the system enters the set A_0 of intervention states of the (y_1,y_2) policy, see section 3 of [4]. Observe that for the present problem the class Z of policies introduced in element 7 of [4] consists only of the (y_1,y_2) policy. Clearly, the assumptions A1-A3 in [4] are satisfied (take $s_z = y_2'$ in A2). Denote by Q the unique stationary probability measure of the above embedded Markov chain and for ease of notation write $Q_0 = Q(\{0\})$, $Q(v) = Q(\{u \mid u > v\})$ for $v \ge y_1$ and $Q_2 = Q(\{y_2'\})$. To determine these probabilities, we define for all $0 < u \le y_1$ and $v \ge y_1$,

$p(u,v)$ = probabiltiy that the state of the first entry of the natural process into the set $\{0\} \cup \{x \mid x > y_1\}$ belongs to the set $\{x \mid x > v\}$ given that the initial state is u,

and we define $p_0(u) = 1-p(u,y_1)$ for $0 < u \le y_1$. Then, by the steady state equation (27) in [4], we have

$$Q(v) = Q_0\,\{1-F(v) + \int_0^{y_1} p(y,v)dF(y)\} + Q_2p(y_2,v), \qquad v \ge y_1 \tag{2.4}$$

$$Q_0 = Q_0 \int_0^{y_1} p_0(y)dF(y) + Q_2p_0(y_2) \text{ and } Q_2 = Q(y_1). \tag{2.5}$$

From (2.4)-(2.5) and $Q_0 + Q(y_1) + Q_2 = 1$, we get

$$Q_0 = \frac{p_0(y_2)}{2}\{1 + \frac{p_0(y_2)}{2} - \int_0^{y_1} p_0(y)dF(y)\}^{-1} \text{ and } Q_2 = \frac{1-Q_0}{2}. \tag{2.6}$$

The stationary distribution Q is now given by (2.4) and (2.6). It remains to determine $p(u,v)$. Using the fact that processing rate 1 is used in the natural process starting from state u, we have for all $0 < u < y_1$ and $v \ge y_1$,

$$p(u+\Delta u,v) = \lambda\frac{\Delta u}{\sigma_1}\{1 - F(v-u) + \int_0^{y_1-u} p(u+y,v)dF(y)\} + $$
$$+ (1-\lambda\frac{\Delta u}{\sigma_1})p(u,v) + O(\Delta u)$$

from which we get for all $0 < u < y_1$ and $v \geq y_1$

$$\frac{\partial p(u,v)}{\partial u} = \frac{\lambda}{\sigma_1}\{1 - F(v-u) - p(u,v) + \int_0^{y_1-u} p(u+y,v)dF(y)\}.$$

It follows from this relation and (2.1)-(2.3) that for all $v \geq y_1$

$$p(u,v) = \phi(u,v) + \int_0^{y_1-u} e^{\delta y}\phi(u+y,v)dM(y), \qquad 0 < u \leq y_1$$

where $\phi(u,v) = c_v + H(v) - H(v-u)$ for some constant c_v. Since $p(u,v) \to 0$ as $u \to 0$, we get for all $v \geq y_1$

$$c_v = - \int_0^{y_1} e^{\delta y}\{H(v) - H(v-y)\}dM(y)/\{1 + \int_0^{y_1} e^{\delta y}dM(y)\}.$$

We can now give a formula for the average cost of the (y_1,y_2) policy. By Theorem 1 of [4] this average cost equals

$$g(y_1,y_2) = \int_{A_0} k(x;1)Q(dx)/\int_{A_0} t(x;1)Q(dx).$$

All quantities appearing in the right side of this formula have been explicitly determined above, but they involve the number δ and the renewal function M.

2.4 NUMERICAL RESULTS

In this section we give some numerical results for the case where F is an exponential distribution function with mean $1/\gamma$. We then find for the case of $\lambda/\sigma_1\gamma \leq 1$,

$$\delta = 0 \text{ and } \frac{dM(x)}{dx} = \frac{\lambda}{\sigma_1} e^{-(\gamma-\lambda/\sigma_1)x} \qquad \text{for } x > 0$$

and for the case of $\lambda/\sigma_1\gamma>1$,

$$\delta = \frac{\lambda}{\sigma_1} - \gamma \text{ and } \frac{dM(x)}{dx} = \frac{\lambda}{\sigma_1} \qquad \text{for } x > 0.$$

Observe that in both cases $e^{\delta x}\, dM(x)/dx$ is the same. We next obtain after elementary but lengthy calculations

$$g(y_1,y_2) = \frac{\alpha_0 r(y_1,y_2) + \alpha_1(y_1^2-y_2^2) + \alpha_2(y_1-y_2) + \alpha_3 y_1 + (\alpha_2+\alpha_3)/\gamma + K}{\beta_0 R(y_1,y_2) + \beta_1(y_1-y_2) + \beta_1/\gamma}$$

where

$$K = K_1 + K_2,\ R(y_1,y_2) = (\sigma_1\gamma-\lambda)^{-1}\left\{\sigma_1\gamma e^{(\sigma_1\gamma-\lambda)y_1/\sigma_1} - \lambda e^{(\sigma_1\gamma-\lambda)y_2/\sigma_1}\right\},$$

$$\alpha_0 = \frac{r_0-r_1}{\lambda} + \frac{r_1\sigma_1\gamma}{\lambda(\sigma_1\gamma-\lambda)} + \frac{h\sigma_1}{(\sigma_1\gamma-\lambda)^2},\quad \alpha_1 = \frac{h\gamma^2(\sigma_1-\sigma_2)}{2(\sigma_1\gamma-\lambda)(\sigma_2\gamma-\lambda)},$$

$$\alpha_2 = \frac{h\lambda}{(\sigma_2\gamma-\lambda)^2} - \frac{h\lambda}{(\sigma_1\gamma-\lambda)^2} + \frac{r_2\gamma}{(\sigma_2\gamma-\lambda)} - \frac{r_1\gamma}{(\sigma_1\gamma-\lambda)},\quad \alpha_3 = \frac{h\gamma(\sigma_1-\sigma_2)}{(\sigma_1\gamma-\lambda)(\sigma_2\gamma-\lambda)},$$

$$\beta_0 = \frac{\sigma_1\gamma}{\lambda(\sigma_1\gamma-\lambda)},\quad \beta_1 = \frac{\gamma^2(\sigma_1-\sigma_2)}{(\sigma_1\gamma-\lambda)(\sigma_2\gamma-\lambda)}.$$

The formula for $g(y_1,y_2)$ applies for any value of $\lambda/\sigma_1\gamma$ except for the value 1 for which the expression for $g(y_1,y_2)$ is obtained from the above one by letting $\lambda \to \sigma_1\gamma$. To save space, we omit the formula for the case of $\lambda/\sigma_1\gamma = 1$. It should be noted that for any (y_1,y_2) policy the average cost $g(y_1,y_2)$ may be larger than the average of the policy that always uses rate 2. The average cost of the latter policy is given by

$$g_2 = r_0(1-\lambda/\sigma_2\gamma) + r_2\lambda/\sigma_2\gamma + h\lambda/\gamma(\sigma_2\gamma-\lambda).$$

Using a computer program based on a unconstrained minimization algorithm of FLETCHER [7], we have computed the values y_1^* and y_2^* for which the function $g(y_1,y_2)$ is minimal for $0 < y_2 \le y_1$. In table 1 we give some numerical results.

TABLE 1. $\mu = 1.25$, $\sigma_1 = 3$, $\sigma_2 = 5$, $h = 1$, $r_0 = 0$, $r_1 = 5$ and $r_2 = 25$

	λ	3	3.5	4	4.5
K = 0	y_1^*	7.038	5.304	4.061	3.127
	y_2^*	7.038	5.304	4.061	3.127
	$g(y_1^*,y_2^*)$	6.776	9.322	12.483	16.184
K = 10	y_1^*	11.870	9.516	7.930	6.739
	y_2^*	5.347	3.482	2.321	1.567
	$g(y_1^*,y_2^*)$	6.925	9.842	13.479	17.567
K = 25	y_1^*	14.213	11.407	9.660	8.425
	y_2^*	4.928	2.925	1.765	1.062
	$g(y_1^*,y_2^*)$	7.007	10.226	14.289	18.726
	g_2	12.738	15.018	17.422	20.057

3. A HOUSE SELLING PROBLEM

3.1 INTRODUCTION

Consider a building-contractor constructing identical houses which may be sold in any stage of the construction. The construction time that is needed to perform a fraction y of the total building of a house has a gamma probability distribution fuction with density

$$g(t|y) = \frac{1}{\Gamma(cy)} a^{cy} t^{cy-1} e^{-ay}, \qquad t \geq 0,$$

where a, c > 0. Observe that this distribution has mean cy/a and that the distribution of the sum of the construction time of a fraction y_1 and that of a fraction y_2 has the same distribution as the construction time of a fraction $y_1 + y_2$, cf. p.46 in FELLER [6].

Potential customers for the houses arrive in accordance with a Poisson process with rate λ. Each potential customer makes an offer where the amount of money offered has a probability distribution function $F(0|y)$ with finite mean when a fraction y of the total construction has been completed. If the offer is accepted, the house is sold and the contractor immediately starts with the construction of a new house. In case the building of a house is completed without any offer having been accepted in the mean time, the house will always be sold for an amount of K. Finally, there are building costs at rate b(y) when a fraction y of the total construction of the house has been completed.

Using the optimality equation (16) of [4] we shall characterize the structure of an average cost optimal policy and show that this policy is in fact determined by an integral-differential equation. This will be done in section 3.3 after in section 3.2 we have specified the elements 1-6 of [4]. Finally, in section 3.4 we give some numerical results.

3.2 THE ELEMENTS

We first note that the state space, the natural process and the feasible decisions must be chosen such that element 4 of [4] applies. To achieve this, a convenient choice of the natural process is one in which the constructor accepts every offer and no new construction is started once a house is sold. This choice involves the introduction of an absorbing state E(say) for the natural process. We now choose as state space

$$X = \{y \mid 0 \le y < 1\} \cup \{(y_1,y_2) \mid 0 \le y_1 < 1,\ y_2 \ge 0\} \cup \{E\}.$$

State y corresponds to the situation where a house is under construction and a fraction y of the total construction has been completed, while no offer is currently made. State (y_1,y_2) corresponds to the situation where a offer of size y_2 is made for a house of which a fraction y_1 of the total construction has been completed. State E corresponds to the situation where no house is under construction. The natural process is chosen as follows. Starting from state y^0 the natural process moves along the states y with $y^0 \le y < 1$ until either a offer is made or the construction of the house is completed. In case of a offer of size y_2 in state y the natural process jumps to state (y,y_2) (i.e. any offer is accepted in the natural process), while in case of completion of the construction the natural process jumps to state E. The natural process starting from state (y_1,y_2) jumps immediately to state E. We take state E as an absorbing state for the natural process (e.g. imagine that in the natural process the contractor closes down his work in state E). We next choose the feasible decisions. For each state y the only feasible decision is the null-decision which leaves the natural process untouched. For any state (y_1,y_2) the feasible decisions consists of the null-decision which prescribes to accept the offer and causes an instantaneous transition to state E and the intervention d = 1 which prescribes to refuse the offer and causes an instantaneous transition to state y_1. The only feasible decision is state E is the intervention d = 1 which prescribes to start with a new construction and causes an instantaneous change to state 0. The following costs are associated to the natural process and the interventions. In the natural process there is incurred a cost at rate b(y) when the natural process is in state y. Further, when the natural process makes a transition to state (y_1,y_2) there is incurred a cost of $-y_2$ and when the natural process makes a transition to state E after completion of a construction there is incurred a cost of -K. Finally, by the above choices, there is no cost associated with any intervention.

Now, for any policy the superimposition of the natural process and the interventions prescribed by that policy agrees with the evolution of the system resulting from the specific control as executed by the decisionmaker. Clearly, element 4 of [4] applies with

$$A_0 = \{E\}.$$

We choose $A_{01} = A_{02} = A_0$ in order to determine the k-and t-functions, see [4]. Clearly, for all $(y_1,y_2) \in X$,

$$t((y_1,y_2);1) = t_0(y_1) - t_0((y_1,y_2) \text{ and } k((y_1,y_2);1) =$$
$$= k_0(y_1) - k_0((y_1,y_2))$$

and

$$t(E;1) = t_0(0) \text{ and } k(E;1) = k_0(0).$$

Since the natural process starting from state (y_1,y_2) immediately jumps to state E, we have for all (y_1,y_2),

$$t_0((y_1,y_2)) = 0 \text{ and } k_0((y_1,y_2)) = -y_2.$$

Further, $t_0(y) = E\min[A,T(y)]$ for all $0 \le y < 1$ where A and T(y) are independent random variables such that A is exponentially distributed with mean $1/\lambda$ and the construction time T(y) has a gamma distribution with density $g(\cdot|1-y)$. We find for $0 \le y < 1$,

$$t_0(y) = \frac{1}{\lambda}\{1 - (\frac{a}{a+\lambda})^{c(1-y)}\}.$$

To determine the function $k_0(y)$, we first make the following observation. The building costs incurred between stages y_0 and y_1 of the construction are given by, for all $0 \le y_0 < y_1 < 1$,

$$\lim_{n\to\infty} \sum_{i=0}^{n} b(y_0 + \frac{i}{n}(y_1-y_0))\frac{c}{a}\left(\frac{y_1-y_0}{n}\right) = \frac{c}{a}\int_{y_0}^{y_1} b(v)dv.$$

Further, for any initial state y with $0 \le y < 1$, let the random variable X_y be equal to 1 when the total construction is completed before a first offer occurs, and let X_y be equal to the stage of the construction at the epoch of the first offer, otherwise. It is routine to verify that, for all $0 \le y < 1$,

$$\Pr\{X_y = 1\} = (\frac{a}{a+\lambda})^{c(1-y)} \text{ and } \Pr\{X_y \le u\} = 1 - (\frac{a}{a+\lambda})^{c(u-y)} \quad \text{for } y \le u < 1.$$

Let $h(u|y)$ be the derivative of $\Pr\{X_y \le u\}$ with respect to u. Then for all $0 \le y < 1$,

$$h(u|y) = c\ln(1+\frac{\lambda}{a})(\frac{a}{a+\lambda})^{c(u-y)} \qquad \text{for } y < u < 1.$$

Now, we have by the choice of the natural process that, for all $0 \le y < 1$,

$$k_0(y) = E[\frac{c}{a}\int_y^{X_y} b(v)dv] - \int_y^1 \{\int_0^\infty vdF(v|u)\}\, h(u|y)du - K\Pr\{X_y = 1\}$$

from which we get after some algebra

$$k_0(y) = \frac{c}{a}\int_y^1 b(u)\left(\frac{a}{a+\lambda}\right)^{c(u-y)}du - \alpha(y) - \left(\frac{a}{a+\lambda}\right)^{c(1-y)}K, \qquad 0 \le y < 1,$$

where

$$(3.1)\qquad \alpha(y) = c\ln(1+\frac{\lambda}{a})\int_y^1 \{\int_0^u vdF(v|u)\}\left(\frac{a}{a+\lambda}\right)^{c(u-y)}du, \qquad 0 \le y < 1.$$

3.3 CHARACTERIZATION OF AN OPTIMAL POLICY

In this section we shall derive from the optimality equation (16) of [4] the existence and the structure of an average cost optimal policy. Moreover, we shall find that in fact such a policy is determined by an integral-differential equation.

Now, let z^* be any policy of Z. Denote by $\{g(z^*), v(z^*;x) \mid x \in X\}$ the unique solution to the equations (8)-(9) with $z = z^*$ of [4] such that

$$(3.2)\qquad v(z^*;E) = 0.$$

Since the intervention d = 1 in state (y_1,y_2) causes an instantaneous transition to state y_1, it follows from relation (11) of [4] and the above formulas for the functions k and t that

$$(3.3)\qquad v(z^*;(y_1,y_2)) = k((y_1,y_2);1) - g(z^*)t((y_1,y_2);1) + v(z^*;y_1) =$$
$$= y_2 + R(z^*;y_1) + v(z^*;y_1) \quad \text{for all } (y_1,y_2) \in A_{z^*},$$

where

$$(3.4)\qquad R(z^*;y) = \frac{c}{a}\int_y^1 b(u)\left(\frac{a}{a+\lambda}\right)^{c(u-y)}du - \alpha(y) - \left(\frac{a}{a+\lambda}\right)^{c(1-y)}K +$$
$$-g(z^*)\left\{1 - \left(\frac{a}{a+\lambda}\right)^{c(1-y)}\right\} \quad \text{for } 0 \le y < 1.$$

By relation (9) of [4] and the fact that the natural process starting from state (y_1,y_2) jumps to the intervention state E, we have

$$(3.5)\qquad v(z^*;(y_1,y_2)) = v(z^*;E) = 0 \quad \text{for } (y_1,y_2) \notin A_{z^*}.$$

Finally, by relation (11) of [4],

$$(3.6) \qquad v(z^*;E) = k(E;1) - g(z^*)t(E;1) + v(z^*;0) = R(z^*;0) + v(z^*;0).$$

Now, let $z \in Z$. Then, by virtue of the fact that the only possible intervention is $d = 1$, it follows from the relations (11) and (13)-(14) of [4] that

$$(3.7) \qquad v([z]z^*;E) = k(E;1) - g(z^*)t(E;1) + v(z^*;0) = v(z^*;E)$$

and

$$(3.8) \qquad v([z]z^*;(y_1,y_2)) = k((y_1,y_2);1) - g(z^*)t((y_1,y_2);1) + v(z^*;y_1) = \\ = y_2 + R(z^*;y_1) + v(z^*;y_1) \quad \text{for all } (y_1,y_2) \in A_z.$$

Further, by definition (14) of [4] and the relations (3.2) and (3.7),

$$(3.9) \qquad v([z]z^*; (y_1,y_2)) = v([z]z^*;E) = 0 \qquad \text{for all } (y_1,y_2) \notin A_z.$$

We shall now prove that a policy $z^* \in Z$ satisfies the optimality equation (see (16) of [4])

$$(3.10) \qquad v(z^*;x) = \min_{z \in Z} v([z]z^*;x) \qquad \text{for all } x \in X_0$$

if and only if for policy z^* holds

$$(3.11) \qquad y_2 + R(z^*;y_1) + v(z^*;y_1) \le 0 \qquad \text{for all } (y_1,y_2) \in A_{z^*}$$

$$(3.12) \qquad y_2 + R(z^*;y_1) + v(z^*;y_1) \ge 0 \qquad \text{for all } (y_1,y_2) \notin A_z .$$

To prove this, we first observe that, by relation (15) of [4] and (3.7), the optimality equation (3.10) is equivalent to

$$(3.13) \qquad v([z]z^*;(y_1,y_2)) \ge v(z^*;(y_1,y_2)) \\ \text{for all } (y_1,y_2) \in X \text{ and all } z \in Z.$$

Suppose first that (3.13) holds. To establish (3.11), we observe that for any state $(y_1,y_2) \in A_{z^*}$ we can find a policy $z \in Z$ such that $(y_1,y_2) \notin A_z$, so, by (3.3), (3.9) and (3.13), we get (3.11). Also, for any state $(y_1,y_2) \notin A_z$ we can find a policy $z \in Z$ such that $(y_1,y_2) \in A_z$, so, by (3.5), (3.8) and (3.13), we get (3.12). Next assume that (3.11)-(3.12) hold. To verify (3.13), fix $z \in Z$. For $(y_1,y_2) \notin A_z$, we get (3.13) from (3.9), (3.5), (3.3) and (3.11). For $(y_1,y_2) \in A_z$, we get (3.13) from (3.8), (3.3), (3.5) and (3.12).

We now have proved that a policy $z^* \in Z$ for which (3.11)-(3.12) hold is optimal. Moreover, we can conclude that such a policy z^* is determined by a function $s(y_1)$, $0 \le y_1 < 1$ such that

$$(3.14) \qquad A_{z^*} = \{(y_1,y_2) \mid y_2 \le s(y_1)\}.$$

Furthermore,

$$(3.15) \qquad s(y_1) = -R(z^*;y_1) - v(z^*;y_1).$$

Since we know the structure of A_{z^*} we can express $v(z^*;y_1)$ in the function $s(\cdot)$. To do this, we first observe that, by (3.3), (3.5), and (3.14)-(3.15), for all (y_1,y_2)

$$(3.16) \qquad v(z^*;(y_1,y_2)) = \begin{cases} y_2 - s(y_1) & \text{for } y_2 \le s(y_1), \\ 0 & \text{for } y_2 \ge s(y_1). \end{cases}$$

Using relation (11) of [4] with $V = \{(y_1,y_2)\} \cup \{E\}$, (3.2) and (3.16), we get

$$(3.17) \qquad v(z^*;y) = \int_y^1 \{ \int_0^\infty v(z^*;(u,v))dF(v|u)\} \, h(u|y)du + v(z^*;E)\Pr\{X_y = 1\} =$$

$$= c\ln(1+\frac{\lambda}{a}) \int_y^1 \{ \int_0^{s(u)} (v - s(u))dF(v|u)\} \, (\frac{a}{a+\lambda})^{c(u-y)}du, \qquad 0 \le y < 1.$$

From this relation and (3.15), we get for $0 \le y_1 < 1$.

$$s(y_1) = -R(z^*;y_1) + c\ln(1+\frac{\lambda}{a}) \int_{y_1}^1 \{ \int_0^{s(u)} (v-s(u))dF(v|u)\} \, (\frac{a}{a+\lambda})^{c(u-y_1)}du.$$

Differentiating this formula and using (3.1) and (3.4), we get after some algebra

$$(3.18) \qquad s'(y_1) = \frac{c}{a}b(y_1) + c\ln(1+\frac{\lambda}{a}) \left\{ \int_{s(y_1)}^{\infty} (s(y_1)-v)dF(v|y_1) - \frac{g(z^*)}{\lambda} \right\}, \quad 0 \le y_1 < 1.$$

Using the fact that $\lim_{y \to 1} v(z^*;y) = 0$ (see (3.17)) and the relations (3.2) and (3.6), we have the boundary conditions

$$(3.19) \qquad s(0) = 0 \text{ and } s(1) = K.$$

The integral-differential equation (3.18) and the boundary conditions (3.19) determine both the curve $s(\cdot)$ giving the optimal policy z^* and the minimal average cost $g(z^*)$.

3.4. NUMERICAL RESULTS.

In this section we give some numerical results for the case where $f(\cdot|y_1)$ is a gamma distribution with density

$$\frac{\{n\lambda(y_1)\}^n}{(n-1)!} v^{n-1} e^{-n\lambda(y_1)v}, \qquad v \ge 0,$$

where n is a positive integer and $\lambda(y_1)$ is a given function. Observe that the mean and the variance of this distribution are equal to $1/\lambda(y_1)$ and $1/n\{\lambda(y_1)\}^2$. By a well-known relation between the Poisson distribution and the gamma distribution, we have

$$\int_{s(y_1)}^{\infty} (s(y_1)-v)dF(v|y_1) = e^{-n\lambda(y_1)s(y_1)} \{s(y_1) \sum_{j=0}^{n-2} [n\lambda(y_1)s(y_1)]^j/j! + - \frac{1}{\lambda(y_1)} \sum_{j=0}^{n-1} [n\lambda(y_1)s(y_1)]^j/j!\}.$$

Hence the relation (3.18) reduces to a differential equations with unknown parameter $g(z^*)$. To solve this differential equations with the boundary conditions (3.19), we have used a computer program developed [17] for parameter estimation in differential equations. In table 2 we give some numerical results.

TABLE 2. $\lambda = 2$, $a = 1$, $c = 1$, $K = 2$, $b(y) = 1$ and $\lambda(y) = 1/(3y+0.01)$

n = 1		n = 5		n = 10	
y	s(y)	y	s(y)	y	s(y)
0.00	0.000	0.00	0.000	0.00	0.000
0.05	0.137	0.05	0.126	0.05	0.124
0.10	0.271	0.10	0.250	0.10	0.245
0.15	0.401	0.15	0.371	0.15	0.365
0.20	0.528	0.20	0.490	0.20	0.483
0.25	0.651	0.25	0.607	0.25	0.598
0.30	0.770	0.30	0.721	0.30	0.712
0.35	0.886	0.35	0.833	0.35	0.823
0.40	0.998	0.40	0.942	0.40	0.931
0.45	1.106	0.45	1.048	0.45	1.037
0.50	1.210	0.50	1.152	0.50	1.141
0.55	1.309	0.55	1.252	0.55	1.241
0.60	1.405	0.60	1.350	0.60	1.339
0.65	1.496	0.65	1.444	0.65	1.434
0.70	1.583	0.70	1.535	0.70	1.526
0.75	1.665	0.75	1.622	0.75	1.614
0.80	1.742	0.80	1.705	0.80	1.699
0.85	1.814	0.85	1.785	0.85	1.780
0.90	1.881	0.90	1.861	0.90	1.858
0.95	1.944	0.95	1.933	0.95	1.931
1.00	2.000	1.00	2.000	1.00	2.000
$g(z^*) = -3.243$		$g(z^*) = -2.816$		$g(z^*) = -2.729$	

4. AN M/M/c QUEUEING PROBLEM WITH A VARIABLE NUMBER OF SERVERS

4.1. INTRODUCTION

We consider the M/M/c queueing problem studied by McGILL [11], where the number of servers operating can be adjusted at arrival and service completion epochs. The customers arrive in accordance with a Poisson process with rate λ and there are c independent servers available each having an exponentially distributed service time with mean $1/\mu$. It is assumed that the lowest possible traffic intensity $\lambda/c\mu$ is less than 1. The cost structure includes a holding cost of $h > 0$ per customer in the system per unit time, an operating cost of $w > 0$ per server turned on per unit time and a switch-over cost of $K(a,b)$ when the number of servers turned on is adjusted from a to b. We assume that

$$K(a,b) = k^{+}.(b-a) \text{ when } a < b \text{ and } K(a,b) = k^{-}.(a-b) \text{ when } a \geq b,$$

where $k^{+}, k^{-} \geq 0$. This problem has been treated amongst others by BELL [1], LIPPMAN [9], McGILL [10], ROBIN [12] and SOBEL [13], cf. also SOBEL [14]. It was shown by LIPPMAN [9] that there is an integer M such that an average cost optimal policy has all c servers turned on or left on when M or more customers are present. We henceforth only consider the following finite class C of stationary policies with this property. A policy in C is characterized by integers $s(i)$, $S(i)$, $t(i)$ and $T(i)$ for $i = 0,1,\ldots$ such that

(a) $-1 \leq s(i) < S(i) \leq T(i) < t(i) \leq c + 1$ for all $i \geq 0$, where $s(i) = c - 1$, $S(i) = T(i) = c$ and $t(i) = c + 1$ for all $i \geq M$,

(b) $s(i) \leq s(i+1)$ and $t(i) \leq t(i+1)$ for all $i \geq 0$.

Under this policy the number of servers operating is adjusted both at arrival and service completion epochs. If there are i customers present and k servers turned on, the number of servers on is adjusted upward to $S(i)$ when $k \leq s(i)$, is kept unaltered when $s(i) < k < t(i)$ and is adjusted downward to $T(i)$ when $k \geq t(i)$.

It is a famous conjecture that there is an average cost optimal policy which belongs to the class C and has the additional property that $S(i) = s(i) + 1$ and $T(i) = t(i) - 1$ for all i.

In this paper a special policy iteration algorithm will be developed which locates an average cost optimal policy. This algorithm generates within the class C a sequence of improved policies, and in all examples tested the algorithm converged to an optimal policy with $S(i) = s(i) + 1$ and $T(i) = t(i) - 1$ for all i (we note that the algorithm may also be used for locating an average cost optimal policy within the class C for the case of general switch-over costs). The algorithm exploits the structure of the particular queueing problem. This appears especially in the value-determination part of the algorithm in which the size of the system of linear equations to be solved is of the order 2M, independent of c. In addition the algorithm does not require any truncation of the state space, i.e. no approximation of the infinite capacity problem to a finite one is needed. These facts compare favourably with the policy iteration algorithm of HOWARD [8] in which Nc linear equations must be solved in the value-determination part, where the integer N arises from the truncation of the state space and denotes the maximum number of customers allowed in the system. We may expect that $N \gg M$, especially when $\lambda/c\mu$ is close to 1 in which case a large choice of N is required in order to obtain a fair approximation of the infinite capacity problem whereas the estimate of M tends to be small since in this case an optimal policy tends to have all c servers on with relatively few customers in the system.

In section 4.2 we specify the basic elements 1-6 of [4] which are crucial for the algorithm and we determine some absorption probabilities which underly the transition probabilities of the embedded decision processes. In section 4.3 we derive the system of linear equations to be solved in the value-determination operation. Finally, in section 4.4 we present the algorithm and give some numerical results.

4.2. THE ELEMENTS

In choosing the state space, the natural process and the feasible decisions, similar considerations as in the first two applications will play a role. In order to obtain a set A_0 which has the desired properties and further allows for computationally tractable k-and t-functions, we will choose the elements 1-3 in such a way that in the natural process always c servers will be turned on when the number of customers is larger than M and, moreover, the states in which no customers are present are intervention states for any policy. The latter can always be achieved by choosing these states absorbing for the natural process, e.g. imagine that in the natural process the system is closed down forever when the system becomes empty.

This choice involves the introduction of interventions for these states and (fictitious) states to which the system is transferred by these interventions.

After these introductory remarks, we now choose as state space

$$X = \{(i,s) \mid i = 0,1,\ldots;\ s = 0,1,\ldots,c\} \cup \{(\bar{0},\bar{s}) \mid s = 0,1,\ldots,c\},$$

where state (i,s) with $i \geq 1$ corresponds to the situation where i customers are present and there are s servers turned on of which min(i,s) servers provide service. The state (0,s) corresponds to the situation where no customers are present, there are s servers turned on and the servers are not available for any future service, while state $(\bar{0},\bar{s})$ corresponds to the same situation except that the servers are now available for future service. We choose the natural process as follows. For both initial state (i,s) with $1 \leq i \leq M$ and initial state (i,s) with $i \neq 0$ and $s = c$ the natural process stays in state (i,s) until the next epoch at which an arrival or service completion occurs after which the natural process assumes either state (i+1,s) or (i-1,s) depending upon whether an arrival or service completion occurs first, so for these initial states the number of servers on is left unaltered in the natural process. For initial state (i,s) with $i > M$ and $s \neq c$ the natural process jumps immediately to state (i,c), i.e. for this initial state the number of servers is adjusted upward to c in the natural process. The states (0,m), $m = 0,\ldots,c$ are chosen as absorbing states for the natural process, whereas the natural process starting from state $(\bar{0},\bar{s})$ stays in this state until the next arrival epoch at which the natural process assumes state (1,s).

We next choose the sets of feasible decisions. For state (i,s) with $1 \leq i \leq M-1$ and $s \neq c$ the set of feasible decisions consists of the decisions $d = 0,1,\ldots,c$ where decision d prescribes to adjust the number of servers turned on from s to d and causes an instantaneous transition to state (i,d). Observe that for this state (i,s) the decision $d = s$ is the null-decision and any decision $d \neq s$ is an intervention. In state (M,s) with $s \neq c$ we choose as only possible decision the intervention $d = c$ which prescribes an upward adjustment of the number of servers to c and causes an instantaneous transition to state (M,c). In each of the states (0,s), $0 \leq s \leq c$ the set of feasible decisions consists of the interventions $d = 0,\ldots,c$ where the intervention d prescribes to "reactivate" the servers and to adjust the numbers of servers on from s to d and causes an instantaneous transition to state $(\bar{0},\bar{d})$. Finally, in the states (M,c) and $(\bar{0},\bar{s})$ for $0 \le s \le c$ we take the null-decision as the only possible decision. The cost structure is as follows. In the natural process a holding cost at rate h.i and an operating cost at rate w.s are incurred when there are i customers present and s servers turned on. There is incurred an intervention cost of K(s,d) when the intervention d is made in any state in which s servers are on.

Now, for any policy the superimposition of the natural process and the interventions prescribed by that policy agrees with the evolution of the system resulting from the specific control as executed by the decisionmaker. Using the fact that $\lambda/c\mu < 1$, it follows that element 4 of [4] applies with

$$A_0 = \{(0,s) \mid s = 0,\dots,c\} \cup \{(M,s) \mid s = 0,\dots,c-1\}.$$

To determine the k-and t-functions introduced in element 5 of [4], we choose $A_{01} = A_{02} = A_0$. From the definitions of these functions, it follows that, for any state (i,s) with $i \neq 0$ and intervention d,

$$t((i,s);d) = t_0((i,d)) - t_0((i,s)),\ k((i,s);d) = K(s,d) + \\ + k_0((i,d)) - k_0((i,s)).$$

Further, for any state (0,s) and $d = 0,\dots,c$,

$$t((0,s);d) = t_0((\bar{0},\bar{d})) - t_0((0,s)),\ k((0,s);d) = K(s,d) + \\ + k_0((\bar{0},\bar{d})) - k_0((0,s)).$$

We shall now calculate the functions t_0 and k_0 as far as needed. Fix s with $s \neq c$. Then

$$(4.1) \qquad t_0((i,s)) = \begin{cases} (\lambda+i\mu)^{-1}[1+i\mu t_0((i-1,s))+\lambda t_0((i+1,s))], & 1 \le i \le s, \\ (\lambda+s\mu)^{-1}[1+s\mu t_0((i-1,s))+\lambda t_0((i+1,s))], & s \le i \le M-1, \end{cases}$$

with $t_0((0,s)) = t_0((M,s)) = 0$. For ease of notation, denote by $h_0((i,s))$ the component of $k_0((i,s))$ in which the expected holding cost are represented, i.e.

$$k_0((i,s)) = swt_0((i,s)) + h_0((i,s)).$$

We have

$$(4.2)\qquad h_0((i,s)) = \begin{cases} (\lambda+i\mu)^{-1}[hi+i\mu h_0((i-1,s))+\lambda h_0((i+1,s))], & 1 \le i \le s, \\ (\lambda+s\mu)^{-1}[hi+s\mu h_0((i-1,s))+\lambda h_0((i+1,s))], & s \le i \le M-1, \end{cases}$$

with $h_0((0,s)) = h_0((M,s)) = 0$. We now discuss briefly the solution of (4.1). The solution of (4.2) proceeds in the same way. We refer to MILLER [12] for details. The equation for $t_0((i,s))$ is a second-order linear difference equation with non-constant coefficients for $i \le s$ and constant coefficients for $i \ge s$. The solution of the equation with constant coefficients is standard. To solve the equation with non-constant coefficients, multiply both sides of this equation by $\lambda + i\mu$ and consider the equation for $\Delta t_0(i) = t_0((i+1,s)) - t_0((i,s))$. This equation is a first-order linear difference equation and a particular solution may be found by using the method of parameter variation. We find for the case of $\lambda/s\mu \ne 1$,

$$t_0((i,s)) = \begin{cases} a_1(i) + \beta_1 b(i) + \alpha_1 & \text{for } 0 \le i \le s, \\ c_1(i) + \delta_1 d(i) + \gamma_1 & \text{for } s \le i \le M, \end{cases}$$

where

$$(4.3)\qquad a_1(i) = -\sum_{t=0}^{i-1} \sum_{j=0}^{t} \frac{t!(\mu/\lambda)^{t-j}}{\lambda j!},\quad b(i) = \sum_{j=0}^{i-1} (\mu/\lambda)^j j!$$

$$(4.4)\qquad c_1(i) = (i-M)/(s\mu-\lambda),\quad d(i) = (s\mu/\lambda)^i - (s\mu/\lambda)^M.$$

By the boundary conditions $t_0((0,s)) = t_0((M,s)) = 0$, we have $\alpha_1 = \gamma_1 = 0$. The constants β_1 and δ_1 follow by considering (4.1) for $i = s$ and substituting the above explicit expressions for $t_0((i,s))$ with $i = s-1$, s and $s+1$ where there are two possibilities for $t_0((s,s))$. To save space, we omit the formulas for these constants. For the same reason, we omit the expression for $t_0((i,s))$ when $\lambda/s\mu = 1$.

Similarly, we find for the case of $\lambda/s\mu \ne 1$,

$$h_0((i,s)) = \begin{cases} a_2(i) + \beta_2 b(i) & \text{for } 0 \le i \le s, \\ c_2(i) + \delta_2 d(i) & \text{for } s \le i \le M, \end{cases}$$

where

$$(4.5) \qquad a_2(i) = -h \sum_{t=0}^{i-1} \sum_{j=1}^{t} \frac{t!(\mu/\lambda)^{t-j}}{\lambda(j-1)!}, \quad c_2(i) = \frac{h(i^2-M^2)}{2(s\mu-\lambda)} + \frac{h(s\mu+\lambda)(i-M)}{2(s\mu-\lambda)^2}.$$

The constants β_2 and δ_2 follow by the same considerations as above.

Next we determine the functions $t_0((i,c))$ and $h_0((i,c))$ where h_0 is defined as above. Clearly,

$$(4.6) \qquad t_0((i,c)) = (\lambda+i\mu)^{-1}[1+i\mu t_0((i-1,c))+\lambda t_0((i+1,c))], \qquad 1 \le i \le c,$$

with $t_0((0,c)) = 0$. To give a recursive relation for $t_0((i,c))$ for $i \ge 1$, we make the following observation. Using the "memorylessness" property of the exponential distribution, it is easily seen that the time needed to reduce the number of customers from $i \ge c$ to $i-1$ by using c exponential servers having each mean service time $1/\mu$ is distributed as the length of one busy period in the M/M/1 queue with arrival rate λ and mean service time $1/c\mu$. This implies

$$(4.7) \qquad t_0((i,c)) = \frac{1}{c\mu-\lambda} + t_0((i-1,c)) \qquad \text{for } i \ge c.$$

Using $t_0((0,c)) = 0$, we get that the solution to (4.6) is given by

$$t_0((i,c)) = a_1(i) + \xi_1 b(i) \qquad \text{for } 0 \le i \le c,$$

where $a_1(i)$ and $b(i)$ are defined in (4.3) and the constant ξ_1 follows by using (4.7) with $i = c$. Next we find

$$(4.8) \qquad h_0((i,c)) = (\lambda+i\mu)^{-1}[hi+i\mu h_0((i-1,c))+\lambda h_0((i+1,c))], \; 1 \le i \le c,$$

with $h_0((0,c)) = 0$. Using the fact that for the above M/M/1 queue the total expected amount of time spent by the customers in the system during one busy period equals $c\mu/(c\mu-\lambda)^2$ (observe that the ratio of this quantity and the expected length of one busy cycle gives the average number of customers present), we find

$$(4.9)\qquad h_0((i,c)) = \frac{h(i-1)}{c\mu-\lambda} + \frac{hc\mu}{(c\mu-\lambda)^2} + h_0((i-1,c)) \qquad \text{for } i \geq c.$$

Using $h_0((0,c)) = 0$, we find that the solution to (4.8) is given by

$$h_0((i,c)) = a_2(i) + \xi_2 b(i)$$

where $a_2(i)$ and $b(i)$ are given in (4.5) and (4.3) and the constant ξ_2 follows by using (4.9) with $i = c$.

We end this section by determining some absorption probabilities which underly the one-step transition probabilities of the embedded decision processes. For any integers i,s,L and R with $0 \leq L \leq i \leq R \leq M$, $R \neq L$ and $0 \leq s \leq c$, define $p(i,s,L,R)$ as the probability that the natural process starting from state (i,s) will assume state (R,s) before state (L,s). Suppress for the moment the dependence of p on L,R and s and write $p(i,s,L,R) = p(i)$. Since in the natural process the number of servers on is not changed as long as not more than M customers are present, we find

$$(4.10)\qquad p(i) = \begin{cases} (\lambda+i\mu)^{-1}[i\mu p(i-1) + \lambda p(i+1)] & \text{for } i \leq s, \\ (\lambda+s\mu)^{-1}[s\mu p(i-1) + \lambda p(i+1)] & \text{for } i \geq s, \end{cases}$$

with $p(L) = 0$ and $p(R) = 1$. We give only the solution when $\lambda/s\mu \neq 1$ and we distinguish between three cases.

Case 1. $L \geq s$. Then we find the solution of the classical ruin problem,

$$p(i,s,L,R) = \{(s\mu/\lambda)^i - (s\mu/\lambda)^L\}/\{(s\mu/\lambda)^R - (s\mu/\lambda)^L\} \quad \text{for all } i.$$

Case 2. $R \leq s$. Then

$$p(i,s,L,R) = \{\sum_{j=L}^{i-1} (\mu/\lambda)^j j!\}/\{\sum_{j=L}^{R-1} (\mu/\lambda)^j j!\} \qquad \text{for all } i.$$

Case 3. $L < s < R$. Then

$$p(i,s,L,R) = \begin{cases} 1 + \eta_1\{(s\mu/\lambda)^i - (s\mu/\lambda)^R\} & \text{for } s \leq i \leq R \\ \eta_2 \sum_{j=L}^{i-1} (\mu/\lambda)^j j! & \text{for } L \leq i \leq s, \end{cases}$$

where the constants η_1 and η_2 follow by the same considerations as before.

4.3. THE SYSTEM OF EQUATIONS FOR A POLICY OF THE CLASS C.

Fix policy $z \in C$. In this section we shall specify for policy z the system of equations (8)-(9) introduced in [4]. We recall that policy z is characterized by integers s(i), S(i), t(i) and T(i) for i = 0,...,M-1 (see section 4.1) and we observe that its set of intervention states is given by $A_z = \{(i,s) \mid i \geq 1,\ s \leq s(i) \text{ or } s \geq t(i)\} \cup \{(0,s) \mid 0 \leq s \leq c\}$. By the structure of policy z we have that after any intervention the system assumes one of the states (i,S(i)), (i,T(i)) or $(\bar{0},\bar{s})$ where $1 \leq i \leq M$ and $s(0) < s < t(0)$. This fact will have as a consequence that in the value-determination procedure we need only to solve 2M + t(0) - s(0) - 2 linear equations. Before showing this, we note that, by the monotonicity properties of policy z, the set A_z will be entered in one of the states (L(s),s) and (R(s),s) with $0 \leq s \leq c$ where

(4.11) $L(s) = \max\{i \mid 1 \leq i \leq M,\ t(i) \leq s\}$ if $s \geq t(0)$, and $L(s) = 0$, otherwise,
$R(s) = \min\{i \mid 1 \leq i \leq M,\ s(i) \geq s\}$ if $s < c$, and $R(c) = \infty$.

That is, for s servers turned on, L(s) denotes that largest queue size for which policy z prescribes either a reduction of the number of servers on or at least their "reactivation", whereas R(s) denotes the smallest queue size for which policy z prescribes an upward adjustment of the number of servers turned on.

We now specify the equations for the average cost g and the relative values v((i,s)) with $(i,s) \in A_z$. By relation (11) in [4], we have for $1 \leq i \leq M$

(4.12)
$$v((i,s)) = k((i,s);S(i)) - gt((i,s);S(i)) + v((i,S(i)),\quad s \leq s(i),$$
$$v((i,s)) = k((i,s);T(i)) - gt((i,s);T(i)) + v((i,T(i)),\quad s \geq t(i),$$

whereas for the intervention states (0,s), $0 \leq s \leq c$, we find

(4.13)
$$v((0,s)) = k((0,s);S(0)) - gt((0,s);S(0)) + v((\bar{0},\overline{S(0)})),\quad s \leq s(0),$$
$$v((0,s)) = k((0,s);T(0)) - gt((0,s);T(0)) + v((\bar{0},\overline{T(0)})),\quad s \geq t(0),$$
$$v((0,s)) = k((0,s);s) - gt((0,s);s) + v((\bar{0},\bar{s})),\quad \text{otherwise.}$$

Letting $p(i,s,L(s),R(s))$ for $s < c$ be defined as in section 4.2 and letting $p(i,c,L(c),R(c)) = 0$, it follows from relation (9) in [4] that, for state $(i,s) \notin A_z$,

$$(4.14) \qquad v((i,s)) = p(i,s,L(s),R(s))v((R(s),s)) + \\ + \{1 - p(i,s,L(s),R(s))\}\ v((L(s),s)).$$

Further, using the fact that $L(s) = 0$ for $s(0) < s < t(0)$, we find

$$(4.15) \qquad v((\bar{0},\bar{s})) = p(1,s,0,R(s))v((R(s),s)) + \\ + \{1 - p(1,s,0,R(s))\}\ v((0,s)) \text{ for } s(0) < s < t(0).$$

The equations for the remaining relative values will not be needed and are omitted.

It now follows that we get $2M + t(0) - s(0) - 3$ linear equations in the $2M + t(0) - s(0) - 2$ unknowns g, $v((i,S(i))$, $v(i,T(i))$ and $v((\bar{0},\bar{s}))$ with $1 \le i \le M-1$ and $s(0) < s < t(0)$ by taking the equations (4.15) and the equations (4.14) with both $s = S(i)$ and $s = T(i)$ and by substituting in the right-hand sides of these equations the corresponding equations for $v((R(s),s)$ and $v((L(s),s))$, cf. (4.12)-(4.13). To determine these unknowns uniquely, we put one of the relative values equal to zero (see Theorem 2 in [4]), e.g. put $v((M-1), T(M-1)) = 0$. Once the above $2M + t(0) - s(0) - 2$ linear equations have been solved, we can next compute any of the required $v(x)$ from (4.12)-(4.14).

4.4. THE ALGORITHM

We shall now present a policy-iteration algorithm which generates a sequence of policies belonging to the class C of structured policies. Before specifying the details of this algorithm, we first give a general outline of the algorithm which is based on the modified policy iteration method given in section 5 of [4].

Algorithm

(a) *Value-determination procedure.* Solve for the current policy $z \in C$ with parameters $s(i)$, $S(i)$, $t(i)$ and $T(i)$ the above described system of $2M + t(0) - s(0) - 2$ linear equations.

(b) *Policy-improvement procedure.* Determine a policy $z' \in C$ with parameters $s'(i)$, $S'(i)$, $t'(i)$ and $T'(i)$ where $s'(i) \geq s(i)$ and $t'(i) \leq t(i)$.

(c) *Cutting-procedure.* Determine a policy $z'' \in C$ with parameters $s''(i)$, $S''(i)$, $t''(i)$ and $T''(i)$ where $S''(i) = S'(i)$, $T''(i) = T'(i)$, $s''(i) \leq s'(i)$ and $t''(i) \geq t'(i)$.

(d) If $z'' = z$, stop, otherwise, go to (a).

We now give in detail the policy-improvement and the cutting procedure.

Policy-improvement procedure

Suppose that we have solved for policy z the system of $2M + t(0) - s(0) - 2$ linear equations as described in section 4.3. For the obtained solution, denote by $g(z)$ the average cost of policy z and denote by $v(z;x)$ the relative value for state x (as already noted, once we have solved the embedded system of equations described in section 4.3 any required $v(z;x)$ follows immediately from one of the relations (4.12)-(4.14)). Since we want to obtain a policy $z' \in C$, we have to apply the policy-improvement procedure of the modified policy-iteration algorithm given in section 5 of [4]. Before doing this, we note that for any state (i,s) with $0 \leq i \leq M-1$ and any decision $d \in D((i,s))$ (cf. definition (13) in [4] and section 4.2),

$$v(d.z;\ (i,s)) = K(s,d) + \psi_i(d) - k_0((i,s)) + g(z)t_0((i,s))$$

where

$$\psi_i(d) = \begin{cases} k_0((i,d)) - g(z)t_0((i,d)) + v(z;(i,d)) & \text{for } i \geq 1, \\ k_0((\bar{0},\bar{d})) - g(z)t_0((\bar{0},\bar{d})) + v(z;(\bar{0},\bar{d})) & \text{for } i = 0. \end{cases}$$

Further, we recall that in the policy-improvement procedure any intervention prescribed by policy z cannot be replaced by the null-decision but only by another intervention. Since in the states $(0,s)$, $0 \leq s \leq c$ the null-decision is not feasible as opposed to the states (i,s) with $i \geq 1$, the two cases have to be considered in a slightly different way.

Fix first $1 \leq i \leq M-1$. Define d_i^* and d_i^{**} as the smallest and the largest integer for which $K(0,d) + \psi_i(d)$ and $K(c,d) + \psi_i(d)$ are minimal on the interval $[s(i) + 1,\ t(i) - 1]$. Observe that d_i^* and d_i^{**} minimize $v(d.z;(i,0))$ and $v(d.z;(i,c))$ for $s(i) < d < t(i)$. It is straightforward to verify that

$d_i^* \le d_i^{**}$. By the same reasoning as on p.258 in SOBEL [13], we find that, for all $0 \le s \le d_i^*$, the number d_i^* minimizes $K(s,d) + \psi_i(d)$ and hence $v(d.z;(i,s))$ for $s(i) < d < t(i)$. Hence, for all $0 \le s \le d_i^*$,

$$(4.16) \qquad v(d_i^*.z;(i,s)) = \min_{s(i)<d<t(i)} v(d.z;(i,s)) \le v(z;(i,s)),$$

where the latter inequality follows from the fact that $v(d.z;x) = v(z;x)$ for $d = z(x)$. Similarly, we have for all $d_i^{**} \le s \le c$,

$$(4.17) \qquad v(d_i^{**}.z;(i,s)) = \min_{s(i)<d<t(i)} v(d.z;(i,s)) \le v(z;(i,s)).$$

For $i = 0$ we determine the numbers d_0^* and d_0^{**} in the same way as above except that we now take $[0,c]$ as the minimization interval instead of $[s(i) + 1, t(i) - 1]$. Similar properties hold for d_0^* and d_0^{**} as for d_i^* and d_{ii}^{**}.
It now follows that we obtain policy $z' \in C$ by taking $s'(i) = d_i^* - 1$, $S''(i) = d_i^*$, $t'(i) = d_i^{**} + 1$ and $T'(i) = d_i^{**}$ for $0 \le i \le M-1$.

The cutting procedure

Suppose we have performed part (b) of the algorithm and obtained policy z'. In addition we have obtained the function $v(z'(x).z;x)$ for $x \in A_{z'}$. For ease of notation, we write $\hat{v}(x) = v(z'(x).z;x)$ for $x \in A_{z'}$.

For the natural process with a cost of $\hat{v}(y)$ for stopping at state $y \in A_{z'}$ we shall now determine a set A with $A_0 \subseteq A \subseteq A_{z'}$ such that (a) the set A is as stopping set at least as good as the set $A_{z'}$ for each initial state $x \in A_{z'}$ (in fact this is trivially met for $x \in A$, so that verification is only needed for $x \in A_{z'} \backslash A$), (b) $A = A_{z''}$ for some $z'' \in C$. This will be done according to the principle outlined in remark s of [4]. For a properly chosen sequence of states $x \in A_{z'}$ with $x \notin A_0$, we shall verify whether $A_{z'} \backslash \{x\}$ is a better stopping set than $A_{z'}$ or not for the natural process starting from state x. Next the intersection of all those sets which are better stopping sets will give the desired set A. Before we demonstrate how this principle can be developed into a simple procedure in our queueing problem, we first evaluate for $x = (i,s) \in A_{z'}$ the quantity $Q'_{is} = E\hat{v}(S_x)$, where S_x is the first entrance state of the natural process into the set $A_{z'} \backslash \{x\}$ when the initial state is x, cf. definition (18) in [4]. Consider first the case where $x = (i,s)$ with $s \le s'(i)$. Then the possible realizations of S_x are the states $(i+1,s)$ and $(i-1,s)$ if $s \le s'(i-1)$ and the states $(i+1,s)$ and $(L'(s),s)$ if $s > s'(i-1)$ where $L'(s)$ is defined by (4.11) with z replaced by z'. Using the definition of the absorption probability p given in section 4.2, we find for state (i,s) with $s \le s'(i)$,

$$Q'_{is} = \begin{cases} [\lambda+\mu\min(i,s)]^{-1}[\lambda\tilde{v}((i+1,s)+\mu\min(i,s)\tilde{v}((i-1,s))], & s \le s'(i-1), \\ p(i,s,L'(s),i+1)\tilde{\tilde{v}}((i+1,s))+\{1-p(i,s,L'(s),i+1)\}\tilde{v}((L'(s),s)), & s > s'(i-1). \end{cases}$$

Similarly, for state (i,s) with $s \ge t'(i)$ we find

$$Q'_{is} = \begin{cases} [\lambda+\mu\min(i,s)]^{-1}[\lambda\tilde{v}((i+1,s))+\mu\min(i,s)\tilde{v}((i+1,s))], & s \ge t'(i+1), \\ p(i,s,i-1,R'(s))\tilde{v}((R'(s),s))+\{1-p(i,s,i-1,R'(s))\}\tilde{v}((i-1,s)), & s < t'(i+1), \end{cases}$$

where $R'(s)$ is defined by (4.11) with z replaced by z' and $p(\cdot,c,\cdot,\cdot) = 0$.

We can now describe the determination of the parameters $s''(i)$, $S''(i)$, $t''(i)$ and $T''(i)$ of policy $z'' \in C$. Recall that in the cutting procedure any intervention prescribed by policy z' cannot be replaced by a different intervention but only by the null-decision. Consequently the states (0,s) for $0 \le s \le c$ need not to be considered in this procedure. Further, we have $S''(i) = S'(i)$, $T''(i) = T'(i)$, $s''(i) \le s'(i)$ and $t''(i) \ge t'(i)$ for all i with $s''(0) = s'(0)$ and $t''(0) = t'(0)$. We determine the numbers $s''(i)$ for $i \ge 1$ by calculating successively $s''(1),\dots,s''(M-1)$ in the following way. For $i = 1,\dots,M-1$, let $s''(i)$ be the largest value of s with $\max(0,s''(i-1)) \le s \le s'(i)$ such that $Q'_{is} \ge \tilde{v}((i,s))$ if such a value of s exists, otherwise let $s''(i) = s''(i-1)$. The numbers $t''(i)$ for $i \ge 1$ are determined by calculating successively $t''(M-1),\dots,t''(1)$. Let $t''(M) = c+1$. For $i = M-1,\dots,1$, let $t''(i)$ be the smallest value of s with $t'(i) \le s \le \min(c,t''(i+1))$ such that $Q'_{is} \ge \tilde{v}((i,s))$ if such a value of s exists, otherwise let $t''(i) = t''(i+1)$. In this way we obtain a policy $z'' \in C$.

REMARK 1. In any iteration step the above policy-improvement procedure yields a policy $z' \in C$ having the additional property that $S'(i) = s'(i) + 1$ and $T'(i) = t'(i) - 1$ for all i. However, except for the final iteration step, the cutting procedure by its very design may generate policies in C without this property.

REMARK 2. The above algorithm needs only a minor modification in order to locate an optimal policy among the class C of policies in case of general switch-over costs with the separability property $K(a,b) = k^+(b) + b^+(a)$ for $b > a$, $K(a,b) = k^-(b) + b^-(a)$ for $b < a$ and $K(a,b) = 0$ for $b = a$, where $k^+(\cdot)$, $k^-(\cdot)$, $b^+(\cdot)$ and $b^-(\cdot)$ are non-negative, $k^+(\cdot)$ is non-decreasing and k^- is non-increasing. Observe that this function $K(a,b)$ includes the case where the switch-over costs consist of a fixed adjustment cost plus linear costs as above. In order to apply the algorithm, only the policy-improvement part needs a slight modification. For all $i \geq 0$ we determine the numbers d_i^* and d_i^{**} as before. We again find $d_i^* \leq d_i^{**}$ for all i. However, we now find for $i \geq 1$ that the relations (4.16) and (4.17) only hold for $0 \leq s \leq s(i)$ and $t(i) \leq s \leq c$, respectively. The parameters of the new policy z' are now obtained as follows. We choose $S'(i) = d_i^*$ and $T'(i) = d_i^{**}$ for all $i \geq 0$ as before. The numbers $s'(i)$ are determined by calculating successively $s'(M-1),\ldots,s'(0)$. For $i = M-1,\ldots,0$, let $s'(i) + 1$ be the smallest value of s with

$$\begin{cases} s(i) + 1 \leq s \leq \min(S'(i) - 1,\ s(i+1)) & \text{if } i \geq 1 \\ 0 \leq s \leq \min(S'(0) - 1,\ s(1)) & \text{if } i = 0 \end{cases}$$

such that $v(S'(i).z;(i,s)) \geq v(z;(i,s))$ if such a value of sexists, otherwise let $s'(i) = \min(S'(i) \dot{-} 1,\ s(i+1))$. The numbers $t'(i)$ are determined by calculating successively $t'(0),\ldots,t'(M-1)$. Let $t'(-1) = 0$. For $i = 0,\ldots,M-1$, let $t'(i) - 1$ be the largest value of s with

$$\begin{cases} \max(t'(-1),\ T'(0) + 1) \leq s \leq c & \text{if } i = 0 \\ \max(t'(i-1),\ T'(i) + 1) \leq s \leq t(i) - 1 & \text{if } i \geq 1 \end{cases}$$

such that $v(T'(i).z;(i,s)) \geq v(z;(i,s))$ if such a value of s exists, otherwise let $t'(i) = \max(t'(i-1),\ T'(i) + 1)$.

TABLE 3. C = 15, λ = 14.25, μ = 1, h = 10, $k^+ = k^-$ = 100.

w = 100			w = 250			w = 400			w = 1000		
i	s(i)	t(i)	i	s(i)	t(i)	i	s(i)	t(i)	i	s(i)	t(i)
0	-1	13	0	-1	8	0	-1	6	0	-1	3
1	0	13	1	0	9	1	-1	7	1	-1	4
2	1	13	2	0	9	2	0	7	2	0	5
3	2	13	3	1	10	3	1	8	3	1	5
4	2	14	4	2	10	4	2	8	4	1	6
5	3	14	5	3	10	5	2	9	5	2	7
6	4	14	6	3	11	6	3	10	6	3	8
7	5	14	7	4	11	7	4	10	7	3	9
8	6	14	8	5	12	8	5	11	8	4	10
9	6	15	9	6	12	9	5	12	9	5	10
10	7	15	10	6	13	10	6	12	10	5	11
11	8	15	11	7	14	11	6	13	11	6	12
12	9	16	12	7	14	12	7	14	12	6	13
13	9	16	13	8	15	13	8	14	13	7	14
14	10	16	14	9	15	14	8	15	14	8	15
15	10	16	15	9	16	15	9	16	15	8	16
16	11	16	16	10	16	16	9	16	16	9	16
17	11	16	17	10	16	17	10	16	17	9	16
18	12	16	18	10	16	18	10	16	18	10	16
19	12	16	19	11	16	19	11	16	19	10	16
20	12	16	20	11	16	20	11	16	20	11	16
21	13	16	21	12	16	21	11	16	21	11	16
22	13	16	22	12	16	22	12	16	22	12	16
23	13	16	23	13	16	23	12	16	23	12	16
24	14	16	24	13	16	24	13	16	24	12	16
			25	13	16	25	13	16	25	13	16
			26	14	16	26	14	16	26	13	16
									27	14	16
$g(z^*)$ = 1782.46			$g(z^*)$ = 3944.60			$g(z^*)$ = 6088.63			$g(z^*)$ = 14644.29		

We were not able to show that the algorithm converges in a finite number of iteration steps to an optimal policy, although any step yields an improved policy. However, convergence appeared in all examples tested. After convergence of the algorithm to a policy z^* (say) we checked a criterion guaranteeing that policy z^* is optimal among the class of all stationary policies when this criterion is satisfied. This criterion is based on Theorem 8 in [4] and requires the verification that (a) $v(d.z^*;(i,s)) \geq$ $\geq v(z^*;(i,s))$ for all (i,s) and all $d \in D((i,s))$, and (b) $Q^*_{is} \geq v(z^*;(i,s))$ for all $(i,s) \in A_{z^*}$ with $1 \leq i \leq M-1$ where Q^*_{is} is defined as Q'_{is} above with z' replaced by z^*.

In all examples tested this criterion was satisfied and, consequently, an optimal policy was found.

In table 3 we give for a number of numerical examples the minimal average cost $g(z^*)$ and optimal values for $s(i)$, $S(i)$, $t(i)$ and $T(i)$ where $S(i)$ and $T(i)$ are given by $S(i) = s(i) + 1$ and $T(i) = t(i) - 1$.

ACKNOWLEDGEMENT

We are indebted to Mr. Rob van der Horst for most valuable help in computer programming.

REFERENCES

[1] BELL, C.E., *Turning off a server with customers present: is this any way to run an M/M/c queue with removable servers?*, Operations Res., Vol.23 (1975), 571-574.

[2] COHEN, J.W., *On the optimal switching level for an M/G/1 queueing system*, Report of the Mathematical Institute, University of Utrecht, The Netherlands, 1975.

[3] COHEN, J.W., *On regenerative processes in queueing theory*, Lecture notes in Economics and Mathematical Systems, Vol.121, Springer-Verlag, Berlin, 1976.

[4] DE LEVE, G., FEDERGRUEN, A. & H.C.TIJMS, *A general Markov decision method, I: model and techniques*, 1976 (submitted for publication).

[5] DOSHI, B.T., *Continuous control of Markov processes on an arbitrary state space*, *MRC Technical Summary Report No. 1468*, Mathematics Research Center, University of Wisconsin, 1974.

[6] FELLER, W., *An introduction to probability theory and its applications*, Vol.2, Wiley, New York, 1966.

[7] FLETCHER, R., *A new approach to variable metric algorithms*, *Comp. J.*, Vol.13 (1970), 317-322.

[8] HOWARD, R.A., *Dynamic programming and Markov processes*, Cambridge, Massachusetts, 1960.

[9] LIPPMAN, S.A., *A new technique in the optimization of exponential queueing systems*, Working Paper No.211, Western Management Science Institute, University of California, Los Angeles, 1960.

[10] McGILL, J.T., *Optimal control of a queueing system with variable number of servers*, Technical Report No.2, Department of Operations Research, Stanford University, 1969.

[11] MILLER, K.S., *Linear difference equations*, W.A. Benjamin Inc., New York, 1968.

[12] ROBIN, M., *Contrôle optimal de files d'attentes*, Rapport de Recherche No.117, IRIA, Rocquencourt, France, 1975.

[13] SOBEL, M.J., *Optimal operation of queues, in: A.B.Clarke (ed), Mathematical methods in queueing theory*, *Lecture Notes in Economics and Mathematical Systems*, Vol.98, Springer-Verlag, Berlin, 1974.

[14] SOBEL, M.J., *Production smoothing with stochastic demand II: infinite-horizon case*, *Management Sci.*, Vol.17 (1971), 724-735.

[15] THATCHER, R.M., *Optimal singel-channel service policies for stochastic arrivals*, Report ORC 68-16, Operations Research Center, University of California, Berkeley, 1968.

[16] TIJMS, H.C., *On a switch-over policy for controlling the workload in a queueing system with two constant service rates and fixed switch-over costs*, 1975 (to appear in Zeitschrift für Operations Res.).

[17] VAN DOMSELAAR, V. & P.W. HEMKER, *Nonlinear parameter estimation in initial value problems*, Mathematical Centre Report NW 18/75, Mathematisch Centrum, Amsterdam, 1975.