J. Appl. Prob. **23**, 708–717 (1986)
Printed in Israel

OPTIMAL SEARCH FOR A RANDOMLY MOVING OBJECT

R. R. WEBER,* *University of Cambridge*

Abstract

It is desired to minimize the expected cost of finding an object which moves back and forth between two locations according to an unobservable Markov process. When the object is in location i ($i = 1, 2$) it resides there for a time which is exponentially distributed with parameter λ_i and then moves to the other location. The location of the object is not known and at each instant until it is found exactly one of the two locations must be searched. Searching location i for time δ costs $c_i\delta$ and conditional on the object being in location i there is a probability $\alpha_i\delta + o(\delta)$ that this search will find it. The probability that the object starts in location 1 is known to be $p_1(0)$. The location to be searched at time t is to be chosen on the basis of the value of $p_1(t)$, the probability that the object is in location 1, given that it has not yet been discovered. We prove that there exists a threshold Π such that the optimal policy may be described as: *search location 1 if and only if the probability that the object is in location 1 is greater than* Π. Expressions for the threshold Π are given in terms of the parameters of the model.

MARKOV PROCESSES; OPTIMAL CONTROL; SEARCH PROBLEMS

1. Introduction

It is desired to minimize the expected cost of finding an object which moves back and forth between two locations according to an unobservable Markov process. When the object is in location i ($i = 1, 2$) it resides there for a time which is exponentially distributed with parameter λ_i and then moves to the other location. The location of the object is not known, and at each instant until it is found exactly one of the two locations must be searched. Searching location i for time δ costs $c_i\delta$ and conditional on the object being in location i there is a probability $\alpha_i\delta + o(\delta)$ that this search will find it. The probability that the object

---

Received 22 August 1984; revision received 30 July 1985.
* Postal address: Engineering Department, Control and Management Systems Group, Mill Lane, Cambridge, CB2 1RX, UK.
This research was conducted at the Department of Electrical Engineering and Computer Sciences and Electronics Research Laboratory, University of California, Berkeley and was supported by the Office of Naval Research Contract N00014-80-C-0507.

starts in location 1 is known to be $p_1(0)$. The location to be searched at time t is to be chosen on the basis of the value of $p_1(t)$, the probability that the object is in location 1, given that it has not yet been discovered.

The above is a continuous-time version of a problem posed by Ross (1983), Section 5.3, in discrete time. In the discrete-time formulation the location of the object follows a two-state Markov chain. Ross (1983) makes the intuitively reasonable conjecture that there exists a threshold Π such that the optimal policy can be described as: *search location* 1 *if and only if* $p_1(t)$ *is greater than* Π. He considers the problem of minimizing the expected time to find the object ($c_1 = c_2 = 1$) and observes that it is not necessarily optimal to search the location where the immediate probability of finding the object is greatest. It may be advantageous to search the location where the object is less likely to be found in the short term, since upon not finding the object in this location one may obtain a better idea of where the object is likely to be thereafter. In other words, it is not necessarily optimal to search location 1 if and only if $\alpha_1 p_1(t)$ exceeds $\alpha_2 p_2(t)$.

Ross's conjecture for the discrete-time formulation of the problem appears remarkably difficult to prove, but it is the result of this paper that the conjecture is true in the continuous-time formulation. The truth of the conjecture and expressions for the value of the threshold are stated in the following theorem. We assume that c_1, c_2, α_1, α_2, λ_1 and λ_2 are all positive.

Theorem. An optimal policy can be described as: *search location* 1 *if and only if the probability that the object is in location* 1 *exceeds* Π. Define

$$\theta_1 = [(\lambda_1 + \lambda_2 + \alpha_1)\alpha_2 c_1 - \lambda_2\alpha_1(c_2 - c_1)]/[(\lambda_1 + \lambda_2)\alpha_1 c_1 + \lambda_2\alpha_1(c_2 - c_1)],$$

$$\theta_2 = [(\lambda_1 + \lambda_2)\alpha_2 c_2 - \lambda_1\alpha_2(c_2 - c_1)]/[(\lambda_1 + \lambda_2 + \alpha_2)\alpha_1 c_2 + \lambda_2\alpha_2(c_2 - c_1)],$$

$$\delta_1 = [-(\lambda_1 + \alpha_1 - \lambda_2) + \{(\lambda_1 + \alpha_1 - \lambda_2)^2 + 4\lambda_1\lambda_2\}^{1/2}]/2\lambda_1,$$

and

$$\delta_2 = [-(\lambda_1 - \alpha_2 - \lambda_2) + \{(\lambda_1 - \alpha_2 - \lambda_2)^2 + 4\lambda_1\lambda_2\}^{1/2}]/2\lambda_1.$$

The threshold Π can be written as $\Pi = \theta/(1 + \theta)$, where there are three distinct possibilities:

(a) If $\theta_1 \leqq \delta_1$ then $\theta = \theta_1$.

(b) If $\theta_2 \geqq \delta_2$ then $\theta = \theta_2$.

(c) Otherwise θ is the unique positive root of the cubic equation

$$\begin{aligned} 0 = {} & [\alpha_1\alpha_2\lambda_1 c_2 + (c_2 - c_1)(\alpha_1 + \alpha_2)\lambda_1^2]\theta^3 \\ & + [\alpha_1^2 c_2(\lambda_1 + \lambda_2 + \alpha_2) + \alpha_1\alpha_2\lambda_1(c_1 + 2c_2) + (c_2 - c_1)(\alpha_1 + \alpha_2)\lambda_1(\lambda_1 - 2\lambda_2)]\theta^2 \\ & - [\alpha_2^2 c_1(\lambda_1 + \lambda_2 + \alpha_1) + \alpha_1\alpha_2\lambda_2(c_2 + 2c_1) + (c_2 - c_1)(\alpha_1 + \alpha_2)\lambda_2(\lambda_2 - 2\lambda_1)]\theta \\ & - \alpha_1\alpha_2\lambda_2 c_1 + (c_2 - c_1)(\alpha_1 + \alpha_2)\lambda_2^2. \end{aligned}$$

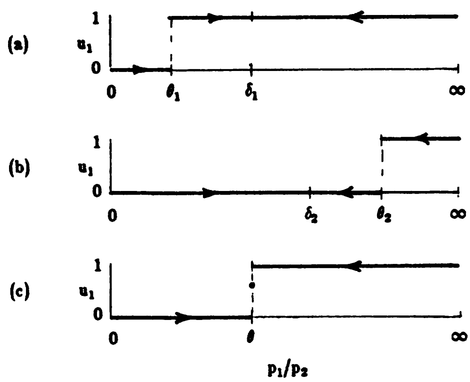

Figure 1. Three types of optimal trajectory

It is helpful to see the trajectories of $p_1(t)/p_2(t)$ for these three cases. Let $u_i(t)$ denote the effort applied to searching location i at time t, given that the object has not yet been found. The vector $u(t)$ is constrained to lie in the set $\Omega = \{v : v \geqq 0, v_1 + v_2 = 1\}$. Figure 1 illustrates the three possible types of optimal trajectory for $p_1(t)/p_2(t)$ and $u_1(t)$. The arrows indicate the direction of change of $p_1(t)/p_2(t)$ with increasing t.

The proof of the theorem proceeds as follows. In Section 2 we formulate a homogeneous variables description of the problem and show that it reduces to a deterministic optimal control problem. We state some standard results of optimal control theory as they apply to this problem. In Section 3 we use these results to demonstrate that the solution of the problem takes the form described by the theorem.

2. Optimal control formulation

As above, we shall let $u_1(t)$ and $u_2(t)$ denote the effort applied to searching locations 1 and 2 respectively at time t. The set of admissible policies is denoted by U and consists of all piecewise continuous functions $u(\cdot): [0,\infty) \to \Omega$. It is tempting to take p_1 as a state variable for the problem. However, this leads to non-linear dynamics of the form $\dot{p}_1 = (\alpha_2 u_2 - \alpha_1 u_1)p_1 p_2 + (\lambda_1 p_1 - \lambda_2 p_2)$. Instead, we formulate a problem in two variables, $x_1(t)$ and $x_2(t)$, where $x_i(t)$ is the probability that at time t the object is in location i and has not yet been discovered. Then $x_1(t) + x_2(t)$ is the probability that at time t the object has not yet been discovered. The problem can be posed as one of determining the control $u(\cdot)$ which achieves the infimum in

$$V(x) = \inf_{u \in U} V(x, u),$$

where

$$(1)\qquad V(x,u)=\int_0^\infty \{c_1u_1(t)+c_2u_2(t)\}\{x_1(t)+x_2(t)\}dt,\qquad x(0)=x,$$

$$(2)\qquad \dot{x}_1(t)=\lambda_2x_2(t)-\lambda_1x_1(t)-\alpha_1u_1(t)x_1(t),$$

$$(3)\qquad \dot{x}_2(t)=\lambda_1x_1(t)-\lambda_2x_2(t)-\alpha_2u_2(t)x_2(t).$$

The state dynamics given in (2) and (3) are derived straightforwardly. We shall write these as $\dot{x}(t)=a(x(t),u(t))$ and let $x(t)=\phi(t,x(0),u)$, $(0\leqq t<\infty)$, denote the trajectory of the state when it starts at $x(0)$ and the control $u(\cdot)$ is employed. Of course we are particularly interested in the solution to the problem for a starting state $x(0)$, such that $x_1(0)+x_2(0)=1$, but we shall solve the problem for any starting state $x(0)\geqq 0$.

Note that in the statement of the problem, $V(x,u)$ denotes the cost resulting from application of a fixed control policy u. We can therefore write

$$V(x,u)=x_1V((1,0),u)+x_2V((0,1),u).$$

This may be proved formally from (1)–(3), or one can note that it is immediate from the nature of the problem that the cost must be $V((1,0),u)$ times the probability that the object starts in location 1, plus $V((0,1),u)$ times the probability the object starts in location 2. Therefore

$$V(x)=\inf_{u\in U}\{x_1V((1,0),u)+x_2V((0,1),u)\}.$$

The following lemma is a consequence of V being the infimum of linear functions of x.

Lemma 1. $V(x)$ is concave in each of the variables x_1 and x_2.

We now appeal to standard results of optimal control theory (as found in Varaiya (1972), Chapters 7 and 8) and state conditions which are necessary and sufficient for a policy to be optimal for the problem.

Lemma 2.

(a) Suppose that for a starting state $x(0)$ the policy $u(\cdot)\in U$ is optimal and the optimal trajectory is $x(t)=\phi(t,x(0),u)$. Then there exists a solution to the adjoint equations

$$(4)\qquad \dot{\eta}_1(t)=\alpha_1u_1(t)\eta_1(t)+\lambda_1\eta_1(t)-\lambda_1\eta_2(t)-c_1u_1(t)-c_2u_2(t)$$

$$(5)\qquad \dot{\eta}_2(t)=\alpha_2u_2(t)\eta_2(t)+\lambda_2\eta_2(t)-\lambda_2\eta_1(t)-c_1u_1(t)-c_2u_2(t),$$

such that $u(\cdot)$ satisfies the maximum principle:

$$(6)\qquad H(x(t),u(t),\eta(t))=\min_{v\in\Omega}H(x(t),v,\eta(t)).$$

$H(x, v, \eta)$ is the Hamiltonian given by

$$H(x, v, \eta) = (c_1 v_1 + c_2 v_2)(x_1 + x_2) + \eta_1(\lambda_2 x_2 - \lambda_1 x_1 - \alpha_1 v_1 x_1) + \eta_2(\lambda_1 x_1 - \lambda_2 x_2 - \alpha_2 v_2 x_2).$$

(b) Conversely, suppose that for a starting state $x(0)$ there exists policy $u(\cdot)$, trajectory $x(t) = \phi(t, x(0), u)$, and function $\eta(\cdot)$ satisfying (4), (5) and (6). Then $u(\cdot)$ is an optimal policy.

Part (a) is Pontryagin's maximum principle (here applied as a minimum principle). It presents a condition which must be satisfied if a policy $u(\cdot)$ is optimal. In general, satisfaction of this condition guarantees only local optimality. However, in the problem considered here, the cost is the integral of a linear function of x, namely $(c_1 u_1 + c_2 u_2)(x_1 + x_2)$, and the dynamics, $\dot{x}_1 = a_1(x, u)$ and $\dot{x}_2 = a_2(x, u)$, are also linear in x. In this special case the existence of a solution to (4), (5) and (6) is sufficient to guarantee that a policy is globally optimal. This is the claim made in part (b).

Let the difference between the terms which multiply v_1 and v_2 in the definition of $H(x, v, \eta)$ be denoted by

$$\Delta(x, \eta) = (c_1 - c_2)(x_1 + x_2) + (\alpha_2 x_2 \eta_2 - \alpha_1 x_1 \eta_1).$$

It follows from the maximum principle (6) that (as a function of x and η) the optimal control must be

$$u(x, \eta) = \begin{cases} (1,0) \\ (0,1) \end{cases} \quad \text{as } \Delta(x, \eta) \lesseqgtr 0.$$

If $\Delta = 0$ then the maximum principle is insufficient to determine the optimal control.

3. Analysis of the optimal policy

We begin by investigating the relationship between $x(t)$ and the adjoint variables $\eta_1(t)$ and $\eta_2(t)$. It is a fact which is established in the proof of Pontryagin's maximum principle that $\eta_1(t)x_1 + \eta_2(t)x_2$ is a tangent hyperplane to $V(x)$ at the point $x = x(t)$. Since time does not enter the objective function or dynamics explicitly, $\eta(\cdot)$ may be viewed as a function of state rather than of time. With a slight abuse of notation we can write $\eta(x)$ to mean $\eta(t)$, evaluated at $t = 0$ when the starting state is $x(0) = x$. Moreover, since $V(x)$ and $H(x, v, \eta)$ are homogeneous in x, the value of $\eta(x)$ depends only on the ratio of x_1 to x_2.

Remark. Much of the following discussion will be in terms of the ratio x_1/x_2. We suppose, without loss of generality, that the starting state is always such that $x_2(0) \neq 0$. This ensures that the denominator of $x_1(t)/x_2(t)$ is never 0.

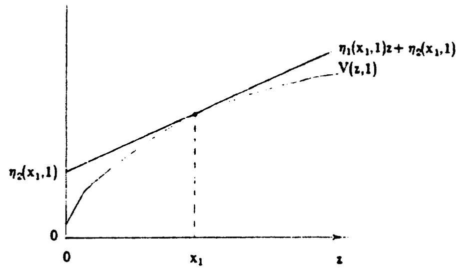

Figure 2. Geometric interpretation of $\eta_2(x_1, 1)$

We fix $x_2 = 1$ and observe by the discussion of the previous paragraph that, as a function of z, the line $\eta_1(x_1, 1)z + \eta_2(x_1, 1)$ is tangent to the function $V(z, 1)$ at the point $z = x_1$. We can think of $\eta_2(x_1, 1)$ as being the intercept of the line $\eta_1(x_1, 1)z + \eta_2(x_1, 1)$ at $z = 0$. This is depicted in Figure 2.

We know by Lemma 1 that $V(z, 1)$ is concave in z. So it follows from consideration of Figure 2 that $\eta_2(x_1, 1)$ is non-decreasing in x_1. Equivalently, $\eta_2(x_1, x_2)$ is non-decreasing in the value of x_1/x_2. A similar consideration applies to $\eta_1(x_1, x_2)$ and therefore the following lemma is proved.

Lemma 3. Considered as functions of x, the adjoint variables $\eta_1(x)$ and $\eta_2(x)$ are respectively non-increasing and non-decreasing in the value of x_1/x_2.

We now prove a useful fact concerning the time derivative of $\Delta(t)$.

Lemma 4. The sign of $\dot{\Delta}(t)$ is positive or negative as $x_1(t)/x_2(t)$ is respectively greater or less than some θ^*.

Proof. Differentiating Δ with respect to time along the optimal trajectory gives

$$\dot{\Delta} = (c_1 - c_2)(\dot{x}_1 + \dot{x}_2) + (\alpha_2 \dot{x}_2 \eta_1 - \alpha_1 \dot{x}_1 \eta_2) + (\alpha_2 x_2 \dot{\eta}_2 - \alpha_1 x_1 \dot{\eta}_1).$$

Substituting for $\dot{x}_1$ and $\dot{x}_2$ from (2) and (3) and for $\dot{\eta}_1$ and $\dot{\eta}_2$ from (4) and (5), we obtain

$$\begin{aligned}\dot{\Delta} = & -(c_1 - c_2)(\alpha_1 u_1 x_1 + \alpha_2 u_2 x_2) + \alpha_2 \eta_2(\lambda_1 x_1 - \lambda_2 x_2 - \alpha_2 u_2 x_2) \\ & - \alpha_1 \eta_1(\lambda_2 x_2 - \lambda_1 x_1 - \alpha_1 u_1 x_1) - (\alpha_2 x_2 - \alpha_1 x_1)(c_1 u_1 + c_2 u_2) \\ & + \alpha_2 x_2(\alpha_2 u_2 \eta_2 + \lambda_2 \eta_2 - \lambda_2 \eta_1) - \alpha_1 x_1(\alpha_1 u_1 \eta_1 + \lambda_1 \eta_1 - \lambda_1 \eta_2) \\ = & (c_2 \alpha_1 x_1 - c_1 \alpha_2 x_2) + (\alpha_1 + \alpha_2)(\lambda_1 x_1 \eta_2 - \lambda_2 x_2 \eta_1).\end{aligned}$$

The lemma follows from the final expression above and Lemma 3.

Lemma 4 implies that there is at most one ratio $x_1(t)/x_2(t) = \theta^*$ for which

$\dot{\Delta}(t)=0$. An immediate consequence of this observation is that it is impossible for an optimal policy to have both $u_1(t)$ and $u_2(t)$ non-zero over some interval of time $I=[t_1,t_2)$, unless $x_1(t)/x_2(t)$ is constant during the interval I. The reason for this is that if $u_1(t)$ and $u_2(t)$ are both non-zero for $t\in I$ then we must have $\Delta(t)=0$ for $t\in I$, and so $\dot{\Delta}(t)=0$ for $t\in I$. By Lemma 4 this implies that $x_1(t)/x_2(t)$ is constant for $t\in I$. Therefore, except at possibly some isolated values of x_1/x_2, it is strictly optimal to take $u(t)$ equal to either $(1,0)$ or $(0,1)$.

We now derive an expression for the cost attained by a constant control $u(t)=u=(u_1,u_2)$, $(0\leqq t<\infty)$. Multiplying (2) and (3) by $(\lambda_1+\lambda_2+\alpha_2u_2)$ and $(\lambda_1+\lambda_2+\alpha_1u_1)$ respectively and adding the results we obtain

$$(\lambda_1+\lambda_2+\alpha_2u_2)\dot{x}_1+(\lambda_1+\lambda_2+\alpha_1u_1)\dot{x}_2=-(\alpha_1\lambda_2u_1+\alpha_2\lambda_1u_2+\alpha_1\alpha_2u_1u_2)(x_1+x_2).$$

Hence by integration we have

$$V(x,u)=\frac{(\lambda_1+\lambda_2+\alpha_2u_2)x_1+(\lambda_1+\lambda_2+\alpha_1u_1)x_2}{\alpha_1\lambda_2u_1+\alpha_2\lambda_1u_2+\alpha_1\alpha_2u_1u_2}(c_1u_1+c_2u_2). \tag{7}$$

Let $u^1(t)=u^1=(1,0)$ be the policy of searching location 1 for ever, and let $u^2(t)=u^2=(0,1)$ be the policy of searching location 2 for ever, $(0\leqq t<\infty)$. Let δ_1 be the unique number such that if $x_1(0)/x_2(0)=\delta_1$ and the control u^1 is employed then $x_1(t)/x_2(t)$ equals δ_1 for all $t>0$. This definition implies that if $x_1(0)/x_2(0)$ is greater than δ_1 and u^1 is employed then $x_1(t)/x_2(t)$ is greater than δ_1 for all $t>0$. The number δ_1 is found from

$$\delta_1=x_1/x_2=\dot{x}_1/\dot{x}_2=(\lambda_2x_2-\alpha_1x_1-\lambda_1x_1)/(\lambda_1x_1-\lambda_2x_2).$$

This gives a quadratic equation in x_1/x_2 whose positive root is the number δ_1 which was defined in the statement of the theorem. We examine the possibility that u^1 is optimal by solving (4) and (5) with $\dot{\eta}(t)=0$ and $u=u^1$. This gives

$$\eta_1(t)=(\lambda_1+\lambda_2)c_1/\alpha_1\lambda_2,\quad\text{and}\quad \eta_2(t)=(\lambda_1+\lambda_2+\alpha_1)c_1/\alpha_1\lambda_2. \tag{8}$$

With these values for the adjoint variables, some simple algebra shows that u^1 satisfies (6) provided $x_1/x_2\geqq\theta_1$, where θ_1 is the number defined in the statement of the theorem. It therefore follows from Part (b) of Lemma 2 that u^1 is optimal if $x_1(0)/x_2(0)\geqq\delta_1$ and $\delta_1\geqq\theta_1$. This is the situation in Case (a) of the theorem.

To complete the analysis of Case (a) of the theorem we must show that in this case it is optimal to search location 2 for all x_1/x_2 less than θ_1. To do this we consider the possibility that for some $\theta_0<\theta_1$ it might be optimal to search location 1 for x_1/x_2 just less than θ_0 and to search location 2 for x_1/x_2 just greater than θ_0. If this is the case then Lemma 3 implies $\Delta(0)=0$ and $\dot{\Delta}(0)>0$ for $x(0)$ such that $x_1(0)/x_2(0)=\theta_0$. Since by hypothesis $x_1(t)/x_2(t)$ is strictly increasing with time along the trajectory starting at $x(0)$, $\dot{\Delta}(t)$ must continue to be positive along this trajectory and therefore $\Delta(t)$ must be strictly positive when after some time t we have $x_1(t)/x_2(t) = \theta_1$. This contradicts the hypothesis that it is optimal to search location 1 for $x_1(t)/x_2(t) = \theta_1$, since this requires $\Delta(t) \leqq 0$. This concludes the argument which establishes a possible solution of the form (a) in the theorem. Case (b) is similar, with θ_2 and δ_2 defined in an analogous fashion.

Case (c) may occur if there is a ratio $\theta_3 = x_1/x_2$ which it would be optimal to maintain by sharing search effort between locations 1 and 2. For this to be possible θ_3 must lie between δ_1 and δ_2 so that searching locations 1 and 2 will cause the value of x_1/x_2 to decrease and increase respectively. Suppose that when the ratio $x_1(0)/x_2(0)$ is θ_3 it is optimal to share effort using a constant control $u(t) = u = (u_1, u_2)$. We derive two relationships between u and θ_3. Firstly, $x_1(t)/x_2(t)$ must have the constant value θ_3 along the trajectory $x(t) = \phi(t, u, x(0))$. So

$$\theta_3 = x_1/x_2 = \dot{x}_1/\dot{x}_2 = (\lambda_2 x_2 - \lambda_1 x_1 - \alpha_1 u_1 x_1)/(\lambda_1 x_1 - \lambda_2 x_2 - \alpha_2 u_2 x_2).$$

Secondly, $\Delta(t)$ must be 0 for all t. Calculating $\eta_1(t)$ and $\eta_2(t)$ from (4) and (5) with $\dot{\eta}(t) = 0$, this requirement is that for $x_1/x_2 = \theta_3$,

$$0 = (c_1 - c_2)(x_1 + x_2) - \frac{(\lambda_1 + \lambda_2 + \alpha_2 u_2)\alpha_1 x_1 - (\lambda_1 + \lambda_2 + \alpha_1 u_1)\alpha_2 x_2}{\alpha_1 \lambda_2 u_1 + \alpha_2 \lambda_1 u_2 + \alpha_1 \alpha_2 u_1 u_2}(c_1 u_1 + c_2 u_2).$$

Straightforward but tedious algebra eliminates u_1, u_2 from the two equations above and results in the requirement that $x_1/x_2 = \theta_3$ be a root of the cubic equation displayed in statement (c) of the theorem. One can check by examining the signs of this cubic and its derivatives at $\theta = 0$ that it has at most one positive root. Alternatively, one can argue that if there were more than one value of x_1/x_2 at which sharing were optimal then there would be at least two values of x_1/x_2 such that $\dot{\Delta} = 0$. But this contradicts Lemma 4.

Consider finally the possibility that there might be a value $x_1/x_2 = \theta_0$ such that it is optimal to search location 1 and cause x_1/x_2 to decrease when x_1/x_2 is just less than θ_0 and it is optimal to search location 2 and cause x_1/x_2 to increase when x_1/x_2 is just greater than θ_0. In this case, it cannot be optimal to employ u^1 in all states x such that $x_1/x_2 < \theta_0$. For if this is so then $V(x)$ can be found from (7) and $V(x)$ is seen to be differentiable. Since $V(x)$ is differentiable $\eta(t)$ is equal to $\nabla V(x)$ evaluated at $x(t)$. Calculation of $\nabla V(x)$ shows that $\eta_1(t)$ and $\eta_2(t)$ are given by (8). If this is a solution to (4), (5) and (6) for all initial states such that $x_1(0)/x_2(0) < \theta_0$ then it is also a solution for initial states such that $x_1(0)/x_2(0) \geqq \theta_0$, and hence u^1 is optimal for all initial states, in contradiction to the assumptions. Similarly, it cannot be optimal to employ u^2 in all states x such that $x_1/x_2 > \theta_0$. Thus the case under consideration must have at least three switching points. One of these is where $x_1/x_2 = \theta_0$. Two others must lie at values of x_1/x_2 which are greater and less than θ_0 and must be such that an optimal trajectory starting at an x with either of these values of x_1/x_2 would maintain that constant ratio of x_1 to

x_2. However, it was noted in the previous paragraph that there is at most one value of x_1/x_2 which would be maintained as constant by an optimal policy. Thus the case envisaged in this paragraph does not arise. This completes the analysis of possible trajectories and establishes that an optimal policy is one of the three types described in the theorem.

4. Concluding remarks

The analysis of the discrete-time version of the problem appears to be more difficult than that of the continuous-time version, but we have found no reason to believe the conjecture is not true. Some of the analysis in this paper can be carried through to a discrete-time model. By demonstrating that a particular cubic equation has at most one positive root it can be shown that there is at most one value of $p_1(t)$ which an optimal policy would maintain as constant by sharing search effort between the two locations. It can also be shown that there exist numbers δ_1 and θ_1 (defined in terms of parameters in the discrete-time formulation) such that if $\theta_1 \leqq \delta_1$ then it is optimal to search location 1 for $p_1(t)/p_2 \geqq \theta_1$. However, we have not been able to prove a result analogous to Lemma 4 or take the analysis of the discrete-time model further.

The continuous-time model may be generalized to one in which the object moves amongst $n > 2$ locations. We conjecture that there exist numbers $\pi_1, \pi_2, \cdots, \pi_n$ (which depend upon the parameters of the problem) such that an optimal policy can be described as: search a location i for which $\pi_i p_i(t) \geqq \pi_j p_j(t)$, all $j \neq i$.

Our model assumed that some location is to be searched at every instant. Suppose we widen the class of admissible policies to those which sometimes search neither location but do find the object in finite expected time. Then there may be some values of p_1/p_2 for which it is optimal to search neither location. At these values it is optimal to wait until the value of p_1/p_2 becomes more attractive. We conjecture that within this class of policies the problem can be treated by similar methods to those in this paper and that there exist thresholds Π_1 and Π_2 with $0 \leqq \Pi_1 \leqq \Pi_2 \leqq 1$ such that an optimal policy is the form: search locations 1 or 2 or do not search either location as $p_1(t)$ is in the interval $[\Pi_2, 1]$, $[0, \Pi_1]$ or (Π_1, Π_2) respectively. For other examples of the use of optimal control theory in stochastic dynamic optimization problems see Nash and Gittins (1977) and Weber (1982).

Acknowledgement

I should like to thank Professors Pravin Varaiya and Jean Walrand of the University of California, Berkeley, and Professor P. R. Kumar of the University of Illinois for some very helpful discussion of this problem. I am grateful to the referee for pointing out that the type of optimal control formulation used in this paper has been discussed by Rishel (1982) in the general setting of a partially observed controlled-jump Markov process.

References

NASH, P. AND GITTINS, J. C. (1977) A Hamiltonian approach to stochastic resource allocation. *Adv. Appl. Prob.* **9**, 55–68.

RISHEL, R. W. (1982) Unnormalized conditional probabilities and optimality for partially observed controlled jump Markov processes, in *Advances in Filtering and Optimal Stochastic Control*, ed. W. H. Fleming and L. G. Gorostiza, Lecture Notes in Control and Information Sciences **42**, Springer-Verlag, New York, 326–343.

ROSS, S. M. (1983) *Introduction to Stochastic Dynamic Programming.* Academic Press, San Francisco, Cal.

VARAIYA, P. (1972) *Notes on Optimization.* Van Nostrand Reinhold, New York.

WEBER, R. R. (1982) Scheduling jobs with stochastic processing times on parallel processors to minimize makespan or flowtime. *J. Appl. Prob.* **19**, 167–182.