CONTRIBUTIONS TO THE THEORY OF OPTIMAL SEARCH

TERUHISA NAKAI

FACULTY OF ENGINEERING SCIENCE

OSAKA UNIVERSITY

CONTENTS

Introduction

Search theory is one of the oldest and important areas of operations research. The initial important investigation was mode by the Koopman's group in the US Navy during World War II to offer efficient methods of detecting enemy submarines. This theory has developed exceedingly after the war and is anticipated as a useful model in many practical applications. For example, possible areas of its application include (1) exploration of petroleum, (2) optimal allocation of law enforcement effort, (3) searching for a criminal, (4) searching for a victim in a sea or a mountain, (5) detecting a broken part of a machine, (6) medical examination for cancer and (7) optimal search for research data. Here we shall provide a simple survey of search theory and a guide for each section of this thesis. Bibliographies and surveys on search can be found in Enslow (1966) and Dobbie (1968). Main published books on search are Stone (1975), Gal (1980) and Koopman (1980).

A search for an object is developed in a search space by a searcher. In some cases a search space consists of discrete points which are usually called "boxes": in others it is a continuum. If an effort expended to search an object can be continuously (discretely) divisible, it is called a continuous (discrete) search effort. One which plays an important part in search theory is a detection function. In the case of continuous effort, let $b(x, z)$ be the conditional probability that an object is detected by the amount z of search effort applied in a position x given it is in x. An important example of detection functions is the exponential detection function $b(x, z) = 1 - \exp\{-\lambda(x) z\}$ where $\lambda(x)$ is a detection rate at a position x. In the case of discrete effort, the unit of effort applied in any box is usually called a "look". Let β_i be the conditional probability that an object is detected by one look in box i given that it is in box i.

Search theory can be divided broadly into two parts. One is one-sided search in which the object cannot take actions of its own free will, that is, it is allowed to move but not to evade. Another is two-sided search in which the object is allowed to take actions of its own free will. In one-sided search, a Bayesian approach is taken, that is, it is assumed that there is a priori distribution of the object's location which is known to the searcher. The problem is to find an allocation of search effort which is optimal under a performance criterion and therefore the main mathematical tools are calculus of variations, dynamic programming and nonlinear programming. Many search models can be considered according to various characters of the searcher and the object. If the object stays at one location and cannot move during a search process, it is called a stationary object, and if it moves in the search space according to a known probability law, it is called a moving target. A fundamental model for a stationary object is as follows: How should the given total effort be allocated in order to maximize the probability of detecting the object located with a priori distribution known to the searcher? This model is solved by Koopman (1957) and deGuenin (1961). Dobbie (1963) laid a foundation for a sequential allocation of search effort and established an important relation between maximizing the detection probability by a given total effort and minimizing the expected total effort until the detection. A list of papers treating a search for a stationary object includes Gluss (1961), Matula (1964), Ross (1969), Kadane (1971), Sakaguchi (1973), Hall (1976) and Barker (1977). Pollock (1970) considered a discrete search for a target moving in two boxes according to a known Markov chain. Dobbie (1974) treats a continuous-time version of the Pollock's model. Stone (1979) obtains necessary and sufficient conditions for a policy to be optimal in a relatively general search for a moving target.

In two-sided search, a game theoretical approach is taken mostly. In this area, there are three problems: (i) Hide-Search Problem. At the beginning of search, player I (Hider) hides in one location and cannot move to other locations during the search process. There are works of Gittins and Roberts (1979) and Suberman (1981). (ii) Evasion-Search Problem. Player I (Evader) can evade (by his own free will) during the search process. There are works of Norris (1962), Sakaguchi (1973), Washburn (1980) and Stewart (1981). (iii) Search-Search Problem. Two friendly players attempt to find each other with limitted information and communications. This type is called a rendezvous problem and there is no paper yet.

The aim of this thesis is to solve (i.e. to find an optimal search policy) some search models and to investigate the learning process by the searcher. Outline of each section is given in the followings:

Section 1.1. A search for multi objects with the same characters, for example, the reward of the searcher for detection of the object and the prior distribution of its location, etc., has been solved in Smith and Kimeldorf (1975). In this section we consider a two-box model of searching one of two objects with different rewards and different forms of the prior distributions. Since search effort is continuous, the model is formulated and solved by calculus of variations. The optimal policy is to search only one box until a threshold time and thereafter to search both boxes in the ratio of inverses of the detection rates. Furthermore the transition of the posterior distribution of the object's location is investigated. The work in this section is based on Nakai (1976).

Section 1.2. Dobbie (1973) formulated a search model with finite false objects by a dynamic programming technique. In this section we solve a two-box model of a search for a true object against an interruption of a false object along the line of the Dobbie's formulation. The analysis is carried out by the same method as in Section 1.1. The optimal policy consists of two stages: the search before the first detection of an object and that for the object after the false object has been detected. This section is based on Nakai (1976).

Section 1.3. Considering a search and rescue for a lost alpinist, we formulate a search model for an object with a random lifetime by means of calculus of variations and obtain necessary and sufficient conditions for a policy to be optimal. In the case that the lifetime density function at each location is differentiable at all time, the optimal search policy is derived from the above conditions. Two numerical examples are given which indicate that the location with the small expected lifetime must be searched first even if the efficiency of search there is bad. Furthermore we analyze a search and stop problem in which the searcher is permitted to stop the search at any time. This section is based on Nakai (1982).

Section 1.4. In all literatures on search theory the conditional detection probability was always constant in time (or period). In this section a search model in which it varies in time is treated by means of a dynamic programming approach. The objective is to maximize the probability of detecting the object during some periods. A search policy depends on the order of searches and a myopic policy (which searches at each period a box maximizing the current detection probability) is not necessarily optimal. The optimal policy is obtained in the two-period case and in the two-box case in which a conditional probability in one box is constant in time. A numerical example of the latter case is given.

This section is based on Nakai (1981).

Section 2.1. The problem in this section is to find a policy minimizing the expected number of looks to detect a target which moves circularly among n boxes according to a known Markov chain. It is shown that the posterior distribution of the target's location can be represented by a point in the posterior simplex which is defined by only the transition matrix. In some three-box cases the optimal policies are derived under the assumption of perfect detection. This section is based on Nakai (1973).

Section 2.2. In the search for a moving target, it is sometimes reasonable to wait, i.e., to expend time without search in anticipation of the transition of the target to the more desirable location for the searcher. To see this point, we allow the searcher to wait in the Pollock's model (1970) in which a target moves between two boxes according to a known Markov chain. The optimal policy has a property that the larger a waiting cost becomes, the narrower the region in which to wait is optimal becomes and that for a sufficiently large waiting cost, no waiting is optimal. We discuss the efficiency of waiting in a search for a moving target. This section is based on Nakai (1980).

Section 2.3. We consider a problem of catching timely a target which appears and disappears randomly. The target moves among three boxes according to a known Markov tansition matrix.

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1-a & a & 0 \\ 0 & 1-b & b \\ 0 & 0 & 1 \end{bmatrix} \end{array}.$$

At each period the searcher must choose one of three actions: to wait, to search box 2 and to stop. The objective is to minimize the expected total loss until the completion of the process. The optimal policy is obtained in the cases $a+b \geq 1$ and $b \geq (\beta R/c-1)/a$ where c, R are a search cost and a reward of detecting respectively. So far, there appears no paper treating this problem.

Chapter III. A two-person sequential evasion-search game is considered in which player I (Evader) can move among n boxes in one direction and player II (Searcher) can search any box at each period knowing the evader's previous position. The optimal solution in a zero-sum case is obtained and a feedback Nash equilibrium solution is obtained in a nonzero-sum case. We give two numerical examples. This chapter is based on Nakai (in submission).

Chapter I Optimal Search for a Stationary Object

§1.1 Optimal Search for One of Two Stationary Objects

Concerning a multi-object search there are two papers. Smith and Kimeldorf (1975) considers the following model: N objects are hidden in m boxes where m is known and N is a random variable having a priori distribution $w = \langle w_1, w_2, \ldots \rangle$; $w_n \equiv \text{prob.}\{N = n\}$. Each object is located with a priori distribution $p = \langle p_1, \ldots, p_m \rangle$ independently of the location of other objects. Associated with box i are search cost c_i and a conditional detection probability β_i. The objective is to minimize the expected total cost expended to find at least one object. Main results are (i) If m = 2 and $w_1 + w_2 = 1$, then an optimal policy prescribes searching a box maximizing $[1-\sum_{n=1}^{2} w_n(1-p_i\beta_i)^n]/c_i$ and (ii) If w is a possitive-Poisson distribution with parameter $\lambda(>0)$, then it is optimal to search a box maximizing $[1-\exp(-\lambda p_i\beta_i)]/c_i$. Kimeldorf and Smith (1979) considers the same problem with one change: the criterion is the expected total cost to find all objects.

In the above models they take account of only search cost. But a reward for detecting an object must be taken into account since each object has usually a respective value. In this section we shall consider a two-box, two-object search problem with rewards. There are two boxes which contain two stationary objects. Let $p_k (k=1, 2)$ be a priori probability that object k is in box 1. Let $c_i (>0)$ be a search cost rate, i.e., a cost required to search box i per unit time $(i=1, 2)$. Assume that if box i containing object k is searched during time t, the searcher can detect it with probability $1 - \exp(-\lambda_i t)$ $(\lambda_i > 0)$ and receive a reward $R_k (-\infty < R_k < \infty)$. Suppose that if we search a box containing two objects, the detection of an object is independent of the existence of each other. The objective is to minimize the expected total loss (i.e., cost minus reward) until at least one object is detected.

A search policy can be described by a function $\phi(t)$ which represents a length of search time in box 1 by time t and satisfies the Condition [A]:
(i) $0 \leq \phi(t) \leq t$ (ii) $\phi(t)$ is piece-wise differentiable in $t(\geq 0)$
(iii) $0 \leq \phi'(t) \leq 1$ for any $t \geq 0$ (iv) $\phi(t)$ and $t-\phi(t)$ converge to infinite as t approaches to infinite. Let $Q_k(t)$ be a probability that object k is not detected until time t by using a policy ϕ. Then we have

$$(1.1.1) \quad Q_k(t) = p_k\exp[-\lambda_1\phi(t)] + (1-p_k)\exp[-\lambda_2\{t-\phi(t)\}] \quad (k=1, 2).$$

Put $P_k(t) = 1 - Q_k(t)$, $Q(t) = Q_1(t)Q_2(t)$ and $P(t) = 1 - Q(t)$. If the first detection occurs at time t, the conditional probability that it is made on object k is

$$(1.1.2) \quad q_k(t) = P_k'(t)Q_{3-k}(t)/P'(t) \qquad (k=1, 2).$$

Let $L(\phi)$ be the expected total loss until the first detection by using a policy ϕ. Then we have

$$(1.1.3) \quad L(\phi) = \int_0^\infty P'(t)[c_1\phi(t)+c_2\{t-\phi(t)\}-q_1(t)R_1-q_2(t)R_2]dt.$$

Therefore the problem is to find a function $\phi^*(t)$ which minimizes $L(\phi)$ subject to Condition [A]. From relations (1.1.2) and (1.1.3) we have

$$(1.1.4) \quad L(\phi) = \int_0^\infty \{(c_1-c_2)\phi(t)+c_2t-R_1\}\{-Q_1'(t)Q_2(t)\}dt + \int_0^\infty \{(c_1-c_2)\phi(t)+c_2t-R_2\}\{-Q_1(t)Q_2'(t)\}dt.$$

Integrating by parts and taking (iv) of Condition [A] into account, we obtain

$$(1.1.5) \quad L(\phi) = \int_0^\infty W[\phi(t), t]dt + (\text{terms independent of } \Phi)$$

where

$$(1.1.6) \quad W[\phi, t] = \begin{cases} c_1(1-p_1)(1-p_2)\exp[2\lambda_2(\phi-t)]+c_2p_1p_2\exp(-2\lambda_1\phi) \\ -(\lambda_1-\lambda_2)^{-1}[(c_1\lambda_2-c_2\lambda_1)(p_1+p_2-2p_1p_2)-\lambda_1\lambda_2(R_1-R_2)(p_1-p_2)] \\ \times \exp[(\lambda_2-\lambda_1)\phi-\lambda_2t] \qquad \text{if } \lambda_1 \neq \lambda_2 \\ c_1(1-p_1)(1-p_2)\exp[2\lambda(\phi-t)]+c_2p_1p_2\exp(-2\lambda\phi) \\ + \lambda[(c_1-c_2)(p_1+p_2-2p_1p_2)-\lambda(R_1-R_2)(p_1-p_2)]\phi\exp(-\lambda t) \\ \qquad \text{if } \lambda_1 = \lambda_2 = \lambda. \end{cases}$$

In (1.1.6), the expression in the case of $\lambda_1 \neq \lambda_2$ does not converge to that in the case of $\lambda_1 = \lambda_2$ as λ_1 approaches to λ_2. But we know that there is no contradiction if the terms independent of ϕ in (1.1.5) are taken into account. It is enident that the solution is a function $\phi^*(t)$ which minimizes $W[\phi(t), t]$ for all $t(\geq 0)$, subject to Condition [A], provided such a function exists. Partially differentiating, we have

(1.1.7) $$\frac{\partial W}{\partial \phi} = 2c_1\lambda_2(1-p_1)(1-p_2)\exp[2\lambda_2(\phi-t)]-2c_2\lambda_1 p_1 p_2 \exp(-2\lambda_1\phi)$$
$$+ [(c_1\lambda_2-c_2\lambda_1)(p_1+p_2-2p_1p_2)-\lambda_1\lambda_2(R_1-R_2)(p_1-p_2)]\exp[(\lambda_2-\lambda_1)\phi-\lambda_2 t]$$
$$= (AU^2-BU-C)\exp(-2\lambda_1\phi)$$

where
$$A = 2c_1\lambda_2(1-p_1)(1-p_2)$$
$$B = \lambda_1\lambda_2(R_1-R_2)(p_1-p_2)-(c_1\lambda_2-c_2\lambda_1)(p_1+p_2-2p_1p_2)$$
$$C = 2c_2\lambda_1 p_1 p_2$$
$$U = U(\phi, t) = \exp[(\lambda_1+\lambda_2)\phi-\lambda_2 t]$$

If $0 < p_k < 1$ for $k = 1, 2$, then A and C are positive and therefore the solution of the equation $\partial W/\partial\phi = 0$ is given by

(1.1.8) $\phi(t) = \alpha(t) \equiv (\lambda_2 t+u)/(\lambda_1+\lambda_2)$

(1.1.9) $u \equiv \log[\{B+(B^2+4AC)^{1/2}\}/(2A)]$.

Since $\partial W/\partial\phi < 0$ for $0 < \phi < \alpha(t)$ and $\partial W/\partial\phi > 0$ for $\alpha(t) < \phi$, $W[\phi(t), t]$ is minimized at $\phi = \alpha(t)$ if there is no restriction on $\phi(t)$. But because of Condition [A], we must find the value of $\phi(t)$ which minimizes $W[\phi(t), t]$ on an interval $[0, t]$ for a fixed $t(\geq 0)$. Hence $W[\phi(t), t]$ is minimized at $\phi(t) = \begin{Bmatrix} t \\ \alpha(t) \\ 0 \end{Bmatrix}$ if $\begin{Bmatrix} 0 \leq t \leq \alpha(t) \\ 0 \leq \alpha(t) \leq t \\ \alpha(t) \leq 0 \leq t \end{Bmatrix}$. It is clear that their values of $\phi(t)$ satisfy Condition [A]. Therefore the optimal function $\phi^*(t)$ is as follows: In the case of $u \geq 0$, if $0 \leq t \leq u/\lambda_1$, then $\phi^*(t) = t$ (i.e., to search only box 1) and if $u/\lambda_1 \leq t$, then $\phi^*(t) = (\lambda_2 t+u)/(\lambda_1+\lambda_2)$ (i.e., to search box 1 and 2 in the rate λ_1^{-1}: λ_2^{-1}). In the case of $u \leq 0$, if $0 \leq t \leq -u/\lambda_2$, then $\phi^*(t) = 0$ (i.e., to search only box 2) and if

$-u/\lambda_2 \leq t$, then $\phi^*(t) = (\lambda_2 t+u)/(\lambda_1+\lambda_2)$ (i.e., to search box 1 and 2 in the ratio $\lambda_1^{-1}:\lambda_2^{-1}$).

Define $t_i^* \equiv (-1)^{i+1}u/\lambda_i$, i.e., t_i^* is the threshold time which indicates the switching time of the optimal search method. Furthermore we put $D_1 \equiv \{(p_1, p_2) | u > 0\}$ and $D_2 \equiv \{(p_1, p_2) | u < 0\}$, i.e., D_i $(i=1, 2)$ is a region of the pair of the prior probabilities in which to search only box i until the threshold time t_i^* is optimal. The boundary between two regions D_1 and D_2 is given by the straight line

(1.1.10) $(r+c)p_1-(r-c)p_2 = 2c_1/\lambda_1$

where $r = R_1 - R_2$ and $c = c_1/\lambda_1 + c_2/\lambda_2$. This line separates the origin and the point (1, 1). (See Fig. 1.1.1.)

Here we investigate how the posterior location probabilities vary. Let p_k' $(k=1, 2)$ be the posterior probability that object k is in box 1 given that no detection occurs until time t by the optimal policy ϕ^*. By the Bayes' rule, we obtain

$$p_k' = \begin{cases} p_k e^{-\lambda_1 t}/(p_k e^{-\lambda_1 t} + 1-p_k) & \text{if } (p_1, p_2) \in D_1 \text{ and } t \leq t_1^* \\ p_k e^{-u}/(p_k e^{-u} + 1-p_k) & \text{if } (p_1, p_2) \in D_1 \text{ and } t_1^* \leq t \\ p_k/[p_k + (1-p_k)e^{-\lambda_2 t}] & \text{if } (p_1, p_2) \in D_2 \text{ and } t \leq t_2^* \\ p_k/[p_k + (1-p_k)e^{u}] & \text{if } (p_1, p_2) \in D_2 \text{ and } t_2^* \leq t. \end{cases}$$

If $(p_1, p_2) \in D_i$ and $t \leq t_i^*$, we have

$$\log(p_k'^{-1}-1) = \log(p_k^{-1}-1) + \lambda_i t$$

which can be transformed to, by eliminating t,

(1.1.11) $(p_2'^{-1}-1)/(p_1'^{-1}-1) = (p_2^{-1}-1)/(p_1^{-1}-1)$.

Hence the pair of two posterior probabilities varies along the curve (1.1.11) until the threshold time and stays there (the point A or B in Fig. 1.1.1) afterwards. We summarize the above discussion in the following theorem.

Theorem 1.1.1. If $(p_1, p_2) \in D_i (i=1, 2)$ in Fig. 1.1.1, the optimal policy is to search only box i until the threshold time $t_i^* = (-1)^{i+1}u/\lambda_i$

and thereafter to search box 1 and 2 in the ratio $\lambda_1^{-1}: \lambda_2^{-1}$ where the boundary between two regions D_1 and D_2 is given by (1.1.10) and u is defined by (1.1.9). Furthermore if the optimal policy is used, the pair (p_1', p_2') of two posterior location probabilities varies along the curve given by (1.1.11).

The following points must be noted.

(i) The ratio with which the optimal policy searches both boxes after the threshold time depends not on search cost c_i but on the search rate λ_i $(i=1, 2)$.

(ii) The optimal policy depends on rewards R_k $(k=1, 2)$ only through their difference $R_1 - R_2$.

(iii) Consider the case that detecting object 2 results in receiving an enormous loss and going bankrupt, i.e., the searcher wants to find object 1 with avoiding to meet object 2. This case occurs when R_2 approaches to infinity. In this case, the optimal policy is to search only box k if $p_k > p_{3-k}$ $(k=1, 2)$.

(iv) Rewriting the equation (1.1.10), we obtain

$$(1.1.12) \quad \frac{(p_1+p_2)\lambda_1}{c_1} = \frac{(\bar{p}_1+\bar{p}_2)\lambda_2}{c_2} - \frac{(R_1-R_2)(p_1-p_2)\lambda_1\lambda_2}{c_1c_2} \qquad (\bar{p}_i \equiv 1-p_i).$$

When $R_1 = R_2$, Theorem 1.1.1 states that if

$$(1.1.13) \quad \frac{\frac{p_1+p_2}{2}\lambda_1}{c_1} \begin{Bmatrix} \geq \\ \leq \end{Bmatrix} \frac{\frac{\bar{p}_1+\bar{p}_2}{2}\lambda_2}{c_2},$$

then the optimal policy is to search only $\begin{Bmatrix} \text{box } 1 \\ \text{box } 2 \end{Bmatrix}$ until the equality holds in (1.1.13) and thereafter to search box 1 and 2 in the rate $\lambda_1^{-1}: \lambda_2^{-1}$. Here the left hand side of the inequality (1.1.13) represents the detection rate in box 1 per unit search cost for one object with the prior distribution $\langle p, 1-p \rangle$ where $p = (p_1+p_2)/2$, i.e., the mean value of two location probabilities in box 1. Therefore if $R_1 = R_2$, we can regard the two-object case as the one-object case. The last term in (1.1.12) represents a bias by the fact that $R_1 = R_2$.

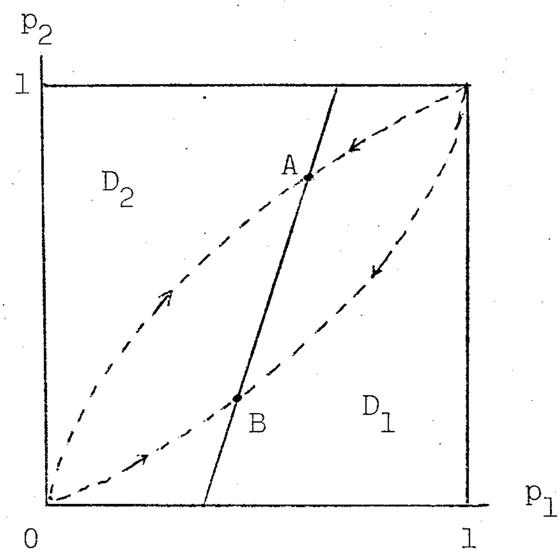

Fig.1.1.1 Regions D_i $(i = 1, 2)$ and the transition of the pair of the posterior location probabilities (p_1', p_2') in Section 1.1.

§1.2 Search Problem with a False Object

When there are false objects which cannot be distinguished from the true object except by a close inspection, the search takes place in two phases. The first phase is carried out by a sensor which can detect an object but not positively identify it. The second phase is to investigate whether the contact is true. A search problem with false objects is studied by Stone and Stanshine (1971). They assume that a search policy does not depend on the number of false objects which were found and eliminated. But if the number of false objects is bounded, this assumption does not hold in general. Dobbie (1973) formulates a search model with finite false objects by a dynamic programming technique under the assumption that a policy is contingent on finding false objects. But he does not solve the problem in general except a simple example of two-box case.

We consider a search problem of a Dobbie's type and obtain its optimal policy by the calculus of variations. There are two boxes and one true object is in one of them according to a priori distribution $\langle p_1, 1-p_1 \rangle$. One false object hides in either box according to a priori distribution $\langle p_2, 1-p_2 \rangle$. Both objects are stationary during search process. Assume that if a box containing a certain object (whether it is true or not) is searched for time t, it is contacted with probability $1-\exp(-t)$. The contact with an object is assumed to be independent of another. If a contact occurs, the search will be interrupted and time ℓ (> 0) will be spent to identify the contact. If the contact is the true object, the search will be stopped. If the contact is the false object, we discard it and reopen the search for the remaining true object. The objective is to minimize the expected time to detect the true object. An example which was solved by Dobbie is the case of $p_2 = 0$ and $\ell = 1$ in our model.

We call the search process until the first contact the first stage and that after discarding the false object the second stage. In the second stage, the problem is to detect one stationary object without false objects and therefore it can be solved by the well-known method as follows: Let p be a probability that the true object is in box 1 at the beginning of the second stage. If $0 \leq p \leq 1/2$ $(1/2 \leq p \leq 1)$, it is optimal to search only box 2 (box 1) until time $|\log(p^{-1}-1)|$ and to search equally both boxes thereafter. The expected time to detect the true object by the optimal policy is expressed by

$$(1.2.1) \quad f_2(p) = (3+|u|+\exp|u|)/(1+\exp|u|)$$

where $u \equiv \log(p^{-1}-1)$.

Now we consider the search process in the first stage. A search policy in the first stage can be described by a function $\phi(t)$ which represents search time in box 1 by time t and satisfies the Condition [A] in Section 1.1. If the first contact by a policy ϕ occurs at time t and is false, a probability that the true object is in box 1 at the beginning of the second stage is given by

$$(1.2.2) \quad \tilde{p}_1(t|\phi) = \frac{p_1\exp[-\phi(t)]}{p_1\exp[-\phi(t)] + (1-p_1)\exp[-t+\phi(t)]} .$$

Substituting (1.2.2) into the relation $u = \log\{\tilde{p}_1(t|\phi)^{-1}-1\}$, we obtain

$$(1.2.3) \quad u = 2\phi(t) - t + \log(p_1^{-1}-1).$$

It is evident that if $p_1 = \left\{\begin{matrix} 0 \\ 1 \end{matrix}\right\}$, the optimal solution is $\phi^*(t) = \left\{\begin{matrix} 0 \\ t \end{matrix}\right\}$ for any $t(\geq 0)$ and therefore we assume that $0 < p_1 < 1$. Let $Q(t|\phi)$ be a probability that the first contact by a policy ϕ does not occur by time t. Then using (1.2.3), we have

$$(1.2.4) \quad Q(t|\phi) = \prod_{i=1}^{2} [p_i\exp\{-\phi(t)\} + (1-p_i)\exp\{\phi(t)-t\}]$$

$$= e^{-t}[p_1(1-p_2)(1+e^{u}) + (1-p_1)p_2(1+e^{-u})].$$

We put $P(t|\phi) \equiv 1 - Q(t|\phi)$. Let $q(t|\phi)$ be a conditional probability that the first contact is false given that it occurs at time t by a policy ϕ. Then

$$(1.2.5)\quad q(t|\phi) = [P'(t|\phi)]^{-1}[p_1\exp\{-\phi(t)\}+(1-p_1)\exp\{\phi(t)-t\}]$$
$$\times \frac{d}{dt}[1-p_2\exp\{-\phi(t)\} - (1-p_2)\exp\{\phi(t)-t\}]$$
$$= [2P'(t|\phi)]^{-1}e^{-t}[p_1(1-p_2)(1-u')(1+e^u)+(1-p_1)p_2(1+u')(1+e^{-u})].$$

Let $f(p_1, p_2|\phi)$ be the expected time to detect the true object by using a policy ϕ in the first stage and the optimal policy in the second stage given that the prior location distributions of both objects are $\langle p_1, 1-p_1\rangle$ and $\langle p_2, 1-p_2\rangle$. Then we have

$$(1.2.6)\quad f(p_1, p_2|\phi) = \int_0^\infty P'(t|\phi)[t+\ell+q(t|\phi)\{f_2(\tilde{p}_1(t|\phi))+\ell\}]dt$$
$$= \ell + \int_0^\infty Q(t|\phi)dt + \int_0^\infty P'(t|\phi)q(t|\phi)\{f_2(\tilde{p}_1(t|\phi))+\ell\}dt.$$

Substituting (1.2.4) and (1.2.5) into (1.2.6) and noting Condition [A], we obtain

$$(1.2.7)\quad f(p_1, p_2|\phi) = \frac{1}{2}\int_0^\infty e^{-t}W[u(t)]dt + (\text{terms independent of } \phi)$$

where

$$(1.2.8)\quad W(u) = \begin{cases} p_1(1-p_2)\{2e^u-u^2/2-(\ell+2)u\}+(1-p_1)p_2\{e^{-u}+(\ell+1)u+1\} & \text{if } u \geq 0 \\ p_1(1-p_2)\{e^u-(\ell+1)u+1\}+(1-p_1)p_2\{2e^{-u}-u^2/2+(\ell+2)u\} & \text{if } u \leq 0. \end{cases}$$

Furthermore Condition [A] becomes Condition [A$'$] defined by

$$\text{Condition [A}'\text{]} \begin{cases} \text{(i)} & u(0) = \log(p_1^{-1}-1) & \\ \text{(ii)} & u(t) \text{ is piece-wise differentiable in } t & \\ \text{(iii)} & |u(t) - \log(p_1^{-1}-1)| \leq t & \text{for } t \geq 0 \\ \text{(iv)} & -1 \leq u'(t) \leq 1 & \text{for } t \geq 0 \end{cases}$$

Therefore the function $u^*(t)$ which minimizes $W[u(t)]$ for all t and satisfies Condition [A$'$] is optimal, provided such a function exists.

$$(1.2.9)\quad W'(u) = \begin{cases} p_1(1-p_2)(2e^u-u-\ell-2) + (1-p_1)p_2(-e^{-u}+\ell+1) & \text{if } u \geq 0 \\ p_1(1-p_2)(e^u-\ell-1) + (1-p_1)p_2(-2e^{-u}-u+\ell+2) & \text{if } u \leq 0. \end{cases}$$

$$(1.2.10)\quad W''(u) = \begin{cases} p_1(1-p_2)(2e^u-1) + (1-p_1)p_2e^{-u} & \text{if } u \geq 0 \\ p_1(1-p_2)e^u + (1-p_1)p_2(2e^{-u}-1) & \text{if } u \leq 0. \end{cases}$$

Since $W''(u) > 0$ for all u, $W'(u)$ is strictly increasing in u. Furthermore $W'(0) = \ell(p_2-p_1)$, $W'(+\infty) = +\infty$ and $W'(-\infty) = -\infty$. Hence the equation $W'(u) = 0$ has a unique root u_0 such that if $p_1 \geq (\leq) p_2$, $u_0 \geq (\leq) 0$. Therefore the optimal function $u^*(t)$ is a function which starts at $u(0)$ at $t = 0$, goes to u_0 as rapidly as the constraints permit and remains unchanged thereafter. Hence if we put $t^* = |u_0 - u(0)|$, $u^*(t)$ is given by

$$(1.2.11)\quad \begin{cases} u(0) \geq u_0 \implies u^*(t) = \begin{cases} u(0) - t & \text{if } 0 \leq t \leq t^* \\ u_0 & \text{if } t^* \leq t \end{cases} \\ \\ u(0) \leq u_0 \implies u^*(t) = \begin{cases} u(0) + t & \text{if } 0 \leq t \leq t^* \\ u_0 & \text{if } t^* \leq t \end{cases} \end{cases}$$

Since $W'(u)$ is strictly increasing in u,

$$(1.2.12)\quad u(0)\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} u_0 \iff W'[u(0)] \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} W'[u_0] = 0.$$

Substituting $u(0) = \log(p_1^{-1}-1)$ into (1.2.9), we have

$$(1.2.13)\quad W'[u(0)] = \begin{cases} (1-2p_1)(2-p_2)+(p_2-p_1)\ell-p_1(1-p_2)\log(p_1^{-1}-1) \\ (1-2p_1)(1+p_2)+(p_2-p_1)\ell-(1-p_1)p_2\log(p_1^{-1}-1) \end{cases} \text{if } u(0)\begin{Bmatrix} \geq \\ \leq \end{Bmatrix} 0.$$

If we define

$$(1.2.14)\quad B(p_1, p_2) \equiv \begin{cases} (1-2p_1)(2-p_2)+(p_2-p_1)\ell-p_1(1-p_2)\log(p_1^{-1}-1) \\ (1-2p_1)(1+p_2)+(p_2-p_1)\ell-(1-p_1)p_2\log(p_1^{-1}-1) \end{cases} \text{if } p_1\begin{Bmatrix} \leq \\ \geq \end{Bmatrix} \frac{1}{2},$$

then $W'[u(0)] = B(p_1, p_2)$ since $u(0) \geq 0$ (≤ 0) is equivalent to $p_1 \leq 1/2$ ($p_1 \geq 1/2$). Hence by (1.2.12), we obtain

$$(1.2.15)\quad u(0)\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} u_0 \iff B(p_1, p_2)\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0.$$

Therefore by (1.2.3) and (1.2.11), the optimal policy can be obtained as follows: If $B(p_1, p_2) \geq 0$,

$$(1.2.16)\quad \phi^*(t) = \begin{cases} 0 \\ \quad \text{(to search only box 2)} \\ \frac{t}{2} - \frac{1}{2}[\log(p_1^{-1}-1)-u_0] \\ \quad \text{(to search equally both boxes)} \end{cases} \text{if } t\begin{Bmatrix} \leq \\ \geq \end{Bmatrix}\log(p_1^{-1}-1)-u_0,$$

and if $B(p_1, p_2) \leq 0$,

$$(1.2.17)\quad \phi^*(t) = \begin{cases} t \\ \quad \text{(to search only box 1)} \\ \frac{t}{2} + \frac{1}{2}[u_0-\log(p_1^{-1}-1)] \\ \quad \text{(to search equally both boxes)} \end{cases} \text{if } t\begin{Bmatrix} \leq \\ \geq \end{Bmatrix}u_0-\log(p_1^{-1}-1).$$

We put $D_1 \equiv \{(p_1, p_2) | B(p_1, p_2) < 0\}$ and $D_2 \equiv \{(p_1, p_2) | B(p_1, p_2) > 0\}$. The boundary $B(p_1, p_2) = 0$ of two regions D_1 and D_2 is described in Fig. 1.2.1. Now we shall investigate how the pair of the posterior location probabilities of both objects in the first stage varies. When the optimal policy is used, let $\bar{p}_1(t)$ ($\bar{p}_2(t)$) be the posterior probability that the true (the false) object is in box 1 given that the first contact does not occur until time t. Then by the Bayes' rule, we have

$$(1.2.18) \quad \bar{p}_i(t) = \frac{p_i \exp\{-\phi^*(t)\}}{p_i \exp\{-\phi^*(t)\}+(1-p_i)\exp\{\phi^*(t)-t\}} \qquad (i=1, 2)$$

which lead us to

$$(1.2.19) \quad (\bar{p}_2^{-1}-1)/(\bar{p}_1^{-1}-1) = (p_2^{-1}-1)/(p_1^{-1}-1)$$

by eliminating $\phi^*(t)$ from (1.2.18). The pair of the posterior probabilities varies along the curve defined by (1.2.19). (See in Fig.1.2.1).
We summarize the above discussion in the following theorem 1.2.1.

Theorem 1.2.1 (a) The optimal policy in the first stage is as follows: If $(p_1, p_2) \in D_i$ $(i=1, 2)$ in Fig. 1.2.1, search only box i until the threshold time $t^* = |u_0 - \log(p_1^{-1}-1)|$ and search equally both boxes thereafter where u_0 is the unique root of the equation $W'(u) = 0$. (b) The optimal policy in the second stage is as follows: Let $\tilde{p}_1$ be a probability that the true object is in box 1 at the beginning of the second stage. If $\tilde{p}_1 \leq (\geq)$ 1/2, search only box 2 (box 1) until time $|\log(\tilde{p}_1^{-1}-1)|$ and search equally both boxes thereafter.

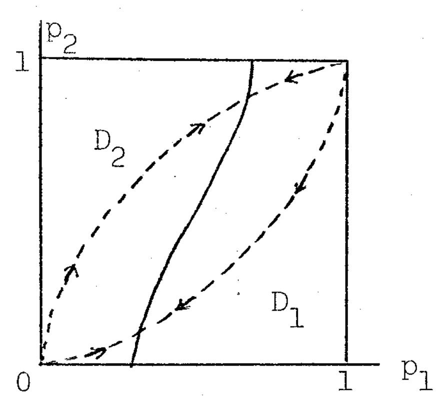

Fig.1.2.1 Region D_i $(i=1, 2)$ and the transition of the pair of the posterior location probabilities (p_1', p_2') in Section 1.2.

§1.3 Optimal Search for an Object with a Random Lifetime

In winter many alpinists are lost in a snowy mountain and the search for them is carried out. In such a search it is important to detect the alpinist alive. Since the capability of his survival depends exceedingly on the geographical feature and the weather at the accident place, we must search first in the place where he cannot live long even if the efficiency of search there is bad. Furthermore if the alpinist dies, his family wants to detect his dead body and therefore the search is continued until the detection of his body whether he lives or not.

The search problem in such a situation is modelled as follows: One stationary object is in one of n boxes with a priori distribution $\langle p_1, \ldots, p_n \rangle$ where p_i is a priori probability that the object is in box i. The lifetime of the object in box i is a random variable according to a probability distribution $F_i(t)$ which is composed of two probability masses α_i at $t = 0$, β_i at $t = \infty$ and a probability density function $f_i(t)$ on the time interval $(0, \infty)$. The mass α_i denotes a probability that the object in box i dies before the beginning of the search and the mass β_i denotes a probability that the object in box i is alive eternally. Note that $\alpha_i + \int_0^\infty f_i(t)dt + \beta_i = 1$ for $i = 1, \ldots, n$. Let c be a search cost per unit time which is assumed to be common to all boxes. Associated with box i is a search rate λ_i (> 0) which means that the object is detected with probability $1-\exp(-\lambda_i t)$ if box i contains the object and is searched for t hours. If the searcher detects the alive object in box i, he can obtain a reward r_i (> 0). But if the dead body of the object is detected, no reward is obtained. The problem is to find the optimal search policy, i.e., the allocation of search time maximizing the expected return (i.e., reward minus cost) until the object is detected whether it is alive or not.

A search policy may be denoted by a funciton $\phi(t) = \{\phi_1(t), \ldots, \phi_n(t)\}$ where $\phi_i(t)$ is search time in box i until time t and satisfies that $\sum_{i=1}^{n} \phi_i(t) = t$ and $\phi_i(t) \geq 0$ $(i = 1, \ldots, n)$ for any t (≥ 0). First we define some quantities.

$g_i(t) \equiv$ the expected return of detecting the object in box i at time t

$\quad = r_i[1-F_i(t)] - ct.$

$P_i(t|\phi) \equiv$ the probability of detecting the object by time t given that it is in box i and that a policy ϕ is used

$\quad = 1 - \exp\{-\lambda_i \phi_i(t)\}.$

$\tilde{R}[\phi] \equiv$ the expected return by using a policy ϕ

$$= \sum_{i=1}^{n} p_i \int_0^{\infty} g_i(t)\, P_i'(t|\phi)dt$$

where $P_i'(t|\phi)$ is the derivative by t. Integrating by parts,

$$\tilde{R}[\phi] = \sum_{i=1}^{n} p_i r_i (1-\alpha_i) + \sum_{i=1}^{n} p_i \int_0^{\infty} g_i'(t) \exp\{-\lambda_i \phi_i(t)\}dt. \tag{1.3.1}$$

Since the first term of the equation (1.3.1) is independent of the policy ϕ, the problem is formulated as follows:

$$R[\phi] \equiv \sum_{i=1}^{n} p_i \int_0^{\infty} g_i'(t) \exp\{-\lambda_i \phi_i(t)\}dt \longrightarrow \max_{\phi} \tag{1.3.2}$$

subject to

$$\sum_{i=1}^{n} \phi_i(t) = t \quad \text{for any } t\ (\geq 0), \tag{1.3.3}$$

(1.3.4) $\phi_i(t)$ is nonnegative, continuous and nondecreasing in t (≥ 0) for $i = 1, \ldots, n$.

Remark 1.3.1 The assumption that the search cost c is independent of the searched box is a very strong restriction, but if it is taken off, the above formulation cannot hold and therefore it seems to be more difficult to analyze the problem. On the other hand, it seems that we can replace the exponential detection function with the more general form.

Remark 1.3.2 The following search problem can be modelled in the same mathematical form as the above-mentioned one. Though the object is alive eternally, the reward is assumed to be diminished with time. Namely the reward for detecting the object in box i at time t is given by $r_i \beta_i(t)$ where $\beta_i(t)$ is a discounted rate in box i at time t and is supposed to be nonincreasing in t. For example, if $\beta_i(t) = 1-F_i(t)$, our model is obtained.

In the next theorem we give necessary and sufficient conditions for a policy to be optimal in the Neyman-Pearson type. The necessary condition is proved by using the de Guenin's method (1961) with respect to time instead of space. The sufficient condition is proved by considering the Gateaux differential.

Theorem 1.3.1 Necessary and sufficient conditions for a policy $\phi^*(t) \equiv \{\phi_1^*(t), \ldots, \phi_n^*(t)\}$ to be optimal are given as follows: There is a nonnegative funciton $\mu(t)$ for any t (≥ 0) such that

$$(1.3.5) \quad p_i \lambda_i \int_t^\infty [-g_i'(s)] \exp\{-\lambda_i \phi_i^*(s)\} ds \left\{ \begin{matrix} = \\ \leq \end{matrix} \right\} \mu(t) \quad \text{if } \phi_i^{*'}(t) \left\{ \begin{matrix} > \\ = \end{matrix} \right\} 0.$$

Proof: The proof of the necessity. Suppose that the policy ϕ^* is optimal. We consider any time t_i (≥ 0) such that $\phi_i^{*'}(t_1) > 0$ and define a policy ϕ for any box j ($\neq i$) and any positive constants ε, Δt as follows:

$$\phi_i(t) = \begin{cases} \phi_i^*(t) & \text{if } 0 \leq t \leq t_1 \\ \phi_i^*(t) - \frac{\varepsilon}{\Delta t}(t-t_1) & \text{if } t_1 \leq t \leq t_1+\Delta t \\ \phi_i^*(t) - \varepsilon & \text{if } t_1+\Delta t \leq t < \infty, \end{cases}$$

$$\phi_j(t) = \begin{cases} \phi_j^*(t) & \text{if } 0 \le t \le t_1 \\ \phi_j^*(t) + \frac{\varepsilon}{\Delta t}(t-t_1) & \text{if } t_1 \le t \le t_1+\Delta t \\ \phi_j^*(t) + \varepsilon & \text{if } t_1+\Delta t \le t < \infty \end{cases}$$

and $\phi_k(t) = \phi_k^*(t)$ for any k ($\neq i,j$) and any t (≥ 0). It can be proved that if $\varepsilon = 0\,(\Delta t)$, the policy ϕ satisfies constraints (1.3.3) and (1.3.4) for a sufficiently small Δt. For the policy ϕ,

$$\text{(1.3.6)}\quad R[\phi^*] - R[\phi] = p_i \int_{t_1}^{\infty} g_i'(t)[\exp\{-\lambda_i\phi_i^*(t)\} - \exp\{-\lambda_i\phi_i(t)\}]dt$$
$$+ p_j \int_{t_1}^{\infty} g_j'(t)[\exp\{-\lambda_j\phi_j^*(t)\} - \exp\{-\lambda_j\phi_j(t)\}]dt.$$

By the mean value theorem, we have

$$\text{(1.3.7)}\quad \exp\{-\lambda_i\phi_i^*(t)\} - \exp\{-\lambda_i\phi_i(t)\}$$
$$= \begin{cases} -\frac{\lambda_i\varepsilon}{\Delta t}(t-t_1)\exp[-\lambda_i\{\phi_i^*(t) - \frac{\theta_1\varepsilon}{\Delta t}(t-t_1)\}] & \text{if } t_1 \le t \le t_1+\Delta t \\ -\lambda_i\varepsilon\exp[-\lambda_i\{\phi_i^*(t) - \theta_2\varepsilon\}] & \text{if } t_1+\Delta t \le t < \infty \end{cases}$$

and

$$\text{(1.3.8)}\quad \exp\{-\lambda_j\phi_j^*(t)\} - \exp\{-\lambda_j\phi_j(t)\}$$
$$= \begin{cases} \frac{\lambda_j\varepsilon}{\Delta t}(t-t_1)\exp[-\lambda_j\{\phi_j^*(t) + \frac{\theta_3\varepsilon}{\Delta t}(t-t_1)\}] & \text{if } t_1 \le t \le t_1+\Delta t \\ \lambda_j\varepsilon\exp[-\lambda_j\{\phi_j^*(t) + \theta_4\varepsilon\}] & \text{if } t_1+\Delta t \le t < \infty \end{cases}$$

where $0 \le \theta_i \le 1$ for $i = 1, 2, 3, 4$. Substituting relations (1.3.7) and (1.3.8) into the equation (1.3.6),

$$\text{(1.3.9)}\quad R[\phi^*] - R[\phi]$$
$$= -p_i\frac{\lambda_i\varepsilon}{\Delta t}\int_{t_1}^{t_1+\Delta t} g_i'(t)(t-t_1)\exp[-\lambda_i\{\phi_i^*(t) - \frac{\theta_1\varepsilon}{\Delta t}(t-t_1)\}]dt$$
$$-p_i\lambda_i\varepsilon\int_{t_1+\Delta t}^{\infty} g_i'(t)\exp[-\lambda_i\{\phi_i^*(t) - \theta_2\varepsilon\}]dt$$

$$+ p_j \frac{\lambda_j \varepsilon}{\Delta t} \int_{t_1}^{t_1+\Delta t} g_j'(t)(t-t_1)\exp[-\lambda_j\{\phi_j^*(t) + \frac{\theta_3 \varepsilon}{\Delta t}(t-t_1)\}]dt$$

$$+ p_j \lambda_j \varepsilon \int_{t_1+\Delta t}^{\infty} g_j'(t)\exp[-\lambda_j\{\phi_j^*(t) + \theta_4 \varepsilon]dt \geq 0$$

where the last inequality follows from the optimality of ϕ^*. Dividing both sides of the inequality (1.3.9) by ε (>0) and letting Δt approach to zero, the first and third terms converge to zero since for any function $z(s)$, $\frac{1}{\Delta t}\int_{t_1}^{t_1+\Delta t} z(s)ds$ converges to $z(t_1)$ as $\Delta t \to 0$. Therefore we obtain

$$(1.3.10) \quad p_i\lambda_i \int_{t_1}^{\infty} [-g_i'(t)]\exp\{-\lambda_i\phi_i^*(t)\}dt \geq p_j\lambda_j \int_{t_1}^{\infty} [-g_j'(t)]\exp\{-\lambda_j\phi_j^*(t)\}dt.$$

If we select box j ($\neq i$) such that $\dot{\phi}_j^*(t_1) > 0$, the discussion obtained by exchanging i for j in the above discussion can be developed and therefore the opposite inequality holds in (1.3.10). Hence if $\dot{\phi}_i^*(t_1)$ and $\dot{\phi}_j^*(t_1)$ are positive, the equality holds in (1.3.10). In other words, if $\dot{\phi}_i^*(t) > 0$, the left hand side of (1.3.10) is independent of i. Hence there is a positive function $\mu(t)$ such that

$$(1.3.11) \quad p_i\lambda_i \int_t^{\infty}[-g_i'(s)]\exp\{-\lambda_i\phi_i^*(s)\}ds = \mu(t) \quad \text{if } \dot{\phi}_i^*(t) > 0.$$

On the other hand, if $\dot{\phi}_j^*(t_1) = 0$, the opposite inequality cannot hold in (1.3.10). Therefore

$$(1.3.12) \quad p_j\lambda_j \int_t^{\infty}[-g_j'(s)]\exp\{-\lambda_j\phi_j^*(s)\}ds \leq \mu(t) \quad \text{if } \dot{\phi}_j^*(t) = 0.$$

The relation (1.3.5) is derived from (1.3.11) and (1.3.12).

The proof of the sufficiency. Suppose that the policy ϕ^* satisties the relation (1.3.5) but is not optimal. Therefore there is a policy ϕ such that $R[\phi] > R[\phi^*]$. We put

$$\xi(t) \equiv \phi(t) - \phi^*(t) = \{\phi_1(t) - \phi_1^*(t), \ldots, \phi_n(t) - \phi_n^*(t)\}$$

and consider the Gateaux differential $R'[\phi^*: \xi]$ of the funciton $R[\phi]$ at the point $\phi = \phi^*$ in the direction of ξ which is defined by

$$R'[\phi^*: \xi] \equiv \lim_{\varepsilon\to 0} \varepsilon^{-1}\{R[\phi^* + \varepsilon\xi] - R[\phi^*]\}.$$

By the mean value theorem, we obtain

$$\varepsilon^{-1}\{R[\phi^* + \varepsilon\xi] - R[\phi^*]\}$$

$$= \sum_{i=1}^{n} p_i \int_0^\infty g_i'(t)\varepsilon^{-1}[\exp\{-\lambda_i[\phi_i^*(t) + \varepsilon\xi_i(t)]\} - \exp\{-\lambda_i\phi_i^*(t)\}]dt$$

$$= \sum_{i=1}^{n} p_i \int_0^\infty g_i'(t)\xi_i(t)(-\lambda_i)\exp\{-\lambda_i[\phi_i^*(t) + \theta\varepsilon\xi_i(t)]\}dt$$

for $0 \le \theta \le 1$. Letting ε approach to zero, we have

$$(1.3.13)\quad R'[\phi^*: \xi] = \sum_{i=1}^{n} p_i\lambda_i \int_0^\infty [-g_i'(t)]\{\phi_i(t) - \phi_i^*(t)\}\exp\{-\lambda_i\phi_i^*(t)\}dt$$

$$= \sum_{i=1}^{n} \int_0^\infty \int_0^t p_i\lambda_i[-g_i'(t)]\{\phi_i'(s) - \phi_i^{*'}(s)\}\exp\{-\lambda_i\phi_i^*(t)\}ds\, dt$$

$$= \sum_{i=1}^{n} \int_0^\infty \int_s^\infty p_i\lambda_i[-g_i'(t)]\{\phi_i'(s) - \phi_i^{*'}(s)\}\exp\{-\lambda_i\phi_i^*(t)\}dt\, ds$$

$$= \sum_{i=1}^{n} \int_0^\infty [p_i\lambda_i \int_s^\infty [-g_i'(t)]\exp\{-\lambda_i\phi_i^*(t)\}dt]\{\phi_i'(s) - \phi_i^{*'}(s)\}ds$$

$$\le \sum_{i=1}^{n} \int_0^\infty \mu(s)\{\phi_i'(s) - \phi_i^{*'}(s)\}ds \quad \text{(by the relation (1.3.5))}$$

$$= \int_0^\infty \mu(s)\{\sum_{i=1}^{n} \phi_i'(s) - \sum_{i=1}^{n} \phi_i^{*'}(s)\}ds$$

$$= 0 \qquad \text{(by the condition (1.3.3))}$$

On the other hand, the functional $R[\phi]$ is concave in ϕ since $g_i'(t) \le 0$ for $i = 1, \ldots, n$. Therefore

$$R[\phi^* + \varepsilon\xi] - R[\phi^*] = R[(1-\varepsilon)\phi^* + \varepsilon\phi] - R[\phi^*]$$

$$\ge (1-\varepsilon)R[\phi^*] + \varepsilon R[\phi] - R[\phi^*] = \varepsilon\{R[\phi] - R[\phi^*]\}.$$

Hence

$$(1.3.14)\quad \varepsilon^{-1}\{R[\phi^* + \varepsilon\xi] - R[\phi^*]\} \ge R[\phi] - R[\phi^*] > 0$$

where the last inequality follows from the assumption of the contradiction method. Letting ε approach to zero in (1.3.14), we obtain that $R'[\phi^*: \xi] > 0$ which contradicts to (1.3.13). Therefore the policy ϕ^* is optimal. (q.e.d.)

In the next theorem we give optimal search rates at any time.

Theorem 1.3.2 If all derivatives $f_i'(t)$ $(i = 1, \ldots, n)$ exist at time t, the optimal search rates at time t are given by

$$(1.3.15) \quad \dot{\phi}_i^*(t) = \begin{cases} \lambda_i^{-1}[\frac{g_i''(t)}{g_i'(t)} + (1 - \sum_{j\in I(t)} \lambda_j^{-1} \frac{g_j''(t)}{g_j'(t)})(\sum_{j\in I(t)} \lambda_j^{-1})^{-1}] & \text{if } i \in I(t) \\ 0 & \text{if } i \notin I(t) \end{cases}$$

where

$$(1.3.16) \quad I(t) \equiv \{i | p_i\lambda_i[-g_i'(t)]\exp\{-\lambda_i\phi_i^*(t)\} = -\mu'(t) \text{ and } \frac{g_i''(t)}{g_i'(t)} > \frac{\mu''(t)}{\mu'(t)}\} .$$

Proof: First we put $K(t) \equiv \{i | \dot{\phi}_i^*(t) > 0\}$ and prove that $K(t) = I(t)$. If $i \in K(t)$, the equality holds in (1.3.5) and therefore by differentiating its both sides in t, we have

$$(1.3.17) \quad p_i\lambda_i[-g_i'(t)]\exp\{-\lambda_i\phi_i^*(t)\} = -\mu'(t)$$

or

$$\phi_i^*(t) = \lambda_i^{-1}\{\log(p_i\lambda_i[-g_i'(t)]) - \log[-\mu'(t)]\}.$$

Differentiating in t once more,

$$(1.3.18) \quad \dot{\phi}_i^*(t) = \lambda_i^{-1}[\frac{g_i''(t)}{g_i'(t)} - \frac{\mu''(t)}{\mu'(t)}].$$

Since $\dot{\phi}_i^*(t) > 0$, we obtain that $g_i''(t)/g_i'(t) > \mu''(t)/\mu'(t)$. Therefore $i \in I(t)$, i.e., $K(t) \subseteq I(t)$. On the other hand if there is a box i such that $i \in I(t)$ and $i \notin K(t)$, then the relation (1.3.17) and $\dot{\phi}_i^*(t) = 0$ are satisfied. Hence by the same method as the above discussion we obtain that $g_i''(t)/g_i'(t) = \mu''(t)/\mu'(t)$ which contradicts to $i \in I(t)$. Therefore $K(t) \supseteq I(t)$. The proof of $K(t) = I(t)$ is completed. Next substituting (1.3.18) into $\sum_{i=1}^{n} \dot{\phi}_i^*(t) = 1$, we have

$$1 = \sum_{i\in I(t)} \dot{\phi}_i^*(t) = \sum_{i\in I(t)} \lambda_i^{-1}[\frac{g_i''(t)}{g_i'(t)} - \frac{\mu''(t)}{\mu'(t)}].$$

or

$$(1.3.19)\quad \frac{\mu''(t)}{\mu'(t)} = -\left(1 - \sum_{j\in I(t)} \lambda_j^{-1} \frac{g_j''(t)}{g_j'(t)}\right)\left(\sum_{j\in I(t)} \lambda_j^{-1}\right)^{-1}.$$

Substituting (1.3.19) into (1.3.18), we obtain (1.3.15). (q.e.d.)

Remark 1.3.3 Since the funciton $\mu(t)$ is not known, the optimal search rate $\dot{\phi}_i^*(t)$ is not explicit in Theorem 1.3.2. Later we shall consider the case in which the optimal search rates can be obtained explicitly.

Remark 1.3.4 Let $p_i(t|\phi)$ be the posterior probability that the object is in box i given that it is not detected until time t by a policy ϕ. By the Bayes'rule, we have

$$p_i(t|\phi) = p_i \exp\{-\lambda_i\phi_i(t)\}\left[\sum_{j=1}^{n} p_j \exp\{-\lambda_j\phi_j(t)\}\right]^{-1}.$$

If $i \in I(t)$, then $p_i(t|\phi^*)\lambda_i[-g_i'(t)]$ is independent of i by Theorem 1.3.2, i.e., the optimal search policy is to allocate search time to equalize the value of $p_i(t|\phi^*)\lambda_i[-g_i'(t)]$ for all boxes which are searched at time t. But we must note that the optimal policy is not necessarily to search all boxes maximizing $p_i(t|\phi^*)\lambda_i[-g_i'(t)]$. Later we shall give an example which indicates this notice. Specially in the case that the object is alive eternally, it is well-known that this property holds, i.e., the optimal policy is to search in all boxes maximizing the value of $p_i(t|\phi^*)\lambda_i$ for all t (≥ 0).

Next for the purpose of obtainning the optimal search rates explicitly, we restrict our attention to the case that the lifetime density functions $f_i(t)$ (i=1, ..., n) are differentiable at all $t \in (0, \infty)$. We put

$$h_i(t|\phi) \equiv p_i\lambda_i[-g_i'(t)]\exp\{-\lambda_i\phi_i(t)\} \qquad (i = 1, \ldots, n)$$

$$J(t) \equiv \{i \mid h_i(t|\phi) = \max_{1\le j\le n} h_j(t|\phi)\} \qquad t \in [0, \infty)$$

and define a search policy ϕ by

$$(1.3.20) \quad \phi_i'(t) = \begin{cases} \lambda_i^{-1}\left[\dfrac{g_i''(t)}{g_i'(t)} + \left(1 - \sum_{j\in J(t)} \lambda_j^{-1} \dfrac{g_j''(t)}{g_j'(t)}\right)\left(\sum_{j\in J(t)} \lambda_j^{-1}\right)^{-1}\right] & \text{if } i \in J(t) \\ 0 & \text{if } i \notin J(t). \end{cases}$$

Lemma 1.3.1 Suppose that $f_i(t)$ $(i = 1, \ldots, n)$ are differentiable at all $t \in (0, \infty)$. If the policy ϕ given by (1.3.20) searches box i at time t, then it searches box i at any time in $[t, \infty)$.

Proof: Suppose that though the above policy ϕ searches box i at time t, it does not necessarily search box i at all time $t \in [t, \infty)$. Hence there is a time s $(\ge t)$ such that $i \in J(s)$ and $i \notin J(s+\Delta t)$ for any small Δt (> 0). Consider box j such that $j \in J(s+\Delta t)$. Therefore

$$(1.3.21) \quad h_i(s + \Delta t|\phi) < h_j(s + \Delta t|\phi).$$

If $j \notin J(s)$, then $h_i(s|\phi) > h_j(s|\phi)$ which is contradictory to the continuity of the function $h_*(t|\phi)$ in t. Hence $j \in J(s)$, i.e.,

$$(1.3.22) \quad h_i(s|\phi) = h_j(s|\phi).$$

On the other hand, we obtain

$$(1.3.23) \quad h_i'(s|\phi) = p_i\lambda_i \exp\{-\lambda_i\phi_i(s)\}[-g_i''(s) + g_i'(s)\lambda_i\phi_i'(s)].$$

Substituting (1.3.20) into (1.3.23) [note that $i \in J(s)$], we obtain

$$\begin{aligned}(1.3.24) \quad h_i'(s|\phi) &= h_i(s|\phi) \times [\text{constant in } i] \\ &= h_j(s|\phi) \times [\text{constant in } j] \quad \text{(by the relation (1.3.22))} \\ &= h_j'(s|\phi)\end{aligned}$$

Two equations (1.3.22) and (1.3.24) lead us to the fact that

$$h_i(s + \Delta t|\phi) = h_j(s + \Delta t|\phi)$$

which contradicts to (1.3.21). Therefore by the contradiction method, the proof is completed. (q.e.d.)

Theorem 1.3.3 If $f_i(t)$ $(i = 1, \ldots, n)$ are differentiable at all $t \in (0, \infty)$, the policy defined by (1.3.20) is optimal.

Proof: If $i \in J(t)$, then $i \in J(s)$ for any s $(\geq t)$ by Lemma 1.3.1. Hence $h_i(s|\phi) \geq h_j(s|\phi)$ $(i = 1, \ldots, n)$ for any s $(\geq t)$. Therefore

$$\int_t^{\infty} h_i(s|\phi)ds \geq \int_t^{\infty} h_j(s|\phi)ds \qquad (j = 1, \ldots, n)$$

which denotes that the policy ϕ satisfies the sufficient condition in Theorem 1.3.1. The proof is completed. (q.e.d.)

Remark 1.3.5 Since $h_i(t|\phi)$ depends not on $\phi_i(s)$ $(s > t)$ but on $\phi_i(s)$ $(s \leq t)$, $\phi_i'(t)$ given by (1.3.20) can be determined by only $\phi_i(s)$ $(s \leq t)$. Therefore the optimal policy $\phi(t)$ can be obtained explicitly by a successive method with respect to time.

Remark 1.3.6 If the object is alive eternally or if all rewards r_i $(i = 1, \ldots, n)$ are zero, then $[-g_i'(t)] = c$ for any i, t. Therefore a policy ϕ which searches at time t all boxes maximizing $p_i\lambda_i \exp\{-\lambda_i\phi_i(t)\}$ in proportion to λ_i^{-1} is optimal.

Numerical Example 1.3.1 We consider the 3-box case with the exponential lifetime distribution, i.e., $F_i(t) = 1-\exp(-\theta_i t)$ $(\theta_i > 0)$ for $i = 1, 2, 3$. We give the values of parameters as follows: $\langle p_1, p_2, p_3\rangle = \langle 0.5, 0.3, 0.2\rangle$, $(\lambda_1, \lambda_2, \lambda_3) = (0.6, 0.9, 1.2)$, $(\theta_1, \theta_2, \theta_3) = (0.1, 0.3, 0.5)$, $c = 1$ and $r_i = 2$ $(i = 1, 2, 3)$. We can obtain the optimal search policy ϕ^* by Theorem 1.3.3 as follows:

$$h_i(t|\phi^*) = p_i\lambda_i\{r_i\theta_i \exp(-\theta_i t) + c\}\exp\{-\lambda_i\phi_i^*(t)\}.$$

$$h_i(0|\phi^*) = p_i\lambda_i(r_i\theta_i + c) = \begin{cases} 0.36 & (i = 1) \\ 0.432 & (i = 2) \\ 0.48 & (i = 3) \end{cases}.$$

Since $h_3(0|\phi^*) > h_2(0|\phi^*) > h_1(0|\phi^*)$, ϕ^* must search only box 3 until time t_1 such that $h_3(t_1|\phi^*) = h_2(t_1|\phi^*)$, i.e., t_1 is the unique positive root of the equation

$$0.24\{\exp(-0.5t) + 1\}\exp(-1.2t) = 0.27\{0.6\ \exp(-0.3t) + 1\}.$$

If $t_1 \leq t \leq t_2$, then ϕ^* must search box 2 and 3 with rates given by (1.3.20), i.e.,

$$\dot{\phi}_i^*(t) = \begin{cases} \frac{4}{7} - \frac{3}{7}[3 + 5\ \exp(0.3t)]^{-1} + \frac{5}{21}[1 + \exp(0.5t)]^{-1} & (i = 2) \\ \frac{3}{7} + \frac{3}{7}[3 + 5\ \exp(0.3t)]^{-1} - \frac{5}{21}[1 + \exp(0.5t)]^{-1} & (i = 3) \end{cases}$$

The time t_2 is given by $h_2(t_2|\phi^*) = h_1(t_2|\phi^*)$, i.e., t_2 is the unique positive root of the equation

$$0.27\{0.6\ \exp(-0.3t) + 1\}\ \exp\{-0.9\ \phi_2^*(t)\} = 0.3\{0.2\ \exp(-0.1t) + 1\}$$

where $\phi_2^*(t)$ is given by

$$\phi_2^*(t) = \int_{t_1}^{t} [\frac{4}{7} - \frac{3}{7}\{3+5\ \exp(0.3s)\}^{-1} + \frac{5}{21}\{1+\exp(0.5s)\}^{-1}]ds$$

If $t_2 \leq t \leq \infty$, ϕ^* must search all boxes with rates given by

$$\dot{\phi}_i^*(t) = \begin{cases} \frac{6}{13} - \frac{7}{78}[1+5\exp(0.1t)]^{-1} + \frac{6}{13}[3+5\exp(0.3t)]^{-1} + \frac{15}{78}[1+\exp(0.5t)]^{-1} & (i = 1) \\ \frac{4}{13} + \frac{4}{78}[1+5\exp(0.1t)]^{-1} - \frac{9}{13}[3+5\exp(0.3t)]^{-1} + \frac{10}{78}[1+\exp(0.5t)]^{-1} & (i = 2) \\ \frac{3}{13} + \frac{3}{78}[1+5\exp(0.1t)]^{-1} + \frac{3}{13}[3+5\exp(0.3t)]^{-1} - \frac{25}{78}[1+\exp(0.5t)]^{-1} & (i = 3). \end{cases}$$

Remark 1.3.7 If the object is alive eternally in the above example, it is a well-known resutl that since $p_1\lambda_1 > p_2\lambda_2 > p_3\lambda_3$, the optimal policy begins to search each box in order of box 1, 2, 3. But in the above example, the order must be box 3, 2, 1 since the expected lifetime in each box is given by $(\theta_1^{-1}, \theta_2^{-1}, \theta_3^{-1}) = (10, 10/3, 2)$, i.e., though the efficiency of search is wrong in box 3, box 3 must be searched first since the expected lifetime is very small there.

Numerical Example 1.3.2 Next we consider the case of the nondifferentiable lifetime density function. Theorem 1.3.2 cannot give the optimal policy explicitly and Theorem 1.3.3 cannot hold. The author made efforts to find an algorithm to obtain the optimal policy explicitly, but could not find an available algorithm. It seems that we can follow Theorem 1.3.3 in the general situation but must make a suitable modification in the neighborhood of the nondifferentiable point of $f_i(t)$ $(i = 1, \dots, n)$. To see this point, we give a numerical example. Consider a two-box problem with uniform lifetime distributions

$$F_i(t) = \begin{cases} t/a_i & (0 \leq t \leq a_i) \\ 1 & (a_i < t \leq \infty) \end{cases} \qquad (i = 1, 2).$$

We give the values of parameters as follows: $\langle p_1, p_2 \rangle = \langle 0.6, 0.4 \rangle$, $(\lambda_1, \lambda_2) = (0.7, 0.9)$, $(a_1, a_2) = (6, 3)$, $(r_1, r_2) = (3, 4)$ and $c = 1$. We put $h_i(t|\phi^*) \equiv p_i\lambda_i[-g_i'(t)]\exp\{-\lambda_i\phi_i^*(t)\}$. Theorem 1.3.1 states that it is optimal to search at time t in boxes maximizing an area of a region generated under the function $h_i(s|\phi^*)$ on the interval $[t, \infty)$. Therefore we must construct the optimal policy such that this property is satisfied at any time. After trial and error, we offer the policy ϕ^* given by Table 1.3.1 as the optimal policy (which will be proved to be optimal later). Table 1.3.1 denotes boxes to be searched in each time interval. If ϕ^* searches both boxes, then the proportion of the search in box 1 is $\lambda_1^{-1}/(\lambda_1^{-1} + \lambda_2^{-1}) = 9/16$ by Theorem 1.3.2.

time interval	$(0, t_1)$	(t_1, t_2)	(t_2, t_3)	(t_3, t_4)	(t_4, t_5)	(t_5, ∞)
boxes to be searched	2	1,2	1	1,2	2	1,2

Table 1.3.1 The policy ϕ^*: $t_2 < 3 < t_4$, $t_4 < 6 < t_5$

For this policy ϕ^*, $h_i(t|\phi^*)$ $(i = 1, 2)$ can be calculated as follows:

$$h_1(t|\phi^*) = \begin{cases} 0.63 & \text{if } 0 \le t \le t_1 \\ 0.63\exp[-0.7\times \frac{9}{16}(t-t_1)] & \text{if } t_1 \le t \le t_2 \\ 0.63\exp[-0.7\{\frac{9}{16}(t_2-t_1) + (t-t_2)\}] & \text{if } t_2 \le t \le t_3 \\ 0.63\exp[-0.7\{\frac{9}{16}(t_2-t_1) + (t_3-t_2) + \frac{9}{16}(t-t_3)\}] & \text{if } t_3 \le t \le t_4 \\ 0.63\exp[-0.7\{\frac{9}{16}(t_2-t_1) + (t_3-t_2) + \frac{9}{16}(t_4-t_3)\}] & \text{if } t_4 \le t \le 6 \\ 0.42\exp[-0.7\{\frac{9}{16}(t_2-t_1) + (t_3-t_2) + \frac{9}{16}(t_4-t_3)\}] & \text{if } 6 \le t \le t_5 \\ 0.42\exp[-0.7\{\frac{9}{16}(t_2-t_1) + (t_3-t_2) + \frac{9}{16}(t_4-t_3) + \frac{9}{16}(t-t_5)\}] & \\ & \text{if } t_5 < t < \infty \end{cases}$$

$$h_2(t|\phi^*) = \begin{cases} 0.94\exp[-0.9t] & \text{if } 0 \le t \le t_1 \\ 0.94\exp[-0.9\{t_1+ \frac{7}{16}(t_2-t_1)\}] & \text{if } t_2 \le t \le 3 \\ 0.36\exp[-0.9\{t_1+ \frac{7}{16}(t_2-t_1)\}] & \text{if } 3 \le t \le t_3 \\ 0.36\exp[-0.9\{t_1+ \frac{7}{16}(t_2-t_1) + \frac{7}{16}(t_4-t_3) + (t-t_4)\}] & \\ & \text{if } t_4 \le t \le t_5 \\ h_1(t|\phi^*) & \text{otherwise} \end{cases}$$

which are described in Fig.1.3.1. Here five threshold times t_i $(i = 1, 2, \ldots, 5)$ are determined as follows: The policy ϕ^* starts at time 0 by searching only box 2 and therefore the function $h_2(t|\phi^*)$ decreases in t and at last intersects the function $h_1(t|\phi^*)$ [$= h_1(0|\phi^*) = 0.63$]. Let t_1 be the first time at which $h_2(t|\phi^*) = h_1(t|\phi^*)$, i.e., t_1 is a unique root of the equation

$$0.94\exp(-0.9t) = 0.63.$$

In the time interval (t_1, t_2), since both boxes are searched, $h_1(t|\phi^*) = h_2(t|\phi^*)$. In the interval (t_2, t_3), only box 1 is searched and therefore $h_1(t|\phi^*)$ decreases in t and $h_2(t|\phi^*)$ is constant except that it falls instantaneously at $t = 3$ since $f_2(t)$ is discontinuous at $t = 3$. See Fig.1.3.1. Let t_3 be the first time at which two functions

$h_i(t|\phi^*)$ $(i = 1, 2)$ coincide with each other again. If we fix the time t_2, then t_3 can be determined as a function of t_2 by the root of the equation

$$0.63\exp[-0.7\{\frac{9}{16}(t_2-t_1) + (t_3-t_2)\}] = 0.36\exp[-0.9\{t_1+ \frac{7}{16}(t_2-t_1)\}].$$

We determine t_2(therefore t_3 also) such that two regions A_1 and A_2 in Fig.1.3.1 have the same areas, i.e.,

$$\int_{t_2}^{t_3}[h_1(t|\phi^*) - h_2(t|\phi^*)]dt = 0.$$

Similarly, we determine t_4 and t_5 such that two regions B_1 and B_2 in Fig.1.3.1 have the same areas. Thus we obtain the following values:

$$t_1 = \frac{10}{9}\log\frac{4}{3}\ (\fallingdotseq 0.32)\ , \qquad t_2 = 3 - \frac{5}{7}\log\frac{7}{3}\ (\fallingdotseq 2.395)$$

$$t_3 = 3 + \frac{5}{7}\log\frac{7}{3}\ (\fallingdotseq 3.605)\ , \qquad t_4 = 6 - \frac{10}{9}\log\frac{21}{10}\ (\fallingdotseq 5.176)$$

$$t_5 = 6 + \frac{10}{9}\log\frac{21}{10}\ (\fallingdotseq 6.824).$$

Finally we prove that the above policy ϕ^* is optimal. Let $\mu(t)$ be a function which is defined by

$$-\mu'(t) = \begin{cases} h_1(t|\phi^*) & \text{if } t \in (t_1, t_4), (t_5, \infty) \\ h_2(t|\phi^*) & \text{otherwise.} \end{cases}$$

The function $\mu(t)$ is expressed by the area of the region which is formed below the function $-\mu'(t)$ on the interval $[t, \infty)$. See Fig.1.3.1. Taking the way of constructing functions $h_i(t|\phi^*)$ $(i = 1, 2)$ and $\mu(t)$ into account, it is evident that ϕ^* satisfies the sufficient condition of Theorem 1.3.1. Therefore ϕ^* is optimal.

Remark 1.3.8 Though $\mu(t) = \max_i \int_t^\infty h_i(s|\phi^*)ds$, it is not satisfied that $-\mu'(t) = \max_i h_i(t|\phi^*)$, for example, see the graphs on the intervals $[t_2, 3]$ and $[t_4, 6]$ in Fig.1.3.1. Hence Theorem 1.3.3 does not hold in this case. See Remark 1.3.4.

In the model treated until now, we supposed that the search is continued until the detection of the object because tha family of the alpinist wants to find him though he is deceased. But if a search cost c is positive, it is reasonable in general to stop the search at a certain time because the expected search cost becomes larger than the reward. Hence we face a stopping problem with permitting to stop a search at any time. We modify the above model about the following points:

1. The searcher is allowed to stop a search at any time.
2. The objective is to maximize the expected return until the termination of the search, i.e., detection of the object or stop the search.

A policy in the modified model can be denoted by (ϕ, τ) where ϕ is a search rule and τ is a stopping time. Let $V(\phi, \tau)$ be the expected return by a policy (ϕ, τ) and therefore we obtain

$$V(\phi, \tau) = \sum_{i=1}^{n} p_i \int_0^{\tau} g_i(t) P_i'(t|\phi)dt + \{1 - \sum_{i=1}^{n} p_i \, P_i(\tau|\phi)\} \, (-c\tau).$$

Integrating by parts, we have

$$(1.3.25) \quad V(\phi, \tau) = \sum_{i=1}^{n} p_i r_i (1-\alpha_i) + \sum_{i=1}^{n} p_i \int_0^{\tau} g_i'(t) \exp\{-\lambda_i \phi_i(t)\}dt - \sum_{i=1}^{n} p_i r_i [1-F_i(\tau)] \exp\{-\lambda_i \phi_i(\tau)\}.$$

Theorem 1.3.4 For a fixed stopping time τ, necessary and sufficient conditions for a search rule ϕ^* to be optimal are given as follows: There is a nonnegative function $\mu(t)$ for any t (≥ 0) such that if $\phi_i^{*\prime}(t) \left\{\begin{smallmatrix} > \\ = \end{smallmatrix}\right\} 0$.

$$(1.3.26) \quad p_i \lambda_i \int_t^{\tau} [-g_i'(s)] \exp\{-\lambda_i \phi_i^*(s)\}ds + p_i \lambda_i r_i [1-F_i(\tau)] \exp\{-\lambda_i \phi_i^*(\tau)\} \left\{\begin{smallmatrix} = \\ \leq \end{smallmatrix}\right\} \mu(t).$$

Proof: We can prove this theorem by following the similar proof to that of Theorem 1.3.1. (q.e.d.)

We can prove that Theorem 1.3.2 and Lemma 1.3.1 hold with no modification. But Theorem 1.3.3 cannot hold since maximizing $h_i(t|\phi)$ does not guarantee the sufficient condition in Theorem 1.3.4.

Theorem 1.3.5 If $f_i(t)$ $(i = 1, \ldots, n)$ are differentiable at all $t \in [0, \infty)$, then we have

(i) For a fixed stopping time τ, a search rule ϕ which searches at time t in all boxes maximizing $h_i(t|\phi)$ with rates given by (1.3.20) is optimal if $\phi_i'(\tau) > 0$ for $i = 1, \ldots, n$.

(ii) If the optimal stopping time τ^* for the search rule ϕ given in (i) satisfies that $\phi_i'(\tau^*) > 0$ for all i, then it is a root of the equation

$$\sum_{i=1}^{n} p_i\{\lambda_i r_i[1-F_i(\tau)]\phi_i'(\tau)-c\}\exp\{-\lambda_i\phi_i(\tau)\} = 0. \tag{1.3.27}$$

Proof: (i) If an optimal search rule ϕ^* satisfies that $\phi_i^{*\prime}(\tau) > 0$ for all i, we obtain by Theorem 1.3.4 that

$$p_i\lambda_i r_i[1-F_i(\tau)]\exp\{-\lambda_i\phi_i^*(\tau)\} = \mu(t) \qquad (i = 1, \ldots, n).$$

Therefore the relation (1.3.26) becomes

$$p_i\lambda_i \int_t^\tau [-g_i'(s)]\exp\{-\lambda_i\phi_i^*(s)\}ds \ \{\overset{=}{\leq}\}\ \mu(t)-\mu(\tau) \quad \text{if } \phi_i^{*\prime}(t)\ \{\overset{>}{=}\}\ 0$$

which has the same form as the relation (1.3.5) except that the integral region $\int_t^\infty$ is replaced by $\int_t^\tau$. Hence Lemma 1.3.1 and Theorem 1.3.3 can hold with a slight modification and therefore ϕ is optimal.

(ii) $\frac{d}{d\tau} \nabla(\phi, \tau) = 0$ leads us to the equation (1.3.27). (q.e.d.)

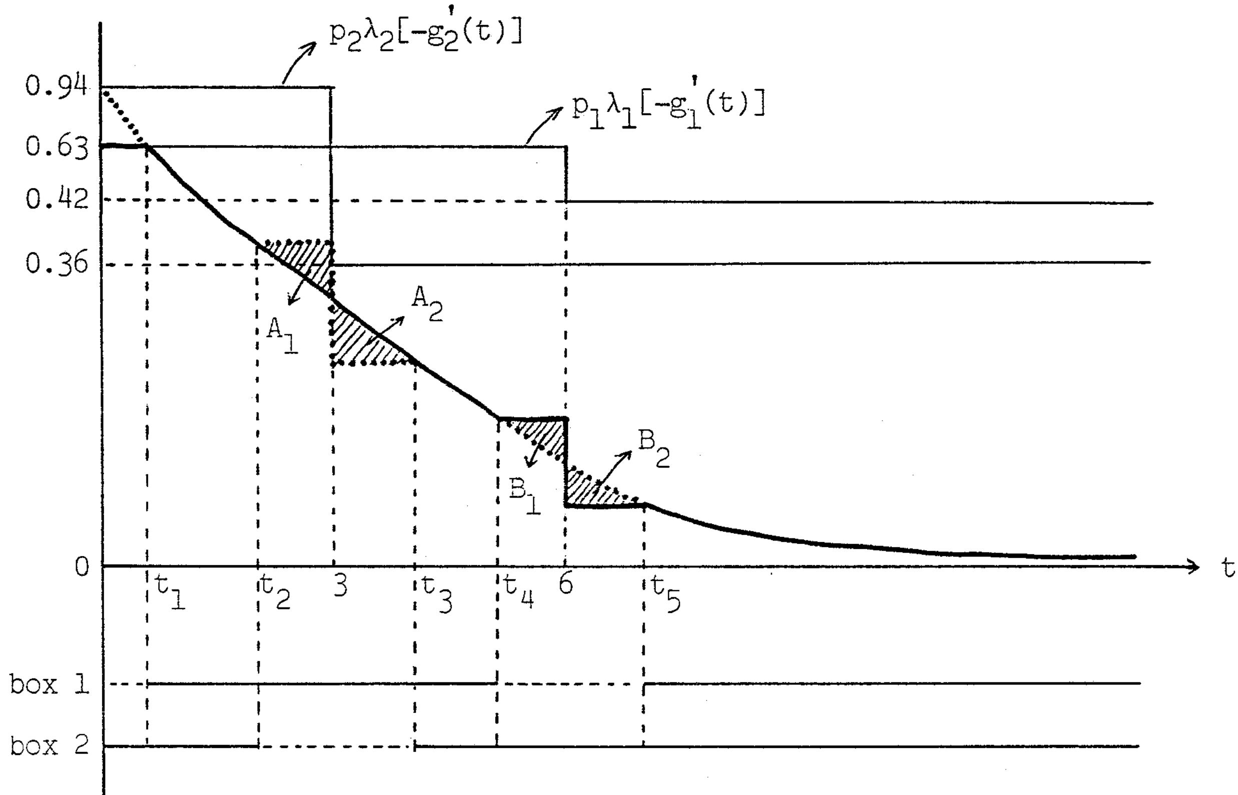

Fig. 1.3.1 ——— $h_1(t|\phi^*)$, ········· $h_2(t|\phi^*)$

§1.4 Discrete Search with Time-Dependent Conditional Detection Probabilities

Various types of allocation problems of search effort to detect a stationary object are analyzed in many literatures, but the conditional detection probability is always constant in time (or period). In this section we consider a fundamental search model with time-dependent conditional detection probabilities and obtain an optimal search policy in some special cases.

There is a stationary object in one of n boxes with a priori distribution $p = \langle p_1, \ldots, p_n \rangle$. Let $\beta_{i,m}$ be a conditional probability that the object is detected by one look in box i at the m-th period given that it is in box i and is not detected until the (m-1)st period. The objective is to find a search policy which maximizes the probability of detecting the object by the M-th period. When we are in the m th period and have the prior location distribution $p = \langle p_1, \ldots, p_n \rangle$ of the object, we express that the search process is in state (p, m). Starting from state (p, m) ($m \leq M$), let $f_m(p)$ be a probability of detecting the object by the M-th period by an optimal policy. By the principle of optimality in dynamic programming, we have

$$(1.4.1) \quad f_m(p) = \max_i [p_i \beta_{i,m} + (1-p_i\beta_{i,m}) f_{m+1}(T_{m,i}p)]$$
$$(m = 1, 2, \ldots, M \; ; \; f_{M+1}(p) \equiv 0)$$

where $T_{m,i}p$ is the posterior location distribution of the object given that the search in box i at the m-th period is unsuccessful. By the Bayes' rule,

$$(1.4.2) \quad (T_{m,i}p)_j = \frac{p_j(1-\delta_{ij}\beta_{i,m})}{1-p_i\beta_{i,m}} \qquad (\delta_{ij}\text{: Kronecker's delta}).$$

In the time-independent case, i.e., the case that $\beta_{i,m} = \beta_i$ for any m, it is well-known that a search policy can be described by only the number of looks in each box and that the myopic policy (which searches at each period in a box maximizing the current detection probability) is optimal.

Kadane (1968) deals with the case that a conditional detection probability in any box does not vary in time so long as the box is not searched. Let $\gamma_{i,k}$ be a conditional probability of detecting the object by the k-th look in box i given that it is in box i and is not detected until the (k-1)st look in box i. When $p_{i,k} \equiv p_i \prod_{\ell=1}^{k-1} (1-\gamma_{i,\ell})\gamma_{i,k}$ is decreasing in k for any i, Kadane proves that there is a number r $(0 < r < \infty)$ such that if $p_{i,k} > (<) r$, the k-th look in box i is (is not) involved in the optimal policy. This result asserts that the myopic policy is optimal. In our model, a search policy depends on the order of looks and the myopic policy is not necessarily optimal. The following example (2-box, 2-period case) points out these facts. Let $\beta_{11} = 2/3$, $\beta_{12} = 7/12$, $\beta_{21} = 1/2$, $\beta_{22} = 1/6$ and $p = \langle 1/2, 1/2 \rangle$. The myopic policy is $\sigma^M = (1, 1)$, namely, searching only box 1. On the other hand, since detection probabilities by four policies (1, 1), (1, 2), (2, 1) and (2, 2) are 31/72, 30/72, 39/72 and 21/72 respectively, the optimal policy is $\sigma^* = (2, 1)$, namely, searching box 2 at the first period and 1 at the second.

A search policy can be denoted by $(\delta_1, \delta_2, \ldots, \delta_M)$ where δ_m indicates a box to be searched at the m-th period. First we obtain an optimal search policy for the two-period case.

Theorem 1.4.1 The solution of the two-period case is given as follows: Suppose that $p_1\beta_{11} \geq p_2\beta_{21} \geq p_i\beta_{i1}$ for any i ($\neq 1, 2$) and that $p_k\beta_{k2} \geq p_\ell\beta_{\ell 2} \geq p_j\beta_{j2}$ for any j ($\neq k, \ell$). If $k \neq 1$, then the policy $\sigma^* = (1, k)$ is optimal and $f_1(p) = p_1\beta_{1,1} + p_k\beta_{k,2}$. If k = 1, then

$$(1.4.3)\quad f_1(p) = \max\begin{cases} p_1\beta_{11} + p_1(1-\beta_{11})\beta_{12} & \sigma^* = (1, 1) \\ p_1\beta_{11} + p_\ell\beta_{\ell 2} & \sigma^* = (1, \ell) \\ p_2\beta_{21} + p_1\beta_{12} & \sigma^* = (2, 1). \end{cases}$$

Proof: Since $f_2(p) = \max_i (p_i\beta_{i2})$,

$$f_1(p) = \max_i \left[p_i\beta_{i1} + (1-p_i\beta_{i1}) \max_j \left\{ \frac{p_j(1-\delta_{ij}\beta_{i1})}{1-p_i\beta_{i1}} \beta_{j2} \right\} \right]$$

$$= \max_i \left[p_i\beta_{i1} + \max \begin{Bmatrix} \max_{j(\neq i)} p_j\beta_{j2} \\ p_i(1-\beta_{i1})\beta_{i2} \end{Bmatrix} \right]$$

$$= \max \begin{cases} \max_{i(\neq k)} (p_i\beta_{i1} + p_k\beta_{k2}) \\ p_k\beta_{k1} + [p_k(1-\beta_{k1})\beta_{k2}] \vee (p_\ell\beta_{\ell 2}) \end{cases}$$

where $a \vee b = \max\{a, b\}$. If $k \neq 1$,

$$f_1(p) = \max \begin{cases} p_1\beta_{11} + p_k\beta_{k2} \\ p_k\beta_{k1} + [p_k(1-\beta_{k1})\beta_{k2}] \vee [p_\ell\beta_{\ell 2}]. \end{cases}$$

Since $p_1\beta_{11} \geq p_k\beta_{k1}$ and $p_k\beta_{k2} \geq [p_k(1-\beta_{k1})\beta_{k2}] \vee (p_\ell\beta_{\ell 2})$, $f_1(p) = p_1\beta_{11} + p_k\beta_{k2}$ and the optimal policy is $\sigma^* = (1, k)$. If $k = 1$,

$$f_1(p) = \max \begin{cases} p_2\beta_{21} + p_1\beta_{12} \\ p_1\beta_{11} + [p_1(1-\beta_{11})\beta_{12}] \vee (p_\ell\beta_{\ell 2}) \end{cases}$$

which leads us to the conclusion of this theorem. (q.e.d.)

Remark 1.4.1 In the above theorem, optimal policies (1, k), (1, 1) and (1, ℓ) are myopic. But the optimal policy $\sigma = (2, 1)$ is not myopic.

Since it is difficult to analyze the n-box, M-period model in general, we shall restrict our attention to the 2-box, M-period case in which a conditional detection probability in box 2 is constant in time. Let $\langle p, 1-p \rangle$ be the prior location distribution of the object. Let α_m (β) be a conditional detection probability in box 1 (box 2) at the m-th period where β is constant in period. See Table 1.4.1.

		m = 1	m = 2	....	m = M
p	box 1	α_1	α_2	....	α_M
$1-p$	box 2	β	β	....	β

Table 1.4.1

The optimality equation is

$$f_m(p) = \max\begin{cases} D_1: & p\alpha_m + (1-p\alpha_m)f_{m+1}\left(\dfrac{p(1-\alpha_m)}{1-p\alpha_m}\right) \\ D_2: & (1-p)\beta + \{1-(1-p)\beta\}f_{m+1}\left(\dfrac{p}{1-(1-p)\beta}\right) \end{cases}$$

$$(m=1,\ 2,\ \ldots,\ M;\ f_{M+1}(p) \equiv 0)$$

where D_i denotes a look in box i. Without loss of generality of the problem we can suppose that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_M$. Let S_k be the set of policies which are composed of k looks in box 1 and (M-k) looks in box 2. Define a policy σ^k by $\sigma^k = (1^k,\ 2^{M-k})$, i.e., σ^k searches only box 1 until the k-th period and only box 2 afterward. Let $f_1(p;\ \delta)$ be a detection probability by using a policy δ.

Lemma 1.4.1 For any p $(0 \leq p \leq 1)$ and any k $(=0,\ 1,\ \ldots,\ M)$, $f_1(p;\ \sigma^k) = \max_{\delta \in S_k} f_1(p;\ \delta)$.

Proof: For the policy σ^k and any policy $\delta = (\delta_1,\ \ldots,\ \delta_M) \in S_k$, we can easily obtain the following expressions in which $\prod_{i \in \phi} \equiv 1$ for a null set ϕ.

(1.4.4) $f_1(p;\ \sigma^k) = p[1-\prod_{i=1}^{k}(1-\alpha_i)] + (1-p)[1-(1-\beta)^{M-k}]$

(1.4.5) $f_1(p;\ \delta) = p[1-\prod_{i \in I}(1-\alpha_i)] + (1-p)[1-(1-\beta)^{M-k}]$

where $I \equiv \{i \mid \delta_i = 1,\ 1 \leq i \leq M,\ \text{integer}\}$. Comparing (1.4.4) with (1.4.5), the conclusion of this lemma is clear by the construction of the policy σ^k.

(q.e.d.)

Let $p^{(k)}$ be the unique root of the equation $f_1(p;\ \sigma^{k-1}) = f_1(p;\ \sigma^k)$, i.e.,

(1.4.6) $$p^{(k)} = \frac{\beta(1-\beta)^{M-k}}{\alpha_k \prod_{i=1}^{k-1}(1-\alpha_i) + \beta(1-\beta)^{M-k}} \qquad (k=1,\ \ldots,\ N)$$

which is nondecreasing in k. We put $p^{(0)} \equiv 0$ and $p^{(M+1)} \equiv 1$.

Theorem 1.4.2 If k is an integer such that $p^{(k)} \leq p \leq p^{(k+1)}$, then the policy $\sigma^k = (1^k, 2^{M-k})$ is optimal and the detection probability by σ^k is given by (1.4.4).

Proof: By Lemma 1.4.1, we obtain

$$f_1(p) = \max_{k=0,1,\ldots,M} f_1(p;\ \sigma^k)$$

which denotes that $f_1(p)$ is the minimum convex function over (M+1) straightlines since $f_1(p;\ \sigma^k)$ is linear in p. Noting that the coefficient of p in $f_1(p;\ \sigma^k)$ is nondecreasing in k, the conclusion of this theorem is clear. (q.e.d.)

Remark 1.4.2 We consider the case that $\alpha_m \equiv \alpha$ for any m. Since $f_1(p;\ \sigma^k) = f_1(p:\ \delta)$ for any p and any $\delta \in S_k$, we obtain that

$$\text{if} \left\{ \begin{array}{l} 0 \leq p \leq p^{(1)} \\ p^{(1)} \leq p \leq p^{(M)} \\ p^{(M)} \leq p \leq 1 \end{array} \right\}, \text{ the decision } \left\{ \begin{array}{l} D_2 \\ D_1 \text{ and } D_2 \\ D_1 \end{array} \right\} \text{ is optimal at the first}$$

period. On the other hand, the myopic policy is to search box 1 (2) if $p \geq (\leq) \beta/(\alpha+\beta)$. Since

$$p^{(1)} = \frac{\beta(1-\beta)^{M-1}}{\alpha+\beta(1-\beta)^{M-1}} \leq \frac{\beta}{\alpha+\beta} \leq \frac{\beta}{\alpha(1-\alpha)^{M-1}+\beta} = p^{(M)},$$

the myopic policy is optimal. This is a well-known result.

Remark 1.4.3 (i) When $\beta = 0$, the policy 1^M (to search only box 1 in all periods) is optimal for any p. (ii) When $\beta = 1$,

$$\text{if} \left\{ \begin{array}{l} 0 \leq p \leq p^{(M)} \\ p^{(M)} \leq p \leq 1 \end{array} \right\}, \text{ the policy } \left\{ \begin{array}{l} (1^{k-1}, 2, 1^{M-k}) \\ 1^M \end{array} \right\} \text{ is optimal}$$

where k is an integer such that $\alpha_k = \min_i \alpha_i$.

Numerical Example 1.4.1 We consider a 2-box, 5-period case where detection probabilities are given by Table 1.4.2. Calculating by (1.4.6), we can obtain that

$$p^{(1)} = 16/78 \doteqdot 0.09 \quad , \qquad p^{(2)} = 16/43 \doteqdot 0.37$$
$$p^{(3)} = 8/11 \doteqdot 0.73 \quad , \qquad p^{(4)} = 8/9 \doteqdot 0.89$$
$$p^{(5)} = 20/21 \doteqdot 0.95$$

		m = 1	m = 2	m = 3	m = 4	m = 5
$\begin{pmatrix} p \\ 1-p \end{pmatrix}$	box 1	1/3	1/5	1/2	2/3	1/4
	box 2	1/3	1/3	1/3	1/3	1/3

Table 1.4.2

By using Theorem 1.4.2, the optimal policy can be obtained. (See Table 1.4.3.)

Condition	Optimal Policy	Condition	Optimal Policy
$0 \le p \le p^{(1)}$	(2, 2, 2, 2, 2)	$p^{(3)} \le p \le p^{(4)}$	(1, 2, 1, 1, 2)
$p^{(1)} \le p \le p^{(2)}$	(2, 2, 2, 1, 2)	$p^{(4)} \le p \le p^{(5)}$	(1, 2, 1, 1, 1)
$p^{(2)} \le p \le p^{(3)}$	(2, 2, 1, 1, 2)	$p^{(5)} \le p \le 1$	(1, 1, 1, 1, 1)

Table 1.4.3

Finally we consider a 2-box, M-period case where the conditional detection probability in box 1 (2) at the m-th period is $\alpha_m(\beta_m)$ which is assumed to be nonincreasing (nondecreasing) in m. We can carry the analysis in this case by the same method as the above discussion and obtain the following theorem.

Theorem 1.4.3 If k is an integer such that $p^{(k)} \le p \le p^{(k+1)}$, then a policy $\sigma^{(k)} \equiv (1^k, 2^{M-k})$ is optimal, where

$$p^{(k)} = \frac{\beta_k\{1-\sum\limits_{i=k+1}^{M}\beta_i\prod\limits_{j=k+1}^{i-1}(1-\beta_j)\}}{\alpha_k\prod\limits_{i=1}^{k-1}(1-\alpha_i)+\beta_k\{1-\sum\limits_{i=k+1}^{M}\beta_i\prod\limits_{j=k+1}^{i-1}(1-\beta_j)\}} \quad \begin{pmatrix} k=1, \ldots, M \\ p^{(0)}=0,\ p^{(M+1)}=1 \end{pmatrix}.$$

Furthermore the detection probability by the optimal policy is given by

$$f_1(p) = p \sum_{i=1}^{k} \alpha_i \prod_{j=1}^{i-1} (1-\alpha_j) + (1-p) \sum_{i=k+1}^{M} \beta_i \prod_{j=k+1}^{i-1} (1-\beta_j).$$

Chapter II Optimal Search for a Moving Tartet

§2.1 Discrete Search for a Circularly Moving Target

In this section we consider a search for a target moving circularly among n boxes according to a Markov transition matrix $Q \equiv [q_{ij}]$ where q_{ij} denotes a probability that the target being in box i at some period moves to box j at next period. We assume that

$$(2.1.1)\quad q_{ij} \equiv \begin{cases} b_i & \text{if } j = i-1 \\ 1 - a_i - b_i & \text{if } j = i \\ a_i & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases} \qquad (i, j = 1, \ldots, n)$$

where subscripts 0 and $n+1$ mean box n and box 1 respectively. Let $p = \langle p_1, \ldots, p_n \rangle$ be a priori distribution of the target at the start of search. Associated with box i ($i=1,\ldots, n$) is a conditional probability β_i that the target is detected by one look in box i given that it is in box i. The objective is to find a policy minimizing the expected number of looks to detect the target.

Let V(p) be the minimum expected number of looks until detection given that the prior distribution is p. By the principle of optimality, the basic equation is given by

$$(2.1.2)\quad V(p) = \min_{i=1,\ldots,n} [1 + (1-p_i\beta_i)V(T_ip)]$$

where T_ip is the posterior location distribution of the target after a look in box i has failed to detect it located with the prior distribution p and after a new movement of the target has occurred. By the Baye's rule, we obtain

$$T_ip = \langle (T_ip)_1, \ldots, (T_ip)_n \rangle$$

$$(2.1.3)\quad (T_ip)_j = \frac{p_{j-1}(1-\delta_{i,j-1}\beta_i)a_{j-1} + p_j(1-\delta_{ij}\beta_i)(1-a_j-b_j) + p_{j+1}(1-\delta_{i,j+1}\beta_i)b_{j+1}}{1 - p_i\beta_i}$$

$i, j = 1, \ldots, n$; δ_{ij} is the Kronecker's delta.

Pollock (1970) considers the two-box case of the above model where a target moves between two boxes according to a Markov transition matrix

$$(2.1.4)\qquad \begin{array}{cc} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \end{array} .$$

His main results are the followings: (i) In the perfect detection case ($\beta_1 = \beta_2 = 1$), there exists a threshold probability p^* such that it is optimal to search box 1 (2) if $p_1 \geq (\leq)\ p^*$ where p^* is given by

$$(2.1.5)\quad p^* = \begin{cases} (a+b+1)^{-1} & \text{if } b(a+b+1) \geq 1 \text{ and } b \geq a(a+b) \\ a(a+b)^{-1} & \text{if } a \leq b(a+b) \quad \text{and } b \leq a(a+b) \\ (b+1)(a+b+2)^{-1} & \text{if } b(a+b+1) \leq 1 \text{ and } a(a+b+1) \leq 1 \\ (a+b)(a+b+1)^{-1} & \text{if } a \geq b(a+b) \quad \text{and } a(a+b+1) \geq 1. \end{cases}$$

(ii) In the null information case ($a+b = 1$), it is optimal to search box 1 (2) if $p \geq (\leq)\ p^0 \equiv \beta_1^{-1}/(\beta_1^{-1} + \beta_2^{-1})$.

If we put $E^i = \langle 0, \ldots, \overset{\overset{i}{\downarrow}}{1}, \ldots, 0\rangle$, then p can be regarded as a point in a simplex S which is defined by a convex hull of the set $\{E^1, \ldots, E^n\}$. The next theorem shows that the posterior distribution $T_i p$ can be regarded as a point of the posterior simplex $T_i S$ which is defined by a convex hull of the set $\{Q^1, \ldots, Q^n\}$ where Q^i is the i-th row of the transition matrix Q.

Theorem 2.1.1 The posterior distribution $T_i p$ can be represented by a linear convex conbination of n points $Q^1, \ldots, Q^n$, i.e.,

$$(2.1.6)\quad T_i p = \sum_{j=1}^{n} \lambda_{ij} Q^{(j)} \quad \text{where } \lambda_{ij} = \frac{p_j(1-\delta_{ij}\beta_i)}{1-p_i\beta_i} .$$

Proof: Let $T_i^0 p$ be the posterior distribution of the target after a look in box i fails to detect it located with the prior distribution p, i.e.,

$$(T_i^0 p)_j = \frac{p_j(1-\delta_{ij}\beta_i)}{1-p_i\beta_i} \qquad (i, j = 1, \ldots, n).$$

Hence by the definition of $T_i p$, $T_i p = (T_i^0 p)\cdot Q = \sum_{j=1}^{n} (T_i^0 p)_j Q^j$ and therefore if we put $\lambda_{ij} = (T_i^0 p)_j$, the relation (2.1.4) is clear. (q.e.d.)

The next corollary is clear without proof.

Corollary 2.1.1 In the perfect detection case (i.e., $\beta_i = 1$ for all i), the posterior simplex $T_i S$ becomes a convex hull of the set $\{Q^1, \ldots, Q^{i-1}, Q^{i+1}, \ldots, Q^n\}$ and the weight λ_{ij} $(j \neq i)$ is in proportion to p_j.

Lemma 2.1.1 The function V(p) is concave in p.

Proof: For any policy σ, let $V_\sigma(p)$ be the expected number of looks to detect the target by the policy σ given that the prior distribution is p. Put $p = \lambda p^1 + (1-\lambda)p^2$ for any $p^1, p^2 \in S$.

$$V(p) = \inf_\sigma V_\sigma(p) = \inf_\sigma\Big[\sum_{i=1}^{n} p_i V_\sigma(E^i)\Big] = \inf_\sigma\Big[\sum_{i=1}^{n}\{\lambda p_i^1 + (1-\lambda)p_i^2\} V_\sigma(E^i)\Big]$$

$$= \inf_\sigma\Big[\lambda\sum_{i=1}^{n} p_i^1 V_\sigma(E^i) + (1-\lambda)\sum_{i=1}^{n} p_i^2 V_\sigma(E^i)\Big] = \inf_\sigma[\lambda V_\sigma(p^1) + (1-\lambda)V_\sigma(p^2)]$$

$$\geq \lambda \inf_\sigma V_\sigma(p^1) + (1-\lambda)\inf_\sigma V_2(p^2) = \lambda V(p^1) + (1-\lambda)V(p^2) \qquad \text{(q.e.d.)}$$

Lemma 2.1.2 In the perfect detection case, the function $V(p; i) \equiv 1 + (1-p_i)V(T_i p)$ is linear in p on any line segment which has the vertex E^i as an end point.

Proof: Let W be a point at which the above line segment intersects the convex hull of the set $\{E^1, \ldots, E^{i-1}, E^{i+1}, \ldots, E^n\}$. Then $W = \langle w_1, \ldots, w_{i-1}, 0, w_{i+1}, \ldots, w_n\rangle$. Any point p on the above line segment can be denoted by $p = \langle kw_1, \ldots, kw_{i-1}, 1-k, kw_{i+1}, \ldots, kw_n\rangle$ $(0 \leq k \leq 1)$. By the relation (2.1.3), we obtain

$$(T_i p)_j = w_{j-1}(1-\delta_{i,j-1})a_{j-1} + w_j(1-\delta_{ij})(1-a_j-b_j) + w_{j+1}(1-\delta_{i,j+1})b_{j+1}$$

which is independent of k. Hence $T_i p$ is constant in p on the above line segment. The proof is completed. (q.e.d.)

Definition: A set A is star-convex with respect to a point $p^0 \in A$ if and only if $p \in A$ implies that $\lambda p + (1-\lambda)p^0 \in A$ $(0 \leq \lambda \leq 1)$.
We define the optimal decision regions D_i^* $(i=1, \ldots, n)$ by $D_i^* = \{p \mid \text{to search box } i \text{ is optimal for } p\}$.

Theorem 2.1.2 In the perfect detection case, the optimal decision regions D_i^* $(i=1, \ldots, n)$ are star-convex with respect to the vertex E^i.

Proof: It is clear that $E^i \in D_i^*$. For any $p \in D_i^*$ and $0 \leq \lambda \leq 1$,

$$\begin{aligned} V[\lambda p+(1-\lambda)E^i] &\leq V[\lambda p+(1-\lambda)E^i;\, i] \\ &= \lambda V(p;\, i) + (1-\lambda)V(E^i;\, i) \qquad \text{(by Lemma 2.1.2)} \\ &= \lambda V(p) + (1-\lambda)V(E^i) \\ &\leq V[\lambda p+(1-\lambda)E^i] \qquad \text{(by Lemma 2.1.1)} \end{aligned}$$

Hence $V[\lambda p+(1-\lambda)E^i] = V[\lambda p+(1-\lambda)E^i;\, i]$ which implies that $\lambda p + (1-\lambda)E^i \in D_i^*$

(q.e.d.)

In the followings, we consider some examples of the three-box case under the assumption of perfect detection. In dealing with the three-box case, it is helpfull to regard $\langle p_1, p_2, p_3 \rangle$ as the barycentric coordinate of the point p and to visualize it in the equilateral triangle of height unity where the distance between the point p and the opposite side of the vertex E^i is p_i. In the next corollary we restate the contents of Corollary 2.1.1 to utilize it in the three-box case.

Corollary 2.1.2 In the three-box case with perfect detection, the posterior distribution $T_i p$ is plotted in the triangular chart as a point which partitions the line segment $\overline{Q^{i-1}\ Q^{i+1}}$ by the ratio $p_{i+1} : p_{i-1}$ where $Q^1 = \langle 1-a_1-b_1,\ a_1,\ b_1 \rangle$, $Q^2 = \langle b_2,\ 1-a_2-b_2,\ a_2 \rangle$ and $Q^3 = \langle a_3,\ b_3,\ 1-a_3-b_3 \rangle$.

Example 2.1.1 We consider the case of $a_i = b_i = a$ $(i = 1, 2, 3)$. In this case the situation is symmetric between any two boxes in all respects and therefore the optimal decision regions are given by $D_i^* = \{p \mid p_i = \max_j p_j\}$ $(i = 1, 2, 3)$. Using Corollary 2.1.2 repeatedly on the triangular chart in which D_i^* $(i = 1, 2, 3)$ and $\Delta Q^1Q^2Q^3$ are described, we can obtain the optimal policy as follows but the detail of the proof is omitted by reason of its complication. If $0 \leq a \leq 1/3$ and $p_i \geq p_j \geq p_k$, then the policy $(i, j, k)^\infty$ is optimal which means to search box i, j, k and to repeat the search in this order periodically until the detection of the target. If $1/3 \leq a \leq 1/2$ and $p \in D_i^*$, then the policy i^∞ is optimal and $V(p) = 1+(1-p_i)/a$ where i^∞ means to search only box i until the detection. This result coincides with our common sense, i.e., if the transition probability is small, we search in the order of the magnitude of the location probability. This optimal policy is the same as in the case of a stationary object. If the transition probability is large, we search only a box having the maximum prior location probability in anticipation of the transition of the target to the box. Note that the above optimal policy is myopic, i.e., it prescribes to search a box having the maximum current location probability at each period.

Example 2.1.2 We consider the case that $a_1 = b_1 = a$ and $a_2 = b_2 = a_3 = b_3 = 0$. Because of the symmetry between box 2 and 3, the boundary between D_2^* and D_3^* is given by $p_2 = p_3$. Hence T_2p (T_3p) cannot be contained in D_2^* (D_3^*) since $Q^1 = \langle 1-2a,\ a,\ a \rangle$, $Q^2 = E^2$ and $Q^3 = E^3$. (i) The case of $1/3 \leq a \leq 1/2$. If $a = 1/2$, then $T_2p \in D_3^*$ and $T_3p \in D_2^*$ since the decision

D_1 is not optimal for $p_1 = 0$. By the continuity of $V(p)$, this property ($T_2p \in D_3^*$ and $T_3p \in D_2^*$ for any p) is seems to be still valid when a is slightly smaller than 1/2. For such a a, by repetitive applications of Corollary 2.1.2, we have

$$V(p;\ 1) = \begin{cases} V[p;\ (1,\ 2,\ 3)] = 3 - 2p_1 - p_2 \\ V[p;\ (1,\ 3,\ 2)] = 3 - 2p_1 - p_3 \end{cases} \quad \text{if } p_2 \begin{Bmatrix} \geq \\ \leq \end{Bmatrix} p_3$$

$$V(p;\ 2) = V[p;\ (2,\ 3)^{\infty}] = 2 + p_1(1-a)/(2a) - p_2 \qquad \text{for any p}$$

$$V(p;\ 3) = V[p;\ (3,\ 2)^{\infty}] = 2 + p_1(1-a)/(2a) - p_3 \qquad \text{for any p.}$$

The equations $V(p;\ 1) = V(p;\ i)$ leads us to the boundary $p_1 = 2a/(1+3a)$ between D_1^* and D_i^* $(i = 2,\ 3)$. To guarantee the above property, the point Q^1 must be below the line $p_1 = 2a/(1+3a)$ and therefore we obtain that $1/3 \leq a$. The solution is given as follows: If $p_2 \geq p_3$ and $p_1 \geq 2a/(1+3a)$, then the policy (1, 2, 3) is optimal and $V(p) = 3 - 2p_1 - p_2$. If $p_2 \geq p_3$ and $p_1 \leq 2a/(1+3a)$, then the policy $(2,\ 3)^{\infty}$ is optimal and $V(p) = 2 + p_1(1-a)/(2a) - p_2$. If $p_2 \leq p_3$, then the solution can be obtained by exchanging p_2 for p_3 in the above solution because of the symmetry between box 2 and 3.

(ii) The case that $a^* \leq a < 1/3$ where $a^* (\fallingdotseq 0.267)$ is a unique root of the equation $12a^3 - 6a^2 - 3a + 1 = 0$ on $[0,\ 1/3]$. We consider the case that the point Q^1 is above the line $p_1 = 2a/(1+3a)$ and that the point T_3T_2p (for any p) is below the line, i.e., $(1-2a)^2/(1-a) \leq 2a/(1+3a)$ or $a^* \leq a$. In this case the region D_2^* is divided into two regions $B \equiv \{p \mid p \in D_2^* \text{ and } T_2p \in D_1^*\}$ and $C \equiv \{p \mid p \in D_2^* \text{ and } T_2p \in D_3^*\}$. We put $A \equiv \{p \mid p \in D_1^* \text{ and } p_2 \geq p_3\}$. By repetitive applications of Corollary 2.1.2, policies (1,2,3), (2,1.3,2) and $(2,3)^{\infty}$ are optimal in regions A, B and C respectively. Furthermore we have

$$V(p) = \begin{cases} 3 - 2p_1 - p_2 & \text{if } p \in A \\ 3 - (1-3a)p_1 - 2p_2 & \text{if } p \in B \\ 2 + p_1(1-a)/(2a) - p_2 & \text{if } p \in C. \end{cases}$$

$$\text{The boundary between} \begin{cases} \text{A and B} \\ \text{B and C} \\ \text{C and A} \end{cases} \text{is given by} \begin{cases} p_2 = (1+3a)p_1 \\ (1-a-6a^2)p_1 = 2ap_3 \\ p_1 = 2a/(1+3a). \end{cases}$$

The solution in the case of $p_2 \leq p_3$ is clear by the symmetry between box 2 and 3.

(iii) The case that $0 \leq a < a^*$ where a^* is given in the case of (ii). Since the point T_3T_2p is above the line $p_1 = 2a/(1+3a)$, $T_3T_2p \quad D_1^*$. Hence the solution is given as follows: We put

$$A \equiv \{p \mid p_2 \leq (1+3a)p_1, \quad 3(1-a)(1+2a)p_1 \geq 1, \quad p_2 \geq p_3\}$$
$$B \equiv \{p \mid p_2 \geq (1+3a)p_1, \quad (1-6a^2)p_1 \geq p_3\}$$
$$C \equiv \{p \mid (1-6a^2)p_1 \leq p_3, \ 3(1-a)(1+2a)p_1 \leq 1, \ p_2 \geq p_3\}.$$

Policies (1,2,3), (2,1,3,2) and (2,3,1,2,3) are optimal in regions A, B and C respectively. Furthermore we have

$$V(p) = \begin{cases} 3 - 2p_1 - p_2 & \text{if } p \in A \\ 3 - (1-3a)p_1 - 2p_2 & \text{if } p \in B \\ 2 + (1+3a-6a^2)p_1 - p_2 & \text{if } p \in C\ . \end{cases}$$

Example 2.1.3 We consider the case that $a_i = 1/2$ and $b_i = 0$ $(i = 1, 2, 3)$. By means of the symmetry of the problem we obtain that $D_i^* = \{p \mid p_i = \max_j p_j\}$. Hence by repetitive applications of Corollary 2.1.2, we can obtain the optimal policy: It is optimal to search first a box having the maximum location probability and to search circularly in the opposite direction of that of the target motion until the detection. Furthermore we consider the case that $a_i = 1$ and $b_i = 0$ $(i = 1, 2, 3)$. The optimal policy can be easily obtained. If $p_i \geq p_{i+1} \geq p_{i+2}$, then the policy (i, i+2, i+1) is optimal and if $p_i \geq p_{i+2} \geq p_{i+1}$, then the policy i^3 is optimal.

§2.2 Efficiency of Wait in an Optimal Search for a Moving Target

In the search for a stationary object, to wait, i.e., to expend time without search is not profitable since the searcher can obtain no information about the target location by waiting. But in the search for a moving target it is sometimes reasonable to wait in anticipation of the transition of the target to the more desirable location for the searcher. In this section we give an example of such a model. Consider a target which moves between two boxes according to a Markov transition matrix

$$(2.2.1) \qquad Q \equiv \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \end{array}.$$

Let $\langle p, 1-p\rangle$ be a priori distribution of the target at the start of search. We suppose the perfect detection, i.e., one look in a box containing the target succeeds necessarily in the detection of the target. At each period the searcher must choose one of three decisions D_i $(i = 1, 2)$ and W where D_i means to search box i and W means to wait without search. Assume that a search cost is unity for both boxes and let w (>0) be a wait cost. The objective is to minimize the expected total cost until the target is detected. If searching box 1 fails to detect the target, the posterior distribution becomes $\langle 0, 1\rangle$ because of the perfect detection assumption. Hence the prior distribution at the next period is $\langle b, 1-b\rangle$ since after the failure the target moves according to the matrix Q. Similarly after the failure of the search in box 2, the prior distribution at the next period is $\langle 1-a, a\rangle$. If the decision W is chosen, the prior distribution at the next period becomes $\langle (1-a-b)p+b, 1-b-(1-a-b)p\rangle$. Let C(p) be the expected total cost until the detection by the optimal policy given that the prior distribution is p. By the principle of optimality, we obtain

$$(2.2.2)\quad C(p) = \min \begin{cases} C(p:\ D_1) = 1 + (1-p)C(b) \\ C(p:\ D_2) = 1 + pC(1-a) \\ C(p:\ W) = w + C[(1-a-b)p + b] \end{cases}$$

where $C(p: \delta)$ ($\delta = D_1, D_2, W$) denotes the expected total cost in the case that the searcher chooses the decision δ first and afterwards follows the optimal policy.

Pollock (1970), which is denoted in detail in Section 2.1, considers the above model under the assumption that at each period the searcher must choose one of two decisions D_i ($i = 1, 2$) and cannot choose the decision W. Our model is an extension of the Pollock's model.

Lemma 2.2.1 The funciton $C(p)$ is concave in p.

Proof: This lemma can be proved by the same method as the proof of Lemma 2.2.1. (q.e.d.)

Define the optimal decision regions by $D_i^* \equiv \{p|D_i \text{ is optimal for } p\}$ ($i = 1, 2$) and $W^* \equiv \{p|W \text{ is optimal for } p\}$. It is clear that $p = 0 \in D_2^*$ and $p = 1 \in D_1^*$.

Lemma 2.2.2 D_i^* ($i = 1, 2$) and W^* are convex sets.

Proof: $C(p: D_1)$ is linearly decreasing in p and $C(1: D_1) = 1$. $C(p: D_2)$ is linearly increasing in p and $C(0: D_2) = 1$. Since $C(p) \geq 1$ for any p, $C(p: W) \geq w + 1 > 1$. Furthermore $C(p: W)$ is concave in p by Lemma 2.2.1. If we describe three curves $C(p: \delta)$ ($\delta = D_1, D_2$ and W) in a plane, the result is clear. (q.e.d.)

Lemma 2.2.3 $p = b/(a+b) \notin W^*$.

Proof: If $b/(a+b) \in W^*$, then $C[b/(a+b)] = \infty$ which is a contradiction and hence the proof is completed. (q.e.d.)

By Lemma 2.2.2, when the set W^* is empty, there is a threshold probability $\hat{p}$ such that the optimal decision is $D_1(D_2)$ if $p \geq (\leq)\hat{p}$. When the set W^* is not empty, there are two threshold probabilities

p^* and p^{**} such that the optimal decision is $\begin{Bmatrix} D_2 \\ W \\ D_1 \end{Bmatrix}$ if $\begin{Bmatrix} 0 \leq p \leq p^{**} \\ p^{**} \leq p \leq p^* \\ p^* \leq p \leq 1 \end{Bmatrix}$.

In the following we consider only the case of $a+b \leq 1$, i.e., $b \leq b/(a+b) \leq 1-a$ since the case of $a+b > 1$ can be analyzed similarly. In this case there are ten possible cases of the optimal decision regions by Lemma 2.2.2 and 2.2.3. But taking the symmetric property into account, it is sufficient to consider only five cases given in Fig.2.2.1. Hereafter we determine the threshold probabilities for each case in Fig.2.2.1 and obtain the conditions under which each case occurs. Put $T_W p \equiv (1-a-b)p+b$.

[A]. In the case [A] of Fig.2.2.1, $b \in D_1^*$ and $1-a \in D_1^*$. Hence $C(b) = 1+(1-b)c(b)$ and $C(1-a) = 1+aC(b)$, i.e., $C(b) = b^{-1}$ and $C(1-a) = 1+ab^{-1}$. From the equation $C(\hat{p}: D_1) = C(\hat{p}: D_2)$, we obtain $\hat{p} = (1+a+b)^{-1}$. The conditions under which this case occurs are $\hat{p} \leq b \leq b(a+b)^{-1} \leq 1-a$ and $C(\hat{p}: D_1) \leq C(\hat{p}: W)$. Hence we obtain $a+b \leq 1$, $b(1+a+b) \geq 1$ and $1-(a+b)/[b(1+a+b)] \leq w$.

[B]. In the case [B] of Fig.2.2.1, $C(b) = b^{-1}$ and $C(1-a) = 1+ab^{-1}$. Since $T_W p^* \in D_1^*$, $C(p^*: W) = C(p^*: WD_1) = w+b[1-(1-a-b)p^*]^{-1}$. From the equation $C(p^*: W) = C(p^*: D_1)$, we obtain $p^* = b(1-w)(a+b)^{-1}$. Similarly since $T_W p^{**} \in D_1^*$, we obtain $p^{**} = 1-b(1-w)$. From conditions that $p^{**} \leq p^* \leq b \leq b/(a+b) \leq 1-a$, we obtain $a+b \leq 1$, $b(1+a+b) \geq 1$ and $1-a-b \leq w \leq 1-(a+b)[b(1+a+b)]^{-1}$.

[C]. In the case [C] of Fig.2.2.1, we assume that

(2.2.3) $T_W^{n-1}b < p^* \leq T_W^n b \quad (n \geq 1: \text{ integer; } T_W^0 b \equiv b)$

where $T_W^n b = b(a+b)^{-1}[1-(1-a-b)^{n+1}]$. Since $b \in W^*$ and $1-a \in D_1^*$,

$$C(b) = C(b: W^n D_1) = (a+b)(1+nw)[b\{1-(1-a-b)^{n+1}\}]^{-1}$$

$$C(1-a) = 1+a(a+b)(1+nw)[b\{1-(1-a-b)^{n+1}\}]^{-1}.$$

From the equation $C(p^*: W) = C(p^*: D_1)$, we obtain

(2.2.4) $p^* = b(a+b)^{-1}-bw\{1-(1-a-b)^{n+1}\}/\ \{(a+b)^2(1+nw)\}.$

By the assumption (2.2.3), either of the following two cases occurs:

(2.2.5) $T_W^{n+k-1}p^{**} < p^* \leq T_W^{n+k}p^{**} \qquad (k = 0, 1)$

where k is determined uniquely if a,b and w are given. Hence

$$C(p^{**}: W) = C(p^{**}: W^{n+k}D_1)$$
$$= (n+k)w+1+C(b)[a(a+b)^{-1}-(1-a-b)^{n+k}p^{**}+b(a+b)^{-1}(1-a-b)^{n+k}].$$

From the equation $C(p^{**}: W) = C(p^{**}: D_2)$, we obtain

(2.2.6) $p^{**} = \dfrac{(n+k)wb[1-(1-a-b)^{n+1}]+(1+nw)[a+b(1-a-b)^{n+k}]}{b[1-(1-a-b)^{n+1}]+(a+b)(1+nw)[a+(1-a-b)^{n+k}]} \quad (k=0,1).$

Substituting (2.2.4) into (2.2.3), we have

(2.2.7) $f(n+1) \leq w < f(n)$

where $f(n) \equiv \dfrac{(a+b)(1-a-b)^n}{1-(1-a-b)^n-(n-1)(a+b)(1-a-b)^n}$.

The relation (2.2.7) is well-defined if $0 < w < 1-a-b$. Since $p^{**} \leq b \leq p^* \leq b(a+b)^{-1} \leq 1-a$, we have $a+b \leq 1$ and $b(1+a+b) \geq 1$.

[D]. In the case [D] of Fig.2.2.1, $C(b) = (1+b)(1-ab)^{-1}$ and $C(1-a) = (1+a)(1-ab)^{-1}$. From the equation $C(\hat{p}: D_1) = C(\hat{p}: D_2)$, we get

$\hat{p} = (1+b)(2+a+b)^{-1}$. Since $b \le \hat{p} \le b(a+b)^{-1} \le 1-a$ and $C(\hat{p}: D_1) \le C(\hat{p}: W)$, we obtain $a+b \le 1$, $b(1+a+b) \le 1$, $a \le b$ and $w \ge (1+b)(b-a)[(1-ab)(2+a+b)]^{-1}$.

[E]. In the case [E] of Fig.2.2.1, $C(b) = (1+b)(1-ab)^{-1}$ and $C(1-a) = (1+a)(1-ab)^{-1}$. Since $T_W p^* \in D_1^*$,

$$C(p^*: W) = w+1+(1+b)[1-(1-a-b)p^*-b](1-ab)^{-1}.$$

From $C(p^*: W) = C(p^*: D_1)$, we obtain

(2.2.8) $p^* = b(a+b)^{-1} - (1-ab)w[(1+b)(a+b)]^{-1}$.

Suppose that

(2.2.9) $T_W^{m-1}p^{**} < p^* \le T_W^m p^{**}$ ($m \ge 1$: integer; $T_W p^{**} \equiv p^{**}$)

where $T_W^m p^{**} = (1-a-b)^m p^{**} + b(a+b)^{-1}[1-(1-a-b)^m]$. Hence

$$C(p^{**}: W) = C(p^{**}: W^m D_1) = mw+1+(1-T_W^m p^{**})C(b)$$

$$= mw+1+(1+b)(1-ab)^{-1}[1-(1-a-b)^m p^{**} - b(a+b)^{-1}\{1-(1-a-b)^m\}].$$

From the equation $C(p^{**}: D_2) = C(p^{**}: W)$, we obtain

(2.2.10) $$p^{**} = \frac{a(1+b)+m(a+b)(1-ab)w+b(1+b)(1-a-b)^m}{(a+b)\{1+a+(1+b)(1-a-b)^m\}}$$

Substituting (2.2.8) and (2.2.10) into (2.2.9), we have

(2.2.11) $g(m) \le w < g(m-1)$

where $$g(m) \equiv \frac{(1+b)(b-a)(1-a-b)^m}{(1-ab)[1+a+(1+b)\{1+m(a+b)\}(1-a-b)^m]}.$$

The relation (2.2.11) is well-defined if $a \le b$ and $0 < w \le (1+b)(b-a)[(1-ab)(2+a+b)]^{-1}$. Since $b \le p^{**} \le p^* \le b(a+b)^{-1} \le 1-a$, we obtain $a+b \le 1$ and $b(1+a+b) \le 1$.

Rearranging the above resutls, we can obtain the solutions in regions A_1 and A_2 in Fig.2.2.2 as follows:

(i) In the case that $(a, b) \in A_1 \equiv \{(a, b) | a+b \le 1, b(1+a+b) \ge 1\}$, the solution is given in Table 2.2.1. In the table, p^* and p^{**} are given by (2.2.4) and (2.2.6) respectively. D_1^∞ denotes an infinite repetition of D_1 until the detection. $p^{(i)}$ is defined by $T_W^i p^{(i)} = p^*$ $(1 \le i \le n+k-1)$ and $p^{(0)} \equiv p^*$, $p^{(n+k)} \equiv p^{**}$. (a), (b) (c) are given as follows:

(a) $1 + p[1+ab^{-1}(a+b)(1+nw)\{1-(1-a-b)^{n+1}\}]$

(b) $iw + 1 + \dfrac{(a+b)(1+nw)}{b[1-(1-a-b)^{n+1}]}\{\dfrac{a}{a+b} + (\dfrac{b}{a+b} - p)(1-a-b)^i\}$

(c) $1 + \dfrac{(a+b)(1+nw)(1-p)}{b[1-(1-a-b)^{n+1}]}$.

Conditions		Optimal Policy	C(p)
$0 < w \le 1 - a - b$	$0 \le p \le p^{**}$	$D_2D_1(W^nD_1)^\infty$	(a)
	$p^{(i)} \le p \le p^{(i-1)}$ $(i=1,2,\ldots,n+k)$	$W^iD_1(W^nD_1)^\infty$	(b)
	$p^* \le p \le 1$	$D_1(W^nD_1)^\infty$	(c)
$1 - a - b \le w \le 1-(a+b)[b(1+a+b)]^{-1}$	$0 \le p \le 1 - b(1-w)$	$D_2D_1^\infty$	$1 + (1+ab^{-1})p$
	$1 - b(1-w) \le p \le b(1-w)(a+b)^{-1}$	WD_1^∞	$w+b^{-1}-p(1-a-b)b^{-1}$
	$b(1-w)(a+b)^{-1} \le p \le 1$	D_1	$1 + (1-p)b^{-1}$
$1-(a+b)[b(1+a+b)]^{-1} \le w$	$0 \le p \le (1+a+b)^{-1}$	$D_2D_1^\infty$	$1 + (1+ab^{-1})p$
	$(1+a+b)^{-1} \le p \le 1$	D_1^∞	$1 + (1-p)b^{-1}$

Table 2.2.1

(ii) In the case that $(a, b) \in A_2 \equiv \{(a, b) | a+b \leq 1,\ b(1+a+b) \leq 1,$ $a \leq b\}$, the solution is given in Table 2.2.2. In the table, p^* and p^{**} are given by (2.2.8) and (2.2.10) respectively. $p^{(i)}$ is defined by

Condition		Optimal Policy	C(p)
$0 < w \leq \frac{(1+b)(b-a)}{(1-ab)(2+a+b)}$	$0 \leq p \leq p^{**}$	$(D_2D_1)^\infty$	$1 + p(1+a)(1-ab)^{-1}$
	$p^{(i)} \leq p \leq p^{(i-1)}$ $(i=1,2,\dots,n)$	$W^i(D_1D_2)^\infty$	(d)
	$p^* \leq p \leq 1$	$(D_1D_2)^\infty$	$1 + (1-p)(1+b)(1-ab)^{-1}$
$\frac{(1+b)(b-a)}{(1-ab)(2+a+b)} \leq w$	$0 \leq p \leq \frac{1+b}{2+a+b}$	$(D_2D_1)^\infty$	$1 + p(1+a)(1-ab)^{-1}$
	$\frac{1+b}{2+a+b} \leq p \leq 1$	$(D_1D_2)^\infty$	$1 + (1-p)(1+b)(1-ab)^{-1}$

Table 2.2.2

$T_W^i p^{(i)} = p^*$ $(1 \leq i \leq n-1)$ and $p^{(0)} = p^*$, $p^{(n)} = p^{**}$. (d) is given by

(d) $\quad iw + 1 + \frac{1+b}{1-ab} \{\frac{a}{a+b} + \frac{b}{a+b} (1-a-b)^n - (1-a-b)^i p\}.$

The optimal decision regions in the case of (i) are described in Fig.2.2.3. In the figure, we observe that the larger w becomes, the narrower W^* becomes and that if $w > 1$, then W^* is empty, i.e., to wait is not optimal for any a, b and p. Hence our resutl in the case of $w > 1$ coincides with the Pollock's result (1970). Furthermore the value of C(p) in our model is not larger than that in the Pollock's model since no waiting is one policy for our model. The difference between the value of C(p) in our model and that in the Pollock's model emphasizes the efficiency of waiting in the search for a moving target. In connection with the matter, it is interesting that the optimal policy does not contain $(W^kD_iD_j)^\infty$ $(k \neq 0,\ i \neq j)$ as its subsequence. In other words, since we choose the decision W in anticipation of the large transition of the target to only one box, it is meaningless to search both boxes successively after waiting. This fact teaches us the meaning of the wait.

For other regions in Fig.2.2.2, we can obtain the solutions by the same method, but their derivation is omitted because of its complication. Since the solution is symmetric about the straight line a = b, the solution in the case of $a \geq b$ can be obtained by exchanging a for b and p for 1-p in the solution of the case of $a \leq b$. Specially if a = b, i.e., the movement of the target is symmetric, then the optimal policy is an infinite repetition of searching box 1 and 2 alternately (See Table 2.2.2) and therefore to wait is not optimal even if w is very small. Furthermore if a+b = 1, i.e., no information case, then the repetitive part of the optimal policy is $\left\{ \begin{matrix} D_1^{\infty} \\ D_2^{\infty} \end{matrix} \right\}$ if a $\{ \begin{matrix} \leq \\ \geq \end{matrix} \}$ 1/2 (See Table 2.2.1).

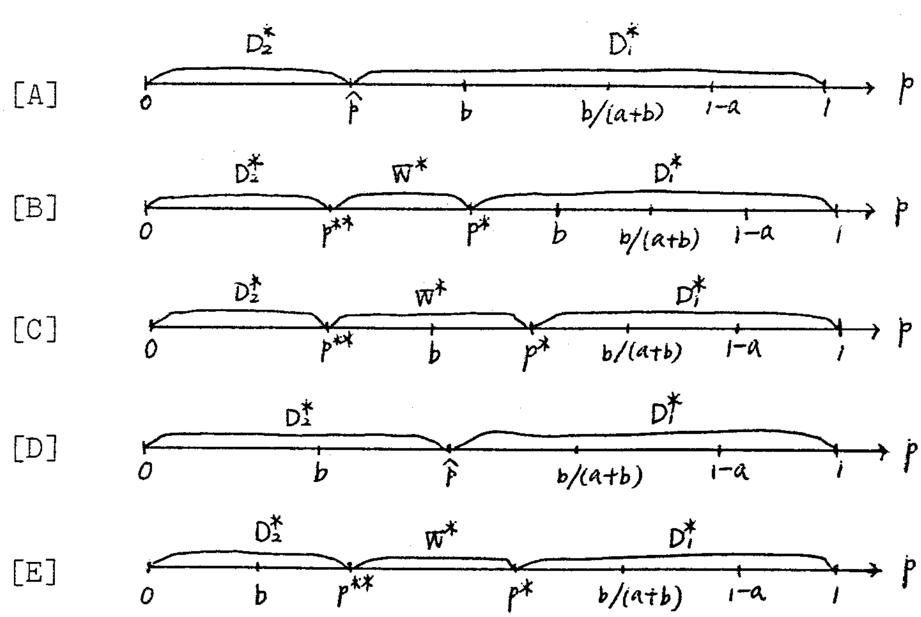

Fig.2.2.1 Five cases of the optimal decision regions when $a+b \leq 1$.

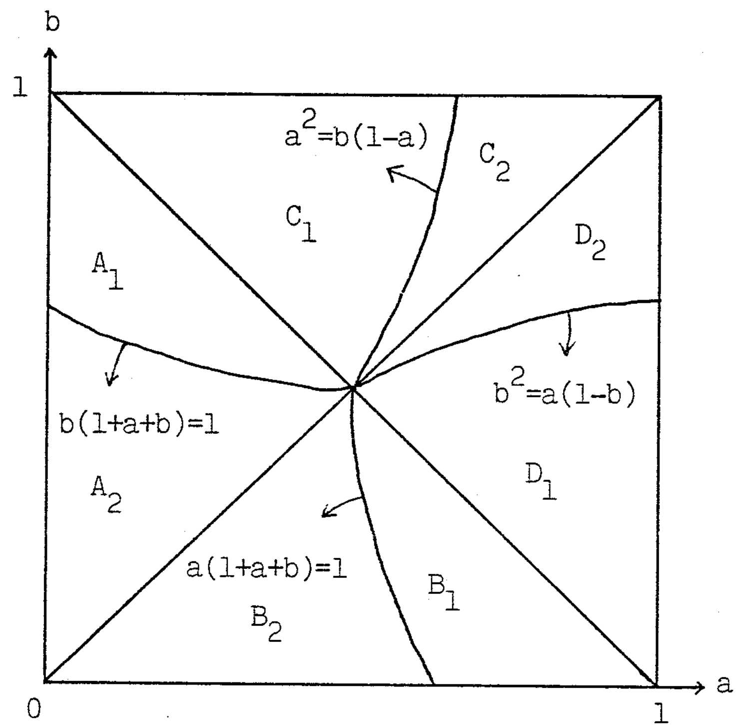

Fig. 2.2.2 The eight regions in the (a, b) plane.

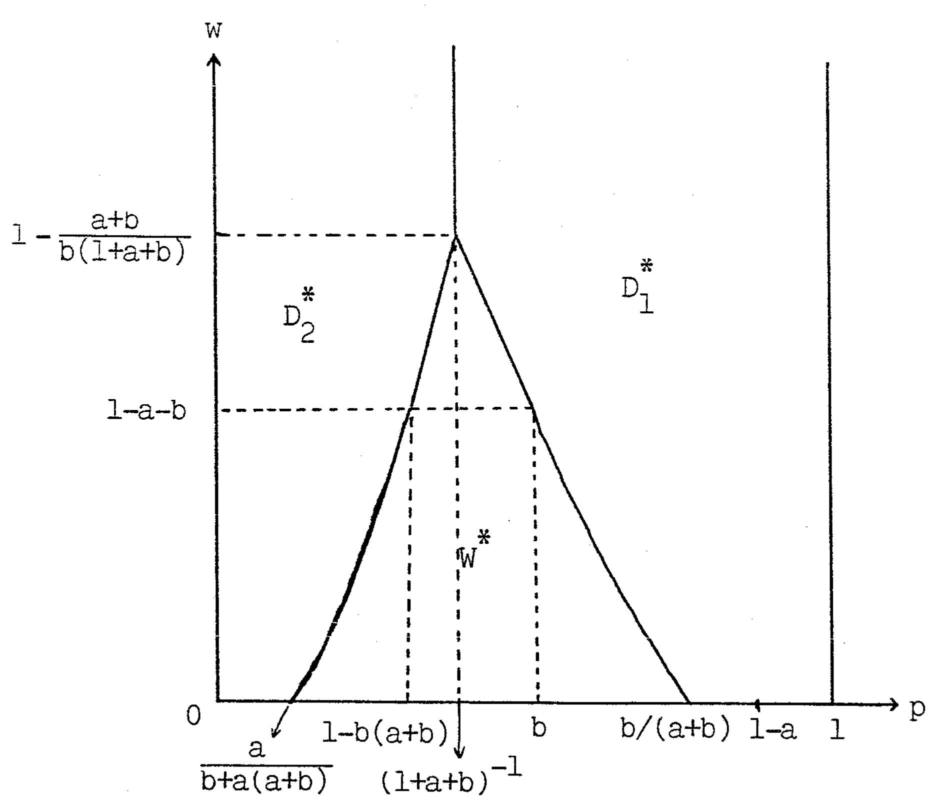

Fig. 2.2.3 The optimal decision regions when $(a, b) \in A_1$.

§2.3 Optimal Wait, Search and Stop for a Passing Target

When we stay in a certain place, it occurs sometimes that we should take an optimal decision to catch a target which comes from another region to our place and goes away to other region. For example, a search for a shoal of fishes which swim in an ocean, and a catching problem for a good chance which occurs and vanishes randomly in time, etc. In this model it seems that we must wait without search at first, search our place if the possibility of its existence in our place becomes high and stop the process if it is seemed to have gone away. In this section we consider a model of such a type.

Pollock (1967) considers a problem of catching as fast as possible a target which appears randomly but does not vanish. One target, which has a priori location distribution $\langle p, 1-p \rangle$ at the beginning of search, moves in two boxes according to a Markov transition matrix

$$\begin{bmatrix} 1-a & a \\ 0 & 1 \end{bmatrix} \qquad (0 < a < 1).$$

The searcher must choose at each period either action: S (to search box 2) and W (to wait). Assume that a search in box 2 is perfect detection. If the target is (is not) in box 2, then the loss of the action S is zero (C_S) and the loss of the action W is C_W (zero). The objective is to minimize the expected loss until the detection. He proves that if n^* is the smallest integer such that $(1-a)^n \leq \{1+a(n+C_S/C_W)\}^{-1}$, then an optimal policy prescibes action W(S) if $p \geq (\leq)\ p^* \equiv 1 - [1-(1-a)^{n^*}]/[a(n^*+C_S/C_W)]$.

Our model is described as follows: There is a moving target which is in one of three boxes with the prior distribution $p = \langle p_1, p_2, p_3 \rangle$ at the start of the process and moves according to a Markov transition matrix

$$Q \equiv \begin{bmatrix} 1-a & a & 0 \\ 0 & 1-b & b \\ 0 & 0 & 1 \end{bmatrix}.$$

At each period the searcher must one of three decisions: W (to wait), D (to search box 2) and S (to stop). Note that box 1 and 3 are unsearchable. The process continues until the detection of the target or the first adoption of decision S. Let w and c be a waiting cost and a searching cost per one period respectively and assume that $0 < w < c < \infty$. Associated with box 2 is a conditional probability β that the target is detected by a search in box 2 given it is in box 2. Let R (> 0) be a reward of the searcher for detecting the target. The problem is to find a policy minimizing the expected total loss (cost minus reward) until the completion of the process, i.e., the detection or the stop. Let f(p) be the minimum expected total loss until the completion given that the prior distribution is p. By the principle of optimality, we have

$$(2.3.1)\quad f(p) = \min \begin{cases} S: & 0 \\ W: & w + f(T_W p) \\ D: & c - p_2\beta R + (1-p_2\beta) f(T_D p) \end{cases}$$

$$(2.3.2)\quad T_W p = \langle (1-a)p_1,\ ap_1+(1-b)p_2,\ bp_2+p_3 \rangle$$

$$(2.3.3)\quad T_D p = \left\langle \frac{(1-a)p_1}{1-p_2\beta},\ \frac{ap_1+(1-b)(1-\beta)p_2}{1-p_2\beta},\ \frac{b(1-\beta)p_2+p_3}{1-p_2\beta} \right\rangle .$$

By the same method as the proof of Lemma 2.1.1, we can prove the following lemma.

Lemma 2.3.1 The function f(p) is concave in p.

Let D^* (or W^*, S^*) be the set of p for which the decision D (or W, S) is optimal. Put $E^1 \equiv \langle 1, 0, 0 \rangle$, $E^2 \equiv \langle 0, 1, 0 \rangle$ and $E^3 \equiv \langle 0, 0, 1 \rangle$.

Lemma 2.3.2 The optimal stopping region S^* is a convex set which contains the point E^3.

Proof: It is clear that $E^3 \in S^*$. Suppose that $f(p^1) = f(p^2) = 0$ and $p = \lambda p^1 + (1-\lambda)p^2$ $(0 \leq \lambda \leq 1)$. By the definition of f, $f(p) \leq 0$. By Lemma 2.3.1, $f(p) \geq f(p^1) + (1-\lambda)f(p^2) = 0$. Hence $f(p) = 0$, i.e., $p \in S^*$. (q.e.d.)

Let Q^i be the i-th row of the transition matrix Q $(i=1, 2, 3)$.

Lemma 2.3.3 For any p, $T_W p$ and $T_D p$ are contained in the triangle $\Delta Q^1 Q^2 Q^3$.

Proof: By definitions of $T_W p$ and $T_D p$, $T_W p = p \cdot Q$ and

$$T_D p = \left\langle \frac{p_1}{1-p_2\beta}, \frac{p_2(1-\beta)}{1-p_2\beta}, \frac{p_3}{1-p_2\beta} \right\rangle \cdot Q$$ which indicate the result of this lemma.

(q.e.d.)

The triangle $\Delta Q^1 Q^2 Q^3$ is called the posterior triangle and is denoted by $T(a, b)$ since it depends on only the transition probabilities a and b. In the next lemma, we give another geometric expression of $T_W p$. For any $P \equiv \langle p_1, p_2, p_3 \rangle$, define $A_1 \equiv \langle 0, 1-p_3, p_3 \rangle$ and $A_2 \equiv \langle p_1, 0, 1-p_2 \rangle$. Let B_1 and B_2 be the inner partition points of the segment $\overline{PA_1}$ and $\overline{PA_2}$ with the rate of a: 1-a and b: 1-b respectively, i.e.,
$B_1 = \langle (1-a)p_1, 1-(1-a)p_1-p_3, p_3 \rangle$ and $B_2 = \langle p_1, (1-b)p_2, 1-p_1-(1-b)p_2 \rangle$.

Lemma 2.3.4 For any p, $T_W p$ can be expressed by the remainning vertex of the parallelogram which has three points P, B_1 and B_2 as its vertices

Proof: The remainning vertex is given by $\langle (1-a)p_1, ap_1+(1-b)p_2, 1-p_1-(1-b)p_2 \rangle$ which coincides with $T_W p$. (q.e.d.)

If $c \geq \beta R$, then the decision S is optimal for any p since the immediate loss of continuing (wait or search) is nonegative for any p. Hence in the following discussion we assume that $c < \beta R$. We define $S_0 \equiv \{p \mid 0 \leq c - p_2\beta R\}$ and $A \equiv \{(a, b) \mid 0 \leq a \leq c/(\beta R)$ and $(\beta R/c-1)\, a \leq b \leq 1\}$.

Lemma 2.3.5 Suppose that $(a, b) \in A$. If $p \in S_0$, then $T_W p \in S_0$ and $T_D p \in S_0$.

Proof: If we put $S_1 \equiv \{p \mid (T_W p)_2 \leq c/(\beta R)\}$ and $S_2 \equiv \{p \mid (T_D p)_2 \leq c/(\beta R)\}$, then in the triangular chart we know that the relation $S_0 \subseteq S_1$ is satisfied if and only if a point $R \equiv \left\langle 1-\frac{c}{\beta R}, \frac{c}{\beta R}, 0\right\rangle$ is contained in S_1, i.e., $(\beta R/c-1)a \leq b$. Similarly the relation $S_0 \subseteq S_2$ is satisfied if and only if $R \in S_1$ and $\langle 1, 0, 0\rangle \in S_1$ and hence we obtain $a \leq c/(\beta R)$. Therefore the result is proved.

(q.e.d.)

Lemma 2.3.6 If $(a, b) \in A$, then $S^* = S_0$.

Proof: Define a sequence of functions $\{f_n(p)\}_{n=0}^{\infty}$ by

$$f_n(p) = \min \begin{cases} S: & 0 \\ W: & w + f_{n-1}(T_W p) \\ D: & c - p_2\beta R + (1-p_2\beta) f_{n-1}(T_D p) \end{cases}$$

and $f_0(p) \equiv 0$. We can prove that if $0 < w < c < \infty$, then the function $f_n(p)$ converges to $f(p)$ as n approaches to infinite. Hence if we prove that $f_n(p) = 0$ for $p \in S_0$, then $f(p) = 0$ for $p \in S_0$. It follows trivially for $n = 0$, so suppose it for $n-1$. By Lemma 2.3.5 we obtain that $f_n(p) = \min[0, w, c-p_2\beta R] = 0$ for $p \in S_0$. Therefore $f(p) = 0$ for $p \in S_0$. On the other hand if $p \notin S_0$, by considering the expected total loss by the policy DS, we obtain $f(p) \leq c-p_2\beta R < 0$. The proof is completed. (q.e.d.)

Let $f(p; \delta)$ $(\delta = W, D, S)$ be the expected total loss by a policy which takes the decision δ first and follows optimally afterward. We consider $p = \langle p_1, p_2, p_3\rangle$ and $p' = \langle p_1', p_2', p_3'\rangle$.

Lemma 2.3.7 If $p_1 \geq p_1'$ and $p_2 \geq p_2'$, then $f(p: \sigma) \leq f(p': \sigma)$ for any policy σ.

Proof: Under the assumption of this lemma, by the definitions (2.3.2) and (2.3.3), it is clear that $(T_W p)_i \geq (T_W p')_i$ and $(T_D p)_i \geq (T_D p')_i$ $(i = 1, 2)$. Hence the assumption of this lemma holds in any period until the termination, i.e., the immediate loss in each period for the initial state p is not larger than that for the initial state p'. The proof is completed.

(q.e.d.)

Theorem 2.3.1 If $(a, b) \in A$, then the optimal decision is D (S) if $p_2 \geq (\leq)\ c/(\beta R)$.

Proof: If $p_2 \geq c/(\beta R)$, either W or D is optimal by Lemma 2.3.6. Moreover if $p_2 \geq c/(\beta R)$, it is clear that $(T_W p)_1 \leq p_1$ and $(T_W p)_2 \leq p_2$ and therefore by Lemma 2.3.7 $f(T_W p: \sigma) \geq f(p: \sigma)$ for any policy σ. Hence $f(p: W, \sigma) = w + f(T_W p: \sigma) \geq w + f(p: \sigma) > f(p: \sigma)$ which denotes that the decision W is not optimal. Therefore if $p_2 \geq c/(\beta R)$, then D is optimal. If $p \leq c/(\beta R)$, then the result is clear by Lemma 2.3.6 (q.e.d.)

Lemma 2.3.8 (i) $(T_W p)_2 \leq p_2$ for any $p \in T(a, b)$ if and only if $a + b \geq 1$. (ii) $(T_D p)_2 \leq p_2$ for any $p \in T(a, b)$ if and only if $a + b \geq 1$.

Proof: (i) If $a + b \geq 1$ and $p \in T(a, b)$, then $ap_1 \leq (1-a)p_2 \leq bp_2$ and therefore $(T_W p)_2 \leq p_2$. Conversely if $(T_W p)_2 \leq p_2$ for any $p \in T(a, b)$, then $\{p|ap_1 = (1-a)p_2\} \subseteq T(a, b) \subseteq \{p|(T_W p)_2 \leq p_2\} = \{p|ap_1 \leq bp_2\}$ and hence $1 - a \leq b$. (ii) If $a + b \geq 1$ and $p \in T(a, b)$, then

$$(T_D p)_2 \leq \frac{(1-a)p_2+(1-b)(1-\beta)p_2}{1-p_2\beta} \leq \frac{\{1-a+(1-b)(1-\beta)\}p_2}{1-a\beta} \leq p_2.$$

Conversely if $(T_D p)_2 \leq p_2$ for any $p \in T(a, b)$, then $(T_D Q')_2 \leq Q_2'$, i.e., $a + b \geq 1$ since $Q' = \langle 1-a, a, 0\rangle \in T(a, b)$. (q.e.d.)

Theorem 2.3.2 When $a+b \geq 1$ and $p \in T(a, b)$, the optimal decision is D (S) if $p_2 \geq (\leq)\ c/(\beta R)$.

Proof: If $p \in T(a, b)$, then $f(T_W p: \sigma) \geq f(p: \sigma)$ for any policy σ by Lemma 2.3.7 and 2.3.8. Therefore $f(p: W, \sigma) = w + f(T_W p: \sigma) \geq w + f(p: \sigma) > f(p: \sigma)$ and hence to wait is not optimal in $T(a, b)$. Moreover since $T_D p \in T(a, b)$ for any p, the optimal policy does not take decision W until the termination. Therefore by excluding decision W, we obtain, for $p \in T(a, b)$,

$$f(p) = \min \begin{cases} S: & 0 \\ D: & c - p_2\beta R + (1-p_2\beta)f(T_D p). \end{cases}$$

Here we consider the OLA policy (one-stage-look-ahead policy) which takes the decision at next period by comparing the expected loss for stopping immediately with that for stopping after one period. The stopping region of the OLA policy is given by $\hat{S} \equiv \{p|0 \leq c - p_2\beta R,\ p \in T(a, b)\}$ which has a property that $p \in \hat{S}$ implies $T_D p \in \hat{S}$, i.e., $\hat{S}$ is closed with respect to the operator T_D. Therefore the OLA policy is optimal by the well-known result and hence the proof is completed. (q.e.d.)

Corollary 2.3.1 If $1-a \leq b < 1$ and $a > c/(\beta R)$, then to stop is not optimal at any p such that $p_2 > c/(\beta R)$.

Proof: The result is clear by Lemma 2.3.2 and Theorem 2.3.2.

(q.e.d.)

We divide the set $T(a, b)$ into a sequence of subsets

$$D_0 \equiv \{p|p \in T(a, b) \text{ and } p_2 \leq c/(\beta R)\}$$

$$D_n \equiv \{p|p \in T(a, b) \text{ and } (T_D^{n-1}p)_2 > c/(\beta R) \geq (T_D^n p)_2\} \quad (n=1, 2, 3, \ldots).$$

By Theorem 2.3.2, a policy $D^n S$ is optimal in region D_n.

Lemma 2.3.9 The expected total loss by the policy $D^n S$ is given by

$$f(p: D^nS) = c - p_2\beta R + \sum_{i=1}^{n-1}\{c-(T_D^i p)_2\beta R\}\prod_{k=0}^{i-1}\{1-(T_D^k p)_2\beta\} \quad (n=1, 2, \ldots).$$

Proof: Because of $f(p: D^nS) = c - p_2\beta R + (1-p_2\beta)f(T_Dp: D^{n-1}S)$,
$f(p: D^{i+1}S) - f(p: D^iS) = \{c - (T_D^i p)_2\beta R\}\prod_{k=0}^{i-1}\{1 - (T_D^k p)_2\beta\}$.

If we take a summation $\sum_{i=1}^{n-1}$ of both sides, the result can be obtained.

(q.e.d.)

Theorem 2.3.3 If $a+b \geq 1$, then the optimal policy can be obtained by the following method. (i) Check which subset D_n contains T_Wp and T_Dp respectively. (ii) Calculate the values of $f(T_Wp)$ and $f(T_Dp)$ by Lemma 2.3.9. (iii) Substitute their values into the basic equation (2.3.1) and solve it.

Proof: The result is clear by Theorem 2.3.2. (q.e.d.)

Numerical Example 2.3.1 We consider the case that $a = 1/2$, $b = 1$, $\beta = 1/2$, $w = 1/2$, $c = 1$ and $R = 10$. By Theorem 2.3.3, we can obtain the optimal policy which is described in Fig.2.3.1. Note that D^* and W^* are not convex.

The case that $b < 1-a$ and $b < (\beta R/c - 1)a$ remains yet unsolved. In this case the optimal policy seems to be more complicated since a is larger and b is smaller relatively. Specially when $\beta = 1$ and $b = p_3 = 0$, this model is similar to the model treated in Pollock (1967). Therefore if $w + c < aR$, then by noting that decision S is not optimal at any p, we can prove that there is a threshold probability γ such that the decision D(W) is optimal if $p_2 \geq (\leq)\ \gamma$.

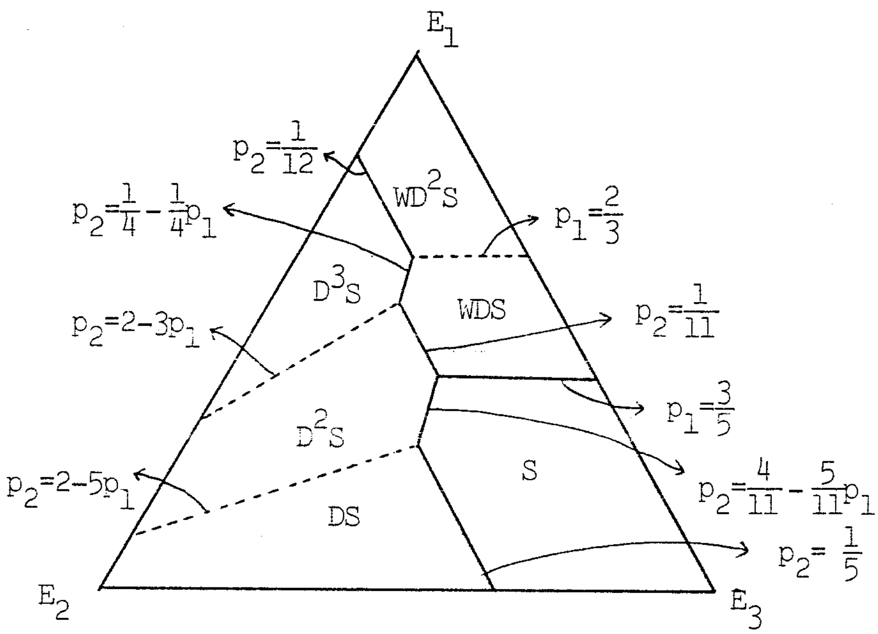

Fig. 2.3.1 The optimal solution of the numerical example.

Chapter III Sequential Evasion-Search Game

In this chapter we consider the following two-person zero-sum m-stage sequential evasion-search game : There are n-boxes (box 1, 2,$\cdots$, n). Player I (Evader) is in box n in the beginning of the game and can move at each period to any box in the "nonincreasing direction", i.e., if he is in box k at some period, he can move to anyone of box 1, 2,$\cdots$, k in next period and cannot go backward. Player II (searcher) can search any box at each period knowing the evader's previous position. Associated with box i(=1,2,$\cdots$,n) is a conditional probability α_i $(0 < \alpha_i < 1)$ that the evader is not detected by a search in box i given it is in box i. To avoid the complication, we assume that $\alpha_i > 0$ (i=1,2,$\cdots$,n), but this limitation can be excluded easily. If the evader is in box i and not detected during one period, he can obtain a reward $r_i(\geq 0)$. Let $R(\geq 0)$ be a reward to the evader when he is not detected during m periods. The payoff is the expected total reward of the evader during m periods.

There are not so many literatures regarding the two-sided search problem and most of them treat a one-stage game or the case that player I (Hider) cannot move during a search process once he hides in some box. Among them, the work of Stewart (1981) is related with our model. There are box 4, 3, 2, 1 and a goal. Player I (Evader) is in box 4 first and must go to goal via box 3, 2, 1 by the m-th period without going backward under the assumption that he can stay in each of box 3 and 1 during a single period. The payoff is a probability that the evader arrives at the goal without being detected. For this two-person zero-sum game, Stewart obtains the optimal search strategy in the "monotone" strategy class. The optimal evade strategy in his model is to run into the goal as fast as possible in order to get out of danger, but in our model the evader must expose himself to danger during m perods against his will and hence he does not go forward so much fast since going forward results in the limitation of the range of his actions. Furthermore Stewart's model is a one-stage allocation game mathematically since the searcher has no information about the evader's position except through learning by search, but essentially our model is a sequential one since in any period the searcher is informed of the evader's previous position.

Washburn (1980) is one of papers which treat a sequential search game. Both players can move freely to any box in each period. Perfect detection is assumed. The evader can know the searcher's previous position. Search cost and travelling cost for the searcher are introduced and an m-period truncated problem is considered. The payoff of this zero-sum game is the expected total cost for the searcher until the search terminates. Washburn discusses about the solution of the limitting problem as m approaches to infinity.

In our model, the state in each period can be described by a pair (k,ℓ) where k is a number of remainning periods and ℓ is the evader's current position. The pure strategies of players I and II in state (m,n) are described by i (i.e., to hide himself in box i : $i=1,\cdots,n$) and j (i.e., to search box j : $j=1,\cdots,n$) respectively and therefore the mixed strategies are $x= \langle x_1,\cdots,x_n \rangle$ (x_i=probability that player I hides in box i) and $y= \langle y_1,\cdots,y_n \rangle$ (y_j=probability that player II searches box j) respectively. Let $G(m,n) \equiv [g(i,j : m,n)]$ $(i,j = 1,\cdots,m)$ be the sequential matrix game starting from state (m,n). The (i,j)-element $g(i,j : m,n)$ is the expected total reward for player I when player I and II choose pure strategies i and j in the first period respectively and play optimally afterwards. Also let $v(m,n)$ be the value of the game $G(m,n)$. Therefore we easily obtain the following relations:

$$\text{(3.0.1)}\quad v(m,n)=\mathrm{val}\begin{bmatrix} \alpha_1\{r_1+v(m-1,1)\} & r_1+v(m-1,1) & \cdots & r_1+v(m-1,1) \\ r_2+v(m-1,2) & \alpha_2\{r_2+v(m-1,2)\} & \cdots & r_2+v(m-1,2) \\ \vdots & \vdots & & \vdots \\ r_n+v(m-1,n) & r_n+v(m-1,n) & \cdots & \alpha_n\{r_n+v(m-1,n)\} \end{bmatrix}$$

$$= \max_x \min_y \left[\sum_{i=1}^{n} \{r_i+v(m-1,i)\}x_i\{1-(1-\alpha_i)y_i\} \right] \quad (m,n=1,2,\cdots)$$

$$\text{(3.0.2)}\quad v(0,n) = R \quad (n=1,2,\cdots).$$

This is a two-dimensional recurrence relation with a boundary condition (1.2). Specially the case of $n=1$ can be easily calculated, i.e.,

$$v(m,1) = \alpha_1 r_1(1-\alpha_1{}^m)(1-\alpha_1)^{-1} + \alpha_1{}^m R \qquad (m=0,1,2,\cdots).$$

Theorem 3.0.1 (i) If the reward R is not smaller (larger) than $v(1,n)$, the value $v(m,n)$ is nonincreasing (nondecreasing) in m for any n.

(ii) The value $v(m,n)$ is nondecreasing in n for any m.

Proof: (i) We prove the assertion in only the case of $R \geq v(1,n)$ by induction in m since the proof in the case of $R \leq v(1,n)$ is similar. First $v(0,n)= R \geq v(1,n)$. Suppose that $v(m-1,n) \geq v(m,n)$ for $m=1,2,\cdots,k$ and any n.

$$v(k,n) = \max_x \min_y \left[\sum_{i=1}^{n} \{ r_i+v(k-1,i)\} x_i\{ 1(1-\alpha_i)y_i\} \right]$$

$$= \max_x \min_y \left[\sum_{i=1}^{n} \{r_i+v(k,i)\} x_i\{ 1-(1-\alpha_i)y_i\} \right] = v(k+1,n).$$

(ii) Consider a $n\times(n+1)$ matrix game $G' = [g'_{ij}]$ defined by

$$g'_{ij} \equiv \begin{cases} g(i,j : m,n) & \text{if } i,j=1,2,\cdots,n \\ r_i+v(m-1,i) & \text{if } i=1,\cdots,n\ ;\ j=n+1. \end{cases}$$

Since the $(n+1)$-st column of G' is dominated by the n-th column, it is clear that val $G'=v(m,n)$. On the other hand, since the game $G(m,n+1)$ is formed by adding the (n+1)-st row to G', player I in the game G(m,n+1) can obtain the expected payoff being not smaller than val G' by taking his optimal strategy for G'. Therefore $v(m,n+1) \geq \text{val } G'$ and hence $v(m,n) \leq v(m,n+1)$.

(q.e.d.)

In order to solve the sequential game G(m,n), we shall slove first a one-stage matrix game G given by

$$(3.0.3)\quad G = [g_{ij}] = \begin{bmatrix} \alpha_1 b_1 & b_1 & \cdots & b_1 \\ b_2 & \alpha_2 b_2 & \cdots & b_2 \\ \vdots & \vdots & & \vdots \\ b_n & b_n & \cdots & \alpha_n b_n \end{bmatrix} \qquad \begin{array}{l} 0 < \alpha_i < 1 \\ 0 < b_i \\ (i = 1,\cdots,n). \end{array}$$

Define a function f on integers $1,2,\cdots,n$ by

$$(3.0.4)\quad f(k) \equiv \{\sum_{i=k}^{n} (1-\alpha_i)^{-1}-1\} \Big/ \sum_{i=k}^{n} \{b_i(1-\alpha_i)\}^{-1},$$

and let ℓ be an integer which attains a maximum of the function f(k), i.e.,

$$(3.0.5)\quad f(\ell) = \max_{k=1,\cdots,n} f(k).$$

Lemma 3.0.1 If we put $b_0 \equiv 0$, then $b_{\ell-1} \leq f(\ell) \leq b_\ell$.

Proof: By a simple calculation, we can prove that

$$(3.0.6)\quad f(k) \{ \substack{\geq \\ =\\ <} \} f(k+1) \text{ if and only if } b_k \{ \substack{\geq \\ = \\ <} \} f(k) \qquad (k=1,2,\cdots,n-1).$$

Since $f(\ell) \geq f(\ell-1)$ and $f(\ell+1)$, we have $b_{\ell-1} \leq f(\ell-1)$ and $b_\ell \geq f(\ell)$ by the relation (3.0.6). Hence $b_{\ell-1} \leq f(\ell) \leq b_\ell$. (q.e.d.)

Theorem 3.0.2 If $b_1 \leq b_2 \leq \cdots \leq b_n$, then the solution of the one-stage game G given by (3.0.3) is given by

$$(3.0.7)\quad x_i^* = \begin{cases} 0 & (i=1,\cdots,\ell-1) \\ [b_i(1-\alpha_i)]^{-1} \Big/ \sum_{j=\ell}^{n} [b_j(1-\alpha_j)]^{-1} & (i=\ell,\cdots,n) \end{cases}$$

$$(3.0.8)\quad y_j^* = \begin{cases} 0 & (j=1,\cdots,\ell-1) \\ [b_j - f(\ell)]/[b_j(1-\alpha_j)] & (i=\ell,\cdots,n) \end{cases}$$

$$(3.0.9)\quad \mathrm{val}\, G = f(\ell).$$

Proof: It is clear that $x_i^* \geq 0$ $(i=1,\cdots,n)$ and $\sum_{i=1}^{n} x_i^* = 1$. From the definition of $f(k)$ and Lemma 3.0.1, we can obtain that $y_j^* \geq 0$ $(j=1,\cdots,n)$ and $\sum_{j=1}^{n} y_j^* = 1$. Let $G(x,j)$ be the expected payoff when players I and II choose a mixed strategy x and a pure strategy j respectively. Also $G(i,y)$ is similar, i.e.,

$$(3.0.10)\quad G(x,j) = \sum_{i=1}^{n} g_{ij}x_i, \quad G(i,y) = \sum_{j=1}^{n} g_{ij}y_j.$$

If the following relation is proved, the proof is completed.

$$(3.0.11)\quad G(i,y^*) \leq f(\ell) \leq G(x^*,j) \quad \text{for } i,j=1,\cdots,n.$$

Substituting (3.0.7) and (3.0.8) into (3.0.10), we have

$$G(x^*,j) = \begin{cases} \sum_{i=\ell}^{n} b_i x_i^* = f(\ell) + \left[\sum_{j=\ell}^{n} \{b_j(1-\alpha_j)\}^{-1}\right]^{-1} > f(\ell) & (j=1,\cdots,\ell-1) \\ \sum_{i=\ell}^{n} b_i x_i^* - (1-\alpha_j) b_j x_j^* = f(\ell) & (j=\ell,\cdots,n) \end{cases}$$

$$G(i,y^*) = \begin{cases} b_i \leq f(\ell) & (i=1,\cdots,\ell-1) \\ b_i - b_i(1-\alpha_i) y_i^* = f(\ell) & (i=\ell,\cdots,n) \end{cases}$$

by Lemma 3.0.1 and the assumption that $b_1 \leq b_2 \leq \cdots \leq b_n$. Thus the relation (3.0.11) is proved. (q.e.d.)

Sakaguchi (1973) considers a one-stage game

$$(3.0.12)\quad M \equiv \begin{bmatrix} c_1-(1-\alpha_1)R_1 & c_2 & \cdots & c_n \\ c_1 & c_2-(1-\alpha_2)R_2 & \cdots & c_n \\ \vdots & \vdots & & \\ c_1 & c_2 & \cdots & c_n-(1-\alpha_n)R_n \end{bmatrix}$$

and obtains a similar result to Theorem 3.0.2. In the following theorem, we state the solution of the sequential game $G(m,n)$ without proof.

Theorem 3.0.3 If we put $b_i=r_i+v(m-1,i)$ in Theorem 3.0.2, the solution of the game G can be regarded as the solution in the first stage of the sequential game $G(m,n)$. The solution of the m-stage game $G(m,n)$ is related with the values of the (m-1)-stage games $G(m-1,i)(i=1,\cdots,n)$ and therefore the solution of $G(m,n)$ can be obtained recurrently by the dynamic programming technique.

Specially we shall consider the case where $r_i=0$ $(i=1,\cdots,n)$ and $R=1$, i.e., the payoff is a probability that the evader is not detected during m periods. The following is an immediate result of Theorem 3.0.1 and 3.0.2.

Corollary 3.0.1 When $r_i=0(i=1,\cdots,n)$ and $R=1$, let ℓ be an integer maximizing a function $f(k)$ given by

$$(3.0.13)\quad f(k) \equiv \left\{\sum_{i=k}^{n} (1-\alpha_i)^{-1}-1\right\} \Big/ \sum_{i=k}^{n} \{(1-\alpha_i)v(m-1,i)\}^{-1}.$$

The value of the sequential game $G(m,n)$ is given by $f(\ell)$ which is nonincreasing in m and nondecreasing in n. Furthermore optimal strategies of players I and II in the first stage of $G(m,n)$ are as follows:

$$(3.0.14)\quad x_i^* = \begin{cases} 0 & (i=1,\cdots,\ell-1) \\ [(1-\alpha_i)v(m-1,i)]^{-1} / \sum_{j=\ell}^{n} [(1-\alpha_j)v(m-1,j)]^{-1} & (i=\ell,\cdots,n) \end{cases}$$

$$(3.0.15)\quad y_j^* = \begin{cases} 0 & (j=1,\cdots,\ell-1) \\ [v(m-1,j)-f(\ell)]/[(1-\alpha_j)v(m-1,j)] & (j=\ell,\cdots,n). \end{cases}$$

By Corollary 3.0.1, the solution of one-stage case can be easily obtained,

$$(3.0.16)\quad x_i^* = y_i^* = (1-\alpha_i)^{-1} / \sum_{j=1}^{n} (1-\alpha_j)^{-1} \qquad (i=1,\cdots,n)$$

$$(3.0.17)\quad v(1, n) = 1 - \{ \sum_{i=1}^{n} (1-\alpha_i)^{-1} \}^{-1}$$

which coincides with the well-known result concerning a one-stage search game. (For example, if we put that $c_i = R_i = 1$ in the game M given by (3.0.12), this case can be obtained.) In the following corollary, we investigate the aspect of the effort allocation by optimal strategies in the multi-stage game as compared with the result (3.0.14) in one-stage case.

Corollary 3.0.2 Suppose thar $r_i = 0$ $(i=1,\cdots,n)$ and R=1.
(i) If α_i is nonincreasing in i, then x_i^* $(i \geq \ell)$ is nonincreasing in i.
(ii) If α_i is nondecreasing in i, then y_i^* $(i \geq \ell)$ is nondecreasing in i and γ_i^* $(\equiv y_i^* - x_i^*)$ $(i \geq \ell)$ is nondecreasing in i.

Proof: (i) Since $v(m-1, i)$ is nondecreasing in i by Corollary 3.0.1, the result is clear by (3.0.14). (ii) Since $y_j^* = (1-\alpha_j)^{-1}[1-f(\ell)/v(m-1,j)]$, y_j^* is nondecreasing in j. Furthermore we obtain

$$\gamma_i^* = (1-\alpha_i)^{-1} \left[1- \sum_{j=\ell}^{n} (1-\alpha_j)^{-1} / \{v(m-1,i) \sum_{j=\ell}^{n} [(1-\alpha_j)v(m-1,j)]^{-1}\} \right]$$

(q.e.d.)

which gives the result.

Numerical Example 3.0.1. We consider the case that $\alpha_i = 1/2$, $r_i = 0$ $(i=1,\cdots,n)$ and R=1. Table 3.0.1 gives optimal strategies in the first period of the games G(m,n) and values v(m,n) (m,n = 1,2,3).

m \ n	1	2	3
1	$x^* = y^* = \langle 1 \rangle$ $v(1,1) = 1/2$	$x^* = y^* = \langle 1/2, 1/2 \rangle$ $v(1,2) = 3/4$	$x^* = y^* = \langle 1/3, 1/3, 1/3 \rangle$ $v(1,3) = 5/6$
2	$x^* = y^* = \langle 1 \rangle$ $v(2,1) = 1/4$	$x^* = \langle 3/5, 2/5 \rangle$ $y^* = \langle 1/5, 4/5 \rangle$ $v(2,2) = 9/20$	$x^* = \langle 0, 10/19, 9/19 \rangle$ $y^* = \langle 0, 8/19, 11/19 \rangle$ $v(2,3) = 45/76$
3	$x^* = y^* = \langle 1 \rangle$ $v(3,1) = 1/8$	$x^* = \langle 9/14, 5/14 \rangle$ $y^* = \langle 1/14, 13/14 \rangle$ $v(3,2) = 27/112$	$x^* = \langle 0, 25/44, 19/44 \rangle$ $y^* = \langle 0, 13/44, 31/44 \rangle$ $v(3,3) = 135/352$

Table 3.0.1 Solutions of the zero-sum sequential games.

Finally, we consider a two person nonzero-sum sequential game which is induced by introducing a reward $Z(\geq 0)$ of the searcher for detecting the evader in the above zero-sum sequential game. The payoff for each player is defined by his expected reward until the termination of the game. As a solution of a nonzero-sum sequential game, there are two types: One is an open-loop solution which choose the entire sequence of strategies for each player at the start of the game. Another is a feedback solution in which both players choose strategies at each period by taking account not of the past history but of the current state. Therefore the feedback solution can be obtained by a dynamic programming technique. It is well-known that the feedback solution is not always better than the open-loop solution in a nonzero-sum sequential game. But since it is difficult to find the open-loop solution, we try to find the feedback Nash solution by the dynamic programming approach. In the sequential game starting from state (m,n), let $M_k(i,j : m,n)$ be the expected payoff for player

$k(=1,2)$ when player I and II use pure strategies i and j respectively in the first period and follow feedback Nash equilibrium strategies afterwards. Let (x^*, y^*) be a feedback Nash equilibrium solution in the first period of this nonzero-sum sequential game. Furthermore let $v_k(m,n)$ be the expected payoff for player k by using the feedback Nash solution.

$$M_1(i,j : m,n) = \begin{cases} \alpha_i[r_i+v_1(m-1,i)] & i=j \\ r_i + v_1(m-1,i) & i\neq j \end{cases}$$

$$M_2(i,j ; m,n) = \begin{cases} (1-\alpha_i)Z + \alpha_i v_2(m-1,i) & i=j \\ v_2(m-1,i) & i\neq j \end{cases}$$

$$M_k(x,y: m,n) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j M_k(i,j : m,n) \qquad (k=1,2)$$

$$v_1(m,n) = \max_x M_1(x,y^* : m,n) , \qquad v_1(0,n) = R$$

$$v_2(m,n) = \max_y M_2(x^*,y : m,n) , \qquad v_2(0,n) = 0.$$

By a simple calculation, we have

$$(3.0.18) \quad M_1(x,y : m,n) = \sum_{i=1}^{n} x_i \{1-(1-\alpha_i)y_i\} \{ r_i+v_1(m-1,i)\}$$

$$(3.0.19) \quad M_2(x,y : m,n) = \sum_{i=1}^{n} x_i y_i (1-\alpha_i) \{Z-v_2(m-1,i)\} + \sum_{i=1}^{n} x_i v_2(m-1,i).$$

We can discuss the problem by the same method as the former model. Define a function g on integers $1,2,\cdots, n$ by

$$(3.0.20) \quad g(k) \equiv \left\{\sum_{i=k}^{n} (1-\alpha_i)^{-1}-1\right\} \Big/ \sum_{i=k}^{n} [(1-\alpha_i)\{ r_i+v_1(m-1,i)\}]^{-1}$$

and let s be an integer maximizing $g(k)$. By the same method as Lemma 3.0.1, we have

$$(3.0.21) \quad r_{s-1}+ v_1(m-1,s-1) \leq g(s) \leq r_s + v_1(m-1,s).$$

Without loss of generality, we can suppose that

$$(3.0.22) \quad r_1 + v_1(m-1,1) \leq r_2 + v_1(m-1,2) \leq \cdots \leq r_n + v_1(m-1,n).$$

Theorem 3.0.4 Under the assumption (3.0.22), a strategy pair (x^*, y^*) given by (3.0.23) and (3.0.24) is a feedback Nash equilibrium solution in the first period of the nonzero-sum sequential game starting from state (m,n).

$$(3.0.23) \quad x_i^* = \begin{cases} 0 & (i=1,\cdots,s-1) \\ [(1-\alpha_i)\{Z-v_2(m-1,i)\}]^{-1} \Big/ \sum_{j=s}^{n} [(1-\alpha_j)\{Z-v_2(m-1,j)\}]^{-1} & (i=s,\cdots,n) \end{cases}$$

$$(3.0.24) \quad y_j^* = \begin{cases} 0 & (j=1,\cdots,s-1) \\ [r_j+v_1(m-1,j)-g(s)]/[(1-\alpha_j)\{r_j+v_1(m-1,j)\}] & (j=s,\cdots,n). \end{cases}$$

Furthermore we obtain

$$(3.0.25) \quad v_1(m,n) = g(s)$$

$$(3.0.26) \quad v_2(m,n) = \left[1 + \sum_{i=s}^{n} \frac{v_2(m-1,i)}{(1-\alpha_i)\{Z-v_2(m-1,i)\}}\right] \Bigg/ \sum_{i=s}^{n} \frac{1}{(1-\alpha_i)\{Z-v_2(m-1,i)\}}.$$

Proof: Substituting (3.0.24) into (3.0.18) we have

$$M_1(x,y^* : m,n) = \sum_{i=1}^{s-1} x_i\{r_i + v_1(m-1,i)\} + g(s) \sum_{i=s}^{n} x_i$$

which is maximized at $x=x^*$ because of (3.0.21) and (3.0.22). Substituting (3.0.23) into (3.0.19), we have

$$M_2(x^*,y : m,n) = \left[\sum_{i=s}^{n} y_i + \sum_{i=s}^{n} \frac{v_2(m-1,i)}{(1-\alpha_i)\{Z-v_2(m-1,i)\}}\right] \Bigg/ \sum_{i=s}^{n} \frac{1}{(1-\alpha_i)\{Z-v_2(m-1,i)\}}$$

which is maximized at $y=y^*$. Hence (x^*, y^*) is a Nash equilibrium solution at the first period. The relations (3.0.25) and (3.0.26) are clear.

(q.e.d.)

Numerical Example 3.0.2 We consider the case that $\alpha_i = 1/2$, $r_i = i-1$ $(i=1,\cdots,n)$ and $R = Z = 4$. Table 3.0.2 gives the feedback Nash solution in the first period of the nonzero-sum sequential game.

m \ n	1	2	3
1	$x^* = y^* = <1>$ $v_1(1,1) = 2$ $v_2(1,1) = 2$	$x^* = <\frac{1}{2}, \frac{1}{2}>$ $y^* = <\frac{1}{3}, \frac{2}{2}>$ $v_1(1,2) = \frac{10}{3}$ $v_2(1,2) = 1$	$x^* = <0, \frac{1}{2}, \frac{1}{2}>$ $y^* = <0, \frac{4}{11}, \frac{7}{11}>$ $v_1(1,3) = \frac{45}{11}$ $v_2(1,3) = 1$
2	$x^* = y^* = <1>$ $v_1(2,1) = 1$ $v_2(2,1) = 3$	$x^* = <0,1>$ $y^* = <0,1>$ $v_1(2,2) = \frac{13}{6}$ $v_2(2,2) = \frac{5}{2}$	$x^* = <0, \frac{1}{2}, \frac{1}{2}>$ $y^* = <0, \frac{85}{344}, \frac{259}{344}>$ $v_1(2,3) = \frac{2613}{688}$ $v_2(2,3) = \frac{7}{4}$
3	$x^* = y^* = <1>$ $v_1(3,1) = \frac{1}{2}$ $v_2(3,1) = \frac{7}{2}$	$x^* = <0,1>$ $y^* = <0,1>$ $v_1(3,2) = \frac{19}{12}$ $v_2(3,2) = \frac{13}{4}$	$x^* = <0, \frac{3}{5}, \frac{2}{5}>$ $y^* = <0, \frac{1105}{18503}, \frac{17398}{18503}>$ $v_1(3,3) = \frac{227373}{74012}$ $v_2(3,3) = \frac{53}{20}$

Table 3.0.2 The feedback Nash solution of the nonzero-sum sequential game.

Acknowledgement

I would like to appreciate Professor Minoru Sakaguchi of Osaka University who conducted me to the area of search theory and gave me very valuable suggestions in warm and repeated discussions.

I wish to thank Assistant Professor T. Kurisu of Osaka University, Assistant Professor Y. Teraoka of Himeji Institute of Technology and Assistant Professor M. Tamaki of Otemon Gakuin University for their valuable discussions.

References

[1] W.H. Barker II, "Information Theory and the Optimal Detection Search", Opns. Res., Vol.25, No.2 (1977), 304-314.

[2] J. deGuenin, "Optimum Distribution of Effort: An Extension of the Koopman Basic Theory", Opns. Res., Vol.9 (1961), 1-7.

[3] J.M. Dobbie, "Search Theory: A Sequential Approach", Nav. Res. Logist. Quart., Vol. 10 (1963), 323-334.

[4] ————, "A Survey of Search Theory", Opns. Res., Vol.16 (1968), 525-537.

[5] ————, "Some Searhc Problems with False Contacts", Opns. Res., Vol.21 (1973), 907-925.

[6] ————, "A Two-Cell Model of Search for a Moving Target", Opns. Res., Vol.22 (1974), 79-92.

[7] P.H. Enslow, Jr., "A Bibliography of Search Theory and Reconnaissance Theory Literature", Nav. Res. Logist. Quart., Vol. 13 (1966), 177-202.

[8] S. Gal, Search Games, Academic Press, (1980).

[9] J.C. Gittins and D.M. Roberts, "The Search for an Intelligent Evader Concealled in One of an Arbitrary Number of Regions", Nav. Res. Logist. Quart., Vol,26, No.4 (1979), 651-666.

[10] B. Gluss, "Approximately Optimal One-Dimensional Search Policies in Which Search Costs Vary Through Time", Nav. Res. Logist. Quart., Vol.8 (1961), 277-283.

[11] G.J. Hall, Jr., "Sequential Search with Random Overlook Probabilities", Ann. Stat., Vol.4, No.4 (1976), 807-816.

[12] J.B. Kadane, "Discrete Search and the Neyman-Pearson Lemma", J. Math. Analysis and App., Vol.22 (1968), 156-171.

[13] ————, "Optimal Whereabouts Search", Opns. Res., Vol.19 (1971), 894-904.

[14] G. Kimeldorf and F.H. Smith, "Binomial Searching for a Random Number of Multinomially Hidden Objects", Manag. Sci., Vol.25, No.11 (1979), 1115-1126.

[15] B.O. Koopman, "The Theory of Search: III The Optimum Distribution of Searching Effort", Opns. Res., Vol.5, No.5 (1957), 613-626.

[16] ————————, Search and Screening, Pergamon Press, (1980).

[17] D. Matula, "A Periodic Optimal Search", Amer. Math. Monthly, Vol.71 (1964), 15-21.

[18] T. Nakai, "A Model for a Target Moving Among Three Boxes: Some Special Cases", J. Opns. Res. Soci. Japan, Vol.16, No.3 (1973), 151-162.

[19] ————, "Two-Box, Two-Object Search", Math. Japonicae, Vol.21, No.5 (1976), 471-479.

[20] ————, "A Search Model for a Moving Target in Which the Decision Wait Is Permitted", Math. Japonica, Vol.25, No.5 (1980), 597-605.

[21] ————, "Search Problem with Time-Dependent Detection Probabilities", Math. Japonica, Vol.26, No.5 (1981), 499-505.

[22] ————, "Optimal Search for an Object with a Random Lifetime", J. Opns. Res. Soci. Japan, Vol.25, No.2 (1982), ~~77-94.~~ 175-192.

[23] ————, "A Sequential Evasion-Search Game". Math. Japonica, in submission.

[24] R.C. Norris, "Studies in Search for a Conscious Evader", M.I.T. Technical Report No.279 (1962).

[25] S.M. Pollock, "Minimum -Cost Checking Using Imperfect Information", Manag. Scie., Vol.13, No.7 (1967), 454-465.

[26] ————————, "A Simple Model of Search for a Moving Target", Opns. Res., Vol.18 (1970), 883-903.

[27] S.M. Ross, "A Problem in Optimal Search and Stop", Opns. Res., Vol.17 (1969), 984-992.

[28] M. Sakaguchi, "An Information Search", Rep. Stat. Appl. Res., JUSE, Vol.20, No.3 (1973), 77-88.

[29] ————, "Two-Sided Search Games", J. Opns. Res. Soci. Japan, Vol.16 (1973), 207-225.

[30] F.H. Smith and G. Kimeldorf, "Discrete Sequential Search for One of Many Objects", Ann. Stat. Vol.3, No.4 (1975), 906-915.

[31] T.J. Stewart, "A Two-Cell Model of Search for an Evading Target", European J. Opnal. Res., Vol.8 (1981), 369-378.

[32] L.D. Stone, Theory of Optimal Search, Academic Press, (1975).

[33] ————, "Necessary and Sufficient Conditions for Optimal Search Plans for Moving Targets", Math. Opns. Res., Vol.4, No.4 (1979), 431-440.

[34] ———— and J.A. Stanshine, "Optimal Search Using Uninterrupted Contact Investigation", SIAM J. Appl. Math., Vol.20, No.2 (1971), 241-263.

[35] E.J. Suberman, "A Hide-Search Game". J. Appl. Prob., Vol.18 (1981), 628-640.

[36] A.R. Washburn, "On a Search for a Moving Target", Nav. Res. Logist. Quart., Vol.27, No.2 (1980), 315-322.