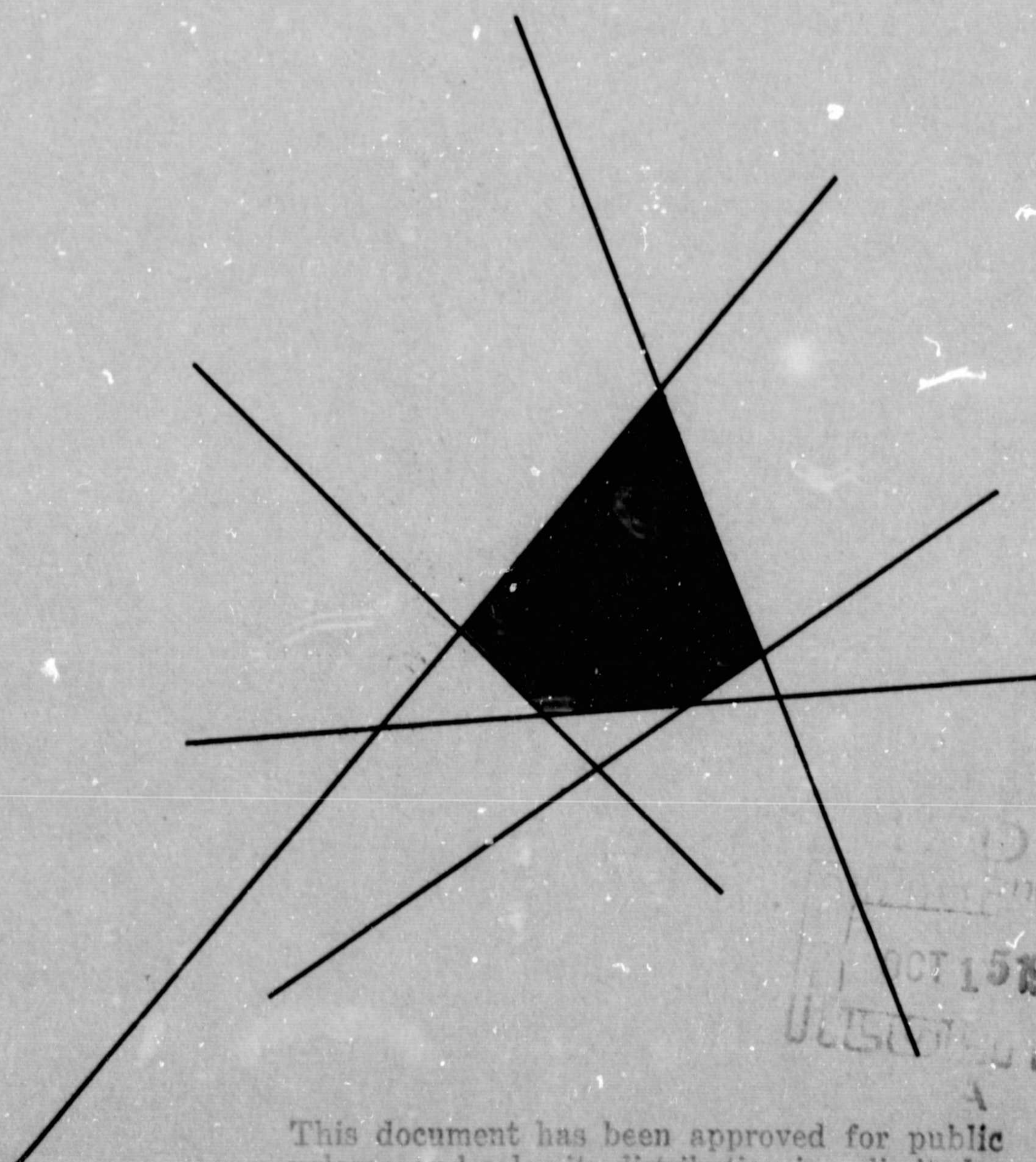
ORC 68-22
JULY 1968

A PROBLEM IN OPTIMAL SEARCH AND STOP

by
Sheldon M. Ross



This document has been approved for public release and sale; its distribution is unlimited. The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

OPERATIONS RESEARCH CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

A PROBLEM IN OPTIMAL SEARCH AND STOP

by

Sheldon M. Ross[†]
Department of Industrial Engineering
and Operations Research
University of California, Berkeley

July 1968 ORC 68-22

This research has been supported by the U.S. Army Research Office-Durham under Contract DA-31-124-ARO-D-331 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[†]Also supported by the Army, Navy, Air Force and NASA under Contract Nonr-225(53)(NR-042-002) with the Office of Naval Research at Stanford University.

ABSTRACT

We are told that an object is hidden in one of $m(m < \infty)$ boxes and we are given prior probabilities p_i^0 that the object is in the i^{th} box. A search of box i costs c_i and finds the object with probability α_i if the object is in the box. Also, we suppose that a reward R_i is earned if the object is found in the i^{th} box. A strategy is any rule for determining when to search and if so which box. The major result is that an optimal strategy either searches a box with maximal value of $\alpha_i p_i / c_i$ or else it never searches those boxes. Also, if rewards are equal, then an optimal strategy either searches a box with maximal $\alpha_i p_i / c_i$ or else it stops.

A PROBLEM IN OPTIMAL SEARCH AND STOP

Sheldon M. Ross
University of California, Berkeley

1. Introduction and Summary

The following model has been considered in the literature: We are told that an object is hidden in one of m boxes and we are given prior probabilities p_i^0 $i=1, 2, \ldots, m$ $(\Sigma p_i^0 = 1)$ that the object is in the i^{th} box. A search of box i costs c_i $(c_i > 0)$, and finds the object with probability α_i if the object is in the box (i.e. $1 - \alpha_i$ is the overlook probability for the i^{th} box). At the beginning of each time period $t = 1, 2, \ldots$ a box is searched; and the process ends when the object is found.

Blackwell (see [5]) has shown that the strategy which at time t searches a box with the largest present value of $\alpha_i p_i / c_i$ minimizes the expected searching cost; (where p_i is the posterior probability at time t that the object is in box i). Chew [3] and Kadane [4] have shown that if $c_i \equiv 1$ then this strategy also maximizes the probability that the searching cost will be less than A for every $A > 0$.

In this paper in order to motivate the search we suppose that a reward R_i $i=1, \ldots, m$ is earned if the object is found in the i^{th} box. We also suppose that the searcher may decide to stop searching at any time (for example he may feel that the rewards are not large enough to justify the searching costs). If the searcher decides to stop before finding the object then from that point on he incurs no further costs and of course receives no reward.

In the second section of this paper we show that an optimal strategy exists and is defined by a functional equation. The optimal strategy is exhibited in a special case. The third section deals with the optimal n-stage return function. The fourth section presents some counterexamples, and in the fifth section we present the major results. Speaking loosely we show that the optimal strategy either searches the box with maximal value of $\alpha_i p_i / c_i$ or else it never searches that box. Also, if rewards are equal, $R_i \equiv R$, then the optimal strategy either searches the box with maximal $\alpha_i p_i / c_i$ or else it stops. In the final section we assume that $R_i \equiv R$ and present a sequence of strategies converging to the optimal.

2. Optimal Strategy

A strategy is any sequence (or partial sequence) $\delta = (\delta_1, \ldots, \delta_s)$ where $\delta_i \in \{1, 2, \ldots, m\}$ for $i=1, \ldots, s$ and $s \in \{0, 1, 2, \ldots\infty\}$. The policy δ instructs the searcher to search box δ_i at the i^{th} period and to stop searching if the object hasn't been found after the s^{th} search. ($s = 0$ means that the searcher stops immediately and $s = \infty$ means that he doesn't stop until he finds the object).

For any strategy δ and any $P = (p_1, \ldots, p_m)$, $p_i \geq 0$, $\Sigma p_i = 1$, let $f(P,\delta)$ be the risk (expected searching cost minus expected reward) incurred when P is the vector of prior probabilities and strategy δ is employed. Also let $f(P) = \inf_{\delta} f(P,\delta)$. Then it follows from standard arguments (see for instance [1] P. 83) that

$$(1) \qquad f(P) = \min\left\{0, \min_{i=1,\ldots,m}\left\{c_i - \alpha_i p_i R_i + (1 - \alpha_i p_i) f(T_i P)\right\}\right\}$$

where $T_i P = ((T_i P)_1, \ldots, (T_i P)_m)$ $\quad i=1, 2, \ldots, m$, and where

$$(2) \qquad (T_i P)_j = \begin{cases} p_j (1 - \alpha_i p_i)^{-1} & j \neq i \\ (1 - \alpha_i) p_i (1 - \alpha_i p_i)^{-1} & j = i \end{cases}$$

Thus $(T_i P)_j$ is just the posterior probability that the object is in box j given that a search of i has not uncovered it. We shall say that the process is in state P at time t if P denotes the posterior probability vector at time t.

In order to show the existence of an optimal strategy let $R = \max_i R_i$ and consider a related process (the prime process) with $c_i' = c_i$, $\alpha_i' = \alpha_i$, but with $R_i' = R_i - R$. However for this new process we suppose that a penalty cost of R units is imposed if the searcher decides to stop searching before finding the object. Now it is easy to see that for any strategy δ which terminates (either by finding the object or by stopping) in finite expected time we have $f(P,\delta) = f'(P,\delta) - R$, and since these are the only strategies we need consider, (any strategy which doesn't terminate in finite expected time has $f(P) = f'(P) = \infty$) it follows that any strategy optimal for the prime process is optimal for the original one.* However, the prime process is a dynamic programming process with a finite number of possible actions available at each stage and with non-positive returns at each stage (since $R_i' \leq 0 \; \forall \; i$). It then follows from Strauch [6] that an optimal strategy exists and also that the optimal strategies may be characterized as those strategies which when the process is in state P chooses one of the actions which minimize the right side of (1), i.e. for such a δ^*, $f(P, \delta^*) = f(P)$ for all P.

The importance of rigorously proving that an optimal policy exists and is determined by a functional equation cannot be overemphasized. For example in the above suppose we relax the condition that $c_i > 0$ and let $c_1 = 0$. Then if $\alpha_1 P_1 > 0$ it is clear that for any strategy $\delta = (\delta_1, \ldots, \delta_s) \neq (1, 1, 1, \ldots)$, $f(P, (1, \delta_1, \ldots, \delta_s)) < f(P, (\delta_1, \ldots, \delta_s))$ (since a search of 1 is free) and thus the only possible optimal strategy would be

---

*The above argument also shows that there is no additional generality gained in assuming that a penalty cost c is incurred when the searcher stops without finding the object, as this process would just be equivalent to the original one with rewards $R_i + c$ instead of R_i.

$\delta_1 = (1, 1, 1, \ldots)$. However $f(P,\delta_1) = p_1R_1$ and it is clear that this need not be maximal. For example if $c_1 = 0$, $\alpha_1 = 1/2$, $p_1 = 1/10$, $R_1 = 10$ and $c_2 = 1$, $\alpha_2 = 1$, $p_2 = 9/10$, $R_2 = 10$ then $f(P, \delta_1) = 1$ while

$$f(P, (1, 1, \ldots, 1, 2, 1, 1, 1, \ldots)) = \frac{1}{10}\left[10(1-(1/2)^n) + 9(1/2)^n\right] + \frac{9}{10}\cdot 9 \uparrow \frac{91}{10}$$

Also the strategy determined by the functional equation turns out to be the (non-optimal) strategy δ_1. (The reason that the existence proof given above breaks down is that since $c_1 = 0$ it no longer follows that all strategies δ with infinite expected termination time have $f(P,\delta) = \infty$).

Now consider the class Λ of strategies $\delta = (\delta_1, \ldots, \delta_s)$ for which $s = \infty$. Any policy $\delta \in \Lambda$ which finds the object with probability 1 will have $f(P,\delta) = E_\delta L - \sum_i p_iR_i$ where L is the searching cost incurred; any $\delta \in \Lambda$ which has positive probability of never finding the object has $f(P,\delta) = \infty$. Thus among the class of policies which never stop searching until the object is found the one with minimal expected searching cost is best. Thus by Blackwell's result the strategy δ_∞ which when in state P searches the box (or one of the boxes) with the maximal value of α_ip_i/c_i is optimal among the policies in Λ.

<u>Lemma 2.1</u>: If $\alpha_ip_iR_i > c_i$ for some i then no optimal strategy stops searching at $P = (p_1, \ldots, p_m)$. If $\alpha_ip_iR_i \geq c_i$ for some i then there is an optimal strategy which doesn't stop at P.

Proof: From (1) we have that

$$f(P) \leq c_i - \alpha_i p_i R_i + (1 - \alpha_i p_i) f(T_i P)$$
$$< 0 + (1 - \alpha_i p_i) f(T_i P)$$
$$\leq 0$$

and so $f(P) < 0$ and thus no optimal policy stops at P. If $\alpha_i p_i R_i \geq c_i$ then $f_i(P) \equiv c_i \ \alpha_i p_i R_i + (1 - \alpha_i p_i) f(T_i P) \leq 0$. Now if $f(P) = 0$ then $f(P) = f_i(P)$ and so searching i is optimal; if $f(P) < 0$ then stopping is not optimal. Q.E.D.

Theorem 2.2: If $\sum_{i=1}^{m} c_i/\alpha_i R_i \leq 1$ then δ_∞ is optimal, i.e. $f(P, \delta_\infty) = f(P)$ for all P.

Proof: For any P, if $\max(\alpha_i p_i R_i - c_i) \geq 0$ then there exists an optimal strategy which doesn't stop at P. So a necessary condition for every optimal strategy to stop at P is for

$$\alpha_i p_i R_i < c_i \qquad \text{for all } i$$
$$\Rightarrow \quad p_i < c_i/\alpha_i R_i \qquad \text{for all } i$$
$$\Rightarrow \quad 1 < \Sigma c_i/\alpha_i R_i$$

So if $\Sigma c_i/\alpha_i R_i \leq 1$ then for every P there is an optimal strategy which doesn't stop at P. Thus an optimal strategy exists in Λ which implies that δ_∞ is optimal. Q.E.D.

3. The Optimal Return $f(P)$

Theorem 3.1: $f(P)$ is a concave function of P.

Proof: Let $f_i(\delta)$ be the conditional risk given that the object is in i and strategy δ is employed, $i=1, \ldots, m$. Then $f(P,\delta) = \sum_i p_i f_i(\delta)$. Now let $P = \lambda P^1 + (1-\lambda)P^2$, then

$$\begin{aligned} f(P) &= \inf_\delta f(P,\delta) \\ &= \inf_\delta f(\lambda P^1 + (1-\lambda)P^2, \delta) \\ &= \inf_\delta \sum_i (\lambda P^1 + (1-\lambda)P^2)_i f_i(\delta) \\ &\geq \lambda \inf_\delta \sum_i P_i^1 f_i(\delta) + (1-\lambda) \inf_\delta \sum_i P_i^2 f_i(\delta) \\ &= \lambda f(P^1) + (1-\lambda) f(P^2) \end{aligned}$$

Q.E.D.

Corollary 3.2: The optimal stop region $S \equiv \{P : f(P) = 0\}$ is convex.

Proof: Suppose $P = \lambda P^1 + (1-\lambda)P^2$ and $f(P^1) = f(P^2) = 0$. Then $f(P) \leq 0$ by (1) and $f(P) \geq 0$ by the above.

Q.E.D.

Let

(3)
$$f_1(P) = \min\left\{0, \min_i\left\{c_i - \alpha_i p_i R_i\right\}\right\}$$

$$f_n(P) = \min\left\{0, \min_i\left\{c_i - \alpha_i p_i R_i + (1-\alpha_i p_i) f_{n-1}(T_i P)\right\}\right\} \qquad n > 1$$

Thus $f_n(P)$ is just the minimal risk incurred if the searcher is allowed at most n searches. Clearly $f_n(P) \geq f_{n+1}(P) \geq f(P)$ for all n, all P, and it seems reasonable that $f_n(P) \downarrow f(P)$ as $n \uparrow \infty$. This is shown in the following.

Letting $c = \min_i c_i$, $D = \max_i (R_i - c_i)$

Theorem 3.3: $f_n(P) - f(P) \leq \frac{D^2}{nc}$ all n, all P.

Proof: Let δ^* be an optimal strategy, let T be the random number of times δ^* searches before terminating, and let δ_n^* be δ^* terminated at n, i.e. $\delta_n^* = (\delta_1^* \ldots \delta_{s \wedge n}^*)$. Then

$$(4) \quad f(P) = f(P,\delta^*) = E_{\delta^*}[X \mid T \leq n]P_r[T \leq n] + E_{\delta^*}[X \mid T > n]P_r[T > n]$$

and

$$(5) \quad f_n(P) \leq f(P,\delta_n^*) = E_{\delta^*}[X \mid T \leq n]P_r[T \leq n] + E_{\delta_n^*}[X \mid T > n]P_r[T > n]$$

where X denotes the total cost incurred (and everything is understood to be conditional on the prior probability vector P). Thus

$$(6) \quad f_n(P) - f(P) \leq \left[E_{\delta_n^*}[X \mid T > n] - E_{\delta^*}[X \mid T > n]\right] P_r[T > n]$$

$$\leq D\, P_r[T > n]$$

To get a bound on $P_r[T > n]$ we use (4) to get

$$(7) \quad 0 \geq f(P) \geq -D\, P_r[T \leq n] + (-D + nc)P_r[T > n]$$

$$= -D + nc\, P_r[T > n]$$

or

$$(8) \quad P_r[T > n] \leq D/nc$$

The result follows from (6) and (8). Q.E.D.

Corollary 3.4: If $\alpha_i R_i < c_i$ for all $i=1, 2, \ldots, m$ then $f(P) \equiv 0$, i.e. the policy which never searches is optimal.

Proof: It follows from (3) that $f_1(P) \equiv 0$, and by induction that $f_n(P) \equiv 0$ for all n, and thus by the above $f(P) \equiv 0$. Q.E.D.

The above Corollary may also be proven directly by letting e^i be the m-vector of all zeroes except for a one in the i^{th} spot. If $\alpha_i R_i < c_i$ for all i then by (1) it follows that $f(e^i) = 0$, $i=1, \ldots, m$; and thus by concavity $f(P) \equiv 0$.

4. Counter-Examples

Consider the following three conjectures:

1. If $c_1 > R_1$ then an optimal strategy will never search box 1.
2. If an optimal strategy doesn't stop at P then it searches a box with maximal $\alpha_i p_i / c_i$.
3. If m is the number of boxes then an m-stage look ahead strategy is optimal; where an m-stage look ahead strategy is defined as any strategy which stops at P if $f_m(P) = 0$, and searches the i^{th} box at P if $f_m(P) = c_i - \alpha_i p_i R_i + (1 - \alpha_i p_i)\, f_{m-1}(T_i P)$.

We shall now give examples showing that each of these conjectures need not hold.

<u>Example 1:</u>

	$\alpha_1 = 1$	$\alpha_2 = 1$
	$P_1 = 3/4$	$P_2 = 1/4$
	$c_1 = 5$	$c_2 = 10$
	$R_1 = 0$	$R_2 = 210$

If the searcher first searches 2 and then acts optimally his risk is $10 - \frac{1}{4}\, 210 = -170/4$; while if he first searches 1 and then acts optimally his risk is $5 - \frac{1}{4}\, 200 = -45 < -170/4$. Thus the optimal strategy starts by searching 1.

<u>Example 2:</u>

	$\alpha_1 = 1$	$\alpha_2 = 1$
	$P_1 = 3/4$	$P_2 = 1/4$
	$c_1 = 10$	$c_2 = 10$
	$R_1 = 0$	$R_2 = 210$

If the searcher first searches 1 then his minimal risk is $10 = \frac{1}{4}\,200 = -40$; while if he first searches 2 his minimal risk is $10 - \frac{1}{4}\,210 < -40$. Thus the optimal strategy starts by searching 2. However $\alpha_1 p_1/c_1 = \frac{3}{40} > \frac{1}{40} = \alpha_2 p_2/c_2$.

<u>Example 3</u>:

$\alpha_1 = 1$	$\alpha_2 = .65$
$p_1 = .4$	$p_2 = .6$
$c_1 = 50$	$c_2 = 50$
$R_1 = 100$	$R_2 = 100$

It can be checked directly that $f_2(.4, .6) = 0$ and so the two-stage look ahead strategy stops. However

$$f_3(.4, .6) = .4(-50) + .6[100 - (.65)100 + .35(50 - 100(.65))] < 0$$

and so the two-stage look ahead strategy is not optimal.

Thus none of the conjectures need be true. We will later show, however, that in a special case ($R_i \equiv R$) conjectures 1 and 2 are in fact true.

5. Main Theorems

For any strategy δ let (i, j, δ) be the strategy which first searches i then j and then follows strategy δ.

We shall need the following

Lemma 5.1: For any strategy δ such that $f(P,\delta) < \infty$

$$f(P,(i,j,\delta)) \begin{matrix} > \\ = \\ < \end{matrix} f(P,(j,i,\delta))$$

$$\text{iff} \quad \alpha_i p_i / c_i \begin{matrix} < \\ = \\ > \end{matrix} \alpha_j p_j / c_j$$

Proof: $$f(P,(i,j,\delta)) = c_i - \alpha_i p_i R_i + (1-\alpha_i p_i)\left[c_j - R_j \frac{\alpha_j p_j}{1-\alpha_i p_i} + \left(1 - \frac{\alpha_j p_j}{1-\alpha_i p_i}\right) f(T_j T_i P,\delta)\right]$$

$$f(P,(j,i,\delta)) = c_j - \alpha_j p_j R_j + (1-\alpha_j p_j)\left[c_i - R_i \frac{\alpha_i p_i}{1-\alpha_j p_j} + \left(1 - \frac{\alpha_i p_i}{1-\alpha_j p_j}\right) f(T_i T_j P,\delta)\right]$$

now since $T_j T_i P = T_i T_j P$ it follows that

$$f(P,(i,j,\delta)) - f(P,(j,i,\delta)) = \alpha_j p_j c_i - \alpha_i p_i c_j$$

Q.E.D.

Notation: For any policy $\delta = (\delta_1, \ldots, \delta_s)$ and $t \le s$, let

$$P_{\delta,t} = T_{\delta_t} T_{\delta_{t-1}} \cdots T_{\delta_1} P.$$

Thus $P_{\delta,t}$ is just the posterior probability vector given that δ is employed and the item has not been found after t searches.

Theorem 5.2: If $\alpha_i p_i^0/c_i = \max_j \alpha_j p_j^0/c_j$ then

(a) If $\alpha_i p_i^0 R_i \geq c_i$ then there is an optimal strategy δ^* having $\delta_1^* = i$.

(b) If there does not exist an optimal strategy with $\delta_1^* = i$ then no optimal strategy ever searches i.

Proof: (a) We first show that there is an optimal strategy δ^* having $\delta_k^* = i$ for some $k \leq s$. For suppose that no optimal strategy ever searched i; then for any optimal strategy δ^*, $\left(p^0_{\delta^*,t}\right)_i \geq p_i^0$ for all t and so by Lemma 2.1 the optimal strategy need not stop. But then δ_∞ is optimal and so there would be an optimal strategy with $\delta_1^* = i$. Thus there is an optimal strategy δ^* which searches i. Let k be the first time δ^* searches i. If

$$k \neq 1 \quad \text{then since} \quad \left(p^0_{\delta^*,k-2}\right)_j = \begin{cases} cp_i^0 & j=i \\ c_j p_j^0 & j\neq i \end{cases} \qquad \text{where } c_j \leq c$$

it follows that $\alpha_i\left(p^0_{\delta^*,k-2}\right)_i/c_i = \max \alpha_j\left(p^0_{\delta^*,k-2}\right)_j/c_j$; and so by Lemma 5.1 there is an optimal strategy with $\delta_{k-1}^* = i$. By induction we see that there is an optimal strategy with $\delta_1^* = i$.

(b) We have shown by the above that if an optimal strategy δ^* has $\delta_k^* = i$ for some k then there is an optimal strategy with $\delta_1^* = i$.

Q.E.D.

Corollary 5.3: If $\alpha_i p_i^0/c_i > \alpha_j p_j^0/c_j$ for $j \neq i$ then

(a) every optimal strategy has $\delta_1^* = i$

or

(b) no optimal strategy every searches i.

Proof: Follows in the same manner as in the previous Theorem.

Note that if the state of the process at time t is P then from that point on we can consider the process as starting anew with prior probability vector P. Thus at time t it is optimal to search the box with the largest present value of $\alpha p/c$ or else that box is never searched from that point on. We are able to prove a stronger result in the special case where all rewards are equal.

Theorem 5.4: Suppose $R_i \equiv R$ for all i. If $\alpha_i p_i^0/c_i = \max_j \alpha_j p_j^0/c_j$ then either

(a) there is an optimal strategy with $\delta_1^* = i$

or

(b) the only optimal strategy is the one which does not search, i.e. $s = 0$.

Proof: Let $\delta^* = (\delta_1^*, \ldots, \delta_s^*)$ be an optimal strategy. If δ^* ever searches i then we can show by successive permutations (as in Theorem 5.2) that there is an optimal strategy with $\delta_1^* = i$. If δ^* never searches i then $s < \infty$, for if δ^* didn't stop and never searched i then it would have infinite risk and so wouldn't be optimal. Suppose now that $s \neq 0$ and let $k = \delta_s^*$. Since k will be the last search made it follows that $\alpha_k \left(P^0_{\delta^*, s-1}\right)_k R \geq c_k$ (or else it would be better not to make the last search). But since δ^* never searches i it follows that $\dfrac{\left(P^0_{\delta^*, s}\right)_i}{p_i^0} \geq \dfrac{\left(P^0_{\delta^*, s-1}\right)_k}{p_k^0}$; and thus

$$\frac{\alpha_i \left(P^0_{\delta^*, s}\right)_i}{c_i} = \frac{\alpha_i p_i^0}{c_i} \frac{\left(P^0_{\delta^*, s}\right)_i}{p_i^0} \geq \frac{\alpha_k p_k^0}{c_k} \frac{\left(P^0_{\delta^*, s-1}\right)_k}{p_k^0} \geq 1/R$$

But then by Lemma 2.1 it would be optimal to search i at time s + 1, and so by the above there would be an optimal strategy with $\delta_1^* = i$.

Q.E.D.

In a similar manner we may prove the following

Corollary 5.5: If $R_i \equiv R$ and if $\alpha_i p_i^0/c_i \neq \max \alpha_j p_j^0/c_j$, then any strategy δ with $\delta_1 = i$ is not optimal.

Proof: Let ℓ be such that $\alpha_\ell p_\ell^0/c_\ell = \max \alpha_j p_j^0/c_j$. If δ searches j at some time then by successively permuting and using Lemma 5.1 it follows that we may (strictly) improve upon δ. If δ never searches j then by the same reasoning as used in the above Theorem it follows that δ can't be optimal.

Q.E.D.

Thus when all rewards are equal it is either optimal to search a box with the maximal value of $\alpha_i p_i/c_i$ or else it is optimal to stop.

In [3] Chew considered the problem where there is no reward given for finding the object but where there is a penalty cost C incurred if the searcher stops without finding the object. He also supposed that $\alpha_1 = 0$ and $p_1^0 > 0$. (Thus there is positive probability that the object is in the first box but with probability one a search would overlook it.)*

*Actually Chew supposed that $\sum_i p_i^0 < 1$. However this is clearly equivalent to having $\sum_i p_i^0 = 1$ and having a box with an overlook probability of one.

He showed that if $c_i \equiv 1$ then the optimal strategy either searches the box with maximal $\alpha_i p_i / c_i$ or else stops. However, as was previously pointed out, this problem is equivalent to the one we've considered with $R_i \equiv C$. Thus Theorem 5.4 may be considered as an extension of Chew's result to non-constant costs and to general overlook probabilities.

6. Approximations to Optimal Strategy

In this section we suppose that $R_i \equiv R$, and exhibit a sequence of strategies which converge to an optimal strategy.

Let $\delta^* = (\delta_1^*, \ldots, \delta_s^*)$ be an optimal strategy which either when in state P stops if $f(P) = 0$ or else searches a box with maximal value of $\alpha_i p_i / c_i$. Let T be the random number of stages δ^* searches before terminating, and recall that $c = \min_i c_i$. We shall need the following:

Lemma 6.1: $P_r(T > n) \leq \left(1 - \frac{c}{\sum_i c_i/\alpha_i}\right)^n$ for all n

Proof:

The minimal value of $\max_i \alpha_i p_i / c_i$ is achieved by that vector P having

(9) $$\alpha_1 p_1 / c_1 = \alpha_2 p_2 / c_2 = \ldots . = \alpha_m p_m / c_m$$

and thus

(10) $$\min_P \max_i \alpha_i p_i / c_i = \frac{1}{\sum_i c_i / \alpha_i}$$

Now each time δ^* searches a box with maximal value of $\alpha_i p_i / c_i$. Thus each time δ^* searches a box (say box j) the probability $\alpha_j p_j$ the item will be found is such that

(11) $$\alpha_j p_j \geq \frac{c_j}{\sum_i c_i/\alpha_i} \geq \frac{c}{\sum_i c_i/\alpha_i}$$

The result follows immediately. Q.E.D.

Now let $\delta^n = (\delta_1, \ldots, \delta_{s_n})$ be the strategy which when in state P stops if $f_n(P) = 0$ or else searches a box with maximal value of $\alpha_i p_i/c_i$, i.e. $s_n = \min\left\{k: f_n\left(P^0_{\delta^*,k}\right) = 0\right\}$. . Since $f_n(P) \downarrow f(P)$ it follows that $s_n \uparrow s$ as $n \uparrow \infty$.

Recalling that $D = \max (R - c_i) = R - c$ we have

<u>Theorem 6.2</u>: $f(P,\delta^n) \leq f(P) + D(1 - c/\Sigma c_i/\alpha_i)^{n+s_n}$ for all P, all n.

<u>Proof</u>:
$$\begin{aligned} f(P,\delta^n) - f(P) &= -f\left(P_{\delta^*,s_n}\right) P_r(T > s_n) \\ &= \left[f_n\left(P_{\delta^*,s_n}\right) - f\left(P_{\delta^*,s_n}\right)\right] P_r(T > s_n) \\ &\leq D\, P_r(T > n)\, P_r(T > s_n) \end{aligned}$$

where the last inequality follows from (6). The result then follows from Lemma 6.1.

Q.E.D.

In order to effectively apply the policies δ^n, $n \geq 1$, we need to be able to characterize the continuation sets $A_n \equiv \left\{P: f_n(P) < 0\right\}$. These sets can be constructed as follows:

(12) $$A_1 = \left\{P: \exists\ i: c_i - \alpha_i p_i R < 0\right\}$$

$$A_2 = A_1 \cup B_2$$

where

(13) $B_2 = \left\{ P: \exists\ i,j:\ c_i - \alpha_i p_i R + (1-\alpha_i p_i)\left[c_j - \alpha_j (T_i P)_j R\right] < 0 \right\}$

Noting that $(T_i P)_j = (1-\alpha_i \delta_{ij}) p_j (1-\alpha_i p_i)^{-1}$ where $\delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

we can write

(14) $B_2 = \left\{ P: \exists\ i,j:\ c_i - \alpha_i p_i R + c_j - \alpha_j p_j R - \alpha_i p_i c_j + \alpha_j^2 \delta_{ij} p_j R < 0 \right\}$

Similarly

$$A_3 = A_2 \cup B_3$$

where

(15)
$$B_3 = \Big\{ P: \exists\ i,j,k:\ c_i - \alpha_i p_i R + (1-\alpha_i p_i)\Big[c_j - \alpha_j (T_i P)_j R + (1-\alpha_j (T_i P)_j)(c_k - \alpha_k (T_j T_i P)_k R)\Big] < 0 \Big\}$$

$$= \Big\{ P: \exists\ i,j,k:\ c_i - \alpha_i p_i R + c_j - \alpha_j p_j R + c_k - \alpha_k p_k R - \alpha_i p_i c_j - (\alpha_i p_i + \alpha_j p_j) c_k + \alpha_j^2 \delta_{ij} p_j (R + c_k) + \alpha_k^2 p_k R(\delta_{jk} + \delta_{ik}) - \alpha_k^3 \delta_{ik} \delta_{jk} p_k R < 0 \Big\}$$

Similarly the other A_n's $= A_{n-1} \cup B_n$ may be obtained. Also we may let

(16) $B_1^1 = A_1$

$$B_2^1 = \left\{ P: \exists\ i \neq j:\ c_i - \alpha_i p_i R + c_j - \alpha_j p_j R - \alpha_i p_i c_j < 0 \right\}$$

$$B_3^1 = \Big\{ P: \exists\ i \neq j \neq k:\ c_i - \alpha_i p_i R + c_j - \alpha_j p_j R + c_k - \alpha_k p_k R - \alpha_i p_i c_j - (\alpha_i p_i + \alpha_j p_j) c_k < 0 \Big\}$$

Then $B_n^1 \subset B_n$ and we may approximate A_n by $\bigcup_{i=1}^{n} B_i^1$. We also note that $B_1^1 = A_1$ and $B_2^1 = A_2$.

REFERENCES

[1] Bellman, R., DYNAMIC PROGRAMMING, Princeton University Press, Princeton, New Jersey, (1957).

[2] Black, W., "Discrete Sequential Search," Information and Control, Vol. 8, pp. 159-162, (1965).

[3] Chew, M., Jr., "A Sequential Search Procedure," Annals of Mathematical Statistics, Vol. 38, pp. 494-502, (1967).

[4] Kadane, J., "Discrete Search and the Neyman Pearson Lemma," Journal of Mathematical Analysis and Applications.

[5] Matula, D., "A Periodic Optimal Search," American Mathematical Monthly, Vol. 71, No. 15, (1964).

[6] Strauch, R., "Negative Dynamic Programming," Annals of Mathematical Statistics, Vol. 37, pp. 871-890, (1966).