The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1446–1455

SEQUENTIAL SEARCH WITH DISCOUNTED INCOME, THE DISCOUNT A FUNCTION OF THE CELL SEARCHED

BY CALVIN W. SWEAT

Naval Undersea Research and Development Center, Pasadena

0. Summary. An object is hidden in one of the cells $1, 2, \cdots, R$ with probability distribution $s = (s(1), s(2), \cdots s(R))$ and remains in that cell while search is conducted. A searcher is informed of s and continues search until the object is found. He is informed also of $p = (p(1), p(2), \cdots p(R))$ where $p(i)$ is the probability of finding the object if i is searched and the object is in i. If the object is found on trial $n+1$ its worth is discounted by the factor $\prod_{i=1}^{R} \beta_i^{N(i)}$, where $N(i)$ is the number of inspections of cell i during the first n trials and $0 \leqq \beta_i < 1$, $i = 1, 2, \cdots, R$ is known by the searcher. For each n, if $s(\bar{f}(n)) = (s(1 \mid \bar{f}(n)), s(2 \mid \bar{f}(n)), \cdots, s(R \mid \bar{f}(n)))$ denotes the conditional location distribution, given the history $\bar{f}(n)$ of failures for n trials, then it is shown that an optimal procedure selects an i achieving

$$\max_i \frac{p(i)s(i \mid \bar{f}(n))}{1-\beta_i}.$$

The limiting behavior of the value is investigated as each element in any collection of components of $(\beta_1, \beta_2, \cdots, \beta_R)$ tends to one.

1. Introduction. The search problem considered in this study is a dynamic programming problem. It is also a two-person, zero-sum, infinite game of perfect recall, in which the minimizing player has only one strategy.

A player called the Searcher is attempting to locate an object which has been hidden in one of R cells according to a known probability distribution $s = (s(1), s(2), \cdots s(R))$. He is informed also of a vector of probabilities $p = (p(1), p(2), \cdots p(R))$, where $p(i)$ is the conditional probability of finding the object, given that it is located in cell i and i is searched. The object is hidden once, at start of play, and search continues until the object is found.

Various writers have assumed that there is a cost c_i associated with searching cell i. In his notes on dynamic programming, Blackwell (see [4]) found a search procedure which minimizes the expected total searching cost; the procedure directs the Searcher to inspect that location for which the ratio of the current conditional detection probability to the cell cost is largest. Staroverov [5] and Matula [4] investigated the question of ultimate periodicity of this procedure, with Matula finding necessary and sufficient conditions. Chew [1] studied this problem for the case in which the cell search costs are all equal and $\sum s(i) < 1$. He introduced a penalty, payable if the Searcher stops looking before the object is found. He determined that the search procedure described above is optimal and obtained some results relative to the optimal stopping rule.

Received June 11, 1968; revised February 18, 1969.

A variation of this model will more adequately describe certain search problems. The physical act of searching may be dangerous, involving some risk that the Searcher will not survive to continue the search. The degree of danger may vary with the area searched, e.g., the cells may be geographical regions containing fortifications or natural barriers. It is assumed that the object being sought does not contribute directly to this danger, i.e., does not attack the Searcher. Survival considerations, coupled with small searching costs relative to the object's worth, may make it more appropriate to replace cell search costs by survival probabilities. A study of this model is the concern of this paper.

Let β_i be the survival probability associated with cell i (the presence or absence of the object does not affect this number). The Searcher wishes to maximize the probability of eventually finding the object. Since the probability of finding the object on a particular day contains as a factor the probability $\prod_{i=1}^{R} \beta_i^{N(i)}$ of survival until this day, where $N(i)$ is the number of times cell i has been searched, an equivalent interpretation of β_i is as a discount associated with cell i. This terminology is employed.

Let the Searcher be informed of the discount vector $\beta = (\beta_1, \beta_2, \cdots, \beta_R)$, $0 \leqq \beta_i < 1$, $i = 1, 2, \cdots, R$. If the object is found on day $n+1$, its worth is discounted $\prod_{i=1}^{R} \beta_i^{N(i)}$, where $N(i)$ is the number of inspections of cell i during the first n days. The Searcher desires to maximize the expected total discounted income (payoff).

We shall show the following:

(i) The search game has a solution. Given the discount vector β, the vector of probabilities p, and the initial location distribution s, the value depends on the history only through the conditional location distribution given that history. An optimal strategy (procedure) consists of searching a cell i for which the ratio of the current conditional detection probability to $1-\beta_i$ is largest, a result bearing analogy with that of Blackwell mentioned above. The value V is the uniform limit of a sequence of functions defined iteratively.

(ii) The value V has a limit, which we obtain, as the elements in any collection of the components of β tend to one.

2. Model. The Searcher is searching sequentially for an object which has been hidden in one of R cells. He is informed only of the three R-tuples $s = (s(1), s(2), \cdots, s(R))$, $p = (p(1), p(2), \cdots, p(R))$, and $\beta = (\beta_1, \beta_2, \cdots, \beta_R)$. A vector $p^{(i)} = (p^{(i)}(1), p^{(i)}(2), \cdots p^{(i)}(R))$ is selected with probability $s(i)$, $i = 1, 2, \cdots, R$, where

$$\begin{aligned} p^{(i)}(j) &= p(i) \quad \text{if} \quad i = j \\ &= 0 \qquad\ \text{if} \quad i \neq j, \end{aligned}$$

and $\sum_{i=1} s(i) = 1$; $s(i)$ is the probability that the object was placed in cell i and $p^{(i)}(j)$ is the conditional probability that the object is found, given that it is in i and j is searched. β_i, $0 \leqq \beta_i < 1$, $i = 1, 2, \cdots, R$, is the discount associated with cell i.

If the object is found on trial $n+1$ and there have been $N(1), N(2), \cdots N(R)$ inspections of cell $1, 2, \cdots, R$, respectively, during the first n trials, then the worth of the object is discounted by the factor $\prod_{i=1}^{R} \beta_i^{N(i)}$.

A sequence of random variables $I(1), X(1), I(2), X(2), \cdots, I(n), X(n), \cdots$ is defined on the infinite Borel space; $I(n)$ is the cell searched on day n (trial n) and $X(n)$ is 1 or 0, depending on whether or not the object is found in cell $I(n)$ on day n. We also denote this sequence by $H(1), H(2), \cdots, H(n), \cdots$, where $H(n) = (I(n), X(n))$ is referred to as the history on day n. The history for n days is $\bar{H}(n) = (H(1), H(2), \cdots, H(n))$. The Searcher has perfect recall, being informed of the past history and retaining this information. According to Kuhn's theorem [2] he may, therefore, restrict himself to the use of behavior strategies. A behavior strategy χ is a function on the space of finite sequences of histories assigning to the history $\bar{H}(n-1)$ a cell to search on day n, $\chi(\bar{H}(n-1)) = I(n)$ ($\bar{H}(0)$ is the history of length zero, corresponding to start of play). Given χ and $p^{(i)}$, the consistency theorem [3] guarantees that a probability measure $P_\chi(\cdot \mid p^{(i)})$ is determined on the infinite Borel space by the specifications of the joint probabilities

$$
\begin{aligned}
(1)\quad & P_\chi\{H(k) = (i(k), 0), k = 1, 2, \cdots, n-1, H(n) = (i(n), x) \mid p^{(i)}\} \\
&= [p^{(i)}(i(n))]^x [q^{(i)}(i(n))]^{1-x} [q(i)]^{N(i,\bar{h}(n-1))} \\
&\qquad\qquad\qquad\qquad \text{if } \chi(\bar{h}(j-1)) = i(j); j = 1, 2, \cdots, n, \\
&= 0 \qquad\qquad\qquad\quad\ \text{otherwise,}
\end{aligned}
$$

where $q^{(i)}(i(n)) = 1-p^{(i)}(i(n))$, $q(i) = 1-p(i)$, and $N(k, \bar{h}(n-1))$, $k = 1, 2, \cdots, R$, is the number of times cell k is searched during history $\bar{h}(n-1)$.

Obviously, any two behavior strategies which agree until the object is located produce the same payoff. It suffices, then, to consider strategies which may be represented as $\chi = (i(1), i(2), \cdots, i(n), \cdots)$, i.e., a specification of the cells to search given that the object is not located. The notation $\bar{h}(n)$, representing a value of the random variable $\bar{H}(n)$, fails to distinguish between $h(n) = (i(n), 1)$ and $h(n) = (i(n), 0)$. To provide this distinction we introduce the notation $f(k) = (i(k), 0)$ to represent the one-day history of failure to find the object on searching cell $i(k)$ on day k, $k = 1, 2, \cdots$. We write $\bar{f}(n) = (f(1), f(2), \cdots, f(n))$. Under $\chi = (i(1), i(2), \cdots, i(n), \cdots)$ the discounted income for n days is $\sum_{k=1}^{n} X(k) D(\bar{f}(k-1))$, where $D(\bar{f}(k-1)) = \prod_{j=1}^{R} \beta_j^{N(j,\bar{f}(k-1))}$. This has expectation under $P_\chi(\cdot \mid p^{(i)})$ of

$$H^{(n)}(\chi \mid p^{(i)}) = \sum_{k=1}^{n} P_\chi(X(k) = 1 \mid p^{(i)}) D(\bar{f}(k-1)).$$

The discounted income for n days converges monotonically to

$$\sum_{k=1}^{\infty} X(k) D(\bar{f}(k-1));$$

by the monotone convergence theorem,

$$
\begin{aligned}
H^{(n)}(\chi \mid p^{(i)}) \uparrow_n H(\chi \mid p^{(i)}) &= \sum_{k=1}^{\infty} D(\bar{f}(k-1)) P_\chi(X(k) = 1 \mid p^{(i)}) \\
&= \sum_{k=1}^{\infty} \delta(i(k), i) D(\bar{f}(k-1)) p(i) [q(i)]^{N(i,\bar{f}(k-1))},
\end{aligned}
$$

where

$$\delta(i(k), i) = 1 \quad \text{if} \quad i(k) = i,$$
$$= 0 \quad \text{if} \quad i(k) \neq i.$$

Then $H(\chi \,|\, p^{(i)})$ has expectation under the distribution s of

$$H(\chi, s) = \sum_{i=1}^{R} s(i) \sum_{k=1}^{\infty} \delta(i(k),i) D(\bar{f}(k-1)) p(i) [q(i)]^{N(i,\bar{f}(k-1))}$$
$$= \sum_{k=1}^{\infty} s(i(k)) D(\bar{f}(k-1)) p(i(k)) [q(i(k))]^{N(i(k),\bar{f}(k-1))}.$$

This is the expected total discounted income under use of behavior strategy χ, when the distribution of the object location is s (when the state is s). It is also referred to as the payoff under χ at state s. The Searcher desires to select χ to maximize $H(\chi, s)$.

It will be convenient to speak of the distribution of object location, s, as the location distribution or the state of the system.

3. The solution. Let the strategy be $\chi = (i(1), i(2), \cdots i(n), \cdots)$. Given failure to locate the object along this sequence, the elements of the sequence $\{s(\bar{f}(n)) : n = 1, 2, \cdots\}$ of conditional location distributions satisfy

$$\text{(2)} \qquad s(i \,|\, \bar{f}(k)) = \frac{s(i \,|\, \bar{f}(k-1)) [q(i)]^{\delta(i(k),i)}}{1 - p(i(k)) s(i(k) \,|\, \bar{f}(k-1))} \qquad i = 1, 2, \cdots, R;\ k = 1, 2, \cdots,$$

where $\delta(i(k), i)$ is 1 or 0 according to whether or not $i(k) = i$ and $\bar{f}(n)$ is as defined previously. Applying this relationship repeatedly to the right side, we deduce

$$\text{(3)} \qquad s(i \,|\, \bar{f}(k)) = \frac{s(i) [q(i)]^{N(i,\bar{f}(k))}}{\sum_{j=1}^{R} s(j) [q(j)]^{N(j,\bar{f}(k))}} \qquad i = 1, 2, \cdots, R;\ k = 1, 2, \cdots.$$

The denominator is just $P_\chi(\bar{f}(k))$, the probability of k failures under strategy χ.

We use (3) to derive a recursive relation for the payoff H:

$$H(\chi, s) = \sum_{k=1}^{\infty} s(i(k))\ D(\bar{f}(k-1))\, p(i(k))\, [g(i(k))]^{N(i(k),\, \bar{f}(k-1))}.$$

Defining $\bar{f}(n \,|\, m) = (f(m+1), f(m+2), \cdots, f(n))$ to be that part of the history $\bar{f}(n)$ following day m $(m < n)$, the contribution to this sum from $k = n+1, n+2, \cdots$ is

$$D(\bar{f}(n)) \sum_{k=n+1}^{\infty} s(i(k)) D(\bar{f}(k-1 \,|\, n)) p(i(k)) [q(i(k))]^{N(i(k),\bar{f}(n))} [q(i(k))]^{N(i(k),\bar{f}(k-1 | n))}$$
$$= P_\chi(\bar{f}(n)) D(\bar{f}(n)) \sum_{k=n+1}^{\infty} s(i(k) \,|\, \bar{f}(n)) D(\bar{f}(k-1 \,|\, n)) p(i(k)) [q(i(k))]^{N(i(k),\bar{f}(k-1 | n))}$$
$$= P_\chi(\bar{f}(n)) D(\bar{f}(n)) H(\chi(\bar{f}(n)), s(\bar{f}(n)))$$

where $\chi(\bar{f}(n)) = (i(n+1), i(n+2), \cdots)$ and $s(\bar{f}(n))$ is the conditional location distribution, given the initial distribution s and the history of failures $\bar{f}(n)$. Hence,

$$\text{(4)} \qquad H(\chi, s) = \sum_{k=1}^{n} s(i(k)) D(\bar{f}(k-1)) p(i(k)) [q(i(k))]^{N(i(k),\bar{f}(k-1))}$$
$$+ P_\chi(\bar{f}(n)) D(\bar{f}(n)) H(\chi(\bar{f}(n)), s(\bar{f}(n))).$$

Specializing to $n = 1$, we obtain

(5) $$H(\chi, s) = p(i(1))s(i(1)) + \beta_{i(1)}[1 - p(i(1))s(i(1))]H(\chi(\bar{f}(1)), s(\bar{f}(1))).$$

THEOREM 1.

(i) *There is a unique bounded function defined on the space of location distributions s satisfying*

$$V(s) = \max_i \{p(i)s(i) + \beta_i[1 - p(i)s(i)]V(s(i, 0))\} \qquad \forall s.$$

(ii) *For any bounded function g on the space of location distributions, the sequence $\{T^n g; n = 1, 2, \cdots\}$ of functions on this space defined by*

$$(T^n g)(s) = \max_i \{p(i)s(i) + \beta_i[1 - p(i)s(i)](Tg^{n-1})(s(i, 0))\} \qquad \forall s;\ T^o g \equiv g,$$

converges pointwise and uniformly to $V(s)$.

(iii) *$V(s)$ is the value of the search game when the conditional location distribution (state), given the past history, is s.*

(iv) *The Searcher has an optimal strategy which consists of selecting an i achieving the maximum in* (i) *at each state s.*

PROOF. The proof differs from that of Lemma 2 and Theorem 1 of [6] only in minor details.

THEOREM 2. *Let $\chi^* = (\alpha(1), \alpha(2), \cdots, \alpha(n), \cdots)$ be a strategy which, in state s, requires selecting an i achieving* $\max_i p(i)s(i)/(1 - \beta_i)$. *Then, χ^* is optimal.*

PROOF. Referring to (2), we conclude that whenever a particular cell is searched and the object is not found, the conditional probability that the object is in that cell decreases, while the remaining probabilities are increased by a constant factor. If i were searched successively, the sequence of conditional probabilities that the object is in i would tend to 0. An obvious consequence is that each cell is searched infinitely often in the sequence $\alpha(1), \alpha(2), \cdots$.

We may suppose, then, that for any cell $i(1) \neq \alpha(1)$, $\chi^* = (\alpha(1), \alpha(2), \cdots, \alpha(d-1), i(1), \alpha(d+1), \cdots)$, where $\alpha(k) \neq i(1)$ for $k < d$, i.e., $i(1)$ is searched for the first time on day d. Consider any strategy starting with search at $i(1)$ and continuing according to the prescription of the theorem. Claim: One such strategy is

$$\chi' = (i(1), \alpha(1), \alpha(2), \cdots, \alpha(d-1), \alpha(d+1), \alpha(d+2), \cdots).$$

PROOF.

By assumption,

$$\frac{p(\alpha(k))s(\alpha(k)\,|\,\bar{f}^*(k-1))}{1 - \beta_{\alpha(k)}} = \max_i \frac{p(i)s(i\,|\,\bar{f}^*(k-1))}{1 - \beta_i} \qquad \text{for} \quad k = 1, 2, \cdots,$$

where $\bar{f}^*(n)$ represents the n-day history of failures under χ^*.

Hence, for $k = 1, 2, \cdots, d-1$,

$$\frac{p(\alpha(k)s(\alpha(k) \mid i(1), 0, \bar{f}^*(k-1))}{1-\beta_{\alpha(k)}} = \frac{p(\alpha(k))s(\alpha(k) \mid \bar{f}^*(k-1))}{[1-p(i(1))s(i(1) \mid \bar{f}^*(k-1))][1-\beta_{\alpha(k)}]}$$
$$= \frac{1}{1-p(i(1))s(i(1) \mid \bar{f}^*(k-1))} \max_i \frac{p(i)s(i \mid \bar{f}^*(k-1))}{1-\beta_i}$$
$$= \max_i \frac{p(i)s(i \mid i(1), 0, \bar{f}^*(k-1))}{1-\beta_i}.$$

Then, $s(i \mid i(1), 0, \bar{f}^*(d-1)) = s(i \mid \bar{f}^*(d))$ for $i = 1, 2, \cdots, R$. Thus, following day d, χ' may call for search of the same cells as χ^* and the claim is established.

Since the cells searched and the conditional location distributions following day d are the same under both χ^* and χ', the payoffs can differ only through the first d days. The income for d days under χ', starting in state s, is

$$H_d(\chi', s) = p(i(1))s(i(1)) + \textstyle\sum_{k=2}^{d} \beta_{i(1)}\beta_{\alpha(1)} \cdots \beta_{\alpha(k-2)} s(\alpha(k-1))p(\alpha(k-1)) \cdot [q(\alpha(k-1))]^{N(\alpha(k-1), \bar{f}^*(k-2))}$$
$$= p(i(1))s(i(1)) + \beta_{i(1)} \textstyle\sum_{k=1}^{d-1} D(\bar{f}^*(k-1))s(\alpha(k))p(\alpha(k)) \cdot [q(\alpha(k))]^{N(\alpha(k), \bar{f}^*(k-1))}.$$

By assumption, for each $k \leqq d$ (since $N(i(1), \bar{f}^*(k-1)) = 0$),

$$\frac{p(\alpha(k))s(\alpha(k) \mid \bar{f}^*(k-1))}{1-\beta_{\alpha(k)}} \geqq \frac{p(i(1))s(i(1) \mid \bar{f}^*(k-1))}{1-\beta_{i(1)}}$$
$$\Leftrightarrow \frac{s(\alpha(k))p(\alpha(k))[q(\alpha(k))]^{N(\alpha(k), \bar{f}^*(k-1))}}{1-\beta_{\alpha(k)}} \geqq \frac{p(i(1))s(i(1))}{1-\beta_{i(1)}} \qquad \text{(using (3))}$$
$$\Leftrightarrow s(\alpha(k))p(\alpha(k))[q(\alpha(k))]^{N(\alpha(k), \bar{f}^*(k-1))} + \beta_{\alpha(k)} p(i(1))s(i(1)) \geqq p(i(1))s(i(1)) + \beta_{i(1)} s(\alpha(k))p(\alpha(k))[q(\alpha(k))]^{N(\alpha(k), \bar{f}^*(k-1))}.$$

Multiplying by $D(\bar{f}^*(k-1))$ and summing from $k = 1$ through $d-1$, we obtain $H_d(\chi^*, s) \geqq H_d(\chi', s)$ since $\alpha(d) = i(1)$. Hence,

$$H(\chi^*, s) \geqq H(\chi', s) = p(i(1))s(i(1)) + \beta_{i(1)}[1-p(i(1)s(i(1))]H(\chi'(i(1), 0), s(i(1), 0))$$
$$= p(i(1))s(i(1)) + \beta_{i(1)}[1-p(i(1))s(i(1))]H(\chi^*, s(i(1), 0)).$$

The last equality is due to the assumption that, following play on day 1, χ' calls for the same sequence as χ^*, starting in state $s(i(1), 0)$. Therefore, the payoff under χ^* satisfies the functional equation of Theorem 1(i), proving that χ^* is optimal.

4. Behavior of V as $\beta \to (1, 1, \cdots, 1)$. Thus far we have assumed that the discount vector was fixed. The payoff under a particular strategy depends on β and, in what follows, we denote this dependence by affixing β as subscript. Let $\mathscr{C}$ be a collection of c components of β, $c \leqq R$. For convenience, suppose $\mathscr{C} = (\beta_1, \beta_2, \cdots, \beta_c)$. Let

$\mathscr{D} = \{i: \beta_i \in \mathscr{C} \text{ and } p(i) > 0\}$ and $\mathscr{D}' = \{i: \beta_i \in \mathscr{C} \text{ and } p(i) = 0\}$. Without loss of generality, let $\mathscr{D} = \{1, 2, \cdots, a\}$ and $\mathscr{D}' = \{a+1, a+2, \cdots, c\}$, $a \leqq c \leqq R$. We investigate the behavior of V_β as

$$\beta = (\beta_1, \beta_2, \cdots, \beta_R) \to \beta^* = (1, 1, \cdots, 1, \beta_{c+1}, \beta_{c+2}, \cdots, \beta_R).$$

LEMMA. *Let* $\alpha = (\beta_{c+1}, \beta_{c+2}, \cdots, \beta_R)$ *and*

$$\tilde{s} = \left(\frac{s(c+1)}{\sum_{k=c+1}^R s(k)}, \frac{s(c+2)}{\sum_{k=c+1}^R s(k)}, \cdots, \frac{s(R)}{\sum_{k=c+1}^R s(k)}\right).$$

Then,

(i) $$(\textstyle\sum_{k=c+1}^R s(k))V_\alpha(\tilde{s}) \leqq V_\beta(s) \leqq \sum_{k\in\mathscr{D}} s(k) + (\sum_{k=c+1}^R s(k))V_\alpha(\tilde{s}).$$

(ii) $$\liminf_{\beta\to\beta^*} V_\beta(s) \geqq \textstyle\sum_{k\in\mathscr{D}} s(k) + (\sum_{k=c+1}^R s(k))V_\alpha(\tilde{s}).$$

PROOF. Consider that strategy which ignores the cells $1, 2, \cdots, c$ and plays optimally among the remaining cells. With probability $\sum_{k=1}^c s(k)$ the income is zero, while with the remaining probability the object is located in cells $c+1$, $c+2, \cdots, R$ with location distribution $\tilde{s}$. Application of Theorem 1 yields the left-most member of the inequalities in (i) as the payoff under this strategy at s; this is dominated at each s by the payoff $V_\beta(s)$ under optimal play.

To prove the right-most inequality, suppose, momentarily, that the searcher's information is modified as follows: He is informed which group of cells contains the object, i.e., either $\mathscr{D}$, $\mathscr{D}'$, or $(c+1, c+2, \cdots, R)$. Let $\gamma = (\beta_1, \beta_2, \cdots, \beta_a)$, $\delta = (\beta_{a+1}, \beta_{a+2}, \cdots, \beta_c)$, $s(\mathscr{D})$, and $s(\mathscr{D}')$ be the conditional location distributions, given that the object is in $\mathscr{D}$ or $\mathscr{D}'$, respectively. Relying on Theorem 1, optimal play in this modified problem earns

$$\begin{aligned}[\textstyle\sum_{i=1}^a s(i)]V_\gamma(s(\mathscr{D})) + [\sum_{i=a+1}^c s(i)]V_\delta(s(\mathscr{D}')) + [\sum_{i=c+1}^R s(i)]V_\alpha(\tilde{s})\\ \leqq \textstyle\sum_{i=1}^a s(i) + [\sum_{i=c+1}^R s(i)]V_\alpha(\tilde{s})\end{aligned}$$

(since the first term is bounded above by $\sum_{i=1}^a s(i)$ and the middle term is zero).

Among the strategies in the modified problem is one which ignores the assumed additional information and is optimal in the complete group of cells $1, 2, \cdots, R$. This strategy earns $V_\beta(s)$. Hence $V_\beta(s)$ is no larger than the value of the modified problem. The proof of the right-most inequality is complete.

To prove (ii), consider a strategy χ' which calls for n inspections in each of the cells $1, 2, \cdots, c$ during the first cn days and subsequently calls for optimal play. By (4),

$$\begin{aligned}H_\beta(\chi', s) = \textstyle\sum_{k=1}^{cn} s(i(k))D(\bar{f}(k-1))p(i(k))[q(i(k))]^{N(i(k),\bar{f}(k-1))}\\ + P_{\chi'}(\bar{f}(cn))D(\bar{f}(cn))V_\beta(s(\bar{f}(cn))).\end{aligned}$$

Using (1) and (3) with the fact that $(s(c+1\,|\,\bar{f}(cn)), s(c+2\,|\,\bar{f}(cn)), \cdots, s(R\,|\,\bar{f}(cn)))$

is a multiple of $(s(c+1), s(c+2), \cdots, s(R))$, we conclude that

$$\begin{aligned} H_\beta(\chi', s) &\geqq \textstyle\sum_{k=1}^{cn} s(i(k))D(\bar{f}(k-1))p(i(k))[q(i(k))]^{N(i(k),\bar{f}(k-1))} \\ &\quad + P_{\chi'}(\bar{f}(cn))D(\bar{f}(cn))[\textstyle\sum_{k=c+1}^{R} s(k\,|\,\bar{f}(cn))]V_\alpha(\tilde{s}) \\ &= \textstyle\sum_{k=1}^{cn} s(i(k))D(\bar{f}(k-1))p(i(k))[q(i(k))]^{N(i(k),\bar{f}(k-1))} \\ &\quad + D(\bar{f}(cn))[\textstyle\sum_{k=c+1}^{R} s(k)]V_\alpha(\tilde{s}). \end{aligned}$$

Obviously, given any $\varepsilon > 0$, n can be fixed so large, independent of β, that, as $\beta \to \beta^*$ the contribution of the first n days to the total income tends to a number within ε of $\sum_{k \in \mathscr{D}} s(k)$ and the factor multiplying $V_\alpha(\tilde{s})$ tends to $\sum_{k=c+1}^{R} s(k)$. Therefore, $\lim_{\beta \to \beta^*} H_\beta(\chi', s)$ is arbitrarily near the right-most member of the inequality in (ii). But, $V_\beta(s) \geqq H_\beta(\chi', s)$ so that (ii) holds.

The lemma leads immediately to

THEOREM 3. $V_\beta(s) \to_{\beta \to \beta^*} \sum_{k \in \mathscr{D}} s(k) + (\sum_{k=c+1}^{R} s(k))V_\alpha(\tilde{s})$.

5. An example. Let $\tilde{s} = (s, 1-s)$, $p^{(1)} = (\frac{1}{2}, 0)$, $p^{(2)} = (0, 1)$, $\beta = (\beta_1, \beta_2)$. Repeating the interpretation of these vectors, s is the probability that the object is located in cell 1 and 1/2 is the conditional probability that the object is located, given that it is in cell 1 and cell 1 is searched; $1-s$ is the probability that the object is located in cell 2 and 1 is the conditional probability that the object is located, given that it is in cell 2 and cell 2 is searched; $\beta_i (i = 1, 2)$ is the discount associated with cell i. It is convenient to say that with probability s the Searcher is playing the game $A = \binom{1/2}{0}$, while with probability $1-s$ he is playing $B = \binom{0}{1}$.

By Theorem 1 the value of the game at each state $(s, 1-s)$ satisfies

$V(s) = \max \binom{s/2 + \beta_1(1-s/2)V(s_1, 0)}{1-s+\beta_2 sV(1)}$. Now, $V(1) = \max \binom{1/2 + \beta_1 V(1)/2}{\beta_2 sV(1)}$. $V(1) \neq 0$; hence

$$V(1) = 1/2 + \beta_1 V(1)/2 \Rightarrow V(1) = \frac{1/2}{1-\beta_1/2}.$$

Then,

$$V(s) = \max \begin{pmatrix} s/2 + \beta_1(1-s/2)V\left(\dfrac{s/2}{1-s/2}\right) \\ 1-s+\beta_2 s\dfrac{1/2}{1-\beta_1/2} \end{pmatrix}.$$

By Theorem 2, when

$$\frac{s/2}{1-\beta_1} \begin{matrix} > \\ < \\ = \end{matrix} \frac{1-s}{1-\beta_2} \qquad \begin{matrix} \text{first strategy is optimal,} \\ \text{second strategy is optimal,} \\ \text{either strategy is optimal.} \end{matrix}$$

Therefore

$$s \leqq s_0 = \frac{1-\beta_1}{3/2-\beta_1-\beta_2/2} \Rightarrow V(s) = 1-s+\beta_2 s \frac{1/2}{1-\beta_1/2}$$

and second strategy is optimal. For $s > s_0$ the first strategy is optimal until there have been n failures, where, using (3), n is the first integer for which $s(1/2)^n/(s(1/2)^n+1-s) \leqq s_0$; i.e.,

$$\begin{aligned}
V(s) &= \sum_{k=1}^{n} \frac{1}{2} s\left(1-\frac{1}{2}\right)^{k-1} \beta_1{}^{k-1} + \beta_1{}^n\left[s\left(\frac{1}{2}\right)^n + 1 - s\right] V\left(\frac{(1/2)^n s}{(1/2)^n s + 1 - s}\right) \\
&= \frac{1}{2} s\left[\frac{1-(\beta_1/2)^n}{1-\beta_1/2}\right] + \beta_1{}^n\left[\left(\frac{1}{2}\right)^n s + 1 - s\right] \\
&\quad \cdot \left\{1 - \frac{(1/2)^n s}{(1/2)^n s + 1 - s} + \beta_2 \frac{(1/2)^n s}{(1/2)^n s + 1 - s} \frac{1/2}{1-\beta_1/2}\right\} \\
&= \frac{s}{1-\beta_1/2}\left\{\frac{1}{2} - \left(\frac{1}{2}\right)^{n+1} \beta_1{}^n + \left(\frac{1}{2}\right)^{n+1} \beta_1{}^n \beta_2\right\} + \beta_1{}^n(1-s).
\end{aligned}$$

As a check on this calculation we verify that

$$\begin{aligned}
&\frac{1}{2} s + \beta_1(1-s/2) V\left(\frac{s/2}{1-s/2}\right) \\
&\quad = \frac{1}{2} s + \beta_1\left(1-\frac{1}{2} s\right)\left\{\frac{s/2}{(1-s/2)(1-\beta_1/2)}\left[\frac{1}{2} - \left(\frac{1}{2}\right)^n \beta_1{}^{n-1} + \left(\frac{1}{2}\right)^n \beta_1{}^{n-1} \beta_2\right]\right. \\
&\qquad \left. + \beta_1{}^{n-1}\left(1 - \frac{s/2}{1-s/2}\right)\right\} = V(s).
\end{aligned}$$

This calculation is valid for $n \geqq 1$.

Summarizing, we have

$$s \leqq \frac{1-\beta_1}{3/2-\beta_1-\beta_2/2} \Rightarrow V(s) = 1-s+\beta_2 s \frac{1/2}{1-\beta_1/2}$$

and the second strategy is optimal;

$$s \geqq \frac{1-\beta_1}{3/2-\beta_1-\beta_2/2} \Rightarrow V(s) = \frac{s}{1-\beta_1/2}\left\{\frac{1}{2} - \left(\frac{1}{2}\right)^{n+1} \beta_1{}^n + \left(\frac{1}{2}\right)^{n+1} \beta_1{}^n \beta_2\right\} + \beta_1{}^n(1-s),$$

where n is the first integer for which

$$\frac{(1/2)^n s}{(1/2)^n s + 1 - s} \leqq \frac{1-\beta_1}{3/2-\beta_1-\beta_2/2},$$

and the first strategy is optimal.

REFERENCES

[1] Chew, M. C. (1967). A sequential search procedure. *Ann. Math. Statist.* **38** 494–502.

[2] Kuhn, H. (1953). Extensive games and the problem of information. *Ann. of Math. Studies* No. 28, 193–216.

[3] Loève, M. (1960). *Probability Theory*, (2nd ed.). Van Nostrand, Princeton.

[4] Matula, D. (1964). A periodic optimal search. *Amer. Math. Monthly* **71** 15–21.

[5] Staroverov, O. V. (1963). On a searching problem. *Theor. Probability Appl.* **8** 184–187.

[6] Sweat, C. W. (1968). Adaptive competitive decision in repeated play of a matrix game with uncertain entries. *Naval Res. Logist. Quart.* **15** 425–448.