M.C.R. Vol. 9, No. 2, May 1984
Printed in U.S.A.

DICHOTOMOUS SEARCH FOR RANDOM OBJECTS ON AN INTERVAL*

R. HASSIN AND M. HENIG

Tel Aviv University

A set of objects to be searched is represented by a set of points lying in an interval of integers. We wish to identify these points within unit-intervals through a dichotomous search, minimizing the expected cost of the search.

The optimal search strategy may depend on the information gained at each stage of the search. It is shown here that for some cases there exists a common optimal strategy. This strategy is generally different from the bisection procedure, and is independent of the number of objects in the interval. As a matter of fact, at each stage of the search a partition of the interval is given and the only information needed for searching a subinterval of this partition is that it still contains an unidentified point.

1. Introduction. Let N be a *counting function* defined on the interval $(0, T]$ of integers.[1] The value $N(t)$ denotes the number of objects in $(0, t]$. Point t is a *jump* if its *increment* $N(t) - N(t-1) > 0$ $(N(0) \equiv 0)$. We wish to identify the jumps of N by the following search scheme. Let I be a function defined on subintervals of $[0, T]$. We call it the *information* function. Initially, an information $I(0, T)$ is known and an a priori joint distribution of the jumps is given. At stage $m = 1, 2, \ldots$ a list of points $0 = t_0 < t_1 < \cdots < t_m = T$ is given and the information $I(t_0, t_1, \ldots, t_m)$ is obtained. Based on the information, the joint distribution function of the jumps is updated. If all the jumps were identified then the search terminates. Otherwise we *split* an interval $(t_{i-1}, t_i]$ for some $i = 1, \ldots, m$ by selecting $t_{i-1} < t^* < t_i$.

A *selection strategy* is a set of rules telling us at each stage, based on the given information, which point to select next. Each selection strategy yields a search whose number of stages is a random variable. A selection strategy is *optimal* if it minimizes the expected number of stages.

We say that an interval $(t_{i-1}, t_i]$ is *resolved* at stage m if $I(t_0, \ldots, t_m)$ identifies all jumps in the interval.

We show that if certain conditions, which we specify later. hold, then the following selection strategy is optimal:

Strategy O: In each stage arbitrarily select an unresolved interval $(t_{i-1}, t_i]$ and split it at $t_{i-1} + 2^{t(d)}$ or at $t_i - 2^{t(d)}$ where $d = t_i - t_{i-1}$ and $t(d)$ is the unique integer satisfying

$$3 \cdot 2^{t(d)-1} \leqslant d < 3 \cdot 2^{t(d)}.$$

Notice that, according to this strategy, the only information needed is whether each interval is resolved or not. Notice also that $d/2^{t(d)}$ does not converge as d increases but fluctuates between 1.5 and 3.

*Received June 1, 1981: revised February 3, 1983.

AMS 1980 subject classification. Primary: 90B40.

OR/MS Index 1978 subject classification. Primary: 751 Search and surveillance.

Key words. Search on intervals.

[1] Interval $(a, b]$ of integers $\equiv \{x \text{ integer} \mid a < x \leqslant b\}$. Unless otherwise stated, we refer to intervals of integers.

0364-765X/84/0902/0301$01.25

As a real life application of the model, suppose that after producing T items in a production line, it was found that there are k types of defects in the last item. Suppose that once a certain type of defect occurs it appears in all of the consecutive items. So by examining the tth item it can be determined whether the appearance of a certain type of defects occurred before producing the tth item, or later. We wish to identify the first appearance of each type of defects. If we can assume that these are independently and uniformly distributed on $(0, T]$ then Strategy O will minimize the expected number of items which are examined.

The problem of locating K balls known to lie within T boxes with a minimax objective was solved by Hassin and Megiddo [5]. The resulting algorithm supplies a bound on the number of questions needed to solve the problem. Other existing literature deals solely with dichotomous search for a single object on a given interval. Morris [8] characterized the solution when the a priori distribution of the object is uniform. Hassin [4] analyzed the problem with geometric a priori distribution and Hu and Tucker [6] and Knuth [7] suggested an algorithm for an arbitrary distribution. Extensions and variations were treated by Cameron and Narayanamurthy [1], Gal [3] and Murakami [9], [10].

2. The main theorem and examples. Suppose that the a priori joint distribution of the jumps is given. Let $R(t_{i-1}, t_i, I(t_0, \ldots, t_m)) = 1$ if $I(t_0, \ldots, t_m)$ indicates that $(t_{i-1}, t_i]$ is unresolved, and 0 otherwise.

Condition A. $R(t_{i-1}, t_i, I(t_0, \ldots, t_m)) = R(t_{i-1}, t_i, I(t_0, \ldots, t^*, \ldots, t_m))$ for every $t_{j-1} < t^* < t_j, j \neq i$.

Condition A means that selection of points outside the interval $(t_{i-1}, t_i]$ does not affect the value of R. This condition implies a common value of $R(a, b, I(t_0, \ldots, t_m))$ for every set of points $t_0 < t_1 < \cdots < t_m$ such that $a = t_{i-1}$, $b = t_i$ for some $1 \leqslant i \leqslant m$. We denote it as $R(a, b)$.

Condition B. $\Pr[R(a, a+d) = 1 \mid I(t_0, t_1, \ldots, t_m)]$ is independent of a and monotone increasing and concave in d for every list $t_0 < t_1 < \cdots < t_m$ such that $t_{i-1} < a \leqslant t_i - d$ for some $i = 1, \ldots, m$.

The condition states that the probability function that an interval $(a, a+d] \subseteq (t_{i-1}, t_i]$ is unresolved given $I(t_0, \ldots, t_m)$ depends only on its length d and is monotone increasing and concave.

THEOREM. *If Conditions A and B are satisfied then Strategy O is optimal.*

We now mention several examples where Conditions A and B are met.

EXAMPLE A. Suppose that several objects are independently uniformly distributed over the interval $(0, T]$. Let $N(t)$ be the number of objects in $[0, t]$. At stage m of the search the information obtained is whether $(t_{i-1}, t_i]$, $i = 1, \ldots, m$, is empty (contains no objects) or not. Thus, $(t_{i-1}, t_i]$ is resolved if it is either empty or $t_i - t_{i-1} = 1$. Condition A is trivially satisfied since splitting at $t^* \notin (t_{i-1}, t_i]$ cannot change the fact that $(t_{i-1}, t_i]$ is empty or nonempty. Let $q(d \mid k)$ be the probability that the interval $(a, a+d] \subseteq (t_{i-1}, t_i]$, $d > 1$, is empty, given that $(t_{i-1}, t_i]$ contains k empty points. Clearly the empty points are uniformly distributed over $(t_{i-1}, t_i]$ and therefore

$$q(d \mid k) = \prod_{j=0}^{d-1} \left[(k-j)/(t_i - t_{i-1} - j) \right] \qquad \text{for} \quad d \leqslant k, \quad \text{and}$$

$$q(d \mid k) = 0 \qquad \text{for} \quad d > k.$$

It is easy to verify that $q(d \mid k)$ is monotone decreasing and convex in d. Therefore, $\sum_{k=0}^{t_i - t_{i-1}} q(d \mid k) \Pr\{(t_{i-1}, t_i]$ contains k empty points$\}$, which is the probability that $(a, a+d]$ is empty, is monotone decreasing and convex in d, which implies Condition B.

EXAMPLE B. Consider the previous example with the exception, $I(t_0, \ldots, t_m) = (N(t_0), \ldots, N(t_m))$. An interval $(t_{i-1}, t_i]$ is resolved then if $N(t_i) - N(t_{i-1}) = 0$ or $t_i - t_{i-1} = 1$, as in Example A. The proof that Conditions A and B are satisfied is similar to that presented in the previous example. Notice that by the theorem the knowledge of $N(t)$ is redundant since the information in Example A leads to the same number of stages. However, $N(t)$ is needed if it is desired to identify the increments and not just the jumps.

EXAMPLE C. Consider again Example A with the exception that $I(t_0, \ldots, t_m) = (J(t_0), J(t_1), \ldots, J(t_m))$, where $J(t)$ is the number of jumps in $(0, t]$. An interval $(t_{i-1}, t_i]$ is resolved if either $J(t_i) - J(t_{i-1}) = 0$ or $J(t_i) - J(t_{i-1}) = t_i - t_{i-1}$. Adding t^* to the list $t_0, t_1, \ldots, t_m$ informs only about $J(t^*)$ but does not change the status of a resolved (or unresolved) interval $(t_{i-1}, t_i] \not\ni t^*$. Therefore Condition A holds. Let $q(d \mid k)$ be the probability that the interval $(a, a+d] \subseteq (t_{i-1}, t_i]$ is resolved, given that $(t_{i-1}, t_i]$ contains k empty points. Let $\delta_1 = \delta_2 = 1$ except for $\delta_1 = 0$ if $d > t_i - t_{i-1} - k$ and $\delta_2 = 0$ if $d > k$. Then

$$q(d \mid k) = \delta_1 \prod_{j=0}^{k-1} \left(1 - \frac{d}{t_i - t_{i-1} - j}\right) + \delta_2 \prod_{j=0}^{d-1} \left(\frac{k-j}{t_i - t_{i-1} - j}\right).$$

This probability function is monotone decreasing and convex which prove Condition B. By the theorem, Strategy O is optimal. Furthermore, obtaining $J(t)$ is redundant and the sufficient information is, whether the interval is resolved or not.

EXAMPLE D. Let N be a counting function defined on the positive real numbers. The value $N(t)$ denotes the number of objects in the interval $(0, t]$ of reals when these objects have been distributed by a compound Poisson process (see Feller [2, Chapter VI.4]). Suppose we want to identify the discontinuity points of $N(t)$ over $(0, T]$ within unit intervals. By restricting N to the integers we face the problem of identifying the jumps of N. Furthermore, the jumps are independently uniformly distributed over the integers.

So far all of the examples assumed independent uniform distribution of the objects over the interval. This assumption can be relaxed by assuming that the jumps are uniformly distributed. For example, the compound Poisson process in Example D can be replaced by a mixture of such processes (Feller [2, Chapter II.5]). Further relaxation may include an a priori joint distribution of the objects which are not necessarily independent but are exchangeable (Feller [2, Chapter VII.4]). That is the joint distribution is invariant under permutation, as in the next example.

EXAMPLE E. Suppose that the a priori joint distribution is generated by sequentially allocating objects on $(0, T]$ according to the following rule: if the points $t_1, \ldots, t_r$ have the same number of objects, then each of them has an equal probability, different from 0 and 1, to obtain the next object. Suppose the same information function as in Example A, then an interval is resolved if it is empty or its length equals 1, and Condition A is satisfied as before. The joint distribution of objects given the information is exchangeable and therefore the empty points are uniformly distributed if their number is given. The proof that Condition B is satisfied is similar to that of Example A. Notice the requirement that the probability for allocating an object is different from 0 or 1, otherwise an interval may become resolved even if it is not empty or has length equal to 1.

3. **The optimal splitting policy.** Let X be a discrete function defined on the integers such that $0 < X(t) < t$, for $t \geq 2$, and $X(1) = X(0) = 0$. We call it a *splitting function*. In this section we restrict ourselves to selection strategies which iteratively select an unresolved interval $(t_{i-1}, t_i]$ and split it at $X(t_i - t_{i-1})$. We call such a strategy a *splitting policy*.

Thus, a splitting policy does not use the information available at each stage of the search to decide what new point in the interval to select. As a matter of fact, the only information it uses is whether an interval is resolved or not. By Condition A this information is not affected by the order by which the unresolved intervals are selected, and therefore the number of stages of the search is independent of the rules used to select the intervals.

For a given T, a sequence of numbers is generated by repeatedly using a splitting function X. Define $x_{0,1} = T$, $x_{1,1} = X(x_{0,1})$, $x_{1,2} = T - x_{1,1}$ and iteratively $x_{l+1,2i-1} = X(x_{l,i})$, $x_{l+1,2i} = x_{l,i} - x_{l+1,2i-1}$. The generated sequence is

$$\{(x_{l,i}) \mid i = 1, \dots, 2^l, l = 0, 1, \dots\}. \tag{1}$$

We can relate the indices (l, i) to the interval $(\sum_{j=1}^{i-1} x_{l,j}, \sum_{j=1}^{i} x_{l,j}]$.

The expected number of stages in the search is equal to the expected number of unresolved intervals occurring during the search and this is equal to

$$E\left\{\sum_{l=0}^{\infty}\sum_{i=1}^{2^l} R\left(\sum_{j=1}^{i-1} x_{l,j}, \sum_{j=1}^{i} x_{l,j}\right)\middle| I(0,T)\right\}$$

$$= \sum_{l=0}^{\infty}\sum_{l=1}^{2^l} \Pr\left(R\left(\sum_{j=1}^{i-1} x_{l,j}, \sum_{j=1}^{i} x_{l,j}\right) = 1 \middle| I(0,T)\right).$$

By Condition B the probabilities depend only on the interval length, so we can denote $p(d) \equiv \Pr(R(a, a+d) = 1 \mid I(0,T))$. Hence the expected number of stages in the search is $\sum_{l=0}^{\infty}\sum_{i=1}^{2^l} p(x_{l,i})$. This is the objective function we want to minimize over all splitting policies.

To each policy there corresponds a sequence (1), and our problem is therefore the following mathematical programming problem:

Program I. $Q(T) = \min \sum_l \sum_{i=1}^{2^l} p(x_{l,i})$ subject to:

$$\left.\begin{aligned} x_{l,i} &= x_{l+1,2i-1} + x_{l+1,2i} \\ x_{l,i} &\geq 0 \quad \text{and integer} \end{aligned}\right\} \quad \begin{aligned} i &= 1, \dots, 2^l, && \text{(I.1)} \\ l &= 0, 1, 2, \dots, && \text{(I.2)} \end{aligned}$$

$$x_{0,1} \equiv T. \tag{I.3}$$

We shall show that when p is concave and nondecreasing on the interval $[0, T]$ with $p(1) = p(0) = 0$ then an optimal solution of Program I satisfies the constraints of Program II.

Let $r = \lfloor \log_2 T \rfloor$ where $\lfloor y \rfloor$ is the largest integer less or equal to y.

Program II. $Q'(T) = \min \sum_{l=0}^{r}\sum_{i=1}^{2^l} p(x_{l,i})$ subject to:

$$\left.\begin{aligned} &\sum_{i=1}^{2^l} x_{l,i} = T \\ &2^{r-l} \leq x_{l,i} \leq 2^{r-l+1}, \qquad i = 1, 2, \dots, 2^l \end{aligned}\right\} \quad l = 0, 1, \dots, r. \quad \begin{aligned} &\text{(II.1)} \\ &\text{(II.2)} \end{aligned}$$

Notice that the objective function of Program I is reduced to that of Program II if constraints II are satisfied since p vanishes for $d = 0, 1$. Program II is composed of r independent problems. Each of them is a Program III problem, which we present now. This problem is also referred to as an allocation problem, where T units of a resource is allocated among n items.

Program III. $\min \sum_{i=1}^{n} p(b_i)$ subject to:

$$\sum_{i=1}^{n} b_i = T \qquad (a \leq T/n \leq c). \tag{III.1}$$

$$a \leq b_i \leq c, \qquad i = 1, \dots, n. \tag{III.2}$$

LEMMA 1. *Let p be a concave function on $[a,c]$. Then $(b_1^*, b_2^*, \dots, b_n^*)$ solves Program* III:

$$b_i^* = c, \qquad i = 1, \dots, m, \quad m \geq 1,$$

$$b_i^* = a, \qquad i = m+2, \dots, n, \quad m \geq 0,$$

$$b_i^* = T - mc - (n - m - 1)a, \qquad i = m+1, \quad m \geq 0$$

where $m = \lfloor (T - na)/(c - a) \rfloor$.

If a, c and T are integers, the solution is also integral.

When p is strictly concave on $[a,c]$ the solution is unique up to a permutation in the indices. If p is linear then every feasible solution is also optimal.

PROOF. A known result in convex programming (see Wagner [11, Chapter 14.9]) states that there exists a basic optimal solution to Program III, i.e., a solution with at most one variable not equal to its bound. By the symmetry of the problem, m arbitrary variables are at their upper bound and arbitrary $n - m - 1$ of the rest of the variables are at their lower bound. ∎

COROLLARY. *Let p be a concave function over $(0, T]$ then the following is a solution of Program* II:

$$x_{l,i} = 2^{r-l+1}, \qquad i = 1, \dots, m(l), \quad m(l) \geq 1,$$

$$x_{l,i} = 2^{r-l}, \qquad i = m(l) + 2, \dots, 2^l, \quad m(l) \geq 0,$$

$$x_{l,i} = T - 2^r - 2^{r-l}(m(l) - 1), \qquad i = m(l) + 1, \quad m(l) \geq 0,$$

where $m(l) = \lfloor (T - 2^r) \cdot 2^{l-r} \rfloor$. Moreover, it satisfies the constraints of Program I.

PROOF. Program II is composed of r independent allocation type problems. So by Lemma 1 the claim is true with

$$m(l) = \lfloor (T - 2^l \cdot 2^{r-l})/(2^{r-l+1} - 2^{r-l}) \rfloor = \lfloor (T - 2^r) \cdot 2^{l-r} \rfloor.$$

Constraints (I.2) and (I.3) are clearly satisfied. To show that (I.1) is satisfied notice that $m(l) = \frac{1}{2} m(l+1)$ if $m(l+1)$ is even, and $m(l) = \frac{1}{2}(m(l+1) - 1)$ otherwise. For every $i \leq m(l)$ $(m(l) \geq 1)$ we have $2i \leq 2m(l) \leq m(l+1)$ and therefore $x_{l+1,2i-1} + x_{l+1,2i} = 2^{r-l+1} = x_{l,i}$. For every $i \geq m(l) + 2$ we have $2i - 1 \geq 2m(l) + 3 \geq m(l+1) + 2$ and therefore $x_{l+1,2i-1} + x_{l+1,2i} = 2^{r-l-1} \cdot 2 = 2^{r-l} = x_{l,i}$. So (I.1) was proved for $i \neq m(l) + 1$ and since $\sum_{i=1}^{2^l} x_{l,i} = \sum_{i=1}^{2^{l+1}} x_{l+1,i} = T$, it must hold also for $i = m(l) + 1$. ∎

LEMMA 2. *Let p be a concave and nondecreasing function on $(0, T]$ which vanishes for 0 and 1. Then there exists an optimal solution of Program I which satisfies the constraint of Program II. Therefore the values $\{x_{l,i}\}$ of the corollary constitute an optimal solution.*

Let $X^0(T) \equiv x_{1,1}$ (and $T - X^0(T) \equiv x_{1,2}$) be an optimal splitting of T.

Let $r = \lfloor \log_2 T \rfloor$ and $l = 1$, then $X^0(T) = T - 2^{r-1}$ (and $T - X^0(T) = 2^{r-1}$) if $T < 2^r + 2^{r-1}$; $X^0(T) = 2^r$ (and $T - X^0(T) = T - 2^r$) if $T \geq 2^r + 2^{r-1}$. In other words, let d denote the unique integer such that $3 \cdot 2^{d-1} \leq T < 3 \cdot 2^d$, then $X^0(T)$ (or $T - X^0(T)$) is equal to 2^d.

PROOF. The proof is by induction on the length of the interval. The claim is true trivially for $T = 2$ since the solution is $x_{1,1} = x_{1,2} = 1$. Suppose the claim is true for lengths $T = 2, \ldots, M$. Given $T = M + 1$, $Q(T)$ in Program I can be written as

$$P(T) + \min_{x_{1,1}, x_{1,2}} \left\{ \min \left[\sum_{l>1} \sum_{i=1}^{2^{l-1}} p(x_{l,i}) + \sum_{l>1} \sum_{i=2^{l-1}+1}^{2^l} p(x_{l,i}) \right] \right\}.$$

Notice that if $x_{1,1}$ and $x_{1,2}$ are given then the variables $\{x_{l,i}\}$ which are constrained by (I.1) can be organized into two independent parts:

$$\{x_{l,i} \mid l = 1,2, \ldots, i = 1,2, \ldots, 2^{l-1}\} \text{ and } \{x_{l,i} \mid l = 1,2, \ldots, i = 2^{l-1}+1, \ldots, 2^l\}.$$

It is obvious that $x_{1,1} \geq 1$ and $x_{1,2} \leq T - 1$. So after arranging the indices and using the induction assumption, we can write Program I as:

Program IV. $Q(T) - P(T) = \min_{x_{1,1}, x_{1,2}} \{ Q'(x_{1,1}) + Q'(x_{1,2}) \}$.

$$x_{1,1} + x_{1,2} = T, \tag{IV.1}$$

$$1 \leq x_{1,1}, x_{1,2} \leq M \quad \text{and integers.} \tag{IV.2}$$

Let $r_1 = \lfloor \log_2 x_{1,1} \rfloor$ and $r_2 = \lfloor \log_2 x_{1,2} \rfloor$.

By the induction assumption there exists a solution of Program I which satisfies the following constraints:

$$\left. \begin{aligned} &\sum_{i=1}^{2^{l-1}} x_{l,i} = x_{1,1} \\ &2^{r_1 - l + 1} \leq x_{l,i} \leq 2^{r_1 - l + 2}, \qquad i = 1, \ldots, 2^{l-1} \end{aligned} \right\} \quad l = 1,2, \ldots, r_1 + 1, \qquad \begin{aligned} &(\text{V.1}) \\ &(\text{V.2}) \end{aligned}$$

$$\left. \begin{aligned} &\sum_{i=2^{l-1}+1}^{2^l} x_{l,i} = x_{1,2} \\ &2^{r_2 - l + 1} \leq x_{l,i} \leq 2^{r_2 - l + 2}, \qquad i = 2^{l-1} + 1, \ldots, 2^l. \end{aligned} \right\} \quad l = 1,2, \ldots, r_2 + 1. \qquad \begin{aligned} &(\text{VI.1}) \\ &(\text{VI.2}) \end{aligned}$$

We shall show first that if r_1 and r_2 are fixed with $s \equiv r_2 - r_1 \geq 1$, then an optimal solution to Program IV exists with either $x_{1,1} = 2^{r_1+1}$ or $x_{1,2} = 2^{r_2}$.

Using the induction assumption and since $p(1) = p(0) = 0$, the objective function of Program IV can be written as

$$Q'(x_{1,1}) + Q'(x_{1,2}) = \sum_{l=1}^{r_1+1} \sum_{i=1}^{2^{l-1}} p(x_{l,i}) + \sum_{l=1}^{r_2+1} \sum_{i=2^{l-1}+1}^{2^l} p(x_{l,i}).$$

By changing the summation index of the first term, splitting the second term and substituting r_2 for $r_1 + s$ we get:

$$Q'(x_{1,1}) + Q'(x_{1,2}) = \sum_{l=s+1}^{r_2+1} \left[\sum_{i=1}^{2^{l-s-1}} p(x_{l-s,i}) + \sum_{i=2^{l-1}+1}^{2^l} p(x_{l,i}) \right] + \sum_{l=1}^{s} \sum_{i=2^{l-1}+1}^{2^l} p(x_{l,i}). \tag{VII.0}$$

Constraints (IV.1), (V.1), (V.2), (VI.1) and (VI.2) can be rewritten as

$$\sum_{i=1}^{2^{l-s-1}} x_{l-s,i} + \sum_{i=2^{l-1}+1}^{2^l} x_{l,i} = x_{1,1} + x_{1,2} = T, \tag{VII.1}$$

$$\left.\begin{aligned} &2^{r_2-l+1} \leq x_{l-s,i} \leq 2^{r_2-l+2}, \quad i = 1, \dots, 2^{l-s-1} &\text{(VII.2)}\\ &2^{r_2-l+1} \leq x_{l,i} \leq 2^{r_2-l+2}, \quad i = 2^{l-1}+1, \dots, 2^l &\text{(VII.3)}\end{aligned}\right\} \quad l = s+1, \dots, r_2+1,$$

$$\left.\begin{aligned} &\sum_{i=2^{l-1}+1}^{2^l} x_{l,i} = x_{1,2} &\text{(VII.4)}\\ &2^{r_2-l+1} \leq x_{l,i} \leq 2^{r_2-l+2}, \quad i = 2^{l-1}+1, \dots, 2^l &\text{(VII.5)}\end{aligned}\right\} \quad l = 1, 2, \dots, s.$$

Consider now the bracketed term of (VII.0) for every $l = s+1, \dots, r_2+1$. Minimizing it under constraints (VII.1), (VII.2) and (VII.3) is an allocation type problem with $n = 2^{l-s-1}(1+2^s)$, $a = 2^{r_2-l+1}$ and $b = 2^{r_2-l+2}$.

By Lemma 1, 2^{r_2-l+2} will be allocated to each of the first $m(l)$ items, and 2^{r_2-l+1} to the last $2^{l-s-1}(1+2^s) - m(l) - 1$ items, where $m(l) = \lfloor (T - (2^{r_1} + 2^{r_2})2^{l-r_2-1}) \rfloor$. We can distinguish between two cases according to the value of $m(l)$.

Case 1. $T < 2^{r_1+1} + 2^{r_2}$. Then $m(l) < 2^{l-s-1}$ and $x_{l,i} = 2^{r_2-l+1}$ for $i = 2^{l-1}+1, \dots, 2^l$. So the amount $2^{l-1}2^{r_2-l+1} = 2^{r_2}$ is allocated among items $2^{l-1}+1, \dots, 2^l$ and the rest $T - 2^{r_2}$ among items $1, \dots, 2^{l-s-1}$. Notice that this allocation is optimal for every $l = 1, 2, \dots, r_2+1$. Moreover, this allocation minimizes the last summation in (VII.0) for every $l = 1, \dots, s$ since $x_{1,2} = 2^{r_2}$ is in the lower bound and p is nondecreasing.

Case 2. $T \geq 2^{r_1+1} + 2^{r_2}$. Then $m(l) \geq 2^{l-s-1}$ and $x_{l,i} = 2^{r_2-l+2}$ for $i = 1, \dots, 2^{l-s+1}$. So the amount $2^{l-s-1} \cdot 2^{r_2-l+2} = 2^{r_1+1} = x_{1,1}$ is allocated among items $1, \dots, 2^{l-s+1}$ and the rest $T - 2^{r_1+1}$ among items $2^{l-1}+1, \dots, 2^l$.

This amount is allocated among items $i = 2^{l-1}+1, \dots, 2^l$ according to the solution of Lemma 2. Moreover, $x_{1,2} = T - 2^{r_1+1}$ is the minimum possible, and since p is nondecreasing it minimizes the last summation of (VII.0) for every $l = 1, \dots, s$.

In both cases the solution satisfies (I.1) and so it minimizes (VII.0) when $s = r_2 - r_1 \geq 1$ is given. So we can rewrite (V.1), (V.2), (VI.1) and (VI.2) with $r_2 - r_1$ decreased by 1.

In Case 2 we can replace r_1 by $r_1' \equiv \lfloor \log_2 x_{1,1} \rfloor = r_1 + 1$. In Case 1 constraints (VI.2) can be rewritten as

$$2^{r_2-l} \leq x_{l,i} \leq 2^{r_2-l+1}, \qquad i = 2^{l-1}+1, \dots, 2^l, \quad l = 1, \dots, r+1.$$

Hence, r_2 in (VI.1) and (VI.2) can be replaced by $r_2' \equiv r_2 - 1$.

This procedure is continued until $r_2 = r_1$. From symmetric reasons an equivalent procedure exists when $r_1 > r_2$. Denote $r = r_1 + 1 = r_2 + 1$, and notice that the optimal solution satisfies

$$\left.\begin{aligned} &\sum_{i=1}^{2^l} x_{l,i} = x_{1,1} + x_{1,2} = T \\ &2^{r-l} \leq x_{l,i} \leq 2^{r-l+1}, \quad i = 1, 2, \dots, 2^l \end{aligned}\right\} \quad l = 1, 2, \dots, r.$$

Hence the induction claim that an optimal solution satisfies Program II was proved, and the theorem is valid. ∎

4. The optimal selection strategy. So far we have proved that splitting according to Strategy O is optimal if we restrict ourselves to splitting policies. Here we show that Strategy O is optimal among all selection strategies.

Let p_m be a *restricted information search problem* where the information I is known only in stages $1, 2, \ldots, m$ of the search while only R, i.e., whether an interval is resolved or not, is known afterwards. Any problem is a p_T problem since no more than T splittings are possible in $(0, T]$.

In a p_1 problem, $I(0, T)$ is known in the first stage while only

$$R(t_{i-1}, t_i, I(t_0, \ldots, t_m)), \qquad i = 1, \ldots, m, \quad \text{is known in stages} \quad m = 2, \ldots, T.$$

By Condition A the order of selecting unresolved intervals is not important. The optimality of a splitting policy among all selection strategies is clear since no information is obtained except whether an interval is resolved or not. Hence Strategy O is optimal for P_1.

We continue by induction. Assume that the theorem is valid for p_m problems and consider now a p_{m+1} problem. At stage $m+1$, $I(t_0, t_1, \ldots, t_{m+1})$ and $R(t_{i-1}, t_i, I(t_0, \ldots, t_{m+1}))$ for $i = 1, \ldots, m+1$ are known. Suppose that $(t_{j-1}, t_j]$ for some $j = 1, \ldots, m+1$ is unresolved and selected to be split. By Condition A, the information obtained by splitting other unresolved intervals does not identify the jumps in $(t_{j-1}, t_j]$. Moreover, although such information may change the probability of a subinterval of $(t_{j-1}, t_j]$ to be resolved, the properties of independence, monotonicity and concavity are valid by Condition B. Hence splitting $(t_{j-1}, t_j]$ is a p_1 problem (since the only information available after the first splitting is whether an interval is resolved or not) and by Lemma 2 the optimal splitting of $(t_{j-1}, t_j]$ is according to Strategy O. Notice that the optimal splitting is independent of $I(t_0, \ldots, t_m)$ except for the information that $(t_{j-1}, t_j]$ is unresolved. Hence the optimal solution of the p_{m+1} problem can be obtained by solving a p_m problem. By the induction assumption a p_m problem is optimally solved by Strategy O. Hence also a p_{m+1} problem is optimally solved by Strategy O. This completes the proof since p_T is the original search problem on $(0, T]$.

References

[1] Cameron, S. H. and Narayanamurthy, S. G. (1964). A Search Problem. *Oper. Res.* 12 623–629.

[2] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Volume II, second edition. John Wiley & Sons, Inc.

[3] Gal, S. (1974). A Discrete Search Game. *SIAM J. Appl. Math.* 27 641–648.

[4] Hassin, R. A Dichotomous Search for a Geometric Random Variable. To appear in *Oper. Res.*

[5] ——— and Megiddo, N. Optimal Search for All the Jumps of a Monotone Step Function. Working paper, Tel Aviv University.

[6] Hu, T. C. and Tucker, A. C. (1971). Optimal Computer-Search Trees and Variable-Length Alphabetical Codes. *SIAM J. Appl. Math.* 21 514–532.

[7] Knuth, D. E. (1973). *Fundamental Algorithms*, Vol. 3. Addison-Wesley, Reading, Massachusetts.

[8] Morris, R. (1969). Some Theorems on Sorting. *SIAM J. Appl. Math.* 17 1–6.

[9] Murakami, S. (1974). A Dichotomous Search. *J. Oper. Res. Soc. Japan* 14 127–142.

[10] ———. (1976). A Dichotomous Search with Travel Cost. *J. Oper. Res. Soc. Japan* 19 245–254.

[11] Wagner, W. M. (1975). *Principles of Operations Research*, second edition. Prentice-Hall, Inc.

HASSIN: DEPARTMENT OF STATISTICS, TEL AVIV UNIVERSITY, TEL AVIV, ISRAEL

HENIG: FACULTY OF MANAGEMENT, TEL AVIV UNIVERSITY, TEL AVIV, ISRAEL