OPERATIONS RESEARCH GROUP

A NOTE ON SEQUENTIAL SEARCH

by

Morton Klein

Columbia University

School of Engineering and Applied Science

NEW YORK, N. Y. 10027

Operations Research Group
Columbia University
New York, New York

A NOTE ON SEQUENTIAL SEARCH

by

Morton Klein

Technical Report Number 36
Aug. 30, 1967

This research was supported by the Army, Navy, Air Force and N.A.S.A. under a contract administered by the Office of Naval Research

Contract Number Nonr 266(55)
Task Number NR 042-099

A Note on Sequential Search*

by

Morton Klein
Columbia University

In [6] Dobbie suggested that problems involving the optimal distribution of search effort for moving objects deserved more attention than they had received in the development of search theory†. Among the several target motion possibilities which he mentioned as being of interest were: (a) "Target motion independent of position and drawn randomly from a population known to the searcher," and (b) "Target motion drawn randomly from a known population that is a function of target position."

Our purpose is to indicate how certain search problems involving such target behavior can be formulated (and consequently solved) by the use of appropriate Markovian decision models. A general description of these problems is given below.

Suppose that an object moves about within a finite number of regions, one per time period, according to known probabilitistic laws (made explicit later). A single searcher, using a detection system whose effectiveness is a function of the amount

---

* This research was supported by the Army, Navy, Air Force, and NASA under a contract administered by the Office of Naval Research; Contract Nonr 266(55). Reproduction in whole or in part is permitted for any purpose of the United States Government.

† An extensive bibliography has been assembled by Enslow [7].

of effort used and the region searched, checks one region at a time until the object is found, his effort budget is exhausted, or he decides that it is "uneconomical" to continue. The problem is to find an optimal sequential search policy; i.e., one which tells the searcher, at each point in time, whether to search, where to search, and how much effort to use.

Definition of the term "optimal policy" will, of course, vary with problem settings. The most frequently used measures of effectiveness in search theory problems appear to be those of maximizing the probability of detecting the target with a given effort or of minimizing the expected search costs needed to find the target. Other possible objectives are to minimize the expected number of periods until the target is found, or, when relevant, to minimize total expected losses until the target is found (where losses may be in different units than search costs). Also, when one of the above is used it may be appropriate to use some or all of the others in the form of constraints.

Markovian decision models will be suggested for two variations of this search problem. The first is characterized (a) by assuming target movements which are independent of the target's location and (b) by assuming that the searcher does not know the target's location until the point at which it is found. In the second model, we assume that the target, if undetected, leaves a trace which is detected by the searcher. Thus, the searcher always knows the target's last location. Here, we also assume that the target's movement is a function of the last region he visited. In both models, we assume that the target's movements are governed by probability laws, known to the searcher. We also assume that the searcher is "noisy," enabling the target to base his movements on knowledge of the searcher's location at the end of each period.

Other sequential search problems have been discussed using closely related models. In [5] a Markovian decision model was used for a non-terminating search problem in which each region had to be visited with preset long run frequencies. Also, Norris [12] studied the structure of some special Markovian search games involving both motionless and moving targets; for the most part he dealt with two region problems and did not consider effort distribution decisions. A sequential search problem for a stationary target was examined by Neuts in [11].

MARKOVIAN SEARCH MODELS

We now give more precise descriptions of the problems and their formulations within the Markovian decision framework.

In both problems, we assume that there are a finite number of regions labeled $0, 1, \ldots, L$. The process starts with the searcher in region 0 (his base) and the object in any region. The budget, of size B, consists of a finite number of discrete units. Travel costs between successive search locations may be counted in the same units.

Model 1: Two classes of states will be used for the Markov chain to be constructed. Members of the first carry labels of the form i_b^0, indicating that region i has just been searched, unsuccessfully (0), and b units of budget remain for future use. Labels for the second class are the same except that a superscript (1) is used to indicate that the object has been found; all such states are stopping states for the process as are other states with labels i_0^0, since the subscript implies that the budget has been used up. Thus, the state space of the decision process is

$$S = \{i_b^\alpha : i = 0,\dots,L;\ b = 0,\dots,\ B-1;\ \alpha = 0,\ 1\} \cup \{\ 0_B^0\ \}$$

where 0_B^0 is its initial state.

We assume that the target discovers the searcher's location at the end of each period. His evasion strategy, based on this information, is assumed to be randomized and represented in the form of a stochastic matrix $H = \{h_{ij}\}$; i.e., $h_{ij} \geqq 0$ and $\sum_{j=0}^{L} h_{ij} = 1$, where i denotes the searcher's current location and j the target's next. Thus, corresponding to each searcher position (i) the target moves to j with probability h_{ij}. One aspect of this assumption is that the target's ability to move is independent of its location. This may not be true of the searcher's mobility.

Now, suppose that the effectiveness of the searcher's detection system depends on the region searched and the amount of effort used. Thus, let $V = \{v_j(e): j = 0,\dots,L\}$ represent the detection system, where $v_j(e)$ is the probability that a search of region j using effort e will find a target if it is in the region.

After each determination of the current state of the decision process, say i_b^{α}, the searcher chooses a decision, j_e, from a finite set $K(i_b^{\alpha})$, i.e., the searcher chooses the next region to be examined (j) and the amount of effort to be used $(e = 1,2,\ldots,b)$. We assume that the decision is made with probability $d(i_b^{\alpha}, j_e)$. Thus, the probability that the system is in some particular next state, say j_f^{β}, is a function of the current state, the decision, and whether the target is present. Since none of these factors depends on anything which has happened in the past, except as these events might be refledted in the current state i_b^{α}, the sequence of successive states that the process follows is a Markov chain. The process is controlled by a randomized statdionary decision rule $D = \{d(i_b^{\alpha}, j_e)\}$ where $d(\quad) \geqq 0$ and $\sum_{j_e} d(i_b^{\alpha}, j_e) = 1$. In general, the problem is to select an optimal rule D from the class of all randomized stationary rules.*

Let A represent all states in which the target is found, that is,

$$A = \{i_b^1 : i = 0, \ldots , L ; b = 0, \ldots , B-1\}$$

and let G contain all states in which the budget is exhausted, i.e.,

$$G = \{i_0^0 : i = 0, \ldots , L\} .$$

Then $T = A \cup G$ is the complete set of stopping states for the

* Derman has shown in [2] that we can restrict our attention to this class of rules.

chain. We suppose that the process starts in state 0_B^0 with probability equal to one. The transition probabilities for the controlled chain, $p(i_b^\alpha, j_f^\beta)$ follow. Let $r_{ij} = 1,2,\ldots$ be the travel effort needed to go from i to j; then, for all integers $b,f : 0 \leqq f = b - e - r_{ij} < b \leqq B$,

$$p(i_b^0, j_f^1) = h_{ij}\, v_j(e)\, d(i_b^0, j_e), \quad i_b \in S\text{-}T, \quad j_f^1 \in A,$$

and

$$p(i_b^0, j_f^0) = (1 - h_{ij})\, d(i_b^0, j_e), \quad i_b^0 \in S\text{-}T, \quad j_f^0 \in S\text{-}A$$

As described, the chain is absorbing by virtue of the stopping states T. For computational purposes, it is convenient to make the chain cyclic by forcing it to return to its initial state whenever T is reached. Hence, we set

$$p(i_b^\alpha, 0_B^0) = d(i_b^\alpha, 0_B^0) = 1, \qquad \text{for} \quad i_b^\alpha \in T$$

It is easy to see that this new chain consists of, at most, one ergodic class of states.

We may now develop formulas for the suggested criterion of constraint functions. Let $C(i_b^\alpha, j_e)$ be the cost if the process is in state i_b^α at the end of a period and decision j_e is made. That is,

$$C(i_b^\alpha, j_e) = e + r_{ij}, \qquad i_b^\alpha \in S\text{-}T,$$

and

$$C(i_b^\alpha, 0_B^0) = 0 \quad , \qquad i_b^\alpha \in T$$

Then the total expected cost $Q(D)$ is

$$Q(D) = E \sum_{t=0}^{\tau(D)} C = E\{\tau(D)\} \; E \; C$$

where $\tau(D)$ is the (random) number of periods taken by the process to reach a stopping state using a specific rule D.

Let $\{ \Pi(i_b^\alpha): i_b^\alpha \in S\}$ represent the (unique) steady state probabilities of the controlled chain. Then the total expected cost can be written in the form

$$(1) \quad Q(D) = [(1/\Pi(0_B^0)) - 1] \sum_{i_b^\alpha} \sum_{j_e} \Pi(i_b^\alpha) \; d(i_b^\alpha, j_e) \; C(i_b^\alpha, j_e)$$

where, from Markov chain theory $1/\Pi(0_B^0)$ is the mean recurrence time for state 0_B^0 and

$$(2) \quad E\{ \tau(D) \} = [1/\Pi(0_B^0)] - 1$$

is the expected duration of the search.

A successful search terminates in state class A; hence it is easy to see that the probability of a successful search, using rule D, is

$$(3) \quad P(D) = (1/\Pi(0_B^0)) \sum_{i_b^\alpha \in A} \Pi(i_b^\alpha) \quad .$$

Various search problem formulations are now at hand, Some examples are indicated:

(a) minimize $Q(D)$ subject to $P(D) \geqq \theta$;

(b) maximize $P(D)$

(c) minimize $E\{\tau(D)\}$ subject to $P(D) \geqq \theta$ and $Q(D) \leqq \Gamma$

(d) maximize $P(D)$ subject to $E\{\tau(D)\} \leqq \Lambda$ and $Q(D) \leqq \Gamma$

Problems without constraints (e.g., (b) above) may be solved using dynamic programming computational methods [9], [10]. However, when restrictions such as those in the other formulations are involved, it is more convenient to transform these problems into linear programs using methods such as were given in [2] for stochastic shortest route problems.

<u>Model 2</u>: As mentioned earlier, here, we assume that an undetected target leaves some trace which is detected by the searcher just after a search has been completed and the target has moved to its next location. We also assume that, at the start of the search, the target is known to be in the region g (say). Thus, the states of the Markov chain are labeled $(i_b^\lambda : j)$ where j is the target's known location after region i has been searched and $(Q_B^0 : g)$ is the initial state.

Now, the second kind of target motion, dependent on target position, can be incorporated. We assume that the target's moments are described by a stochastic matrix

$$Y = \{y(i,j;k) : i,j,k = 0,\ldots,L\} \text{ where } \sum_{k \in R(j)} y(\quad) = 1$$

Thus, the conditional probability that the target's next state is (k) depends on the last location of the searcher (i) and the target (j). The set R(j) represents regions accessible to the target in one period from location j.

As in the first model, given V, Y, and a rule D, it is easy to see that the sequence of successive states form a Markov chain consisting of at most one ergodic class. Expressions for criterion and constraint functions similar to those used in the first model can be written for this chain.

It is now also possible to give a more explicit accounting of losses due to an undetected target than was possible in Model 1. Since the target's location is now part of the state space description, a loss function, perhaps in different units than search or effort costs, can be defined. A typical loss function would have zero value for all stopping states associated with a successful search and various positive values for the other states. Its expected value would have a linear form similar to (1).

ADDITIONAL REMARKS

In some search problems, the time permitted for the search process may be limited, say, to N periods. A simple device for handling such a constraint is to enlarge the state space of the process by adding a time counting term $n = 0,1,\ldots,N$ to the state labels as in [4]. This keeps track of the number of periods taken to reach the current region and budget status of the process. All states carrying the label N would then be stopping states. This also permits the inclusion of time-dependent target movements.

Optional stopping can be included in either model by the addition of an artificial state, say *, which would be included in the set of stopping states. In terms of the first model, the transition probabilities for the chain related to this state would be set as $p(i_b^\alpha, *) = d(i_b^\alpha, *)$, for $i_b^\alpha \in S-T$. The cost associated with such a transition might be the travel cost incurred by the searcher to return to base, plus the expected loss associated with such an action.

Simpler Markovian decision models can be formulated if the bound on the budget is relaxed somewhat and changed from a deterministic to some stochastic form. For example, this could be done by imposing either a constraint on the expected total effort used or a probabilistic bound on actual total effort used. In such cases the budget term in the state labels can be dropped and the dimensionality of the state space reduced by the multiplicative factor B. This also would result in a corresponding reduction in the number of restraining equations in a linear programming formulation of such a problem. In order to be sure that the process would terminate, some constraint, such as an upper bound on the expected time to find the target or on the expected total cost, would have to be imposed. If optional stopping is included, this could also be done by setting $p(i^\alpha, *) = d(i^\alpha, *) = \delta$, for all $i^\alpha \in S-T$, where δ is a small positive number; such restrictions would not only assure us that the chain will only have one ergodic class, but that it will also be irreducible.

In each of the proposed models, we have assumed that the stochastic laws governing the target's movements were given. Since this may not be the case, the question of finding an optimal search rule against an unknown target strategy is of interest. A computational investigation against a limited number of target strategies is, of course, one way of treating such a problem; however, this would rapidly become impractical if one wanted to sweep out a reasonably sized class of such strategies. Algorithms proposed by Charnes and Schroeder[1], and Hoffman and Karp [8], for solving multimove stochastic games may, under certain circumstances, provide a useful approach. For our purposes, a stochastic game is one in which both searcher and target seek optimal strategies, and in which the cost incurred by the searcher can be construed as the target's gain. In these terms a minimum cost Markovian search problem is a stochastic game against a "dummy", since such a problem has the target's strategy specified. Assuming, as indicated, that the game is zero-sum and that there are no added constraints, the Hoffman-Karp algorithm would involve solving an alternating sequence of Markovian search problems* and standard (single move) games. The computation would start with a Markovian decision problem against an arbitrary target strategy; then, given an optimal search rule, another target strategy is generated by solving a certain standard game, etc. The Charnes-Schroeder algorithm which involves solving a sequence of linear programs, can be used for problems involving certain additional constraints.

* A mild requirement is that each such problem should involve a controlled chain with an irreducible state space.

REFERENCES

1. Charnes, A., and R. G. Schroeder, "On Some Stochastic Tactical Antisubmarine Games," Naval Research Logistics Quarterly, Vol. 14, No. 3 (1967).

2. Derman, C., "On Sequential Decisions and Markov Chains," Management Science, Vol. 9, No. 1 (1962).

3. Derman, C., "Optimal Replacement under Markovian Deterioration with Probability Bounds on Failure," Management Science, Vol. 9, No. 3 (1963).

4. Derman, C., and Klein, M., "Some Remarks on Finite Horizon Markovian Decision Models," Operations Research, Vol. 13, No. 2 (1965).

5. Derman, C., and Klein, M., "Surveillance of Multi-Component Systems," Naval Research Logistics Quarterly, Vol. 13, No. 2 (1966).

6. Dobbie, J. M., "Search Theory: A Sequential Approach," Naval Research Logistics Quarterly, Vol. 10, No. 4 (1963).

7. Enslow, P. H., Jr., "A Bibliography of Search Theory and Reconnaissance Theory Literature," Naval Research Logistics Quarterly, Vol. 13, No. 2 (1966).

8. Hoffman, A. J., and Karp, R. M., "On Nonterminating Stochastic Games," Management Science, Vol. 12, No. 5 (1966).

9. Howard, R. A., Dynamic Programming, Wiley, New York, 1960.

10. MacQueen, J., "A Modified Dynamic Programming Method for Markovian Decision Problems," J. Math. Anal. and Appl., Vol. 14, No. 1 (1966).

11. Neuts, M. F., "A Multi-Stage Search Game," J. of SIAM, Vol. 11, No. 2 (1963).

12. Norris, R. C., "Studies in Search for a Conscious Evader," Lincoln Lab., M.I.T. Tech. Report No. 279 (A D 294 832) 14 Sept. 1962 .

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Columbia University

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
A NOTE ON SEQUENTIAL SEARCH

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Technical Report

5. AUTHOR(S) *(First name, middle initial, last name)*
KLEIN, Morton

6. REPORT DATE
August 30, 1967

7a. TOTAL NO. OF PAGES
14

7b. NO. OF REFS
12

8a. CONTRACT OR GRANT NO.
Nonr 266(55)

b. PROJECT NO.

c. Task No. NR 042-099

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
Operations Research Group, Columbia University, School of Engineering and Applied Science

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT
Releasable without limitations on dissemination.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Logistics and Mathematical Statistics Branch, Office of Naval Research
Washington, D.C. 20360

13. ABSTRACT

Optimal distribution of search effort problems for certain kinds of moving targets are considered. It is shown that they can be formulatedby the use of appropriate Markovian decision models.

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Distribution of Effort						
Sequential Search						
Markovian decision process						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)

Security Classification