CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11

LIMITATIONS IN REPRODUCTION QUALITY

ACCESSION # AD606938

☒ 1. WE REGRET THAT LEGIBILITY OF THIS DOCUMENT IS IN PART UNSATISFACTORY. REPRODUCTION HAS BEEN MADE FROM BEST AVAILABLE COPY.

☐ 2. A PORTION OF THE ORIGINAL DOCUMENT CONTAINS FINE DETAIL WHICH MAY MAKE READING OF PHOTOCOPY DIFFICULT.

☐ 3. THE ORIGINAL DOCUMENT CONTAINS COLOR, BUT DISTRIBUTION COPIES ARE AVAILABLE IN BLACK-AND-WHITE REPRODUCTION ONLY.

☐ 4. THE INITIAL DISTRIBUTION COPIES CONTAIN COLOR WHICH WILL BE SHOWN IN BLACK-AND-WHITE WHEN IT IS NECESSARY TO REPRINT.

☐ 5. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED, DOCUMENT WILL BE AVAILABLE IN MICROFICHE ONLY.

☐ 6. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED DOCUMENT WILL NOT BE AVAILABLE.

☐ 7. DOCUMENT IS AVAILABLE IN MICROFICHE ONLY.

☐ 8. DOCUMENT AVAILABLE ON LOAN FROM CFSTI ( TT DOCUMENTS ONLY).

☐ 9.

NBS 9 64

PROCESSOR: [illegible]

AD 606938

# ON k-TH BEST POLICIES


Richard Bellman
Mathematics Division
Robert Kalaba
Engineering Division
The RAND Corporation



P-1417

3 July 1958

Revised 26 February 1960

## SUMMARY

It is shown how the functional equation technique of dynamic programming can be used to determine the optimal, second best, third best, etc., policies for various deterministic and stochastic multistage decision processes.

This is of importance in various problems in combinatorial analysis, network and switching theory, feedback control, and sensitivity analysis. A routing problem is discussed in some detail.

# ON K-TH BEST POLICIES

Richard Bellman
Robert Kalaba

## 1. Introduction

In recent years, a good deal of effort has been devoted to the study of the theory of multistage decision processes, or dynamic programming, see [1]. The emphasis has been upon analytic determination of optimal policies and upon the numerical determination of these policies and the associated return functions through the use of certain algorithms carried out by means of high speed computers.

In a number of situations where only a finite number of possible decisions are possible, there is no question as to the existence of an optimal policy. However, if the number of possibilities is large, then no straightforward enumeration of cases is feasible, and one is forced to develop more elegant, if less simple, techniques. In the course of doing this, the question arises as to whether or not it is possible to determine not only the optimal policy, but the next best policy, and so on, i.e., the preferred suboptimal policies.

Not only is this a challenging mathematical question, but as we shall discuss below, it has significance in connection with "sensitivity analysis" and a variety of network problems.

Before considering the general problem, we shall discuss an interesting particular problem, that of "optimal routing."

## 2. An Optimal Routing Problem

Consider a network, plane or otherwise, consisting of $N$ nodes and interconnecting links. Associate with any two nodes, the i-th and j-th, a quantity, $t_{ij}$, which we can for intuitive purposes call the time required to travel from $i$ to $j$ along the connecting link.

It is important to keep in mind that $t_{ij}$ need not equal $t_{ji}$, and that some of these quantities need not be finite. The topological meaning of this last comment is that $i$ and $j$ need not be connected. Finally, $t_{ij}$, the time of traverse, need not necessarily be proportional to the actual "physical distance" between the nodes $i$ and $j$.

If we think of node $i$ as the i-th state of a system which can only be in one of $N$ states, and if $t_{ij}$ is taken to be the energy required to transform the system from state $i$ to state $j$, then we are seeking the control decisions to be made in order to bring the system from an initial state $i$ to a desired terminal state $N$ with minimal expenditure of energy. This is a fundamental problem of automatic control theory.

The problem of tracing a path of shortest "time" between two given points of the network, 1 and $N$, has been considered by a number of authors. Some published results are contained in Minty [9], Ford [6], Dantzig [4,5], and Bellman [2]. Bock, Kantner and Haynes [3], have discussed the determination of the k-th shortest path, as have Hoffman and Pavley [7]. For a general discussion of this topic and related optimization problems, see Kalaba [8].

Our aim here is to discuss this latter problem using the functional equation technique of dynamic programming.

Although the original question is that of tracing minimal paths from 1 to N, we imbed this problem within the family of problems requiring the determination of minimal paths from a generic point $i$ to the fixed point N. This apparent complication of the problem enables us to employ functional equations. First we determine shortest paths, then second shortest.

We introduce the sequence of quantities $\{u_i\}$, where

(2.1) $u_i$ = the time required to go from $i$ to N using an optimal policy, $i = 1,2,\ldots,N-1$,

$u_N = 0$.

Observe now that if the initial point is $i$ and if the initial decision is to go directly from $i$ to $j$, then the remainder of the route must certainly be selected to minimize the time required to go from $j$ to N. This is an application of the principle of optimality [1].

We are led by this observation to the system of equations

(2.2) $$u_i = \min_{j \neq i} (t_{ij} + u_j), \quad i = 1,2,\ldots,N-1,$$

$$u_N = 0.$$

Although these equations are interlinked in such a fashion that they cannot be solved recursively, there are several quite efficient ways of obtaining the solution, discussed in [2,8].

Here we shall merely observe that if we define the new sequence $\{u_i^k\}$ by means of the relations

$$(2.3) \qquad u_i^0 = t_{iN}, \quad i = 1,2,\ldots,N-1,$$
$$u_N^0 = 0,$$

and

$$(2.4) \qquad u_i^{k+1} = \min_{j \neq i} (t_{ij} + u_j^k), \quad i = 1,2,\ldots,N-1,$$
$$u_N^{k+1} = 0,$$

for $k = 0,1,2,\ldots,$ then the sequence $\{u_i^k\}$ will converge in a monotonically decreasing fashion to the sequence $\{u_i\}$ as $k \to \infty$. In actuality, it is easy to see that $k$ need never be determined beyong the value $N - 2$.

Since only additions and comparisons are required, and since only a small memory is needed, this is a feasible computing scheme for either hand or machine techniques.

With this background, let us treat the problem of determining a second shortest path from 1 to N. Let us introduce the new notation

$$(2.5) \qquad \min_k(x_1,x_2,\ldots,x_N) = \text{the k-th smallest value of the quantities } x_i.$$

This function is not defined for all $k$: for example, if all the $x_i$ are equal, there is no second smallest value.

Let us introduce the quantities $v_i$ defined as follows:

(2.6) $v_i$ = the second shortest path length from $i$ to $N$,

$i = 1,2,\ldots,N-1$, if it exists $(v_i < u_i)$;

$v_N = 0$.

In order to obtain equations connecting the various members of the sequence $\{v_i\}$, we argue as follows. If a path from $i$ to $N$ is to be second shortest, then whatever the initial choice, the continuation must be either a shortest path or a second shortest path. It follows that $v_i$ must be equal to one of the expressions

$$(2.7) \qquad \min_{j \neq i} {}_2 (t_{ij} + u_j),$$

$$\min_{j \neq i} {}_1 (t_{ij} + v_j).$$

Since it must equal the smaller of these, we obtain finally the desired relation

$$(2.8) \qquad v_i = \min \begin{bmatrix} \min_{j \neq i} {}_2 (t_{ij} + u_j) \\ \\ \min_{j \neq i} {}_1 (t_{ij} + v_j) \end{bmatrix}.$$

Once the sequence $\{u_i\}$ has been computed, the sequence $\{v_i\}$ can be determined using successive approximations in the fashion outlined above.

## 3. General Discrete Deterministic Processes

Let us generalize the foregoing considerations by considering dynamic programming processes of the following special type:

(3.1) (a) The state of the system is specified by a finite dimensional vector, the components of which can assume only a finite set of values.

(b) At each stage, we have a choice of only a finite set of decisions.

The problem of determining an N-stage policy which maximizes a prescribed function of the final state is then of completely finite nature, and it is sensible to ask not only for an optimal policy, but also a next best policy, and so on. We consider all policies leading to a maximum return as optimal or first best, all policies leading to a return that is less then optimal but at least as great as all others as second best, and so on. Why we are interested in ordering policies will be discussed below in Section 6.

Let $g(x)$ denote the criterion function measuring the value of the final state, and let $\{T_i(x)\}$ denote the set of allowable decisions, resulting in transformations of the state of the system at each stage. Then, if we introduce the function

(3.2) $f_N(x)$ = the return from an N-stage process obtained using an optimal policy, starting with a system in state $x$, $N = 1,2,\ldots,$

we obtain in the usual fashion the relations

$$(3.3) \qquad f_1(x) = \max_i g(T_i(x)),$$

$$f_N(x) = \max_i f_{N-1}(T_i(x)), \quad N = 2,3,\ldots.$$

Next, let us introduce the functions

(3.4) $f_N^{(k)}(x)$ = the return from an N-stage process with the system initially in state $x$, using a k-th best policy. A k-th best policy produces a return which is <u>smaller</u> than all 1-st, 2-nd, ..., (k − 1)-st best policies, but which is at least as great as the return produced by all other policies.

In particular, we have

$$f_N(x) = f_N^{(1)}(x). \tag{3.5}$$

Another application of the principle of optimality leads to the relations

$$f_N^{(k)}(x) = \max_1{}_k \left\{ f_{N-1}^{(1)}(T_1(x)), f_{N-1}^{(2)}(T_1(x)), \ldots, f_{N-1}^{(k)}(T_1(x)) \right\}, \quad N = 2, 3, \ldots, \tag{3.6}$$

$$f_1^{k}(x) = \max_1{}_k\, g(T_1(x)). \tag{3.7}$$

It follows that the terms in the sequence $f_N^{(1)}(x), f_N^{(2)}(x), \ldots$, may then be determined recursively, for suitable ranges of $N$ and $x$. At the same time, the appropriate decision in a k-th best policy is determined in terms of the state of the system and the time remaining before termination of the process.

As $k$ increases, the dimensionality of the problem increases. The memory requirements are directly proportional to $k$.

## 4. Paths Most Likely to be Available in a Network

Once again, let us consider a network of $N$ nodes numbered from 1 to $N$. The link from $i$ to $j$ is assumed to be available for service with probability $p_{ij}$, and the availabilities for the various links are assumed to be independent. Consequently, we are involved in a stochastic situation.

We first show how to determine which paths from $i$ to $N$ have greatest probability of being available for service, and then indicate how the second, third, and other greatest paths can be calculated. This problem is an important one in telephony where one point in a switching network can be connected to another via several different paths; if the most likely paths are unavailable, then the second most likely may be scanned, etc.

Let

(4.1) $u_i$ = the probability that a path from $i$ to $N$ is available for service, the path being an optimal one.

Once again, upon employing the principle of optimality [1], we have

$$(4.2) \qquad u_i = \max_{j \neq i} p_{ij} u_j, \quad i = 1,2,\ldots,N-1.$$

These equations can be resolved using successive approximations, and, as before, the method is a method of exhaustion; i.e., it converges after a finite number of steps bounded in advance.

Next, let us introduce

(4.3) $\quad v_j$ = the probability that a second best path is available for service.

Then we have

$$(4.4) \qquad v_i = \max \begin{cases} \max\limits_{j \neq i} p_{ij} v_j \\ \max\limits_{j \neq i}{}_2\, p_{ij} u_j \end{cases}, \quad i = 1, 2, \ldots, N-1.$$

The solution may now be obtained by successive approximations.

## 5. Stochastic Decision Processes

Let us return to the deterministic decision process discussed in Section 3. We wish to modify this process, though, in that we shall now assume that the result of making decision i, with the system in state x, is no longer precisely known. In its place, we merely know that there is a certain probability that the system is transformed into the state y, which we denote by dG(y;x,i). The objective of the process will now become that of maximizing the expected value of a given function g(x) of the final state.

Let

(5.1) $\quad f_N(x)$ = the expected value of the return from an N-stage process, beginning in state x, and using an optimal policy.

For a one-stage process, we find

$$(5.2) \qquad f_1(x) = \max_i \int g(y)\,dG(y;x,i),$$

where the integration is over all states $y$. For the N-stage process, we have

$$(5.3) \qquad f_N(x) = \max_i \int f_{N-1}(y)\,dG(y;x,i), \quad N = 1,2,\ldots.$$

For the determination of suboptimal policies, we introduce the functions

(5.4) $f_N^{(k)}(x)$ = the expected value of the return from an N-stage process, beginning in state $x$, and using a k-th best policy.

Using the same reasoning as earlier, we are led to the formulas

$$(5.5) \qquad f_N^{(k)}(x) = \max_k \begin{cases} \int f_{N-1}^{(1)}(y)\,dG(y;x,i), \\ \int f_{N-1}^{(2)}(y)\,dG(y;x,i), \\ \int f_{N-1}^{(k)}(y)\,dG(y;x,i), \end{cases}$$

$$(5.6) \qquad f_1^{(k)}(x) = \max_k \int g(y)\,dG(y;x,i),$$

which permit the recursive determination of the functions defined in Equation (5.4) along with the appropriate policies. In effect, we have to compute a sequence of functions of the variable, $x$, which represents a basic simplification if $x$ is of dimension three or less.

## 6. Sensitivity Analysis

Let us now explain the significance of the foregoing results in connection with the actual solution of physical problems.

As we know, whenever we construct a mathematical model of a real situation, we make certain compromises, or approximations, as they are more diplomatically called. It follows that an optimal solution to a mathematical problem may not be an optimal solution to the engineering or economic problem under consideration. There are now two alternatives. We can either complicate the mathematical model to remove this difficulty, or we can look for approximate solutions of the mathematical problem which more nearly solve the physical problem. Which step we take depends upon the available time, the cost, the utility of improved solutions, and so on.

If we are interested in finding approximate solutions of the mathematical problem, then the foregoing techniques are useful.

In somewhat the same context, we are frequently forced, because of the limited memories of computers and their slowness of computation, to use much coarser grids in both space and time than we would like. Sometimes, we are forced to retain this type of solution for want of better, while, occasionally, we can use these crude solutions as initial approximations to be successively improved.

One way of evaluating the meaningfulness of a coarse approximation is to examine the behavior of the neighborhood of the optimal policy. If there is too drastic a change, we can be assured that the formulation is too crude. If the change is slight as we go from optimal to second best, from second best to third best, and so on, then there is a chance that we are getting worthwhile results.

The numerical solution of any physical problem must always be subjected to a stability, or sensitivity, analysis of this type.

## REFERENCES

1. R. Bellman, Dynamic Programming, Princeton Univ. Press, Princeton, N. J., 1957.

2. R. Bellman, "On a routing problem," Q. Appl. Math., vol. 16, 1958, pp. 87-90.

3. F. Bock, H. Kantner, and J. Haynes, An Algorithm (the r-th Best Path Algorithm) for Finding and Ranking Paths Through a Network, Armour Research Foundation Research Report, May 19, 1958.

4. G. B. Dantzig, "Discrete-variable extremum problems," Oper. Res., vol. 5, 1957, pp. 266-277.

5. G. B. Dantzig, On the Shortest Route Through a Network, The RAND Corporation, Paper P-1345, 1958.

6. L. R. Ford, Jr., Network Flow Theory, The RAND Corporation, Paper P-923, August 14, 1956.

7. W. Hoffman and R. Pavley, "A method for the solution of the N-th best path problem," J. Assoc. for Computing Machinery, vol. 6, 1959, pp. 506-514.

8. R. Kalaba, On Some Communication Network Problems, The RAND Corporation, Paper P-1325, 1958.

9. G. J. Minty, "A comment on the shortest route problem," Oper. Res., vol. 5, 1957, p. 274.