CHAPTER 8

THEORIES OF BOUNDED RATIONALITY

HERBERT A. SIMON

1. Introduction

Rationality denotes a style of behavior that is appropriate to the achievement of given goals, within the limits imposed by given conditions and constraints. Theories of rational behavior may be normative or descriptive – that is, they may prescribe how people or organizations should behave in order to achieve certain goals under certain conditions, or they may purport to describe how people or organizations do, in fact, behave. This essay will be concerned with the structure of theories of rational behavior, whether they are intended prescriptively or descriptively.

Individual and organizational rationality. A theory of rational behavior may be concerned with the rationality of individuals or the rationality of organizations. In fact, the two bodies of theory are not wholly distinct.[1] One plausible distinction between them is that a theory of organizational rationality must treat the phenomena of goal conflict, while a theory of individual rationality need not. This is only partly correct, for goal conflict may be important in individual as in group behavior – it is a major theme of so-called "dissonance theory" in psychology. (See N. P. CHAPANIS and J. A. CHAPANIS (1964).) A theory of individual behavior microscopic enough to concern itself with the internal organization (neurological or functional) of the central nervous system will have a significant organizational component. A theory of organizational behavior macroscopic enough to treat the organization as a monolith will be a theory of an "individual." Although this chapter will be aimed primarily at understanding individual rationality, I shall not hesitate to use the theory of the firm – classically, the theory of a

---

[1] This point was made by J. MARSCHAK (1955) in his first paper on teams, "Elements for a Theory of Teams." I shall follow his good precedent.

monolithic entrepreneur – as a convenient and enlightening illustrative example.

From the standpoint of this chapter, then, the distinction between individual and organization will not be very important. A more significant taxonomy of theories of rational behavior, for our purposes, differentiates them by the assumptions they make about the "givens" – the given goals and given conditions. Particularly important is the distinction between those theories that locate all the conditions and constraints in the environment, outside the skin of the rational actor, and those theories that postulate important constraints arising from the limitations of the actor himself as an information processor. Theories that incorporate constraints on the information-processing capacities of the actor may be called *theories of bounded rationality*.

Rationality in the classical theory of the firm. The classical theory of the firm in its simplest form provides a useful standard for comparing and differentiating theories of rationality. In the theory of the firm, the given objective is to maximize profits, where profit is defined as the difference between gross receipts from sales and cost of production. The given conditions are two in number:

(I) *the demand function*: the quantity demanded is a function of price:

$$(1) \qquad q_d = D(p), \text{ or } p = D^{-1}(q_d).$$

Since gross receipts equal price times quantity, the demand function determines gross receipts:

$$(2) \qquad R = pq_d.$$

(II) *the cost function*: the cost of production is a function of the quantity produced:

$$(3) \qquad C = C(q_s).$$

If the quantity produced equals the quantity demanded,

$$(4) \qquad q_s = q_d,$$

then the profit, to be maximized, is simply the difference between gross receipts and the cost of production:

$$(5) \qquad \text{Profit} = R - C = pq - C(q),$$

and, under appropriate assumptions regarding differentiability, we will

have for the maximum profit:

(6) $d(R-C)/dq = p+qd(D^{-1}(q))/dq - dC(q)/dq = 0$.

The constraints in this theory, the demand and cost functions, D and C, are both located in the actor's environment. He is assumed to find the solution of equation (6). To do this, he must have perfect knowledge of these constraints, and must be able to perform the necessary calculations – to set the derivative of profit with respect to quantity equal to zero and to solve the resulting algebraic equation.

The limits of rationality. Theories of bounded rationality can be constructed by modifying these assumptions in a variety of ways. *Risk and uncertainty* can be introduced into the demand function, the cost function, or both. For example, certain parameters of one or both of these functions can be assumed to be random variables with known distributions. Then the assumption of the actor's perfect knowledge of these functions has been replaced by the assumption that he has perfect knowledge of their distributions. This change in assumptions may, in turn, make it easier or more difficult to carry out the calculations for finding the optimum – usually it becomes much more difficult than in the corresponding case of certainty.

Another way in which rationality can be bounded is by assuming that the actor has only *incomplete information about alternatives*. Fewer models have been constructed to deal with this situation than with the situation in which he has incomplete information about consequences. However, in certain search models it is assumed that the actor knows the probability distribution of profits in a population of possible alternative actions. Specific actions become available to him – say, by random sampling from this population – as a function of the amount of resources he devotes to search. His task is to find the alternative that maximizes his expected profit net of the search cost. In this class of models, selecting the best alternative from among those already discovered is assumed to be a trivial problem; the decision question has been switched to the question of how much of the actor's resources should be allocated to search.[2]

---

[2] For an example, see STIGLER (1961). Theories of the allocation of resources to search can also be constructed to deal with incomplete information about consequences. Sequential sampling theory falls into this category, for it answers the question: shall I make a decision now, or wait until I have gathered additional information? The question is answered by comparing the incremental cost of enlarging the sample with the expected gain through the resulting average improvement in the decision.

Finally, rationality can be bounded by assuming *complexity* in the cost function or other environmental constraints so great as to prevent the actor from calculating the best course of action. Limits on rationality stemming from this source have not been prominent in classical theories of rational behavior. However, in numerical analysis, the theory of approximation provides analogues, for it is concerned with the rate at which an approximation can be expected to improve as a function of amount of computational effort. By introducing explicitly into that theory the cost of computational effort, it can be transformed into a theory of optimal approximation.

Alternatives to the classical goals. The classical theory can be modified not only by altering the nature of the conditions and constraints, but also by altering the nature of the given goals. Some modern theories of the firm depart from the classical theory, not along any of the dimensions mentioned above, but by postulating different goals from the classical goal of profit maximization.

BAUMOL (1959, pp. 45–53), for example, has developed a model in which the firm maximizes sales subject to the constraint that profit should not be less than a specified "satisfactory" level. According to this theory of Baumol, equation (6) in the classical model should be replaced by:

$$(6') \qquad dR/dq = p + q \, d(D^{-1}(q))/dq = 0,$$

subject to the constraint that

$$(7) \qquad P = R - C \geq P^*.$$

It may be observed that the informational and computational requirements for applying Baumol's theory to concrete situations are not very different from the requirements of the classical model.

This essay will not be concerned with variants of the theory of rationality that assume goals different from profit or utility maximization, except to the extent that there is significant interaction between the assumptions about goals and the assumptions about conditions and constraints. We shall see, however, that this is a very important exception. In actual fact, most of the variants of the theory that make significant modifications in the assumptions about conditions and constraints also call for assumptions about goals that are different from the classical assumptions of profit or utility maximization. The reasons for this interaction will appear as we proceed.

2. Approaches to Rational Choice in Chess

A number of the persons who have engaged in research on rational decision-making have taken the game of chess as a microcosm that mirrors interesting properties of decision-making situations in the real world. The research on rational choice in chess provides some useful illustrations of alternative approaches to rationality.

The problem confronting a chess player whose turn it is to move can be interpreted in either of two ways. First, it can be interpreted as a problem of finding a good (or the best) strategy – where "strategy" means a conditional sequence of moves, defining what move will be made at each successive stage after each possible response of the opponent.

Second, the problem can be interpreted as one of finding a set of accurate evaluations for the alternative moves immediately before the player.

From a classical standpoint, these two problems are not distinguishable. If the player has unlimited computational power, it does not matter whether he selects a complete strategy for his future behavior in the game, or selects each of his moves, one at a time, when it is his turn to play. For the way in which he goes about evaluating the next move is by constructing alternative complete strategies for the entire future play of the game, and selecting the one that promises the best return (i.e., the best return under the assumption that the opponent will also do *his* best to win). This is the approach taken in the von Neumann-Morgenstern theory of games (VON NEUMANN and MORGENSTERN (1953)).

The game-theoretical definition of rationality in chess. As von Neumann and Morgenstern observed, chess is a trivial game. "... if the theory of Chess (i.e., the complete tree of possible games) were really fully known there would be nothing left to play" (*ibid.*, p. 125). Each terminus of the tree of possible games represents a win, loss, or draw for White. Moving backward one branch on the tree, the player whose move it is at that branch can examine the termini to which it could lead by his choice of move, and can choose the move having the preferred terminus. The value of that terminus becomes, then, the value of the branch that leads to it. Working backward in this way, a value – win, lose, or draw for White – can be assigned to each position, and ultimately to each of the initial legal moves for White. Now each player can specify an optimal strategy – a strategy that will guarantee him at least as good an outcome as any other – by specifying which move he would select at each branch point in the tree whenever it is his move.

Unfortunately, as von Neumann and Morgenstern also observed, the triviality of chess offers no practical help to a player in actually choosing a move. "But our proof, which guarantees the validity of one (and only one) of these three alternatives [that the game must have the value of win lose or draw for White], gives no practically usable method to determine the true one. This relative, human difficulty necessitates the use of those incomplete, heuristic methods of playing, which constitute 'good' Chess; and without it there would be no element of 'struggle' and 'surprise' in that game" (*ibid.*).

What "impracticality" means becomes more vivid when we calculate how much search would be involved in finding the game-theoretically correct strategy in chess. On the average, at any given position in a game of chess, there are about 30 legal moves – in round numbers, for a move and its replies, an average of about 10^3 continuations. Forty moves would be a not unreasonable estimate of the average length of a game. Then there would be perhaps 10^{120} possible games of chess. Obviously the exact number does not matter: a number like 10^{40} would be less spectacular, but quite large enough to support the conclusions of the present argument.

Studies of the decision-making of chess players indicate strongly that strong players seldom look at as many as one hundred possibilities – that is one hundred continuations from the given position – in selecting a move or strategy. One hundred is a reasonably large number, by some standards, but somewhat smaller than 10^{120}! Chess players do not consider all possible strategies and pick the best, but generate and examine a rather small number, making a choice as soon as they discover one that they regard as satisfactory (see DE GROOT (1965)).

Before we consider in detail how they do it, let us return to the classical model and ask whether there is any way in which we could make it relevant to the practical choice problem, taking account of the size of the problem space, in a game like chess. One possible way would be to replace the actual problem space with a very much smaller space that approximates the actual one in some appropriate sense, and then apply the classical theory to the smaller approximate space.

This approach was taken in some of the early computer programs for playing chess and checkers. In the Los Alamos program, for example, the computer generated all legal moves, all legal replies to each, and so on, two moves deep. Each of the terminal positions thus generated (about a million in a two-move analysis) was evaluated, and the minimax procedure applied, working backwards, to find the best first move. Thus, a space of about

10^6 elements was substituted for the space of 10^{120} elements that represents the "real" world of chess.

The scheme was approximate, because the actual chess values of the million terminal positions were not known, and could not be known accurately without returning to the space of 10^{120} elements – that is, returning to the game-theoretical analysis of the full game. In place of these unknown true values, approximate values were computed, using rules of thumb that are commonly employed by chess players – conventional numerical values for the pieces, and measures of mobility. Thus, the approximate scheme was not guaranteed to select the objectively best move, but only the move leading to the positions that appeared best, in terms of these heuristic criteria, after an analysis two moves deep. Experience indicates that it is not possible to make such approximate evaluations accurately enough to enable the program to play good chess. The optimal decision in the approximated world is not necessarily even a good decision in the real world.

Satisficing processes in chess thinking. Chess programs now exist that take the alternative course, trying to emulate the human chess player in looking at only a very few continuations. The effectiveness of such a scheme depends critically on three components: the *move generators*, processes that select the continuations to be explored; the *evaluators*, processes that determine how good each continuation is; and the *stop rules*, criteria that determine when the search should be terminated and a move selected.

By scanning a chess position, features of the position can be detected that suggest appropriate moves. To take an extreme case, suppose a chess player discovers, when it is his move, that one of his Pawns attacks the opponent's Queen. Obviously, the capture of the Queen by the Pawn is one move that deserves consideration. It may turn out to be a poor move – another piece will checkmate him, say, if he captures the Queen – but its superficial merits are obvious, and its deficiencies can only be detected by considering it and evaluating it dynamically. A simple process that would generate this move, and others like it, would consist in determining which of the opponent's pieces were attacked by a piece of lesser value, or were undefended and attacked by any piece. Thus, a suitable set of move-generating processes might identify for further analysis all or most of the moves deserving serious consideration. If the generators were ordered appropriately, they might usually identify first the most promising moves, then the ones slightly less promising, and so on.

Possible moves, produced by the move generators, can be evaluated by a combination of static and dynamic criteria. Static criteria are features of the

position, or differences between successive positions. Thus, one of the important static evaluators used by all chess players is the piece count: each piece is assigned a conventional value (say, Pawn = 1, Knight and Bishop = 3, Rook = 5, Queen = 9), and the sums of the values for the two players are compared. In general, if the piece count of one player exceeds that of the other by more than one point (or even, in many cases, by a single point), the player with the higher count can find a winning continuation *unless* the balance is very quickly redressed by a sequence of forceful moves. (Thus, it does not matter being 5 points down if you can capture the opponent's Queen on the next move without further reprisals.)

The short-run tactical considerations are handled by carrying out dynamic analysis of plausible continuations until a position is reached that is sufficiently quiet or "dead" that it can safely be evaluated by means of the static evaluators. These static evaluators are then propagated backwards to the move under consideration by the familiar minimax procedure.

Two kinds of stop rules are needed in a program having this structure: rules to stop exploration at dead positions that can be evaluated statically, and rules to stop the entire process and select a move when a satisfactory one has been found. The former class of stop rules has already been discussed; the latter needs to be examined more closely. If the alternatives in a choice situation are not given, but have to be discovered or invented, and if the number of possible alternatives is very large, then a choice has to be made before all or most of them have been looked at. It was precisely this difficulty in the classical requirement of comparing all alternatives that led to the approach described here. But if all alternatives are not to be examined, some criterion must be used to determine that an adequate, or satisfactory, one has been found. In the psychological literature, criteria that perform this function in decision processes are called *aspiration levels*. The Scottish word "satisficing" (= satisfying) has been revived to denote problem solving and decision making that sets an aspiration level, searches until an alternative is found that is satisfactory by the aspiration level criterion, and selects that alternative (SIMON (1957), Part IV).

In satisficing procedures, the existence of a satisfactory alternative is made likely by dynamic mechanisms that adjust the aspiration levels to reality on the basis of information about the environment. Thus, in a chess-playing program, the initial aspiration level can be set (preferably with a little upward bias) on the basis of a static evaluation of the position. As alternative moves are considered and evaluated by dynamic and static analysis, the evaluation of the position can gradually be reduced until the

best move discovered so far reaches or exceeds in value the aspiration level.

The limits of rationality in chess. In the introductory section of this paper, three limits on perfect rationality were listed: uncertainty about the consequences that would follow from each alternative, incomplete information about the set of alternatives, and complexity preventing the necessary computations from being carried out. Chess illustrates how, in real world problem-solving situations, these three categories tend to merge.

If we describe the chess player as choosing a *strategy*, then his difficulty in behaving rationally – and the impossibility of his behaving as game theory says he should – resides in the fact that he has incomplete information as to what alternatives (strategies) are open to him. He has time to discover only a minute fraction of these strategies, and to specify the ones he discovers only incompletely.

Alternatively, if we describe the chess player as choosing a *move*, his difficulty in behaving rationally lies in the fact that he has only rough information about the consequences of adopting each of the alternatives (moves) that is open to him. It would not be impossible for him to generate the whole set of his legal moves, for they seldom number more than about thirty. However, he can evaluate them, even approximately, only by carrying out further analysis through the immense, branching, move tree. Since only a limited amount of processing time is available for the evaluation, he must allocate the time among the alternative moves. The practical facts of the matter are that it is usually better to generate only a few of the entire set of legal moves, evaluating these rather thoroughly, than it is to generate all of them, evaluating them superficially. Hence the good chess player does not examine all the moves open to him, but only a small fraction of them. (Data presented by DE GROOT (1965) suggest that typically a half dozen to a dozen of a set of thirty legal moves may be generated and explored by the chess player.)

From still a third standpoint, the chess player's difficulty in behaving rationally has nothing to do with uncertainty – whether of consequences or alternatives – but is a matter of complexity. For there is no risk or uncertainty, in the sense in which those terms are used in economics or statistical decision theory, in the game of chess. As von Neumann and Morgenstern observe, it is a game of perfect information. No probabilities of future events need enter the calculations, and no contingencies, in a statistical sense, arise.

From a game-theoretical standpoint, the presence of the opponent does not introduce contingencies. The opponent can always be counted on to do

his worst. The point becomes clear if we replace the task of playing chess with the task of proving theorems. In the latter task, there is no opponent. Nor are there contingencies: the true and the derivable theorems reside eternally in Plato's heaven. Rationality in theorem proving is a problem only because the maze of possible proof paths is vast and complex.

What we refer to as "uncertainty" in chess or theorem proving, therefore, is uncertainty introduced into a perfectly certain environment by inability – computational inability – to ascertain the structure of that environment. But the result of the uncertainty, whatever its source, is the same: approximation must replace exactness in reaching a decision. In particular, when the uncertainty takes the form of an unwieldy problem space to be explored, the problem-solving process must incorporate mechanisms for determining when the search or evaluation will stop and an alternative will be chosen.

Satisficing and optimizing. The terms satisficing and optimizing, which we have already introduced, are labels for two broad approaches to rational behavior in situations where complexity and uncertainty make global rationality impossible. In these situations, optimization becomes approximate optimization – the description of the real-world situation is radically simplified until reduced to a degree of complication that the decision maker can handle. Satisficing approaches seek this simplification in a somewhat different direction, retaining more of the detail of the real-world situation, but settling for a satisfactory, rather than an approximate-best, decision. One cannot predict in general which approach will lead to the better decisions as measured by their real-world consequences. In chess at least, good players have clearly found satisficing more useful than approximating-and-optimizing.

A satisficing decision procedure can often be turned into a procedure for optimizing by introducing a rule for optimal amount of search, or, what amounts to the same thing, a rule for fixing the aspiration level optimally. Thus, the aspiration level in chess might be adjusted, dynamically, to such a level that the expected improvement in the move chosen, per minute of additional search, would just balance the incremental cost of the search.

Although such a translation is formally possible, to carry it out in practice requires additional information and assumptions beyond those needed for satisficing. First, the values of alternatives must be measured in units comparable with the units for measuring search cost, in order to permit comparison at the margins. Second, the marginal productivity of search – the expected increase in the value per unit of search time – must be estimated on some basis or other. If one were designing a chess-playing program, it is

doubtful whether effort spent in attempting to imbed the program in such a dynamic optimizing framework would be nearly as worthwhile as equivalent effort spent in improving the selectivity of the program's move-generating and move-evaluating heuristics.

Another quite different translation between optimizing and satisficing schemes has also been suggested from time to time. A chess program of the "classical" type, which makes optimal decisions in an approximated world, can be regarded as a particular kind of satisficing program, in which "satisfactory" is defined by the approximating procedure that is used. Hence, it is difficult to draw a formal distinction between optimizing and satisficing procedures that is so iron-clad as to prevent either from being reinterpreted in the frame of the other. The practical difference, however – the difference in emphasis that results from adopting one viewpoint or the other – is often very great.

In research on optimizing procedures, considerable attention has been paid to the formal properties of the evaluation functions, to the existence and efficiency of procedures for computing the optimum, and to procedures for reducing uncertainty (e.g., forecasting methods). The nature of the approximations that are necessary to cast real-world problems into forms suitable for optimization, and the means for choosing among alternative approximations, have been less fully and less systematically studied. Much effort, for example, has gone into the discovery of efficient algorithms for solving linear programming problems. Finding an appropriate way of formulating a concrete real-world decision problem as a linear-programming problem remains largely an art.[3]

Research on satisficing procedures has focussed primarily on the efficiency of search – on the nature of the heuristic methods that enable the rare solutions in enormous spaces of possibilities to be sought and found with moderate amounts of search effort. Since moderate changes in heuristics often make order-of-magnitude changes in search effectiveness, highly accurate means for assessing the quality of solutions or the effort required to find them may be relatively unimportant. It probably does not require delicate evaluation functions or stop rules to change a duffer's chess play to a reasonably effective move-choosing program.

---

[3] The work of A. CHARNES and W. W. COOPER (1961) is full of sophisticated examples of this art. See, for instances, Appendix B and Chapter 11 of Volume I.

3. Bounded Rationality in Design

The engineering activities usually called "design" have not been much discussed under the heading of rational decision-making. The reason for this should be clear from the foregoing discussion: classical decision theory has been concerned with choice among *given* alternatives; design is concerned with the discovery and elaboration of alternatives. Our exploration of the microcosm of chess has indicated, however, how the theory of design can be assimilated to a satisficing theory of rational choice. Let me spell the point out a little more fully.

Consider that interpretation of chess which views the task as one of choosing a strategy, and not just a single move. Specifically, consider a situation where a player is searching for a combination (a strategy) that will definitely checkmate his opponent, even though it may require sacrifices of pieces along the way. A chess player will ordinarily not enter into such a course of action unless he can see it through to the end – unless he can *design*, that is, a water-tight mating combination.

As we have seen already, the evaluations and comparisons that take place during this design process are not, in general, comparisons among complete designs. Evaluations take place, first of all, to guide the search – the elaboration of the design itself. They provide the basis for decisions that the design should be elaborated in one direction rather than another. Complete designs (in this case, mating combinations), when they are finally arrived at, are not generally evaluated by comparing them with alternative designs, but by comparing them with standards defined by aspiration levels. In the chess situation, as soon as the player discovers a strategy that guarantees a checkmate, he adopts it. He does not look for all possible checkmating strategies and adopt the best (H. A. SIMON and P. A. SIMON (1962)).

In the design of complex objects – a bridge, say, or an airplane – the process has an even more involved search structure. Here, the early stages of search take place in highly simplified spaces that abstract most of the detail from the real-world problem, leaving only its most important elements in summarized form. When a plan, a schematized and aggregated design, has been elaborated in the planning space, the detail of the problem can be reintroduced, and the plan used as a guide in the search for a complete design.

More than two spaces may be used, of course; there may be a whole hierarchy of planning spaces, leading from a highly abstract and global design to successive specification of detail. At each of these levels of abstrac-

tion, the design process, too, may be differently structured. Since the more abstract spaces tend to be "smoother," it is often possible to use optimization models for planning purposes, reverting to satisficing search models to fill in the detail of the design. Thus, linear programming or dynamic programming may be used for general planning of factory operations, while more heuristic techniques are used for scheduling of individual jobs. In other situations, the overall design process may employ satisficing search procedures, while optimizing techniques may be used to set parameters once the general design has been fixed.[4]

4. Bounded Rationality in Management Science

Most of the formal techniques that constitute the technical backbone of management science and operations research are procedures for finding the best of a set of alternatives in terms of some criterion – that is, they fall in our category of "classical" procedures. Linear and dynamic programming are among the most powerful of these techniques. The dominant approach to problems in this sphere has been to simplify the real-world problems to the point where the formal optimizing models can be used as approximations.

Some industrial problems of a combinatorial sort have not yielded easily to this approach. Typically, the recalcitrant problems involve integer solutions, or, what usually amounts to the same thing, the consideration of possible permutations and combinations of a substantial number of elements. Warehouse location is a problem of this kind. The task is to "determine the geographical pattern of warehouse locations which will be most profitable to the company by equating the marginal cost of warehouse operation with the transportation cost savings and incremental profits resulting from more rapid delivery" (KUEHN and HAMBURGER (1963), p. 643).

A heuristic program devised by KUEHN and HAMBURGER (1963) for locating warehouses has two parts: "(1) the main program, which locates warehouses one at a time until no additional warehouses can be added to the distribution network without increasing total costs, and (2) the bump and shift routine, ..., which attempts to modify solutions ... by evaluating the profit implications of dropping individual warehouses or of shifting them from one location to another" (*ibid.*, p. 645).

---

[4] Some modern semi-automated procedures for the design of chemical processing plants proceed from heuristic techniques for selecting the unit operations and their flow, then employ linear programming to determine the parameters of the system so specified.

This program fits our earlier characterization of design procedures. A possible plan is gradually built up, step by step, through a search procedure, and then possible local modifications are investigated before the final plan is settled upon. In building up the initial plan, locations are tried that are near concentrations of demand, adding at each step the warehouse that produces the greatest cost savings for the entire system. Only a fraction of the possible warehouse sites, which preliminary screening selects as "promising," are evaluated in detail at each stage. Finally, a so-called "bump-shift" routine modifies the programs tentatively arrived at by (1) eliminating warehouses no longer economical because new warehouses have been introduced at later steps of the program, (2) considering shifting warehouses to alternative sites within their territories. The flow diagram of the warehouse location programs, which will serve to illustrate the typical structure of heuristic programs when they are formalized, is shown in Fig. 1.

Kuehn and Hamburger have carried out some detailed comparisons of the heuristic program with optimizing techniques. They conclude that "in theory, a linear programming approach ... could be used to solve the problem. In practice, however, the size and nonlinearities involved in many problems are such that application is not currently feasible" (*ibid.*, p. 658). They attribute the superior performance of the heuristic program to two main causes: "(1) computational simplicity, which results in substantial reductions in solution times and permits the treatment of large-scale problems, and (2) flexibility with respect to the underlying cost functions, eliminating the need for restrictive assumptions" (*ibid.*, p. 656).

Perhaps the technique most widely used in management science to deal with situations too complex for the application of known optimization methods is simulation. The amount of detail incorporated in the simulation of a large system is limited only by computational feasibility. On the other hand, simulation, unaided by other formal tools of analysis, provides no direct means for discovering and evaluating alternative plans of action. In simulation, the trial and error is supplied by the human investigators rather than by the technique of analysis itself (see FORRESTER (1961)).

5. Conclusion

The theory of rational decision has undergone extremely rapid development in the past thirty years. A considerable part of the impetus for this development came, during and since World War II, from the attempt to use formal decision procedures in actual real-world situations of considerable

A HEURISTIC PROGRAM FOR LOCATING WAREHOUSES

1. Read in:

    a) The factory locations.

    b) The M potential warehouse sites.

    c) The number of warehouse sites (N) evaluated in detail on each cycle, i.e., the size of the buffer.

    d) Shipping costs between factories, potential warehouses and customers.

    e) Expected sales volume for each customer.

    f) Cost functions associated with the operation of each warehouse.

    g) Opportunity costs associated with shipping delays, or alternatively, the effect of such delays on demand.

2. Determine and place in the buffer the N potential warehouse sites which, considering only their local demand, would produce the greatest cost savings if supplied by local warehouses rather than by the warehouses currently servicing them.

3. Evaluate the cost savings that would result for the total system for each of the distribution patterns resulting from the addition of the next warehouse at each of the N locations in the buffer.

4. Eliminate from further consideration any of the N sites which do not offer cost savings in excess of fixed costs.

5. Do any of the N sites offer cost savings in excess of fixed costs?

Yes

6. Locate a warehouse at that site which offers the largest savings

No

7. Have all M potential warehouse sites been either activated or eliminated? | No

Yes

8. *Bump-Shift Routine*

    a) Eliminate those warehouses which have become uneconomical as a result of the placement of subsequent warehouses. Each customer formerly serviced by such a warehouse will now be supplied by that remaining warehouse which can perform the service at the lowest cost.

    b) Evaluate the economics of shifting each warehouse located above to other potential sites whose local concentrations of demand are now serviced by that warehouse.
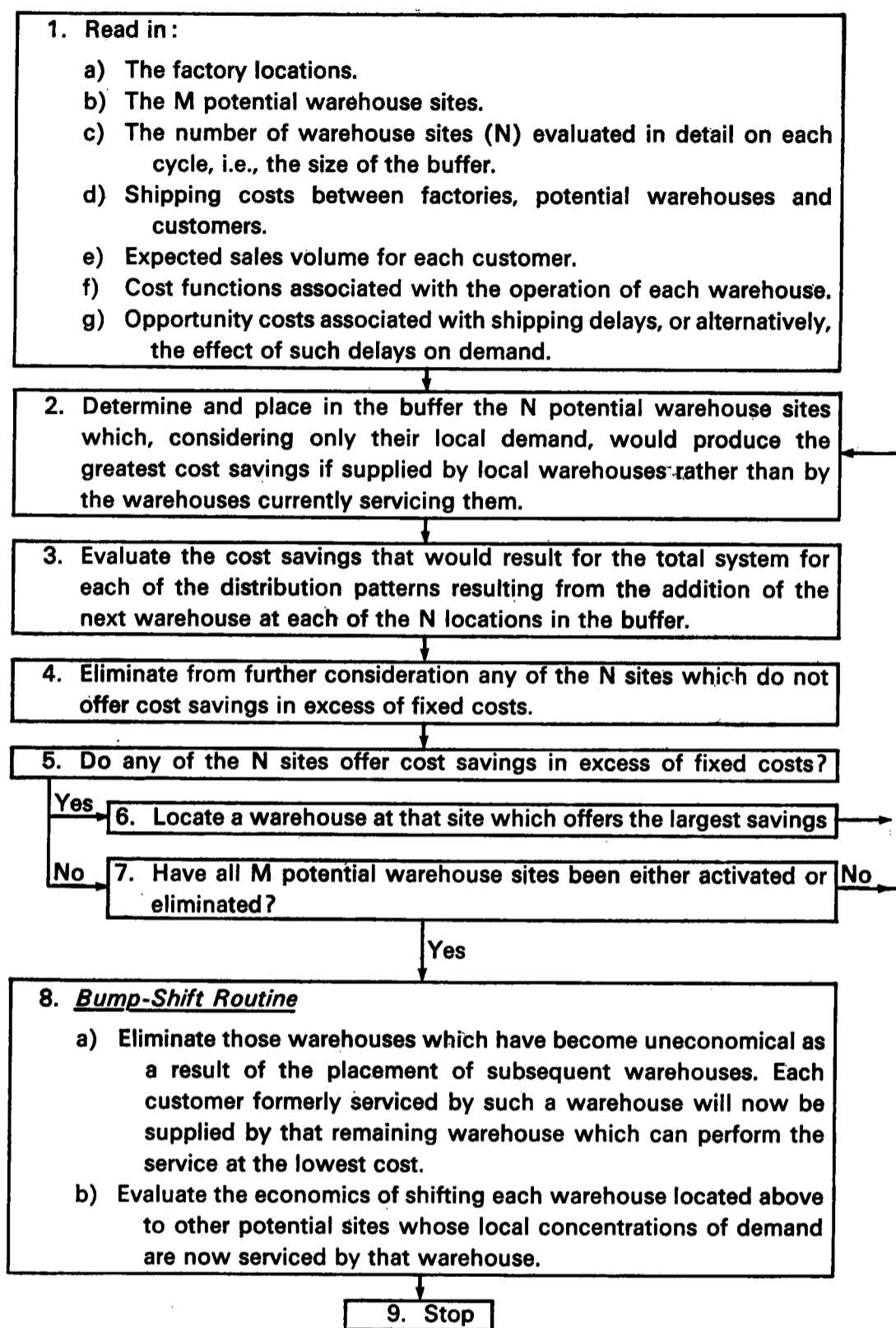
9. Stop

Fig. 1.   Flow diagram

complexity. To deal with this complexity the formal models have grown in power and sophistication. But complexity has also stimulated the development of new kinds of models of rational decision that take special account of the very limited information-gathering and computing capacity of human beings and their associated computers.

One response to the concern with uncertainty, with the difficulties of discovering or designing alternatives, and with computational complexity has been to introduce search and information transmission processes explicitly into the models. Another (not exclusive) response has been to replace optimization criteria with criteria of satisfactory performance. The satisficing approach has been most often employed in models where "heuristic" or trial-and-error methods are used to aid the search for plausible alternatives.

As a result of all these developments, the decision maker today, in business, government, universities, has available to him an unprecedented collection of models and computational tools to aid him in his decision-making processes. Whatever the compromises he must make with reality in order to comprehend and cope with it, these tools make substantially more tractable the task of matching man's bounded capabilities with the difficulty of his problems.

References

BAUMOL, W. J. (1959), *Business Behavior, Value and Growth,* Macmillan, New York, pp. 45–53.

CHAPANIS, N. P. and J. A. CHAPANIS (1964), "Cognitive Dissonance: Five Years Later," *Psychological Bulletin, 61,* 1023.

CHARNES, A. and W. W. COOPER (1961), *Management Models and Industrial Applications of Linear Programming,* Wiley, New York, (2 volumes).

DE GROOT, A. (1965), *Thought and Choice in Chess,* Mouton, The Hague.

FORRESTER, J. W. (1961), *Industrial Dynamics,* M.I.T. Press, Cambridge.

KUEHN, A. A. and M. J. HAMBURGER (1963), "A Heuristic Program for Locating Warehouses," *Management Science, 9,* 643-666.

MARSCHAK, J. (1955), "Elements for a Theory of Teams," *Management Science, 1,* 127–137

SIMON, H. A. (1957), Part IV in *Models of Man,* Wiley, New York, pp. 196–279.

SIMON, H. A. and P. A. SIMON (1962), "Trial and Error Search in Solving Difficult Problems," *Behavioral Science, 7,* 425–429.

STIGLER, G. J. (1961), "The Economics of Information," *Journal of Political Economy, 69,* 213–225.

VON NEUMANN, J. and O. MORGENSTERN (1953), *Theory of Games and Economic Behavior,* (3rd ed.), Princeton University Press, Princeton, pp. 125.