77 16 [1972]

Scheduled Tuesday, May 16, 1972

PLANNING, MANAGEMENT, POLICIES AND STRATEGIES: FOUR FUZZY CONCEPTS

Robert Rosen
Visiting Fellow
CSDI

I. *Introduction*

I think it is fair to say that the mood of those concerned with the problems of contemporary society is apocalyptic. It is widely felt that our social structure is in the midst of crisis, certainly serious, and perhaps ultimate. It is further widely felt that the social crises we perceive have arisen primarily because of the anarchic, laissez-faire attitude taken in the past towards science, technology, economics and politics. The viewpoint of most of those who have written on these subjects revolves around the theme that if we allow these anarchies to continue, we are lost; indeed, one way to make a name nowadays is to prove, preferably with computer models, that an extrapolation of present practices will lead to imminent cataclysm. The alternative to anarchy is management; and management implies in turn the systematic implementation of specific plans, programs, policies, and strategies. Thus it is no wonder that the circle of ideas centering around the concept of *planning* plays a dominant role in current thought.

However, it seems to an external observer (such as myself) that the net effect of the current emphasis on planning has been simply to shift the anarchy we perceive in our social processes into our ideas about the management of these processes. If we consider, for example, the area of "economic development" of the under-developed countries (a topic which has been extensively considered around the table during my residence here), we find (a) that there is no clear idea of what constitutes "development"; (b) that the various definitions employed by those concerned with development are incompatible and contradictory; (c) that even among those who happen to share the same views as to the ends of development, there are similarly incompatible and contradictory views as to the means whereby the end can be attained. Yet in the name of developmental planning, an enormous amount of time, ink, money and even blood is in the process of being spilled. Surely no remedy can be expected if the cure and the disease are indistinguishable.

If it is the case that planning is as anarchic as the social developments it is intended to control, then we must ask whether there is, in some sense, a "plan for planning", or whether we face an infinite and futile anarchic regress. It may seem at first sight that by putting a question in this form we gain nothing. However, what we shall attempt to argue in the present paper is that, in fact, this kind of question is "well-posed" in a scientific sense, that it can be investigated in a rigorous fashion and its consequences explored. Moreover, we would like to argue that, in the process of investigating this question, some useful and potentially applicable insights into planning itself are obtainable.

The fact that so many distinct and contradictory views of planning, management, policy-making, etc., can be held by able people means nothing else but that the concepts named by these words are fuzzy in the extreme. Thus a first step in dealing with such concepts is to try to identify and remove the source of the fuzziness. We have already indicated one way in which this can be done in the paper prepared for the conference concerned with B. F. Skinner's book, "Beyond Freedom and Dignity". In that conference, another fuzzy concept, namely that of "behavior", was at issue. We argued that it was possible to clarify this concept by constructing a variety of worlds, in each of which the concept of "behavior" could be given a precise and unambiguous meaning. in each such world, the properties and consequences of behavior could be rigorously investigated. The fuzziness of the term "behavior" in ordinary parlance, and the resulting equivocation which made many commentators unhappy with Skinner's use of the term, arises then because the properties of several such worlds are being indiscriminately mixed. Thus we believe that the first step towards using any fuzzy concept properly is to separate out various worlds of this kind, understand each of them in detail, and then in any specific real situation to use the concept in the exact sense manifested by the model world which appropriately represents that real situation.

In what follows, we are going to attempt to analyze the concepts of planning, management, strategies, policies, design, etc., in this way. We are going to construct a particular model world (or rather, a class of model worlds all constructed on the basis of a common set of principles) in such a way that these concepts can be reasonably and unambiguously defined. We will then try to indicate, at least in a preliminary way, how we may use these model worlds to obtain useful insights into the planning process itself.

In Section II below we shall consider a few preliminary ideas regarding system modelling in general, which will be useful to have before us in order to motivate the construction of our model worlds. In Section III these worlds will be defined, and the concepts of planning, design, policies, etc., will be identified in them. The remaining sections are devoted to a preliminary exploration of the properties of these worlds and how they may perhaps be used for a variety of useful purposes.

II. *The Philosophy of Modelling Systems*

The class of worlds we are going to construct seem to be rather different than any which have previously been considered. It will be important to understand how and why they are different; in order to see this, it will be helpful to review some of the implicit assumptions which traditionally underlie our making of models for the study of system behavior.

At present, the languages which we use to construct system models of whatever kind have their roots in the mechanics of Newton. However much these languages may differ in detail and emphasis, they all represent paraphrases of the language of Newtonian mechanics. The salient ideas are as follows. Two separate ingredients are necessary for the process of system description; they are: (a) a specification of what the system is like at any particular instant of time, with the associated concept of the *instantaneous state* of the system, and (b) a specification of how the system changes state, as a function of present or past states and of the forces imposed on the system. The characterization of the instantaneous state involves the specification of an appropriate set of state variables, while the characterization of how the system changes state involves a specification of the *equations of motion* of the system.

A mode of system description which looks at first sight quite different from this involves "input-output analysis", or "black box analysis". In this situation, we admit that we cannot find a state description for our system, which is accordingly called a black box. We can, however, manipulate the inputs to this box, and for each input we can observe the corresponding output. The box is thus characterized in functional terms, by specifying for each input what the corresponding output shall be. In Newtonian terms, however, the inputs to a black box correspond to the various *forces* which can be imposed upon the system; the outputs represent certain system observables (which are simply numerical functions on the set of states of the system). The emphasis in Newtonian mechanics is to find out how the states change when the forces on the system are kept fixed; i.e. to look upon the state as the output and specify it as a function of time. In black box analysis the forces are varied and we seek the output (which is a function of state, and which is known if the state is known) as a function of input or force. Thus the two modes of system analysis, at first sight so different, are in fact essentially the same, and differ primarily in emphasis. Indeed, one way to carry out a black box analysis is simply to introduce a formal set of state variables for the black box and proceed entirely within a Newtonian framework.

As we have noted above, *all* of our modes of system description are derived from mechanics in one way or another, and share all of the implicit prescriptions and presumptions of mechanics. One of the most basic of these presumptions, so basic indeed that it is hardly ever mentioned explicitly, is the following: we must *never allow future states of the system to affect the present changes of state*. Typically in physics only the *present* state, and the *present* force, enter into the equations which govern the changes of state; i.e. physics typically deals with memoryless systems. In biological and technological systems we must consider systems possessing a memory (in some sense), and for these systems we must allow the dynamical equations to involve past states and past forces as well. But conventional causality requires that the future must not enter into the equations of motion. Systems which do allow future states to determine present changes of state are called *anticipatory*; if they are ever mentioned in theoretical works it is only to explicitly exclude them from consideration.

As a result of this implicit restriction, all of our dynamical theories, in biology and human sciences as well as in physics, are based on the dynamics of non-anticipatory systems. Consequently we have not the faintest idea of how anticipatory systems behave formally. This is a most amazing fact, when we consider that once we reach a certain level of complexity, say in biological systems, many types of frankly anticipatory behaviors are encountered. All sensory mechanisms are anticipatory in effect; the present behavior of an organism is modified as a function of a future state implied, in some sense, by present sensory data (such mechanisms, as we have noted in a previous note, involve feed-forward control in an essential way; we shall have more to say about feed-forwards and their relation to anticipatory dynamics below). The kinds of behavior we like to call “intelligent” are anticipatory in this sense. Biological examples of this kind of anticipatory response could be multiplied endlessly. Yet because our only modes of system description implicitly exclude anticipatory behavior, all of our attempts to model such behavior has been in terms of non-anticipatory models. And indeed, it is true that, given any *particular* system behavior, we can always model an anticipatory system with a non-anticipatory model. But we require a different model for each new behavior, and the models we obtain thereby are unrelated and contradictory, leading to endless acrimony (the interested reader may consult, for example, the literature on intelligence, both natural and artificial). Thus it can be argued that the very basis we have used for the understanding of biological behaviors is inadequate in principle, and we can open thereby a door to a potentially most fruitful area of research.

Insofar as "planning" is a manifestation of intelligence, and insofar as the goal of planning is to cause present changes of state in accordance with anticipated future states of the system of interest, it becomes in particular important to ask whether we can try to understand the planning process itself most effectively within the context of the heretofore forbidden class of anticipatory systems. This indeed is what we are going to do; we are going to construct our worlds in such a way that anticipation is built intrinsically into them. There are many ways in which this can be done; one of them in particular will be described in detail in the following section.

III. *A Class of Anticipatory Systems*

We are now ready to construct our model world, which will consist of a class of systems of definite structure, involving anticipation in an essential way, and in which the fuzzy terms associated with "planning" can be given a concrete meaning.

Let us suppose that we are given a system S, which shall be the system of interest, and which we shall call the *object system*. S may be an individual organism, or an ecosystem, or a social or economic system. For simplicity we shall suppose that S is an ordinary (i.e. non-anticipatory) dynamical system.

With S we shall associate another dynamical system M, which is in some sense a *model* of S (the details of how M and S must be related in order for M to actually be a model of S will be discussed separately in the Appendix). We require, however, that if the trajectories of S are parameterized by real time, then the corresponding trajectories of M are parameterized by a time variable which goes faster than real time. That is, if S and M are started out at time $t = 0$ in equivalent states, and if (real) time is allowed to run for a fixed interval T, then M will have proceeded further along its trajectory than S. In this way, the behavior of M *predicts* the behavior of S; by looking at the state of M at time T, we get information about the state that S will be in at some time later than T.

We shall now allow M and S to be coupled; i.e. allow them to interact in specific ways. For the present, we shall restrict ourselves to ways in which M may affect S; later we shall introduce another mode of coupling which will allow S to affect M (and which will amount to updating or improving the model system M on the basis of the activity of S). We shall for the present suppose simply that the system M is equipped with a set E of *effectors*, which allow it to operate either on S itself, or on the environmental inputs to S, in such a way as to change the dynamical properties of S. We thus have a situation of the type diagrammed below:

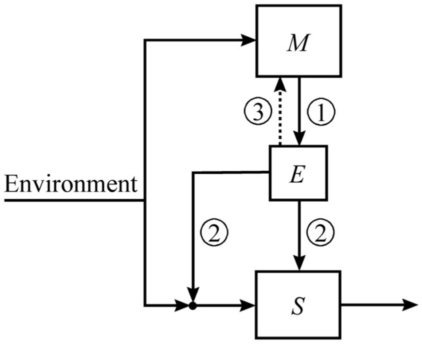

Figure 1

If we put this entire system into a single box, that box will appear to us to be an adaptive system in which prospective future behaviors determine present changes of state. It would be an anticipatory system in the strict sense if M were a *perfect* model of S (and if the environment were constant or periodic). Since in general M is not a perfect model, for reasons to be discussed in Section V below, we shall call the behavior of such systems *quasi-anticipatory*.

We have said that M "sees" into the future of S, because the trajectories of M are parameterized faster than those of S. How is this information to be used to modify the properties of S through the effector system E? There are many ways in which this can be formalized, but the simplest seems to be the following. Let us imagine the state space of S (and hence of M) to be partitioned into regions corresponding to "desirable" and "undesirable" states. As long as the trajectory in M remains in a "desirable" region, no action is taken by M through the effectors E. As soon as the M-trajectory moves into an "undesirable" region (and hence, by inference, we may expect the S-trajectory to move into the corresponding region at some later time, calculable from a knowledge of how the M- and S-trajectories are parameterized) the effector system is activated to change the dynamics of S in such a way as to keep the S-trajectory out of the "undesirable" region.

From this simple picture, a variety of insights into the nature of "planning", "management", "policies", etc., can already be extracted. Let us review these in sequence.

A. *Choice of M*

The first essential ingredient in the planning process in these systems involves the choice of the model system M. There are many technical matters involved in choosing M, which will be discussed in more detail in Section V below. We wish to point out here that the choice of M involves paradigmatic aspects as well, which color all future aspects of the "planning" process. One simple example of this may suffice. Let us suppose that S is a simple early capitalist economic system. If we adopt a model system which postulates a large set of small independent entrepreneurs, approximately equivalent in productive capability and governed by "market forces", we find that the system S is essentially stable; coalitions are unfavored and any technical innovations will rapidly spread to all competitors. On the other hand, if we adopt a model system S in which there are positive feedback loops, then we will see the same situation as *unstable*, much as an emulsion of oil and water is unstable. That is, initially small local accretions of capital will tend to be amplified, and the initially homogeneous population of many small entrepreneurs will ultimately be replaced by enormous cartels. This, in a highly oversimplified way, seems to me to represent the difference between laissez-faire capitalism and Marxian socialism, proceeding from two different model systems of the same initial economic system S, and hence predicting two entirely different futures for S.

B. *Selection of the effector system E*

Once the model M has been chosen, the next step of the planning or management process for S is to determine how we are to modify the dynamics of S according to the information we obtain from M. This problem involves several stages. The first stage involves a selection of "steering" variables in S, or in the environment of S, through which the dynamical properties of S can be modified. In general, several different kinds of choices can be made, on a variety of different grounds. In empirical terms, this choice will most often be made in terms of the properties of the model system M; we will consider how M can be most effectively steered, and use the corresponding state variables of S (if possible) for the control of S. Thus again the initial choice of the model system M will again tend to play a major role in determining the specifics of the planning process.

C. *Design of the effector system E*

Once having chosen the control variables of S, we must now *design* a corresponding effector system. This is a technological kind of problem, governed by the nature of the control variables of S and their response characteristics. We may wish, for example, to employ only controls which are easily reversible.

D. *Programming of the effector system E*

The final aspect of the planning process involves the actual programming of the effector system; i.e. the specification of a dynamics on E which will convert the input information from M (i.e. information about the future state of S) into a specific modification of the dynamics of S. This transduction can be accomplished in many ways, and involves a mixture of "strategic" and "tactical" considerations.

E. *Identification of "desirable" and "undesirable" regions*

Ultimately the programming of the effectors E will depend heavily on the character of the regions we consider "desirable" and those we consider undesirable. This choice too is arbitrary, and is in fact independent of the model system M which we have chosen. It represents a kind of constraint added from the outside, and enters into the planning process in an equally weighty fashion as does the model M and the effector system E.

F. *Updating the states of M*

In Figure 1 above we have included a dotted arrow (labelled ③) from the effector system back to the model. This is for the purpose of resetting the states of the model, according to the controls which have been exerted on the system S by the effector system. Unless we do this, the model system M becomes useless for predictions about S subsequent to the implementation of control through E. Thus, the effector system E must be wired into M in a fashion equivalent to its wiring into S.

The enumeration (A)–(F) above seems to be a useful atomization of the planning process for the class of systems we have constructed. Within this class, then, we can proceed further and examine some of the consequences of planning, and in particular the ways in which planning can go wrong. We shall sketch these analyses in the subsequent sections.

IV. *An Interesting Property of Anticipatory Systems*

Before we proceed further with our analysis of the planning process, an interesting property of anticipatory models may be noted, especially for its bearing on "retrospective futurology". We have pointed out in a previous note that non-anticipatory open systems, in the neighborhood of any (stable or unstable) steady state, can be decomposed into a variety of feedback loops between subsystems which may be regarded as controllers and controlled systems according to the conventional views of feedback control theory. The anticipatory systems which we have constructed, and indeed anticipatory systems in general, exhibit an interesting duality with the more familiar non-anticipatory ones. Namely, it is possible in anticipatory systems, in the neighborhood of steady states, to decompose the system into *feed-forward* loops of the kind shown in Figure 1 above. These too may stabilize or destabilize the controlled system S, just as feedbacks may be positive or negative.

I would like to conjecture that one can use such decompositions, when considering historical questions, to obtain insight into the plans, strategies, policies, etc., which underlie specific historical developments. This amounts to taking the total historical system under consideration and analyzing its dynamics to obtain diagrams like that shown in Figure 1. For such a diagram, the overall dynamics determines the properties of the model system, the effectors, and the interaction between these and the controlled system S. Unfortunately this kind of decomposition is not unique, but if intelligently done, it can throw interesting light on historical questions, just as the corresponding analysis done in non-anticipatory systems (e.g. biochemical cell models) have displayed entirely new control properties in these systems.

V. *How Planning Can Go Wrong*

The atomization of the planning process, at least within the class of anticipatory systems which we have constructed, gives some interesting indications of how planning can fail. A quantitative analysis of these questions might yield some interesting diagnostics which could be applied to specific planning problems. We cannot undertake this here; however, a number of qualitative remarks on how planning may fail might be of interest, as they touch on some basic methodological considerations.

1. *Bad Models*

Clearly one important failure of planning arises from a poor choice of the model system M. There are a number of distinct ways in which a model can be bad, some of which are more serious than others.

One large class of bad models arises from purely technical considerations. These involve the ignoring of important state variables, or an incorrect specification of the equations of motion. A rather more insidious class of bad models arises from an incorrect paradigm; an attempt to model a particular kind of dynamical behavior through inappropriate means. As an example of a possibly incorrect paradigm, we may offer the remarks made in Section II, regarding our attempts to model anticipatory behavior by means of non-anticipatory models — no amount of technical refinement of the modelling process will enable the essence of the process of interest to be captured. A third class of bad modelling practices arises from an incorrect correspondence between the states of the object system S and the states of the model system M; this also gives rise to an incorrect prediction of the future of S, even if the model itself is technically and paradigmatically correct.

2. *Bad Effectors*

Any one of the steps involved in designing and installing the effector system E will lead to a failure of planning. Specifically: (a) the state variables chosen to steer S may not in fact be capable of so doing; (b) the effector system itself may not be capable of making the appropriate manipulations of the state variables; (c) the effector system may be programmed badly, so that the changes made in S are not the changes which should be made in S on the basis of the information coming from M.

3. *Side Effects*

There is, however, a class of planning difficulties which do not arise from such obvious considerations, and which merit a fuller discussion. This class of difficulties has to do with the problem of *side effects*; as we shall see, these will generally arise, even if the model system is perfect, the effectors perfectly designed and programmed, because of inherent system-theoretic properties. Let us see how this comes about.

In a previous paper presented here I enunciated a conjecture which I believe to have general validity: namely, that in carrying out any particular functional activity, a system S typically only uses a few of its degrees of freedom. This proposition has several crucial corollaries, of which we noted two in the preceding paper:

(1) The same structure can be involved simultaneously in many different functional activities, and conversely,

(2) The same functional activity can be carried out (or "realized") by many different kinds of structures.

We stressed in that paper how the fact that *all* of the state variables defining any particular system S are more or less strongly *linked* to one another via the equations of motion of the system, taken together with the fact that the many state variables not involved in a particular functional activity were free to interact with other systems in a non-functional or dysfunctional way, implied that any particular functional activity tends to be modified or lost over time. This, we feel, is a most important result, which bears directly on the "planning" process under discussion. The easiest way to see this is to draw another corollary from the fundamental proposition that only a few degrees of freedom of a system S are involved in any particular functional activity of S.

(3) Any functional activity of a system S can be *modelled* by a system whose structure is simple compared to that of S (simply by neglecting the non-functional degrees of freedom of S). Indeed, it is largely because of this property that science is possible at all. Conversely,

(4) No one model is capable of capturing the full potentialities of a system S for interactions with arbitrary systems.

The corollary (4) is true even of the best models, and it is this corollary which bears most directly on the problem of *side-effects*. Let us recall that S is by hypothesis a real system, whereas M is only a model of a particular functional activity of S. There are thus many degrees of freedom of S which are not modelled in M. Even if M is a good model, then, the capability for dealing with the non-functional degrees of freedom in S have been abstracted away. And these degrees of freedom, which continue to exist in S, are generally *linked* to the degrees of freedom of S which are modelled in M, through the overall equations of motion which govern S.

Now the planning process requires us to construct a real system E, which is to interact with S through a particular subset of the degrees of freedom of S (indeed, through a subset of those degrees of freedom of S which are modelled in M). But from our general proposition, only a few of the degrees of freedom of E can be involved in the interaction. Thus both E and S have in general many "non-functional" degrees of freedom, through which other, non-modelled interactions can take place. Because of the linkage of all observables, the actual interaction between E and S specified in the planning process will in general be affected. Therefore, we find that the two following propositions are generally true:

a. An effector system E will in general have other effects on an object system S than those which are planned;

b. The planned modes of interaction between E and S will be modified by these effects. Both of these propositions describe the kind of thing we usually refer to as *side effects*. As we see, such side effects are unavoidable consequences of the general properties of systems and their interaction. They are by nature unpredictable, and are inherent in the planning process no matter how well that process is technically carried out. As we pointed out in our previous paper, there are a number of ways around this kind of difficulty, which we have partially characterized, but they are only applicable in special circumstances.

We would like to conjecture further that, for any specific planning situation (involving an object system S, a model M, and suitably programmed effectors E) each of the ways in which planning can go wrong will lead to a particular kind of syndrome in the total system (just as the defect of any part of a sensory mechanism in an organism leads to a particular array of symptoms). It should therefore be possible, in principle, to develop a definite diagnostic procedure to "trouble-shoot" a system of this kind, by mimicking the procedures used in neurology and psychology. Indeed, it is amusing to think that such planning systems are capable of exhibiting syndromes (e.g. of "neurosis") very much like (and indeed analogous to) those manifested by individual organisms.

VI. *Some Further Considerations*

There are many ramifications of the class of systems developed above, for the purpose of studying the planning process, which deserve somewhat fuller consideration than we have allowed. In this section we shall consider two of them: (a) how can we update and improve the model system M, and the effector system E, on the basis of information about the behavior of S itself?, and (b) how can we avoid a number of apparent infinite regresses which seem to be inherent in the planning process?

These two apparently separate questions are actually forms of the same question. We can see this as follows. If we are going to improve, say, the model system M, then we must do so by means of a set of effectors E'. These effectors E' must be controlled by information pertaining to the effect of M on S; i.e. by a model system M' of the system $\left(S+M+E\right)$. In other words, we must construct for the purpose of updating and improving M a system which looks exactly like Figure 1, except that we replace M by M', E by E', and S by $S+M+E$. But then we may ask how we can update M'; in this way we see an incipient infinite regress.

There is another infinite regress inherent in the discussion given of side effects in the preceding section. We have seen that the interaction of the effectors E with the object system S typically give rise to effects in S unpredictable in principle from the model system M. However, these effects too, by the basic principle that only a few degrees of freedom of S and E are utilized in such interactions, are capable of being modelled. That is, we can in principle construct a new model system M_1 of the interaction between S and E, which describes interactions not describable in M. If these interactions are unfavorable, we can construct a new set of effectors, say E_1, which will steer the system S away from these side effects. But just as with E, the system S will typically interact with E_1 in ways which are in principle not comprehensible within the models M or M_1; these will require another model M_2 and corresponding new effectors E_2. In this way we see another incipient infinite regress forming.

Indeed, this last infinite regress is highly reminiscent of the “technological imperative” which we were warned against by Ellul and many others. Thus the question arises: can such infinite regresses be avoided?

These kinds of questions are well-posed, and can be investigated in system-theoretic terms. I have considered questions like these in a very different connection; namely, under what circumstances is it possible to add a new functional activity to a biological organization like a cell? It turns out that one cannot simply add an arbitrary function and still preserve the organization; we must typically keep adding functions without limit. But under certain circumstances, the process does indeed terminate; the new function is included (though not *just* the new function in general) and the overall organization is manifested in the enlarged system. On the basis of these considerations, I would conjecture that (a) it is possible in principle to avoid the infinite regresses just alluded to in the planning process, and in particular to find ways of updating the model M and the effectors E; (b) *not every way of initiating and implementing a planning process allows us to avoid the infinite regress*. The first conjecture is optimistic; there are ways of avoiding this form of the “technological imperative”. The second can be quite pessimistic in reference to our actual society. For if we have in fact embarked on a path for which the infinite regresses cannot be avoided, then we are in serious trouble. Avoiding the infinite regresses means that developmental processes will stop, and that a stable steady-state condition can be reached. Once embarked on a path for which the infinite regresses cannot be avoided, no stable steady-state condition is possible. I do not know which is the case in our own present circumstances, but it should at least be possible to find out.

I hope that the above few remarks on the planning process will provide food for thought for those more competent to investigate such problems than I am.