LIBRARY
OF THE
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
PLANNING AS AN ITERATIVE HIERARCHICAL PROCESS

AND ITS FORMALIZATION IN COMPUTER MODELS

James C. Emery, Asst. Prof.

October, 1964

Draft of a paper to be delivered to the
Second Congress on the Information System Sciences
Hot Spring, Virginia, November 22-25, 1964

The study was supported in part by a grant to M.I.T. from the National Aeronautics and Space Administration. It is receiving limited circulation in order to elicit comments, and should not be quoted or reproduced without permission. I received valuable suggestions for improving the paper from my colleagues, Billy Goetz, William Pounds, Charles Ying, and Zenon Zannetos; and from Dr. Norman Waks, Mitre Corp. Computer time and programming assistance is being provided by M.I.T.'s Project MAC.
PLANNING AS AN ITERATIVE HIERARCHICAL PROCESS
AND ITS FORMALIZATION IN COMPUTER MODELS

ABSTRACT

This paper attempts to provide a conceptual framework for designing computer planning models. Planning is viewed as an iterative hierarchical process. Each level in this hierarchy involves essentially the same five-step sequence: 1) determine planning variables; 2) propose alternative plans and generate their consequences; 3) select the best alternative analyzed; 4) translate the plan into a form for lower-level planning; and 5) control the plan. Planning at each level is subject to constraints imposed by the organizational structure and goals, resources and information available, and various forms of programmed behavior. Efforts to improve the planning process (and hence plans) must act primarily through these planning constraints. The role of computers in planning is examined in light of these concepts. It is concluded that a formalized computer planning model will, like all planning, have a hierarchical structure and involve an iterative process. Finally, the paper discusses some of the characteristics of such a model.

Introduction

Planning is a universal component of management; it necessarily precedes all action. And yet, to a considerable extent, planning is not well understood by those who practice it. Many managers fancy themselves men of action, and view planning with a certain amount of distaste. In fact, however, only through the planning process can a manager exert any lasting influence on the behavior of the organization he commands.

Planning has been given a variety of meanings. All essentially agree, however, with Newman's definition (1951, p. 15): "Planning is deciding in advance what is to be done." The simplicity of the definition hides the difficulty of planning.
The complete set of plans within an organization constitutes an exceedingly elaborate network. This network has both an organizational and a time span hierarchy. Plans at one level in the organization should (within "reasonable" limits) be consistent with the plans of higher-level units; and short-range plans should be consistent with long-range plans. Furthermore, plans throughout the entire organization should be such that the planned production and acquisition of resources meshes in quantity and time-phasing with the planned resource requirements. The generation of a set of plans meeting these consistency criteria represents one of the major difficulties of planning.

But consistency alone is not sufficient; plans must also promote purposeful behavior. The planned activities within each organizational subunit must lead to the satisfactory achievement of its goals. The goals assigned to each unit are the result of a process of dividing--or "factoring"--the organization's overall goals into a hierarchy of subgoals. This factoring process has as its aim the creation of a set of (seemingly) independent subgoals whose separate achievement will result in the accomplishment of the global goals.

**Purpose of a plan**

A plan has two primary purposes: 1) it defines actions to be taken and outcomes desired, and 2) it serves as a vehicle of coordination. It would be useful, I think, to examine these in somewhat more detail.

**The plan as a definition of actions and outcomes.** A plan communicates actions to be taken. The form of the communication may vary. A "standing plan" (Newman, 1951, p. 18) may persist over a considerable time span, to be evoked under specified conditions or by an explicit message that triggers execution.
Such a plan constitutes a continuing "action module" or "performance program" (March and Simon, 1958, pp. 141-150) by which higher-level units govern the behavior of lower-level activity. A "single-use" plan, on the other hand, describes--normally unconditionally--actions to be taken over a specified planning horizon. The total plan network of an organization consists of a hierarchy of single-use plans that incorporate or reflect actions described by the standing plans.

Plans--like computer programs--may be expressed in either a procedural or a declarative language. A procedural plan specifies a step-by-step sequence of actions that presumably will lead to a known (and desired) outcome. Alternatively, a plan may only declare a desired outcome, and leave to those executing it the responsibility for specifying the sequence of actions necessary to achieve it.

A plan is described in terms of variables. A manufacturing schedule, for example, is described by such variables as units of production, manpower levels, costs, and so forth. These provide only an abstract representation of planned behavior. This is particularly true of high-level plans which largely deal with gross aggregates. Such plans require eventual amplification into lower-level plans defined in terms of more detailed variables. But even in its most detailed form, a plan specifies only a relatively tiny set of variables out of the infinite set possible.

Both "action" and "outcome" variables are used to describe a plan. An action variable describes the controllable actions by which a plan is implemented.

---

1Procedural and declarative descriptions correspond, respectively, to process and state descriptions (Simon, 1962, p. 479) or to activity and product specifications (March and Simon, 1958, pp. 144-146).
(although further translation into more detailed action variables may often be required). Examples include employment level, capital expenditures, aggregate inventory investment, advertising expenditures, and so forth.

An outcome variable describes the consequences stemming from a given choice of action variables—profit, return on investment, and predicted inventory stock-out rate, for instance. These variables need not be completely independent. If the planner does not know the relationship existing between two variables, he might wish to describe a plan in terms of both even if one is clearly subsumed under the other. For example, "profit" and "inventory stockout rate" might be used as outcome variables, although presumably the latter is of no inherent interest except insofar as it has some (generally unknown) effect on long-range profit.\(^1\)

It is useful—but unfortunately often not feasible—to assign probability estimates to outcome variables. Alternatively, outcomes may sometimes be described in terms of ranges, standard deviations, and other variables that show the probabilistic nature of a plan.

In choosing a plan, the planner\(^2\) naturally tries to use a set of action variables that will lead to a satisfactory outcome in terms of the organization's "true" goals. The variables ignored are assumed to be irrelevant, insignificant, or implied by the accomplishment of the stated variables. For example, a satisfactory outcome in terms of the explicitly stated variable "profit" may

---

1If the indirect consequences of an action cannot be estimated, it may be thought of as both an action and an outcome variable. For example, "research and development expenditures" describes an action and the direct consequences of that action.

2I use the term "planner" throughout this paper to describe a manager having decision-making authority for planning.
imply a satisfactory outcome in terms of the unstated variable "survival" (both the organization's and the planner's); and a production schedule defined in terms of units completed may imply satisfactory performance in terms of stability of employment, delivery performance, and so forth. The stated variables must thus close "loopholes" sufficiently well that achievement of the defined plan will in fact have desirable consequences. Ordinarily this requires the specification of plans in physical units, cost units, and units that measure capital investment (Goetz, 1949, pp. 92-115).

Since a plan communicates actions required of each unit, it furnishes the channel through which management guides the behavior of the organization. Under this philosophy, the act of choosing a plan (or, equivalently, approving one proposed by a lower-level unit) constitutes one of the most important functions of management. Upon approval, a plan enters the plan network that serves as the basis for execution and coordination. Any significant addition, deletion, or modification to a plan should be effected only through the same official approval mechanism that authorized the plan in the first place. Failure to do this may rob the plan of its integrity and thwart the objectives of the original planner. This concept is very much in evidence in the programming system of the Department of Defense (D.O.D., 1962).

In order for the approved plan network to play this central role, it must be based on the best current predictions of future events (within allowed error tolerances). This, in turn, calls for a control system that takes periodic samples of actual events and compares them with the predictions on which the current plans are based. If a significant deviation occurs, the control system signals the need for new planning that takes into account the most recent available
information about the environment. With such a control system to "close the feedback loop," planners at all levels in the organization can look to the latest plan network as a reliable source of information for execution and further planning.

Management under these circumstances largely focuses on the planning process. Far from being just a sterile paperwork or bureaucratic activity, the planning function represents the central nervous system of the organization. At the risk of overdramatizing, I think it is not too unreasonable to view a manager as a planner who lives largely in an analogue, abstract world involving the predicted future. He is affected by the real world only to the extent that its essence is captured in his analogue planning world. The planner, like the brain, is forever cut off from the world "out there," and can only perceive things through highly filtered information channels (Bishop, 1960, pp. 122-146).

The plan as a vehicle of coordination. The plan network reflects the factoring of the organization's goals into a hierarchy of subgoals. One of the primary purposes of the goal factoring process is to generate a hierarchy of relatively independent subtasks. Only with such independence can each organizational unit pursue its own subgoals without paying constant attention to the activities of other units.

The method of factoring has a major effect on the degree of independence among subunits. In general, closely interacting activities should be hierarchically "close" to each other. This reduces the amount of coordination required across hierarchical lines, and thus partially insulates each unit from other parts of the organization.

Unfortunately, this principle often conflicts with other aims, and one must strike a balance among them in establishing hierarchical structure. For example,
a project (or "purpose") organizational structure in its purest form assigns to a single branch of the hierarchy all of the tasks, and only those tasks, connected with an independent objective. This provides a high degree of independence. Alternatively, a functional (or "process") organizational structure groups together complementary activities that serve multiple purposes— all of the electrical engineering work associated with several different development projects, say. This allows greater specialization, with all its attendant efficiencies. It does it, however, at the cost of assigning closely interacting tasks to different branches of the organizational tree, thereby increasing the amount of coordination required across hierarchical lines.

Although interactions among subunits can be reduced by the choice of organizational structure, neither economics nor the complex nature of organizational activities permit them to be eliminated entirely. Interactions exist all but universally in organizations composed of hierarchical subunits pursuing a common set of global goals. They arise through elaborate couplings of physical inputs and outputs flowing between units, through the allocation of scarce common resources, and through effects on a common environment. It is the presence of these interactions that makes managing an organization such an enormously complex task.

Various devices exist for coping with interactions. One approach aims at creating greater short-term independence among subunits by partially decoupling closely related tasks. This is accomplished by standardization of material flowing from one unit to another, by the use of buffer inventories to provide a short-term cushion between the rate of production of a supplying unit and the rate of consumption of a using unit, and by making available excess resources and time in order to increase the probability of meeting scheduled physical output requirements (March and Simon, 1958, pp. 158-160).
All of these devices serve to reduce interactions among subunits. A price is paid for this, however: tighter tolerances increase manufacturing costs, buffer inventories take money to acquire and maintain, and slack resources reduce capital and labor utilization.

Thus, a reduction in interactions has costs associated with it, whether it is effected through the organizational structure or through decoupling. Closer coordination of separate subunits provides a partial remedy. But coordination, too, costs money—in the form of data collection, data transmission, and computation. The compromise between greater independence and greater coordination represents a basic organizational issue.

Plans provide the primary vehicle for coordination. The plan network describes the actions or outcomes expected of each unit in the organization. If the process that generates the network explicitly considers the effects of significant interactions, then the ensuing behavior will of course recognize these interactions. In this way the plan network coordinates activities over the life of the plan.

Even if it is infeasible to formulate plans that explicitly consider all important interactions simultaneously, each planner must have some information about the expected behavior of other closely related units. For example, a subassembly planner must know the assembly schedule before he can establish his own schedule. In other words, even if the magnitude of the task precludes treating the actions of others as variables, the more important of them should at least be recognized as constraints.

Obviously, a problem of circularity arises here. If A bases his plans on B's, and B does conversely, who plans first? Normally a planner has considerable prior information about the activities of the other units with which he interacts strongly. With this, he can partially predict the behavior of the other units.
Plans in their various forms provide the basic source of information for predicting the actions of others. This is true of both standing and single-use plans. Standing plans—in the broadest sense, policies, procedures, and all other types of programmed activity—play much the same role as habit in biological organisms. They provide organizational stability, and thus increase the accuracy with which one unit can predict the behavior of another. This allows a unit to formulate its own plan based on an assumed or provisional level of activity in other parts of the organization. The plan may then require iterative modification if predicted behavior proves to be significantly in error.

A single-use plan describes explicitly the anticipated activity of units included in the plan. Even though the plan gives only an abstract description of behavior, its precision is normally perfectly adequate for purposes of coordination. This is so because the interactions among hierarchically distinct units are generally governed by aggregate, not detailed, behavior—for otherwise the activities should be combined into a single unit.

The weak point in this seemingly happy state of affairs lies in the inability of a planner to generate plans that really take into account all of the interactions that occur throughout the organization. Each action has almost endless ramifications, and the planner faces an impossible—and expensive—task if he attempts to trace them all.

Fortunately, he need not try. He can simply ignore interactions—or at least the vast bulk of them. For example, a production control planner can often schedule the assembly section without detailed consideration of its effect on the various subassembly sections. Less strongly interacting units—a service department
such as the plant cafeteria, say—can certainly be ignored. The planner thus creates an artificial independence among units by basing his plans on the fiction that most of the interactions do not exist. The partitioning and decoupling of the activities within the organization must be such that this fiction represents a legitimate and useful approximation of reality.

The advisability of considering a given interaction obviously depends on the costs involved. Failure to consider an important interaction carries the penalties associated with "suboptimization"—the achievement of local subgoals inconsistent with the global goals of the organization (Hitch and McKean, 1960, pp. 128-131 and 158-181). On the other hand, increasing the scope of planning to include additional effects of interactions can add enormously to the complexity of planning. Once again, the planner must attempt to balance these conflicting costs.

Let me summarize here the discussion on coordination. Plans serve as the basic vehicle of coordination. Comprehensive plans—in terms of their scope, detail and frequency—can guide the organization toward consistent, purposeful behavior in seeking its global goals. But the more comprehensive the planning, the more it costs. The organization can reduce the need for coordination by its choice of organizational structure; it can create short-term independence through various decoupling devices; and it can simply ignore secondary interactions. These, too, have costs associated with them. The planner faces a trade-off between the costs of increased coordination and the costs of real or assumed independence.
**The planning process as a simulation**

Planning generates the plan network by "thinking through" organizational activities prior to their actual execution. This can be viewed as a form of simulation, whether or not it is explicitly recognized as such. From this simulation process emerges a network of plans that describes desired behavior on the part of all units within the organization. With these as a guide, then, the organization carries on its activities that eventually result in some outcome.

One must justify planning on the grounds that the outcome reached by this indirect route is somehow superior to the results that would be achieved using an informal plan generated concurrently in "real time" as actual events take place.\(^1\) Although the logic of making decisions is very much the same in simulated time as it is in real time, the simulated world has at least three distinct advantages. Let us examine them.

**Access to information about activities throughout the organization.** The ability of the concurrent planner to perceive existing conditions throughout the organization is severely restricted by the limited capacity of his sensory channels. As a result, he must normally treat the activities of other units as fixed parameters (at least in the short run). His limited computing rate imposes a further severe restriction. Time moves inexorably forward; it cannot be stopped to allow for additional gathering and processing of information before making an urgent decision.

A planner in a simulated world has much greater access to essential information about his (imaginary) environment than does his real time counterpart. Simulated time moves as the planner directs. The time required to make a decision about a simulated event bears no relation to the time required to execute it—the two are not coupled. Therefore, a planner is constrained not by his

---

\(^1\)Even in the absence of a formal plan, at least some planning occurs in the minds of those participating in an activity. For want of a better term, I will call this "concurrent" planning.
instantaneous information processing rate, but rather by the total available capacity over the period during which he must complete the planning.

The simulation planner enjoys a similar advantage in viewing the future. The ability of a concurrent planner to gaze into the future is limited by his capacity to perceive the current state of nature and make predictions based on his perceptions. The simulation planner suffers from no such limitations: in the simulated world, the timing of an event can be specified.\(^1\) As a result, the planner can incorporate in a plan the earlier activities required to implement the event. In the absence of planning, no allowance will be made for the lead times required to accomplish these antecedent steps.\(^2\)

**Generation of alternatives.** Another advantage of planning is that simulated history is not irrevocable: The moving finger writes; and, having writ, may be cancelled. A planner can simply discard an alternative that fails to satisfy his goals, and can continue to generate plans until he finds it advantageous to terminate. The final plan that results from this process must, of course, be chosen from the set of alternatives considered. Other things being equal, then, the wider the choice presented, the better the final plan selected.

To be sure, copious generation of alternatives will not by itself guarantee a satisfactory plan. In order to avoid choosing among a plethora of mediocre alternatives, the alternatives should be generated through a sequential search

---

\(^1\)The degree of correspondence between these *simulated* events and later *real* events is, of course, another matter, and represents the ultimate test of the usefulness of planning. This matter will be discussed in a later section.

\(^2\)Experimental results in physical control systems demonstrate dramatically the importance of anticipating future actions. See, for example, Sheridan, et. al., 1964.
that provides useful information during each cycle of the process. Using such information, the planner can (hopefully) converge toward better and better plans. Ideally, he terminates when the expected cost of further search exceeds its expected marginal gains. (Manheim, 1964, p. 89). Thus, an optimal plan is, in general, not a "perfect" one.

The selection of a given plan in preference to the others considered implies that there exists some abstract measure for comparing the relative merits of alternative plans. For brevity and consistency with economic literature, I will use the somewhat discredited term "utility" to describe this index.¹

Utility is a function of the outcome variables by which the consequences of a plan are defined. If—as is typically the case—some of the outcome variables are not explicitly commensurable, the planner can then only assign subjective utility values. By so doing he implicitly establishes bounds on the trade-offs between variables. In particular, the utility values reflect the planner's attitude toward the risk associated with each plan (as defined by the probabilistic outcome variables and his own "intuition").

The utility function need not be continuous. In fact, if the planner is a "satisficer," his utility function may assign only two discrete values—"acceptable" or "not acceptable" (Cyert and March, 1963, pp. 9-10).

The utility function may change in response to changes in the planner's perceptions, aspiration level, and "intuition." Furthermore, there is no reason to suppose that the planner is always able to assign consistent utilities (in the sense that they exhibit transitivity).²

¹"Criterion of effectiveness" is often used in management science literature, and conveys the same idea.

²Some "inconsistencies" may be more apparent than real. If the formal outcome variables do not include important but unquantified characteristics of a plan (pertaining to risk, say), then what appears to be formally "irrational" merely reflects differences in the submerged variables (Bowman, 1963).
Simulation in an abstract space. Planning deals with an abstraction of the real world that greatly facilitates the search process. The variables used in planning furnish only the barest outline of the underlying real world. The vast bulk of variables are omitted altogether or aggregated with other variables, most interactions are ignored, and functional relationships are greatly simplified. The resulting abstract "model" of the real world may be embodied in a formal mathematical or computer language, or—much more commonly—it may be represented merely by a simplified view of the world as perceived by the human planner.

The usefulness of these abstractions depends on the fulfillment of two conditions: First, the planner must find the abstract world somehow more congenial to finding satisfactory plans than he does the real world; and second, a satisfactory plan in the abstract world must translate into a satisfactory plan in the real world.

The first requirement offers no great problem. The real world is, by design, abstracted in a way and to a degree that the planner does find it easier to manipulate the resulting model. Within the abstract world, then, the planner searches for a satisfactory plan by whatever means seem appropriate. In particular, he might compute an "optimum" plan if his abstract model permits the application of available optimization techniques (and if it is optimal to compute the "optimum").

The second condition presents more serious difficulties. The success of a plan hinges on whether it leads to satisfactory performance judged in "real" terms. The planner, in developing his abstraction of the real world, attempts to achieve a close correspondence between performance predicted by a plan and the resulting behavior in the real world. Only if his model passes this test
will a "good" plan necessarily lead to "good" behavior. This is by no means a simple requirement. 1

The correspondence between a planning model and the ultimate outcome achieved depends on three factors: the realism of the model, the accuracy in predicting planning variables, and the fidelity with which the plan is carried out.

A planning model is realistic to the extent that the transformation of action into outcome, when measured in terms of the abstract planning variables, is the same in both the model and the real world. This notion is diagrammed in Fig. 1.

Fig. 1 - Correspondence Between a Model and the Real World

Ideally, the model should represent a homomorphism of the real world (Minsky, 1963, pp. 441-443). If it does, a one-to-many transformation exists between any given plan (as described by a specific state of the model) and a set

It might be added parenthetically that planning in an organization plays precisely the same role as planning in a heuristic program. In both cases an inductive transformation first takes place from a complex world into a simplified model of it. The model is then manipulated in order to find a "solution" in abstract terms. Finally, this solution is transformed back into the complex form. The transformed version of the abstract solution may or may not provide a real solution. Failure signifies that some essential characteristics of the real world were lost in the transformation.
of detailed real actions that are consistent with the plan. The resulting real outcomes stemming from the alternative real actions will all have the same outcome in terms of the variables used in the abstract model.

The second factor affecting the correspondence between a plan and the eventual real outcome is the accuracy of predicting the values of the parameters used in the planning model. Because of prediction errors, a model may provide a perfect structural homomorphism of the real world and yet fail to predict outcomes. This problem can be mitigated by increasing the accuracy of predictions through more elaborate computation or by having more detailed and timely information; by maintaining a control system that detects significant deviations from the most recent predictions; or, if the other approaches prove infeasible, by changing the planning model to conform to the predictability of parameters.

The outcome achieved ultimately rests on the persons responsible for executing it. Although plans are often not followed with precision, the failure usually lies with imperfections in the plan and not in human perversity. A plan obviously cannot be adhered to if it is infeasible. But a feasible plan is only a necessary condition, and not a sufficient one. A plan may be "perfect," but if it is not followed it fails in its purpose.

A job shop schedule illustrates these concepts. The scheduling model used is often unrealistic, such as the one used in the standard backdating scheme that implicitly assumes infinite capacity. Under these circumstances, the plan can provide only a loose guide to action and a rough approximation of outcomes (in terms of such variables as scheduled delivery times, machine and labor utilization, and work-in-process inventory levels).

Alternatively, the schedule might be generated through a detailed simulation of the shop (Emery, 1961). The model itself may be a perfect abstraction, but
the predicted values of capacities, processing times, and so forth may--and, in general, will--be somewhat in error. As a result, the shop may not be able to follow the schedule precisely. Hopefully, however, the predictions will be good enough to serve as a useful guide between scheduling intervals.

Hierarchical nature of planning

Man can tolerate just so much complexity. The evidence available suggests that his limits are reached very quickly when dealing with interacting variables. When he encounters a problem that exceeds his threshold of complexity, he universally resorts to breaking the total problem into a hierarchy of subproblems. The ubiquity of hierarchical structures reflects our particular affinity for viewing complex systems in hierarchical terms rather than as an amorphous mass. It is highly questionable whether we could come to grips with complex systems in any other way--or, indeed, even recognize them as systems (Simon, 1962, p. 477).

The factoring of the total problem into subproblems is accomplished through a "means-ends" analysis. The end result--the solution of the problem--is decomposed into the means to achieve it. In order to effect these means, lower-level subproblems are generated. These, in turn, are factored through a similar analysis into still lower-level subproblems. Ultimately, tractable problems are generated that can be solved without further factoring, and each of these is linked with the global problem through a means-ends chain. In composite, the chains form an elaborate hierarchical structure of problems within problems (Newell, Shaw, Simon, 1959).

Planning, like any other complex problem-solving, exhibits a hierarchical structure. High-level plans establish requirements and constraints for planning at the next lower level. This process continues down through a means-ends chain to the lowest-level plans. At each level, plans are formulated without detailed
consideration of the lower-level plans used to implement it. "Strategists" need only assume "reasonable"-performance on the part of "tacticians" in conformity with higher-level constraints (Starr, 1964, pp. 67-75).

In order to achieve the simplifications that justify the means-ends analysis, each planning model generated by the factoring process is treated as being more or less independent of other planning models. If this were not the case—if each planner had to consider all ramifications of his actions—then the factoring serves no real purpose. A relatively high degree of real or assumed independence effectively isolates a problem sufficiently well that it can be handled within the planner's limited ability to cope with complexity.

Despite this need for independence, all interactions need not be neglected. The hierarchical structure of planning provides a method of dealing with the more important ones. To accomplish this, high-level plans must recognize interactions among lower-level planning models. For example, a monthly schedule of a chemical plant can take into account capacity constraints that cause interactions among planners preparing weekly schedules for each of the processing units within the plant. The high-level planner in selecting a given schedule tries to balance conflicting effects of the more important interactions.

The result of this compromise is communicated to lower-level planners in the form of an aggregate high-level plan. If the high-level plan is a "good" one, "suboptimization" within the constraints it imposes introduces only insignificant penalties. Thus, communication between levels is typically confined to aggregate variables. This provides a means of recognizing important
interactions, while at the same time allowing each planner to formulate his plans without considering the detailed actions of others.\(^1\)

A network of plans provides a hierarchical description of intended behavior. If consistency exists among different levels in the hierarchy, plans at all levels essentially describe the same behavior. The precision of the description, however, increases as high-level plans are amplified into lower-level plans. Each level adds information, so that at the end of the process the plan network (hopefully) describes desired behavior unambiguously enough to achieve the planned outcome within reasonable limits.\(^2\)

Plans are described in terms of variables having unit and time dimensions. High-level plans normally employ gross aggregate units and relatively long time periods. A high-level plan, for example, might deal in quarterly forecasts of major product lines, or— in the case of the Defense Department—fiscal year estimates of "program element" costs (e.g., Minuteman missile costs)(D.O.D., 1962). Each level in the planning process typically breaks down higher-level variables into more detailed variables that extend over shorter time intervals. Thus a quarterly product line forecast may be amplified into weekly item

\(^1\)The hierarchy of plans constitutes a "nearly decomposable" system (Simon, 1962, pp. 473-477). In such a system, each component is approximately independent of other components, and those interactions that do exist largely depend on aggregate, rather than detailed, variables.

\(^2\)Even at the lowest planning level, however, plans by no means spell out every detail. Conceptually, one should continue the hierarchical planning process until the cost of additional planning exceeds the expected marginal improvement over "concurrent" planning.
schedules, and program element plans are expanded into much more detailed "program authorizations."¹

Flow of planning information among levels

The strong downward bias of planning in the presence of interactions suggests that the more critical issues are resolved at the higher-level planning stages, and that the success of the organization depends less and less on plans generated at lower and lower levels. If this were not the case, the lowest levels could plan their own destiny unconstrained by higher-level considerations.² The fact that a high-level planner selects for execution one alternative plan over the others available to him implies that the preference ordering among alternatives is rather insensitive to lower-level planning.

¹I have been using the concept of planning level without attempting to define it. My notion of level is similar to Manheim's (1964, pp. 39-47). According to his definition (loosely translated), a plan at level B lies below level A if it partitions the behavior described by plans at level A into finer detail. A high-level plan can be amplified into two or more consistent but distinct lower-level plans. Therefore, a one-to-many transformation exists between a high-level plan and its lower-level plans. Consistency implies that the different lower-level plans are indistinguishable in terms of the variables used in defining the high-level plan. For example, an aggregate quarterly schedule of each product line has a higher-level than a detailed weekly item schedule because the weekly schedules normally include detail which is irrelevant in terms of the variables used in quarterly schedules. A vast number of alternative weekly schedules may be perfectly consistent with a given quarterly schedule.

²In a purely competitive economy, interactions are resolved in the market place. Adam Smith's "invisible hand," acting through the price mechanism, provides the "higher-level" constraints within which each firm attempts to optimize. If significant interactions--or "externalities"--exist, then prices alone do not contain enough information to achieve an efficient allocation of resources (in the Pareto sense). Additional information might be provided, for example, by a central planning agency that assigns production quotas. Even if this "solution" were politically acceptable, the technical and computational problems involved make it infeasible under most circumstances.

The "transfer prices" associated with intrafirm movement of goods serve an analogous role within the firm. They also suffer from the same shortcomings. Within the firm, however, interactions are likely to be relatively more significant, and higher-level planning as a means of coping with them becomes more feasible.
This notion of the relative superiority of high-level plans seemingly conflicts with the common view that an organization's total behavior is governed by the behavior of its lowest-level units. "For want of a nail, a kingdom was lost." "The Army is governed by its sergeants." "Take care of the pence, for the pounds will take care of themselves." These aphorisms ascribe the success or failure of the organization to its lowest levels.

This has the ring of truth about it. The behavior of the organization does ultimately depend on the composite activity at the lowest levels. However, if the organization as a whole is to achieve purposeful behavior in the face of interactions, lower-level activity must be guided by a hierarchy of higher-level planning constraints. Otherwise, lower-level success tends to be local rather than global. The want of a nail may--although it is highly unlikely--cause the loss of a kingdom, but the best way to reduce this risk is to develop an improved inventory control system. Generals may organize logistic operations; they do not serve as blacksmiths.

Regardless of the organization's commitment to decentralization, higher-level planners inevitably impose constraints on lower-level planning. They do this through the specification of organizational structure and resources, goals, policies, procedures, programmed behavior, transfer prices, and various informal and unstated requirements. These constraints are designed to induce lower-level behavior that is consistent with the means-ends decomposition of global goals.

To be sure, planning information has an upward as well as a downward flow. A higher-level plan may seemingly be generated merely as a composite of lower-level plans. However, even in extreme cases the higher-level planner has some power of veto over the proposed plans submitted to him. This represents a
potent instrument for guiding lower-level planning, as Secretary McNamara has demonstrated.¹

Geometric interpretation of hierarchical planning

It is useful, I think, to give a geometric interpretation to planning (Sisson, 1960; Manheim, 1964). In such terms, a composite lowest-level plan constitutes a point in an n-dimensional abstract space containing all alternative plans open to the organization.² Each detailed action or outcome variables used to describe a plan represents a dimension of this space.

Every point has an associated utility value, a function of the outcome variables that serve to define the point. The mapping into the utility value is, in general, many-to-one—that is, two or more points may be indistinguishable from the standpoint of the goals of the planner.

A high-level plan defines a relatively large set of points that are consistent with its specified aggregate, low-resolution variables. In general, not all points included in such a set have the same utility. Therefore, a high-level plan typically does not have a single utility value, but rather a distribution of values (Manheim, 1964, p. 48).

¹A planner with veto power acts, in effect, as a "trainer," exercising his authority through selective "reinforcement" of lower-level behavior. He "rewards" a planning process by accepting its output, and "extinguishes" a process by rejecting its plans (Minsky, 1963, pp. 426-430). This mechanism offers the distinct advantage of not requiring the higher-level planner to have detailed knowledge of lower-level activities; he must only possess a means of distinguishing "acceptable" from "unacceptable" plans. Its great disadvantage is that it may be a much less efficient method of guiding lower-level behavior than more direct intervention. With only one-bit "go, no-go" information, lower-level units may move very slowly—if at all—toward improved planning. In practice, of course, additional information is provided in the form of specific requests for modifications to plans, veto messages, and so forth.

²The plan space does not include all "real" alternatives, but only those that can be described in terms of the abstract variables used in formulating plans. More will be said about this issue later in the paper.
The utility ultimately achieved depends on the way in which a high-level plan is amplified into the lowest-level plans, and typically this information is not available to the high-level planner at the time he selects a plan. He must therefore make a choice (normally not consciously) on the basis of a subjective estimate of the probability distribution of utility values (or its parameters) associated with each alternative high-level plan. The hierarchical planning process then successively narrows down the set of points remaining in the region constrained by higher-level plans. Ultimately there remains only a single point having a single utility value (but which reflects, of course, the probabilistic nature of outcomes and the planner's attitude toward risk and uncertainty).

The point finally selected in the abstract plan space eventually is used to guide organizational behavior. From this certain actions and outcomes will result. The control system measures these actions and outcomes, and translates them into the abstract variables used in planning. In general the "actual" point will not coincide with the planned point, since neither actions nor outcomes will necessarily occur as planned. This should be recognized in the control system by allowing some tolerance for deviations from the plan. A deviation is thus treated as significant only if the actual point falls outside the region surrounding the planned point.

A simple example will clarify these concepts. Suppose that an organization has a three-level planning process. The highest level fixes the budget. Production and marketing plans are then generated at the intermediate level, consistent with the higher-level budget. Finally, at the lowest level detailed schedules are formulated to implement the plans.

Assume that plans at all levels can be described in terms of only two dimensions, "sales" and "profit." However, a lower-level plan extends over a
shorter time span than the corresponding higher-level plan. Confined to two dimensions, we can represent a composite schedule as a point in a plane. A higher-level plan is represented by a region containing the set of points consistent with its plan variables. This is shown in Fig. 2.¹

![Geometric Representation of Hierarchical Planning](image)

Fig. 2--Geometric Representation of Hierarchical Planning

A given budget includes only certain portions of the total plan space. Points outside this region cannot be reached within the constraints imposed by the budget. In Fig. 2, for example, budget $B_1$ restricts the set of points to the region surrounded by the solid curve labeled $B_1$. A different budget, $B_2$, has associated with it a different region, the one included within the broken curve. Suppose that the highest-level planner chooses budget $B_1$ over budget $B_2$, presumably because he perceives it to include a more satisfactory set of points. "Satisfactory" is defined in terms of the characteristics of the distribution of utility values associated with each set--its mean, variance, range, or some other statistic.

¹Fig. 2 is similar to Manheim's representation of hierarchical planning (Manheim, 1964, p. 21).
Intermediate planning further narrows the region of alternative schedules available to the organization. In order to be consistent with budget $B_1$, the points included within the feasible region of the production and marketing plan must all lie within region $B_1$. Scheduling, the final steps in the planning process, then selects a point within the region constrained by this plan.

Consistency and iterative planning

For a plan to be perfectly consistent with a higher-level plan it must lie wholly within the region of the latter. Therefore, a set of consistent hierarchical plans is represented by a set of nested regions, with the innermost region associated with the most detailed plan. A higher-level region serves merely to confine the next lower-level region, which in turn confines the next region, and so on down to the most detailed plan. It is this detailed plan that guides behavior during the actual execution of a hierarchy of plans.

A higher-level plan thus guides lower-level planning, and only indirectly effects actual events. It fails in this if the detailed plan does not conform to it, since the detailed plan shapes behavior whether or not it is consistent with higher-level plans. A high-level planner must therefore take great interest in the consistency of the hierarchy of plans.

Consistency requires realistic higher-level planning and a control system that encourages compliance with plans. Without these, no logical assurance exists that any point included in a lower-level plan necessarily falls within the region defined by a higher-level plan. Conceivably, a higher-level plan may be completely infeasible. In this case no point is included in its (null) region, and therefore no feasible lower-level plan can be consistent with it.
Realism in planning can be achieved in two ways. The most common approach is to confine plans to well explored regions of the organization's plan space. An organization typically pursues relatively stable goals, and therefore each new plan differs only modestly from previous plans formulated under similar conditions. This continuity enormously simplifies high-level planning, since it permits the generation of plans with a minimum amount of information—namely, information about past performance and the significant internal or external changes that have occurred subsequently. (Determining the implications of such changes may be exceedingly difficult, however.) (Cyert and March, 1963, pp. 111-112).

Unfortunately, such conservatism bars the possibility of obtaining really fundamental improvements. To overcome this disadvantage, the planner must make a more far-reaching search into uncharted territory within the organization's plan space. To achieve realism with this more radical approach, the planner can no longer rely principally on information about past performance; he must have access to a much greater source of information in order to determine the probable consequences of untried plans.

This information is acquired primarily through an iterative dialogue with lower-level planners. Based on broad higher-level guidance, lower-level planners propose plans which they consider realistic (perhaps as a result of proposals submitted to them by still lower-level planners). The higher-level planner may well modify such proposals, but hopefully in a way that does not do violence to their realism. The process continues until "convergence" is achieved.

Lower-level planners have the responsibility to point out unrealistic aspects of higher-level plans. If the planning system is strongly biased against this upward flow of information, the organization will end up with unrealistic
high-level plans and inconsistent lower-level plans. The lower levels may bury the inconsistencies by an elaborate shuffling of their resources, by bookkeeping slight-of-hand, or by postponing as long as possible the inevitable report of deviations from the higher-level plan. The inconsistencies exist none the less, and they are apt to prove costly.

**Characteristics of the planning process**

Planning has very much the same character regardless of its level. It is an iterative process, and so defies any simple scheme for describing the sequence of steps involved. Nevertheless, it seems useful to try. The following sequence is admittedly somewhat arbitrary, but I hope it gives the essential flavor of the planning process (Cyert and March, 1963, pp. 84-86).

**Determine planning parameters.** The first step in planning requires retrieval of data from the organization's data base. This is followed by operations on the data to transform them into predicted values of planning parameters. These parameters include higher-level plans and constraints, functional relationships among variables, physical and cost standards, and appropriate exogenous variables such as forecasted sales over the planning horizon.

**Propose alternative plans and generate their predicted outcomes.** The second step involves manipulation of the planning parameters in order to determine the consequences of alternative plans. Each plan proposed by the planner (components of which may have been proposed by lower-level planners) defines

---

1The issue is of course not as clear-cut as I might seem to imply. A lower-level planner obviously has his own biases, and probably they do not run in favor of a plan that presses him too energetically. A certain amount of gaming takes place when planners negotiate. Stedry discusses some of the issues involved (1962).
behavior in terms of controllable action variables. The consequences stemming from these actions are described in terms of outcome variables related to the goals of the planner. The mapping of actions into outcomes typically requires, of course, an extremely complex model (which may be either a formal or informal—that is, "intuitive"—model).

**Selection of the "best" alternative.** Based on a display of alternatives analyzed in the previous step, the planner selects the one that appears most suitable. His choice is straightforward if he has available a formalized utility function of the outcome variables: he merely chooses the alternative having the highest value.\(^1\) In the absence of a formalized utility function, the planner must exercise his judgment by selecting that plan that seems most suitable. If he feels that the expected improvements justify the cost of further search, he may generate additional alternative plans.

**Translate into a form for lower-level units.** The variables used in a planning model need not be the same ones employed to communicate with lower-level planners. In some cases a higher-level planner may factor variables into more detailed form before assigning them to lower-level planners. For example, plans formulated in terms of aggregate production might be broken down into a standard mix of individual products. (This obviously assumes that the outcome as predicted by the higher-level planning is essentially invariant with respect to the factoring.) In other cases, a planner might choose to aggregate variables before passing them down to lower-level planners. He might do this, for example,

---

\(^1\)This will not represent the "optimum" plan unless all alternatives have either been considered explicitly or have been rejected by some logical means as being inferior to the set of explicitly considered plans. The true—but unknown—optimum recognizes the cost of searching for improved alternatives, and therefore rejects a stubborn hunt for the "perfect" plan.
in the spirit of decentralization, relying on the lower-level planners to develop their own detailed plans constrained by aggregate variables that have been found feasible through a more detailed analysis.

**Control of the plan.** The final step, control of the approved plan, is not a part of the planning process itself, but is a vital ingredient to its success. The control system compares actual performance against the plan, and determines those deviations outside control tolerances. "Exception" messages identify these deviations and trigger any necessary re-planning. Such planning involves the same sequence of steps used in the creation of the earlier plan. It examines the current situation anew, constrained only by higher-level plans and the consequences stemming from past actions.

The feedback information provided by the control system serves two important functions. For one thing, it encourages more realistic initial planning and closer adherence to plans once approved.\(^1\) Secondly, the control system guards against excessive deviations from current plans that cause a partial breakdown in coordination within the organization. In setting control limits, ideally one would like to balance these costs against the costs involved in increasing the frequency of planning. Even though in practice these costs are extraordinarily difficult to estimate, the concept offers a useful guide nevertheless.

\(^1\)The encouragement to adhere to current plans should not be so great that it causes lower-levels in the organization to stick with an outmoded plan. This induces the typical organizational behavior associated with unrealistic plans (discussed in the previous section). Nevertheless, there is ample evidence to suggest that without adequate control planning tends to become a superficial exercise. One need not search very far back within the archives of the Defense Department, for example, to discover past instances where original plans proved to be wildly unrealistic. Part of this is due to the advanced technology involved in modern weapons; but at least part of the blame rests with a system that positively encouraged over-optimistic initial plans and insufficiently discouraged cost escalations and schedule slippages. These problems have been widely recognized, of course, and many of them are currently being corrected.
The steps involved in a planning process appear to be largely independent of its level. To be sure, the scope of the process varies greatly with level. High-level processes deal with aggregate variables over relatively long intervals, while low-level processes deal in much finer detail. The models used to analyze alternatives, and the utility function by which an alternative is measured, are typically much easier to define for low-level planning than for high-level planning. Nevertheless, the similarities, rather than the differences, best characterize the planning process at all levels.

Planning constraints

The fact that a high-level plan normally provides a low-resolution description of behavior does not imply that the organization necessarily can exercise only imprecise direction over low-level behavior. In amplifying its plan into plans for the next lower level, each organizational unit must conform to constraints of various forms. The specification of the controllable constraints provides a powerful tool for directing lower-level activity.

These planning constraints are the result of higher-level decisions (which, as we shall see, are themselves the result of planning). The more important ones appear to be the following: 1) organization structure and goals; 2) resources available; 3) technology available; 4) performance programs; and 5) information available. Let me discuss these briefly.

The organization structure reflects the hierarchical partitioning of the activities within the organization as a whole. Each unit has assigned to it certain responsibilities. In carrying out its responsibilities the unit is guided by its goals. The goals provide a utility function by which the planner measures the relative attractiveness of alternatives available to him. Presumably
the planner attempts to maximize the utility function subject to higher-level plans and the planning constraints (but, in practice, with only partial success).

Plans must recognize the availability of resources. These include capital, managerial, manpower, material, and energy resources. Instead of "congealed" physical resources, a unit may be granted a flexible resource in the form of an aggregate money constraint (which of course congeals upon use). Closely related to resources is the technical knowledge made available to a unit. Although the fund of technical knowledge is by no means completely controllable by the organization, it is certainly subject to some control through the organization's allocation of resources to research and development activities.

"Performance programs" of one form or another serve to limit the ways in which a planner can transform a high-level plan into lower level plans (March and Simon, 1958, pp. 141-150). Included within this constraint are standing plans, standard operating procedures, policies, computational algorithms, and so forth. These programs constitute a language by which higher levels within the organization partially prescribe the planning process to be used at lower levels.¹

Information represents the raw material for planning. Each unit plans not on the basis of the state of the environment, but on the perceived state of the environment as represented in its data base. The data base may or may not provide a suitable analogue of the environment; the data it contains are certainly subject to errors, delays, and transformations of varying usefulness.

Higher levels within the organization can partially govern behavior at lower levels by means of the information they make available. For example, the

¹Performance programs are somewhat analogous to computer subroutines.
organization employs "uncertainty absorption"--in which variables subject to uncertainty are communicated as "stipulated facts"--as a means of achieving some direction and consistency over lower-level planning (March and Simon, 1958, p. 165). Transfer prices provide another means for guiding lower-level planning.

Each of the planning constraints exhibits its own hierarchy. This is clearly true in the case of organization structure, goals, resources (including technical), and information. Performance programs are also hierarchical, since the output from a high-level programmer tends to constrain lower-level programmers.

**Adjustment of planning constraints**

A single-use plan gives only an ephemeral description of behavior; it serves only over the planning interval or until superseded by a revised plan. A planning process, on the other hand, has a much more persistent effect, since it continues to govern behavior as long as the planning constraints remain unchanged. An organization can therefore achieve long-run success by developing suitable planning processes. The planning constraints provide the vehicle for doing this.

The organization can improve planning by making appropriate adjustments to the planning constraints through a continuous process of hierarchical adaptation. The adjustments should normally be "small" ones in order to increase the stability of planning and make it easier to discern the effects of the changes. With such knowledge, those adjustments that bring improvements can be continued and strengthened, and those that worsen behavior can be rescinded.

The frequency of these adjustments varies with hierarchical level. A high level constraint--such as organization structure and goals--requires infrequent changes, while low-level procedures may be revised relatively often.\(^1\) (Simon, 1962, p. 166)

---

\(^1\)The situation is perhaps analogous to nested loops in a computer program. The loop index of an outer loop is changed less frequently than the indices of the inner loops that it controls.
Each execution of a planning cycle provides little evidence for making changes in the process, especially in the presence of random variations. Since high-level planning ordinarily has a relatively long cycle time, changes in high-level planning constraints should be correspondingly infrequent. The stability thus provided contributes to the predictability of behavior and allows time for persons within the organization to learn their jobs and specialize their activities (Zannetos).

The planning process is itself subject to planning: an adjustment to a planning constraint constitutes an output from a higher-level process. The search for improved inventory control procedures, for example, involves planning—the comparison of alternative procedures and the selection of the best one. This planning is in turn subject to constraints. The operations research group developing inventory procedures is constrained by its organizational structure, goals, resource budget, information, policies, and procedures. One can develop recursively still higher-order processes (procedures to change procedures to change procedures, say).

The notion that adaptation comes through changes in the planning process has exceedingly important implications for management. If a manager hopes to exert a lasting and persistent influence over the behavior of the organization, he must work primarily through planning. He should not focus attention on ad hoc adjustments to plans, but rather with the process that generates them (and with the process that generated that process).

**Computer planning models**

The planning process discussed up to this point represents a conceptual viewpoint largely independent of the methods used to implement it. The topic
I now want to address is the role that computers might play in this process—particularly at the higher-levels within the organization. Despite a few caveats (Deardon, 1964), there seems to be a general consensus that computers will contribute increasingly to planning at all levels.

Skeptics would certainly have no difficulty in marshalling ample empirical evidence to support a less sanguine view. The results achieved to date in using computers for high-level "strategic" planning have been relatively drab and meager compared with their widespread use for low-level "operational control" (Deardon, 1964, pp. 129-130). Nevertheless, I can see in this experience no compelling evidence that supports a view that inherent limitations preclude the use of computers in strategic planning.

Developing formal computer models to aid the planner obviously presents a task of great difficulty—but not a unique one. A planner cannot avoid the use of a model. Whether he relies on a formalized model or his own intuition, he necessarily forms an abstract representation of reality (Craik, 1943). There is every reason to suppose that formalized computer models can assist the planner in doing this.

It may not be possible—and certainly not desirable—to develop formal models of all high-level planning processes. Initial effort should be concentrated on decisions that are in some sense repetitive, especially important to the success of the organization, and relatively amenable to a formalized description. In a manufacturing organization, activities concerned with the allocation of resources often offer the most attractive rewards.

The use of computers in planning does not imply that the planner must abrogate his responsibilities to the machine.¹ A man-computer system, rather

¹This is so obvious a point that I would hesitate to make it were it not for the fact that the contrary is sometimes suggested.
than one based on either component alone, offers the best hope of achieving a major improvement in high-level planning. To the human is relegated the job of proposing alternative plans and judging their suitability; to the computer is left the computational task of transforming the action variables into outcome variables (by use of analytical or simulation models).

I hasten to point out, however, that even in a man-machine system of the type described the computer inevitably makes some "decisions." In transforming action variables into outcome variables the computer employs any feasible algorithm that appropriately embodies the planner's decision rules. For example, in determining the consequences of a proposed increase in the aggregate production rate, the computer must make some assumptions about the allocation of production capacity. In doing this, the computer should distribute production in an "optimal" fashion subject to the specified aggregate capacity constraint (Emery, 1960). Based on these detailed "decisions," the computer can then determine such outcome variables as manufacturing costs, expected stockouts, and so forth.¹

Does this encroach on the planner's authority for decision making? Exactly the opposite is the case. The planner must participate in the development of the algorithms by which his aggregate decisions are amplified into detailed plans. If the algorithms incorporate the planner's goals, he can then formulate high-level plans with a reasonable confidence that the resulting detailed plans will be satisfactory.

In the absence of such a formalized model the high-level planner has little assurance that his plans will be translated faithfully. Although he may

¹The computer should also provide the planner with information helpful in making the aggregate production rate decision--the relation of the projected inventory level to the "optimal" level, say.
participate in the development of procedures and personnel to generate detailed plans, nevertheless a great deal of "noise" is introduced during the hierarchical translation process. Lower levels in the organization have a vast catalogue of devices for frustrating the execution of grand strategy, most of which are evoked merely out of misunderstanding and confusion. Computer planning will certainly not eliminate these problems, but it can mitigate them by transferring part of the translation process to a formalized model.

**Design and use of formalized computer models**

Computer planning models, like the less formal ones, must have a hierarchical structure (Dalkey, 1962). Factoring is necessary, as before, in order to break up the global task into manageable subtasks. This obviously introduces problems of coordination. However, the problem is much less severe when employing computer models than it is with conventional planning.

For one thing, the global task need not be fragmented to the same extent as before, since the computer can handle much greater complexity than the unaided planner. In effect, each model can have a wide "span of control," thereby reducing the number of submodels and the hierarchical depth of global planning. "Sub-optimization" will still occur (if a submodel happens to "optimize"), but it can be made more comprehensive and hence less subject to the well-known (but often overlooked) penalties associated with suboptimization.

Furthermore, the enormous input-output capacity of modern computers permits a close link of a computer model with the data base of the organization.¹ This facilitates communication among models (via the data base).

---

¹The weakest link between the computer and the data base is no longer the input-output hardware, but rather the language by which wanted data are described and retrieved. Lombardi (1964) argues this point cogently, and suggests possible approaches to the problem.
Computer models will not change the iterative nature of planning. This scheme provides an efficient means of communication between hierarchial levels in formalized as well as conventional planning. By such means the planner can reduce discrepancies existing between the highest-level and lower-level models.

Discrepancies can arise when lower-level planning models do not explicitly recognize the constraints imposed by a higher-level plan. Feasible algorithms often do not exist for amplifying a plan subject to such constraints. In this case, an iterative modification of the lower-level variables can sometimes be used to achieve closer consistency with the high-level plan.

Discrepancies may also arise if the planner chooses an infeasible plan. A high-level planning model gives only an approximation of lower-level models. The more detailed analysis performed by the lower-level models reveals information submerged from view of the high-level model. Upon closer lower-level scrutiny, a plan that appears attractive in terms of the high-level model may be shown to be infeasible in some respects.

The differences between a high-level and lower-level analysis may or may not be significant. A discrepancy is significant if it is large enough to bring to the attention of the planner so that he has the option of revising his original plan based solely on the high-level model.\(^1\) If the planner rejects this option, he implies acceptance of the amplified (and presumably more realistic) version. If, on the other hand, the planner does revise his plan in light of the additional information gained from the more detailed analysis, the new plan undergoes a detailed amplification similar to the previous one. Through a (presumably converging)

---

\(^1\)Tolerance limits thus play the same role here as they do in a conventional control system. In the man-machine system, however, an explicit "exception report" comes back during the iterative planning process rather than during the attempted execution of an unrealistic plan.
man-machine iterative process, consistency within required tolerances will hopefully be reached between outcomes as predicted by the high- and lower-level models.

Computer models offer the obvious advantage of speeding up the planning process. Greater speed in planning permits quicker response to changes in existing plans if that should prove necessary. Of far greater importance, the computer's speed permits relatively quick response to the planner's proposed plans. The planner can therefore evaluate a larger number of alternative plans, and consequently stands a better chance of finding a superior one. Furthermore, with a short response time it is reasonable to expect better decisions, since the planner can retain a closer grasp of a complex problem over the reduced response interval. Efficiency of a sequential, hierarchical search

The highest-level computer model deals in aggregates and gross approximations, and therefore the space of its perceived alternatives is very much smaller than that of the lower-level models. This space remains, nevertheless, enormous and forever largely hidden by the "curse of dimensionality." However, a rapid man-

---

1 Information acquired during the lower-level analysis should be incorporated into the higher-level model (Dalkey, 1962). Hopefully, this might involve nothing more than modifying the high-level parameters. In more extreme cases the information can be added to the high-level model only by more basic changes in its algorithms or structure. In either case, the interplay between the two levels constitutes an adaptive or learning process, with the composite lower-level models serving as teacher (Newell, Shaw, and Simon, 1960; Minsky, 1963, pp 425-435). Adaptation reduces the differences in the outcome of a plan as perceived at the different hierarchical levels. Its goal is to increase the capability of the higher-level models to predict the outcomes as measured by the lower-level models. This permits the planner to choose a high level plan with confidence that it can be amplified into low-level plans having (at least roughly) the predicted consequences.

2 The need for quick response applies almost exclusively to low-level plans extending over a short planning horizon. Seldom is it desirable--and often it is extremely undesirable--to alter high-level plans on the basis of the scanty evidence furnished by very recent and "timely" data. To do so is likely to introduce more "noise" than "signal" into the planning process, since the data would normally represent too small a sample of the environment to contain much information.
machine sequential search of this highest-level space will, hopefully, prove both efficient and effective in locating a high-payoff region. Its success rests on the usefulness of information acquired through preceding probes of the space in directing the planner to regions with improved outcomes.

The highest-level plan selected by the planner in this sequential fashion confines the next lower-level search to a tiny fraction of the total plan space. Within this highly restricted region, the planner again employs a sequential search process to locate a still smaller region having attractive outcomes. This hierarchical process continues down to the lowest level. Each level slashes the remaining search region by a high factor, reducing the total plan space exponentially. Thus, by permitting the designer to concentrate on more or less independent aspects of the total problem, sequential hierarchical planning becomes both feasible and efficient.

Adjustments to the planning model

The search scheme outlined above has as its purpose the location of a satisfactory plan out of the set of all alternative plans that can be generated and described by the algorithms and variables used in the composite hierarchical planning model. This set by no means includes all "real" alternatives, since not every possible alternative can be mapped into a given planning model. The plan space defined by a model's action and outcome variables represents only an infinitesimal portion of the organization's "real" plan space. Furthermore, only a very small fraction of the points included in even this reduced space remains accessible to the planner. For example, the model will not generate an inventory distribution that is "nonoptimal" in terms of the model's algorithm for allocating aggregate production.
The planner naturally seeks improved plans out of the real alternatives available to him, and not just the potential candidates found in the space of the existing planning model. This does not mean, however, that he should have complete access to every conceivable alternative. On the contrary, the very purpose of the planning model is to confine the planner's search to a "good" region of the real space, and to exclude "bad"--and therefore irrelevant--portions. The planner must aim at improving the model so that it provides access not to a greater number of alternatives, but rather to fewer--albeit, better--alternatives. The ideal model makes accessible only a single plan--the optimum one in terms of the real alternatives and real goals of the organization.

This philosophy leads the planner to employ modifications in the planning model as the predominate mechanism for generating improved plans. The planning model should be highly "parameterized" to facilitate these changes. For example, an inventory control algorithm should include a carrying cost parameter that can be modified readily. If a temporary shortage of funds within the organization forces a reduction in inventory, the planner can accomplish this by increasing the carrying cost parameter. The "knob" to adjust the parameter can simply be turned until the desired level of inventory is found through a trial-and-error process.\textsuperscript{1}

A change in a algorithm represents a more basic type of modification in the planning model than does a change in its parameters. For example, an inventory distribution algorithm may be revised to include some refinement--the explicit consideration of capacity constraints, say. The planning model should be designed

\textsuperscript{1}The distinction between a parameter of this sort and an action variable is essentially arbitrary. The planner uses either as a means of exerting influence over lower-level planning (in the example, the generation of specific inventory order quantities). Normally, however, the value assigned to a parameter persists over several planning cycles, while an action variable may change with each cycle.
to permit great flexibility in making such changes (Lombardi, 1964).

Still more fundamental changes are possible, of course. For example, the structure of the model may be altered by combining two or more activities that previously were planned independently. This might occur, say, when the scheduling of two factories is combined in order to find joint optimum schedules instead of independently suboptimized schedules.

**Search for improved planning models**

Making adjustments of the type described requires planning, and therefore involves a search among alternative plans. Once again, the planner confronts a vast space of alternatives—in this case, the space of alternative planning models rather than alternative single-use plans. In order to find a fundamentally improved model he must employ efficient search techniques.

Ad hoc adjustments to an existing planning model provide a means of searching for improved models. However, even the most efficient sequential search techniques reveal only a minute part of the abyss of alternative models. The planner must employ more powerful techniques in order to find really fundamental improvements. A hierarchical search can again prove useful here. This requires the development of a higher-level "metamodel" designed specifically for the purpose of exploring for improved models.

Industrial dynamics models seem particularly well suited for such experimentation (Forrester, 1961). There are a number of reasons for this. First of all, industrial dynamics models facilitate an investigation of systems having many interacting variables. This is precisely where one must focus attention when seeking fundamental systems improvements. The structure of the organization and links between its components are basic determinants of organizational behavior.
When doing any modeling, one must choose appropriate variables for describing outcomes. In the case of the metamodel, its outcome variables must provide a means of comparing the performance of alternative planning models. The dynamic characteristics of each alternative model are among the most important measures for this purpose. The behavior induced by a given planning model must be described in terms of its stability, the dynamic response to changes in exogenous variables, and so forth. Industrial dynamics models—as the name implies—are ideally suited for such investigations, since they provide a trajectory through time of any desired variable.

Finally, the language of industrial dynamics—DYNAMO (Forrester, 1961, pp 369-381)—makes it relatively easy to develop and modify a metamodel. Because of this, an analyst can employ what amounts to a hierarchical exploration of planning systems. A "first-level" model can be developed to ascertain the more sensitive variables and parameters. Those portions of the system that prove especially important can then be modeled in greater depth. This process can continue as long as it yields worthwhile improvements in the metamodel.

By this means the planner can explore the space of the metamodel until he judges that the cost of further search would exceed its expected gain. The terminal values of the metamodel variables and parameters—decision rules, delays, the flows of resources and information, and so forth—are then transformed into the corresponding values in the planning model. If the metamodel provides a sufficiently accurate abstract representation of the planning model, then the planning model should have roughly the same dynamic characteristics as the chosen metamodel.

Conclusion

I would not deny for a minute that the scheme outlined here—a system of
hierarchical models—requires an extremely ambitious program to implement. A prudent manager would do well to treat very gingerly any suggestion that such a system be developed for his organization. I am convinced, however, that systems roughly of the type described here (for which I claim no great originality, of course) will play a central role in the management of organizations.

An organization currently pays an inordinate price for planning. It pays in the form of the costs required to sustain its present planning "model"—the organizational hierarchy engaged in the amplification of high-level plans into more detailed form. It also pays in the form of unnecessarily poor performance.

The generation of substantially better plans is so complex a task that only through an elaborate computer system can we hope to come to grips with it. It appears that the system will take the form of a hierarchy of planning models. Planning with the formalized models will become a man-machine process involving a sequential, iterative search through the hierarchical plan space of the organization.
REFERENCES


<table>
<thead>
<tr>
<th>Date Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCT 16 '75</td>
</tr>
<tr>
<td>SEP 11 '75</td>
</tr>
<tr>
<td>APR 10 '76</td>
</tr>
<tr>
<td>SEP 5 '76</td>
</tr>
<tr>
<td>JUL 2 '80</td>
</tr>
<tr>
<td>JUL 12 '80</td>
</tr>
<tr>
<td>DEC 15 1936</td>
</tr>
<tr>
<td>JUL 7 0 1989</td>
</tr>
</tbody>
</table>

Lib-26-67