Modeling the Evolution and Breakdown of Adaptive Systems: A Review

Rasmussen, Jens

Publication date:
1993

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Rasmussen, J. (1993). *Modeling the Evolution and Breakdown of Adaptive Systems: A Review*.

MODELING THE EVOLUTION AND BREAKDOWN OF ADAPTIVE SYSTEMS: A REVIEW

Jens Rasmussen

HURECON, Smorum Bygade 52, Dk 2765 Smorum, Fax (45) 42 653867.

Work funded by Risø National Laboratory

TABLE OF CONTENTS:

INTRODUCTION

Understanding the evolution and breakdown of adaptive systems is becoming increasingly important as modern information technology is changing work systems from being a set of separate, functionally rather de-coupled units or organizations to a number of overlapping, functional networks that cut across units and organizations and thus integrates work activities into tightly coupled functions. At the same time, the pace of technological change is very fast. The effects of the introduction of a new tool or of the invention of a new product will not have settled, before the effects of another change starts to propagate through the integrated functional networks.

In this situation, careful pre-planning of work processes, instructing the staff in proper work procedures, and coordinating work by a hierarchical organization, as it is proposed by 'scientific management' and Tayloristic theories, are no longer effective. Such open-loop, feed-forward approaches to design of work systems has to be replaced by a feedback design approach. That is, to make it possible locally to respond rapidly and effectively to the multiple, propagating changes, work processes are not preplanned in detail but the local actors are informed about the objectives and goals of the work system; they are supplied with adequate resources (capabilities) in terms of knowledge, tools and materials; and a measuring system is put in place to feed back the necessary performance information. In other words, the design strategy is to rely on the ability of the individual staff member *to adapt effectively to the changing conditions.*

MODELS OF MODERN WORK SYSTEMS

This situation, in turn, changes the requirements to models of work systems used to support our understanding of the behavior of such systems and to predict behavior of new conceptual system designs. In case of practically separated and stable work systems when work processes are preplanned, system models and system design in terms of task procedure representations are adequate and verification can be found by prototype experiments with properly trained actors and the actual system can be controlled and its stability (reliability) ensured by monitoring the adherence to the instructions.

This is not the case with work systems based on the feedback design principle. Such a system is designed to change with the changing conditions of the activities of the individual staff member, conditions which cannot be predicted during system conception. System design can be based on the creation of resource envelopes around the individual actors and of proper mechanisms for propagation of objectives and performance measurements. In highly integrated, tightly coupled systems, small scale, local experiments to evaluate the implications of new technologies are not always possible and understanding the properties of adaptive systems in order to predict the effects of changes become increasingly important. In contrast, prediction of the stability of the resulting systems and their properties when capability limits are surpassed, calls for models of the performance of the entire system.

The problem can be illustrated by James Watt's design of the steam engine which could be based on the aggregation of functional elements which were designed and explained by the interaction among the functional components. However, when feedback principles were introduced in terms of the 'flying ball governor' the occasional instability of steam engines were only understood and mastered when Maxwell studied the phenomenon by means of an abstract model of the entire relational network by means of differential equations. We are facing the problem that models of work systems created by chaining causal input-output properties of the elements of a work system are not capturing the higher level properties of the closed loops embedded in the systems. Or, in conclusion, even if we can design effective, complex adaptive work systems, we do not have models which can support our understanding of the stability of such systems. This modeling problem is clearly demonstrated by the analysis of recent major industrial accidents.

Causes of System Breakdown

Recent accounts from analysis of major system failures, that is, accidents invariably identify 'human error' as a major contributor in 70-90% of the cases and design efforts are typically focused on the control of human errors. However, as it has been argued elsewhere (Rasmussen, 1990 a,b), the notion of 'human error' is ambiguous. The identification of a person and a particular act causing an accident is very much a question of the stop-rule applied for the causal backtracking during analysis. Furthermore, whether the contributing act is considered 'an error' or if it is a quite normal way of behaving caught by unlucky circumstances, is often a matter of taste. Analyses of most major accidents reveal many conditions that, in hindsight, are considered accident precursors and often it is found that a potential accident has only been waiting for an additional triggering event. Unfortunately, however, an unlimited number of conditions are quite harmless under the usual work conditions but become dangerous in certain combinations. Accidents are not caused by stochastic coincidence of several human errors and technical faults, they are shaped by basic, structural properties of work systems.

When a particular person contributes to an accidental course of events we have an occurrence of a mismatch between human behavior and system requirement. Either human behavior has changed from normally successful performance to a degree that is unacceptable to system function, or the system requirements have changed in a way that makes the usual human behavior unacceptable. In both cases, we are faced with a human-system interaction that is too narrowly adapted to the normally successful conditions, either due to system design, work planning, or local experience. The basic problem then is an adaptation to that close match between work requirements and behavioral patterns which is also the hallmark of expertise.

This close relationship between human error mechanisms and the adaptation to work requirements necessary for the attainment of expertise has been discussed in detail elsewhere (Rasmussen, 1993, 1990 a,b). In most work situations, the individual actor is faced with many degrees of freedom with respect to the composition of acceptable work procedures. Expertise at all levels, manual work, work planning, and strategic decision making, depends on an adaptive process by which an effective work practice evolves. During this process 'errors' are intimately connected to exploration of the boundaries of acceptable performance. Performance will be optimized according to the individual's subjective process criteria within the boundaries of the available resources. Unfortunately, perception of the qualities of the work process itself is immediate and unconditional and will govern the local adaptation to subjective performance criteria, while the ultimate product of work from these adaptive trials can be considerably delayed, obscure and frequently conditional with respect to other multiple factors. Short-cuts and tricks-of-the-trade will frequently evolve and be very efficient under normal conditions while they will be judged serious human errors when they, under special circumstances, lead to severe accidents.

If it is accepted that patterns of human behavior in any system is shaped dynamically by situation depending factors and subjective performance criteria and that 'human errors' intimately connected to effective adaptation to the work requirements, then we cannot evaluate the stability of a system and guard it from breakdown by studying its performance at the individual task level. We have to study the nature of the closed loops involved in adaptive behavior at a higher, more system related level.

'Thermo-dynamic Description of Behavior

The objectives and constraints which must be respected by the actors for work performance to be successful define the boundary conditions of a work space within which the human actors can navigate freely. The choice among several possible work strategies for navigation within the envelope specified by these boundaries depends on subjective criteria, related to process features such as time spent, work load, pleasure, excitement of exploring new territory, etc. Activity, therefore, will show great variability due to local, situational features which, in turn, leads to frequent modifications of (and shifts among) strategies. Activity will be characteristic by local, situation induced variations within the work space calling to mind the "Brownian movements" of the molecules of a gas. Such variability will give ample opportunity for the actors themselves to identify "an effort gradient" and management tends to always make a "cost gradient" very visible to the staff.

The result very likely will be a systematic migration toward the boundary of acceptable performance and, when crossing an irreversible boundary, a system failure may occur, see figure 1. This analogy indicates the need to explain the breakdown of adaptive systems at a conceptual level higher than that used for description of the behavior of human individuals. Metaphorically speaking, we have to shift our level of consideration from particle physics to thermo-dynamics.

Survival of the Fittest

When considering the self-organizing and adaptive nature of work performance of individuals and organizations in a dynamic environment in which trial-and-error leads to work performance composed by 'survival of the fittest' work procedure, the analogy to biological evolution is clear. The normal optimization of work procedures and the switch between different regimes of behavior characterizing 'ultra-stable' adaptive systems call to mind Gould's distinction between modes of evolution such as a) Variations within populations, b) speciation, and c) very long-term macro-evolutionary trends. The questions thus is: Can the recent development of biological models of adaptation and evolution give inspiration to new approaches to the present problem?

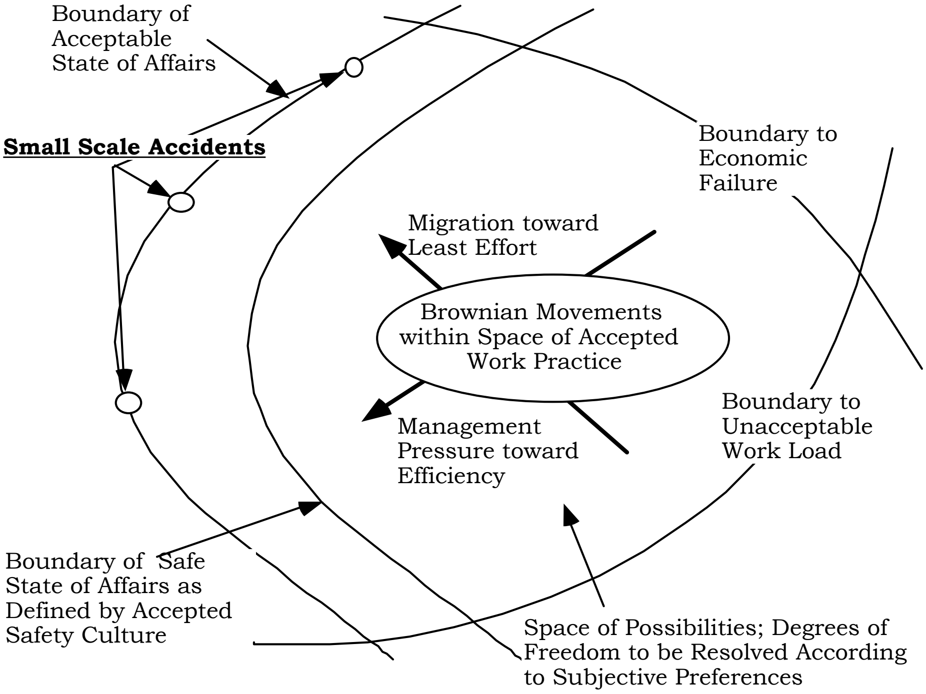


Figure 1 illustrates an analogy between migration toward boundaries to loss of control of human activities under pressure to optimize effectiveness and the Brownian movements of particles in a field subject to strong gradients.

Conclusion

It follows from this discussion that the modeling approaches taken by several disciplines are relevant for the present purpose and should be reviewed. The discussion, however, also raises the epistemological question of the basic nature of the models required. To what extent can we rely on decomposition of a system into elements that can be studied separately within the various disciplines and when are systemic models derived by functional abstraction and isolation necessary to study the behavior-in-the-loop? In addition, when can we develop quantitative, mathematical relationships for representation of system behavior and when are qualitative models in cause-and-effect terms useful? The following sections set the stage for such distinctions.

QUALITATIVE VERSUS QUALITATIVE MODELS

A distinction is often made between pre-scientific, qualitative models and proper scientific, quantitative models. In the conventional view, qualitative models are premature representations which should ultimately be replaced by mathematical representations. This view was advocated by Warren Weaver (1948) who introduced the distinction between *problems of simplicity* which are approached by physical sciences; *problems of disorganized complexity*, the domain of statistical mechanics and thermodynamics and, finally, *problems of organized complexity* which include

> "a wide range of problems in the biological, medical, psychological, economic and political sciences" which "cannot be handled with the statistical techniques so effective in describing average behavior in problems of disorganized complexity."

He concludes that

> "science must, over the next 50 years, learn to deal with these problems of organized complexity"

and he finds two bits of evidence promising a achievement of this target: the emergence during the second world war of the cross-disciplinary approach to *Operations Analysis* and the invention of the *electronic computer*.

Models are, however, tools for a purpose and different kinds of models serve different ends. Interestingly enough, qualitative models in terms of verbal descriptions, imagery or pencil sketches on the back of an envelope seem to be the representation serving inventors and designers during the creative phase of their activity. Once the design has been established and the basic functions verified by experiments, quantitative scientific representations may be used to optimize the design and to find the limits of possible refinements, this is the basis for engineering analysis based on mathematical representations. The steam engine was invented by manipulating qualitative representations, but optimized by scientific mathematical representations based on thermo-dynamic theories. In fact, the evolution of thermo-dynamics was paced by the need of steam engineers. The 'flying-balls' governor of Watt's steam engine - in modern terms: its automatic speed controller - was conceived by Watt from qualitative reasoning, but the efforts to understand the occasional instability of its control of engines forced Maxwell (1868) into a mathematical representation which led him to the differential equation representation used throughout modern control theory.

The problem when applying qualitative reasoning to a steam engine with a governor is that such reasoning is based on a decomposition of the machine into its parts and its behavior into a linear cause-and-effect chain of events, that is, it depends on decomposition and reduction. In contrast, the functioning of the machine with its governor depends on their interaction in a closed loop, any disturbance of any part of the machine will immediately affect the state of all other parts in the loop. In consequence, the function of the loop cannot be studied by decomposition and reduction, only by abstraction and separation of the characteristics *of the loop function.* That is exactly what Maxwell did by is mathematical model.

A problem similar to Maxwell's problem faced with the instability of the steam engine governor is found in the analysis of many modern technical and social systems. Efforts to control the safety of large-scale industrial systems appear to be less effective due to adaptation by the socio-technical system under competitive pressure in a turbulent environment, a stability problem similar to the governor problem (Rasmussen, 1988, 1990). The stability of large-scale financial systems also appears to be a problem (the Wall Street collapse, Waldrop, 1987, and the EEC currency turbulence, 1992) which is difficult to handle by the usual if-then arguments of present causal models.

Functional Abstraction versus Structural Decomposition

The interaction between the two kinds of model is particularly important when studying and designing human-machine systems. When we conceive of a new work system, we normally use the structural perspective. To serve a particular purpose, we chose among the available parts, tools, and productive processes and we select staff members having an appropriate background and education. We then aggregate these elements into a productive structure and instruct the actors how to apply the tools and productive processes for the purpose of work. In other words, we arrange the elements in cause-and-effect chains according to their individual input-output characteristics so as to have the intended overall effect. The problem we face in modeling systems incorporating human actors is, however, that humans do not have stable input-output characteristics which can be studied in isolation and we cannot develop models of the *actual functioning* of human machine systems by aggregating input-output models developed in isolation. When the system is put to work, the human elements change their characteristics; they adapt to the functional characteristics of the working system, and they modify system characteristics to serve their particular needs and preferences. In other words, to understand system behavior *when adaptation has taken* *place*, we have to look at the entire system and instead of decomposing functions according to the structural elements, we have to abstract from these elements and, at a purely functional level, to identify and to separate the relevant functional relations, see figure 2.

Basically, we have to identify the adaptive mechanisms that generate the observed behavioral trajectories and work practices. In other words, design by aggregation of input-output relations identified for the individual structural elements in isolation makes it *possible* for the resulting work system to function. However, to make the functioning *effective*, the design should not result in a system that constrains the behavior of human actors to only one possible work process. In stead, the design should define for the actors a space bounded by the goal-and resource- constraints, Within this space they should be allowed to adapt freely according to subjective criteria, such as e.g., effort, time spent, or joy of discovery. Analysis of existing work systems to understand their behavior and to get a basis for design of new systems therefore should not be focused on decomposition into their structural elements, but on a functional abstraction and separation of functional patterns and the criteria that *generate* behavior.

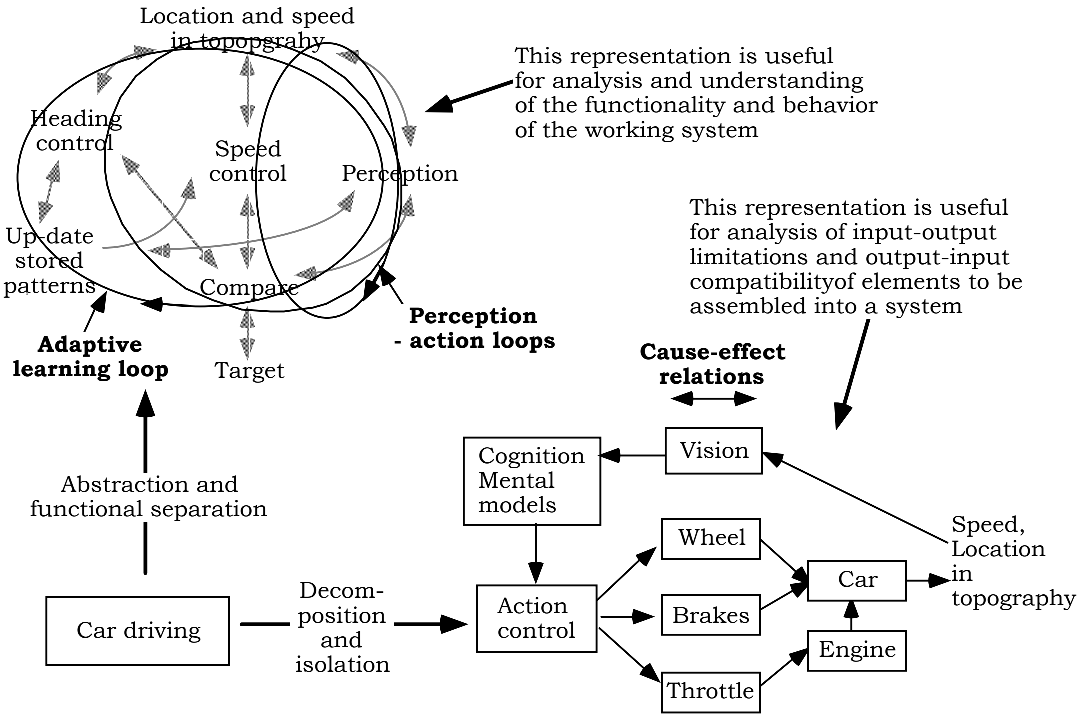


Figure 2. The behavior of a system can be analyzed in two ways. One is by decomposition into parts the behavior of which is known in isolation. The behavior of the system is described in terms of cause-and-effect relations. Another is by abstraction to a functional level and separation of the loops of interaction of interest.

A familiar example illustrates this point. When a novice is driving a car, it is based on an instruction identifying the controls of the car and explaining the use of instrument readings, that is, when to shift gears, what distance to maintain to the car ahead depending on the speed, and how to use the steering wheel. In this way, the function of the car is controlled by discrete rules related to separate observations and navigation depend on continuous observation of the heading error and correction by steering wheel movements. This aggregation of car characteristics and instructed input-output behavior makes it *possible* to drive; it initiates the novice by synchronizing him/her to the car functions. However, when driving skill evolves, the picture changes radically. Behavior changes from a sequence of separate acts to a complex, continuous behavioral pattern. Variables are no longer observed individually; complex patterns of movements are synchronized with situational patterns; and navigation depends on the perception of a field of safe driving. The driver is perceiving the environment in terms of his driving goals. At this stage, the behavior of the system cannot be decomposed according to the structural elements. A description must be based on abstraction into functional relationships.

As a background for the modeling review, a brief discussion of the basic properties of quantitative and qualitative models may be useful.

Quantitative, Relational Representations

For problems such as the stability of adaptive systems, it appears that we should look out for quantitative, relational models. Relational representations depend on *abstraction and isolation* of selected relational structures which connect quantitative variables. They represent 'practically isolated relationships' (Russell, 1913) which are valid for a variety of systems. The internal consistency can be proved mathematically, their validity in the world can be tested experimentally in a variety of experimental configurations with controlled conditions. This type of model does not necessarily represent the actual, in-the-world behavior of the phenomena of interest, but is effective for understanding basic mechanisms and *to define limits of performance and conditions for optimal function.*

In relational models, the objects of the actual system are only present in terms of a set of parameters distributed across equations. It is therefore relatively difficult to modify the model in response to changes in the physical world. That is, the propagation of the effect of changes, errors, and faults through the system is rather obscure within this type of representation.

Qualitative Cause and Effect Representations

Qualitative models are based on representations in terms of objects which interact through causal relations among events. Events represent changes of states and configurations of objects. Whereas relational models are based on *abstraction and isolation* of relational patterns, causal models are based on *decomposition* of the system into objects and *reduction* of their behavior into typical events (Rasmussen, 1988, 1990a).

The objects and events of causal models are prototypes, that is, they are typical examples representing classes identified by verbal labels. The prototypical nature of the representational elements make them very effective for representation of new configurations because the particular example of a class entering a thought process will be shaped by the context into which it is implanted.

Accepting the view, that design involves matching an object which is not yet in existence to a context which cannot be completely described (Alexander, 1964) implies that qualitative models and causal representations are necessary for that creative activity. Being expressed in terms of types, classes, the elements of a particular model and the patterns of the context are adaptive and able to shape themselves according to the changing context of the mental experiments of the designer.

This feature very likely is the reason why early industrial inventors and designers were actually from professions well versed in visual representation and manipulation of the physical world (map makers, painters, etc.) because they were experts in shaping new visual complexes from modified samples of a visual alphabet (Hindle, 1981).

On the other hand, even if qualitative models are very effective for the individual inventor and designer of new machinery, experimental equipment, and other artifacts, they have some serious drawbacks when applied for scientific analysis for which the are used extensively when the object of analysis is too complex for exhaustive definition, such as analysis of accidental chains of events in complex technical systems (industrial complexes) as well as the behavior of humans and social systems. Two problems are particularly relevant in the present discussion: 1) Its prototypical nature, that is, the validity of qualitative, causal representations depends on the acceptance of the proper context, and 2) its linear structure, that is, causal reasoning in causal nets including closed loops is ineffective (cf. Watt's governor).

The qualitative models are very effective for representation of changes in the world, due to the analog, one-to-one mapping of objects and events and to their prototypical nature which allow reinterpretation of the objects and events of the model when the context changes.

In the following sections, a review is presented of the approaches to modeling complex systems as they presently are found within different disciplines. The selection of approaches to consider in a review is biased by the objective under which the review is made and in the present case, focus is on models that may serve analysis of adaptive, self-organizing systems.

QUANTITATIVE, RELATIONAL MODELS

To recapture, this kind of models represents the behavior of systems in terms of relationships between quantitative variables that are identified by isolation and abstraction, not by decomposition. Consequently, the result is a *behavioristic representation of the input-output relation* that omits consideration of the structure and internal organization of the object of study. In this way, models generalize across systems and types of behavior found in different kinds of systems can be represented by one model (Newton's model of gravitation is valid for heavenly bodies as well as for falling apples).

It follows from the discussion in the introduction, that an evident approach to modeling the behavior of purposive organizations will be one based on *control theory.* We will, therefore, begin the review with some well-known, control theoretic models.

CYBERNETICS: THE CONTROL THEORETIC PERSPECTIVE

Wiener coined the term *cybernetics* in his classic book "Cybernetics or Control and Communication in the Animal and the Machine" from 1948, but an earlier and influential statement of the quantitative, control-theory approach to modeling the behavior of living organisms is found in Rosenbluth, Wiener and Bigelow (1943). They applied a control theory perspective on purposive behavior and in their classification of behavior they equate purposeful behavior with simple feedback:

> "All purposeful behavior may be considered to require negative feed-back. If a goal is to be attained, some signals from the goal are necessary at some time to direct behavior. By non-feed-back is meant that in which there are no signals from the goal which modify the activity of the object *in the course of behavior."* (Emphasis in original).

The authors define teleology as synonymous with "purpose controlled by feedback" and find that:

> "Teleology has been interpreted in the past to imply purpose and the vague concept of a "final cause" has often been added. This concept of final causes has led to the opposition of teleology to determinism. ---However, purposefulness, as defined here, is quite independent of causality, initial or final."

For a brief discussion of Aristotle's actual conception of 'final cause', see the discussion in a later section.

This simplistic definition of purposeful behavior, does not consider its evolution by selection of behavior found successful in the past, and has biased much of the later 'general system theory' work. (The arguments of Rosenbluth et al. were later subject to severe criticism by Taylor (1950) who found their arguments (that only active feedback systems were purposeful) to be totally metaphorical). Furthermore, even if the future selection of behavioral patterns found to be successful in the past can be considered feedback at longer time scale, it is not feedback "in the course of behavior."

Mesarovic's Models of Complex Systems

Rosenbluth, Wiener and Bigelow were concerned with the basic application of the control theoretic concepts to purposive systems. Later, the control theoretic representation was applied to complex, hierarchical systems by Mesarovic and his group from Case Western Reserve University in a number of papers and books through late the 60's and early 70's (see Mesarovic, 1970 a,b). This approach is particularly interesting in the present context because the qualitative part of Mesarovic's framework representing an adaptive control system by a hierarchy of functions match the concepts of our framework (see Rasmussen, 1991).

In this theoretical framework special emphasis is on the formulation of the hierarchical concepts from a variety of approaches "which are termed hierarchical with more or less justification." The aim of this conceptual framework was to structure the representation of complex systems in a way suitable for a quantitative, mathematical description. This interest in basic structural system problems of complex systems later decayed considerably in the control theoretic literature.

The essential characteristics of hierarchies are taken to be: vertical decomposition, priority of action or right of intervention, and (vertical) performance dependence. These distinctions are closely related to the classical notion of a command hierarchy:

> "Any hierarchy involves a vertical arrangement of subsystems. .... The operation at any level is influenced directly and explicitly from the higher levels. .... The influence is binding for the lower levels, reflecting a priority in importance of actions and goals of the higher levels..." (Mesarovic, 1970a).

Different notions of hierarchies of importance for process control are discussed which are similar in nature to the distinctions made in our taxonomy. Distinctions are made between: 1. Levels of description or abstraction, 2. Levels of decision complexity and 3. Organizational levels. In order to keep the different hierarchical descriptions separate, Mesarovic introduces different terms for the levels: *strata, layers, and echelons*, respectively. This terminology is, however, not maintained consistently by the various authors of the period.

1. Strata: Levels of Description

According to Mesarovic (1970 a),

> "an effective method to describe a complex system is in terms of a hierarchy of models referred to as a stratified description of the system. Each stratum refers to a different aspect of system operation and for a complete description one has to consider the totality of all strata."

This stratification has many aspects in common with the means-ends levels discussed in our taxonomy (Rasmussen, Pejtersen, and Schmidt, 1991). One major difference is, however, that Mesarovic does not define explicitly the dimension along which 'abstraction' takes place and in the examples given, abstraction in terms of part-whole relations and in terms of the functional abstraction, that is the language used for description, is not kept separate:

> "Several features of a stratified description are particularly important in the process control area; they are: 1. For each strata there usually exists a different set of rules, principles, and laws in terms of which the system operation is described; this contributes to reducing the coupling between strata. 2. On the lower strata the description is more detailed than on the higher strata."

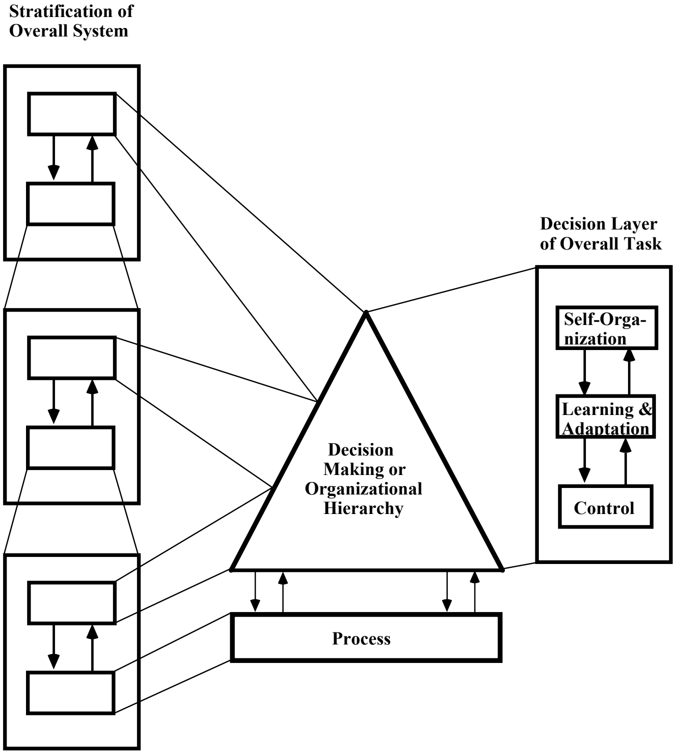


Figure 3. In a system with an organizational hierarchy, the tasks of the decision units on different levels can be specified in terms of a stratified description of the system as well as with reference to decision layers of the overall problem (reproduced from Mesarovic, 1970).

In general, there is an emphasis of the part-whole dimension:

> "-- on any strata one concentrates on the functioning of a subsystem while the effects of interactions between subsystems are considered on the higher strata."

This point of view reflects the requirements during control system design for a representation structured according to the basic physical functioning. Note, that in the cognitive framework proposed in our taxonomy, the part-whole decomposition is considered to be orthogonal to the abstraction (means-ends) dimension.

2. Layers: Levels of Decision Complexity

For advanced control systems, in particular computer-based systems, the objective is to

> "maximize profit by taking into account both economic and technological factors and in reference to both short term and long term effects in a changeable environment. .... In practice, there is no solution method which will accomplish such a global objective because of both complexity and lack of knowledge and information.

> A method of approaching such a solution, however, can be designed based on constructing a hierarchy of sub-goals defined in practical operational terms and which then will enable synthesis of the corresponding decision-making units arranged in a hierarchical fashion and charged with pursuing the respective sub-goals. .... In other words, the solution of the overall problem results from the solution of the hierarchically arranged sub-problems which are simpler and for which solution methods and algorithms already exists."

The levels in this type of hierarchy are called layers.

As being of special interest to process control is mentioned the

> "three fundamental aspects of the general control (decision) problem under conditions of true uncertainties",

and the corresponding three layers of control: 1) The lowest layer, the selection layer determining the control to be actually applied to the physical process and which actually can be divided into two layers, the regulator layer (including feedback control) and the optimizing layer. 2) The next layer of adaptation and learning (including statistical and logical techniques), and 3) the highest layer of self-organizing control (e.g., based on heuristics). Note, that these three levels are similar in nature to the skill-, rule-, knowledge based levels of cognitive control.

3. Echelons: Levels of the Organizational Hierarchy

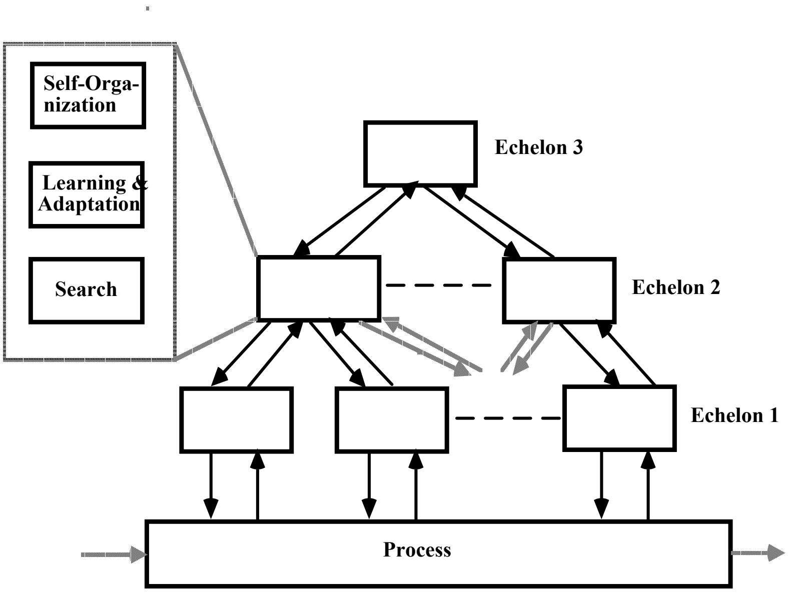


Figure 4. The figure illustrates how the layered organization of control functions can be found embedded in an organizational echelon. Furthermore, any decision unit in an organizational hierarchy can use a stratified description for the model of the system under its control (Reproduced from Mesarovic, 1970).

> "For a variety of practical and conceptual reasons ... a vertical decomposition is accompanied by a horizontal decomposition"

forming an organizational hierarchy:

> "On each level a decision unit (controller) is concerned with a larger portion of the system and its primary task is to co-ordinate the actions of the sub-ordinate units."

The difference between this organization in echelons and that in decision layers is not defined explicitly. Organization in echelons, however, appears to be related to decomposition of a complex task for allocation to a number of similar functional units, whereas the layered decision structure is related to basically different decision functions.

4. Relationships between the Notions of Strata, Layers, and Eschelons

In practice, a multilevel system may be described in reference to all three notions.

> "For example, in a system with organizational hierarchy the tasks of the decision units on different levels can be specified in terms of a stratified description of the system or in reference to decision layers of the overall problem. "

Another example (figure 4) illustrates how the layered organization of control functions can be found embedded in an organizational echelon. Furthermore, any decision unit in an organizational hierarchy can use a stratified description for the model of the system under its control. In other words, each of the notions are related to a particular aspect of control system design. The different notions are not conceptually related in a consistent framework and their use for a theoretical discussion of the properties of different types of complex systems and their organization is not feasible. This is probably the reason why the theoretical analysis proposed by Mesarovic is focused only on a mathematical algorithms for co-ordination of a number of similar units in organizational echelons (e.g., for control of electric power dispatch in a pool consisting of a number of similar power plants).

5. Co-ordination Strategies.

In control theory, distinction is frequently made between several different strategies for co-ordination of hierarchical systems, depending on the decomposition approach (Schoeffler, 1969). Co-ordination can be obtained through:

- 1. manipulation of *goals* of lower levels or parameters of these goals;
- 2. manipulation of *constraints* of lower levels, i.e., restriction of the domain of their control actions;
- 3. manipulation of *interaction variables* between units from estimation or prediction, or
- 4. manipulations of the *information* available to the units at lower levels, etc.

It appears to be possible to apply similar distinctions for a formalization of the co-ordination of the social organization of decision makers representing the evolution of the various role allocation criteria as illustrated by the categories of our taxonomy.

Conclusion

In conclusion, Mesarovic's theory of hierarchical control structures is basically a qualitative categorization and description of the elements and the structure of such systems. It includes adaptive systems but not features of self organization.

Mesarovic applies the structure as a basis for a quantitative, mathematical theory of the performance of hierarchical control systems, but only considers systems including multiple, but similar elements, such as electrical supply grids with many power plants, etc.. The quantitative part of Mesarovic's framework is, therefore, not relevant for the present problem because of the great variety of components found in socio-technical systems.

Mesarovic's General Systems Theory

Mesarovic has, however, explicitly discussed the generalization of control theory to include biological phenomena (Mesarovic, 1968). He argues in favor of

> 'the adequacy of systems theory to provide both a conceptual basis and a working method for the scientific explanation of biological phenomena.'

In this context he focuses on different basic aspects of systems theory, such as

> 'the problems of foundation (general systems theory), teleological explanations (goal seeking behavior), and hierarchy or multilevelness.'

He defines 'systems theory' as being a formal, mathematical theory of real life or conceptual systems, using the concepts of information and signal processing as well as decision making and control. The properties of formal models can be investigated either by mathematical deduction or computer simulation. Two aspects of study are important: Informal, dealing with meaning, interpretations, significance, objectives, values, etc., and formal, dealing with the form, that is, structure, in which the relationship between attributes appears. The latter is the topic of systems theory. Note that Mesarovic, by this definition, excludes the formulation of objectives and values from the systems theoretic modeling and, furthermore, excludes the dynamic evolution of the form and structure of the functionality of an organism or organization.

System theory is, according to Mesarovic based on two premises: a). A theory of any real life phenomena (biological or otherwise) is always based on an image, a model and b). without introducing any constraints whatsoever the formal, invariant aspects of that model can be represented as a mathematical relation, which is termed 'a system.' Once more, note that systems theory is concerned with 'formal, invariant aspects of the model,' that is, the behavior of stable structures.

To generate a 'system,' that is, *a mathematical relation* which is model of a selected phenomenon the following steps are taken:

a) Select the attributes to be observed,

b) specify (by experiments or assumption) what values the attributes could take,

c) indicate the combinations of the values for the attributes which are observed,

d) collect all these combinations from all experiments conducted to represent the phenomenon as a system.

The specification found is usually in a form that some attributes (outputs, 'effects') can be determined when some other attributes (inputs, 'causes') are given.

> "This is the so-called input-output (terminal) specification of the systems which models the stimulus-response behavior of biological systems."

Mesarovic finds it necessary to place his

> '"brief discussion of systems theory into proper perspective (especially regarding its relation to biology)"

by some additional considerations:

> "1) Numerical algorithms are traditionally used almost exclusively for specification of systems.
>
> It is to be expected, however, that non-numerical systems will become increasingly important in the near future for the following reasons: a) Complexity: With the concern for the increasingly complex biological systems the mathematical models used for the systems specification will have to be restricted only to the important features of the biological phenomena while many specific details will have to be ignored to arrive at a manageable description. Algebraic and topological mathematical structures will, of course, be used for that purpose. b) Individual-variations: While the basic form of a relationship between the attributes might be invariant for any member of a given species, the actual values of the parameters in that relationship depend on the individual characteristics of each particular member. When this effect is pronounced, the non-numerical structures are particularly useful in describing the basic cause-effect relationships which are invariant over the entire species.
>
> 2) A given system can have many alternative constructive specifications. [Note: remember M's definition of system as being the mathematical representation] Some of them might reveal some important aspects of the 'true biological mechanisms' responsible for the observed behavior. More often, however, the constructive specifications should be viewed primarily as means for effective description of experimental findings, and special care should be taken not to be lured into unwarranted generalizations on the basis of a selected constructive specification.
>
> 3) Constructive specification can also be given by using goal-seeking concepts."

This approach is discussed in more detail in the following paragraphs.

Mesarovic remarks that the controversy over the role and acceptability of teleological explanations in biology is very old and yet still outstanding and he finds that system theory has a significant contribution to the controversy. He reminds us that systems theory is basically a method to provide a constructive theory for the system (the formal model) so that whenever an input (and a state) is given, the output can be obtained using the procedure. This formulation calls to mind Russell's notion on deterministic representations in contrast to causal representations. We can then accept Mesarovic's premise, that

> "for an important class of situations, one can develop an effective constructive specification of the system (formal model) only if one is using a goal-seeking (i.e., teleological) description; furthermore, this can be done so that by using such a description the basic character of the system as a (mathematical) relation has not been changed."

He emphasizes that

> "Goal-seeking here implies a whole gamut of systems descriptions which are given in terms of concepts expressing the purpose of the systems behavior such as adaptation, evolution, control, homeostasis, etc."

He is referring to the 'system', that is, in his terminology, the formal mathematical representation in which, following Russell, the arrow of causality has disappeared. Mesarovic's system theory is aimed at a description of the input-output relationship, not at an explanation of the underlying mechanisms generating the relationship. For an effective feedback loop, being it a functional or an evolutionary loop, a representation of the function in terms of the feedback relationship is the most economic; compared to a representation in terms of the often very vacillating functional mechanism of the forward path. As an illustrative example he uses Lettvin et al.'s (1959) study of the vision of frogs. In this study, they found that the most efficient representation of the physiological data was in terms of the goal in terms of survival value, that is more specifically, the search for food.

Mesarovic then remarks that

> "the same line of reasoning might be used to explain the function of the brain rather than studying the neural network functions. Even if the individual neuron can be modeled rather accurately, the sheer complexity of their interactions prevents a reliable modeling of brain functions."

In other words, he uses the approach suggested also by Anderson 20 years later (see the discussion below). The problem from our point of view is, that a model of the functions of an organism adapting dynamically to the requirements of a changing environment should also account for the adaptive process itself, not only for the input-output relationship of the adapted system.

Mesarovic concludes his discussion by a review of the reasons why systems theory has not quite lived up to the expectations at that point in time. He notes, from the systems science point of view, the following reasons:

> "A) One important reason is that system science has not been directly concerned with some of the problems of vital importance to biology: 1) How to deal with the interactions among subsystems at the same level, and how to account for the interactions when subsystems are studied in isolation? 2) The interaction among levels in hierarchical, goal-seeking systems. Biologists traditionally tend to ignore the effects of sub-systems from both higher and lower levels. 3) In representing biological systems as goal-seeking the key problem is, how to identify the internal references or goals from the external input-output data. This problem has not at all been approached on a formal basis. These three problems in biology has not received proper attention by systems theoreticians.
>
> B) from the point of view of a biologists: The reasons for the gap between systems theory and the needs of biology is caused by the fact that systems theory has primarily been dealing with problems of interest to other disciplines, such as various kinds of engineering."

In conclusion, in Mesarovic's presentation of general system theory, the aim is to establish an economic description of the input-output transformation of a system adapting its behavior to an explicitly stated goal or purpose as it is found represented in a set of observed data. That is, the use of goal-seeking as a modeling approach is argued from the economy-of-modeling perspective.

Forrester's World Model

Jay W. Forrester argues along similar lines when developing the system simulation model DYNAMO (Forrester, 1968) that later was used for the Club of Rome world simulation model (Forrester, 1971; Meadows et al. 1972). Forrester argues:

> "Gradually over the last hundred years it has become clear that the barrier to understanding systems has been, not the absence of important general concepts, but only the difficulty in identifying and expressing the body of universal principles that explain the success and failures of the systems of which we are a part. Economics has identified many basic relationships within our industrial system, Psychology and religion have described some of the interactions between systems of people, Medicine has treated biological systems. Political science has explored governmental and international systems. But most such analysis has been verbal and qualitative."

He then argues:

> "But now the concepts of "feedback" systems seem to be emerging as the long-sought basis for structuring our observations of social systems." ... "Around the system principles discussed in this book it should be possible to structure our confusing observations about political and business systems. When a structure and governing principles for systems have been accepted, they should go far to explain the contradictions, clarify ambiguities, and resolve controversies in the social sciences"

Basically, Forrester's approach is based on a set of first order differential equations representing conservation models of supply/flow-rate/inventory relationships together with function generators used to represent non-linear empirical relationships among measured variables. Models of elementary relationships, as studied within economics, sociology, and psychology, are fitted to empirical data and then connected into complex system models. Parameters are then adjusted to give system responses matching observed data. The strength of the approach is that complex non-linear models can be developed and used for experiments for identification of sensitive parameters and for development of an understanding (intuition) regarding the behavior of complex closed loops. The problem with the approach when used for extrapolation (as e.g., the Club of Rome did) is that it does not include the basic adaptive mechanisms, it is, as Mesarovic emphasized, only an economic way of representing the input-output relationships of a closed loop feedback system. Forecasting only depends on projecting the future behavior of the mechanisms behind data from the past, sometimes supplemented by asymptotic limiting mechanisms.

Basically, the system model is developed by aggregating process models as found outside the closed-loop interaction. The adaptive reshaping of processes and structures when operating in the closed loop is not included. That is, reorganization according to the changing value structures of the individuals and organizations are not included, as Mesarovic also noted.

Mesarovic later applied the hierarchical system concepts for an extension of the Forrester-Medows simulations (Mesarovic and Pestel, 1975). The main difference being that Forrester-Medows simulated world dynamics as being one uniform, global system and argued that this system will collapse some time in the middle of the next century if the present trends continue and that an immediate slow-down of economic growth must be initiated to reach equilibrium in a relatively short period of time. Mesarovic-Pestel hold that such a uniform model is misleading. Instead a regional model is required and their simulations point to catastrophes and collapses at the regional level could occur much earlier, in different regions, for different reasons, and at different times.

In this way, the use of simulation models has been extremely important in the efforts to make politicians and the general public understand the nature of exponential growth and to make them aware of the long term ecological implications. They are also very effective for demonstration of the influence of different action scenarios on the projections of the trends found in data from the past.

Ashby's Design for a Brain

Ashby's book (1960) is an early and well-known attempt to use control theory to model human behavior, including higher level structural adaptation. It is, however, strongly influenced by the behavioristic tradition of psychology and, consequently, is focused on stimulus-response or input-output representation of reflexes and responses learned through adaptation. The brain is treated as an organ developed through evolution as a specialized means to survival. He explicitly states that neither psychological concepts nor teleological explanations are used:

> "never will the explanation be used that an action is performed because it later will be advantageous to the animal'. -- 'Consciousness and its related subjective elements are not used for the simple reason that at no point [he] has found them necessary."

His model is based on mathematical relations between variables, represented as differential state space equations and a system is defined

> "as any set of variables that he [the observer] selects from those variables on the real machine."

The modeling activity is based on experiments by which

> "the experimenter can control any variable he pleases: that he can make any variable take any arbitrary value at any arbitrary time."

He forces the surrounding conditions of the system and

> "observes to what state the system goes as it moves under the drive of its own dynamic nature."

A line of behavior is then

> "specified as a succession of states and the time intervals between them."

The behavior of a system can be represented in a 'phase-space' and 'a system's field is the phase-space containing all the lines of behavior found by releasing the system from all possible initial states in a particular set of surrounding conditions.

Ashby introduces the concept of *'requisite variety'* which has been very influential, see e.g., Beer's approach discussed below. In essence, the concept states that to be able to control a system, the controller should have available a variety of control alternatives which matches the variety of states of the system environments.

Ashby finds it necessary to stress that an organism's interaction with its environment basically is a feedback system because

> "most physiological experiments are deliberately arranged to avoid feedback."

He generalizes boldly from his feedback concepts:

> "Thus, machines with feedback are not subject to the oft-repeated dictum that machines must act blindly and cannot correct their errors.' -'Once it has been appreciated that feedback can be used to correct any deviation we like, it is easy to understand that there is no limit to the complexity of goal-seeking which may occur in machines quite devoid of any 'vital' factor."

This statement is very much in line with the position of Rosenbluth, Wiener and Bigelow and one may wonder whether not the 'trial and error' adaptation of Ashby's ultra-stable systems represents behavior of 'machines acting blindly.' In addition, introduction of 'subjective' internal process criteria to account for resource-demand matching depends on internal process alternatives. Such alternatives can be formulated without 'introducing psychological concepts' and the system-related criteria do not represent 'vital factors.'

This discussion points to the modeling problem of keeping concepts related to the functions of living organisms and those used for representation of such functions quite separate. 'Consciousness' is a human function, a psychological concept, computational models do not have 'consciousness' but may very well have processes representing selected functions of consciousness, such as evaluation of own performance according to internal process criteria.

Ashby defines adaptation as a form of behavior that maintains the essential variables within physiological limits. A stable system is a system which return to an equilibrium within the state space field after disturbances. An ultra-stable system is a system that, when no stable equilibrium exists under the influence of the environment, switches to another state-space configuration having a field containing an equilibrium. This step-function switch is activated by a variable in the environment that has only a few, discrete values. That is, in our terminology, a sign reconfiguring the control law governing the lower level signal-controlled movement patterns. In fact, Ashby's framework includes the skill-and rule-based levels of our model, but only in terms of a behavioristic input-output representation and adaptation is only represented in terms of 'trial and error' learning.

Stafford Beer's Viable System Model

Ashby's approach and, in particular, his concept of 'requisite variety' have been the basis for Stafford Beer's extensive writing on 'management and control' (Beer, 1959, 66, 79, 81) which is concluded in his 'viable system model' (Beer, 1984):

> "it has always seemed to me that Ashby's law [on requisite variety] stand to management science as Newton's laws stand to physics; it is central to a coherent account of complexity control: ' Only variety can destroy variety.' " (Beer, 1989, p. 18).

Beer (1989) describes the basis of his model in this way:

> "The model of any viable system, VSM, was devised from the beginning (the early fifties) in terms of a set of interlocking Ashbean homeostats. An industrial operation, for example, would be depicted as homeostatically balanced with its own management on one side, and with its market on the other. But both of these loops would be subject to the Law of Requisite Variety. Since the variety generated by the market would obviously be greater than the industrial operation could contain, then 'this part must be blocked at all costs' as Ashby says." (p. 18).

Beer developed several versions of his theory through the years:

> "So the viable systems model (VSM) dates back 30 years. I pursued it through neurocybernetics and social science, through the invention and study of cybernetic machines, through the mathematics of sets and stochastic processes, and all the times through the OR fieldwork in industry and government. " ... "The set theoretic model proved difficult for people to understand, and eventually a streamlined version of the model appeared called *Brain of the Firm,* using neurophysiological terminology instead of mathematics." (p.12). ... "Moreover, I developed a topological version of the original set-theoretic algebra that it seemed no-one would study properly. The drawings were now rigorous mathematics in themselves, in that they offered explicit homomorphic mappings of any one recursion on to the next - as may be seen in the simplified version at Figure [5]." (p. 13).

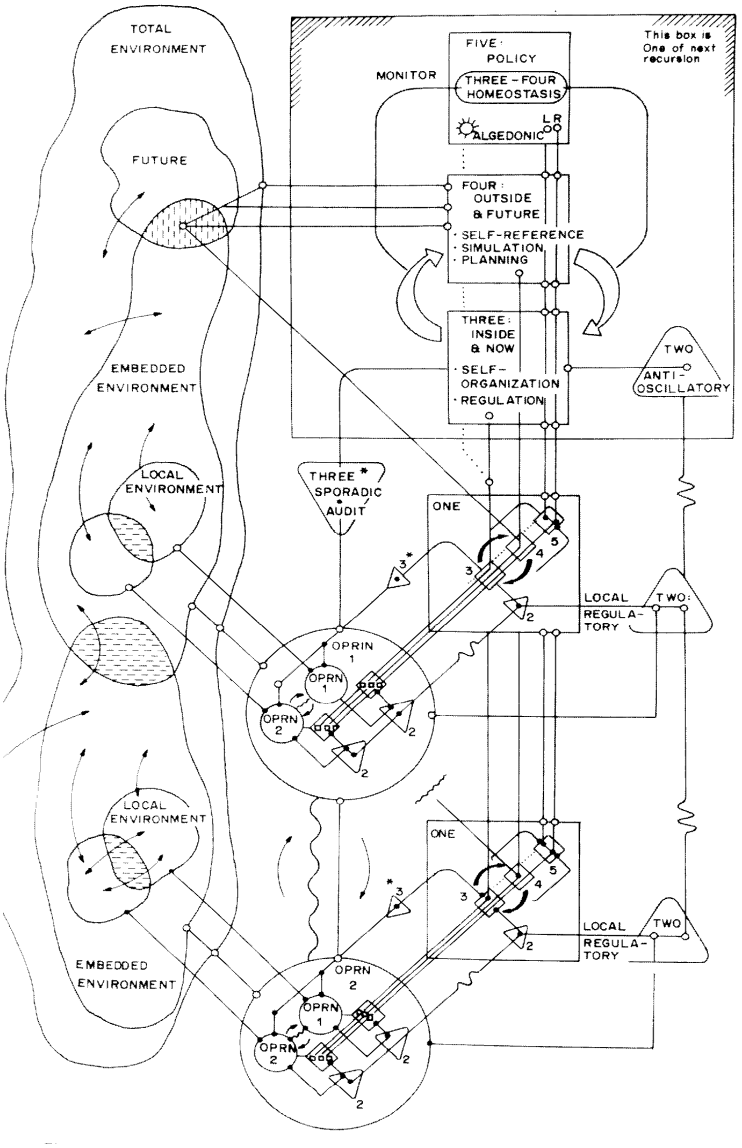


Figure 5. Beer's Viable Systems Model, reproduced from Beer, 1989

As the figure indicates, Beer's VSM concept five subsystems, 1) system ONE, representing the elements of the system, a level of description representing decomposition into parts. This representation should be used recursively to represent individuals, organizations, and societies. each system ONE includes a system TWO, a local regulatory function, a system THREE taking care of self-organization and regulation "inside & now;" a system FOUR taking care of self-reference, planning and simulation "outside & future" and, finally, a system FIVE coping with policy. That is, his concepts operates with orthogonal decomposition and functional abstraction in a way similar to our framework.

Beer's presentation of his theories is often very impressionistic and the formal versions have had little influence. The greatest impact has been through his extensive 'intervention' research. Through the discussions of Beer and his students with managers and decision makers of companies and institutions during field work, the framework has served well to give managers a fundamental understanding of the complexity and the closed-loop interactions within their system and thus better to cope with the variety of their dynamic environment, that is, they learn to study their work functions in the loops, in contrast to the simulations of Forrester. During the interaction, decision makers will be able to re-interpret and re-shape Beer's formal and stereotypical model structures within their own, actual context. As Beer says:

> "At any rate, anyone who has taken a really good look at Figure [5] and its infrastructure is surely in a position to understand why purely hierarchical models of management are useful for little more than appointing blame, and why the familiar debates about centralization and decentralization (based as they usually are on ideology) are powerless to resolve vital questions of autonomy." (p. 24).

A recent review of the application of Beer's theories for active intervention in management development is found in Espejo and Harnden (1989).

Conclusion

Models based on control theoretic representations of closed loop performance have been very important in Mesarovic's and Forrester's formulation for the understanding of the behavior of complex socio-technical systems. In addition, their simulation models have been very effective in representing the non-linear relationships observed in the past behavior of complex systems, for projection of present trends, and for studying the influence of different decision scenarios.

In Beer's interpretation they have, in particular, served to communicate to decision makers an understanding of the complexity and closed-loop nature of management problems. The quantitative, control theoretic models developed within 'general system theory' and 'cybernetics' have not, however, included the adaptive mechanisms and changing value systems that govern the evolution of system structures and the resulting behavior, including their breakdown when effective adaptation fails.

Manual and Optimal Control Models

Thus, quantitative, control theoretic models in terms of 'general systems theory" have not been particularly successful for prediction of the behavior of complex, sociotechnical systems. For modeling human control of more well defined and predictable, technical systems, however, the picture is different. Models of human performance in closed-loop control tasks have been very important for representation of the properties and limitations of humans in vehicle control, in particular in aviation, and a separate school of modeling based on control theoretic tools has evolved.

Manual Control Models.

A review of models of human sensori-motor performance has been given by Pew (1974). Such models have been developed, in particular, from laboratory tracking tasks, and will typically give information on signal-to-noise ratio, maximum bandwidth when tracking unpredictable wave forms, prediction capabilities in sine-wave tracking, etc.

The role of predictive feed-forward control for an industrial control task has been demonstrated experimentally by Crossman and Cooke (1962). Models of manual control of dynamic systems based on differential state-space equations have been reviewed by Sheridan and Ferrell (1974). This approach have led to the 'optimal control models,' which are based on the observation of Leonard (1960) and Roig (1962) that the mean-square error from human tracking data approximated the mean-square error of various optimal controllers.

Vehicle Control Models

Optimal control models have in particular been developed by Baron and Kleinman (1969) and used for describing pilot performance (Kleinman et al., 1971). This model is based on the assumption that a well-trained, well-motivated human operator will act in a near-optimal way subject to certain internal constraints that limit behavior. The internal dynamic world model necessary to account for human anticipation is represented by a Kalman-Bucy optimal filter. The model also includes observation noise and time delays depending on the instrument scanning strategy. The criterion used is typically minimum square deviations from desired output as well as squared control effort according to a chosen trade-off ratio. Thus, the models include parametric adaptation to changing environmental conditions.

Strategic Decision Making.

The model has been developed beyond the simple man-in-a-loop case, in that higher-level sequential functions for parameter and criterion control have been added in order to include multi-variable control, monitoring, and decision making and, in this way, some self-organizing features are included in a predictive model. This effort has proved successful for flight control and landing approach planning (Muralidharan and Baron, 1980; Baron et al., 1981). Efforts have also been made to extend this model to process control, that is, to control of systems with a less bounded behavioral space. Vehicle dynamics is a well defined phenomenon, which is not the case for process dynamics when also disturbance control is to be included.

Baron et al., (1982), have proposed a quantitative model based on control theory for simulation of the dynamic performance of a nuclear power plant including the operating staff. One important aspect of this approach is that human behavior at all three levels (skill-, rule-, and knowledge-based) as well as their interactions are considered in one integral model. At present, however, this approach pose some fundamental problems because of the present interface technology which, in contrast to vehicle control, separates the operators from the time-space properties of the process to be controlled.

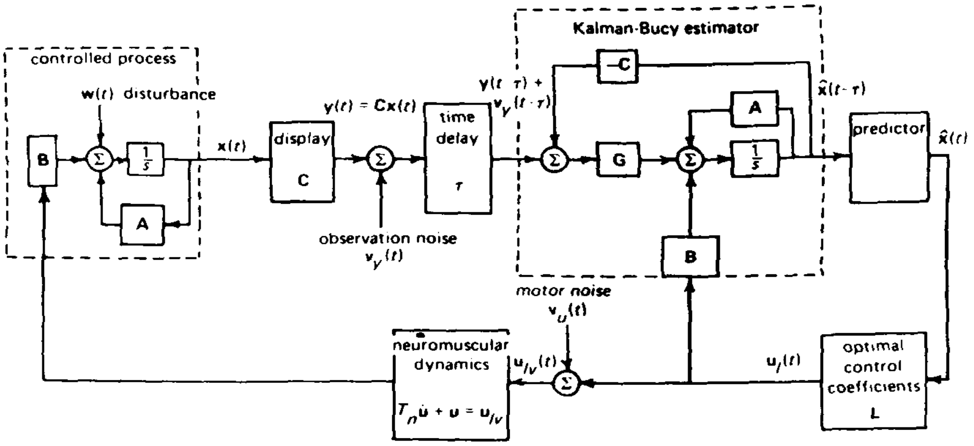


Figure 6 illustrates the structure of the optimal control model of vehicle control proposed by Baron et al.

Conclusion

At present, integrated, quantitative models are relevant for design of aviation and other vehicle systems because the decision and manual control tasks of the operator form one integrated task in direct coupling with system dynamics. The task of a pilot or driver is a direct space-time control of a moving physical object, the vehicle. The sensori-motor level of the human information processing in this task serves for control signal processing; i.e., the output manual actions are continuous signals. For process plants of the present levels of automation the continuous control signal processing is, however, automated. This means that human output actions will typically be related to switching and valving, and will be interpreted as stereotyped signs by the plant systems. Models at the manual skill level is related to interface manipulation at the system surface while performance at the rule based choice level and the knowledge-based decision level is related to the deep functional structure of the system. Thus, the two levels of modeling are not tightly related and the interface separates the part of the system, that is accessible to quantitative, control theoretic modeling and the part, which has to be represented by qualitative, decision making concepts.

This situation may, very likely, change if 'direct perception/direct manipulation' interfaces are developed representing a direct mapping of the deep structure of the work system. In that case, the three levels of skill, rule, and knowledge based control will add up to an integrated control structure.

THE PERSPECTIVE OF PHYSICS

Another approach to quantitative models of complex systems has been made from the perspective of theoretical physics which has a long tradition for modeling the behavior of physical system at several levels of abstraction such as quantum mechanics, particle physics, thermodynamics, etc.. The observation that stable, regular patterns evolve in a turbulent flow of fluids under certain conditions has lead to physical theories of the evolution of organizations of matter at several levels.

Iberall and Soodak: Physical Basis for Complex Systems

In a series of papers, Iberall and Soodak (1978), Yates and Iberall (1982) have promoted a theory of complex systems based on a multi-level presentation developed by a recursive application of the micro- and macro-theories of Newtonian physics and thermodynamics. They argue that complex systems can be comprehended dynamically through an extended statistical mechanics, irreversible thermodynamics, and non-linear mechanics, in the form of an electrodynamic field physics. Iberall and Soodak (1978) propose to extend the application of the theories of 'pure physics' to complex systems including 'planetary atmosphere, biological organisms, human societies, ecology, and galaxies' by means of a number of 'bridging propositions':

1. A deterministic continuum description of homogeneous matter must include dissipation for consistency. In this sense, mechanics implies thermodynamics.
2. An ensemble of interacting atomistic entities ('atomisms') at any organizational level acts like a continuum, at an appropriate time scale.
3. A fluid-like continuum at any organizational level becomes dynamically unstable at some sufficient scale of stress, creating a spectrum of patterned structures of super-atomisms that are freely mobile in broadly extended media.
4. The physics of interacting super-atoms, including their internal processes, appears as ad hoc at their level, even though derivable from the physics of the lower atomistic level.
5. The dissipative nature of a continuum, as required by Proposition 1, implies fluctuations at a lower level of organization.

They claim that

> "these propositions link the hierarchical systems found in nature, and imply a statistical mechanics-irreversible thermodynamics for each level."

and offer a

> "rationale for the extension of statistical mechanisms of interacting atoms or super-atoms to all levels of organization."

They argue that structure emerges when stress in the abstract continuum reaches a level given by a generalized Reynold's number. The Reynolds number, usually the ratio of inertial to viscous forces, is generalized as a critical condition to express the ratio of convective velocity across a field to the diffusive velocity representing energy or momentum being absorbed into the field.

> "The generalized Reynolds number provides the basis for creating hierarchies in the universe. If a homogeneous field receives an input flux such that high Reynolds number conditions are created, inhomogenities will appear that become atomisms in a new, emergent level of a statistical mechanical field. At each level in a hierarchy, the same physics is re normalized around the scales of the emergent atomisms and processes."

They find that

> "living systems at the individual, organismic level have another (approximate) summation invariant, beyond mechanical conservations: population density or number. That is, organisms behave (reproduce) as if to preserve the species population in their ecological niche, albeit with fluctuations."

In addition to some physical examples, Iberall and Soodak apply the approach to formation of settlements in a human society. Based on assumptions about parameters such as equilibrium daily energy available to humans, energy released in internal modes, human time constants of action modes, time constants of adaptation, translational viscosity and generalized Reynold's number, etc., they develop trading constellations of primitive societies which they compare to published estimates and note:

> "Thus the application of physical notions to society is not far fetched."

They conclude:

> "The net effect of our Propositions is to provide a formal recognition of an 'intuitive' wisdom that there seems to be interaction from dei ex machina above and below. We identify the mechanisms from below with their basic atomistic fluctuations (illustrated at different levels as nuclei, atoms, cells, individual organisms, social constellations, turbulent air masses, rivers, lakes, earth plates, or galaxies). In their hidden variable covariations, they make up the gross formed or mean status of the macroscopically isotropic and homogeneous 'material' systems above. We identify the mechanisms from above as the convective stresses which provide organizing competence."

Yates and Iberall (1982) introduce the concept of information and notes that we need a richer, semantic concept than the formal information theory of Shannon to be able to account for the coordination within complex systems. They note, that

> "any definition of information is more static than dynamic it does not depend strongly on rates. Nevertheless, it can lead to modulation of dynamic (rate) processes. That is the chief function of information."

Discussing the 'modulation' of the activity modes of flowering plants, they consider information of chemical languages as being

> "catalytic linkages among mechanisms and thermodynamic engines in 'soft-coupled' (complex) systems."

Conclusion

The problem we face with this approach in our context is its entirely behavioristic basis, with no consideration of the prehistory of the organism, that is, its memory and ability to learn from previous activities. The approach is probably fruitful for the initial evolution of organisms, but not for the phases when *functional competition* and *selection* are in action. This raises some questions of 'Vitalism' in modeling complex systems.

Yates on Vitalism

In a review paper, Yates (1982) discusses the concept of function and intention and the effort found in some approaches to theoretical biology to avoid Vitalism. He refers to Dennett's observation (1978):

> "Anytime a theory builder proposes to call an event, state, structure, etc., in any system (say the brain of an organism) a signal or message or command or otherwise endows it with content, he takes out a loan of intelligence. He implicitly posits along with his signals, messages, or commands something that can serve as a signal reader, message understander, or commander, else his 'signals' will be for nought, will decay unreceived, uncomprehended."

Yates concludes from this:

> "If physiologists have any mission that we all should share, it seems to me that it must be to account for the richness of behavior of the organisms, at all levels of organization, by a level-independent physical theory of selforganization and dynamic stability, that doesn't require a loan of intelligence."

This attitude shapes the entire effort towards a physical theory of living systems. He reacts e.g., to Waterman's comments:

> "Biophysicists and molecular biologists often seem to believe that physical and chemical explanations of life are the final goal of biology. One may as well say that a Fourier analysis as such can provide *the* explanation of a Bartok quartet. The ultimate scientific explanation of life must indeed be biological. Both historically and functionally live organisms share a number of basic organizational and behavioral properties not known in the nonliving world."

Yates objects to this as being 'organismic biology' and he refers to Claude Bernard's (1878) theories of an organism which were the first to introduce the notions of regulation and stability, concepts which were later elaborated by Cannon (1929) in terms of 'homeostasis.'

Yates concludes the historical overview by the statement that, in his view:

> "-living systems are yet another, e.g., a fourth physical state of matter. Gas is matter that fills a container, liquid is matter that conforms to the shape of the container, solid is matter that supports its own shape; life is matter involved in an actual or potential '--self-perpetuating open system of linked organic actions, catalyzed stepwise and almost isothermic by ... specific organic catalysts which are themselves produced by the system' (Perrett, 1952). Life is further characterized by autonomous morphogenesis, nearly invariant reproduction, and teleonomic behaviors (Monod, 1972).' ----'The local conditions that constrain the ordinary laws of physics to produce the biochemical and physiological operations of life have been thought to lie outside the possibility of reduction to physics (Polanyi, 1968). It is true that physics is not traditionally a science that has sought to extend its explanations to historical, evolutionary processes, even though it was concerned early with both motion and origins of the stellar and cosmic firmament. Now it has extended its scope and revived concern about origins of all objects in the universe."

Following the historical review, Yates presents an extension of the theory of 'homeostasis' by means of a physical theory of dynamic regulation called 'homeokinesis,' based on the work of Iberall and Soodak (1978).

He asks:

> "What would a physical theory of an organisms have to account for?"

and he answers that

> " --any physical theory of an organism that would be useful in physiology would have to describe and predict the peculiar stability characteristics we see in multicellular forms of life, for example, in mammals."

He cites Cannon's propositions from 1925:

> 1. In an open system such as our bodies represent, compounded of unstable material and subjected continually to disturbing conditions, constancy is itself evidence that agencies are acting, or ready to act to maintain this constancy.
>
> 2. If a state remains steady it does so because any tendency towards change is automatically met by increased effectiveness of the factor or factors which resist the change.
>
> 3. Any factor which operates to maintain a steady state by action in one direction does not also act at the same point in the opposite direction.
>
> 4. Homeostatic agents, antagonistic in one region of the body may be cooperative in another region.
>
> 5. The regulating system which determines a homeostatic state may comprise a number of cooperating factors brought into action at the same time or successively.
>
> 6. When a factor is known which can shift a homeostatic state in one direction it is reasonable to look for automatic control of that factor or for a factor or factors having an opposing effect."

With regard to Cannon's statement, Yates notes:

> "Unfortunately, since the announcement by Norbert Wiener of 'Cybernetics' in the 1940's, this sentence (item 6) is likely to be read as a call for application of an engineering theory of feedback or servomechanisms. It is exactly that inappropriate notion that has become a prominent part of the community of ideas now called 'biomedical engineering,' and it has seductive appeal to engineers who begin to look at biological systems. Fundamental to such feedback analysis is the decomposition of organisms into factory plants and controllers extrinsic to those plants. But it was precisely in recognition of the limitations of those notions that homeokinesis was born, out of a more physically oriented thermodynamics and statistical mechanics. Its central image is of statistical fluctuations of atomistic-like entities at various levels (molecules, cells, tissues, organs, human bodies in societies) that themselves create and maintain *both* the *plant* and *regulatory* functions. The emphasis is on circular causality, not on arbitrary feedback schemes imitating an inapplicable technology vision."

Yates continues:

> "The greatest difficulty in reducing biology to physics is that we can't use a representation theorem to relate the terms and statements of biological theory (e.g., evolutionary theory or genetics) to physical theory, partly because our biological theories seem so inherently unphysical in character, i.e., not deterministic with respect to any state variables (footnote: Instead, biology presents a picture to us that seems to call for a statistical mechanics for multiple atomistic levels, such as might apply to genetic material, organelles, cells, individuals, etc.). Therefore, the science of biology seems not simply lawful in the physical sense, but instead seems to consist of adherence to rules. Rules, in contrast to laws, are local, require a corporeal entity for their execution, and are rate independent (Pattee, 1977). They are more syntactical than dynamical."

Conclusion

Yates' arguments seem to depend on a rather simplistic perception of control theory in an 'engineering sense.' He seems to perceive control theory as being concerned with separable control mechanisms, as represented by the usual feedback system diagrams, not as abstract, 'practically isolated' functional relationships which, very often, are embedded in what he calls the 'plant.' Furthermore, he seems to have a strong bias with respect to represent biological systems at the level of 'particle physics' (cf. his footnote statement). The real problem in representation of living, that is, biological systems, seems to be to define deterministic, practically isolated relationships (in Russell's sense) among variables in behavioral, not physical, terms at several levels of abstraction. The basic problem is to represent the mechanisms and criteria of selection among alternatives, during evolution (survival value) and during learning (storing successful paths to purposes, goals). The problem is not to avoid the 'loan of intelligence,' but to represent the implicit and explicit function behind directional and convergent selection.

The main body of Yates' presentation includes discussions of 'the physical scales of life' in terms of scales of size, mass, time, and energy. He considers life as a constellation of coupled physical and chemical processes carried out largely internally, in gas, liquid, gel, or solid phases, with plastic elastic fluid processes dominating. The chief processes are chemical transformations, and mass, charge, energy, and information transports. Diffusion and convection dominate the transportation modes, with convection and chemical reactions providing important non-linearities. He then reviews the potentials that support terrestrial life: Physical, such as e.g., chemical, electrical and gravitational potential, mechanical pressure, genetic, such as 'information content of genome,' and epigenetic, such as 'value potentials of culture or society or tribe.' The members of this set seem to be of different conceptual categories, in particular the 'information content of the gene' differs from the rest. Yates recognizes this difficulty: it

> "is not yet treatable as a physical potential, even though it clearly is a kind of chemical potential."

He discusses in detail his physical theory of homeokinesis for metabolism and the cardiac system, after a demonstration of the basis of homeokinesis by the simple system including a battery with a load resistance:

> "The battery is also a regulated system. It involves internal chemical reactivity, and internal resistances. It modifies its voltage output to the resistance load, but keeps its voltage output (as its potential) relatively constant over a wide range of current flows in the circuit and for a long time. We intend it to regulate itself!"

It is not quite clear, why we need a special theory of homeokinesis to account for that kind of mechanisms. Representation of intrinsic feedback mechanisms, active without a separate 'controller,' is trivial within control theory.

Yates' position seems to be an over-reaction to the fear of Vitalism, and appears as a fundamental regress to radical behaviorism. If living systems have evolved from basic physical matter, a kind of implicit 'intelligence' emerges from the simple fact that organisms exist and survive, a kind of 'intelligence' (in terms of reward of survival values) is implicit in the functional competition. The mere selection of organisms most suited for survival can be considered some kind of 'final cause' or 'intelligence.' Who is doing the selection? The global system itself, it can be embedded in the dynamic description of the system, as argued by the physicists, but this fact does not imply that a concept of purpose is not useful for modeling the functions of an organism once it has evolved. The search of a quantitative, relational model of the behavior of living systems should not be limited to search for mathematical models phrased in concepts and terms of physics. There is no reason to assume that recursion of a model in terms of mechanics and thermodynamics up through the levels of living systems and organizations is the most fruitful approach. In contrast, there is good reason to believe that particular concepts and terms are required for each level, just as different concepts and terms are used for different levels of physics (Quantum physics, Newtonian physics, thermodynamics). An entirely physical theory is necessary to explain the organization of dead matter into living organisms through the early stages of evolution, but not necessarily through the later, competitive stages, and not at all for the behavior of the organisms, once they function and consciously formulate goals and subjective performance criteria. *Typically, all the specific examples used by Yates in his review are focused on the functioning of physiological mechanisms, not on higher level purposive behavior.*

THE PERSPECTIVE OF MATHEMATICAL BIOLOGY

Mathematical biology has primarily been concerned with the origin of organisms in terms of self-organization. The question posed is: How can an adaptive and autonomous organism evolve and how can the purposive control of the behavior of an organism be modeled? The writings of Pattee are illustrative for this approach.

Pattee on Modeling Adaptive Organisms

In a number of publications Pattee (1973, 77, 83), has been concerned with the analysis and description of 'natural control systems' in living organisms and their origin. His approach is based on the concepts of physics and is particularly relevant in the present context.

Pattee on Natural Controls

Pattee (1973) describes the physical embodiment of a control system as 'non-integrable constraints' on an organism's behavior and notes that constraints, in general, require alternative descriptions to the microscopic description of the organism:

> "This alternative description selectively ignores detail which corresponds in the physical system to some form of dissipation process which gives rise to the new coordinations of the constraint and a simplified collective behavior we recognize as function. The origin of such control constraints must begin with low selectivity and imprecise function and gradually sharpen up to high specificity and narrow, precise function."

The basic idea in Pattee's formulation of control theory is to produce a desirable or predetermined behavior in a physical system by imposing additional forces or *constraints*. Pattee's problem is to explain the origin of control that arise naturally without human intervention.

Pattee' first essential point is: Since a control system cannot be distinguished by its material basis, we need for identification an alternative description which for some purpose is more useful than the microscopic, dynamical description. In physical terms, a control device is a time-dependent constraint which alters the path of selected degrees of freedom of the system it controls in a variable, but regular way. In mechanics, such path-dependent, non-integrable relations are called non-holonomic constraints. The basis of such constraints is to make the number of dynamical degrees of freedom less than the number of coordinates necessary to specify the configuration of the system. Pattee notes that:

> " ----any alternative description is either redundant, covering the same level of detail in an equivalent way or else the alternative description is an abstraction covering less detail in certain degrees of freedom but realizing in return some additional regularity in other aspects of the system."

If these 'other aspects' are taken to be the 'final cause' of evolution, his approach is analogical to the alternative representations at the various levels of the abstraction hierarchy of our framework. A similar point was raised by Rosen in the commentary following Pattee's presentation:

> "The problem of alternative descriptions by the author in connection with the origin of life problem is a profound one in many other areas of biological investigation (or indeed wherever one finds a complex system or organization). It appears even in engineering, in the distinction between a 'state variable' description and an 'input-output' or 'black-box' description of the same system."

He considers

> "the essential part of the control constraints" to be "that part of the configuration which internally executes some form of averaging process over selected degrees of freedom while at the same time, through its dissipative process, establishes new correlations between other variables which then appear as the control variables."

The basic problem of the evolution of control is that

> "each element of the coding system is highly selective and simple in its function but loses all significance without the preexistence of all other coordinated elements. Thus no partial code seems to make sense, so that there does not appear to be any obvious gradual process leading to such a highly coordinated set of control molecules. On the other hand, there seems to be no reasonable likelihood of a sudden, spontaneous occurrence of such a complete, integrated collection of control elements."

Pattee perceives this difficulty as a reflection of the representation:

> "In most physical theories, we define the system we are talking about by fixing a number of degrees of freedom and then expressing the forces or interactions acting on these variables. It is easy to conceive of these forces being gradually removed so we can see how the behavior of the fully interacting system grows continuously out of the non-interacting particles. We treat equations of constraint in quite another way. They are either present or absent there is no gradualness about it. This is because the equations are an alternative description."

From our point of view, this relates to the fact that the higher level representations act as final causes which can only be defined after the fact, that is, after survival.

Pattee on Levels of Description

Adaptability requires self-description. Pattee takes up the question of levels of description in subsequent papers (e.g., Pattee, 1977). In this paper, he approaches the question of representing 'complex systems.' He notes here:

> "Furthermore, defining various numerical measures of complexity is an exercise that seems to leave out the *functions* of complex systems, their adaptive behavior and their developmental and evolutionary potentials. These are characteristics of the large class of complex systems we associate with life and its artifacts."

He emphasizes that 'living systems contain their own description.' Von Neuman conjectured a threshold of complication below which systems spontaneously degenerates into simpler systems and above which there is evolution of even greater complication. Above the threshold, the system comprises a kind of universal constructor and a description of what is to be constructed, along with a suitable executor or control program, and environment consisting of the necessary parts. The meaning of 'description' should be taken quite literally, as a *representation that is written* rather than an image or model, that is, a physical analog or likeness of the system. This, in turn, implies a symbol system, syntactical rules for combining symbols into interpretable statements, and a mechanism for effectively writing, reading, and executing such statements. If an analog was implemented, the model in order to anticipate the behavior of the system should run faster than the system and if this model should be adaptive, another model should again run faster, and so on until infinite regress. (This argument is not too convincing, because adaptation depends on trial and error and selection can be possible without this regression).

The illusion of non-biological machines. Pattee argues against the idea that 'non-biological machines' exist:

> "The fact that we design machines that can, for a time, operate as physically separate and informationally autonomous systems does not make them less biological or less dependent for their design, construction, repair, adaptation and evolution on their ultimate interaction with humans and the languages of the brain."

In this way, he is in line with our conception of the representation of intentionality within all purposive systems.

Description is not simple:

> "The concept of description has its own threshold of complexity. A large part of the mystery of the origin of language as well as the origin of life results from our inability to reduce the number of elements and the number of rules that appear essential for the writing and reading of any structure that could be called a description."

Von Neuman developed formal descriptions of self-reproduction and self-construction, but only by the loss of most of the physical and chemical concepts and languages that are considered fundamental to both physics and chemistry. He relates this difficulty to the different complexity of mapping changes in the physical world onto different levels of description. The mapping of changes onto representation of laws of motion involving time, space, and energy is rather straight forward whereas for abstract, formal descriptions, a particular representation is required for each change. In addition, linguistic rules of formal descriptions can produce only discrete serial operations rather than parallel continuous operations (cf. the discussion in the introductory sections of the advantage of causal models for representing change). Construction, on the other hand, can often be performed most efficiently with parallel and coordinated dynamics. Finally, Pattee points to the interest in formal programs and machines depending on the predisposition of mathematicians for generality and universality whereas real devices necessarily are special purpose and explicit representations of logical operations that depend on an informal interpretation or measurement process that does not itself have a detailed description.

> "It is the characteristic of real measuring devices that a complete dynamical description of the device is not only unnecessary but even inconsistent with the process of measurement."

This once more stresses the difference between function and implementation processes:

> "The difficulty in relating dynamical laws of systems and the process of measurement on systems has never been clearly resolved even for elementary physical systems. The remarkable fact is that our tacit or intuitive concepts of measurement serve us so well in establishing our models of the world. The only plausible explanation for this is that natural selection has provided us with a brain structure that needs no linguistic description to interpret measurements properly, just as the enzyme needs no description to fold and recognize substrates properly. But of course this begs the question of how internal descriptions of systems interact with the operating system they describe How do we know when physical structures like a DNA molecule or a magnetic tape are also descriptions of systems?"

This question raises the concept of top-down intentionality, either by human design or the final cause of evolution, that is, survival value.

Laws and Constraints, Rules and Syntax. Pattee distinguishes laws by the fact that they have

> "a primary existence that does not require a description or a physical embodiment for their execution."

while descriptions

> "require special structures that constrain the motion of matter beyond laws, which they must always obey."

An evident example from our context is the steam engine: the work of the steam in the cylinder depends on thermodynamical laws, whereas the sequential process of steam intake and exhaust depends on the constraints, that is the description of the cycle, embedded in the valve linkages. The brain has evolved so that people can manipulate strings of symbols according to complex rules without being aware of the processes of the brain that are constrained by the statements of the rules. In philosophy, this is the mind-body paradox, in physics it is phrased as the measuring problem: A measuring device results in detailed record of an event that does not depend on any detailed knowledge of the measuring device itself. From the point of view of our framework, this does not appear to be a problem, measuring is a function defined independently of the underlying physical process, that is, an input-output relationship calibrated in symbolic terms.

This leads to Pattee's definition of complex systems as being systems able to write and read their own messages.

The Physics of Reading and Writing. Writing a symbol

> "requires a rate-dependent dynamics augmented by non-holonomic auxiliary conditions or constraints that result in freezing out specific degrees of freedom leaving a rate independent structure."

Reading is not the inverse process but the

> "generation of the meaning of the symbols from their structure, which in our physical context means the active control or constraint of a rate-dependent dynamics according to the coded description contained in the rate-independent symbol vehicles. Reading requires additional non-holonomic devices totally different from writing devices."

Discussing in some detail the biological processes of a living cell, he concludes:

> "The essential process of interpretation is a dynamic, highly parallel interaction of all the elements constructed by the linear, sequential, rate-independent reading process." --"Only when syntactical rules act as non-holonomic constraints on physical laws can the adaptive process operate efficiently in complex systems. If this is the case, then we would not expect a complete, formal description of a complex system to adapt or function as rapidly or reliably as the partially self-describing, tacit dynamic system it simulates."

His concern is that

> "over-formalization in a single language may obscure rather than illuminate nature's most successful adaptive and evolutionary strategies."

Dynamic and Linguistic Complementarity. From Pattee's point of view, the distinction between rate-dependent dynamical processes and the rate-independent linguistic description is essential to development of models of complex systems and he finds that relatively isolated development of dynamic and linguistic descriptions has caused the key problems of complex systems to be largely neglected.

He concludes that biological evolution can be characterized by the increasing elaboration of internal descriptions and models which are usually called by other names such as goals, plans, policies, strategies which are only higher levels of self-description in a linguistic mode which Rosen calls 'anticipatory' and 'feed forward' systems which cannot be represented by conventional dynamical theory.

The Evolutionary Process.

In a later paper Pattee (1987) takes up the discussion of the basis of the modeling of the evolutionary process. He starts noting that evolution has been approached from the point of view of several different disciplines, and that every one sees self-organizing behavior within the framework of their own particular discipline. His discussion is focused on two classes of self-organizing systems, namely the statistically unstable systems and the information-dependent systems which have been discussed separately in the languages of physics and physical chemistry respectively the language of biology. The former is formulated in terms of laws of the basic laws of nature which is not at all the case for the latter. The difference between the two approaches is partly caused by the differences of the two disciplines, partly by the enormous difference of the complexity of the systems considered.

He notes that

> "even though these laws of physics are a foundation for all organization, including self-organization, we recognize in the self-organizing behavior of both non living and living systems many entirely new forms and patterns that are not simply the perturbation of stable systems or the probabilistic behavior of unstable systems. The novelty and persistence of emergent forms characteristic of living systems do not fit our definition of either stable or unstable behavior. Therefore, we may expect theories of self-organization also to require complementary subjective or functional modes of description."

After discussing Bohr's theory of complementarity, he concludes,

> "I simply find no alternative but to accept multiple formally incompatible descriptions as a satisfactory explanation of many types of biological events."

Considering the micro and macro-levels of description of physics based on reversible (time-symmetric deterministic laws, respectively irreversible statistical laws, he suggests

> "that chance is displaced in some optimal sense by informational constraints that efficiently control the higher levels of dynamic behavior."

He finds a fundamental difference between the types of self-organization found in dissipative structures of macroscopic chemical systems and even the simplest living systems, that is between system that develop by chance and those constrained by symbolic information. Physical laws do not change with time and have no symbolic memory of past events. The basic difference between the physicist's and the biologist's approach to self-organization is that the physicist's theory recognizes no symbolic restrictions and no historical regularities whereas the biologist's theory assumes genetic symbol systems and a vast prehistory of selected historical structures. The physicist's instability-based concepts and the biologist's information-based concepts are closely linked to the two levels of the function-form complementarity.

Poincare defined an unstable system as a system in which a small change creates a large effect. In that sense, the effect of a symbol depending on the arbitrary or conventional symbol-referent relation basically requires a physical system which exhibits some form of instability. Stable dynamics generates no alternatives and there is no need for informational control except for control of boundary conditions.

However, there is a basic discrepancy between the characteristics of dissipative structures of physics and the informational, symbolic structures of biology. Dissipative structures are dynamic, that is, they depend on the rate of matter and energy flows. In contrast, symbol systems exist as rate-independent constraints.

Symbols have no meaning outside the context of a complex dynamical organization for which the informational constraints have evolved. It is without meaning to search for the meaning of symbol strings since the symbolic constraints are only significant near the dynamical instabilities. From this point of view, the form-function relationship of modern computers is an extreme example, since the computer hardware has no stable dynamics at all. From this it appears that the central epistemological problem of cognitive simulation is to distinguish the part of the simulation that simulates the neural dynamics from the part that simulates the neural symbolic information, since in a general purpose computer, all dynamics are simulated by information,

Conclusion.

In conclusion, the models derived bottom up from laws of physics as advocated by Iberall, Soodak and Yates are inadequate for representing conscious and purposive systems. For that purpose, models representing functionality at several levels expressed in different languages are necessary, as concluded by Pattee. Furthermore, it is important to consider the basic difference between the physicist's and the biologist's approach to self-organization, that is, that the physicist's theory recognizes no symbolic restrictions and no historical regularities whereas the biologist's theory assumes genetic symbol systems and a vast prehistory of selected historical structures. Therefore, when modeling the evolution of sociotechnical systems, the ability for generating value structures and performance criteria, for memory and learning from previous activities must be explicitly represented.

QUALITATIVE, CAUSAL MODELS

Whether or not we succeed in developing quantitative relational models of the behavior of purposive systems, qualitative, causal models are important for representation of human reasoning and decision making at the higher levels of system representation. Commonsense reasoning depends on a decomposition of the complex environment by categorization and concept formation and the study of this process is biased by a tradition having its roots in the classic Greek philosophy (see e.g., the discussion in Bruner et. al. 1956). It will, therefore, be natural to introduce a review of causal models of complex systems by a reference to the Aristotelian physics.

ARISTOTELIAN PHYSICS

Causal modeling of purposive systems goes back to Aristotle who distinguished between four kinds of causes which he illustrates by the process of building a house: 1) the *material cause, i.e.,* what it is made of; 2) the *formal cause,* i.e., the architects conception of its shape or form; 3) the *efficient cause*, i.e., the process of building the house; and 4) the *final cause*, i.e., the purpose of the whole operation. This latter concept has been taken to imply a cause subsequent to its effect. His concepts, therefore, in general have discredited teleological approaches. This position, however, according to Bambrough (1963) is based on a misinterpretation of Aristotle:

> "Because of modern developments in the physical sciences, we have come to think of causation primarily as a relation between events; but for Aristotle, the four causes were primarily causes of things or substances. The doctrine is intended as an account of how particular substances originate or "come to be" and why they have those properties that we recognize in them. When we think in these terms, and escape from our customary preoccupation with events, we can see at once how natural and how closely interconnected are the four questions to which Aristotle's four causes indicate the relevant types of answer: What is it? What is it made of? How was it made? Why was it made?"

From this point of view, Aristotle's different causes are more closely related to the relationships among concepts within the means-ends hierarchy of our taxonomy, rather than to modern causality in terms of relations between events or, in other words, his concept of cause is related to the creation of systems, not to their functioning. Hence:

> "it is no accident, but an essential feature of his causal theory, that the four causes should be illustrated so often and so effectively in terms of examples drawn from biology and from human skills of manufacture."

That is, from systems evolving through selection by a designer according to some intended purpose, or by nature, according to some 'survival value.'

Aristotle also elaborates upon the prototypical nature of his categories:

> "The sense in which all things have the same causes might be paraphrased by saying that there is a single framework of causal conceptions into which the causes

> of all individual substances can be fitted. Every substance has a form, but not every substance has the same form: every man has a father, but not every man has the same father." (Bambrough (1963).

It appears, in this interpretation, that the classic Aristotelian physics is quite close to the means-ends representation of the environment we use to model common-sense reasoning and there is actually no teleological problem in his 'final cause' concept.

THE PERSPECTIVE OF HISTORIANS

One of the basic problems found in the modeling perspective of physics in our context is the lack of the historical dimension. Physics is concerned with modeling dynamical system, that is, the systems are invariant over time in the sense that given the state equations and the initial conditions, behavior is determined irrespective of the prehistory of the system. The physical theories do not explain how the state equations evolved through history, as Pattee argues. For the modeling of socio-technical systems evolving through their normal purposive activity it will, therefore, be relevant to have a look at some contemporary historical theories on the evolution of societies. From this point of view, Toynbee's theory is particularly interesting. A brief review of his extensive writings is found in Toynbee, 1972.

Toynbee on Cultural Evolution

Toynbee introduces his discussion by an analysis of the problems found in the traditional approach to historical research. He notes that our western culture is governed by two institutions: The industrial system of economics and the system of politics, that is, a representative democracy in an independent national state (cf. the argument (Rasmussen, 1991, Rasmussen et al. 1991), that work organizations are shaped bottom-up by the material work environment and top-down by the social 'style').

In the present, cross-disciplinary research context a brief review of Toynbee's discussion of the cross-disciplinary problems he has met in his research may be relevant. He observes that an industrial system depends on the division of labor and the application of a scientific approach to the operation on the physical environment. Articles are manufactured from raw materials by the mechanically co-ordinated work of a number of people. He then argues that a serious problem is caused by introduction of the same approach to research. He finds that books in the industrial society have been replaced by periodicals

> 'with its division of labor and its sustained maximum output of articles manufactures from raw materials mechanically.'

He accepts that

> "this may well be the right way of handling any branch of physical science in its early stages."

But regrets that

> "the same method has latterly been applied in many realms of thought beyond the bounds of physical science."

The prestige of the industrial systems has imposed itself upon the 'intellectual workers' of the Western world and historians have given their best energies to the 'assemblage' of raw materials. The term 'original work' is restricted to the discovery or verification of facts not previously published. The priorities for choice of research topic are not related to the importance of the potential finding but are instead

> "what is the richest mine of unworked raw material in the field."

Toynbee discusses the two driving forces of modern history, industrialism and democracy and notes that, in fact, rather than democracy, nationalism has been influential being

> "the sour ferment of the new wine of democracy in the old bottles of tribalism."

The national point of view has been attractive to historians because the richest mine of data has been the public archives of governments. He argues that up to the end of the 19th century these two influences have been cooperating in creating 'world powers' but later they have been acting against each others because trade and industry has evolved across national borders. (Consequently, there has been a pronounced tendency of nations breaking up into smaller units having a particular and individual historical cohesion. He actually predicts the present break-up of Czechoslovakia and Yugoslavia).

The field of historical inquiry is, consequently, societies, not national states, and he defines societies as being

> "the total network of relations between human beings. The components of society are thus not human beings but relations between them." -- "A visible and palpable collection of people is not a society; it is a crowd."

He focuses his study upon developed societies in terms of civilizations and argues the necessity of studying the history of particular civilizations and to compare the results to understand the driving mechanisms.

Considering the study of a particular civilization, he strongly emphasizes that history does not

> "run along a single line and single line charts of history does not work. Multiple-track charts are the only kind that will fit the phenomena as we find them."

This view implies that it is no longer possible for an observer to take events as he finds them and to represent them by a narrative, instead a model must be devised, and

> "if a symbol is to work effectively as an instrument for intellectual action that is to say, as a 'model' it has to be simplified and sharpened to a degree that reduces it to something like a sketch-map of the piece of reality to which it is intended to serve as a guide."

He goes on noting that the question is not whether the model is valid in itself, but whether it is useful for getting insight.

Discussing the transitional societies, Toynbee emphasizes the interaction between technology and religion. Technology shapes the need for coordination, religion serves to motive society to actually respect this need.

He exemplifies this point by the transition from a hunting to agriculture required introduction of foresight and long term planning, exploitation of metallurgy is based on specialization and cooperation among professions. This evolution, in turn, require production of agricultural products in surplus and, in some regions, therefore irrigation and large-scale production planning leading to a complex social structure. Since the average individual can not be expected to have foreseen in his imagination the fruits of his efforts, he "must have been induced to work in faith or coercion, or both." Therefore, the influence of religious ideas is crucial. Toynbee notes that the stages of evolution of technology and religions are not corresponding one-to-one but are considerably off-set and he interaction is complex. Technological changes are both causes and consequences of social changes. Once the social organization and religious value structure needed to control a technology, they tend to arrest the advance of technological change.

An important perspective resulting from Toynbee's comparative studies of cultures is the conception of social structures as evolving from a functional interaction between top-down propagation of value structures and bottom-up propagation of material opportunities and constraints. This is very similar to our view of the dynamic evolution of the organization of cooperative work (Rasmussen et al. 1991) which is the result of comparative studies of different work systems .

THE PERSPECTIVE OF BIOLOGICAL EVOLUTION

Given the review of mathematical biology found in previous sections, a brief reference to the present theories of biological evolution may be relevant..

Gould's Theory of Evolution

Theories of biological evolution have traditionally been depending on the concepts of 'extrapolationism' as being the process underlying change (that is, the gradual substitution of different 'alleles' in many genes) and the reliance on selection leading to adaptation through a gradual change of species. Presently, this view is being challenged by a theory of hierarchical change. Gould (1987) reviews this trend. He considers evolution as a hierarchical process with complementary, but different modes of change at three large scale levels: a) Variations within populations, b) speciation, and c) very long-term macro-evolutionary trends. Speciation is not always an extension of gradual adaptive allelic substitution, but may represent a different style of genetic change, a rapid reorganization of the genome, perhaps non-adaptive. Macro evolutionary trends do not arise from the gradual, adaptive transformation of populations, but usually from a higher order selection operating upon groups of species. Individual species generally do not change much after their, in geological perspective, instantaneous origin. Gould labels the two discontinuities in the evolutionary hierarchy the Goldsmith break (change in speciation is different from change in populations) and the Wright break (speciation is different from macro-evolutionary trends).

This differentiation between evolutionary changes may be of interest for further study, considering our approach to learning automata, for which we have considered different adaptation mechanisms such as adaptation (optimization) by local hill climbing and by global stochastic exploration of a work space. For such evolution of heuristics and goal-directed adaptation, some system of performance criteria and self-observation is important and, consequently, the approach of Maturana and his cooperators seems to be relevant.

Maturana's Autopoesis

The concept of 'Autopoesis,' has been formulated by biologists, to represent the 'ability of living systems to maintain their unitary continuity of pattern, i.e., their autonomy, individuality and distinctiveness, despite of the ceaseless turnover of their components. It has been adopted by other professions, in particular the social sciences.

It has been defined by Maturana and Varela (1980):

> "An autopoetic system is a distinguishable complex of component-producing processes and their resulting components, bounded as an autonomous unit within its environment, and characterized by a particular kind of relations among its components and component-producing processes: the components, through their interaction, recursively generate, maintain, and recover the same complex of processes which produced them. It is important to note that the unity-characterizing complex of processes is assumed to be invariant: it is being "the same" complex of processes. Further, the components, as the terminal products of component-producing processes, are the prerequisites for the activation of the component-producing processes themselves.' A particular complex of processes are referred to as 'the organization' whereas the particular spatio-temporal configuration of components is called 'the structure.'"

These two concepts maps directly onto our means-ends hierarchy, the structure referring to the physical form level, the organization to the process or the function level, depending on the language of representation used to represent the autopoetic process.

In the example describing an autopoetic model of a simple cell, found in Zeleny (1980), a general, functional language is used. It assumes the minimal organization of components that can demonstrate autopoesis: A catalyst interacts with a substrate so that membrane-forming components can continuously be produced. The components are: 1. holes, 2. substrate, 3. free link, 4. single-bonded link, 5 fully-bonded link, 6. catalyst. The processes are: 1. production, 2. bonding, 3. disintegration. (The behavior of this system is similar to the well known game of 'life'). When the catalyst is added to the substrate, a 'structure' evolves characterized by a membranous boundary that identifies the system as a separate and autonomous unit.

The point is, that the 'system' is generated by its components, no control structure is needed. Maturana and Varela argue against anthropomorphism:

> "In our terminology, although the concept of a gene is useful for elucidating structural differences, its role in explaining the organization and the autopoesis of the living systems is still to be demonstrated. There is neither logical nor factual support for the supposition that organization can be explained in reference to 'gene' interaction alone."

With reference to Weiss (1925) they find that

> "genetic methodology can only test structural differences between living organisms. There may be many "genes" referring to differences in color and specific architecture of eyes and hair, for example. But there are no "genes for eyes" or "genes for hair" that could explain the basic formative dynamics by which these systems attain and retain their generic configurations".

The approach appears to be an analog to the physical bottom-up approach by Iberall et al. It appears as a transfer to the causal modeling domain of the physics perspective and has very behavioristic overtones. The example used is based on an object-and-event representation analyzed by object oriented simulation. Furthermore, no specific examples of the materialization of the theory on a complex organization has been described. The aim is to get as far as possible without invoking the linguistic (rule-based) control of the evolutionary choice.

Zeleny presents a review of the history of the approach almost exclusively including French and Russian sources. He quotes as precursors: Claude Bernard (1864), Giovanni Battista Vico (1774), Bronislaw Trentowski (1843), A .A. Bogdanow (1912), Stephane Leduc (1911), and General Smuts (1926).

Maturana and Varela discuss 'machines, living and otherwise.' The relations that define a machine as a unity are the 'organization.' This organization can be realized in many different manners by different kinds of components that, given in space constitute the 'structure.'

> "The use to which a machine can be put by man is not a feature of the organization of the machine but of the domain in which the machine operates, and belongs to our description of the machine in a context wider than the machine itself." - "However, we use the notion of purpose when talking of machines because it calls into play the imagination of the listener and reduces the explanatory task in the effort to convey to him the organization of a particular machine. In other words, with the notion of purpose, we induce the listener to invent the machines we are talking about. [Note: the economy of intentional explanations] This, however, should not lead us to believe that purpose, or aim, or function, are constitutive properties of the machine we describe with them; such notions are intrinsic to the domain of observation, and cannot be used to characterize any particular type of machine organization.

This argument appears to be stressed to avoid teleological explanations of the function of an existing, invariant system organization which is in focus of the autopoetic analysis. If, however, 'organization' is taken to include the functional relations which can result in the evolution (by survival of the fittest organism selected by a competitive environment or by choice following self-evaluation of performance according to a performance criterion) the case is not that clear cut.

As discussed by Zeleny, Maturana and Varela define autopoetic systems in terms of systems without input and output, but maintaining their invariant organization. They stress, that teleology and teleonomy are notions which

> "are unnecessary for the understanding of the living organization."

The focus of Maturana's approach is on the self-reproduction of an 'invariant' organization. Therefore, concepts and results from their approach do not appear to be relevant to the present context of functional evolution of goal directed systems.

ARTIFICIAL INTELLIGENCE PERSPECTIVE

The basic weakness of the bottom-up approaches taken by physics and general system theory is the lack of representation of symbolic, 'written' control structures and of an explicit representation of processes, functions, and goals at several levels of abstraction. Considering this difficulty, it is pertinent to have a look at the models of human reasoning developed within Artificial Intelligence for a supplement to a mathematical representation of the lower level physical processes.

Modeling within AI has been strongly influenced by the research of Simon and Newell. Research is primarily directed toward the development of artificial reasoning systems, not on the development of models of human behavior albeit it draws heavily on analogies with the functioning of human intelligence. The proponents of the approach has been highly optimistic and have generalized beyond reason, but has been very effective in the development of tools for qualitative representation of human behavior in terms of sequences of events and arguments. In Artificial Intelligence as well as in cognitive sciences in general, the focus has been on the level of representation and algorithms with little consideration of the biological level below and most concepts have been argued from computer metaphors.

The basis of Newell's theory is the 'physical symbol hypothesis' which Newell has stated as follows:

> "The necessary and sufficient condition for a physical system to exhibit general intelligent action is that it be a physical symbol system" (Newell 1980).

For computer processing, Newell operates with four levels of representation. In-between the program or symbol level and the device level he finds a level of 'register transfer.' Newell (1982). He also developed with respect to computer systems a theory of a level of representation above the program and symbol level. He formulated the 'knowledge level hypothesis' which he extended also to human reasoning:

> "There exist a distinct computer system level, lying immediately above the symbol level, which is characterized by knowledge as the medium and the principle of rationality as the law of behavior." (Newell, 1982, p.99).

The concept of rationality in Newell's sense reflects the hypothesis that humans engage in logically correct reasoning when deciding what to do. Again, this reflects the basic AI effort to develop artificial reasoning systems which largely has been focused on tasks governed by sets of formal rules, such as games, cryptograms, and mathematics. This also explains why the existence of alternative modes of reasoning and the need for subjective process criteria for choice among such modes are not given any emphasis by Newell. He does not take into account the findings of research within decision theory and social judgment demonstrating that people are irrational and inconsistent (in Newell's sense) in an actual task context.

This concept of rationality makes his theory particularly interesting in the present context of a top-down modeling approach to adaptive systems. Newell formulates a principle of rationality:

> "If an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select that action."

That is, the behavior of an agent is explained by intentional reasoning. The principle appears to be self-evident or circular but, in the present context, the interesting point is that behavior can be predicted by analysis of the work environment under the simple assumption that the relevant goal is known and that the agent has adequate knowledge. In fact, from this point of view, it is possible to predict behavior without any consideration of the processes at the lower symbol and algorithm and the device level. What Newell introduces is in fact a kind of feedback loop representation of the human agent: Given a goal (set-point) and the necessary resources (knowledge), it can be predicted that the agent will reach the goal, and the route taken can be predicted from an analysis of the task, given the agent is rational (in Newell's sense. We will return to this point below). As usual in the AI business, Newell refers to rather abstract tasks governed by formal rules such as chess and he notices some problems met by the rational analysis. In the example of chess, the knowledge level analysis would imply that someone who knows the rules of chess would play a perfect game, because such a game would logically follow from such knowledge. Newell acknowledges to this objection that a rational analysis is 'a radical approximation' and that prediction from rational analysis is often overridden by consideration from a lower level such as the impossibility of searching the game tree within a finite task time. Actually, most of Newell's efforts have been spent on the symbol and algorithm level and analysis of the knowledge and the device level has only been used to derive further constraint on the processing.

From the point of view of our framework one can wonder why Newell has not considered the fact that many degrees of freedom with respect to human symbolic processing normally meet the requirements of a task. A rational analysis will, at best, only predict the behavior in terms of task elements, that is, *what* to do in the task environment, not *how* to do it, that is, the process to use in terms of symbolic information processing. When several different mental strategies having different resource demands are available for a given task, knowledge about the agent's resource profile and performance criteria is necessary to predict the sequence of mental states and processes. These features, in fact, reflect the characteristics of the lower 'register transfer' and 'device' levels. Therefore, significant differences are to be expected between the modeling used when the aim is to predict human behavior in an actual task situation, compared to the approach when the aim is to develop artificial reasoning systems by matching a rational analysis to the potential of a device level in terms of the program and configuration of a computer.

Despite these problems, Newell's basic approach has some very important features in our context, and a closer look at the development of the model SOAR is relevant.

THE MODEL SOAR

AI approaches to modeling human cognition often are like complex tool boxes and new tools are added ad hoc when the model proves unable to perform a task. Furthermore, most AI models are directed toward simulation of performance in formal and well structured problem spaces. In other words, they are typically complex mechanisms operating in a rather simple problem space while our observations from analysis of actual work performance often indicate a successful performance of simple mechanisms in a complex environment.

A particular line of development therefore is of interest in the present context, being based on a restricted set of cognitive processes for operation in a complex problem representation. The approach follows Simon's edict that the complexity of adaptive behavior generally reflects the complexity of the environment, not of the adaptive mechanism. The start of this approach is the Logical Theorist by Newell, Shaw and Simon (1958). This program was able to prove logical theorems by heuristic search in a space of propositional expressions that is generated by applying rules of deduction to primary axioms. In this case, we are in a well structured, logic domain. Later, the GPS (General Problem Solver) system (Newell and Simon, 1963) was an attempt to extend the basic strategy so as to also be applicable to more than logic problems.

The core of the GPS program was a means-ends analysis applied recursively. In this respect, the approach is very compatible with the concepts of our work taxonomy when the means-ends hierarchy is applied as the context of knowledge-based action planning. GPS compares a well specified present problem state to a well defined goal state and then heuristically looks for an operator that can reduce the distance between these A basic problem found for the program was its single-minded reliance on means-ends analysis which does not capture people's expertise in terms of direct cue-action matching. Newell (1973) introduced the formalisms of 'production systems' into cognitive psychology to represent stimulus-response mechanisms as a general representation of knowledge. This opened the door to an extension of the GPS concepts as they are now found in the SOAR modeling program.

Description of SOAR

The stated aim of the SOAR project is to understand the functionality required to support general intelligence (see Rosenbloom et. al., 1991). In our context, this basis give some fundamental limitations to the applicability of the results. For instance, to 'understand general intelligence' and to demonstrate the feasibility of designing artificial reasoning systems, it may suffice to incorporate just one human mode of learning (in SOAR, learning by 'chunking') while models serving an understanding of actual work performance will have to consider explicitly all of the essential modes of human learning.

The SOAR approach is based on a multi-level representation of human cognition, each level providing a description at some level of abstraction. Newell's preference for quantitative features leads to a definition in terms of the time frames of the various levels, not primarily the various cognitive modes of control. The 'cognitive band' includes the four levels between 10 msec. and 10 sec. The lowest cognitive level - at 10 msec. - is the symbol accessing level, where knowledge referred to by symbols is retrieved. The second cognitive level - at 100 msec. - is the level at which elementary deliberate operations occur; that is, the level where the most elementary choices are made. The third and fourth cognitive levels - at 1 sec. and 10 sec. are the simple-operator-composition and goal-attainment levels. Above these cognitive levels is the rational band, at which the system can be described as being goal-oriented, knowledge-based, and strongly adaptive (Note, that adaptation is only given emphasis at this level, a fact related to the limited representation of learning functions in terms of chunking, not including trial-and-error learning during interaction). Below the cognitive band, the neural band is found.

The development of the SOAR program is focused on the cognitive band, as opposed to representation at the neural and the rational levels which are considered the domains of the connectionist research respective the logicist and expert-system efforts. Understanding the cognitive band is seen as being effective to constrain models at the neural and the rational bands. For instance, the rational-band models need the heuristic adequacy provided by the cognitive band to become computationally feasible.

The modeling is based on 'experimental data, theoretical justifications, and comparative studies in both artificial intelligence and cognitive psychology' with a preference for quantitative relationships: 'Human experiments provide data about performance universals and limitations that may reflect the structure of the architecture. For example, the ubiquitous power law of practice - the time to perform a task is a power function of the times the task has been performed - was used to generate the model of human practice.'

One basic assumption is that the architecture should consist of a small set of orthogonal mechanisms and all intelligent behavior should involve all, or nearly all, of these mechanisms. This assumption 'biases the development of SOAR in the direction of uniformity and simplicity, away from modularity and tool kit approaches.'

Structure of SOAR

SOAR is structured in three levels: memory, decisions, and goals. At the memory level, a variety of types of knowledge are stored, including declarative, procedural, and episodic knowledge. This knowledge is found in a global memory store. he result of memory access is the retrieval of information into a working memory consisting of an interrelated set of objects with attribute-value pairs. Objects are related by using the identifiers of some objects as attributes and values of other objects. A special type of memory structure, preference, is used to encode control knowledge about the acceptability and desirability of action according to a fixed semantics of preference types. SOAR's productions are neither operators nor implications in the usual sense, but performs parallel memory retrieval. The right hand side of a rule represents a long-term datum, the left hand side represent the situations in which it is appropriate to retrieve the datum into the working memory. All control in SOAR is performed at the decision level, all productions execute in parallel. This results is retrieval into the working memory of all the accessible knowledge that is relevant to the current situation. After retrieval quiescence has occurred, the decision procedure selects one of the retrieved actions based on the preferences that were retrieved into the working memory and their fixed semantics.

In addition to making decisions, intelligence implies the ability to direct behavior toward some end. This function is served by the third level of the SOAR structure. Goals are set whenever, a decision cannot be made by the level below, that is, when the decision procedure reaches an impasse. Impasses occur when there are no alternatives that can be selected or when there are multiple alternatives that can be selected, but insufficient discriminating preferences exist to allow a choice to be made among them. All symbolic, goal oriented operations are formulated in problem spaces, consisting of a set of states and a set of operators. The states represent situations, and the operators represent actions which when applied to states yield other states.

Given a goal, a problem space should be selected in which the goal achievement can be pursued. Then an initial state should be selected that represents the initial situation followed by the selection of an operator to be applied to the initial state. Another state is then selected (that is, normally the state resulting from the previous operation) and this process continues until a state is met, matching the goal state.

Each problem solving decision, that is, the selection of a problem space, a state, and operators, is based on knowledge accessible in the global production memory.

All learning by the SOAR program occurs by the acquisition of chunks, that is, productions that summarize the problem solving that occurs when a problem has been decomposed into sub-goals before processing. The actions of a chunk represent the knowledge generated as the result of the subgoal processing. The primary effect of chunking is thus to allow relevant similar future decisions to be based on direct retrieval of knowledge from memory rather than on problem solving within a subgoal.

Conclusion

The structure of the SOAR model is very interesting in our context considering the separate formation of a global memory bank, a situation related working memory bank, and a separate set of general cognitive mechanisms which add up to a structure very similar to the structure of our framework for work analysis. However, the SOAR modeling concepts also pose several problems in the present search for a framework for modeling adaptive systems:

Production rules are never deleted. This appears to be a very distinct difference from most adaptive systems since the decay of rules seldom used appears to be a problem in systems when performance during infrequent situations is an important qualification.

Furthermore, the model is limited to the 'cognitive' level of 'immediate behavior' in the time span of 10 msec. to 1 sec., that is, to the rule based level. Learning is represented only by the function of 'chunking' production rules from analysis of subgoals when lessons learned from new problems are stored. in consequence, discovering cue-action correlates, that is, new heuristic rules, by observation of the environment when behavior is internally planned by other means is not included as learning mode. Nor is the discovery of new heuristics by trial and error processes operating dynamically on the environment.

In addition, perception and motor control are conceived as separate information input- and out-put functions. Considering the perception and motor control as being functions separate from the cognitive control eliminates the role of the dynamic mental model as an active representation of temporal and spatial properties of the environment. Consequently, direct perception in the Gibsonian way is not possible, tacit knowledge is replaced by scene analysis. due to this limitation, SOAR is ineffective for dynamic activities linked to the dynamics of the environment. Lindsay (1991) has voiced this critique through the years. He argues that specialized representations are required for specialized tasks and, in addition, some thinking depends on 'mental simulation' on constructing and 'running' models of the environment.

To conclude, even when the basic concepts of the GPS-SOAR models are promising, the actual SOAR approach does not offer a useful avenue for development of models of adaptive systems as found under actual work conditions.

COGNITIVE SCIENCE PERSPECTIVE

Several cognitive science approaches are focused on the interaction between different levels of representation and are of interest in the present context.

Marr's Study of Vision

A basic feature of Marr's (1982) approach is the representation of vision systems at three distinct levels. The top level is considered to represent the 'abstract computational theory' of the system of vision. At this level, the performance of vision is characterized as

> "a mapping of one kind of information to another, the abstract properties of this mapping are defined precisely, and its appropriateness and adequacy for the task at hand are demonstrated."

In other words, the *functions* of the vision system is characterized with reference to the goal at hand. At the center level are represented

> "the representation for the input and output and the algorithm to be used to link one to the other".

That is, the *process* of the vision system is represented. The bottom level represents

> "the details of how the algorithm and representation are realized physically the detailed computer architecture, so to speak."

This level is similar to our *physical configuration* level. Marr notes:

> "The three levels are coupled, but only loosely. The choice of an algorithm is influenced, for example, by what it has to do and by the hardware in which it must run. But there is a wide choice available at each level, and the explication of each level involves issues that are rather independent of the other two."

A close analogy the our means-ends level representation is evident.

Furthermore, as we have argued elsewhere and as already Rubin noted (see below), Marr stresses the effectiveness of a top-down explanation within this means-ends hierarchy:

> "--an algorithm is likely to be understood more readily be understanding the nature of the problem to be solved than by examining the mechanism (and the hardware) in which it is embodied."

This conclusion is similar to our view that human performance is more readily understood by studying the nature of the problem people are trying to solve in a work context than from considering psychological mechanisms identified in the laboratory.

Marr's motivation to his analysis of the three levels is to promote cognitive science. He finds that within vision research,

> "for far too long, a heuristic program for carrying out some task was held to be a theory of the task and the distinction between what a program did and how it did it was not taken seriously."

The advantage of Marr's approach in the present context is that he studies a cognitive function, vision, in its function of controlling the interaction with the environment, not in an artificial laboratory task.

Pylyshyn's Approach

Also Pylyshyn distinguishes between three levels: The semantic, the symbolic, and the biological levels. Within the symbolic level, he distinguishes between mental algorithms and the functional architecture (Pylyshyn, 1984). The algorithm level is an abstract representation of the steps the system must go through to perform a task. The level below, represents the functional architecture that actually implements the steps of the algorithm, that is

> "it includes the basic operations provided by the biological substrate, say, for storing and retrieving symbols, comparing them, treating them differently as a function of how they are stored, (hence, as a function of whether they represents beliefs or goals), and so on, as well as such basic constraints of the system, as a limited memory."

It is clear from this, that Pylyshyn's level of functional architecture is expressed in terms of elementary psychological processes and constraints at a level above the biological processes. It is Pylyshyn's assertion that a particular, distinguished algorithm level exists where the

> "interpretation of symbols is in the intentional, or cognitive, domain or in the domain of the objects of thought" (Pylyshyn, 1984).

The examples, Pylyshyn uses to illustrate the characteristics of the various levels are based on syllogistic reasoning, that is, they are entirely abstract. His examples, therefore, do not involve *action* on part of the person and the complex dynamic representations at the level of skilled movements are not activated. This may be the reason for the insistence that the algorithmic representation level 'actually exists.' Pylyshyn's interpretation of the algorithmic level seems to be the specific programming language that control behavior. Pylyshyn uses the principle of 'cognitive impenetrability' to distinguish between the content of the level of functional architecture and the algorithmic level: "The operations at the functional level are not affected by the organism's goals and beliefs." That is, only the symbolic level should be influenced by the semantic content of our knowledge. Once more, this distinction seems to be colored by the academic focus on abstract reasoning tasks. If highly skilled (symbolic) object manipulation, performance depends on pattern matching based on parallel processing of the neural net, then the algorithmic representation is not the only realistic model.

However, Pylyshyn's two levels of algorithms and functional architecture are the kind of representation which is in focus of most research within cognitive psychology.

McClelland and Rummelhart's Constructionist Models

An exception from this tradition was introduced by McClelland and Rummelhart (1986) in terms of their PDP models, that is representation of cognitive processes by parallel, distributed processing by which the model is embedded in the structure of the neural net. The neural net models are not intended to be models of the biological hardware implementation, but rather functional models at the level of Pylyshyn's functional architecture. In contrast to the usual cognitive science position, McClelland and Rummelhart argue that the algorithm level is not a real level. Instead, the emergent properties of the PDP functional architecture approximate the production rules that cognitive scientists propose for the algorithmic level. This view is very compatible with our position that the basic function of the neural net is a dynamic simulation of the interaction of the body with objects (and symbols) in a dynamic environment based on a distributed processing function that Lashley (1942) proposed to be based on interference patterns among propagating excitation waves. The point being that a set of production rules can be a very valid algorithmic representation of the continuous behavior of the neural system derived by decomposition and discretization in the same way as a causal representation is used for the continuous behavior of the dynamic physical environment. Propositional, verbal representations can very well be the conscious reports of continuous changes of the states of the mind as verbal reports describe the observations of the behavior of the environment.

McClelland and Rummelhart compare the algorithmic and their PDP levels of representation to the distinction between Newtonian and Quantum physics. In this sense, the algorithmic, symbol processing level is a macroscopic account whereas the PDP level is a microscopic account. The advantage of the Newtonian representation is that it is much simpler to use for prediction in many cases because it describes the world in terms of entire objects and their interaction without much concern about their internal structure. Both levels are representations of the same real world but are focused on different levels of decomposition. In many cases, decomposition beyond a certain level requires a fundamental change of representation and the discussion which level is 'real' is without meaning. McClelland and Rummelhart in a sense consider the macro theory to be an approximation of the micro theory:

> "Through a thorough understanding of the relationship between the Newtonian mechanics and quantum theory we can understand that the macroscopic level of description may be only an approximation to the more microscopic theory. Moreover, in physics, we understand just when the macro theory will fail and the micro theory must be invoked. We understand the macro theory as a useful formal tool by virtue of its relationship to the micro theory. In this sense the objects of the macro theory can be viewed as emerging from interactions of the particles described at the micro level."

This discussion of the approximate nature of higher level representations again focus the question on the problem whether the 'real thing' has been represented. Actually, however, different levels of description are representing different relationships for different purposes, and one representation cannot be considered an approximation of a lower level account and representations at two levels can never be identical. *All levels are equally 'real.'*

Anderson's Study of Adaptation in Cognition

Considering the present focus on adaptive systems, John Anderson's (1990) recent research is an important contribution. It is, however, important to consider the bias from his particular research context. His book on ATC* was motivated by the learning theories that account for the acquisition of procedural knowledge and the possible identification of a neurally plausible basis for an implementation of the theory. The context of his theories was set by his analysis of knowledge acquisition within well defined domains such as high-school mathematics and college-level programming courses. Both of these domains are characterized by being abstract in the sense that the manual sensori-motor skill is decoupled from the 'deep structure' of the problems considered in the task. The patterns of movements required for manipulation of the symbols map onto the syntax of the representation, not onto the semantics of the problem. This bias shines through in the conclusions of Anderson in a very substantial way when comparing with the results of our analysis of actual work performance.

Another basic feature of Anderson's approach is his emphasis on the identification of the properties of human cognitive mechanisms as they actually exist, not on the development of models of the cognitive system which supports prediction of human behavior in a complex environment. His theory reflects the general bottom-up attitude of cognitive science that, given the individual basic processes of the cognitive system are identified, the behavior in a given task can be predicted. His approach to adaptation is very different from ours. He is considering the evolutionary adaptation of the cognitive functions and processes to the requirements of the environment in order to identify their basic characteristics and constraints without having to infer such characteristics from behavioral data, whereas we are considering the short term adaptation of behavior to the immediate behavior shaping constraints of the environment, in order to be able to predict human behavior in a kind of 'rational' top-down fashion ('rational' in our sense of matching sensible process criteria).

In the introduction it is argued that conceptualization in terms of causal models is shaped by the intuitive bias of the analyst. This is demonstrated by Anderson's comments to Marr's descriptions of the levels of representation. Anderson finds Marr's terminology confusing:

> "His (Marr's) level of computation theory is not really about computation but rather about the goals of the computation." -- "The issue here (at the computation level) is not how it should be done, but what should be done."

He does not seem to realize that the term *'function'* can be used to specify the effect of a lower *process* level. He concludes that Marr's levels are 'fine for vision,' but that the representation and algorithm level is the fundamental level for study of human cognition. The problem with Anderson's theory in our context is, that it is separated from the actual interaction with a physical world through the cognitive control of movements, not that 'the physical base of cognition is still unclear.'

He also has a backward perception of McClelland and Rummelhart's concepts. He accepts that they

> "must be right that the brain only approximates the symbolic rules we ascribe to it. However, the interesting question is, "under what interpretation?". A computer only approximates the program it is implementing-there are failures of memory, interrupt device processes, overhead operating systems, small surges of voltages, and so on. The Rummelhart and McClelland enterprise is based on the belief that the brain approximation of the algorithm is neither good nor faithful."

This passage seems to indicate that Anderson defines the algorithm level as the 'real thing' and the brain processes as being approximations. That is not what is implied in the statement of McClelland and Rummelhart about algorithms being emergent features of the PDP processes. (Rather, the concepts at the algorithm level are approximations of world and brain states?). He continues by a

> "bold assertion: When all is said and done and we know the truth about what is happening in the human brain, there will turn out to be only two levels of analysis that are psychologically real (i.e., in the brain). They are the algorithm level and the biological level. Marr had it right, and, despite the practice in cognitive psychology, there is no intermediate implementation level except as an approximation useful in calculation. That is to say, Rummelhart and McClelland have it just wrong it is not the algorithm level that has the status of Newtonian physics; it is the implementation level."

(Findings from neuro-biology could be interpreted the other way round and support the PDP position, cf. Lashley's (1942) conception of brain activities in terms of interference patterns from propagating excitation waves).

> --"There is no convincing evidence or argument that there is any real level between the biological and the algorithm level."

Compared to our position, we are facing an epistemological problem. We do not consider the levels of a multi-level representation to be more or less real, but the different levels cannot be approximations but necessary conceptual transformations to match brain processes and states to mental concepts and onto human goals and interpretations of physical processes and states in the environment.

Anderson's particular contribution to cognitive psychology is his emphasis on the fact that humans are effectively adapted to the environment by their long evolutionary past and that 'a rational analysis' of the environment can be used to identify features of the cognitive system assuming optimal adaptation. Anderson's concept of 'rationality' is different from Newell's, not being tied to logically correct reasoning but more like the economists' theory of the 'rational man.' Being rational in this sense only implies human behavior being optimal with respect to achieving the agent's goals. His basic assumption is that

> "human cognition is likely to be one of the aspects of the human species that is most completely optimized and optimized in a clean, simple way so that it will yield to scientific analysis."

He formulates a

> "General principle of rationality: The cognitive system operates at all times to optimize the adaptation of the behavior of the organism."

He realizes that the adoption of the concept of rationality is a controversial position, mainly because of the ambiguity of the concept of being rational and he stresses the difference between the concept as used by economists and by Newell. He does not, however, mention the possible dependence of the reference for judging 'rationality' upon the resource demands of the task and the resource profile of the agent. This is a direct consequence of his focus on the long term evolutionary adaptation of the cognitive machinery to the general nature of the environment, not the short term adaptation of behavior to its immediate behavior shaping constraints. He therefore exclusively consider optimality with respect to general criteria, for instance with respect to the properties of memory functions, optimal performance is defined in terms of a rational stop-rule for retrieval search: terminate search when value of items are less than cost of search.

He discusses at length the critical comments to be expected from decision theorists and social judgment theorists who find people to be irrational and inconsistent and argues that this position is caused by the normative reference criteria of such research. He does not, however, elaborate on the difference between the mechanisms as being generally optimal, and the choice of behavior being optimal in the short term. One can wonder why Anderson does not draw the full consequence of the adaptive point of view and, therefore, consider the use of 'a rational analysis' to predict the actual behavior from an identification of the immediate behavior shaping constraints and optimality with reference to short term product and process criteria. This can be a consequence of his bias from the use of abstract, formal task spaces and his focus on the identification of properties of separate cognitive mechanisms.

Anderson's rational analysis involves the following steps:

> "1. Precisely specify what are the goals of the cognitive system."
>
> "2. Develop a formal model of the environment to which the system is adapted (almost certainly less structured than the standard experimental situation)."
>
> "3. Make minimal assumptions about computational limitations. This is where one specifies the constraints of the evolutionary history. To the extent that these assumptions are minimal, the analysis is powerful."
>
> "4. Derive the optimal behavior function given items 1 through 3."
>
> "5. Examine the empirical literature to see if the predictions of the behavioral function are confirmed."
>
> "6. If the predictions are off, iterate."

From here, Anderson examines a set of separate cognitive functions, that is, memory, categorization, causal inference, and problem solving. In all, for our purpose of predicting human responses to system concepts, Anderson's approach appears to be somewhat circular and to reflect his bias from the study of performance in formal task situations.

He decomposes complex human behavior into separate cognitive functions that reflect the usual academic research categories, his 'formal model of the environment' is limited to some short, general reflections about the environment from which he suggests formal models of the cognitive functions. Given his traditional decomposition, these functions match academic research categories well. When prediction fails, he iterates which implies a kind of empirical parameter fitting to the experimental evidence. Because his is focusing on the long term evolutionary adaptation of cognitive functions to the general features of the environment he defines optimal performance for separate cognitive functions in a very general way such as, for memory functions, that optimal retrieval requires the retrieval search to be terminated when value of items is less than cost of search. In this way, his approach is largely independent on the properties of mechanisms such as computational cost, simply because the mechanism is supposed to have adapted to optimality criteria such as the one cited. Because he does not consider the short term behavioral adaptation to immediate environmental constraints, he does not discuss the optimal adaptation involved in matching for a proper choice the multidimensional resource requirement of alternative mental strategies to the multidimensional product as well as process criteria of the organism. Therefore, he is able to derive general characteristics of 'optimal' cognitive functions which can be tested separately against experimental evidence. As Anderson states:

> "Because the rational level offers a different cut at cognition (rather than a higher level of abstraction), it allows one to pursue issues of the temporal and reliability properties of the human cognition in a way that is free from problems of approximation and identifiability."

His approach makes it possible for him to identify 'the optimal input-output mapping' of cognitive functions with 'minimal assumptions' about the limitations of the implementation level. He thus seems to circumvent the identifiability problem. However, he does not discuss how he will avoid the identifiability problem when he, for prediction of actual performance, has to match resource requirements of various mental strategies to capacity limitations of the various mechanisms.

He is very much concerned about the 'identifiability problem' at the implementation level:

> "Basically, what we are trying to induce is the function that maps input to output. We choose to specify this function as a set of mechanisms, but this should not obscure the fact that these mechanisms compute an input-output function, and it is this function we can empirically test. Said in another way, if two different sets of mechanisms compute the same input-output function, there is no way to discriminate among them."

He does not consider he possibility of testing the way in which the functions fail when the mechanisms become stressed to their limits (Simon: only when adaptation breaks down, do the internal mechanisms reveal themselves). In addition, Anderson tests his prediction by comparing successful performance to empirical data, not to the limits, because he wants to avoid the implementation issue. Maybe he plans to take up this issue when he, as he announces, will look at the implementation of his approach for the ACT framework.

In other words, Anderson's approach is a very interesting test of the hypothesis, that evolution has granted humans a cognitive systems which is optimal in some sense, but his model does not define the optimality criteria which are effective in actual, complex situations outside the psychological laboratory.

DeKleer and Brown's Qualitative Physics

An interesting, albeit indirect, demonstration of the importance of means-end relations for functional reasoning is the difficulties met by AI attempts to model the function of mechanical devices 'bottom-up' from the function of the components. De Kleer and Brown find that determining the function of a device like an electric buzzer solely from its structure and the behavior of the parts require complex reasoning. The inference model proposed is based on an examination of the propagation of events through the structure. In an earlier presentation, a basic principle was the 'no-function-in-structure' assumption (Brown et al. 1981). In a later discussion (De Kleer and Brown, 1983, 86), however, inference is guided by 'class-wide' assumptions and 'functional evidence' which in fact appear to be a representation of purpose in disguise. The resulting inference process appear to be very artificial, compared to the top-down inference process guided by functional considerations such as those described by Rubin. In the De Kleer-Brown model, it will be difficult to see the wood for trees, while Rubin's description (see below) appears to be guided by a birds-eye perspective.

Qualitative Simulation:

Qualitative simulation has recently been explored by several research groups as a method of causal modeling. Two main streams of applications are found: 1) simulation of human reasoning and 2) qualitative, that is, causal representations of physical processes for simulation. In the present context, qualitative simulation is promising because these two developments offer tools for a compatible simulation of work environments and the cognitive processes of the actors interacting with this environment. The qualitative simulation has largely been developed for simple systems and problems such as those studied by De Kleer and Brown, but attempts to use the approach for complex systems have been made by Govindaraj and by Schryver (1992). The most comprehensive theoretical framework so far has been developed by De Kleer and Brown (1983, 86) in terms of a calculus for qualitative variables and a method for satisfying constraints. De Kleer and Brown emphasize the basic advantage of causal modeling:

> "Causality as a theory of how devices function provides many advantages. Because it is a theory of how a device achieves its behavior rather than just what its behavior is, -it is now possible to ask what functional changes result from hypothetical structural changes."

This argument is analogous to the argument that causal models are tied to objects and events and, therefore, maps physical changes very easily. This advantage is not necessarily tied to De Kleer and Brown's bottom-up approach. Their models are based on their concept of "mythical causality" which is a method for describing the time ordered trajectories of non-equilibrium states a system passes through in response to a disturbance. This definition basically matches the definition of an event as being the change of state of an object. However, they do not appear to take the full consequence of the causal approach, since they base their analysis on an equilibrium represented by the state-space equations among flow and balance variables and events by incremental changes of variables, not by qualitative changes of the states of objects. In this way, their approach appears to be an awkward way of simplifying the usual quantitative simulation model based on conservation laws. This approach, therefore, appears not to be adequate, neither for simulation of human reasoning, which is not, according to our findings, based on propagation of changes of variables, nor for simulation of work systems. If work systems are based on physical laws, quantitative simulation is preferable for stability reasons, if they are not, events cannot be defined by changes of the state of variables.

Iwasaki and Simon, (1986) have also criticized the approach. However, also their approach is based on a rigorous representation by a set of equations, that is, they display a 'category mistake' by mixing relational and causal modeling concepts.

EUROPEAN DENK-PSYCHOLOGY PERSPECTIVE

Problem solving includes an exploration of the means-ends structure of the problem space in order to identify available alternative means-ends relations between the levels of the hierarchy. The selection of the applicable relations requires an analysis of the compatibility once the chosen element is embedded in the network. When a model has been established of the hypothetical solution, mental simulation of the performance can be used for test of the feasibility of the solution.

The role of a multilevel abstraction hierarchy in problem solving is most explicitly seen in Duncker's (1943) research on practical problem solving related to physical, causal systems (radioactive tumor treatment and functioning of a temperature-compensated pendulum). Based on verbal protocols, Duncker describes how subjects go from the problem to a solution by a sequence of consideration where the items proposed can be characterized by a "functional value" feature pointing upwards to the problem, and a "by means of which" feature pointing downwards to the implementation of a solution; the relation to the means-end hierarchy is clear.

Yet another observation on the role of an abstraction hierarchy on understanding a mechanical device has been reported by Rubin (1920), who reports an analysis of his own efforts to understand the function of a mechanical shutter of a photographic camera. He finds that consideration of purpose or reason plays a major role in the course of arguments: he conceived all the elements of the shutter in the light of their function in the whole. He did not perceive the task to explain how the individual parts worked, but rather what their functions were in the whole. How they worked was immediately clear when their function was known. He mentions that he finds it an analytical task to identify the function of parts, the direction of thought being from overall purpose to the individual function (top-down considerations). The hypothesis necessary to control the direction is then readily available. This approach was found to have additional advantages: solutions of sub-problems have their place in the whole picture, and it is immediately possible to judge whether a solution is correct or not. In contrast, arguing from the parts to the "way they work" is much more difficult as a result of being a synthesis. Solutions of sub-problems must be remembered in isolation and their correctness is not immediately apparent.

PIAGET ON ADAPTATION

Piaget's research may a particularly interesting psychological approach in the present context. For a very readable review, see Boden (1980) on which he present account is based.

Piaget was basically not a psychologist, but a biologist becoming interested in the epistemiological questions raised when he studied biological aspects of adaptation. His basic concern was the adaptation of living organisms to their environment. As a special case, he became interested on the role of human knowledge and, in particular the evolution of knowledge, such as logic and mathematics, that transcends the limitations contingent upon the spatio-temporal embodiment of the organism. He actually, in his early writings described adaptation in terms of self-regulating structures depending on the equilibratory processes of assimilation and accommodation.

Piaget employed the different approaches to the study of knowledge which are characteristic of biology (how does knowledge contribute to the adaptation of the organisms to the environment?), epistemology (how is knowledge possible, and how does it contribute to our view of reality?), and psychology (what sorts of knowledge do humans have and how do they develop from birth to maturity?).

Piaget argues a 'dialectic approach' in which behaviorism and gestaltism are synthesized into Piaget's developmental theories (compare the efforts to synthesize behaviorism and cognitive psychology within the SRK framework), in the same way he tries to synthesize Lamarckism and Darwinism into 'epigenetic assimilation.' Philosophically, he tries to synthesize empirism and rationalism into his genetic epistemology. At the general level he finds the thesis to represent genesis without structure, the antithesis structure without genesis, while his synthesis offers genesis with structure in the terms of 'self-regulated development of increasingly equilibrated structures.' (Boden, page 7).

He distinguish between two opposite adaptive mechanisms: Assimilation, that is, the modification of incoming stimuli or input information by the activity of a pre-existing structure, and accommodation, that is, the active modification of the structure itself as to adapt to the input. Equilibrium, another central concept, is a rather stable, but dynamic state of some structure, such as it can accept and adapt to varied input without any essential change. One problem is that Piaget uses these terms in a very abstract way to cover developing structures of all kinds, such as accommodation of the pupil to light, the assimilation of food by the digestive system and the homeostatic control of blood temperature. Piaget widened such notions of biological control to include psychological phenomena, at all stages of intellectual growth (the baby that assimilates the teddy bear's ear to her sucking patterns while her mouth accommodates to nipple and teat to suck milk).

For Piaget, learning is not so much discovery as construction, a dialectic process of building structures through interaction with the environment. Piaget's three key ideas in the notion of structure are wholeness, transformation, and self-regulation. A structure is a unified, dynamic whole whose parts can be identified only within the mutual relationship. Structural changes are orderly transformations and only through active transformations is knowledge possible. In order to know objects, a subject must act upon them, and thereby transform them: he must displace, connect, combine, take apart and reassemble them. Structures are self-regulating in that their whole is conserved through adaptive compensatory transformations among the parts.

In order to describe the dynamic process of structural transformations, he had available only the formalisms of algebra and logic (Boden, p. 14) and he had to deform them, since algebra and logic are not, in contrast to some recent cybernetically based, computational formalisms, well suited to the representation of dynamic self-regulating systems. "

In conclusion, since Piaget's theory is a conceptual framework at a very general level and "lacks specification of detailed procedural mechanisms competent to generate the phenomena it describes." (Boden, p. 14), it does not add new insight with respect to the present modeling problem.

PERSPECTIVES OF ORGANIZATION THEORY

The present fast pace of change of technology has caused a change in the paradigm of many researchers within the discipline of organizational theory and management science.

Barley (1988) discusses the approaches to the analysis of technological change and finds that investigations of technologically driven social changes typically treat one of three topics: 1. economic issues such as productivity and unemployment, occupational shifts, technical obsolescence; 2. technical change and alienation, de-skilling, and quality of work; 3. technology as a determinant of organizational structure in terms of horizontal and vertical differentiation or centralization and decentralization. Quote:

> "By focusing continually on employment, alienation, and formal organization we may have unwittingly truncated the number of social parameters we hold responsible to technical change. In the existing literature on technology and work, potential registers of social change are nearly exhausted once one considers the division of labor, routinization, distribution of authority and control, levels of staffing and the nature of tasks and skill."

He finds it necessary, in the present situation, to analyze *in detail* the changes caused by technological innovations.

Organizational studies can be divided into ''micro" and macro'' analyses. The former typically includes the study of individual, interpersonal, and inter-group behavior, such as in the study of leadership, motivation, and job design while the latter concerns systemic aspects such as structure, relations with the environment, effects of technology, and so forth. This latter perspective is clearly the focus we are concerned with in the present context.

This work organization perspective conceives of the system of work as a rational system in the sense that it, by and large, is *functional* to the environment by producing a product, providing a service, or whatever, under the specific conditions and constraints characterizing the environment. As an open, rational system the work organization is conceived of as permeated by its environment.

Open and Closed System Models

Two lines of reasoning seem to be underlying most of the literature within management and organizational science. Already Gouldner (1959) distinguished clearly between two fundamental models the 'rational' and the 'natural system' models, which later Thompson (1967) relates to 'closed-system' and 'open-system' strategies of analysis. Much of the literature about organizations has, however, been focused on search for improved efficiency and performance and, consequently, it is based on closed-system assumptions about organizations. That is, the 'rational' model perspective is adopted.

Closed system models. Thompson (1967) brings a discussion of the three major closed system strategies of analysis: 1) Scientific Management (Taylor, 1911) which is focused primarily on manufacturing and similar production activities, and which employs economic efficiency as ultimate criterion and seeks to maximize efficiency by rational planning procedures. Conceptual closure is obtained by assuming that goals are known, tasks are repetitive, output of the production processes will 'somehow disappear' and resources are available in uniform qualities. 2) Administrative Management (Gulick and Urwick, 1937) which is focused on structural relationships among production, personnel, supply, and other service units of the organization. Administrative management achieves closure by assuming that ultimately a master plan is known, against which specialization, departmentalization, and control are determined. 3) Bureaucracy (Weber, 1947), follows similar patterns, focusing on staffing and structure as means of handling clients and disposing of cases. Efficiency is maximized by defining offices according to jurisdiction and place in a hierarchy, by appointing experts to offices, establishing rules for categories of activities and clients and by motivating proper performance of expert officials by providing salaries and patterns of career advancement. Closure of the bureaucratic concept is obtained: the possible influence of policy makers above the bureaucracy is 'set aside'; the complexity of human components is controlled by divorcing private life and office roles by means of rules, salary and career, effects of outsiders, 'clientele', are nullified by depersonalization and categorization.

Open system models. Thompson illustrates two examples of open system strategies of analysis: the informal organization view with attention focused on sentiments, cliques, and social control by informal norms seen as patterned, adaptive responses in problematic situations (Roethlisberger and Dickson, 1939), and the interaction view in which organizations are not autonomous entities but interacting intimately with other organizations and the public: Even the best laid plans of managers have unintended consequences and are conditioned or up-set by other social units (Barnard, 1938; Selznik, 1949; Clark, 1956)

In reality, this dichotomy is artificial, both points of view are necessary together and an elaboration of Barnard's work by Simon, Cyert, and March (Simon 1957a, March and Simon, 1958, Cyert and March, 1963) have produced a newer tradition which evades the dilemma: Organizations must, to cope with complexity, develop processes for searching and learning, as well as for deciding; decisions are satisficing rather than maximizing, and are based on 'bounded rationality' (Simon, 1957b). In his book, Thompson "seeks to extend this newer tradition", but ask why so many people maintain the older, more extreme strategies of analysis and he therefore tries to combine rather than to compromise.

Thompson's Approach. The basis of Thompson's approach is Parson's (1960) suggestion that organizations exhibit three distinct levels of responsibility and control: technical, managerial, and institutional. This distinction is very similar to the stratification in terms of means-ends relations proposed in the previous sections: Thompson argues that every formal organization has sub-organizations whose 'problems' are focused around effective performance of the technical functions - teachers conduct classes, the bureau processes income tax, and handle recalcitrants, workshops process material and supervise operation. The managerial level services the technical sub-organizations. It mediates between them and their customers, pupils, etc., and procures the resources and supplies. The managerial level controls. i.e., administers the technical level. Finally, the institutional level is the source of the 'meaning' of the entire enterprise, it supplies the higher-level support to make the organization's goals possible. Parson's reasoning leads to the expectation that different technical functions or technologies cause significant differences among organizations and, consequently, Thompson stresses the need to include the control of the physical work domain in the organizational model:

> "The technical parts of the system provide a major orientation for the social structure. There are both instrumental and economic reasons but the instrumental question is prior to that of efficiency."

Technical rationality can be evaluated by two criteria: instrumental and economic. Instrumental is concerned whether an action does in fact result in the desired output, the economic whether the most beneficial way is in fact chosen. Present literature on organizations gives considerable attention to the economic criteria,

> "but hides the importance of the instrumental question, which in fact take priority. Complex organizations are built to operate technologies which are impossible or impractical for individuals to operate. The instrumentally perfect technology would produce the desired outcome inevitably. Less perfect technologies needs organizations."

Corresponding to the way in which the means-ends levels in our framework represent qualitatively different system properties, Parson stresses the qualitative break in the simple continuity of 'line' authority because the functions at each level is qualitatively different.

> "Those on the lower levels are not simply lower-order spellings-out of the higher level functions. The articulation of levels and of functions rest on a two-way interaction with each side, withholding its important contribution, in a position to interfere with the function of the other and of the larger organization."

The idea of Thompson's approach is to combine all this: It will be advantageous for an organization subject to the criteria of rationality to remove uncertainty from its technical core. Hence, if both resource-acquisition and output-disposal problems (which are in part controlled by environmental elements) can be removed from the technical core, the logic can be brought closer to closure. Uncertainty would appear to be greatest at the institutional level. At this level, the closed system of logic is inappropriate.

Since technology is an important variable in understanding the actions of complex organizations, a typology of work domains is necessary. Thompson apply only a rather crude distinction:

1. 'Long-linked technology' covers sequentially dependent acts, e.g., assembly line technology for standard products, based on repetitive processes. This is the domain of scientific management.
2. 'Mediating technology' serving a population of clients; insurance companies, banks, etc., requiring operating in ways, i.e., bureaucratic techniques of categorization and impersonal application of rules.
3. 'Intensive Technology' in which a variety of techniques are brought into action on some specific object, hospital, construction industry, military combat teams, etc.

Technical rationality is an abstraction only perfect in closed systems. The closure varies in the three categories, most perfect in the first. Under the norms of rationality, organizations seek to seal off their core technologies from environmental influences. This is done by different means: buffering with input and output components, by smoothing transactions, by forecasting and planning, and when none of these works by rationing. (Note: these are all control theoretic measures.)

The approach to organizational and management models represented by Thompson is compatible with our framework. In order to characterize the adaptive behavior of an organization, Thompson finds the most important characteristics of task environments along dimensions in terms of degree of homogeneity and stability.

During adaptation homogeneity influence organization structure in the following way:

> "Under norms of rationality,
>
> - organizations facing heterogeneous task environments seek to identify homogeneous segments and to establish structural units to deal with each.

> - boundary-spanning components facing homogeneous segments are further subdivided to match surveillance capacity with environmental action."

In the discussion of the influence of environmental stability on the adaptation of an organization to the environment, distinctions are made which have many parallels to our distinction between skill, rule, and knowledge-based control of individual behavior:

> "-The organization component facing a stable task environment will rely on rules to achieve its adaptation to that environment.
>
> -When the range of variation presented by the task environment segment is known, the organization component will treat this as a constraint and adapt by standardizing sets of rules. (This is an empirical categorization routine typically used by bureaucratic organizations: bureaucratic procedures are based on categorizing events and selecting appropriate responses)."
>
> -When the range of task-environment variations is large or unpredictable, the responsive organization component must achieve the necessary adaptation by monitoring that environment and planning responses, and this calls for localized units."

The close relationship with the evolution of work organization as represented in our framework is clear when Thompson describes how the actual, informal organization evolves as a joint result of the strive to cope with a task environment:

> "The basis for grass-root groups, and for successive combinations of groups into clusters, clusters into larger clusters, etc., eventually result in an overall structural pattern for the complex organization. ... To a large extent they [the organizational varieties] can be accounted for as attempts to solve the problems of concerted action under different conditions, especially conditions of technological and environmental constraints and contingencies." (p. 74).

As an example of the evolution of an organization, Thompson mentions the synthetic organization emerging for emergency management after a major accident:

> "In a surprisingly short time and with little of the random, aimless behavior sometimes attributed to disasters, resources designed for other purposes are disengaged from their normal employment and adapted to disaster-recovery activities. ... Initial efforts at disaster recovery occur whenever resources and an obvious need or use for them occur simultaneously. .... In a relatively short time, usually, two things happen... and bring about a synthetic organization: (1) uncommitted resources arrive...and (2) information regarding need for additional resources begins to circulate. When knowledge of needs and resources coincide at a point in space, the headquarters of the synthetic organization has been established. Such headquarters only occasionally emerge around previously designated officers, indicating that their power rests not on authority in any formal sense but on scarce capacity to co-ordinate..... Authority to co-ordinate is attributed the individual or group which by happenstance is at the crossroads of the two kinds of necessary information, resources available and need. ... What is does have, compared with normal organizations, is (1) consensus among participants about the state of affairs to be achieved and (2) great freedom to acquire and deploy resources, since the normal institutions of authority, property and contract are not operating." (p.52)."

This example of an organization evolving by a process controlled by 1. the information available to decision makers and 2. the part of the work domain under their control illustrates the approach suggested for an evolutionary simulation model described in the section below.

Thompson's approach to modeling the mechanisms behind the self-organizing evolution of a work organization appears to be compatible with the structure of our taxonomic framework with respect to the need for representation at several levels and the interaction between top-down propagation of intentionality and bottom-up propagation of material opportunities and constraints.

Pondy and Mitroff's Extensions.

Pondy and Mitroff (1979) have presented a discussion and extension of Thompson's approach which is particularly relevant in the present context. They find that Thompson's models are too narrowly focused on the control perspective: How do organizations change to remain stable confronted with changes in the environment? They therefore adopt Boulding's classification (Boulding, 1968) of systems to identify the necessary extensions of Thompson's concepts.

They review Boulding's categories ranking 9 levels of complexity as follows:

Level 1: Frameworks. Only static structural properties are represented in frameworks, as in descriptions of the human anatomy, etc.

Level 2: Clockworks. Stable dynamic properties are represented in clockwork systems, as in open-loop mechanical automata.

Level 3: Control systems. Control system models describe regulation system behavior according to an externally prescribed target or criterion as in heat-seeking missiles, thermostats, economic cycles in centrally controlled economies, and the physiological process of homeostasis.

Level 4: Open systems. Whereas a control system tends toward the equilibrium target provided to it and therefore produces uniformity, an open system maintains its internal differentiation by exploiting environmental change.

Level 5: Blueprinted growth systems. Systems at this level do not reproduce through a process of duplication, but by producing 'seeds or 'eggs' containing preprogrammed instructions for development, as in the acorn-oak-system, or egg-chicken system, or other 'dual level' systems. Explaining level 5 systems means discovering the generating mechanisms that produce the observed behavior. And models of level 5 systems will exhibit this dual level structure as well.

Level 6: Internal image systems. Level 3, 4, and 5 models incorporate only primitive mechanisms for absorbing and processing information. The essential characteristic of level 6 systems is a detailed awareness of the environment acquired through differentiated information receptors and organized into a knowledge structure or image. Level 6 systems do not exhibit the property of self-consciousness. They do not know that they know. That enters at level 7.

Level 7: Symbol processing systems. Level 7 systems are self-conscious language users, like individual human beings. What is not so obvious is that human groups can be level 7 systems. The best example of what it means for a group to have an image of its environment is the process of the social construction of shared models of reality. That is, a group can be said to be a symbol-processing entity if its members share a common definition of reality.

Level 8: Multi-cephalous systems. These are literally systems with several brains. Boulding's term for this level is 'social organization,' a collection of 'individuals' (any acting unit) acting in concert. What is at issue is that the collection or assemblage of 'individuals,' whether they be genes, humans, human groups, or computers, creates a sense of social order, a shared culture, a history and a future, a value system.

Level 9: To avoid premature closure, Boulding adds a ninth, open level to reflect the possibility that some new level of system complexity not yet imagined might emerge

Pondy and Mitroff concludes their discussion of Boulding's levels:

> "Generalizing from the above conclusion, our worst fears are that the field of organization theory will take its task for the next decade to be the refinement of analysis at levels I through 4. Our greatest hope is that ways will make an effort at moving up one or two levels in our modeling (both conceptual and formal) and begin to look at, for example. phenomena of organizational birth, death, and reproduction, the use of language, the creation of meaning, the development of organizational cultures. and other phenomena associated with the types of complexity in the upper half of the hierarchy.
>
> Specifically, we offer five major reasons in support of this position:
>
> I . By focusing on maintenance of the organization's own internal structure, open-system theory has directed us away from ecological effects— broadly defined— of the organization's actions, to the ultimate detriment of the organization itself.
>
> 2. We should be directing our efforts to understanding massive dysfunctions at the macro level, not just explaining order and congruence. How do organizations go wrong?
>
> 3. We need to reflect in our own model conceptions of people in other fields, especially those that picture persons as having the capacities for self-awareness, for the use of language, for creative growth, and for learning from their experience.
>
> 4. Troublesome theoretical questions ignored by open-system theory are suggested by other models. For example, do organizations reproduce themselves? If so, how?
>
> 5. For the purpose of maintaining organization theory's adaptability as an inquiring system (Churchman, 1971; Mitroff, 1974), we need to discredit what we know, to change for the naked sense of change to prevent ossification of our ideas."

The topics of item 1 o 3 in particular appear to match the requirements to modeling adaptive systems in a turbulent environment. Pondy and Mitroff elaborate on these issues in the following way:

1. The Ecology of Organizational Action

They note, that authors often claim to be using an open-system strategy whereas they are in reality using level 3 control system models. They have failed to make the distinction between 'natural' and 'open' system models. They quote Thompson (1967):

> "Central to the natural-system approach is the concept of homeostasis. or self-stabilization, which spontaneously, or naturally governs the necessary relationships among pans and activities and thereby keeps the system viable in the face of disturbances stemming from the environment" (Thompson 1967, p. 7).

And they comment:

> "In other words, the environment is a source of disturbance to be adapted to, instead of the source of ''information" that makes internal organization possible. Self-stabilization referred to by Thompson is a level 3 process. The equivalent level 4 process is self-organization. What Thompson calls a closed system is equivalent to Boulding's clockwork (level 2). Thompson made a major contribution by formalizing organization theory at a higher level than it had been at. But we argue that it was not at the level of open systems as understood by Boulding and other systems theorists."

With respect to the ecological consequences of control system thinking, Pondy and Mitroff stress that

> "the aim of a control system is to produce uniformity, i.e., to decrease variety, if it can. To the extent that the system environment is highly varied in its texture over time, the regulator part of the system must match the variety of the environment so that it can control that variety and produce a uniform environment for its operator part. This is the essence of Ashby's Law of Requisite Variety (Ashby, 1956). In Thompson's language, this means creating the conditions necessary for rational operation at the technical core by controlling environmental uncertainty. The ecological implication of control system thinking, both theoretical and practical, is that environments as well as organizations will become more uniform."

They then formulate the ecological consequences of open system thinking in this way:

> "The ecological consequences of open system thinking are quite different from those of control system thinking. An open system is at such a level of complexity that it can maintain that complexity only in the presence of throughput from a differentiated environment. If an open system insulates itself from environmental diversity and differentiation, or if it attempts actually to kill environmental diversity, then it will have only a uniform, gray soup to feed on, and eventually its own internal structure will deteriorate to the point that open system properties can no longer be maintained. If control system models are used to manage open systems, the system will be led to take precisely the wrong actions! The organization will attempt to insulate itself from the very diversity that it needs to survive as an open system. Suppose an open system does not attempt to buffer out variability, but exposes itself to the uncertainties of the environment. If environments are plentiful, and the system is agile, it may still extract the needed organizing information from the immediate and present environment, leave it depleted (i.e., undifferentiated) and move on to another."

2. Dysfunctions in Organization Theory

Pondy and Mitroff find one of the striking differences between organizational behavior and organization theory to be that organizational behavior defines much of its research effort in terms of dysfunctions of the system, such as absenteeism, turnover, work dissatisfaction, etc. In contrast, organization theory has been a theory of order, that is the proper match between structure and technology, between environment and structure, between forms of involvement and forms of control, etc.

Even though some of the more spectacular dysfunctions have been documented in case and accident analyses, organization theory offers no systematic typology of dysfunctions at the macro or systemic level. They conclude by arguing that we need to develop a theory of error, pathology, and disequilibrium in organization.

3. Models of human behavior

Pondy and Mitroff find that:

> "A third reason for needing to go beyond the prevailing model of organization is that it excludes many fruitful models of human behavior. Organization theories seem to have forgotten that they are dealing with human organizations. not merely disembodied structures in which individuals play either the role of ''in-place metering devices'' (..) designed to register various abstract organizational properties (e.g., complexity, formalization, etc.), or the role of passive carriers of cultural values and skills."

They refer to Thompson's conception of the individual:

> ". . . if the modern society is to be viable it must sort individuals into occupational categories; equip them with relevant aspirations, beliefs. and standards; and channel them to relevant sectors of 'the labor market.'" (Thompson. 1967, p. 105).

and argue that

> "--following Thompson, many macro organization theorists have downplayed man's higher capacities, including his ability to use language. his awareness of his own awareness, and his capacity to attribute meaning to events, to make sense of things. These capacities are characteristic of Boulding's level 5 through level 8."

Pondy and Mitroff concludes with some important research issues:

> "One important class of theoretical questions addresses the phenomena of organizational birth and reproduction. Extant open system models, for the most part, are about mature organizations. Although Thompson discusses some aspects of growth, his analysis is about continued growth of adult organizations. And it is growth whose patterns are shaped by external forces, not the blueprinted growth of Boulding's level 5. ---- Consider the following model.
>
> 1. The development of organizations is constrained by environmental forces, but it is directed by fundamental rules for organizing which are stored inside the organization itself. Those governing rules, or generative mechanisms, produce the observed patterns of differential functioning that make up the organization.
>
> 2. The organizing rules are stored in the brains of some, perhaps all individuals in the organization. Those rules result from a previous process of negotiating the organizational order. Some organizing rules are also stored on paper (e.g., job descriptions, standard operating procedures), so that the content of the rules may transcend the tenure of any organization member.
>
> 3. When a person leaves the organization, he carries with him those organizing rules. Should he be the founder of a new organization, those rules would find expression through unfolding in a new environment."

These 'governing rules' or 'generative mechanisms' are in fact what has been formulated in our framework as mandatory for any attempt to model work systems as being adaptive organisms. In particular, the focus on interpretation of the environment, on modeling man-in-the-environment, and the need to use organizational failures as a window of evolutionary mechanisms match very well the essence of our framework.

Mitroff and Linstone's Extensions

Mitroff and Linstone (1993) have reviewed organizational decision making with reference to Churchman's (1971) categories of 'inquiring systems.' The topic of the discussion is the problem of decision making in highly integrated and dynamic socio-technical systems. The main thesis of the book is not to use bounded thinking to unbounded problems and stresses the cross disciplinary and complex nature of problems in modern systems. Distinctions are made between several different modes of inquiry:

'Traditional' Thinking:

1. Inductive-concensual, e.g., the Delphi method, consensus by expert judgment, the tails of the distribution are discarded. This approach raises the 'Humean' problem: there is no underlying justification of the empirical 'causality.'

2. Analytic deductive, the rational approach based on models.

'Complex' Thinking:

3. Multiple realities, multiple perspectives are applied, often from different disciplines, interpretation of facts depends on the perspective, that is, the model context adopted. The main purpose of this mode of inquiry is to act, the decision maker in this mode is not interested in truth per se. The model and the related hypothesis are judged by the outcome of the actions.

> "Through action, the decision maker comes to learn of the 'working truth' of the model, where 'truth' is defined by the ability to respond not only to one's immediate problem, but to anticipate future problems, and thus gain an invaluable edge on working on them before they have gotten out of control."

This approach raises the 'Kantian' problem: Which representational structure is relevant for the interpretation of the observations in the world? Kant was, in particular, considering the temporal-spatial frameworks, presently more elaborate bases for representations must be considered. In a generalized framework, this simply implies that the interpretation of observations and communications depends on the model available for interpretation. This relates closely to the concept of 'hypothesis testing' at the knowledge-based level and the related error types.

4. Conflict and dialectical thinking, opposite positions are formulated, for instance by removing judgment at the center of the distribution in Delphi inquiries and the implications are investigated. In the examples discussed, the opposition is between the goals and priorities derived from established practice and a basic analysis of the potential goals as deduced from an analysis of the environment and the system characteristics. This approach is very similar to an analysis within the means-ends analysis of our taxonomy. Resolution is based on a process similar to the interaction between the district attorney and the defense lawyer as found in a court room.

'New' Thinking:

5. Unbounded thinking implies 'multiple perspectives' that are intimately interacting. No science (natural science or physics) or perspective has a priority over the others. This leads to a request for systems thinking which, according to the authors, has been founded by Singer and Churchman. One basic proposition by Singer is that no discipline can offer an adequate set of elementary or simple acts into which a complex situation can be decomposed. Three different perspectives must normally be taken into account: The *technical* (science and technology), the *organizational* (social entities, small to large, formal to informal), and the *personal* (individuation, the self). 'Perspectives' are normally much broader than 'realities' involved in the multiple realities mode. The 'technical perspective' thus can be based on several sets of models and data interpretations, 'realities.' Within a 'reality' and the related model, the system considered is normally broken down to sub-problems, treated separately. In the unbounded mode, the system is normally treated as an integrated whole. The unbounded mode is another approach to solving 'Kant's problem.' Each perspective reveals insight which, in principle, cannot be obtained from any other perspective. The perspectives, therefore, often interact in a dialectic way. The emphasis on the different perspectives is a matter of one's ethical values and judgments.

Some basic characteristics of the different world views, perspectives are:

	Technical	**Organizational**	**Personal**
Goal	Problem solving	Action, stability, process	Power, influence prestige
Mode of inquiry	Data, modeling, analysis	Consensual, adversary	Intuition, learning, experience
Ethical basis	Logic, rationality	Justice, fairness	Individual values, morality
Planning horizon	Far	Intermediate	Short

The interesting feature of Mitroff and Linstone's discussion in our context is the emphasis placed on the need to have an analysis including multiple realities and perspectives which are very analogous to our emphasis on a simultaneous analysis of the propagation of functionality and intentionality through the levels of a means-ends hierarchy. In addition, their request for consideration of the subjective process and situation criteria is closed to our focus on criteria for choosing among action alternatives. The discussion presented by Pondy, Mitroff, and Linstone serves very well to explain the lack of success of 'general systems theory' for modeling adaptive organizations.

CONCLUSION

GENERAL OBSERVATIONS

A number of general statements about the options for modeling adaptive systems can be made from this review of the approaches taken by the various disciplines:

The Bias of Academic Research

All approaches to modeling adaptive systems will be very much shaped by the context of the underlying research, that is, the tasks and problems that has been used as a source of data and the purpose of the modeling effort. This is not always evident in the sources reviewed. Two tendencies are clearly appearing:

One is a tendency to generalize the validity of the models beyond the domains which are data sources. For instance, cognitive science and artificial intelligence sources focused on formal tasks constrained by normative rules, tend to generalize widely to less constrained tasks without analyzing explicitly the differences in the nature of behavior shaping constraint found in the various domains.

Another one is the academic tradition to focus research on a comparative evaluation of research paradigms and the resulting effort of each author to introduce their results by a comparative overview of other modeling approaches. Taken together with the general tendency to generalize without explicitly stating the basis, this tradition often leads to unreliable judgment of 'competing' approaches because the models normally are based on different (implicit) assumptions and, therefore, not directly comparable (see., e.g., the discussion in a previous section of Anderson's (1990) comparison of different modeling approaches).

In our context, we are focusing on models that will support design of information systems for adaptive work organizations. This objective is different from most (all?) academic modeling efforts, and this particular context must be considered. Judgment of the relevance of a particular modeling approach in the present problem oriented context, therefore, must be based on a very careful evaluation of the objective and the actual data base of the various approaches.

The conclusion is that our modeling objective, being focused on the design and evaluation of advanced information systems, is basically different from the objective of academic modeling efforts within the individual disciplines. The solution of the design problem appears to be to consider carefully the different models available from academic sources for use in separate, well-bounded problems and to develop further the taxonomy for work analysis (Rasmussen et. al, 1990) into a framework for integration of the detailed models imported from extern sources.

The Modeling Objective in System Design

As it is discussed in the introduction, it was clear from the outset that we need different models for the design of work systems and for the evaluation of their stability To serve the first objective, object-event oriented, causal models are useful. For the latter purpose, we have been looking for quantitative, relational models capturing the characteristics of closed-loop, adaptive systems.

For this purpose, it was natural to look for solutions from the quantitative, control theoretic models developed within 'general system theory' and 'cybernetics.' Such models, however, do not adequately represent the dynamic propagation of intentionality and objectives in interpretative, linguistic terms and the resulting dynamic reconfiguration of the closed control loops (Pondy and Mitroff). In particular, they are not suited to model system break-down when effective adaptation fails. Therefore, the lack of convincing results from 'general systems theory' seems to be caused by the behavioristic modeling approach based on input-output transformations (Mesarovic).in terms of feedback loops.

In consequence, a wider review of models of adaptive systems and their evolution appeared to be necessary.

Models of Evolutionary Processes

Evolution of organisms from inanimate matter. Approaches based on physics and thermodynamics may be relevant for the early evolution of structured organisms from inanimate matter (Iberall, Sodac and Yates), but they do not capture evolution at the later functional stages because laws of nature do not have memory and do not include linguistic representations (Pattee).

In addition, *biological evolution at the functional level* of an organism in a competitive environment cannot be represented bottom-up, as suggested by physicists. Evolution at this level is, in fact, teleological by nature. Selection for survival of successful species does not involve vitalism nor loan of intelligence as suggested by Dennet, but only sources of change and memory of successful features. Evaluation of success is done by the competitive environment and does not involve self-evaluation or consciousness.

Modeling *evolution of work practice.* On the other hand, evolution of the behavior of an individual and *survival* of *successful patterns of behavior* depend on short-term process criteria and self-evaluation by the individual. Whether this function involves consciousness and intelligence, is a matter of definition. Modeling functional evolution in the latter sense, involves at least two levels of representation, a functional, physical one and an interpretative, linguistic one (Pattee, Pondy and Mitroff). This approach to modeling can be used for the individual and an organization (Pattee, Pondy and Mitroff). The need for representation of the rule based, linguistic mode and causal reasoning in models of adaptive organisms (Pattee) implies that quantitative relational models are not adequate. For fundamental reasons, object/event oriented causal models are also required to capture the human interpretation and reasoning.

In general, the various modeling approaches appear to be very much biased by the rather narrow span of attention of the academic disciplines and, therefore, not directly applicable in the present design context.

Multi-level, Multi-language Models

It is clear from the review that for our purpose we need a multi-level model framework including system representation in quite different representational languages to capture the characteristics of physical processes, abstract functions, human interpretations, and system objectives and values. For the lower levels, quantitative models based on concepts of physics and control theory can be relevant, depending on the nature of the system. For the higher levels, qualitative, causal models are necessary to capture human interpretation and decision making.

In the work systems considered here, the human actors are considered to act as controllers of the state of affairs of the work environment. Also for the representation of the functions of the human controller, a multi-level model ranging from quantitative representations at the sensori-motor level to qualitative, causal models at the conscious reasoning level. Many useful models are available at all levels from cognitive science and artificial intelligence. No one level represents the 'real thing' and, practically speaking, all the modeling paradigms found in the literature are useful for particular problems. It is, however, necessary to go beyond the generalizations and comparative judgments of the individual author to identify the biases resulting from the tasks modeled and the objectives for modeling.

One major difficulty in the present context is, that most AI and cognitive science models are based on a complex 'tool-box' which is modified and augmented when comparison of behavior of the model and humans uncover discrepancies (a basic difficulty in test of causal models (Brehmer et al., 1991). Therefore, present models developed within cognitive science and AI cannot serve to evaluate performance limits and breakdown of adaptation of actual systems.

WHERE TO GO NEXT?

The basic conclusion of the review of modeling approaches is that a multi-level framework is needed to integrate the available models of the separate functions found in a complex work system. Such a framework has been developed by the Risø group through the analysis of work performance in different domains (Rasmussen, Pejtersen and Schmidt, 1990) and its use for the design and evaluation of advanced information systems has been demonstrated with respect to the normal, functional properties of the work systems (Rasmussen and Pejtersen, 1993, Rasmussen, Pejtersen, and Goodstein, in press).

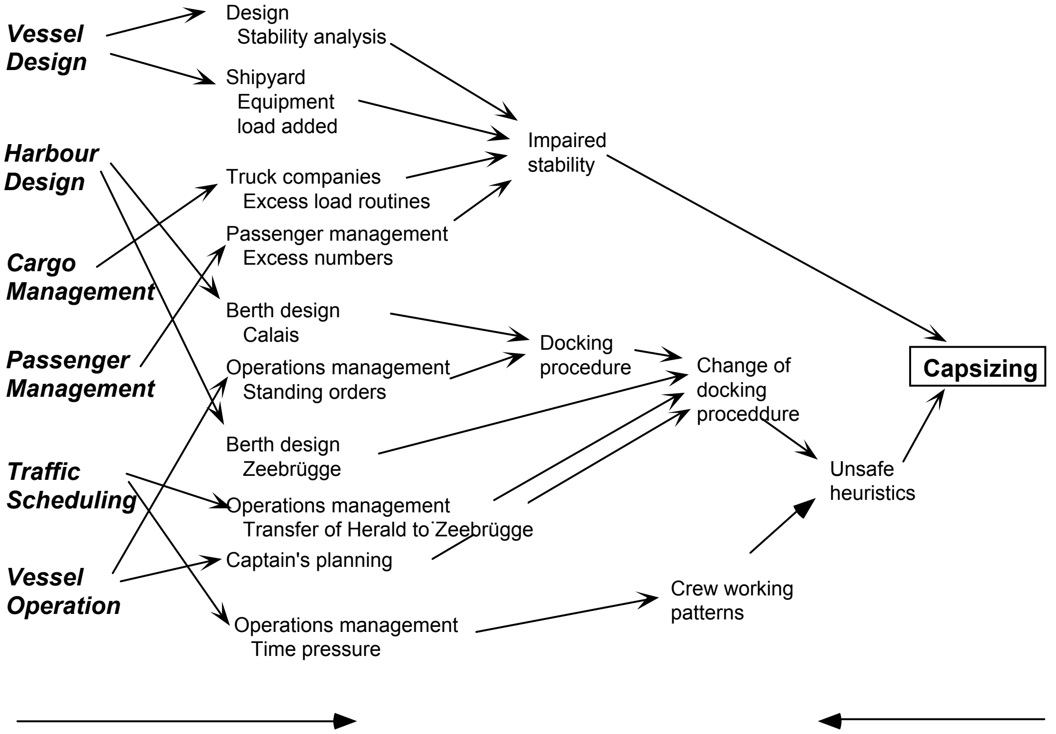
Further development should, however, be considered for the analysis of the properties at the limits of normal performance and evaluation of systems which pose major hazards and for which, therefore, performance during rare, disturbed states is mandatory. In the following sections the developments to consider for analytical, respective empirical evaluation of system properties at the performance boundaries are considered.

Analytical Evaluation of Performance at the Limits

For traditional work systems with stable work procedures, performance limits are identified by a task analysis and an evaluation of the resources required for proper task performance. Performance at the limits is then analyzed by an error-mode-and-effect analysis, that is, by analyzing the effect of the various error types found in an error taxonomy on each of the actions in the task sequence.

This approach is only realistic for stable task sequences, that is, for tasks in which the task sequence is heavily constrained by the task environment, e.g., by the functionality of technical equipment to be controlled. In general, this approach is unrealistic in a modern work system because the number of possible task sequences leads to a combinatorial explosion and it is unlikely that all relevant, low probability - high hazard combinations will be identified.

Instead, an analysis must start by an identification of the hazard sources present in the work system by a kind of morphological analysis. Since all major accidents, by necessity, are caused by the release of energy or hazardous substances, hazard sources can be identified by a proper search in the physical part of the work system. From here, the relevant accident sequences are 'designed.' That is, the processes, necessary to release the hazards and to channel them toward the sensitive targets are identified and the total inventory of components,



Human error analysis:
Failure-mode-and-effect
analysis of the activities of
individual decision makers is
ineffective due to combinatorial
explosion: Very likely, one will not
see the wood for the trees

Hazard Analysis:
Combinatorial
structure of possible
accidents can easily
be identified

Figure 7 illustrates the complex pattern of the Zeebrügge accident. It will be difficult to identify the actual combinatorial structure of the course of events of the accident from an analysis of task activities in a failure-mode-and effect analysis. However, from analysis of the system, the hazard is clearly visible and the necessary constituents of an adaptive breakdown, that is, an accident can readily be identified in the activities of people involved.

equipment, and people is considered together with their potential processes and actions to judge whether the necessary constituents of an accident are present. In this way, the hazardous situations and the relevant accidental courses can be identified for further analysis applying the framework for work analysis with respect to the potential for recovery by the human actors involved.

The difference in visibility of the hazard potential from adaptation under pressure as seen from the perspective of the individual actor's activities and from an analysis of the hazard involved and a search for possible contributors is clearly illustrated by figure 7 representing the causal tree of the Zeebrügge accident.

This kind of analysis of the work system can serve to identify the singularities within the work space which will cause system breakdown if activated by adaptive migration of task performance. In addition, it can serve to focus the more detailed analysis of the visibility of the singularities to actors, of potential defenses against activation and of the nature of the gradients governing migration.

Simulation of Adaptive Systems

The analytical approach discussed in the previous paragraphs will only serve to identify in a qualitative way the location of hazardous limits of adaptation and the presence of potentially threatening, migrating work behavior. The present review has not identified quantitative models that can serve an analytical evaluation of the dynamic behavior of the complex, non-linear systems involved.

One alternative approach that may be worth a more thorough investigation is to extend the 'cybernetic' simulation models applied by Mesarovic and Forrester for their World Model studies by an integration of their use of differential state-space equations for representation of the input-output relations of lower level processes and functions with a representation of the higher level decision making and adaptive choice among alternatives by object/event based causal models. The representation used to describe the *options for choice* by actors will, by nature, be a qualitative, causal model based on prototypical tokens drawn from the categories applied by the staff and labeled by the category names they use. A representation of the *functional interaction* among the means for action chosen (for instance for simulation of a task environment) will normally be based on a causal, qualitative representation when dealing with human work activities and on a quantitative, relational representation when considering the function of technical equipment.

The modern, object-oriented programming languages appear to offer new opportunities for simulation experiments based on multi-level, multi-language models. In particular, since now formal representation of purposive systems in terms of multiple levels of means-ends relations are being developed, such as Lind's Multi-level Flow Model concepts (Lind, 1992, 1993). Representing organizations as closed loop, purpose seeking systems implies development of models of mechanisms able of exploration, evaluation, and choice action alternatives according to certain performance criteria. Solution of this problem involves the identification and representation of 'generative mechanisms' stressed by Pondy and Mitroff.

It is clear from the review, that such an integrated simulation model must include models at several levels, probably formulated in different conceptual languages, and that models representing system behavior at the individual levels are proposed from the various academic disciplines. A difficult modeling problem appears to be that we have to consider the *interaction between models* selectively suited to represent particular perspectives, functions or cognitive modes of control which have been developed from selective studies. Due to the strong influence by the task domain from which the models are derived and by the modeling objective, an integrated model cannot be developed by adding together the separate models found in the literature. In addition, from a theoretical point of view, this difficulty is increased by the non-linear nature of socio-technical systems, that is, the models are not simply 'additive' and changes made in the conceptual structure of the model at one level can drastically change the validity of the models at other levels.

In conclusion, the development of a simulation model for the study of the properties of complex adaptive work systems is a long term project which must be based on a problem oriented research strategy that can ensure that the part-models developed for particular system functions are compatible. In addition, the development is only realistic in a cross-disciplinary research community sharing an integrating over-all strategy.

Development of a Simulation Model

Simulation of adaptive work systems pose some fundamental problems which are not met in the usual cognitive science approach to development of simulation models.

The aim of cognitive simulation models is normally to verify a model of a particular cognitive process. Often, this process is involved in a well bounded activity calling for knowledge-based reasoning (the definition of a cognitive process) within a domain that can be formally described , such as physics (deKleer and Brown), mathematics or calculation (Anderson), cryptograms and puzzles (Shaw and Simon).

Compared such cognitive tasks, simulation of actual work performance raises several basic problems:

1. Modeling expertise: The basic activity takes place in the context of heuristics and expertise. Only occasionally will the behavior raise to the knowledge-based level. Therefore, the model must include mechanisms for formation of heuristics and skilled behavior, that is, a repertoire of cue-action rules that depends on a particular actor's level of expertise and subjective performance criteria. In consequence, the model parameters must be identified that can be varied to match the model to different categories of human actors and different levels of expertise.

2. Interacting control structures: The role of the knowledge-based level is to guide the rule-based level in a supervisory control mode. Similarly, the rule-based level controls the sequence of acts for which the skill-based level exerts the necessary continuous control of the physical movements of the actor. Ideally, therefore, development of a simulation model should begin by a model of the skill-based manipulation of the work environment, not by a model of the cognitive processes involved in formal, knowledge-based problem solving.

One of the basic modeling problems is, that different kinds of models are needed for the different levels and that their dynamic interaction during skill acquisition and rule learning must be faithfully captured by the simulation model. The ultimate model, consequently, should not only represent the cognitive control functions during performance at the three levels during work in a complex environment, but also their interaction during learning and skill acquisition. The model, in this way, will have to represent theories of human behavior from several different psychological paradigms.

3. Evolutionary modeling approach: Development of a computational model of work performance is a very ambitious and long term goal and an evolutionary approach that can be tested by comparison of laboratory experiments with human actors with simulated work scenarios will be necessary. The phases of model development, therefore, have to be carefully chosen from consideration of the normal cognitive control structure and of a series of progressively more complex experimental tasks which can serve to test the model as it grows.

4. Level of representation chosen for a model. A 'computational' simulation model will be formulated in information processing terms, that is, behavior is decomposed into standardized, recurrent elementary information processes. Match to behavioral evidence will be analyzed at the level of strategies, rules, and actions, not on the level of physiological or neurological mechanisms.

One basic consequence of this choice of modeling at the information processing level is that there will be no one-to-one mapping of model processes or functions onto psychological, physiological or neural processes and, consequently, the decomposition of functions into processes to apply in the modeling domain will be different from the decomposition of human behavior applied by psychologists, simply because the 'technology' of the underlying mechanisms is basically different. However, we must, by necessity, seek a consistent mapping even when this will turn out a complex relationship between the processes and parameters of the model and the psychological concepts.

In consequence, it will be without meaning to compare higher level functions belonging to the two domains separately. To ask, for instance, whether a simulation model has consciousness or intelligence is a category mistake. In the modeling domain, selected effects of concepts such as consciousness or intelligence are represented by sets of information processes such as those involved in problem solving, respectively 'self-evaluation' according to individual performance criteria. This does not, however, imply that the model has e.g., 'consciousness' because, as mentioned, there will be no one-to-one mapping between the meta-process controlling the lower level processes of the two domains.

5. Simulation by Computational, Causal Models. A simulation model for the present purpose will include causal models, based on categories of objects (material as well as abstract), states and events represented by prototypical examples.

For simulation it is necessary to replace the classes by particular members, but then the entire exercise will be an ad-hoc demonstration of selected examples. There will be no formal stop rule to terminate the additions of new relations or objects in order to match simulated performance to observed real life performance. Empirical testing of the internal consistency of a theory in causal terms by simulation of scenarios is difficult (see Brehmer et a., 1991).

Simulation should be based on principles that will dynamically generate the activity scenarios. One approach can be from a control theoretic point of view in terms of a self organizing, distributed control system in which the actors adopt control of a sub-space of a loosely coupled work domain defined by the available means-ends relations. The actual organization of actors can then evolve from a specification of the conventions and constraints defined for their co-operation and communication.

Given the individual simulation runs are only trajectories of a particular case, that is, a demonstration of one possible piece of behavior, further experiments are needed to test an object/event based simulation model. One possibility appears to be to vary systematically the parameters of the model in a series of simulation runs followed by a cluster analysis of the resulting trajectory samples to see whether the categories of simulated behavior match the categories observed in the actual work system. This turns out as a kind of prototypical significance analysis of the model, replacing the usual statistical significance test.

In conclusion, the validity of the model depends on the presence of a context by which the prototypes can be shaped. This context, in a complex world, cannot be exhaustively defined but only characterized by a verbal description referring to a generally accepted terminology. This probably supports the use of computer games with widely accepted 'cover stories' as experimental vehicles.

In addition to parameter variation, an approach to model validation is to stress its performance until it fails and then to compare the limiting properties and the 'error' categories to actually observed human behavior.

Use of Computer Games for Model Development

The evolutionary modeling strategy suggested here involves an iterative, progressive modeling sequence comparing simulation experiments with the analysis of performance in experimental scenarios of increasingly complexity.

For such an approach, each experimental session will be designed to selectively identify central behavior shaping features of the work system. The individual experiments will be designed so as to define an constraint envelope, see figure 8, around the experimental subject and the 'controlled variables' of the experiments will be chosen so as to study some higher level abstract, structural feature of the work system, such as the influence on error types, learning speed, resource requirements, etc.

The vast repertoire of computer action games available representing different levels of difficulty and different interface complexity appears to be a useful source of simulated task repertoires for development of a cognitive simulation automaton.

Simulation of Behavior at the Rule-based Level

The best starting point will be to consider first well formed tasks at the rule-based level. This regime has been in focus of cognitive psychology of the information processing paradigm and much AI research has been based on 'production systems.' Well developed tools are therefore available for modeling. Furthermore, task environments can be found for which the skill-based movement control is decoupled from the task content by general, sign-based command interfaces (key board, joystick and mouse-based action controls). In addition, displays with well defined cues can be applied. In this case, the simulation model does not have to include the subjective, perceptive feature-formation and sensori-motor movement patterns that are the basic ingredients of expertise for activities in 'natural' task environments.

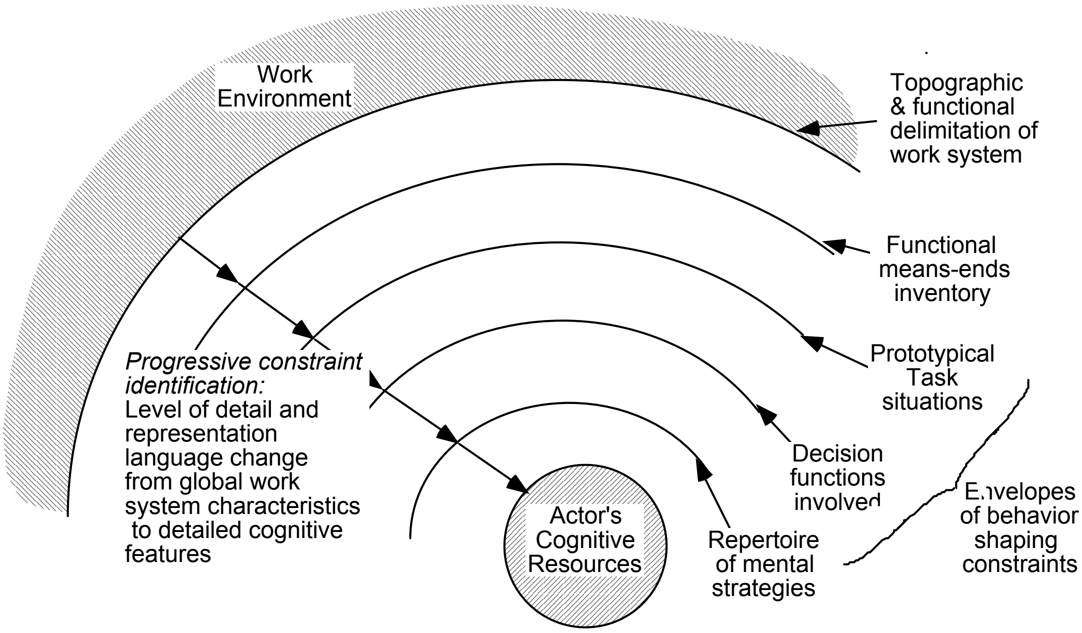


Figure 8. The figure illustrates the different envelopes of behavior-shaping constraints that are defined by the framework for work analysis and which are well suited to defined the boundary conditions for experiments in a model development program, See Rasmussen and Pejtersen, 1993.

A starting point can be a computer game such as 'the gymnast girl' studied by Hansen et al., (1991). This game has a number of features which makes it a promising first choice, and a rule learning automaton can be based on general trial-and-error learning and optimization features as illustrated schematically in figure 9.

The Gymnast Girl Game. This game involves the control of actions of a gymnast who can jump a horse in a number of different ways including one or more somersaults and/or body twists. The score depends on the complexity of the jump, its height and the elegance of the final landing. Several strategies can therefore be chosen by the player according to different subjective criteria such as playing safe (aiming at a high average score across sessions), taking risk (aiming at highest possible score in an individual session), curiosity (searching for new jump patterns), etc. The game, therefore, is well suited to study discovery and learning of new rules and strategies.

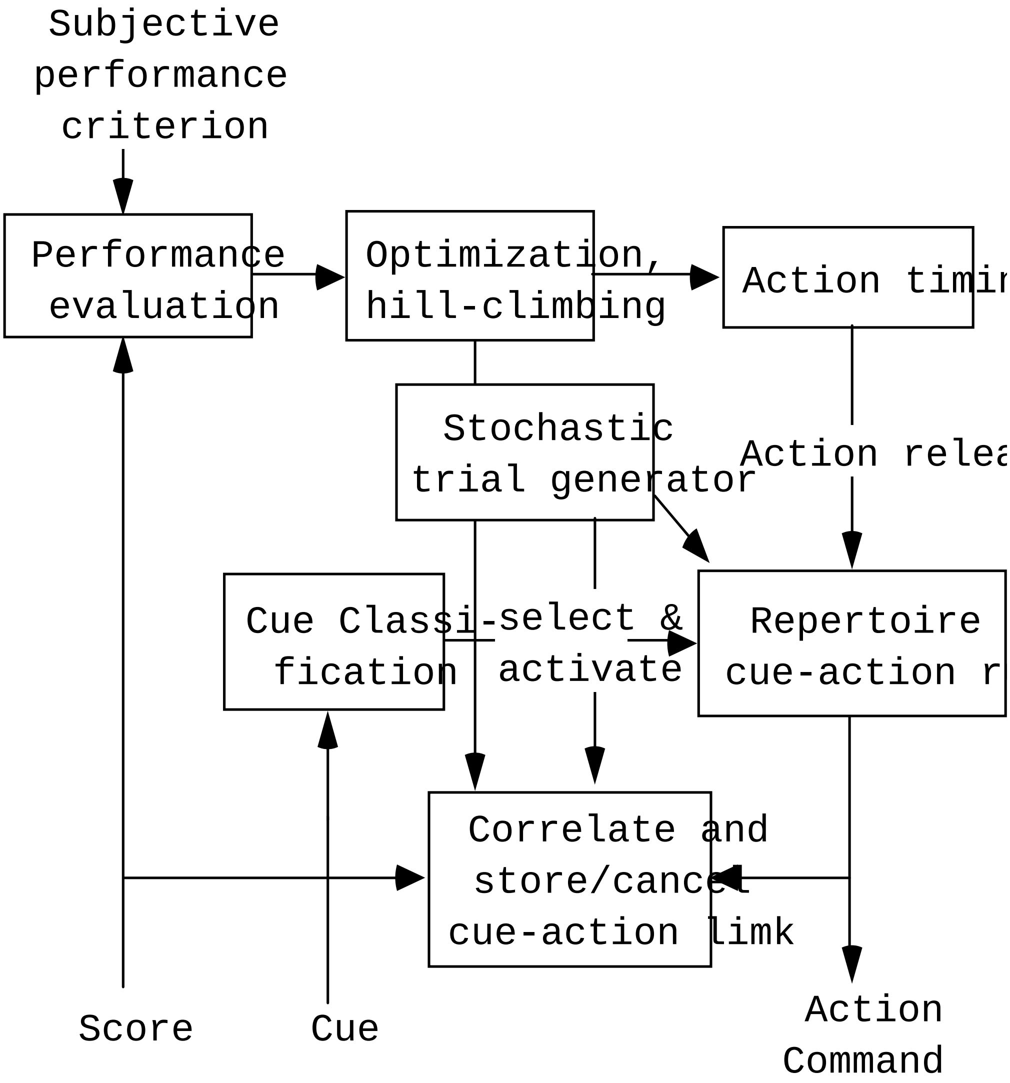


Figure 9. A schematic illustration of the structure and functions implied in simple trial-and-error rule-learning and optimization in a simulated task (game) having explicit cues presented by the interface and actions controlled by selection from a set of commands.

Learning strategies. Different learning strategies can be studied by comparing performance of subjects playing the game with an automaton by a sequence of sessions with different levels of a priori knowledge. On one hand, the subjects can be instructed initially regarding all the possible jumps and the related action rules. In this case, the simulating automaton will be programmed to include all action sequences. The learning process then merely involves the discovery of the temporal patterns of each jump followed by optimization of the accurate timing. On the other hand, the subjects can be introduced to only one simple jump rule and asked to discover other jump sequences by trial and error. In that case, some stochastic search for successful action sequences must precede the optimization of timing. Proper sequencing of actions and optimization of timing is a learning task which is involved in most activities. Even if the timing requirements of the gymnast-girl game is close the limits of normal sensori-motor performance, the game is useful for experiments with automata due to the short temporal length of a game session which facilitates the use of a large number of experimental sessions.

Subjective Performance Criteria. Even if the game is simple, the repertoire of jumps of different levels of difficulty makes the game well suited for the initial experiments to identify automata characteristics which can serve to represent 'individual differences' among the experimental subjects. Lõvborg's data (see Hansen et al., xyz) show a clear difference between the cautious subjects favoring a good average score and the risk-takers, favoring an experimental approach with low average score but with occasional very high score of a single, successful and complex jump. In the representation and processes of an automaton, such a difference will require different parameters of the hill-climbing and the explorative stochastic search.

Action-by-Command Interface. The game is controlled by simple actions by a joy stick. Therefore, the spatial aspects of the movement patterns are decoupled from the subject matter of the game and need not to be included in the cognitive model. That is, the spatial part only include stereotype commands (signs) that can be replaced by programmed commands when the automaton is coupled to the game (designing robotic control of joystick positions by the automaton will not contribute to the modeling aim). The temporal aspects of the movement patterns, however, are directly involved in the task and skill acquisition here involves the optimization of the temporal pattern for each individual jump type. This part of the adaptation can be represented by a hill-climbing algorithm optimizing scores when first the rule, that is, the action sequence, corresponding to the individual type of jump has been learned (the automaton has been synchronized to the jump pattern). Actually, it will only be of interest to include the spatial aspects of the skilled movement patterns in the model when direct manipulation interfaces to a task are considered (computer based or 'natural' in a manual object manipulation tasks).

Cues for Action. The representation of the state of the game on the interface of the 'gymnast girl' includes a stationary background and a representation of the jumping girl which is generated in the game program by combination of a limited number of icons, 'sprites,' representing body parts in different positions. The cues for action adopted by the subject, which will be derived from the position and the posture of the jumping girl, and can be derived from the program statements controlling the display. It will, consequently, be possible to connect the input of an automaton directly to the display commands of the computer and, in this way, to develop a simulation of the rule learning mechanisms without being involved in the difficult problem of identifying and modeling the perceptive and subjective feature formation involved in games with a more complex and dynamic interface representation.

Tacit Knowledge Context. The context of the 'gymnast girl' considered as a 'work domain' is familiar to most people. This is an advantage when used as a cover story to define the experimental context met by the subjects. It is, therefore, possible by analysis of the game interface to infer the rules that subjects will know implicitly (that is, in addition to the instructed rules). These rules must, of course, be given explicitly to the automaton. This opens the opportunity to use the experiments to explore experimentally which rules a subject actually derive from the context presented by an interface 'cover-story.'

Simulation of Perceptive Cue Formation

Simulation of the perceptive feature formation that is an important function of the cue-action links involved in sensori-motor patterns in a natural environment is probably the next important development of an automaton. This involves the process of correlating certain invariant features of the interface to the success of certain patterns of actions as experienced during explorative activities. For this phase of the model development, games with more dynamically complex interfaces will be necessary, such as, e.g., the NewFire game described by Løvborg (xyz). The interface of this game is a simple two dimensional map of a forest indicating burning areas and location of firefighters in a 20x20 partition of the forest area. For this game, several different fire-fighting strategies have been identified as being successful for distinct different fire patterns. Consequently, different configural, visual patterns are acting as cues for selection of strategies and actions. Such complex cues are identified by a subconscious feature formation and correlation process and the automaton must be interfaced to the game by a learning feature classifier. An interesting extension of the simple automaton of figure 9 would be to interface it to more complex interfaces by the use of an adaptive classifier based on a neural-net algorithm.

Simulation of Skill-based Sensori-motor Performance

Another necessary extension of the simple rule-learning automaton will be to include a simulation of the skill-based movement control. This extension requires use of games or simulated work environments having 'direct manipulation features' and a well structured temporal-spatial constraint source.

A promising candidate for a first extension in this direction would be to use a vehicle control task for which an extension of the rule-learning automaton could be based on the optimal-control paradigm shown in figure 9. Such a task is studied by Hansen (xyz), who analyzes the performance in ship/harbor simulators of sea captains docking ships in harbors of different topographic design. From the simulator studies, cue-action sets have been identified representing different phases of learning and different subjective performance criteria as adopted by the individual captains. The task appears to be well suited for simulation because the temporal-spatial characteristics of ships are well defined, and the cue-rule sets adopted by captains are well described.

This task scenario is particular interesting for the development of a learning automaton because there is a potential for immediate application of the automaton for simulator evaluation of different harbor designs. At present, new designs are tested in ship simulators by professional captains which is an expensive approach. Possibly, the initial tests could be performed by simulation of captains' stereotype cue-action rules.

Simulation of Knowledge-based Problem Solving

An extension of the simulation model to also include knowledge-based reasoning appears to be less realistic until a reliable simulation of rule- and skill-based performance in some well structured work domains have been developed. 'Natural' reasoning and decision making can probably by simulated most realistically in the first phase by a recursive application of the rule-skill-simulation model on a symbolic object-event environment representing the mental model used for reasoning. First when rule-based reasoning on an established mental model is modeled does it appear to be realistic to enter simulation of the generation of mental models for unfamiliar environments.

For the simulation of knowledge-based cognitive control, much inspiration can be found in the cognitive and artificial intelligence research, It will, however, be necessary to modify the models available so as to match the function of a higher level supervisory function of the lower cognitive levels.

REFERENCES

Reviews of Trentowski, Bogdanov, Smuts and Leduc are found in: International Journal of General Systems, Vol.5, No.1, 1979, pp. 63-71.

Alexander, C. (1964). Notes on the Synthesis of Form. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1990): The Adaptive Character of Thought. Hillsdale, New Jersey: Lawrence Erlbaum.

Ashby, W. R. (1960): Design for a Brain: The Origin of Adaptive Behavior. London: Chapman and Hall

Bambrough, R. (1963): The Philosophy of Aristotle. New York: Mentor Books.

Barley, S. R. (1988): On Technology, Time, and Social Order: Technically Induced Change in the Temporal Organization of Radiological Work. In: F. A. Dubinskias (Ed.): *Making Time*; Philadelphia: Temple Univ. Press.

Barnard, C. I. (1938): The Function of the Executive. Cambridge, Mass.: Harvard University Press.

Baron, S., and D. L. Kleinman. (1969). The Human as an optimal controller and information processor. IEEE Trans. Man-Mach. Syst. MSS-10, (1), 9-17.

Baron, S., C. Feehrer, R. Muralidharan, R. Pew, and P. Horwitz. (1982). An Approach to Modeling Supervisory Control of a Nuclear Power Plant. NUREG/CR-2988. ORNL/Sub/81-70523/1.

Baron, S., G. Zacharias, R. Muralidharan, and R. Lancraft. (1981). PROCHU: A model for analyzing flight crew procedures in approach to landing. Proceedings of the Eighth World Congress of the International Federation of Automatic Control, Vol 7, 3481-3487. Oxford, England: Pergamon

Beer, S. (1959): Cybernetics and Management. New York: John Wiley

Beer, S. (1966): Decision and Control. New York: John Wiley

Beer, S. (1975): Platform for Change. New York: John Wiley

Beer, S. (1979): The Heart of the Enterprise. New York: John Wiley

Beer, S. (1981): The Brain of the Firm. Chichester: John Wiley

Beer, S. (1989): The Viable System Model. In: R. Espejo and R. Harnden (Eds.): The Viable System Model: Interpretations and Applications of Stafford Beer's VSM. Chichester: John Wiley

Bernard, C. (1878): Leçons sur les phenomenes de la vie. Communs aux animaux et aux vegetaux. Vol. 1. pp. 113-121. Paris: Balliere.

Bogdanov, A.A. (1922): See Gorelik, G. (1980): Essays in the Universal Organizational Science. InteSsystems Publishers.

Boulding, K. (1968): General Systems Theory - The Skeleton of a Science. In: Walter Buckley (Ed.): Modern Systems Research for the Behavioral Scientist. Chicago: Aldine Pub. Co.

Brehmer, B. and Leplat, J. and Rasmussen, J. (1991): Use of simulation in ths study of complex decision making. In: Rasmussen, J., Brehmer, B. and Leplat, J. (Eds.): Distributed Decision Making: Cognitive Models for Cooperative Work. London: John Wiley and Sons.

Brown, I. S., and J. DeKleer. (1981).Towards a theory of qualitative reasoning, in, J. Rasmussen and W. B. Rouse, eds. Human Detection and Diagnosis of System Failures, New York: Plenum Press, 317-335.

Bruner, J. S., Goodnow, J. J. and Austen, G. A. (91956): A Study of Thinking. New York: Wiley.

Cannon,W. B. (1929): Organization for Physioloical Homoestasis. Physiol. Rev. 9: 399-431.

Churchman, C. W, (971): The Design of Inquiring Systems: Basic Concepts of Systems and Organization. New York: Basic Books

Clark, B.R., (1956): Adult Education in Transition, Berkeley: University of California Press.

Crossman, E. R. F. W., and J. E. Cooke. (1962). Manual control of slow-response systems. International Congress on Human Factors in Electronics, Long Beach, California, 1962; Also in (1974). W. Edwards and F. P. Lees eds. The Human Operator in Process Control London: Taylor & Francis.

Cyert, R. M. and March, J.G., (1963): A Behavioral Theory of the Firm, Englewood Cliffs, N.J.: Prentice-Hall, Inc.

De Kleer, J. and Brown J. S. (1986): A Qualitatiive Physics Based on Confluences. In: Bobrow, D. G. (Ed.): Qualitative Reasoning about Physical Systems. Amsterdam: Elsevier

De Kleer, J. and Brown, J. S. (1983): Assumptions and Ambiguities in Mechanistic Mental Models. In: Gentner, D. and Stevens, A. L. (Eds.): Mental Models. Hillsdale, N.J.: Lawrence Erlbaum.

Dennet, D. C., (1978): Brainstorms. Montgomery, VT: Bradford Books.

Duncker, K. (1943). On problem solving. Psychological Monographs. Vol. 58, No. 5, Whole no. 270.

Espejo, R. and R. Harnden (Eds.) (1989): The Viable System Model: Interpretations and Applications of Stafford Beer's VSM. Chichester: John Wiley

Forrester J.W. (1968): Principles of Systems. Cambridge, Ma.: Wright-Allen Press

Forrester J.W. (1971): World Dynamics. Cambridge, Ma.: Wright-Allen Press

Gould, S. J. (1987): Is a New and General Theory of Evolution Emerging? In: Yates, F. E., Garfinkel, A., Walter, D.O., and Yates, G.B. (Eds.): Self-Organizing Systems: The Emergence of Order. New York: Plenum.

Gouldner, A.W., (1959): Organizational Analysis, in Robert K. Merton, Leonard Broom, and Leonard S. Cottrell, Jr. (eds.), Sociology Today, New York: Basic Books.

Govindaraj, T. (1987): Qualitative Approximation Methodology for Modeling and Simulation of Large Dynamic Systems: Applications to a Marine Steam Power Plant. IEEE Trans. SMC, Vol. 17, pp. 937-955

Gulick, L. and Urwick, L. (eds.), (1937): Papers on the Science of Administration, New York: Institute of Public Administration.

Hansen, J.P., Løvborg, L. and Rasmussen, J. (1991): Simulation of Cognitive Behavior in Computer Games. Proceedings of the Second MOHAWC workshop, Manchester, November 1990. Roskilde: Risø National Laboratory

Hindle, B. (1981): Emulation and Invention. New York: New York University Press.

Iberall, A. and Soodak, H. (1978): Physical Basis for Complex Systems Some Propositions Relating Levels of Organization. Collective Phenomena, Vol.3, pp. 9-24, 1978.

Iberall, A. S. (1978): A Field and Circuit Thermodynamics for Integrative Physiology. I. Introduction to the General Notion. Invited Opinion in: Am. J. Physiol. 233, R171-180.

Iberall, A. S. (1978): A Field and Circuit Thermodynamics for Integrative Physiology. III. Keeping the Books A General Experimental Method. Invited Opinion in: Am. J. Physiol. 235, R85-R97.

Iberall, A. S. Soodak, H. and Hassler, F. (1978): A Field and Circuit Thermodynamics for Integrative Physiology. II. Power and Communicational Spectroscopy in Biology. Invited Opinion in: Am. J. Physiol. 234, R3-R19.

Iwasaki, Y. and Simon, H. A. (1986): Causality in Device Behavior. Artificial Intelligence, Vol. 29 pp. 3-32.

Iwasaki, Y. and Simon, H. A. (1986): Theories of Causal Ordering: A reply to de Kleer and Brown. Artificial Intelligence, Vol. 29 pp.63-73.

Kleinman, D. L., S. Baron, and W. H. Levison. (1971). A control theoretic approach to manned-vehicle systems analysis. IEEE Trans. Automat. Control. AC-16, 824-832.

Lashley, K. S. (1942). The problem of cerebral organisation in vision, in Visual Mechanisms. H. Kluver, ed.Lancaster Cattell: (1969): Biological Symposia, 7, 301-322. Reprinted in K. H. Pribram, ed. Perception and Action. (Brain Behaviour, Vol. 2), 575. Harmonthsworths, England: Penguin Books.

Leduc, S. (1911): The Mechanisms of Life. London: Rebman Ltd.

Leonard, T. (1960). Optimizing linear dynamics for human operated systems by minimizing the mean squared tracking error. WESCON, 4. (Pt. 4), 57-62.

Lettvin, J. V., H. R. Maturana, and W. S. Mcculloch. (1959). What the frog's eye tells the frog's brain. Proc. IRE, 47, 1940-1951.

Lind, M. (1993): Interactive Planning for Integrated Supervision and Control in Complex Plant. Tech. Report, Project 4937-92-08-ED ISP DK. Lyngby, Dk.: Institute of Automatic Control; Tech. Univ.

Lind, M. (1993): Multilevel Flow Modeling. AAAI Workshop on 'Reasoning about Function.' Washington, DC. July, 1993

Lind, M. (1992): A Categorization of Models and its Application for the Analysis of Planning Knowledge. Proc. POST ANP'92 Conference on Human Cognitive and Cooperative Activities in Advanced Technological Systems. Kyoto, 1992

March, J.G. and Simon, H.A., (1958): Organizations, New York: Wiley.

Marr, D. (1982): Vision. San Francisco, CA: Freeman.

Maturana, H. R. and Varela, F. J. (1980): Autopoesis: The Orgaanization of the Living. In Maturana and Varela: Autopoesis and Cognition. Boston: Reidel Publ.

Maxwell, J. C. (1868): On Governors. Proc. Royal Soc. London. 16-1868, 270-283.

McClelland and Rummelhart (1986): Parallel Distributed Processing. Cambridge, MA: MIT Press

Meadows et al. (1972): Limits to Growth. Washington, D.C.; Potomac Ass.

Mesarivic, M and Pestel, E. (1975): Mankind at the Turning Point: The Second Report to the Club of Rome. London: Hutchinson

Mesarovic, M. D. (1968): Systems Theory and Biology: View of a Theoretician. In: Mesarovic, M. D. (Ed.): Systems Theory and Biology, Proc. 3rd System Symp. Berlin: Springer Verlag, pp. 59-87

Mesarovic, M.D. (1970a). Multilevel Systems and Concepts in Process Control, *Proceedings of the IEEE*, vol. 58, No. 1. New York: The Free Press of Glencoe.

Mesarovic, M.D. (1970b). Theory of Multi-Level, Hiearachical Systems. New York: Academic Press.

Mitroff, 1974

Mitroff, I. I. and Linstone, A. I. (1993): The Unbounded Mind: Breaking the Chains of Traditional Business Thinking . New York: Oxford University Press, 1993.

Monod, (1972): Chance and Necessity. New york: Knopf.

Newell A, (1982): The Knowledge Level. Artificial Intellignce 18, 87-127

Newell, A (1973): Production Systems: Models of Control Structures. In Chase, W. C. (Ed.): Visual Information Processing. New York: Academic Press.

Newell, A. (1980): Physical Symbol Systems. Cognitive science, 4, pp. 135-183.

Newell, A. and Simon, H. A. (1963): GPS: A Program that Simulates Human Thought. In: Feigenbaum, E. A. and J. Feldman (Eds.): Computers and Thought. New Yprk: McGraw-Hill.

Newell, A., Shaw, J. C. and Simon, H. A. (1958): Elements of a Theory of Human Problem Solving. Psychological Review, 65, 151-166.

Parsons, T. (1960): Structure and Process in Modern Society, New York: The Free Press of Glencoe.

Pattee, H. H. (1973): Physical Problems of the Origin of Natural Controls. In Locker, a. (Ed.): Biogenesis, Evolution, Homeostasis. Berlin: Springer Verlag.

Pattee, H. H. (1977): Dynamic and Linguistic Modes of Complex Systems. Int. J. General Systems, 1977, Vol. 3. pp. 259-266.

Pattee, H. H. (1987): Instabilities and Information in Biological Selforganization. In: Yates, F. E., Garfinkel, A., Walter, D.O., and Yates, G.B. (Eds.): Self-Organizing Systems: The Emergence of Order. New York: Plenum.

Pew, R. W. (1974). Human perceptual-motor performance, in Human Information Processing: Tutorials in Performance and Cognition. B. H. Kantowitz, ed. New York: Erlbaum.

Pondy, L. R. and Mitroff, I. I. (1979): Beyond Open System Models of Organization. Research in Organizational Behavior, Vol. 1, pp 3-39. JAI Press

Polanyi, M. (1968): Life's Irreducible Structure. Science, 160, 1308-1312.

Pylyshyn, Z. W. (1984): Computation and Cognition: Toward a foundation of Cognitive Science. Cambridge, MA: MIT Press

Rasmussen, J. (1988): Coping Safely with Complex Systems. American Association for Advancement of Science, invited paper for Annual Meeting, Boston, Februar 1988; In: Risø-M-2769.

Rasmussen, J., (1993): Risk Management, Adaptation, and Design for Safety. In Sahlin, N. E. and B. Brehmer (Eds.): Future Risks and Risk management. Dordrecht: Kluwer. In press.

Rasmussen, J.. (1990a): Human Error and the Problem of Causality in Analysis of Accidents. Phil. Trans. R. Soc. Lond. B 327, 449-462.

Rasmussen, J. (1990b): The role of Error in Organizing Behaviour. Ergonomics, 1990, vol. 33, nos 10/11, 1185-1190.

Rasmussen, J. (1991): Modelling Distributed Decision Making. In: Rasmussen, J.: Brehmer, B. and Leplat, J. (Eds.) (1991): Decision Making: Cognitive Models for Cooperative Work. London: John Wiley and Sons.

Rasmussen, J. and Pejtersen, A., M., (1993): MOHAWC Taxonomy: Implications for Design and Evaluation. Final report, CEC Basic Research Project Mohawc. Also in: Risø-R-673; Roskilde: Risø National Laboratory.

Rasmussen, J. and Pejtersen, A., M., (1993): Virtual Ecology of Work. Invited contribution in: Flach, J., Hancock, P., Caird, J., and Vicente, K. (Eds.): Ecology of Human-Machine Systems: Hillsdale, N.J.: Lawrence Erlbaum.

Rasmussen, J., A. M. Pejtersen & K. Schmidt (1991): *Taxonomy for Cognitive Work Analysis.* Risø-M-2871

Rochlin, G.I., La Porte, T.R., and Roberts, K.H., (1987): The Self-Designing High-Reliability Organization: Aircraft Carrier Flight Operations at Sea, Naval War College Review, Autom 1987.

Roethlisberger, F.J. and Dickson, W.J. (1939): Management and the Worker. Cambridge, Mass.: Harvard University Press.

Roig, R. M. (1962). A comparison between human operator an optimum linear controller RMS-Error Performance. IRE Trans. Hum. Fact. Electron.. HFE-3 (1), 18-21.

Rosch, E. (1975): Human Categorization. In: N. Warren (Ed.): Advances in Cross-Cultural Psychology. New York: Halsted Press.

Rosenbloom, P. S., Laird J. E., Newell, A., and McCarl, R. (1991): A Preliminary Analysis of the Soar Architecture as a Basis for General intelligence. Artificial Intelligence, 47, 1991, 289-325.

Rosenbluth, A., Wiener, N, and Biegelow, J. (1943): Behavior, Purpose, and Teleology. Philosophy of Science, Vol. 10, pp. 18-24.

Rubin, E. (1920). Vorteile der Zweckbetrachtung fur die Erkentnis. Z. Psychol.. 85, 210-223. Also in Experimenta Psychologica, 66-81. Copenhagen: Munksgaard.

Russel, B. (1913): On the notion of cause. Proc. Aristot. Soc. 13, 1-25.

Schoeffler, J.D., (1969): Decomposition and Multilevel Methods for On-Line Computer Control of Industrial Processes. Technical Report; Systems Research Center, Case Western Reserve University.

Schriver, J. C. (1992): Object-Oriented Qualitative Simulation of Human Mental Models of Complex Systems. IEEE Traans SMC Vol. 22, No. 3, 526-541

Selznick, P. (1949): TVA and the Grass Roots. Berkeley, Calif.: University of California Press.

Sheridan, T. B. and Ferrell, W. R. (1974): Man-Machine Systems: Information, Control, and Decision Models of Human Performance. Cambridge, MA: MIT Press.

Simon, H. A. (1969): *The Sciences of the Artificial*, 2nd ed., MIT Press, Cambridge, Mass., and London, 1981.

Simon, H.A. (1957a): Administrative Behavior. New York: Macmillian.

Simon, H.A. (1957b): Models of Man, Social and Rational. New York: John Wiley and Sons.

Smuts, J. Ch. (1926): Holism and Evolution. Reprinted: Wesstport Conn.: Greenwod Press, 1973.

Taylor, F.W. (1911): Scientific Management. New York: Harper and Row.

Taylor, R, (1950): Comments on a Mechanistic Conception of Purposefulness. Philosophy of Science. Vol. 17, pp 310-317.

Thompson, J.D. (1967): Organizations in Actions. New York: McGraw-Hill.

Toynbee, A.J. (1972): A Study of History. London: Oxford University Press and Thames and Hudson.

Trentowski, B. (1843): Stozunek Filozofil do Cybernetyki Czyll Sztuky Rzadzenia Narodem. Poznan: J.K. Zupanski.

Vico, G. (1774): See: Bergin T. G. and N. H. Fisch (1970): The New Science of Giambattista Vico. Abridged Translation of the third edition, 1774. Ithacha, N.Y.: Cornell Univ. Press.

Waldrop, M. M. (1987). Computers Amplify Black Monday. Science, Vol. 238, p. 602-604

Weaver, W. (1948): Science and Complexity. American Scientist, Vol. 36, No. 4, pp 536-544.

Weber, M., (1947): The Theory of Social and Economic Organization. A.M.Henderson and Talcott Parsons (trans.)and Talcott Parsons (ed.), New York: The Free Press of Glencoe.

Weiss, P.A. (1925): Tierische Verhalten als 'Systemreaktion.' Die Orientierung der Ruhestellungen von Schmetterlingen (Vanessa) gegen Licht und Schwerkraft. Biologia Gen., Vol. 1, 1925, pp. 168-248.

Wiener, N. (1948): Cybernetics or Control and Communication in the Animal and the Machine. New York: John Wiley & Sons.

Yates, F. E. (1982): Outline of a Physical Theory of Physiological Systems. Can. J. Physiol. Pharmacol. 60: 217-248.

Yates, F. E. and Iberall, A. S. (1982) A Skeleton of Physical Ideas for the Dynamics of Complex Systems. Mathematics and Computers in Simulation, XXIV (1982), pp. 40-436.

Zelenyy, M. (1980): Autopoesis: A Paradigm Lost. In: Zeleny, M. (Ed.): Autopoesis, Dissipative Structures, and Spontaneous Social Order. 1980. Boulder Colorado: Westview Press.