[11] D. B. Lenat, "Beings: Knowledge as interacting experts," in *Proc. 4th Int. Joint Conf. Artificial Intelligence*, Sept. 1975, pp. 126–133.
[12] V. R. Lesser, R. D. Fennell, L. D. Erman, and D. R. Reddy, "Organization of the HEARSAY II speech understanding system," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 11–24, Feb. 1975.
[13] V. R. Lesser *et al.*, "Working papers in distributed computation I: Cooperative distributed problem solving," Dep. Comput. and Inform. Sci., Univ. of Massachusetts, Amherst, July 1979.
[14] V. R. Lesser and L. D. Erman, "Distributed interpretation: A model and experiment, *IEEE Trans. Comput.*, vol. C-29, pp. 1144–1163, Dec. 1980.
[15] H. P. Nii and E. A. Feigenbaum, "Rule-based understanding of signals," in *Pattern-Directed Inference Systems*, D. A. Waterman and F. Hayes-Roth, Eds. New York: Academic, 1978, pp. 483–501.
[16] N. J. Nilsson, *Problem Solving Methods In Artificial Intelligence*. New York: McGraw-Hill, 1971.
[17] R. G. Smith and R. Davis, "Applications of the contract net framework: Distributed sensing," *Proc. ARPA Distributed Sensor Net Symp.*, pp. 12–20, Dec. 1978.
[18] R. G. Smith, "A framework for problem solving in a distribu processing environment," Dep. Comput. Sci., Stanford Univ., St ford, CA, STAN-CS-78-700 (HPP-78-28) Dec. 1978.
[19] ——, "The contract net protocol: High-level communication : control in a distributed problem solver," *IEEE Trans. Comput.*, C-29, pp. 1104–1113, Dec. 1980.
[20] D. L. Waltz, "Generating semantic descriptions from drawings scenes with shadows," Massachusetts Inst. Technol., Cambric MIT AI-TR-271, Nov. 1972.
[21] P. H. Winston, *Artificial Intelligence*. Reading, MA: Addis Wesley, 1977.
[22] S. W. Zucker, R. A. Hummel, and A. Rosenfeld, "An applicatio relaxation labelling to line and curve enhancement," *IEEE Tr Comput.*, vol. C-26, pp. 394–403, Apr. 1977.
[23] S. W. Zucker, Y. G. Leclerc, and J. L. Mohammed, "Continu relaxation and local maxima selection," Dep. Elec. Eng., Mc Univ., Rep. 78-15R, Dec. 1978.

# An Organizational View of Distributed Systems

MARK S. FOX, MEMBER, IEEE

*Abstract*—The relationship between organization theory and distributed systems is studied. By viewing distributed systems as analogous to *human organizations*, concepts and theories germane to the management science field of organization theory can be applied. Task *complexity, uncertainty*, coupled with *resource constraints* are shown to be important factors in deciding how a system is to be distributed.

## I. Introduction

DISTRIBUTED SYSTEMS are difficult to design. At least two approaches should be taken to alleviate this problem. We should build more distributed systems (learning by doing) and draw upon ideas from other fields (e.g., biology, management science) that have considerable experience with their own distributed systems (learning by analogy). This paper is an example of the latter. By viewing distributed systems as being analogous to human organizations, concepts and theories germane to the management science field of organization theory can be applied.

To begin our discussion, a distributed system is defined as a particular organization—task decomposition and control regime—resulting from the distribution of a set of tasks over a set of logically or physically disjoint processing elements. Research in distributed systems has focused analyzing classes of systems. Systems where multiple ta exist but share the same goal have been termed *teams* Team tasks where the data (resource) are (initially) ph cally distributed, e.g., weather monitoring, have come der the study of *distributed sensor nets* [2]. Team tasks t require the sharing of data (raw or processed) have b termed *cooperative* [3]. Tasks where control is hierarchi are called *organizations* [4]. Though a variety of structu forms have been distinguished, there is a paucity of kno edge available to guide system design. Moreover, the sign of a distributed system's hardware often precedes organizational analysis, resulting in inadequate archi tures. The major problem with designing distributed tems is deciding how the task should be decomposed the control regime to be used, and this choice of org zation is determined by features of the task (domain) some measurement criteria.

Economics and management science have had a hist of designing and analyzing distributed systems, i.e., hu organizations. Beginning with Adam Smith's theory labor division [5], through the more current organizati work of March and Simon [4], Galbraith [6], and Willi son [7], much research has centered around construc distributed systems that best suit a particular task. Gi the existence of this body of knowledge, can it be app to distributed systems design in the computer sciences? stated another way, is the distributed computer syste human organization metaphor valid? Distributed syst

Manuscript received December 5, 1979; revised September 10, 1980 and September 25, 1980. This work was supported in part by the Defense Advanced Research Projects Agency (DOD), Arpa Order 3597, monitored by the Air Force Avionics Laboratory Contract 533615-78-C-1551 and in part by a National Research Council of Canada Post-Graduate Scholarship.

The author is with The Robotics Institute and Department of Computer Science, Carnegie-Mellon University, Pittsburgh PA, 15213.

0018-9472/81/0100-0070$00.75 ©1981 IEEE

ce the problem of allocating tasks, resources, and infor-ation to a set of processors. These are precisely the oblems of human organizations. Hence it appears valid the macro level, but does it hold at the micro level, i.e., oes the computer processor share characteristics with the iman mind? Simon's theory of *bounded rationality* seems oplicable here:

> The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world—or even for a reasonable approximation to such objective rationality [8].

he human mind's processing capacity is limited. This esult is what Simon calls *bounded rationality.* Bounded ationality implies that both the information a person can bsorb and the detail of control he may wield is limited. As asks grow larger and more complex, means must be found o effectively limit the increase of information a person ees and the complexity of control. Bounded rationality is prime factor in the evolution of multiperson organi-ations from an unregimented group to more structured lternatives.

A computer processor also has a limited processing apacity. A processor can execute only a limited number of nstructions per second. This limits the amount of informa-ion a processor may process and the amount of control it nay exercise within a given time period. Hence, pro-grammed systems, whether centralized or distributed, may exhibit symptoms similar to the bounded rationality ex-hibited by humans when capacities are exceeded. The metaphor appears viable.

The purpose of this paper is to attempt a technology transfer from the field of organization theory; to transfer both descriptive and analytical information to aid in the analysis and design of distributed systems. Fig. 1 depicts the overall approach to be taken in analyzing such systems. The design of a distributed system (organization) requires the selection of an organizational structure, i.e., processes (modules) and communication paths, and a control regime. The efficacy of a selected organization is dependent upon complexity and uncertainty features of the task. Determining how uncertainty and complexity affects the organization requires measurement techniques, for example, transaction analysis. Finally, the measurement cannot be interpreted without some reference criteria called organizational goals.

The following sections elucidate the levels in Fig. 1. First, the Hearsay-II speech understanding system architecture which will be used as an example throughout the paper is described. Next, the analysis of organization theory is begun by surveying the set of organization structures and control found in the literature. The requirements for a theory of organization analysis are then described, followed by an analysis of two task features, uncertainty and complexity, which affect the organization (level 2). For most task features, an analysis technique is outlined (level 3) and an organizational solution specified (level 1).

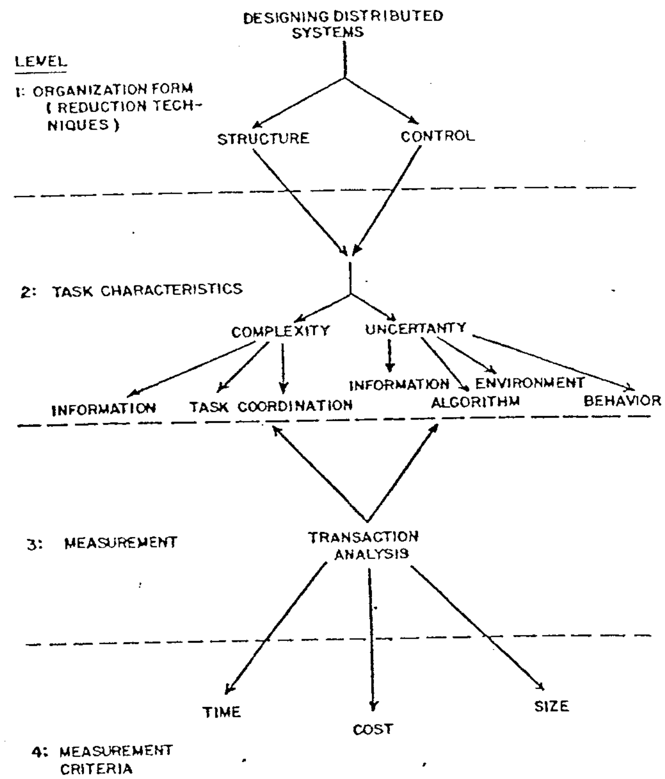


Fig. 1. Organization analysis.

## II. An Example: Hearsay-II

Throughout this paper, organization theoretic concepts will be introduced. To aid the reader in interpreting their applicability, we apply and contrast these ideas with the architecture of the Hearsay-II speech understanding system. The reader should refer to the following description throughout the rest of the paper.

Hearsay-II [9], [10] is a system designed to understand connected speech.[1] Utterances, without artificially introduced pauses between words, are spoken to the system. Hearsay-II must interpret, understand, and reply to the utterance. The current version of Hearsay-II retrieves and answers questions about abstracts stored in its data base [11].

The process of understanding utterances requires the application of many sources of knowledge: acoustic, syllabic, lexical, prosodic, syntactic, semantic, pragmatic, etc. Each source of knowledge can be used to interpret the utterance at its own particular level of representation. Each source of knowledge only partially represents the knowledge a human brings to bear when parsing speech. These sources represent the state of the art of our knowledge of the speech understanding process. Because of the incompleteness of the knowledge, the understanding process is saturated with error. Thus speech understanding is a search in a large space of possible interpretations for the utterance that best fits the input data, i.e., the speech waveform.

The design of a speech understanding system must allow the integration of sources of knowledge in such a way that they may gracefully interact. The errorfulness of the

[1] See [30] for a good introduction to the problem.

processing requires that the program have the ability to redirect its attention whenever the current best interpretation of the utterance proves implausible.

The approach taken in Hearsay-II is as follows. The knowledge in the system is represented in separate processes called knowledge sources (KS's). Each KS contains a separate portion of knowledge such as syntax and semantics: SASS [12]; lexical:POMOW [13]; semantics: SEMANT [14]. The knowledge is integrated by allowing the knowledge sources to communicate via a blackboard (BB). The BB is a common dynamic data structure. Each KS can be viewed as an expert in its particular field and contributes to the "discussion" among the experts by reading and writing information on the BB. The mode of BB interaction is *hypothesize and test* [15]. Each KS can either place an hypothesis, describing its interpretation of BB data (i.e., other hypotheses), on the BB, or test (i.e., accept or reject) BB hypotheses produced by other KS's. As mentioned above, the knowledge in the different KS's can be used to interpret the utterance at different levels of representation. Specifically, the levels of representation (knowledge) form a hierarchy. Each level is built upon a lower level. The lexical level is built upon the syllabic, and the syntactic upon the lexical. The job of a KS is to construct an interpretation (hypothesis) at its level of expertise by postulating (or testing) hypotheses constructed from hypotheses at a lower level or by elaborating hypotheses from a higher level.

The processing of the system is data-directed. It is directed by the current state of the BB data.[2] Each KS can view BB hypotheses at its level(s) of expertise. Whenever a change is made to an hypothesis or the BB by a KS, other KS's react through further hypothesization and testing. At any time there are many possible KS's capable of executing. The choice of which KS to execute is controlled by policy modules and the scheduler. Together they provide a focus of control (FOCUS) mechanism [16] capable of directing the system's attention to the currently best hypotheses, or redirecting the system when the current hypothesis proves unfruitful.

Fig. 2 shows the organization of the Hearsay-II system. Fig. 3 shows the blackboard hypotheses for interpreting the utterance "Tell me about beef." Initially, the segmentation KS segments the speech input and places on the BB hypotheses for all possible labels of the sound in each segment. The syllable KS (MOW) looks at the label hypotheses and transforms them to syllables. The lexical KS (POM) transforms syllable hypotheses to words and the syntax and semantics KS (SASS) transforms words to phrases. SASS also predicts, top down, missing words which causes the lower level KS's to attempt to verify the predictions. At any time over 100 possible KS executions are queued. FOCUS chooses the best KS to work on the more highly rated hypotheses.

[2]Each hypotheses is rated. It is a function of the ratings of the hypotheses it is constructed from and the knowledge used in the construction.

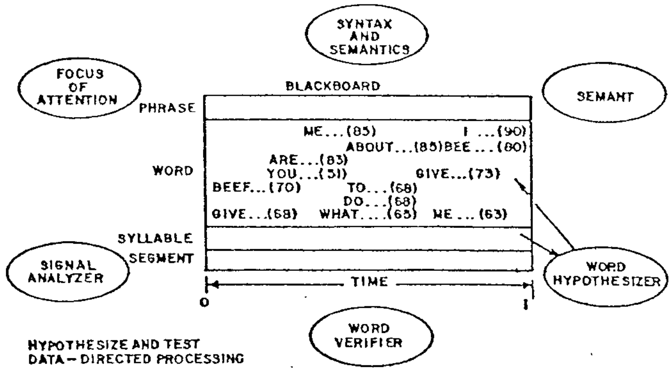


Fig. 2. Hearsay-II architecture.

## III. Organizational Structure and Control

An organization was defined as a composition of stru ture and control regime. The set of possible structure range from strict hierarchies to heterarchies. Possible co trol regimes range from an employment relation whe people (processes) are directly controlled by others vi agreed upon commands, to a price system where servic are contracted. In the following, a variety of organizatior found in the organizational literature are defined an related to the Hearsay-II architecture.

The simplest organization is the single person. The pe son performs all tasks, reacting to information and th environment when necessary. A single-person organizatio suffices as long as the person has the resources to achie the goal. *If Hearsay-II had only a single KS, it could l viewed as a single-person organization.*

As the task requires more resources (mental or physical the size of the organization increases, requiring more con plex control and an increase in information processir capabilities. Organizational forms such as a group result. group allows the cooperative coordination of individu members to achieve a shared goal. The task is divided u and subtasks are allocated to members who are best able t execute them. Coordination in a group is achieved throug mutually agreed upon decisions. To achieve this type c coordination, group members must share all available i formation. They must understand it and be able to con municate their views. Finally, they must arrive at a decisio that satisfies all. Each step in the coordination problem is series of transactions. *In its simplest interpretatio Hearsay-II is a group. Though all KS's are theoretical independent, their tasks are well integrated for achieving th problem of understanding speech. The PUP6 system* [17] *another example of a group organization.*

As the size of the group increases, collective decisio making becomes costly. The cost of information distrib tion and communication required to converge to a con mon decision increases. Hence a simple hierarchy evolve A simple hierarchy has two levels. The top level contains single decisionmaker who coordinates the efforts of th persons on the lower level. Complete information must b made available only to that person, and he must have th authority to effect changes in the organization's behavio

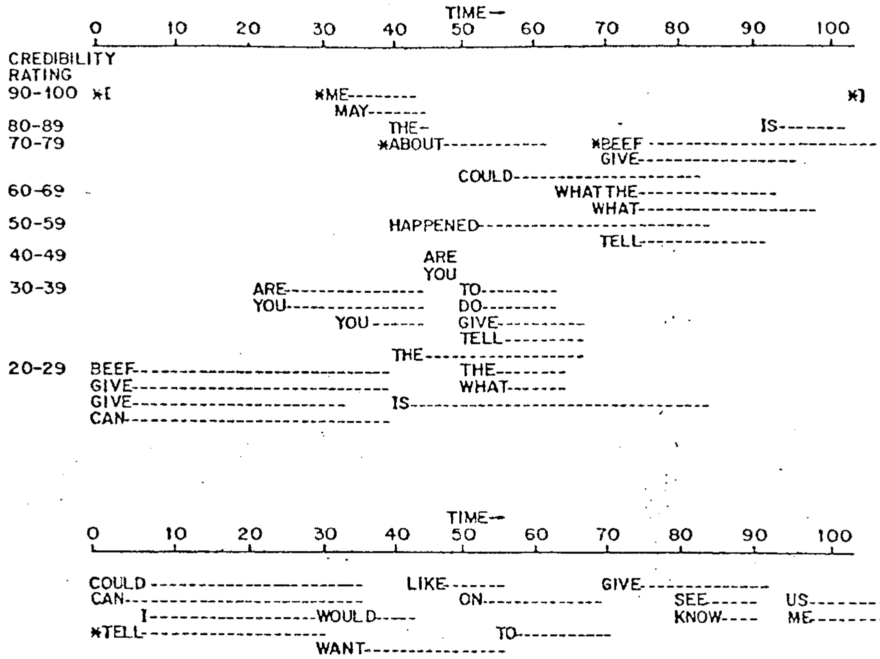


Fig. 3. Hearsay-II blackboard.

Proper coordination, implying employment and authority relations, and distribution of information, is required for the organization to be effective. *The final version of Hearsay-II fits this definition. The FOCUS KS made overall scheduling and directional decisions. (Also see SU/P[18], a derivative of Hearsay-II.)*

As the size of the single product organization grows, groups and simple hierarchies fail due to bounded rationality. The decisionmaker is unable to process all information (saturation). A uniform hierarchy results. Multiple levels of management are created to insure proper and centralized decisionmaking. Each level of the hierarchy acts as a filter on the information and decisions that are propagated up the hierarchy. Decisions are made at the lowest level in the hierarchy that has both the information to make the best decision and the authority to execute it. *Such an organization could be constructed with a hierarchy of Hearsay-II's. Each subsystem is a separate Hearsay-II system with a manager (focus) and employees (the rest of the KS's). Such a structure can be found in [19].*

As the uniform hierarchy increases in size and the number of products, powers of control are impaired, resulting in transactional diseconomies. With multiple products being produced, competition for the same resources arises among units. The problem of allocating resources so that enough are available to produce the products on schedule is quite complex.

One approach to reducing these effects is the multidivisional hierarchy. The organization is split along product lines. Each division is in full control of the tactics involved in producing their product. Hence control is situated locally where the information that enables control is available. Strategic control is vested in an elite staff assigned to a general office. The general office is concerned with strategic planning, appraisal and control including resource allocation. *If one viewed the interpretation of each area of the utterance time line as a separate "product," then the distributed Hearsay-II system [20] closely approximates this structure. Distributed Hearsay-II differs from a multidivisional hierarchy in that there is no strategic planning KS.*

As the organization grows, so does the coordination problem and the amount of the information to be processed. A significant step towards solving the problem is the price system. A number of disjoint organizations are available to produce a product or supply a service. Actions are initiated after the successful negotiation of a contract. This system eliminates all forms of control between units. All communication is contained in a contract to purchase some product or service. Control is exerted through the price of the product. Price (should) reflects the marginal cost of the product. The assumption is that through marginal pricing of goods, all resources will be utilized without waste. If a product is priced too high, it will not be purchased, if too low, the unit supplying it will go bankrupt. *Communication and control in Hearsay-II is noncontractual. The system was designed so that KS's share the same goal and blackboard communication process. Contracting assumes that a set of independent processes exists. Once processes enter into a contract, an organization is instantiated. Hence, contracting can be viewed as the dynamic creation of a system architecture. One could interpret a pre-Hearsay-II architecture as a heterarchy of unrelated KS's. A focus-like KS would let contracts to other KS's to do various speech understanding functions, with the resulting postcontractual architecture being similar to the Hearsay-II architecture.*

With the introduction of the price (or market) system, an organization does not have to create a new unit for each new function but can contract for the function in the marketplace. The next step in the evolution of an organization is a collective organization. The hierarchy is split into separate organizations who cooperate to achieve a shared goal. In one sense a collective can be viewed as a set of organizations that share long-term contracts.

The next step in successive reduction of control and information flow is the introduction of competition. Competing approaches to goal achievement are allowed (in the marketplace) with many organizations available to achieve any goal. Hence each organization peruses its own goals which correspond to another organization's needs. This is the general market situation. Services are contracted for in the marketplace for short or long periods of time.

## IV. Choosing an Organization

The previous section presented a variety of possible organizations. The question remains: which organization is best suited for the task at hand? Developing an analytical theory of organization choice is the ultimate goal. Sadly, none exists yet,[3] but a descriptive approximation to the analytical theory can be sketched. It contains three parts:

1) a delineation of diagnostic features of a task,
2) a set of organizational structures and control regimes that can cope with these features, and
3) a way of measuring how well the organization performs.

In the descriptive theories developed by organization theory, two features have received major attention. The first is complexity. Complexity is defined as excessive demands on rationality. That is, task requirements exceed current bounds on computational capacity. For example, a manager receives more information than he or she can possibly read, or must coordinate more workers than he possibly can. Three types of complexity are described in the following sections: information, task, and coordination.

The second feature is uncertainty. Galbraith [6] in his explication of the contingency theory[4] approach to organization, states that organization structure is dependent upon uncertainity and diversity of task. Uncertainity is defined as the difference between information available and the information necessary to make the best decision. Variation in organizational structure results from diverse attempts in reducing uncertainty.

Contingency theory design strategies are based on the manifestation of uncertain information. We depart from the contingency view by defining three additional types of uncertainty. First, uncertainity can manifest itself in information. This means that the correctness of the information can be represented by a certainty measure;[5] the organization may not fully believe the information (stimuli) it perceives. For example, the correctness of a survey of consumer desires concerning a product the company produces is always under scrutiny. The second manifestation of uncertainty is found in the algorithm. No matter how certain the information a decision is based on, the decision itself may be uncertain due to knowledge lacking in the possible outcomes of the decision. Optimal decisions based on analysis of possible outcomes have been extensively studied [1].

[3] Systems science [31], computer science [32], [33] artificial intelligence [22], [34], and decision theory [1] have done some work towards this ideal.

[4] Contingency theory has two premises: 1) there is no best way to organize, and 2) all ways of organizing are not equally effective.

[5] Lesser *et al.* [34] propose a measure called reliability.

The contingency theory interpretation of information uncertainty is that at each state in the organization all information is assumed true (e.g., how much material is around, how many machines broken, etc.) but is a poor predictor of future states. Hence the information is uncertain in the sense that it is a poor estimator of environmental change. The focus is actually the changing environment and not the information. We call this environmental uncertainty. The environment changes over time (from state to state), and the organization must adapt to these changes in state. A fourth type of uncertainty commonly found in organizations is behavioral uncertainty. An employee, unit (department), or another organizations cannot always be depended on to produce the contracted for products or services.

The second part of the theory would associate organization structures and control regimes that can adequately cope with the complexity and uncertainty features of the task. Such associations are described in the next two sections.

Finally, the theory requires a way of measuring how well an organization performs its task. The task goal provides the measurement criteria. Goals such as minimizing resource consumption, maximizing production, minimizing quality, etc., are commonly found in business organizations. These are analogous to reducing time and space increasing processing power, and producing results o higher certainty (validity) in distributed systems. One measurement technique that has experienced renewed interes in management science is transaction analysis [7]. Transactions take on a rather broad definition. They encompas normal contractual agreements, communication of information, monitoring, delegation and control, and most othe activities that require interaction among participants withi an organization or market. The handling of transaction requires the consumption of resources. Hence transaction are too complex when they require more resources than ar available (bounded rationality); complexity reduction be comes the problem of minimizing resource consumptio The distributed system analog is that processing is resourc limited. Memory, cycles, and bandwidth are limited. Henc transactions among processes have to be well-structured t minimize resource consumption. Transactions can also b characterized by the assumed differences in informatio motivation, and behavior amongst the parties of the tran action. By detailing the transactions among organizatic participants, the efficacy of alternative organization structures can be measured.

In the next two sections, the second part of the theory–associating structure and control with task features—elaborated. A variety of complexity and uncertainty typ and organizations for reducing their effects are describe Where possible, a transactional approach for measurir uncertainty and complexity is defined. And the suggested organizations are interpreted in the Hearsay-II architectural framework.

## V. Complexity Reduction Techniques

### A. *Information Complexity*

Bounded rationality limits the amount of information a human or processor may process within a given time period (or other resource constraint). Information becomes too complex when it requires more processing than available in order to be properly analyzed and understood. Ways must be found to reduce the complexity of information so that humans and processors can be more effective.

Recognizing information complexity requires the characterizing of information flow as transactions. For each communication channel the average number of transactions and processing per transaction must be compiled and compared to the processing power of the receiving person (or module).

How is the information complexity reduced? By abstraction and omission. Abstraction is attained by use of several levels of representation. A purchasing manager does not care how much material is used every minute in a department, or how it is used. He is interested only in gross usage of materials. Material/minute is a detail that can be abstracted into material/day or material/week, while its usage can be ignored completely, since the former is actually an abstraction of it. A second approach to information reduction is computer-based summarization techniques. Whether statistical or graphical in nature, information can be reduced to a few meaningful parameters.

Information complexity, due to a processor's bounded rationality, also exists in distributed systems. Accordingly, similar techniques have been used. Examples of these techniques can be found in the Hearsay-II and Baseball [21] systems. In Hearsay-II information is represented at many levels (Fig. 4). Each level is an abstraction of the lower levels. A KS that makes a decision at the syntax level only uses information at the syntax and lexical levels, without going into the more detailed levels. To do so would require a greater information processing capacity and the ability to interpret information at those levels. To analyze information complexity, each hypothesis would be considered a transaction. Then knowing the average number of hypotheses at each level, the average number of hypotheses produced by a KS, and the average number of resources consumed by a KS when processing a hypothesis, complexity limitations can be calculated. (See [22] for more detail.) The concept of omission requires that a module's task be analyzed to deduce what information must be communicated, and what should be stored or hidden locally and communicated on demand.

### B. *Task Complexity*

Task complexity is concerned with the volume of actions (disjoint or coupled)

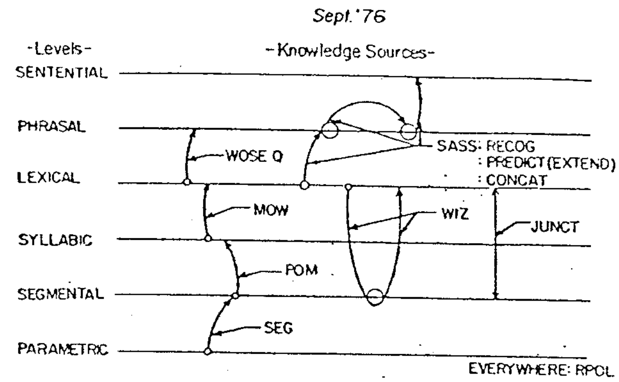


Fig. 4. Levels of representation.

volume exceeds a manager's ability to grasp the tas "gestalt," then this complexity must be reduced. The so tion to this problem is the division of labor [5]. T requires the partitioning of resources (men and materia modules and computer resources) into units. Each unit assigned a specific task related to the organizational gc The manager delegates jobs to these units, viewing them primitives[6] in the organizational plan. Each unit then int prets the control instructions and expands upon them control the primitives within the unit.

The encapsulation of both mechanism and informati is primary to the proper structuring of an organization. is necessitated by bounded rationality. The following scribes the characteristics of units that satisfy the cc straints imposed by bounded rationality.

1) View the numerous actions (programs) contained a unit as a single action (abstraction).
2) Control units as if they were primitive actions (pl ning in abstraction spaces).
3) Delegate authority. Commands to a unit elaborated by and within the unit.
4) Reduce information flow. Information within a u can be summarized by the unit.
5) Hide detail. Information and control not needed other units is hidden within the unit.

Interpreting the above constraints in the design of distr uted systems requires that

1) the products of the process must be well defined;
2) the interaction between processes must be minin (near decomposability);
3) the effects of a process upon other processes must understood;
4) clear lines of authority must be recognized;
5) clear lines of information flow must be recognize

The first interprocess constraint is a minimum requi ment. We must know what a process produces before it c be used. The second reduces the control complexity of t organization which may reduce its effectiveness. The thi is necessary so that one process's action will not undermi

[6]Viewing units as primitives is another example of an abstraction.

another process's. The last two points insure proper control and the information upon which to base that control.

The architecture of Hearsay-II conforms to many of these requirements. A KS is a unit that can be viewed as a single action. In Hearsay-II stimulus-response frames are used to describe the effect of a KS as a single action to the scheduler. A KS hides its algorithm and temporary results from other KS's and restricts its communication to a few hypotheses on the blackboard. The blackboard structure mechanism clearly defines how KS's communicate and how they interact (simply posting an hypothesis may cause another KS to react).

### C. Coordination Complexity

Task and coordination complexity are strongly linked. Once a task has been decomposed to a point where it is comprehensible, coordination must be considered. If the units cooperate in the completion of a task, then, there is usually resource dependence between them, i.e., information, partial products, etc. The actions of each unit must be coordinated so that each produces the proper resource at the proper time.

At present there are few heuristics that guide the division process so that coordination complexity is reduced. One of these is the definition of near decomposability of a system [23] which implicitly appears in the contingency theory approach to design: construct the units so that the interaction between units is minimal.

Analyzing coordination complexity is similar to analyzing information complexity. The set of transactions necessary to carry out a task are quantified and compared to the controlling person or module's limited resources.

It has been thought that task size (space and time) can be overcome by adding more processors and memory. Such simplistic views ignore the problem of dependence in task decomposition. Empirical and analytical results on multiprocessor systems such as C.mmp [24] and CM* [25] have shown that linear speedups are not always attainable [26], [27]. In some problem situations there is an upper bound on the effect of added resources. Conversely, linear speedups can be obtained if algorithms and information storage are carefully analyzed and properly structured. Consequently, greater attention to task decomposition in program organizations and problem solving must be paid.

Following is a set of organizational substructures that reduce the complexity of coordination.

*1) Slack Resources:* One aspect of coordination complexity is the coordination of coupled tasks. Tasks are coupled when the input of one depends on the output of another. Tasks are tightly coupled when state changes in one task immediately affect the state of another task. To reduce the tightness of the coupling, slack resources are introduced. Buffer inventories are inserted between coupled tasks so that if one task has something go wrong with it, the other tasks are not immediately affected.

Two interpretations of slack in distributed systems are 1) the replication of tasks (processes, modules) on alternate processors in case of a processor failure, and 2) the replacement of procedure calls by message queues. Requests and messages to a task are placed in a queue to reduce the synchronization (tight coupling) of tasks.

Slack has also been used extensively in computer hardware. The reliability of hardware has been extended by the duplication of functional units. Space vehicles, for example, duplicate essential units. Slack appears in a number of ways in Hearsay-II. KS's can be duplicated on many processors so that the same task can be carried in parallel. Secondly, the creation and placement of multiple hypotheses on the blackboard can be viewed as creating a buffer inventory. A KS has many hypotheses to work on and can continue processing without being aware of what another KS is doing or is not doing.

*2) Function Versus Product Division:* The coupling of tasks can also be reduced by proper decomposition. Organization theory distinguishes between two types of organization partitioning. The first is a product or self-contained division. This division requires that units be centered around the product that is to be produced by the organization. The second type is a functional division. A functional division orients the units to the functions necessary to produce the products (e.g., purchasing, marketing, materials, etc.). Why do we have these alternate forms of division? Depending on characteristics of the problem being solved by the organization (e.g., producing a plane), one division reduces complexity while the other increases complexity. An important measure of complexity is the amount of coordination. Any division of a problem assumes that there is greater interaction within a unit than between units (interaction locality). A system that exhibits interaction locality is called a nearly decomposable system [23]. When interaction locality no longer exists, the coordination of units becomes too complex.

An example of the division methodology is the restructuring of an operating system. The creation of scheduler, memory manager, I/O controller, and file manager modules is a functional decomposition. Each job interacts with the modules and indirectly with the other jobs. Thus coordination between jobs is important and time-consuming. If the machine was restricted to Basic and APL programming, the generality would not be necessary. The operating system could be divided into two modules, an APL module and a Basic module. Each would contain a memory manager, file manager, etc., plus certain physical resources such as disk, a portion or main memory, etc. Coordination would concern resources shared by both modules only. Economy of specialization and reduced coordination is achieved. Hearsay-II is a functional decomposition. Each KS embodies a function required for each level of speech analysis.

*3) Contracting:* In an organization there may exist functions that are too costly to carry out. This cost may be due to

- lack of experience within the organization,
- small usage, hence economy of size is not afforded,
- coordination problems,
- information processing problems.

It is simpler for the organization to contract for this service in the marketplace, hence information is reduced to a single price, control to contractual terms. Conditions exist under which contracting is not achievable; this is usually due to the idiosyncratic nature of the job. Williamson views most positions in organizations as being idiosyncratic. Though a position may be characterized by a general job classification, the organization, methods of communication, people interacted with, on the job learning, etc., make positions idiosyncratic. A primary consequence is that the cost of replacing a person is not negligible, nor is the service easily contracted for in the marketplace.

It is unclear how transaction analysis would detect situations where contracting is applicable. A simple approach would be to tag all transactions (communication, control, resources, etc.) with the task that produces and consumes them. Tasks whose transaction costs exceed some function of the market cost for the same product or service should be contracted.

The problem of market versus hierarchical organizations is important when more than one task is competing for the same resources (element, memory, etc.). Hence ownership (employment contract) of a processor and associated process by one task is uneconomical if it is under-utilized while external demand is high. Greater utilization is afforded when the element is a free agent, able to contract services to many different tasks.

### D. *The Complexity Shift*

It is clear that complexity has a definite effect on organiation structure. The cause of the transition from single erson to group to simple hierarchy was ascribed to capacty excesses stemming from bounded rationality. As tasks ecome larger, processing capacities are exceeded, requirng task decomposition and allocation. Groups fail when nformation sharing exceeds capacity (resources), and deisionmaking and coordination becomes too complex. Simle and uniform hierarchies utilizing information reduction echniques and authority relations appear. Uniform ierarchies usually suffice until the organization grows to a oint of producing multiple products. As a result, interacon among modules reaches a thrashing level. This is due ) competition among organization participants (departents, people, modules) for other participant's attention .g., competition among different products for the same sources). The multidivision hierarchy deals with this oblem by decomposing the organization along product es so that key departments (modules) are duplicated for ch product division. The next step in reducing informaon and control complexity is to contract for products and rvices in the marketplace. Information is reduced to a gle variable: price.

## VI. Uncertainty Reduction Techniques

In the contingency theory approach to organization action analysis [7] also focuses on uncertainty in transactions as a cause of organizational change. The emphasis in the former is information uncertainty, while in the latter it is behavioral uncertainty. Because of the fuzziness of the definition of these two types, Section IV introduced four types of uncertainty: information, algorithm, environment, and behavior. A detailed discussion of each of the uncertainty types and their effects can be found in [22], [10]. In following, we analyze how uncertainty affects the structure of organizations.

Information uncertainty can be reduced using verification techniques such as synthesis or prediction. To reduce or recover from algorithm uncertainty distributed systems must maintain multiple approaches to achieving a task, e.g., multiple hypotheses in Hearsay-II and relaxation techniques [37], or formulate new programs of action: data gathered, programs constructed and executed, results evaluated and program repaired (feedback). Some work in the latter approach has been done in artificial intelligence [38], but much research remains.

It is often the case that the existence of information, algorithm, and/or environment uncertainty results in behavioral uncertainty such as opportunism. Opportunism occurs when a party in a transaction takes advantage by making self-disbelieved threats or promises, or withholds information. The opportunistic party secures a contract that is less favorable to the other party than might be obtained otherwise. The following are factors that lead to behavioral uncertainty.

Information impactedness [7] is a differential of information between parties of a transaction. Impactedness may be due to bounded rationality considerations because of the amount of information, unavailability of information due to one party's inability to communicate, or a party's deliberate hiding of information. Impactedness would be of little concern if the cost of achieving parity was not prohibitive in most cases. Information impactedness is a recurring condition for opportunistic behavior.

Small numbers is a market condition where the number of market participants is small, circumventing the marginal pricing behavior of competition. Contracting under a small numbers condition may result in opportunistic behavior due to participant's lack of competitive pressure.

First-mover opportunity occurs when a person or organization has idiosyncratic knowledge of a particular function, which is unattainable (due to cost and information impactedness) by other market participants. As a result, a small numbers market condition results, enabling opportunistic behavior. A first-mover condition can appear when a person in an organization attains idiosyncratic knowledge of his particular job, or an initial contractor attains idiosyncratic knowledge of the contracted job. In subsequent contracting for the same or similar job, or searches for new personnel, the previous person or contractor has a considerable advantage due to their superior (idiosyncratic) knowledge.

Whenever uncer...

tion is, "What organizational structures facilitate the adaptation process?" In contrast to complexity, uncertainty, especially environmental and behavioral, appears to shift organization structure in the opposite direction, from heterarchy to hierarchy.

The smooth operation of markets requires marginal pricing, little uncertainty, and no opportunistic bargaining in small-number situations. Market uncertainty may invalidate contracts before they expire: materials may not be available, prices increase, a strike occurs, etc. Insuring uncertain events requires complex contingent contracts, but at some point accounting for all possible contingences becomes too costly. Bounded rationality also limits contingent contracts. Contract writers cannot foresee all possible contingencies or assimilate the necessary information to do so. Hence uncertainty in market environments must be small.

Reducing information to a single signal, price, results in information impactedness which opens the door for opportunistic behavior. Opportunism will succeed only if a small-numbers bargaining situation obtains. Barring that, competition among market participants should attenuate opportunistic behavior since no advantage can be attained.

Long-term contracts in the marketplace are not feasible due to uncertainty and bounded rationality. Under such conditions it is better to make spot (short-term) contracts. Spot contracting allows organizations to sequentially adapt to changing environments. This reduces the information and certainty required in writing a contract. A problem immediately occurs negating the satisfiability of spot contracting. That is, the initial contractor obtains first-mover advantages resulting in small-numbers bargaining.

As uncertainty increases in the marketplace, bounded rationality reduces the market's ability to contract accordingly, and as opportunism appears in first-mover advantages, an alternate form of organization must appear. Such a form is the collective organization. Market participants are integrated into a single organization to cooperate in achieving a single goal. Hence bargaining costs are lowered and idiosyncratic tasks are undertaken without risk of exploitation. A collective approach allows greater adaptability to uncertain environments since formal contracts do not exist and the collective jointly decides in a sequential fashion how best to adapt to the current situation. Opportunism does not appear since it is a joint venture. Information impactness is reduced by sharing information among the group.

The transition from the market (as typified in the above structures) to a hierarchical organization has been called vertical integration. Vertical integration is the process of adding to a hierarchical organization a product or service that was originally contracted for in the market. Uncertainty, information impactness, opportunism, etc., are the attributes in the market that have to be considered when analyzing the cost of contracting versus integrating into the firm.

As uncertainty increases, the transition from a heterarchy to a hierarchy becomes preferable. The attributes of a hierarchy that support this are [7, p. 40] as follows.

1) In circumstances where complex contingent claim contracts are infeasible and sequential spot market are hazardous, internal organization facilitates adap tive sequential decisionmaking, thereby economizin on bounded rationality.
2) Faced with present or prospective small-numbe exchange relations, internal organization serves t attenuate opportunism.
3) Convergent expectations are promoted, whic reduces uncertainty.
4) Conditions of information impactedness are mo easily overcome and, when they appear, are le likely to give rise to strategic behavior.
5) A more satisfying trading atmosphere sometim occurs.

More importantly, hierarchies are not bound to particul courses of action. Due to the employment relati employees have with the firm and the firm's ability control resource allocation, unexpected situations can dynamically adapted to: employees assigned new tasks a resources assigned to different products. This flexibility not typically available in contracted relations. Finally, t type of hierarchy created is dependent upon size a number of products produced, which in turn determi the organization's complexity.

The problem of behavioral uncertainty has not be considered in computer systems. Such a problem can oc in hierarchical organizations but is more prevalent market situations. The question can be asked whet distributed systems will ever be organized as markets. Su a possibility is near.

Recent work has demonstrated the feasibility of co puter networks supporting a market organization. example, assume that each host on an ARPANET- network is a separate organization offering services products to other hosts. Each host has a software pack that allows it to decide what services to acquire and under a variety of conditions. In their discussion of h level protocols for networks, Sproull and Cohen [28] scribe a network plotter protocol (NPP), a language describing graphics plotting tasks which allows ma participants to communicate about plotting tasks. work of Smith [29] defines a protocol for contrac among modules. Hence the mode of market interact contracting and bargaining, and the language for desc ing the task, NPP, have been created for the task contracting for plotter printings by market participant a commercial network (nonsubsidized), the price h charge for their services must reflect the costs of main ing them, and if the same service is provided by more one host (e.g., color graphics output, Illiac-4-like servi competition may occur. Hence the equivalent of a ma organization will result. Problems of complexity and un tainty can arise, requiring the application of the reduc techniques described in the previous sections.

Reduced communication in markets (i.e., price) implies luced information inviting information impactedness nditions. If modules can be hard-wired to share the same al, then opportunistic advantages due to an information balance will not occur, but when modules have the edom to choose when and where to do processing, :thods of prerating module performance (e.g., similar to :redit agency) and monitoring of module performance : required to reduce opportunistic behavior. Transaction alysis suggests the latter can be too costly in the market, nce requiring the integration of a market function into organization.

## VII. Conclusion

:arly computer programs bore little resemblance to hu- n organizations, but as the problems attacked grew in :, resource limitations appeared and prevented the cess of programmed solutions. Resource limitations can viewed as the cause of bounded rationality whose effects ear in programs as in businesses. One can view the k in artificial intelligence as attempts circumvent re- rce limitations. Hence it is not surprising to find pro- ms that (attempt to) exhibit "intelligence" that also play characteristics of human organizations. Systems h as Hearsay-II have shown trends that reflect human anizations and human problem solving methods (also [17], [35], [3]). These trends have resulted in modules : contain problem solving characteristics similar to hu- ıs.

n this survey of organization theory ideas, we have nd that some of the solutions to organizational prob- s have already been discovered and used in computer ems, and some have not. What organization theory s contribute is another way of looking at distributed ems. That complexity and uncertainty are two im- tant factors in deciding how to structure an organiza- . It appears that complexity and uncertainty are two osing forces; complexity forcing a distribution of task nately resulting in a heterarchical structure; uncer- ty pushing in the opposite direction, vertically integrat- tasks into a more hierarchical structure (Fig. 5). By zing this view, the distributed system designer has an ınization tool kit that contains structures and control mes that reduce the effect of the particular factor. It provides a pointer to transaction analysis—which at time is still poorly developed—as a tool for recogniz- these factors.

n important view underlying this paper is that mod- , processes, and tasks act like humans in an organiza- . Consequently, the problem of motivation, a cause of ıvioral uncertainty, must be considered. Little attention been paid to motivation in computer programs, but the :nt of the module approach to constructing complex ;rams and the limitations on resources (e.g., processors) necessitate a module's ability to decide when and what lems to work on. Once self-motiviation appears in a

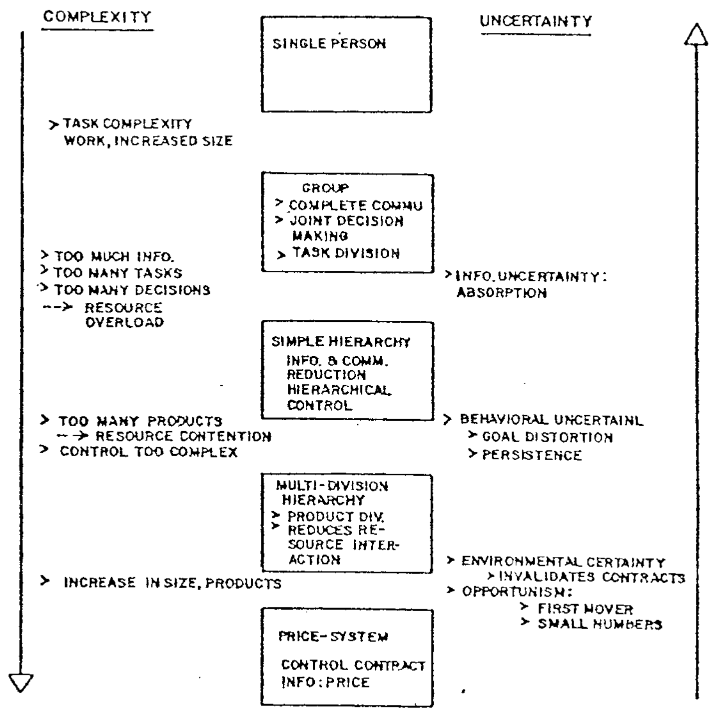


Fig. 5. Organization structure continuum.

This paper lays part of the groundwork for constructing a theory of distributed system organization design. The approaches described by organization theory are interesting and useful but not rigorous. Better methods of measuring complexity and uncertainty must be found. Whether these measures will be derived from organization theory, system science, or computer science remains to be seen.

## Acknowledgment

The author wishes to thank Herb Simon for his introduction to and guidance in organization theory. I would also like to thank Victor Lesser, Karsten Schwans, Reid Smith, Dan Corkhill, Raul Medina-Mora, Rich Korf, B. Chandrasekaran, and the referees for their comments on various aspects of this work.

## References

[1] J. Marschak and R. Radner, *Economic Theory of Teams*. New Haven, CT: Yale Univ. Press, 1972.
[2] DSN, *Distributed Sensor Nets*, Workshop Proc. Comput. Sci. Dep., Carnegie-Mellon Univ., Pittsburgh, PA, Dec. 1978.
[3] V. R. Lesser and D. D. Corkhill, "Cooperative distributed problem-solving: A new approach for structuring distributed systems," Tech. Rep., Dep. Comput. Inform. Sci., Univ. Massachusetts at Amherst, Amherst, MA.
[4] J. G. March, H. A. Simon, and H. Guetzkow, *Organizations*. New York: Wiley, 1958.
[5] A. Smith, *Wealth of Nations*, 1776.
[6] J. Galbraith, *Designing Complex Organizations*. Reading, MA: Addison-Wesley, 1973.
[7] O. E. Williamson, *Markets and Hierarchies: A Transactional and Antitrust Analysis of the Firm*. New York: Free Press, 1975.
[8] H. A. Simon, *Models of Man*. New York: Wiley, 1957.
[9] L. D. Erman, and V. R. Lesser, "The Hearsay-II system: A tutorial," in *Trends in Speech Recognition*, W. A. Lea, Ed. Englewood Cliffs

[10] L. D. Erman, F. Hayes-Roth, V. R. Lesser, and D. R. Reddy, "The Hearsay-II speech understanding system: Integrating knowledge to resolve uncertainty," Computing Surveys, vol. 12, pp. 213–254, 1980.

[11] F. Hayes-Roth, G. Gill, and D. J. Mostow, "Discourse analysis and task performance in the Hearsay-II speech understanding system," in *Speech Understanding Systems: Summary of Results of the Five-Year Research Effort*, Dep. Comput. Sci., Carnegie-Mellon Univ., Pittsburgh, PA, 1977.

[12] F. Hayes-Roth, D. J. Mostow, and M. S. Fox, "Understanding speech in the Hearsay-II system," in *Speech Communication with Computers*, L. Bolc, Ed. Berlin: Springer-Verlag, 1978.

[13] A. R. Smith, "Word hypothesization in the Hearsay-II speech system," in *Proc. 1976 IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 578–581, 1976.

[14] M. S. Fox and D. J. Mostow, "Maximal consistent interpretations of errorful data in hierarchically modelled domains," presented at the 5th Int. Joint Conf. Artificial Intelligence, Cambridge, MA, 1977.

[15] A. Newell, "Heuristic programming: Ill-structured problems," in *Progress in Operations Research 3*, J. Aronofsky, Ed. New York: Wiley, 1969, pp. 360–414.

[16] F. Hayes-Roth and V. R. Lesser, "Focus of attention in a distributed logic speech understanding system," in *Proc. 1976 IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 1976, pp. 416–420.

[17] D. Lenat, "Beings: Knowledge as interacting experts," in *Proc. 4th Int. Joint Conf. Artificial Intelligence*, 1975.

[18] P. Nii and E. Feigenbaum, "Rule based understanding of signals," in *Pattern-Directed Inference Systems*, D. Waterman and F. Hayes-Roth, Eds. New York: Academic, 1977.

[19] M. S. Fox, "Knowledge structuring: Knowledge accommodation and discovery using specialization and planning," Ph.D. thesis, Comput. Sci. Dep., Carnegie-Mellon Univ., Pittsburgh, PA (to appear, 1980).

[20] V. R. Lesser and L. D. Erman, "An experiment in distributed interpretation," in *Proc. 1st Int. Conf. Distributed Computing Systems*, 1979, pp. 553–571.

[21] E. Soloway and E. Riseman, "Levels of pattern description in learning," presented at the Int. Joint Conf. Artificial Intelligence, Cambridge, MA, Aug. 1977.

[22] M. S. Fox, "Organization structuring: Designing large, complex software," Comput. Sci. Dep., Carnegie-Mellon Univ., Pittsburgh, PA, Tech. Rep. 1979.

[23] H. A. Simon, "The architecture of complexity," Proc. Amer. Phil. Soc., vol. 106, pp. 467–487, 1962. (Also in Simon 1968.)

[24] S. H. Fuller and S. P. Harbison, "The C.mmp multiprocessor," Comput. Sci. Dep., Carnegie-Mellon Univ., Pittsburgh, PA, Tech. Rep. 1978.

[25] R. Swan, S. H. Fuller, and D. P. Siewiorek, "Cm*: A modular, multi-processor," in *Proc. Nat. Computer Conf.*, 1977.

[26] P. Oleinick and S. H. Fuller, "The implementation and evaluation of a parallel algorithm on C.mmp." Comput. Sci. Dep., Carnegie-Mellon Univ., Pittsburgh, PA, Tech. Rep. 1978.

[27] L. Raskin, "Performance evaluation of multiple processor systems," Ph.D. thesis, Comput. Sci. Dep., Carnegie-Mellon Univ., Pittsburgh, PA, Tech. Rep., 1978.

[28] R. F. Sproull, and D. Cohen, "High-level protocols," *Proc. IEEE*, vol. 66, Nov. 1978.

[29] R. G. Smith, "A framework for problem solving in a distributed environment," Ph.D. thesis, Comput. Sci. Dep., Stanford Univ., Stanford, CA, Memo HPP-78-28, 1978.

[30] D. R. Reddy, "Speech recognition by machine: A review," *Proc. IEEE*, vol. 64, 1976.

[31] M. D. Mesarovic, D. Macko, and Y. Takahara, *Theory of Hierarchical Multi-level Systems*. New York: Academic, 1970.

[32] P. N. Chanon, "On a measure of program structure," Ph.D. thesis Comput. Sci. Dep., Carnegie-Mellon Univ., Pittsburgh, PA, 1974.

[33] C. L. McClure, "A model for program complexity analysis," in *Proc. 3rd Int. Conf. Software Engineering*, 1978.

[34] V. R. Lesser, S. Reed, and J. Pavlin, "Quantifying and simulating the behavior of knowledge-based interpretation systems," in *Proc 1st Annu. Nat. Conf. Artificial Intelligence*, pp. 111–115.

[35] C. Hewitt, "The ACTOR formalism," in *Proc. 3rd Int. Joint Conf Artificial Intelligence*, 1973.

[36] B. Chandrasekaran, F. Gomez, S. Mittal, and J. Smith, "An approach to medical diagnosis based on conceptual structures," in *Proc. 6th Int. Joint Conf. Artificial Intelligence*, 1979, pp. 134–142

[37] S. W. Zucker, "Relaxation labelling and the reduction of local ambiguities," in *Pattern Recognition and Artificial Intelligence* C. H. Chen, Ed. New York: Academic, 1976.

[38] G. Sussman, *A Computational Model of Skill Acquisition*. New York: American Elsevier, 1975.