RESEARCH PROPOSAL

COMPUTER LABORATORY FOR CLINICAL DECISION-MAKING

G.A. GORRY

MARCH 25, 1974

Introduction

As Schwartz has noted (1):

"Many discussions during the past decade have considered the use of computers as an adjunct to medicine. Few, however, have fully explored the possibility that the computer as an intellectual tool can reshape the present system of health care, fundamentally alter the role of the physician, and profoundly change the nature of medical manpower recruitment and medical education -- in short, the possibility that the health-care system by the year 2000 will be basically different from what it is today.

"Much has, of course, already been said about the role of the computer in improving the efficiency of the health-care system. These now familiar projections envision the computer performing a wide variety of functions such as the scheduling of hospital admissions, the keeping of medical records and the operation of laboratory and pharmacy. Such developments in the area of "house-keeping" activities offer considerable hope for the improvement of both hospital and outpatient operations but do not come to grips with the more fundamental problems of the health-care system -- the increasing shortage of physician manpower and the geographic maldistribution resulting from the reluctance of today's doctor to practice in rural or depressed urban communities. Even less do they give hope of dealing with the difficult challenge of maintaining a high level of physician competence in the face of a continued expansion of medical knowledge that tends to widen progressively the gap between what a doctor should know and what he can retain and utilize. The computer thus remains (in the light of conventional projections) as an adjunct to the present system, serving a palliative function but not really solving the major problems inherent in that system. There is, in fact, little reason to believe that any of the current proposals for solving these problems, technologic or other, will do more than mitigate their severity".

One radical and intriguing possibility for improving the efficiency and effectiveness of the health care system is to use the computer as an "intellectual" or "deductive" instrument -- a consultant that is built into the very structure of the health care system and augments the abilities of physicians and paramedical personnel. Clearly, however, considerable intellectual and technological resources must be marshalled and a long term research commitment must be made if this possibility is to be realized.

We will argue in the body of this proposal, that the principal impediment to the realization of this exciting prospect is the lack of a good theory of clinical cognition. Despite successes in certain areas of clinical medicine, no theory of clinical decision-making has been developed which can explain the richness of the problem-solving behavior of

> ***experts. Further, we will argue that the computer is the key to the development of such a theory. The computer provides an environment in which ideas about process can be expressed in a quite natural way. Such environment is essential if we are to advance our understanding of clinical cognition.***

Although the idea of using computers and computer programs in the development of cognitive theories is not new (2), recent developments in computer science and technology make this idea more powerful. We have organized a team of computer scientists and medical scientists in a concerted attack on the problem of understanding clinical decision-making in new and profound ways. The Computer Laboratory concept is one which fits well into our current activities, and indeed, it offers us real leverage with respect to the growth of our efforts.

<u>What We Propose To Do</u>

Various approaches to the problems of automating processes for clinical decision-making have been employed by researchers in the field, and considerable success has been achieved. We believe that an expert program which can deliver advice and consultation with respect to serious clinical problems will make use of many of these approaches. At present, however, none of these approaches is sufficiently powerful to offer the <u>integrative</u> or <u>administrative</u> capability required to organize the variety of problem solving approaches necessary for the full range of clinical problems. Thus while other researchers continue with the development and refinement of existing techniques, we propose to devote our efforts to the problem of defining and implementing the framework within which these techniques can be organized and controlled.

The only examples we have of the integrative abilities which are required come from the performance of clinical experts. Clearly they possess the administrative problem-solving knowledge to shift from one approach to another as the case merits. For this reason, the principal focus of our efforts will be on gaining a better understanding of the behavior of experts.

We propose to undertake a program of research which will result in a new and significantly better theory of clinical cognition, with special emphasis on the administrative aspects of the problem-solving behavior. The computer will play a central role in the formulation and testing of this theory. Further, because the concepts upon which the theory will be based will be expressed in a form which is programmable, we will have a new technological framework within which efforts to create <u>distributable</u> <u>expertise</u> can proceed in concert. This in turn will speed the realization of that revolutionary role of the computer in the health-care system suggested above.

The activities of the Laboratory initially will be centered on several specific research projects which are related to our overall goal. These projects are discussed in detail in the body of this request. Here we will simply mention them and their relation to our primary goal.

1) Taking the Present Illness

The present illness is the initial point of contact between the patient and the physician, and for this reason, it represents a logical starting point. More importantly, however, the cognitive demands of taking the present illness, establishing the facts, drawing inferences about the facts and about the patient, dealing with discrepant information and uncertainty,. etc. are central to all clinical decision-making.

One of our major projects will be to develop a computer simulation of an expert taking the present illness. Such a simulation will be based on specific mechanisms for solving the various cognitive problems involved. These mechanisms, in turn, will be central to a variety of other decision-making programs. The knowledge gained from this effort and the results of the next project discussed will allow us to attack the problems of differential diagnosis and the risk/benefit analysis of management.

2) The Formalization of Clinical Knowledge

A second major project of the Laboratory will involve the development of new ways to formalize medical knowledge. Initially, this knowledge will be primarily that which appears in texts or journal articles on clinical problems, augmented and refined by clinical experts.

The criteria by which proposed representations of this knowledge will be judged include:

a) clarity
b) parsimony
c) completeness
d) capacity for expressing relations among "pieces" of knowledge
e) the ease with which it can be assimilated by a computer

Loosely speaking, the present illness project can be said to be concerned with how knowledge is used, whereas this project is concerned with formalizing what knowledge is required.

The result of this effort will be a methodology for building a knowledge base for programs such as the cognitive simulation of the present illness, and it should be viewed as being intimately connected with that project. Further it should provide a basis for research on the construction of diagnostic and management programs for various problems by providing a framework within which the basic knowledge required can be organized. The Laboratory will also develop programs for diagnosis and management when a good base of understanding has been achieved.

3) Model-Based Decision-Making

There are a number of important areas of clinical medicine in which a formal (generally mathematical) model is available upon which certain diagnostic or management decisions could conceivably be based. In many of these cases, however, the model in question is of little clinical use. Although the model often surpasses the ability of even the best physician to deal with certain aspects of the problem, or with "classic" cases, it cannot cope with a variety of patient-specific factors which should be factored into the decisions, or certain emergency conditions which should cause a re-ordering of the priorities in the model. In general physicians understand how to alter and refine their approach to a problem in the light of such factors, but computer programs unfortunately remain very rigid in this regard.

If we look to the day when various models and techniques are combined in a single system, it is clear the new flexibility must be built into the component pieces so that they can be "tailored" to fit a certain situation, and so that component pieces work coherently under the same assumptions about the patient.

To achieve this aim, we need new ways to combine medical "common sense" with mathematical models. The models themselves must be represented in such a way as to allow this common sense to be applied. Hence it must be clear to some supervisory program what the basis for a particular model is, and how changes in assumptions about the patient affect this basis, and hence the model.

We will begin to investigate these problems in the context of a model for the administration of digitalis/digoxin. This problem is a good one, because the "best" strategy for any patient depends in part on the use of a model, and in part on a basic understanding the medical problems of the patient.

Some Recurrent Problems

There are several problems which arise in almost all phases of clinical decision-making, and these will be the focus of a continuing research activity of the Laboratory. We mention them separately here, but we want to emphasize that they really represent threads which run through all our work.

4) <u>Dealing with Discrepant Information</u>

One of the important problems in clinical medicine is the amount of discrepant information which must be dealt with. Some of this difficulty arises because patients are not always accurate observers of their symptoms, or because they wish to conceal facts from the physician. Other problems arise from errors in laboratory tests or medical records. In addition tnere are many problems in which the discrepancy is not <u>absolute</u>, but rather <u>relative</u> to some currently believed hypothesis about the patient.

The question of <u>belief</u> is thus central to clinical decision-making. We plan to study this problem in a variety of contexts, with the intention of answering such questions as:

How is the credibility of a piece of information established?
How are potential discrepancies among facts detected?
How are conflicts between facts resolved?
What strategies are employed to resolve ambiguities or discrepancies?

5) <u>The Representation of Time</u>

Time plays a key role in clinical medicine. Diseases and their manifestations evolve through time. The interpretation of facts is often affected by the place of these facts in time. Often time-based relationships are crucial in making diagnoses or management decisions.

If we are to capture clinical expertise in a machine, we must equip the machine with an understanding of time and events which take place in time. Thus the machine needs a minimal ability to place events and intervals on some form of "time-line", and to make appropriate deductions about this arrangement. But much more is required. For example, we must develop ways to capture the concept of <u>episodes</u>. The machine needs to understand such fragments as "the gradual onset of the disease" and "an abrupt cessation of symptoms".

This is an area where substantive progress can probably be of direct use to other researchers in the field who to date have employed rather <u>ad hoc</u> methods to solve the problems of time representation or who have had to skirt the issue entirely to the detriment of their efforts.

6) <u>Inquiry and Explanation</u>

Another area in which we will be working is the development of mechanisms which allow a user to employ a natural and direct mode of interaction with a program and will allow a program the ability to explain its behavior in terms which are readily understandable to a clinician.

To a large extent, we will rely on research and development of natural language capabilities by others, in particular some of our colleagues at M.I.T, but we will play an active role in adapting their work to the medical context.

We will play a more central role in the development of the technology which will allow a program to generate explanations. Such explanations may be based on a variety of principles such as the use of physiological models. The point is that such a capability must be developed to meet several needs:

a) the need for users to understand the basis for a program's advice, particularly when the clinical problem is a serious one.
b) the need for clinicians working in our group to have access to facts and procedures used by the program in arriving at a particular conclusion.
c) the need for students to interrogate the program to learn about its strategies

Here again, progress in the development of these facilities, coupled with progress on our other projects should have an immediate and direct impact on the work of other researchers in the field, as well as a longer term impact on the delivery of health care.

<u>Summary</u>

In summary, we are proposing some projects which we believe will provide the proper direction for the Laboratory. The problems addressed by these projects are all basic problems for computer-aided clinical decision-making. Our emphasis on the study of clinical experts and on the use of the latest concepts of computer science to express the results of this study will provide a unifying theme for members of the Laboratory.

We have already formed a group of computer scientists, clinicians, and graduate students, which has begun work on these problems. The Laboratory would greatly facilitate and accelerate collaborative efforts of this kind, and it would be a link between the impressive computer science resources of M.I.T. and the equally impressive clinical resources of the Tufts-New England Medical Center. It would also provide a center into which researchers from other institutions could be drawn. In all, we envision that the Laboratory would be the center of new, vital, and important combinations of research and education. Its activities should have a significant impact on the computer-aided delivery of health care, as well as on medical education.

Background

The Laboratory we are proposing here will bring together experts from the computer sciences and from medicine for the purpose of gaining a new and deep understanding of the processes of clinical cognition and developing the mechanisms to translate this understanding into improvements in health care delivery. Here we want to give a brief history of the development of the research group, and then because of our involvment in both medicine and computer science, we want to briefly review important concepts and developments in both computer-aided clinical decision-making and in the relation of computer science to psychology and to theories of problem-solving.

The Development of Our Research Group

In order to put our application into perspective, we want to include a brief history of the development of the research group.

The nucleus of the group was formed several years ago, and it consisted of Drs. Schwartz and Kassirer and Professor Gorry. Schwartz and Kassirer had been working on the problem of encoding the protocols of experts in computer programs, and had developed a program for acid-base problems {1}. Gorry had developed a program which used statistical decision theory to solve diagnostic problems {3}. Because of the common interest in automating processes for clinical decision-making, the three joined forces.

The initial efforts of the group were directed along the lines suggested by the decision theory program. The work was considerably deepened and expanded during the two years following the initial formation of the group. A series of papers describing the work were published, most notably two recent articles ({4 } and {5}). These two papers consider in detail the application of decision analysis to clinical decision making, both insofar as the automation of the process is concerned, and with respect to the use of this formalism by clinicians.

Dr. Pauker joined the group in 1971, bringing to it a rare combination of expertise both in medicine and in computer science.

During the latter stages of our work on decision analysis, we began to see certain difficulties in using decision analysis as the sole basis for a system to deal with real problems of crisis medicine. After further definition of these difficulties, we were given a research grant from HISMA under which we explored these problems. From this exploration emerged a recognition of the need for a close cooperation with skilled computer scientists.

In order to promote a closer union between researchers in computer science and the workers in our group, we held last summer a week long conference on the problems of clinical decision-making and the relevance of advances in computer science to these problems. Attending the conference were five members of the M.I.T. computer science faculty (including Professor Marvin Minsky, the director, of the Artificial Intelligence Laboratory and Professor Edward Fredkin, Director of Project MAC) and the members of our group already mentioned. The major result of this conference was the recognition of the potential benefits to medicine of a strong computer science supported research program, and the complementary benefit to computer science of a close invólvment in medicine.

At this meeting, we resolved to organize a research program which would bring together first rate computer scientists and clinicians in a coordinated study of the problems of clinical decision-making. This proposal and the work upon which it is based is the result of that collaboration.

Since that meeting, we have been actively pursuing research in this area. We have funded our activities through small amounts of money from various sources. Despite this limitation of resources, however, we are proceeding at a rapid rate. In addition to the research discussed in this proposal, we are attracting graduate students in computer science. Five graduate students are already working with us, and we would have more if more funds were available.

Professor Gorry has joined the faculty of the Electrical Engineering Department at M.I.T. and is working at Project MAC. Professor Sussman of the Artificial Intelligence Laboratory is taking an active role in our research efforts, and other faculty, notably Professors Fredkin and Minsky are advising us and encouraging our efforts. Most notably, Dr. Schwartz will be a Visiting Professor at M.I.T. next year where he can devoted increased energy to the research program.

All this causes us to be very optimistic about our ability to mount an excellent program of research and education in computer science and medicine. The critical problem now is not the people or the ideas, but simply that we lack funds. Because our work seems so well in line with the intention of the Computer Laboratory Program, we hope to obtain the needed funds from that program.

Previous Research on Clinical Decision-Making by Computer

Broadly speaking, work on computer-aided clinical decision-making falls into two categories. In the first category are efforts to develop computer-based mechanisms for assuring orderly and complete acquisition of data concerning the patient. Examples of such efforts are Weed's problem-oriented approach {6} and work in history-taking, physical examination , and laboratory testing procedures (See, for example {7}.) It is believed that with improvements in the data acquisition and data structuring processes will come improvements in either the effectiveness or the efficiency of the clinical decision-making process, and in general, this belief seems well-founded.

In the second category fall all the efforts which are directed at developing computer realizations of procedures for making diagnostic and/or management decisions. In general, activities of this type have paid less attention to the orderly acquisition of facts than to the problems of interpreting the facts as presented. Within this category, however, a further division of efforts can be made. This division is based on the view which the researchers take of the decision-making procedures they are developing -- whether these are thought to be <u>descriptive</u> or <u>normative</u>. In the former case, the researchers have attempted to codify the way in which experts actually make diagnostic or therapeutic decisions. In most cases, the determination of exactly how an expert behaves has been rather <u>ad hoc</u>, involving a mix of introspection, interview, and various forms of observation. Some notable successes have been achieved in this way. {8} (Here we are measuring <u>success</u> in terms of providing <u>distributable</u> expertise about some problem domain.)

Those workers with a more normative bent have emphasized the development of models and procedures for decision-making which are thought (under certain assumptions) to be the basis for <u>optimal</u> decisions. In almost all cases, the assumptions are met only loosely, and no real claim of optimality can be made. Still, the general flavor of the work suggests that computers <u>ought</u> to make decisions in this way, without regard to the way in which humans make the same decisions. The more normative approach has also yielded success in certain areas (e.g. {3}, {5}, and {9})

Although work in both of these categories has shown considerable promise, and research continues actively on both approaches, no program has been produced which can cope with the real complexities of the clinical situation, e.g. time dependent changes in disease, multiple disease in the same patient, and a variety of patient specific factors which have an influence on both diagnostic and management strategies.

We believe that these approaches and the techniques which they have produced will enter into an expert system in an important way. We do not believe, however, that either of these approaches, as currently employed, can be the basis of the kind of administrative and integrative structure required in such an expert system.

For this reason, we want to explore in some detail the methodological limitations of the approaches which have been used to date. It should be remembered that our criticisms of these approaches are in the context of trying to provide an overall framework for clinical decision-making.

1) Flow Charting

The 'descriptive' approach is to construct a flow chart to represent the way in which a particular problem is to be handled (e.g., {7}, {8}). As was noted above, the manner in which the flow chart is obtained is usually ad hoc. Sometimes the flow chart represents the opinion of an expert as to the process he believes he uses. In other cases, it is based on a mixture of introspection and more formal modeling of aspects of physiology or pathophysiology. In any event, the resulting flow chart is an encoding of a decision procedure which is deemed to be a good one to follow in the particular clinical area in question.

There are two major difficulties with this approach insofar as complex clinical problems are concerned. First, a rigid definition of the logic to be used in a given situation may be impossibly cumbersome if it attempts to account for time dependencies, multiple interacting problems, patient specific constraints, etc. Even if such flow charts can be constructed for subproblems of a clinical problem, the decision as to how and when they should be combined, modified, and applied to a given situation remains. The representation of knowledge in flow charts makes this latter decision exceedingly difficult. Medical knowledge about a given clinical situation is implicit, not explicit in a decision flow chart. Because the reasons for a particular branching are not available to the program, in general it cannot make even simple deductions about them. Thus, unless the clinical situation matches exactly a series of branches in the flow chart, the program is helpless, because its lack of underlying knowledge prevents it from adjusting its approach to a non-standard problem.

Further, with this kind of structure, a user cannot inquire about the basis for a decision or suggestion from the program. And, an expert cannot add new knowledge to the program except through a laborious search through the programs or frames of the flow chart to ascertain what the program already knows a given subject, and how the new knowledge should be related to it.

2) Decision Analysis:

Another approach to the problem of computer-aided decision-making is to give a program an explicit description of the relations between findings and diseases and between actions and outcomes. Then one can incorporate an inference procedure into the program for sequentially deducing the path it should take with respect to a given problem. This approach is the basis for the decision analysis program we built for acute renal failure ({3} and {5}), and has been used by others in different contexts. (e.g., {9})

By explicitly recognizing the uncertainty in the relationships and by generating a decision tree for each new situation, a decision analysis program for balancing costs and benefits can deal with the equivalent of a very large number of flow charts.

This work has demonstrated that decision analysis is a very powerful approach to problems of balancing risks and benefits in the clinical context.

With this approach, however, there are limitations which pose very serious problems when real-world complexities are introduced. Our current methods for the explicit description of the probabilistic relationships, the courses of diseases, action-consequence relationships, etc. are very rigid and to a large degree, artificial, and although these forms of description are well-suited for the decision analysis algorithm, they are very cumbersome for the expression of medical facts in medical terms. Thus, a time-consuming and error-prone process must be undertaken to translate descriptive statements (made by experts, for example) into material which the program can use correctly.

A second problem is that it is very difficult to give procedural advice to a program based solely on decision analysis. For example, an expert might want to suggest a logical procedure (perhaps a "flow-chart") by which a specific situation can be efficiently and effectively handled. He may have processed (in some way) all the uncertainties, risks, and benefits associated with the situation, and he knows that the procedure is useful. He cannot, however, add the procedure to the program directly. The options are either to reprogram the system or to determine some parameters which, when used by the decision analysis program, cause it to do the "right" thing. Both alternatives are unsatisfactory if much knowledge is to be added to the program.

Finally, to the extent that explicit descriptions of diseases, etc., are formulated in terms of probabilities, the knowledge of the program is basically a mass of numbers, and the explanation of decisions or suggestions made by the program will be very difficult for an expert (and more so for the average user) to understand. Concepts and language naturally employed by the expert to express his knowledge have to be converted to a set of numbers which when coupled with some decision produced the same results.

> *To summarize, neither the flow chart approach nor decision analysis can be the basis for a program which deals with complex clinical diagnosis and management problems. Both approaches have value in certain circumstances and should be used as appropriate, but new techniques are required for a program to be able to deal with the full range of complexities which arise in serious clinical situations. Advances are also required if it is to be possible for an expert to interact with a program in such a way that the program can assimilate the expert's knowledge, and for a user of that program to be able to have natural and direct access to that portion of the knowledge which is most relevant to the clinical problem he is considering.*

The need for these innovations is underscored by the diversity of knowledge which experts used. They use descriptive, causal, procedural, and administrative knowledge along with common sense. It seems apparent that current formalisms are suited for only one or two types of knowledge, and that a new framework for organizing and using these diverse kinds of knowledge is required. More recent work, such as that of the Rutgers Special Research Resource on Computers in Biomedicine is directed to the solution of some of these problems. We hope that the proposed Laboratory would establish close relationships with such activities.

The Relevance of Advances in Computer Science

Advances in computer technology, including dramatic increases in information storage capacity and the development of remote access capabilities in the form of time-sharing systems, suggest the possibility mentioned above, that computers would serve as a repository for medical expertise and as a means for disseminating that expertise to points of need within the population. If such 'knowledge-based' systems could be built to serve as consultants for clinical problems, they could be replicated (either in fact, or effectively through multiple remote access to one system) as needed.

Unfortunately, this computer power alone is not enough to carry us to our goal. As we noted in the introduction, the major impediment to progress is our lack of understanding of the processes of clinical cognition. Therefore, advances in computer programming and technology, alone, will not solve the problems. It is important, however, to recognize the role which advanced computer science and technology play in research such as that being proposed here.

It is an unfortunate fact that although advances in computer science and technology cannot solve the problems, deficiencies in either can pose a serious hindrance to progress. Until recently, various attempts to formulate behavioral theories of complex processes would have suffered from a serious lack in the existing technology, the technology which had to be the testing ground for these theories. As a result, the development of theories of intelligence in certain domains was retarded.

In recent years, there has emerged from research in computer science a new 'technology' for representing some kinds of knowledge in computer systems. This capability is relatively new, dating from the late 1960's, and we believe that its availability will greatly ameliorate the problems of formulating and testing cognitive theories. This in turn will have a very beneficial effect on research into clinical decision-making. We are not claiming that there are no technological problems in our path; on the contrary, there are many. It is our opinion, however, that this new technology permits us to begin to explore new forms of procedures which simulate aspects of clinical cognition.

The advances and ideas to which we are referring are concerned with new techniques for programming computers and new techniques for representing knowledge and meanings in programs. In the old style of making 'computer models', things were very rigid. In the new style, it is much easier to include knowledge about how contingencies and side conditions affect, not only the states of the models, but especially how the models are to be applied in various situations. (Later we will describe some of our ongoing research in applying some of these ideas to the problem of digitalis/digoxin administration.)

In the new style, communication between programs is more flexible and direct. Some kinds of knowledge can be represented as procedures, able to intervene actively in the control of other programs when specified 'patterns' arise in the other programs' operations.

Goal-Directed Programming Languages

Rather than being organized as a step-by-step sequence of actions to be performed, specified in advance by the programmer, programs in these programming languages are controlled by the activation of certain statements called goals. When a goal is activated, the system retrieves from a data base of knowledge statements those that match the 'pattern' of the goal. (A pattern is a description of a state of affairs in a model, or an encoding of some fact about the world, etc.) These retrieved statements then serve as advice about what should be done to achieve the goal; they may dictate that a certain program be run, that the goal be replaced by one or more subgoals, or that certain priorities be re-arranged., and then control be returned to an earlier, superior goal system.

Understanding Natural Language

For twenty years, the public has been titillated by promises that computers would understand natural language and even translate from on language to another. A justifiable skepticism has resulted from such promises. Although progress in the theory of 'syntax', both formal and informal was steady, this progress did not lead to the anticipated improvement in the computer's ability to handle language. The trouble, of course, is that syntax is not enough. A deeper understanding of the semantics of language was required. Only in the late 1960's with the work of such people as Winograd, Woods and currently Martin, were the earlier skirmishes with the problems of syntax and semantics sharpened into serious attacks on the problems of the meaning of language. (See, for example, [10].) Thus although real problems remain to be solved, there is now justifiable optimism that a natural and direct interface between a user and a knowledge-based system can be built.

We want to underscore the importance of research on natural language to the kind of work we are currently doing, and to the proposed work of the Computer Laboratory. Of course, there is the obvious advantage of having a natural language interface with a program which contains clinical knowledge about some domain. Such an interface will permit the direct involvment of various experts (some not actively involved with the research of the Laboratory) with the program. This involvment will provide invaluable feedback with respect to the 'facts' in the program and with respect to the theories upon which the program is based.

A second benefit, perhaps, is less obvious. It has become clear that in large part the major impediment to progress in natural language research has been in semantics rather than syntax. The recent progress has built on new and better schemes for representing meanings. Further, as this research progresses, these representational schemes will be further developed and refined.

Even a cursory study of the kinds of knowledge employed by experts in solving clinical problems shows how much use is made of conceptual frameworks which at present are receiving increasing attention in language research. Such concepts as time, causality, change, etc. require deep analysis if machine representations of their meanings are to be found. The central role that such concepts play in medical knowledge means that progress by natural language researchers will almost certainly benefit our research directly. In fact much of our current thinking about representation of medical knowledge is strongly influenced by our colleagues (e.g. Martin) who are working on English.

Recognition and Analysis of Conflicting Goals

In many problem-solving applications, the recognition of conflicting goals is an important problem. Further, once these conflicts are recognized, it is important to have some means for resolving them. In earlier problem-solving programs, the recognition of goal conflict was generally difficult, because the goal structure of a program was implicit in the program itself. As we noted, the use of goal-directed programming languages lessens this problem considerably.

The analysis of conflicting goals, although still a significant problem, is also an area where improvements have been made. In the past, conflict between goals was handled by very crude strategies: either the goals were assigned simple priorities, or a trial-and-error search procedure would be tried first on one goal and then on the other in the hope that both would be achieved in some attempt.

Only recently have programs been developed which monitor their own performance sufficiently well to recognize and describe conflicts as they occur.Such monitoring is made possible in large part by the use of the goal oriented languages mentioned above to make the intention of a program more clear. (See, for example, [11]). Once in the open, problems of conflict can be faced (perhaps by special purpose programs) instead of being hidden in the rather arbitrary control structures of conventional programming systems.

Although we cannot say with any certainty exactly what processes would be needed for a computer simulation of the clinical cognitive process, it seems certain the performance monitoring and the analysis of conflicting goals would play important roles. Therefore advances from computer science research in this area are undoubtedly important for our proposed research efforts.

The Role of Computer Science Methodology

Perhaps the most important contribution which computer science research can make to the activities of the proposed laboratory is methodological in nature. The major reason that cognitive psychology has made relatively little progress with respect to understanding behaviors as complex as that involved in clinical decision-making is because there was a serious shortage of ways to describe the more procedural aspects of that behavior. As has been argued in [12]:

> *"The community of ideas in the area of computer science makes a real change in the range of available concepts. Before this, we had too feeble a family of concepts to support effective theories of intelligence, learning, and development. Neither the finite-state and stimulus-response catalogs of the Behaviorists, the hydraulic and economic analogies of the Freudians, or the holistic insights of the Gestaltists supplied enough technical ingredients to develop*

> ***such an intricate subject. It needs a substrate of debugged theories and solutions to related but simpler problems. Computer science brought with it a flood of such ideas, well defined and experimentally implemented, for thinking about thinking; only a fraction of them have distinguishable representations in traditional psychology.***

It is this rich set of ideas which we plan to exploit in the description and analysis of clinical cognition. From this effort will come a new theory of the behavior of clinical experts and new concepts for the realization of this behavior in a computer.

Research Plan

Introduction

In order to provide a context for our discussion of the research plan for the Laboratory, we want to re-iterate our goals, and to relate these goals to our perceptions of the needs of the health care system.

> ***We propose that the major activity of the Laboratory will be the use of the computer and advanced computer science methodology in the study of clinical decision-making. From the activities of the Laboratory will come two major results: 1) a deeper and better-articulated theory of expert clinical cognition, and 2) mechanisms for realizing the concepts of the new theory in computer programs for clinical decision-making.***

The reasoning underlying the organization of the Laboratory around these themes is as follows. We start from the premise that there is a need for distributable expertise concerning a number of clinical problems. Our particular interest is in the domain of <u>serious medical problems</u>, problems which are often potentially life-threatening. If we can make progress in understanding the way in which serious and complex problems should be dealt with by a clinician, and hence by a computer, we will be able to develop new technology of considerably improved flexibility and power which will be applicable across a broad range of medical decision-making applications. It can be anticipated, for example, that these advances will have an impact on the ability of the practicing physician to deal with complex or serious medical problems, placing the consultant as near as the nearest console. Such expertise should make far more effective the performance of allied health personnel, such as nurse practitioners and MEDEX personnel. In remote rural areas, for example, the availability of expert consultation should make it possible for allied professionals to deal competently with problems more serious than they otherwise could care for. In addition, the computer should be able to serve an important triage function, assisting the non-physician in his decisions concerning referral - in effect telling him when he should transfer the patient to a physician for care.

At present, however, the techniques for providing computer-based consultation are limited in application and remain generally incompatible with one another because no mechanisms for organizing and integrating them in a more general clinical context.

> ***It is this lack of integrative mechanisms which is one of the principal impediments to the realization of the full potential of the computer in health care delivery.***

Our goal is to undertake the research which will produce these integrative mechanisms, and to do this, we have turned to the study of the behavior of clinical experts, because these experts have demonstrated their abilities to combine various approaches into a coherent strategy suitable to a given situation. We should begin by understanding how they achieve their performance. Recent advances in computer science provide us with new building blocks from which we can construct a better theory of clinical cognition. This theory will be developed through extensive use of computers and computer programs as a medium for expressing the theory, and as the means by which the theory is tested.

Below we will outline a set of research projects which we believe have the proper orientation to yield major progress toward the understanding we are seeking. As our work progresses, of course, new paths will become apparent, and our ability to define problems more sharply will increase.

In what follows, we have listed the principal participants in each project. Each group of principal participants contains computer scientists and clinicians, and the activities of the groups are fully collaborative. In a real sense, everyone mentioned in any project has an active interest in all the projects, but we thought it might be of some interest to the readers of this proposal to know who currently plays a major role in each project.

Taking The Present Illness

Principals
Professor G. Anthony Gorry
Peter B. Miller
Dr. Stephen G. Pauker
Dr. William B. Schwartz

'Will, as Schwartz has suggested, the computing science "largely replace the intellectual functions of the physicians?". I think not. The subtle process of the patient-physician interaction and the input we receive from this interaction has not yet been reduced to precise mathematical terms. Attempt as we will to analyze this subtle process, it appears that (despite) our best efforts to penetrate (it), this mystery will elude us for some time.'

(Warren Glaser, Professor of Medicine, in a comment on a forthcoming article on computer-aided diagnosis)

The sentiment expressed in this quotation is shared by many physicians. Those who have thought carefully about the interaction between a patient and a physician realize the complexity of the behavior involved. When a physician is confronted with a patient with one or more presenting problems, he enters into a mode of data acquisition and problem solving known as 'taking the present illness'. This activity is one in which virtually all clinicians participate every day. When we try to understand this process in detail, however, we find that it assumes a very complex and often subtle character. In fact, virtually all the problems of clinical cognition arise in this context. The process is like a puzzle for which some of the pieces can be rather easily found and described, but for which others remain quite vague and apparently ill-formed, while some appear to be missing entirely. The question of interest here is to what extent can we identify the pieces of that puzzle and put them together to form some coherent picture.

On the other hand, if a machine is to understand the process of clinical problem-solving, it must understand the taking of the present illness, because it is this process which provides much of the underpinning of the rest of the decision-making activities. Therefore, a deep understanding of the behavior of the clinician in this setting would provide a great deal of knowledge about how to support clinical decision-making. Additionally, we chose to begin work on the present illness because it represents the initial point of contact between patient and doctor, and because of the richness it presents with respect to cognitive processes and the integrative demands it places on the clinician. Further, it has the advantage that issues of risk and benefit such as those we addressed in our work on decision analysis can be ignored. Later, as our understanding increases, we can move the boundaries our our work to include these issues as well.

<u>Preliminary Work</u>

Our activities began with the analysis of <u>protocols</u>, tape-recorded records of the verbalized problem-solving behaviors of clinicians. These protocols are augmented by in-depth questioning of the clinicians regarding their approach to specific clinical problems and by criticisms offered by these same clinicians of preliminary computer realizations of our hypotheses concerning their cognitive processes. The purpose of all these efforts is to gain a deeper understanding of the way in which clinicians <u>actually</u> deal with the complexities of the clinical environment.

We have developed our hypotheses to the point where it has been possible to implement a rudimentary computer simulation of the process of taking a present illness. Though very detailed studies of the problem solving behavior of that program, we have gained new insights into the process. The use of the computer as the medium for the expression of the theory has aided enormously the advancement of that theory. This close man-machine exploration of the behavior of the simulation of the theory will be a key aspect of our research "style". Of course, this style has the additional benefit that when a satisfactory theory has been developed, a program which takes an excellent present illness for the given problem domain will also be available.

A further aspect of our style has been our emphasis on a "complete" examination of the issues involved in taking a present illness for a single complaint (in this case, edema). By forcing ourselves to consider even "minor" differences between the behavior of the program and the behavior of the clinician as problems for investigation, we have considerably sharpened our understanding of the process the doctor uses.

Now we want to present our first, rather rudimentary understanding of the problems and processes associated with the present illness. Then we will describe our first theory and the computer realization of that theory. Finally we will discuss our research plans for this project.

<u>Observations of the Present Illness</u>

The physician, when taking the present illness, asks the age and the sex of the patient, and elicits a chief complaint. The latter is the problem which caused the patient to seek medical attention, but it will often be closely followed by mention of other problems the patient has. In fact, one interesting problem which is currently of concern to us is how a clinician links several presenting problems together. For simplicity of discussion, however, we will assume that the patient presents with a single chief complaint.

The response of the physician to the chief complaint will vary in details, but the principal thrust of it will invariably be at elucidating and refining the description of the complaint as given by the patient. For example, if the patient's chief complaint is 'swelling of the face', the physician's questions generally will explore the duration of the swelling, its specific location (e.g. around the eyes), the symmetry of the swelling (is it only on one side of the face?), etc.

The characterization of the presenting complaint is important because it is this characterization which along with the age and sex of the patient gives the clinician his initial framework or <u>context</u> within which to work.

> ***The rapid selection of a context is vital for the clinician. The clinician is about to hear a reasonably large amount of information from the patient, and if he is to be able to organize that information and to deal with it effectively, he must have a framework into which it can be fitted. Because of the breadth and diversity of medical problems and the scope of knowledge concerning these problems, a failure to focus attention and to narrow drastically the domain under consideration will prevent the clinician from <u>understanding</u> what he will be told.***

Note that this is the reason physicians require the age and sex of the patient at the outset of the history; because these facts, in conjunction with the chief complaint provide a great deal of <u>focus</u> for what follows. Consider the difference in your reaction to the chief complaint of 'severe, progressive weakness' in the case of an 80 year old man, and that of a 13 year old girl.

Therefore the initial goal of the physician in taking the present illness is to get an <u>adequate</u> description of the chief complaint of the patient. What constitutes an adequate description, however, is determined by another fundamental goal, namely that of gaining a framework within which to understand the information which will be forthcoming from the patient.

In some cases, fragments of this investigation will appear to be a rote recitation of a standard sequence of questions (e.g. in the case of abdominal pain, 'Is the pain made worse by lying down?', 'Is it made worse by eating?', 'Is it made better by eating?', etc.). Other fragments will be strongly influenced by the responses of the patient. For example, if the swelling of the face is periorbital and symmetric, the physician might want to know whether it appears in the morning and disappears during the day. If the answer is yes, then he might well transfer his attention to an investigation of possible pedal edema. On the other hand, if the swelling is in one cheek and is painful, the investigation might switch to questions of recent dental work on the patient.

Clearly, then, the path of the investigation of the chief complaint taken by the physician is in part a function of the responses given by the patient to the former's questions. This path is equally well a function of the clinical knowledge of the physician. Only a doctor who recognizes the periorbital edema described above as very likely the result of renal disease (specifically acute glomerulonephritis or less often, nephrotic syndrome) would follow the path suggested. So underlying the observable behavior of the physician is a knowledge base, the use of which is only implicit in the process of investigation.

That the investigation of the chief complaint follows a path determined by both the medical (and other) knowledge of the clinician and the responses and descriptions given by the patient is apparent to anyone who has looked at the present illness in even the most cursory manner. Thus it is non-controversial that these two factors are pieces of our puzzle. What remains unclear is how these pieces interlock in any given situation.

The exploration of the chief complaint generally results in a much sharper characterization of it than originally offered by the patient, although usually only certain additional features of the complaint have been elicited, i.e., the exploration of the complaint has been stopped short of exhausting all the properties which this problem might conceivably have. This of course raises the possibility that some aspect of the patient's problem has been overlooked, and the need for further investigation may arise in later in the session.

The characterization of the chief complaint as elaborated by this process can prompt a number of different behaviors on the part of the physician. In certain cases, the description of the complaint suggests little to him, and so he may simply encourage the patient to volunteer more information ('Have you had any other difficulties lately?') or he may begin a 'review of systems' type of investigation of the system involved in the patient's problem.

If the latter approach is used, however, it will seldom persist as the basic modus operandi, because it is too passive for use in taking the present illness, and it is used here only as a temporizing measure. As soon as it yields some additional information, the physician will assume a more aggressive stance with respect to information gathering.

The purpose of this excursion into the review of systems is the same as that underlying the original attempt to refine the characterization of the chief complaint, namely to get just enough information to glean a good suggestion of a context for further discussion of the patient's problems.

The initial context chosen will of course be further refined as the present illness is taken. It may be an organ system (in the sense that the chief complaint is strongly suggestive of a problem with that organ system); it may be much more specific in that the chief complaint might suggest a specific disease. (Of course, there may be more than one disease or organ system suggested.) In any event, the extent to which the clinician pursues the characterization of the chief complaint depends on the search for an appropriate context and the potential availability of contexts which are quite specific. For example, the facial edema described by the patient above would be pursued to establish its specific location and temporal pattern because of the specificity of the renal disease context which would result if the appropriate characterization could be achieved.

At its most macroscopic level, the taking of the present illness can be described as the clinician moving from context to context with occasional returns to previously-invoked contexts. At each context, the activities of the present illness can be thought of as being under the control of that context. By this we mean that the questioning of the patient is directed at either the confirmation of details associated with the context (such as asking about pedal edema because it is generally found when periorbital edema is present) or at the selection of a more 'specific' context (as when the clinician asks a patient with exertional dyspnea whether he has paroxysmal nocturnal dyspnea in order to choose between the contexts of lung disease or heart disease).

__Present Illness or Diagnosis?__

Before we continue our discussion, we want to comment on the role which __diagnosis__ plays in the present illness. Clearly, the present illness is 'driven' by the desire to establish an understanding of the patient's problems and their interrelations with one another; hence the clinician is seeking a diagnosis which is suitable as a basis for management decisions. There is a very real sense, however, in which the present illness is more than diagnostic process as the latter is conventionally construed.

Normally we think of __a diagnosis__ as an inference about the state of the patient which is based on his signs and symptoms, and we call the activities associated with the collection of information (identification of signs and symptoms) the __diagnostic process__. We have noted that the taking of the present illness is also an information gathering activity, but it is directed as much toward the problem of ascertaining __what the facts are__ as it is toward the problem of __what the facts mean__.

Although we admit that there is a level at which one can view the present illness as part of the diagnostic process and the process of diagnosis as an integral part of the taking of the present illness, we feel that the distinction we have made has some merit. It helps expand our view of the problems of clinical cognition.

For example, when we think only of 'the diagnostic process' we tend to think of such questions as 'What inferences can you draw concerning a 28 year old man with dyspnea and orthopnea who had an attack of acute rheumatic fever when he was 15, and.. . etc.' We tend to view the problem as understanding the meaning of a constellation of findings __as given__. We assume that the patient indeed does have dyspnea and orthopnea and that the attack of rheumatic fever actually took place. In taking the present illness, however, the clinician often is not given these facts, but must 'dig them out', and even then he may be left with significant doubts concerning the facts themselves. It is this additional aspect of establishing and characterizing the facts and assessing their reliability which we are emphasizing in our rather arbitrary distinction between the process of diagnosis and that of taking the present illness.

Now that we have made the point that the two activities of <u>establishing</u> the facts and <u>interpreting</u> the facts are central to clinical cognition, we will now explore some of the ways in which these two activities interact, and we will drop our distinction between taking the present illness and working toward a diagnosis.

<u>Prerequisites for Clinical Cognition</u>

Although many of the details of the processes employed by the clinician in taking a present illness or in proceeding to a further diagnosis are still obscure, it is possible to identify some major aspects of the general cognitive process. We can do this by analyzing the <u>task environment</u> of clinical medicine. A physician who is well adapted to that environment will <u>necessarily</u> possess cognitive processes for dealing with each of the major demands placed upon him by the environment. Although we may not be able at present to give much detail concerning these processes, we will have made a first step by recognizing the necessity of their existence. (In the following discussion, we make use of some terms borrowed from Minsky [13].)

1) <u>Expectation and Focusing</u>

The first problem that a clinician faces when he is dealing with a patient is that both the number of disease states and the number of possible findings which may have some relevance are extremely large. This means that the clinician faces a search through a potentially bewildering maze of possibilities. Because his cognitive capacities are limited (especially with respect to the number of 'simultaneous' paths he can explore), he must use the facts as presented to <u>drastically</u> reduce the number of possibilities which he will consider in any detail.

As we noted in our brief discussion of the present illness, this rapid focusing serves the principal purpose of providing the clinician with a context for his further problem solving activities. In our studies of expert clinical decision-making, we have been struck by the rapidity with which experts achieve such a framework. When they are presented with only a few (two or three) facts, experts almost always have one or two working hypotheses. It may very well be that the hypothesis first chosen will later be discarded. Our point is not that this first choice is an <u>accurate</u> or <u>optimal</u> one. It is a <u>good</u> working hypothesis, however, in that it brings important <u>structure</u> to the problem.

Because the stimuli for this focusing are the presenting signs and symptoms of the patient, it is reasonable to infer that the expert remembers patterns of findings which 'point to' good working hypotheses or contexts for those findings. Our current speculation is that these patterns contain relatively little detail, and they serve only as a first rough cut at the problem of classifying the patient. This speculation is based primarily on the experts' descriptions of the patterns they are using and on the rapidity with which this focusing takes place.

When a context has been selected, the clinician appears to match the findings of the patient against a more detailed description of the prototypical pattern of findings associated with the context. For example, 'shortness of breath in a 50 year old man' immediately suggests the contexts 'heart disease' and 'lung disease'. (Notice in fact how focused these contexts are relative to the total number of disease states which could be presented by the patient.) Most clinicians would proceed immediately to the characterization of the shortness of breath in order to focus on either heart disease or lung disease.

This attempt to match the presenting findings or the chief complaint to a more detailed pattern for a context is typical of the activities which underlie much of the present illness. For example, consider the presenting problem of periorbital edema. It immediately suggests (among a few other things) acute post-streptococcal glomerulonephritis. A renal expert would very likely move directly to a series of very detailed questions concerning the temporal pattern of the edema. The context of AGN has already been 'suggested'; the detailed examination of the characteristics of the edema will determine whether this context will govern the succeeding questions of the clinician.

2) Elaboration

Once a context has been chosen, the clinician faces the problem of confirming his choice. This confirmation requires two steps: first, he must convince himself that the rest of the signs and symptoms presented by the patient conform to his understanding of the disease state or the physiological state represented by the context, and second, he must assure himself that these findings are not better associated with one another in some other context.

One of the fundamental principles which we have observed in out studies is that experts use the principle of parsimony. The expectation that all the patient's findings are related to the same problem is strong in the clinician's mind. He yields this idea only grudgingly. In our discussion below, we will see examples of the major role this idea plays.

The process of elaboration is very complex, involving several distinct, but interacting activities.

a) <u>Filling in the Details</u>

When the clinician has chosen a working hypothesis, he is faced with the problem of confirming the details of that hypothesis. Major research question are:

* How does he select the details to explore?
* What facts should he seek from the patient?
* How should he try to establish the facts he desires?
* In what sequence should he seek these facts?
* How does he assess the validity of information?

b) <u>Assessing the 'goodness' of Fit</u>

The clinician faces another problem when a more detailed piece of information concerning the patient has been obtained, regardless of the means. He must assess how well the new information 'fits' the current context. Further this assessment must be merged with similar assessments of the 'goodness of fit' of other facts. In the face of poorly fitting facts, how far should he pursue the current context before abandoning it?

One aspect of the assessment of the goodness of fit for a finding which is particularly interesting is the process by which alternative explanations are constructed for facts which appear to be discrepant with a given hypothesis. In such cases, the poor fit of a fact to a hypothesis does not immediately cause the rejection of a hypothesis, but rather it triggers a search for a way to 'explain away' the problem. In a later section, we will discuss in more detail the problem of discrepant information.

c) <u>Rejecting Contexts</u>

Above we mentioned that under certain circumstances, a context which was chosen by the clinician may be discarded by him, because of a 'poor fit' with the facts. In this case, the clinician is giving up the working hypothesis despite his initial desire to confirm it. Here, however, the principle of parsimony may make him reluctant to give up a particular hypothesis. For example, in abandoning the current hypothesis, he may be forced to hypothesize more than one disease. Although he is often <u>forced</u> to do this, the clinician, in general, is reluctant to do so, and so he may continue with a hypothesis which fits the facts rather poorly for longer than would otherwise be expected.

In other circumstances, however, the clinician may actively want to reject contexts. The most obvious example of this occurs when the clinician has found the working hypothesis to be a good fit to the presenting facts, and he now wants to reject any other competing hypotheses.

In many cases, the clinician remembers a specific pattern of the presence or absence of various signs and symptoms which virtually precludes the presence of a particular disease. In other cases, no such specific pattern is known to the clinician, and he must use other arguments (such as the relative likelihood of two hypotheses) to exclude the hypothesis in question. Of course, in certain cases, no such exclusion can be achieved, and he must base subsequent decisions on consideration of more than one hypothesis.

It should be noted that this process of confirming one hypothesis by matching the hypothesis and then rejecting other, competing hypotheses is one which is generally interwoven throughout the process of clinical cognition. For example in the present illness, the working hypothesis might concern the 'facts' concerning some piece of the history, with competing hypotheses providing alternative interpretations of what really happened to the patient at the time in question. The same issues of confirmation, rejection, and weighing likelihoods are relevant here even though the hypotheses are not about diseases, but rather about the facts themselves.

3) <u>Alteration</u>

It was noted above that the initial context chosen by the clinician is often not supported by the information subsequently gathered. Hence the context must be replaced by a new one. If the clinician is to operate effectively and efficiently in the clinical environment, he must generally be able to shift smoothly from one hypothesis to another. The process by which this replacement occurs is an important and interesting one.

One hypothesis is that the facts are again sifted through the pattern matching processes mentioned above, and from this re-examination of the data, a new hypothesis emerges as the working context. There seems little doubt that this happens in some situations, but as a general rule, such a process seems more characteristic of a medical student or a new intern than of an experienced clinician. For the latter, a more much directed move to a new hypothesis seems appropriate. That is the expert, because of his richer and more extensive experience uses certain 'failures' in matching findings to hypotheses as direct pointers to new hypotheses. Thus, for example, the working context might be 'glomerulitis', and a questionable fit of the facts has been found; the patient has heavy proteinuria but no significant hematuria. The expert responds to this 'mismatch' by moving <u>directly</u> to the 'nephrotic syndrome' context, because he has been in this situation a sufficient number of times to have stored the 'contingency' pointer.

The importance of these direct 'pointers' arises from the amount of <u>structure</u> which they preserve. In general, a reasonable amount of cognitive effort has gone into the 'fleshing-out' of the working hypothesis, and a lot of information has been gathered. If the hypothesis is simply abandoned, and no other one is directly taken up in its place, the clinician may lose track of certain pieces of information. If the new hypothesis can be obtained directly from the old one, then this smoother transition is apt to disrupt less severely the information structure he has built.

4) Dealing with Novelty

What does the clinician do when none of his working hypotheses seems consistent with the facts at hand? Such a situation can easily occur. For example, the might be one or more facts which are in error. Alternatively, the patient might be suffering from more that one disease, and the findings cannot all be attributed to one of them.

Because such situations clearly arise in clinical practice, the good clinician will have developed strategies for dealing with them. We do not know much about these strategies at present, but we will offer a few observations. First, there is always the possibility that the clinician is facing a situation which is truly novel in certain very important regards. In this case, he will have to fall back on general intelligence and 'creativity', but we cannot offer much detail about how this is done. Undoubtedly he begins his search for an understanding of the situation by trying to understand what modifications of contexts which 'almost fit' would be required. From these necessary modifications he may be able to move to a better grasp of the situation.

In other cases, the working hypothesis seems basically sound, but certain facts cannot be fitted into the framework it provides. At face value the situation may *appear* novel, but the clinician suspects that either one or more 'facts' are in error, or there is some alternative 'explanation' of the facts which will fit into the current context. This situation is discussed in more detail in a later section which considers how clinicians deal with discrepant information.

5) Learning

The abilities described above are in some sense a minimal set for an expert to have if he is to *perform* as an expert. We know that he possesses cognitive mechanisms to realize these abilities because we can observe him successfully dealing with the problems of clinical medicine, and this task environment *requires* these skills.

Because experts are not created *de novo*, however, they must possess the skills required to *become* experts. They must possess the ability to *learn*. In terms of our above discussion, they must be able to assimilate new contexts, recognition patterns, explanations of discrepancies, and administrative strategies. This assimilation draws from a variety of sources: school, books, clinical experience, introspection, etc. Further, it is clear that simple assimilation is not sufficient for expert behavior. The knowledge that is assimilated must be *organized by the learner* so that it is *effectively* available to him in the task environment of clinical practice.

The question of whether a piece of information has been effectively assimilated into the knowledge structure possessed by the clinician can be judged only with respect to the way in which the new knowledge is used in the above processes. Hence it seems that a prerequisite for understanding learning as it relates to clinical expertise is the understanding of performance in the clinical environment.

The Initial Theory

Our theory of the cognitive behavior of clinicians is an amalgam of the ideas of a number of the workers in our group and was strongly influenced by Minsky [13]. Particularly notable contributions to the structure of this theory were made by Sussman, Pauker, and Rubin. Although our current theory is primitive and incomplete, we believe that it represents a good beginning. Here we will present it in some detail. Basically this presentation is a re-working of the above discussion in terms of the computer-based model we have implemented. The concepts used in that model are introduced at appropriate points in the discussion.

Frames

It seems that the knowledge possessed by a clinician is grouped into chunks, which, after Minsky [ref], we call frames. When he begins to entertain a certain diagnostic possibility, be it a disease, like acute post-streptococcal glomerulonephritis, a clinical state, like nephrotic syndrome, or a physiological state, like sodium retention, he bring many facts about this possibility to mind at once. It appears that physicians behave as if certain findings, which he have called triggers serve to awaken the frame into our consideration. (This is the basic mechanism for dealing with the problem of expectation discussed earlier.) At that point, any of its findings or slots can relate to presented data, but when it was in its dormant state, most of these slots could not react to presented data. For example, when told of fever, one would not immediately think of cellulitis (a kind of skin infection), but if told that there was a red, painful swelling of one cheek, the additional finding of fever fits in neatly.

Frames appear to have other types of data associated with them besides slots. There appear to be relational pointers to other frames, so that when one is considering one frame as a possibility, one is "sort of" thinking about other related frames. This relationship may be of several varieties, but a neat grouping of many of them can be made by considering the causes-of, things caused-by, complications-of, and things complicated-by the frame. For example, when one is considering acute glomerulonephritis, one "sort of" thinks about acute renal failure and acute hypertension, both of which are complications of AGN, but they are not thought of in the same detail as AGN, e.g., one usually does not consider their complications, like encephalopathy, hyperkalemia,etc, unless other data suggests them or reinforces the hypotheses of acute renal failure and acute hypertension.

Differential Pointers

In addition, there appear to be some special kinds of slots which function as lateral or differential pointers to other frames. These are meant to handle unexpected finding in a fashion that makes backing-up (a relatively costly procedure) less necessary. Rather than going back to the beginning and 'reshuffling' all the facts when a hypothesis is rejected, it appears that the physician has certain heuristics which point in specific directions when certain inconsistencies are encountered. This is a part of their response to the problem of alteration discussed earlier. For example, when presented with a patient with massive edema and heavy proteinuria, the expert can leap to a hypothesis of nephrotic syndrome. If he later discovers the patient has jugular venous distension, he can move directly to considering constrictive pericarditis, realizing that the two entities can be confused. This lateral motion is not based on reconsideration of all the data at hand, but on the differential pointer that says:

> "If you are considering nephrotic syndrome, and there is neck vein distension, then consider constrictive pericarditis."

Similarly, a young man with facial edema and hypertension can be hypothesized to have acute glomerulonephritis, since the unexpected findings of hypertensive retinopathy or ventricular hypertrophy on electrocardiogram would immediately lead to consideration of chronic glomerulonephritis.

Pruning Frames

It also seems that the physician does not maintain multiple copies of diseases having certain variations, but rather he has a general knowledge and certain rules about how to tailor-make this to the case at hand. We call this process pruning. Pruning is related to the problems of elaboration and alteration discussed earlier. Pruning may involve findings (slots), evaluations or relationships to other frames. Thus, the general picture of cirrhosis must be modified in that one cannot expect to consider gynecomastia in a women. Sodium retention may be manifested by pedal edema, facial edema, ascites and the like, but ascites is rare in renal edema and facial edema is rare in cardiac edema, even though both are part of the physician's general knowledge about sodium retention. Sodium retention may be caused by cirrhosis in the adult, but rarely in children, so when considering sodium retention one should not "sort of" consider cirrhosis, if it is a child.

Translation Frames

Another type of knowledge which physicians often bring to bear on their diagnoses relate not so much to the specific disease entities, but to a general knowledge about the world in general and medicine in particular.

Much of this knowledge can be expressed in a special kind of frame which we have called a translation frame. In some ways this can be viewed as a simple stimulus-response set:

> "If one is told the patient served in the army, it means he most likely did not have hypertension or proteinuria at that time (he passed an army physical), he probably did not have a murmur (army physicals are not known for careful observation), and probably had reasonable exercise tolerance."

> "If the patient attended summer camp, he was likely exposed to plant allergens,snake bite, other children and therefore common childhood diseases of summer (like the enteroviruses)."

Hypothesis Generation

There appears to be a hierarchy of hypothesis in so far as how actively they are being considered and in comparing them to each other. There appears to be several general classes of consideration which he have called happy, active, semi-active, and dormant.

When beginning consideration of any problem, all hypotheses are dormant; that is to say, only their trigger slots can grasp incoming data. Under specified conditions, usually finding a datum to satisfy a trigger slot, the frame moves into active state. This means that any of its slots can match findings (with the constraint that they may be pruned in fitting the frame to the case at hand). The neighbors (e.g., causes-of, complications-of, etc.) of the frame are "sort of" made active. We call their level of activity semi-active. It differs from full activity in that its "awakening" does not awaken its neighbors, thus avoiding the explosive awakening of too many frames. Finally, under certain conditions, frames become happy, that is to say, they are convinced beyond reasonable doubt that they are true and they assert that they are indeed true so that other conclusions may proceed from this assertion.

Hypothesis Testing

As findings are gathered, each frame is evaluated in several ways:

1) A check is made to see if the new datum excludes that frame. For example, the absence of proteinuria virtually denies the existence of a glomerulitis.

2) A check is made whether data is sufficient to establish the hypothesis. For example, if one finds red cell casts in the urine sediment, this virtually establishes the presence of a glomerulitis.

3) A measurement is made of how well the data fit the hypothesis and how much of the data are explained by the hypothesis. These are two complementary measures and the clinician considers some combination of

them. If the goodness of fit exceeds a certain level, he might say that the "weight of evidence" would allow the frame to become happy. On the other hand, if the fit is sufficiently poor, one might drop the hypothesis from active consideration. In doing this <u>scoring</u>, the physician allows for propagation through relations, i.e., if one is considering aortic stenosis and congestive heart failure, the finding of rales in the chest examination is very helpful to the congestive heart failure hypothesis, but by helping that hypothesis, it "sort of" lends weight to aortic stenosis also.

This then represents the substrate of the initial theory of the response of the clinician to the presentation of information about the patient. The theory has certain additional features which we can call <u>heuristic</u> rules, or what to do in certain situations. An example might be how to handle contradictory data:

> If one is told there are both red blood cell casts on urine sediment and no hematuria, then consider that there are probably no red cell casts (they are often confused with other casts) , but at some later time, see how your conclusions would be altered if red cell casts were present.

> If renal function is normal but you are told that there are no kidneys on x-ray of abdomen, consider the possibility that there are really large kidneys present, but the radiologist did not see them (as often happens with really large kidneys).

<u>Information Seeking</u>

At present, our theory of how the clinician chooses what facts to seek out is somewhat underdeveloped. We do have some understanding of this process, however, and this is a problem which is currently under study.

First it is clear that what may appear to be a "fact" to an outside observer may be less than that to a clinician. By this we mean that clinicians seem to deal in "chunks" of information which are, strictly speaking, composed of more than one fact. For example, a clinician tends to follow rather set patterns of questions until he has gotten a chunk of information about the patient. If the complaint is edema, a renal specialist will react by invoking a small "subroutine" to further characterize the edema. We call this a subroutine because clinicians themselves seem to recognize the questioning net they use as an automatic response to the stimulus "edema".

The rationale for the particular sequence of questions employed is understood by the physician, and he can readily explain it. But in practice, he does not "derive" this sequence, but rather simply remembers and invokes it.

Once a suitable chunk of information has been gained, the triggering and matching processes described above are invoked.

For the selection of which chunk of information to seek next, the clinician appears to make use of the frames themselves, trying to fill in the slots of his current hypothesis. Our understanding of the details of this process is inadequate at present, but we have been able to get some interesting results in our computer simulation by following this simple strategy.

The following few sections discuss specific projects which we have undertaken in support of the development of this theory. The first is the computer simulation of the present illness. The second project is concerned with style differences among clinicians insofar as their approach to the present illness is concerned and with measuring the effectiveness and efficiency which these differences promote. The third project is concerned with the development of orderly and concise means for identifying and codifying clinical knowledge, particularly of the kind found in medical textbooks. This work is aimed at filling some of the gaps which the present illness project must necessarily leave as it concentrates on strategy.

<u>Initial Computer Simulation of Cognitive Process</u>

In conjunction with our explorations of the knowledge and problem-solving behavior of clinicians described in the preceding sections, we have developed some preliminary computer programs to simulate aspects of the observed process of taking a present illness.

We will provide only some of the details of the operations of the computer programs involved to give the reader the flavor of our work. It should be understood, however, that these details will almost certainly be changed. In fact, much of the work discussed below in the section on supporting computer science research is aimed at refining and improving the mechanisms upon which this rudimentary simulation is built.

The basic operation of the simulation program is as follows. The age and sex of the patient is presented to the program along with the chief complaint. The program responds to this information by formulating hypotheses about the patient's condition. These hypotheses are the result of patterns of signs and symptoms which the program recognizes as suggestive of particular diseases, clinical states, or pathophysiological states. For example, the pattern "middle-aged man with pedal edema" might suggest idiopathic nephrotic syndrome, sodium retention, etc. The pattern currently known to the program were identified in our studies of experts, and the program makes the same use of them that the experts do, namely to immediately get one or more working hypotheses around which it can structure the initial phases of the present illness.

In the current simulation, the program must seek out all additional information about the patient. Therefore, once it has "digested" the age and sex and presenting complaint of the patient, it undertakes questioning of the user to learn more about the patient. Whenever a new fact is learned, the program revises its assessment of various hypotheses, and then seeks more information in accordance with its latest "opinion" of the situation. To understand the simulation, then, we need to understand two basic functions of the program:

1) how hypotheses are generated and tested
2) how questions are selected.

Here we will briefly investigate each of these questions. As noted, the emphasis will be on the concepts involved, not on the technical details of the program.

Hypothesis Generation

Stored in a data base used by the program are a great many patterns of signs and symptoms. Associated with each pattern is some action which the program is to take if the pattern is found during the present illness. Some of the actions affect hypotheses, in that they cause hypotheses to be formed, modified, or deleted. Other types of patterns and their uses will be discussed below.

The patterns of findings which cause hypotheses to be promoted to active consideration are called triggers. At the beginning of the present illness, all hypotheses are dormant in that although the program has descriptive knowledge about them (See the discussion of frames below.), it is not actively considering any of them. The triggers are used to promote some hypotheses to the active state when the chief complaint is entered. (Triggers are used at other points in the present illness also, as we shall see.) While a hypothesis is active, the program matches new facts to the description of the hypothesis (the frame) which is has been given, and it uses the frame for the hypothesis in its question selection activities. On the other hand, dormant hypotheses are ignored in both these activities.

So a trigger moves a hypothesis from the dormant state to the active state. In doing so, it may cause other hypotheses to move from the dormant state to a state which we have called semi-active. To understand the purpose of this third state, consider the above example, namely the presenting problem of massive pedal edema in a middle-aged man. There are triggers which cause the hypotheses of nephrotic syndrome, idiopathic nephrotic syndrome, and sodium retention (among other things) to become active. To reflect the fact that at this point a clinician would "sort of" be thinking of congestive heart failure (because it is a cause of sodium retention), the program moves congestive heart failure to the semi-active. The simulation program matches findings to semi-active hypotheses, but it does not use them in its question selection activities.

The specific rule which the program uses to determine which hypotheses to move into the semi-active state when a trigger is matched is as follows. The program looks at the description (frame) for the hypothesis denoted by the trigger, and finds all hypotheses related to the hypothesis in question by such relations as "causes", "complication-of", etc., and makes these hypotheses semi-active (assuming, of course, that they are not already active).

Hypotheses can move from the semi-active state to either the active state or to the dormant state as the present illness proceeds. For example, if a later finding is a trigger for a semi-active hypothesis, the latter will move to the active state. In addition, a hypothesis can move from semi-active to active if more than one other hypothesis, in becoming active, tries to move the hypothesis in question to semi-active status.

In fact, throughout the present illness, there is continual movement of hypotheses from one state to another. Active hypotheses may be "demoted" to dormant by the hypothesis testing function because it deems them to be very poor fits to the facts. The important point, however, is that hypotheses are being re-evaluated and re-ranked by the program in light of the most recent set of facts about the patient.

Consider Figure 1. Here is the trace of the simulation program as it responds to the presentation of massive pedal edema in a middle-aged man. The age and sex descriptor are translated into internal format, where each property is labeled by type. When massive pedal edema is entered, we see that this <u>triggered</u> sodium retention and nephrotic syndrome, which in turn, cause their "relatives" (for example, congestive heart failure and acute tubular necrosis are causes-of sodium retention) to go into the <u>semi-active</u> state. When <u>idiopathic</u> nephrotic syndrome became semi-active, it discovered that a prior fact (the age descriptor) fitted neatly into its description, and this second match allowed the frame (<u>idiopathic</u> nephrotic syndrome) to rise to full activity). This did not occur when the age descriptor was intially given because that finding was not a trigger for the frame. The frame had to be at least semi-active (rather than dormant) before the match could occur.

Similar interactions occur with chronic renal failure and chronic glomerulonephritis, but the reason that they come to full activity is not that they find a supporting finding, but rather that they are "sort of" thought about by more than one other frame (in this case, sodium retention and nephrotic syndrome).

In Figure 2 is a tabulation of the state of the hypotheses considered by the program. It is easy to see how this might be transformed into a "problem list" with relatively little effort. Each frame has two associated measures: its score is a normalized measure (from -1 to 1) of how well the data fits the frame, and is EXPL is the fraction of findings explained by the frame and its possible associated subframes.

[redacted]

```
=========================================================================

>  (MIDDLE-AGED MAN)

>>>>>>> (MAN (AGE MIDDLE-AGED) (TIME NOW))

>  (MASSIVE PEDAL EDEMA)

>>>>>>> (EDEMA (LOCATION PEDAL) (SEVERITY MASSIVE) (TIME NOW))

((EDEMA (LOCATION PEDAL) (SEVERITY MASSIVE) (TIME NOW))
 ==TRIGGER==>
 ((SODIUM RETENTION)<-- ACTIVE
  ==>
  ((CONGESTIVE HEART FAILURE) <-- SEMI-ACTIVE
   (CIRRHOSIS) <-- SEMI-ACTIVE
   (ACUTE TUBULAR NECROSIS) <-- SEMI-ACTIVE
   (NEPHROTIC SYNDROME) <-- SEMI-ACTIVE
   (ACUTE GLOMERULONEPHRITIS) <-- SEMI-ACTIVE
    ==>
    ((NEPHROTIC SYNDROME) <-- ACTIVE
      ==>
     ((DIABETES) <-- SEMI-ACTIVE
      (SYSTEMIC LUPUS) <-- SEMI-ACTIVE
      (IDIOPATHIC NEPHROTIC SYNDROME) <-- SEMI-ACTIVE))))))

((AGE (AGE MIDDLE-AGED) (TIME NOW))
 ==TRIGGER==>
 ((IDIOPATHIC NEPHROTIC SYNDROME) <-- ACTIVE))

((EDEMA (LOCATION PEDAL) (SEVERITY MASSIVE) (TIME NOW))
 ==>
 ((NEPHROTIC SYNDROME)
   -->
   ((INSECT BITE) <-- SEMI-ACTIVE
    (NEPHROTOXIC DRUGS) <-- SEMI-ACTIVE
    (CHRONIC GLOMERULONEPHRITIS) <-- SEMI-ACTIVE
    (GLOMERULITIS) <-- SEMI-ACTIVE
    (CELLULITIS) <-- SEMI-ACTIVE
    (HYPOVOLEMIA) <-- SEMI-ACTIVE
    (CHRONIC GLOMERULONEPHRITIS) <-- ACTIVE
     ==>
     ((CHRONIC RENAL FAILURE) <-- SEMI-ACTIVE)))

 ==>
 ((CHRONIC RENAL FAILURE) <-- ACTIVE
  ==>
  ((UREMIA) <-- SEMI-ACTIVE
   (HYPERKALEMIA) <-- SEMI-ACTIVE))
 ==>
 (CHRONIC GLOMERULONEPHRITIS) --> ((CHRONIC HYPERTENSION) <-- SEMI-ACTIVE
                                   (FOCAL GLOMERULONEPHRITIS) <-- ACTIVE))

=========================================================================
```

FIGURE 1. HYPOTHESIS GENERATION
(NOTE: User input preceded by single '>'.)

[redacted]


```
==========================================================================

 (MAN (AGE MIDDLE-AGED) (TIME NOW))
 (SEX (GENDER MALE) (TIME NOW))
 (AGE (AGE MIDDLE-AGED) (TIME NOW))
 (EDEMA (LOCATION PEDAL) (SEVERITY MASSIVE) (TIME NOW))
 (BOUND (EDEMA (LOCATION PEDAL) (SEVERITY MASSIVE) (TIME NOW))
 (SODIUM RETENTION)
 (EDEMA SODIUM RETENTION))
 ((SODIUM RETENTION) ACTIVE)
 (PRUNED-SLOTS (SODIUM RETENTION) ((DIURETIC SODIUM RETENTION)))

HAPPY-FRAMES
  NONE

ACTIVE-FRAMES

   (IDIOPATHIC NEPHROTIC SYNDROME) SCORE 0.165 EXPL 0.5 AVG 0.332
   (NEPHROTIC SYNDROME) SCORE 0.151 EXPL 0.5 AVG 0.325
   (SODIUM RETENTION) SCORE 0.102 EXPL 0.5 AVG 0.301
   (CHRONIC RENAL FAILURE) SCORE 0.071 EXPL 0.5 AVG 0.285
   (FOCAL GLOMERULONEPHRITIS)
   (CHRONIC GLOMERULONEPHRITIS)

SEMI-ACTIVE-FRAMES

   (ACUTE GLOMERULONEPHRITIS) SCORE 0.097 EXPL 0.0 AVG 0.048
   (CHRONIC HYPERTENSION)
   (HYPERKALEMIA)
   (UREMIA)
   (HYPOVOLEMIA)
   (CELLULITIS)
   (GLOMERULITIS)
   (NEPHROTOXIC DRUGS)
   (INSECT BITE)
   (SYSTEMIC LUPUS)
   (DIABETES)
   (ACUTE TUBULAR NECROSIS)
   (CIRRHOSIS)
   (CONGESTIVE HEART FAILURE)

==========================================================================
```

FIGURE 2. FACTS AND HYPOTHESES

The details of the scoring scheme are discussed below in connection with hypothesis testing.

Hypothesis Matching

In the above discussion, we ignored the representation of knowledge about diseases, clinical states, etc. used by the simulation program. We did not need this detail in our discussion of the triggering mechanism and the various states for for hypotheses.

One of the major activities of the present illness simulation program, however, is assessing how well the facts in hand at any point in time match a given hypothesis. Therefore, we need to examine the way in which descriptions of hypotheses are stored and used.

Each description is represented by a frame. A frame is an organized collection of facts about the hypothesis, what its findings are, how it is caused, what complications can arise from it, etc.

Because medical knowledge generally is organized about diseases or clinical states, and not about the implications of specific findings, this system allows for data input as its is available in standard medical texts. The necessary cross referencing for the appropriately useful associations is taken care of automatically by a frame compiler. Figure 3 is an example of a typical frame. This frame might be paraphrased as:

> Nephrotic syndrome is a clinical state characterized by hypoalbuminemia, heavy proteinuria (usually over 5 grams in a 24-hour urine), massive edema, symmetrically distributed, often involving the face, especially ther area about the eyes. There is associated elevation of serum cholesterol and urine lipids are present. It may be caused by acute or chronic glomerulonephritis, nephrotoxic drugs, some insect bites, diabetes, systemic lupus, diabetes, or may be idiopathic. It may be complicated by hypovolemia (intravascular) or infection of the massively swollen extremities. There is almost never facial edema in the absence of pedal edema, and massive edema associated with over 5 grams of protein loss daily is enough to establish the diagnosis. It may be confused with constrictive pericarditis, but in that case there is neck vein elevation. It may also be confused with cirrhosis, but in that case, ascites are usually present. If there is flank pain, one must consider renal vein thrombosis as a possible cause of the renal protein loss.

Now we can explore the scoring or hypothesis matching performed by the simulation program. Consider the scoring data shown in Figure 3, under the titles MAJOR and MINOR.

```
(DEFRAME
 $(NEPHROTIC SYNDROME)
 (TYPE CLINICAL-STATE)
 (SLOT ALB (TRIGGER) $(ALBUMIN LOW))
 (SLOT PRO NIL $(PROTEINURIA HEAVY))
 (SLOT PROQ (TRIGGER) $(PROTEINURIA >5GRAMS))
 (SLOT EDEMA (TRIGGER) $(EDEMA MASSIVE (NOT ASYMMETRICAL)))
 (SLOT FACED (TRIGGER) $(EDEMA (OR FACIAL PERI-ORBITAL) (NOT ASYMMETRICAL)))
 (SLOT CHOL NIL $(CHOLESTEROL HIGH))
 (SLOT URFAT NIL $((URINE LIPIDS) PRESENT))
 (CAUSED-BY $(ACUTE GLOMERULONEPHRITIS)
     $(CHRONIC GLOMERULONEPHRITIS)
     $(NEPHROTOXIC DRUGS)
     $(INSECT BITE)
     $(IDIOPATHIC NEPHROTIC SYNDROME)
     $(SYSTEMIC LUPUS)
     $(DIABETES))
 (COMPLICATED-BY $(HYPOVOLEMIA) $(CELLULITIS))
 (MAJOR #(($$(ALBUMIN LOW) 1.0)
   ($$(ALBUMIN HIGH) -1.0))
 #(($$(PROTEINURIA >5GRAMS) 1.0)
   ($$(PROTEINURIA HEAVY) 0.5)
   ($$(PROTEINURIA (OR ABSENT LIGHT)) -1.0))
 #(($$(EDEMA MASSIVE (NOT ASYMMETRICAL)) 1.0)
  ($$(EDEMA (NOT ABSENT) (NOT ASYMMETRICAL) (NOT ASYMMETRICAL)) 0.3)
   ($$(EDEMA ERYTHEMATOUS (NOT ABSENT)) -0.2)
   ($$(EDEMA ABSENT) -1.0)))
 (MINOR #(($$(CHOLESTEROL HIGH) 1.0)
   ($$(CHOLESTEROL (NOT HIGH)) -1.0))
 #(($$((URINE LIPIDS) PRESENT) 1.0)
   ($$((URINE LIPIDS) ABSENT) -0.5)))
 (MUST-NOT-HAVE $*(AND (EDEMA FACIAL (NOT ABSENT)) (EDEMA PEDAL ABSENT)))
 (IS-SUFFICIENT $* (AND (EDEMA MASSIVE) (PROTEINURIA >5GRAMS)))
 (DIFFERENTIAL-DIAGNOSIS
  ($((NECK VEINS) ELEVATED)
   (SEMI-ACTIVATE '$(CONSTRICTIVE PERICARDITIS)))
  ($(ASCITES PRESENT) (SEMI-ACTIVATE '$(CIRRHOSIS)))
  ($(FLANK-PAIN)
   (SEMI-ACTIVATE '$(RENAL VEIN THROMBOSIS)))))
```

FIGURE 3. NEPHROTIC SYNDROME FRAME

The score information given in each frame consists of a list of various tests, associated with a number between -1 and 1. If the test is true, that number is added to an accumulating sum. The maximum sum is the total number of such items, so a normalized score is the actual sum divided by the maximum. If no data is known about the fact sought, zero is added to the actual sum, so this weighs somewhat against the score, but less so as more data is known since the sum is divided by a larger normalizing factor. Major and Minor scores just specify factors by which there respective sums are multiplied, so the major factors count more. Score propagation is accomplished by passing the score of the related frame (not its sum), which is therefore normalized already, as an additional test. Frames may move from one state to another (e.g., from active to semi-active) when certain logical criteria are met. (A positive throat culture is sufficient to establish a streptococcal infection), but we also allow changes based on weight of evidence. For example, is the score of any active frame exceeds a pre-established threshold, then it becomes happy, whereas if it falls below a different pre-established threshold, it may lapse into the semi-active state.

At this point we might digress to mention score-propagation . It is clear that when a frame gains evidence in its behalf, its relatives must also become more convinced of their truth also. For example, acute glomeronephritis is related to (by "complicated-by") acute hypertension. If we learn, that there is hypertension in the absence of hypertrophy on the electrocardiogram, this must add weight to acute glomerulonephritis. If we then learn that there is no chronic hypertensive retinopathy, acute hypertension gains more credence, and this gain must be propagated up to acute glomerulonephritis.

The inverse effect is equally true, i.e., since a low urine sodium is explained by sodium retention, and since sodium retention can be caused-by acute glomerulonephritis, then acute glomerulonephritis can explain the abnormal finding of low urine sodium if we can invoke sodium retention. In this program, both scores and "explanations" of findings can be propagated through frames which are either happy or active.

Question Selection in the Present Illness

Now we can turn our attention to the way in which the program seeks additional information during the present illness. Here we have implemented procedures which are first approximations to those the program will need if it is to behave in the style of a physician in so far as its choice of and ordering of questions is concerned

From our detailed study of the way in which a particular expert took a present illness, we concluded that he used two distinct modes of questioning. At times, he invoked a rather rigid, "compiled", sequence of questions, particularly to sharpen the characterization of a particular finding. This sequence seemed aimed at quickly, but narrowly, focusing the problem solving. Such questions can be thought of as filling a pattern which if matched will trigger a very specific hypothesis. An example of such a sequence is shown in the first part of Figure 4.

The program is first told that the patient is a young boy with facial edema (at this point, it might be well to say that the patient who is being questioned in this example actually has acute glomerulonephritis). The program attempts to further characterize the facial edema, asking about duration, recurrence, temporal pattern, etc. The edema fits so well into the typical pattern of renal edema, that the program does not pursue details such as pain and erythema. At this point, the chances that this is anything other than renal edema are so remote that the program is willing to pay a "reprocessing penalty" if it is wrong.

Next the program asks about associated pedal edema. This occurs because of a simple heuristic rule which states: "if you are told of facial edema, see if there is associated pedal edema ." (This fits with the MUST-NOT-HAVE rule in the Nephrotic Syndrome frame <Figure 3>). Pedal edema is likewise explored in depth, but note the additional questions about severity, pain and erythema which are relevant for this kind of edema. It should be noted that the determination of what is relevant here is the behavior of the expert himself. He asks different questions about pedal edema than about facial edema, and in order for the program to appear to behave "naturally", it must do the same.

This additional question can also be thought of as being more or less "automatic" in that the common conjunction of pedal and facial edema makes the investigation of the former a natural next step for the clinician, and hence for the program.

It is at this point that the second form of questioning makes its appearance. Here the clinician must weigh various hypotheses and possible questions and seek to plot a good strategy.

We have suh()essed the detailed output of the program as it does the hypothesis generation and matching described above. Suffice it to say that at this point, the program examines the states and scores of the various frames in deciding on its next question. It will pursue the active frame (and its related subframes) which has the highest score. It will not pursue happy frames, since it is already reasonably convinced of their applicability.

The score used in selecting the frame to explore is not just the "goodness of fit" score mentioned above. Since the choice is guiding further questioning, we are also concerned that our hypothesis (and its subframes) explain as much as possible of the data. Hence the program averages the SCORE and EXPL values, and pursues the hypothesis with the highest average. In this case, that is acute glomerulonephritis, and the program turns to one of the latter's subframes , streptococcal infection. When it finds that there is indeed a pharyngitis, it explores that in depth as to character and severity.

```
==============[redacted]=====================================
>> (YOUNG BOY)
>> (FACIAL EDEMA)
ABOUT (NOW FACIAL EDEMA),
CONCERNING DURATION,IS IT:
1.    DAYS
2.    WEEKS
3.    MONTHS
4.    YEARS
>> 1
CONCERNING RECURRENCE,IS IT:
1.    FIRST-TIME
2.    INFREQUENT
3.    OCCASSIONAL
4.    FREQUENT
>>
CONCERNING PERI-ORBITAL?,IS IT:
1.    PERI-ORBITAL
2.    NOT-PERI-ORBITAL
>> 1
CONCERNING SYMMETRY,IS IT:
1.    SYMMETRICAL
2.    ASYMMETRICAL
>> 1
CONCERNING TEMPORAL-PATTERN,IS IT:
1.    NO-TEMPORAL-PATTERN
2.    WORSE-IN-MORNING
3.    WORSE-IN-EVENING
>> 2

IN THINKING
ABOUT (PEDAL EDEMA),
CONCERNING TIME,IS IT:
1.    NOW
2.    RECENT-PAST
3.    PAST
>> 1
CONCERNING STATUS,IS IT:
1.    ABSENT
2.    PRESENT
>> 2
CONCERNING DURATION,IS IT:
1.    DAYS
2.    WEEKS
3.    MONTHS
4.    YEARS
>> 1
CONCERNING RECURRENCE,IS IT:
1.    FIRST-TIME
2.    INFREQUENT
3.    OCCASSIONAL
4.    FREQUENT
>> 1
CONCERNING TEMPORAL-PATTERN,IS IT:
1.    NO-TEMPORAL-PATTERN
2.    WORSE-IN-MORNING
3.    WORSE-IN-EVENING
>> 3
=============================================================
```

FIGURE 4. PRESENT ILLNESS SIMULATION

(Note: user inputs preceded by '>>')

```
************************************************************************
CONCERNING SEVERITY,IS IT:
1.     1+
2.     2+
3.     3+
4.     4+
5.     MASSIVE
>> 3
CONCERNING SYMMETRY,IS IT:
1.     SYMMETRICAL
2.     ASYMMETRICAL
>> 1
CONCERNING TYPE,IS IT:
1.     PITTING
2.     NON-PITTING
>> 1
CONCERNING ERYTHEMA,IS IT:
1.     ERYTHEMATOUS
2.     NOT-ERYTHEMATOUS
>> 2
CONCERNING PAIN,IS IT:
1.     PAINFUL
2.     NOT-PAINFUL
>> 2

PLEASE TELL ME ABOUT
 PHARYNGITIS,I.E.,
 IS THERE (NOW (NOT ABSENT) PHARYNGITIS) ?
>> YES
ABOUT (NOW (NOT ABSENT) PHARYNGITIS),
CONCERNING APPEARANCE,IS IT:
1.     EXUDATIVE
2.     NON-EXUDATIVE
>> 2
CONCERNING SEVERITY,IS IT:
1.     MILD
2.     SEVERE
>> 2

HAPPY-FRAMES
    NONE

ACTIVE-FRAMES
   (ACUTE GLOMERULONEPHRITIS) SCORE 0.298 EXPL 0.75 AVG 0.524
   (SODIUM RETENTION) SCORE 0.195 EXPL 0.75 AVG 0.472
   (STREPTOCOCCAL INFECTION) SCORE 0.181 EXPL 0.75 AVG 0.465
   (IDEOPATHIC NEPHROTIC SYNDROME) SCORE 0.161 EXPL 0.75 AVG 0.455
   (NEPHROTIC SYNDROME) SCORE 0.068 EXPL 0.75 AVG 0.409
   (ACUTE RENAL FAILURE) SCORE 0.066 EXPL 0.75 AVG 0.408

PLEASE TELL ME ABOUT
 STREPTOCOCCI,I.E.,
 IS THERE (NOW EXPOSURE STREPTOCOCCI) ?
>> ?

************************************************************************
```

FIGURE 4. Continued

```
===========================================================================

PLEASE TELL ME ABOUT
 SCHOOL,I.E.,
 IS THERE (NOW ATTENDED SCHOOL) ?
>> YES

PLEASE TELL ME ABOUT
 PENICILLIN,I.E.,
 IS THERE (NOW GIVEN PENICILLIN) ?
>> ?

PLEASE TELL ME ABOUT
 FEVER,I.E.,
 IS THERE (NOW MILD FEVER) ?
>> NO

PLEASE TELL ME ABOUT
 (THROAT CULTURE),I.E.,
 IS THERE (NOW BETA (THROAT CULTURE)) ?
>> YES

PLEASE TELL ME ABOUT
 HEMATURIA,I.E.,
 IS THERE (NOW (NOT ABSENT) HEMATURIA) ?
>> YES
ABOUT (NOW (NOT ABSENT) HEMATURIA),
CONCERNING AMOUNT,IS IT:
1.     MICROSCOPIC
2.     GROSS
>> 1

PLEASE TELL ME ABOUT
 PROTEINURIA,I.E.,
 IS THERE (NOW (NOT ABSENT) PROTEINURIA) ?
>> YES
ABOUT (NOW (NOT ABSENT) PROTEINURIA),
CONCERNING AMOUNT,IS IT:
1.     LIGHT
2.     HEAVY
>> 1
CONCERNING QUAN-AMOUNT,IS IT:
1.     <100MGRAMS
2.     100MGRAMS-5GRAMS
3.     >5GRAMS
>> ?

PLEASE TELL ME ABOUT
 WEIGHT,I.E.,
 IS THERE (NOW (OR HIGH RISING) WEIGHT) ?
>> NO

PLEASE TELL ME ABOUT
 RALES,I.E.,
 IS THERE (NOW PRESENT RALES) ?
>> YES

===========================================================================
```

FIGURE 4. Continued

[redacted]

```
========================================================================

HAPPY-FRAMES
  (STREPTOCOCCAL INFECTION) SCORE 0.348 EXPL 0.538 AVG 0.443
ACTIVE-FRAMES
  (ACUTE GLOMERULONEPHRITIS) SCORE 0.477 EXPL 0.538 AVG 0.508
  (GLOMERULITIS) SCORE 0.287 EXPL 0.538 AVG 0.413
  (SODIUM RETENTION) SCORE 0.208 EXPL 0.538 AVG 0.373
  (IDEOPATHIC NEPHROTIC SYNDROME) SCORE 0.177 EXPL 0.538 AVG 0.358
  (CONGESTIVE HEART FAILURE) SCORE 0.110 EXPL 0.538 AVG 0.324
  (ACUTE RENAL FAILURE) SCORE 0.075 EXPL 0.538 AVG 0.307
  (ATHEROMATOUS EMBOLI) SCORE 0.005 EXPL 0.538 AVG 0.271
  (NEPHROTIC SYNDROME) SCORE -0.043 EXPL 0.538 AVG 0.247
  (STONE) SCORE 0.25 EXPL 0.076 AVG 0.163

 (NOW YOUNG BOY)


 (NOW FACIAL DAYS FIRST-TIME PERI-ORBITAL SYMMETRICAL WORSE-IN-MORNING EDEMA)
 (PEDAL NOW  PRESENT DAYS FIRST-TIME WORSE-IN-EVENING 3+ SYMMETRICAL PITTING
        NOT-ERYTHEMATOUS NOT-PAINFUL EDEMA)
 (NOW (NOT ABSENT) EXUDATIVE SEVERE PHARYNGITIS)
 ((STREPTOCOCCI (EXPOSURE EXPOSURE) (TIME NOW)) UNKNOWN)
 (NOW ATTENDED SCHOOL)
 ((PENICILLIN (GIVEN? GIVEN) (TIME NOW)) UNKNOWN)
 ((NOT MILD NOW) FEVER)
 (NOW BETA (THROAT CULTURE))
 (NOW (NOT ABSENT) MICROSCOPIC HEMATURIA)
 (NOW (NOT ABSENT) LIGHT PROTEINURIA)
 (NOW (NOT (OR HIGH RISING)) WEIGHT)
 (NOW PRESENT RALES)

========================================================================
```

FIGURE 4. Continued

Next in Figure 4, we see the state of the various hypotheses that the program is considering.

Continuing its pursuit of streptococcal infection, the program looks for possible exposure. When it is told that no information about this is available, the program seeks indirect confirmation of the presumed exposure. he program finds that school attendance can result in streptococcal exposure. It makes this connection by tracking through a series of <u>translation</u> frames (streptococcal exposure goes along with childhood disease exposure, and the latter might occur in school or summer camp).

Finally the program is told that the throat culture was positive. With this fact, streptococcal infection becomes "happy", e.g. the program is , convinced that streptococcal infection is present, and asserts it as a finding. The program then turns to the next subframe of acute glomerulonephritis, since more data about streptococcal infection would at this point be moot. This next subframe is glomerulitis, and the issues of hematuria and proteinuria are pursued. At this point, sodium retention is explored.

In the bottom lines of Figure 4, we see a summary of the data in order of acquisition. To paraphrase:

> This is a young boy, who presents for the first time with symmetrical, peri-orbital edema, worse in the morning, for the past few days. It is associated, over the same time period, with 3+ symmetrical, pitting, pedal edema, which is worse in the evening. The patient has a severe, exudative pharyngitis, has been attending school, has not received penicillin and has not been febrile. Throat culture was positive for beta-hemolytic strep. There is also microscopic hematuria and light proteinuria. His weight is normal and has not been increasing, but rales are present on chest examination.

This is not an unreasonable first-pass present illness for a patient with acute glomerulonephritis.

A contrasting strategy is shown in Figure 5 where a different case is considered. After exploring the finding of pedal edema in a middle-aged man, a specific avenue of exploration is not clearly evident. Rather that beginning with in depth exploration of a particular frame, the program asks, in the manner of a clinician, about dyspnea, alcohol consumption, past proteinuria, varicose veins and jaundice, attempting to pin down an area for profitable exploration. Once the existence of past proteinuria has brought chronic glomerulonephritis, nephrotic syndrome and chronic renal failure to mind, these are then explored.

```
========================================================================

>> (MIDDLE-AGED MAN)
>> (SYMMETRICAL PEDAL EDEMA)
ABOUT (NOW SYMMETRICAL PEDAL EDEMA) ,
CONCERNING DURATION, IS IT:
1.     DAYS
2.     WEEKS
3.     MONTHS
4.     YEARS
>> 2
CONCERNING RECURRENCE, IS IT:
1.     FIRST-TIME
2.     INFREQUENT
3.     OCCASSIONAL
4.     FREQUENT
>> 3
CONCERNING TEMPORAL-PATTERN, IS IT:
1.     NO-TEMPORAL-PATTERN
2.     WORSE-IN-MORNING
3.     WORSE-IN-EVENING
>> 3
CONCERNING SEVERITY, IS IT:
1.     1+
2.     2+
3.     3+
4.     4+
5.     MASSIVE
>> 5
CONCERNING TYPE, IS IT:
1.     PITTING
2.     NON-PITTING
>> 1
CONCERNING ERYTHEMA, IS IT:
1.     ERYTHEMATOUS
2.     NOT-ERYTHEMATOUS
>> 2
CONCERNING PAIN, IS IT:
1.     PAINFUL
2.     NOT-PAINFUL
>> 2

IN THINKING
ABOUT (NOW DYSPNEA) ,
CONCERNING STATUS, IS IT:
1.     ABSENT
2.     PRESENT
>> 1

IN THINKING
ABOUT (NOW (ALCOHOL CONSUMPTION)) ,
CONCERNING AMOUNT, IS IT:
1.     NONE
2.     RARE
3.     SOCIAL
4.     HEAVY
5.     ALCOHOLIC
>> 3

========================================================================
```

FIGURE 5. ANOTHER PRESENT ILLNESS SIMULATION

(Note: user inputs preceedded by '>>')

```
========================================================================

IN THINKING
ABOUT (PAST PROTEINURIA) ,
CONCERNING STATUS, IS IT:
1.    ABSENT
2.    PRESENT
>> 2
CONCERNING AMOUNT, IS IT:
1.    LIGHT
2.    HEAVY
>> 1
CONCERNING QUAN-AMOUNT, IS IT:
1.    <100MGRAMS
2.    100MGRAMS-5GRAMS
3.    >5GRAMS
>> ?

IN THINKING
ABOUT (NOW (VARICOSE VEINS)) ,
CONCERNING STATUS, IS IT:
1.    ABSENT
2.    PRESENT
>> 1

IN THINKING
ABOUT (NOW JAUNDICE) ,
CONCERNING STATUS, IS IT:
1.    ABSENT
2.    PRESENT
>> 1

HAPPY-FRAMES
     NONE
ACTIVE-FRAMES
  (CHRONIC GLOMERULONEPHRITIS) SCORE 0.213 EXPL 0.285 AVG 0.249
  (SODIUM RETENTION) SCORE 0.204 EXPL 0.285 AVG 0.245
  (NEPHROTIC SYNDROME) SCORE 0.166 EXPL 0.285 AVG 0.226
  (IDIOPATHIC NEPHROTIC SYNDROME) SCORE 0.166 EXPL 0.285 AVG 0.225
  (CHRONIC RENAL FAILURE) SCORE 0.086 EXPL 0.285 AVG 0.186
  (FOCAL GLOMERULONEPHRITIS)

 PLEASE TELL ME ABOUT
 KUB ,I.E.,
 IS THERE (NOW KIDNEYS-BOTH-SMALL KUB) ?
>> NO

PLEASE TELL ME ABOUT
 HEMATURIA ,I.E.,
 IS THERE ((NOT ABSENT) PAST HEMATURIA) ?
>> NO

PLEASE TELL ME ABOUT
 HYPERTENSION ,I.E.,
 IS THERE (NOW (NOT ABSENT) HYPERTENSION) ?
>> NO

========================================================================
```

FIGURE 5. Continued

Protocol Collection and Analysis

Principals
Professor G. Anthony Gorry
Dr. Jerome P. Kassirer
Peter B. Miller

In conjunction with our studies of the clinical decision making process, we have undertaken the collection and analysis of tape-recorded protocols of a number of clinicians taking present illnesses. We really have two purposes in mind with respect to this study.

In the present illness project discussed above, we have relied on the observation of and introspection by a single clinical expert for the most part. Although this has proved very productive, we want to know if major variations in "style" exist, and whether some styles are more efficient and/or effective than others. Therefore, we need to broaden the base of the observed problem solving behavior upon which we are constructing our cognitive theory.

The second purpose of this study is to collect protocols which can be used in testing the computer simulations we are employing. With detailed protocols in hand, we can compare the behavior of programs with that of clinicians on a "step by step" basis. Such comparisons will undoubtedly suggest refinements and improvements in the theories, and this form of testing will be a central methodological tool of the Laboratory.

We have already initiated this collection and analysis of protocols. Our current study involves the presentation of a case to renal experts. The clinician is asked to take a present illness from the patient. (The part of the patient is played by another physician.) The basic procedure of the experiment is as follows:

1) The renal expert is first told the age, sex, and chief complaint of the patient.

2) The renal expert then can ask questions concerning the patient, one at a time.

3) For each question, he must say <u>why</u> he is asking the question.

4) After receiving the answer to a question, the expert must say what the answer "means" to him insofar as his current view of the case is concerned.

In the current study, we are presenting the same case to five renal experts on the staff of the New England Medical Center Hospital. This group was chosen for several reasons: 1) they are indeed experts, and we are interested in expert behavior; 2) they are kidney specialists, and their protocols on a kidney problem can be used in testing the simulation programs we are developing; and 3) because they are all in the same specialty and in the same hospital, they are apt to show some common behavior, and this will make our first pass at modeling their behavior somewhat easier.

As we become more experienced in the collection and analysis of protocols, and as our understanding of the clinical process becomes more highly developed, we will expand our efforts to include clinicians from other specialties.

As an example of a problem in which "style" differences might play an important role, consider the following.

Because the physician is often interested in historical information about the patient, he must often rely on the patient himself for this information. In many cases, the patient cannot (sometimes will not) remember the exact circumstances in question, or the recollections of the patient are suspect. In such a situation, the clinician may search for witnesses to the patient's past condition. Consider, for example this brief excerpt from a protocol in which the patient is a young boy with symptoms of heart disease and a possible episode of acute rheumatic fever some five years ago.

Pat. "Well, 4 or 5 years ago, I was out of school for 3 or 4 months. I had pain in my joints...."
Doc. "Tell me a little more about this episode. Were you hospitalized?"
Pat. "No. The doctor took care of me at home."
Doc. "What did he say was wrong with you?"
Pat. "St. Vitus dance."
Doc. "Did he treat you with anything?"
Pat. "He just gave me aspirin."
Doc. "He gave you aspirin? Did you take it frequently?"
Pat. "He said...you know... I don't even remember."
Doc. "Did you have a sore throat that started the whole thing off? Did anyone ever mention it to you? Did the doctor ask you whether you had a sore throat?"
Pat. "I don't know doc. I get a lot of sore throats."
Doc. "Did the doctor inject you with penicillin back in that time? Do you remember?"
Pat. "No he didn't inject me."
Doc. "You don't remember if you took any penicillin by mouth?"
Pat. "Oh, maybe he gave me some pills."
Doc. "Where's your mother?"

Now in this brief excerpt, we see the clinician trying to establish whether the patient in fact had an attack of acute rheumatic fever four or five years ago. The patient gives evidence which is not conclusive on the matter. The clinician turn his attention on the quest for witnesses, people or their actions, which would confirm the acute rheumatic fever.

Did the attending physician tell the patient he had acute rheumatic fever? Did he treat the patient as though he had ARF? Did he ask the patient the questions on would expect a doctor to ask if that doctor thought the patient had ARF? Can the mother be found, and will her recollections of the time in question prove more decisive?

A central question to ask is whether this behavior is typical of experts in similar situations. Perhaps this kidney expert reverts to this behavior because the problem of acute rheumatic fever is out of his domain of expertise. Will he use the same approach to a problem of acute glomerulonephritis that occurred five years ago?

A cardiologist with whom we discussed this specific protocol, said that he did not believe that he would have followed this line of investigation. He felt he would have questioned the patient more carefully about his remembrance of the symptoms. The cardiologist conjectured that he would pursue this line because he was very familiar with the symptoms of acute rheumatic fever.

If this were the case, then the difference in style would really reflect a difference in knowledge. In other cases we have studied, however, real style differences seem to arise. Some clinicians work backward in time in that they move in a rather strict line from a problem to its antecedents. Others seem to move across all the problems which occured at a particular time before moving back in time with any one of them. Still other Clinicians seem to "jump around" quite a bit.

This study will proceed with these experiments, attempting to identify differences in style, and to devise measures of the efficiency and effectiveness of these style variations. We do not feel that important new cognitive processes will be uncovered here that have been overlooked in the present illness project (although certain aspect of the process may receive attention sooner). What will be different here will be the characterization of the various ways in which different clinicians assemble and apply the building blocks of the present illness.

To bolster our ability to maximize what we learn from this study, we are planning to include a cognitive psychologist in our group for consultation on issues of cognitive style.

The Formalization of Clinical Knowledge

Principals
Dr. Jerome P. Kassirer
Ann D. Rubin
Professor Gerald J. Sussman

Introduction

One of the obvious problems facing researchers in computer-aided clinical decision-making is how to identify and codify the knowledge which is relevant to a given clinical area. In the present illness project, we face this problem, but we have chosen to skirt some of the major (and difficult) problems of codification and representation in order to rapidly push forward into the process of the present illness. In this project, we are taking a more careful look at the problem of identifying and coding expert knowledge in an orderly way. This problem is difficult for several reasons:

1) It is often unclear, even to the expert, exactly what knowledge he uses in a given situation.

2) For many clinical problems, there seems to be a very large amount of knowledge which is relevant (at last potentially) .

3) Much of the knowledge seems to be very diverse, consisting of pieces of knowledge which are quite diverse in form.

These problems make the development of a concise, orderly way for representing clinical knowledge very important.

Above we commented on the limitations of previous formalisms for representing clinical knowledge. Basically, each has its virtues, and each can be fruitfully applied in certain circumstances; but none is sufficiently flexible and powerful to cope with the diversity and complexity of clinical knowledge.

The most obvious example of an attempt to deal with this problem of organization and presentation is a book abut a particular clinical problem. Although the book serves certain purposes well, it is inadequate in many respects. First, a book is an intrinsically <u>linear</u> form. That is, the author must choose a central theme around which his facts or opinions must be organized. Consider the following passage from a chapter about acute glomerulonephritis. (13)

> "Typically the illness with pharyngitis or tonsillitis accompanied by fever and malaise. Whether or not specific antibiotic therapy is given, respiratory symptoms and fever disappear after a few days, and the patient feels entirely well. One or two weeks after the onset of the illness, weakness and anorexia return, and the patient notices that his urine is scanty in amount and smoky in appearance. Upon

awakening the next morning, he notes swelling around the eyes and complains of shortness of breath and headache."

The text continues in this vein with a discussion of the remainder of the scenario for the "classic" patient with acute post-streptococcal glomerulonephritis. Later in the chapter, in a discussion of clinical features of the disease, we find:

> "Gross hematuria, one of the most common initial symptoms, occurs in more than one-third of the patients. The urine is often described as reddish-brown, smoky, rusty, tea-colored, or cloudy. In most cases, gross hematuria disappears after a few days, but it may continue for one or two weeks. Microscopic hematuria can, of course, be found for a much longer period, and often persists even after significant proteinuria is no longer present."

In the first quotation, it is clear that the authors have chosen to organize the information they are presenting around the time course of the evolution of the disease in the "classic" patient. The discussion mentions a number of signs and symptoms, but only in passing. The objective is to provide a coherent picture of the course of the disease, and too much attention to details will obscure that picture. There can be only one major line to the discussion at one time.

In the second quotation, the focus of attention has been shifted to hematuria, one of the 'details' of the earlier discussion. Now much about hematuria that was passed over in the first discussion is presented. In this discussion, proteinuria is treated as a detail, but later in the chapter, it, too, becomes a main theme around which other facts are organized. In fact, in that discussion, hematuria is treated as a detail.

The point is a rather obvious one, but it is very important. The conventional presentation of information in a book places a real cognitive burden on the reader. The reader must organize the information in his memory, and he must create the associative links implicit in the text. For example, he should associate the 'smoky urine' of the first discussion with the 'smoky urine' in the hematuria discussion. Links must be formed from the details of the first discussion to more extensive knowledge structures about these details.

For knowledge such as this to be <u>clinically</u> useful, it must be digested by the clinician. The demands of the clinical environment are such that the linear organization (as in the book) is inadequate. At a minimum, the clinician must be able to access this knowledge from the 'entry point' of the patient's presenting problems (e.g. smoky urine) and from the entry point of particular disease hypotheses (e.g. Does the patient match the picture of AGN?).

A second cognitive demand which information presentation such as this places on a reader is the need for re-coding. Clearly the clinician does not remember such text verbatim. His memory of it is coded in terms of a (perhaps very large) number of symbolic structures. Part of this re-coding probably is essential if he is to remember the material; another part probably is idiosyncratic and helpful in efficiently retrieving the facts contained in the material.

Although our knowledge of these matters, particularly with respect to details of the mechanisms involved, is limited, our interest in gaining an understanding of these questions is very great. Few would argue against our contention that knowledge such as that presented in the quotes from the chapter is an essential ingredient of clinical expertise. It is also certain, that such knowledge is not organized in the expert's memory the way it is organized in a book.

We have undertaken a research project aimed at the identification of the knowledge structure of an expert in a particular area of clinical medicine, the differential diagnosis of hematuria. The advantage of working with an expert is that he has already digested material such as that cited above and he has organized it in a way which is clinically useful (at least to him). By working primarily with him, and supplementing this work with studies of books and papers such as the one mentioned, we can proceed most efficiently and effectively. Our goals are several:

1) First, we want to catalog what the specific knowledge is.

2) Second, we want to understand how much knowledge is required for expert performance in this problem.

3) Third, we want to develop a formalism for representing this knowledge including the appropriate associations.

4) Fourth, we want to understand how this knowledge is employed by the expert to solve clinical problems.

This project is closely related to the present illness project discussed above, and it is also closely tied to the efforts to develop good computer representations of medical knowledge which we will discuss below. Further, we expect these projects to move in close concert in the future, with a major activity of the Laboratory centering on the merging of fruits of these efforts.

For the near future, however, we feel that by maintaining different emphasis in these projects, we can best bring the research issues into focus. Continuity and cooperation among the projects will be maintained by the participation of key researchers in more than one project each.

Preliminary Work

To gain a better understanding of the knowledge possessed by an expert about the problem of hematuria, we undertook a series of experiments in what we called "CPC mode". Each experiment consisted of presenting a case from a Clinical Pathology Conference to a clinician.

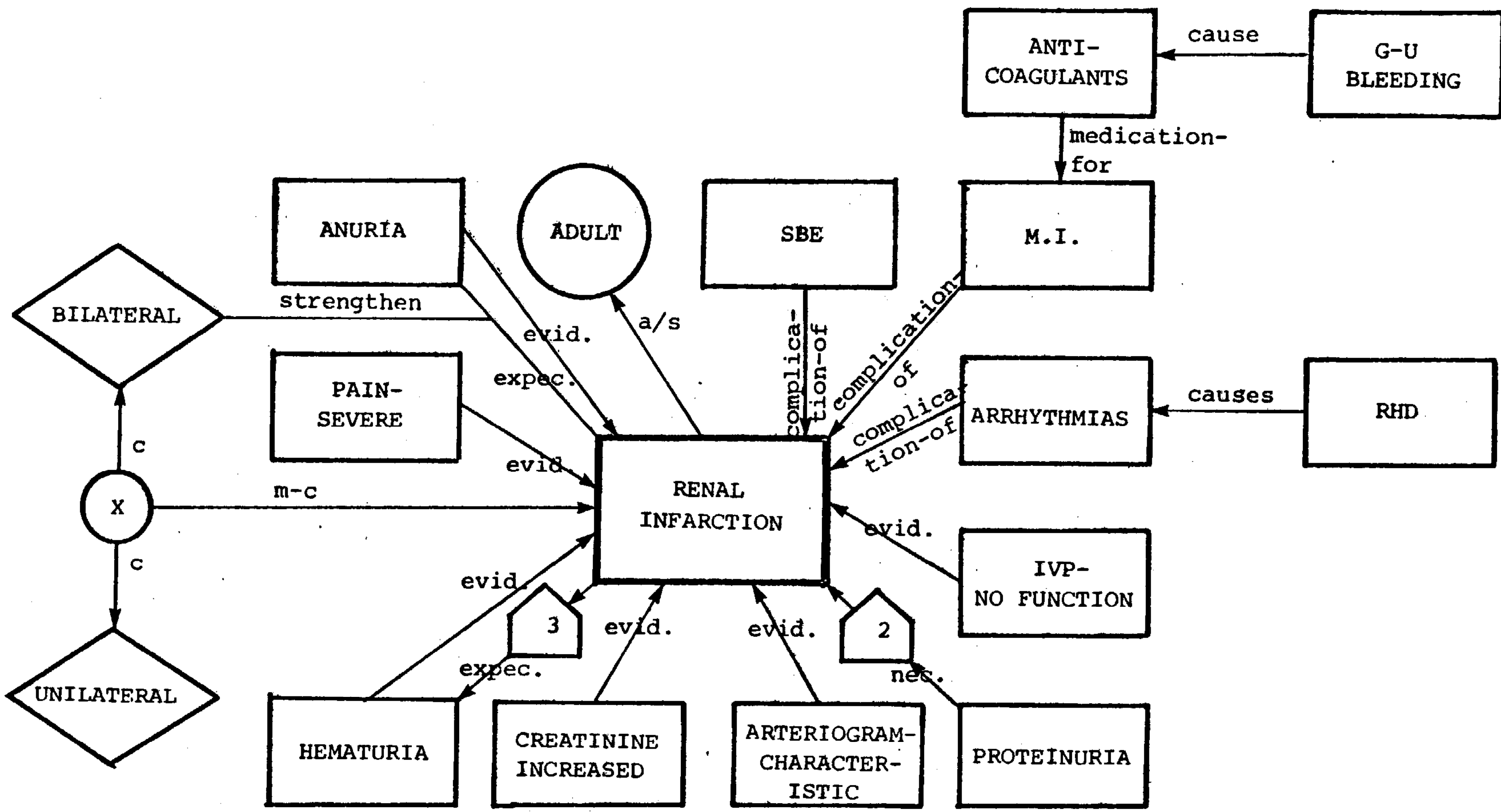
The CPC was presented to him one fact at a time. After each fact was given to him, he was asked to discuss the "meaning" of the fact. The meaning of the fact to him included the immediate conclusions which he could draw from it, its effects on hypotheses currently being considered, its suggestions of new hypotheses, etc. He was questioned in detail to make certain that the observers understood the reasons for his interpretation of the fact. When a satisfactory understanding of his reaction to the fact had been obtained, another fact was given to him, and the process was repeated.

From the observations of several such sessions, a first representation of the inferred knowledge base was constructed. This was discussed in detail with the clinician, and he was able to make many alterations and suggestions for additions. The knowledge structures discussed below result from many iterations of this process.

There are certain problems which arise during this kind of observation of behavior. Most are minor. One problems is that the clinician generally finds this mode of information acquisition somewhat uncomfortable and unnatural. Another problem is that it is sometimes necessary to ask him questions to clarify the details of his response. This raises the possibility that the clincian may alter his behavior in response to the additional questioning.
In addition, there is a question as to the validity and completeness of introspective statements concerning the knowledge employed. Even if we acknowledge all these problems, however, we still can report that these experiments were very successful. From them we gained new insights into the structure of clinical knowledge, and we gained some new ideas about how to represent this knowledge and its structure.

Consider the diagrams in Figure 6 and Figure 7. These are slices of clinical knowledge, the first organized about the central concept of renal infarction; and the second, about pyelonephritis. These slices are typical of the large number of such diagrams which have been constructed during the course of this project. The purpose is to identify and structure a sufficient amount of knowledge about a given problem (here, hematuria) to form the basis for a program to do differential diagnosis.

As is apparent from these sample diagrams, the same problems of organization of information remain. The construction of such slices requires the selection of a central theme.



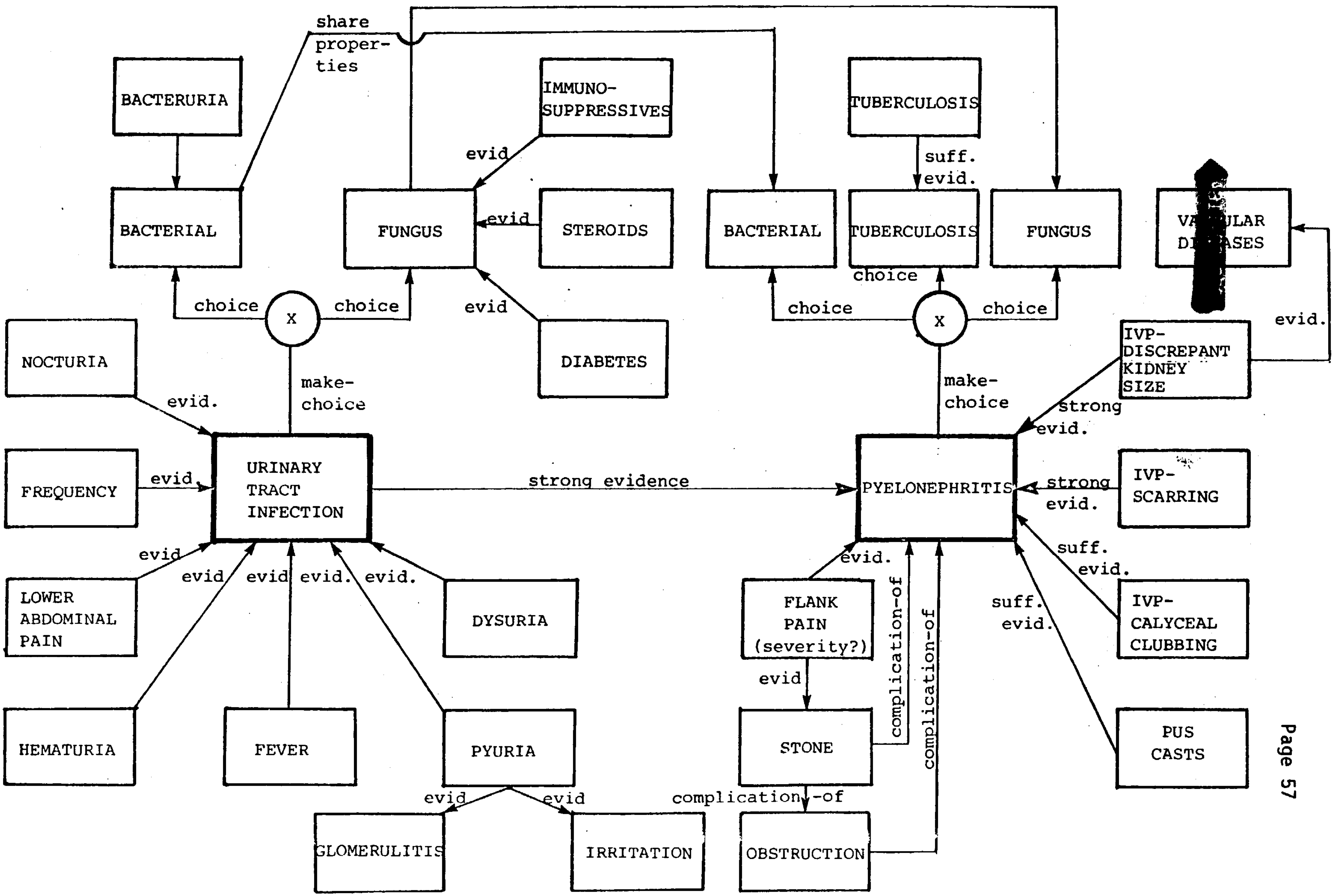
RENAL INFARCTION

ABRUPT ONSET OF ALL SYMPTOMS

FIGURE 6

FIGURE 7

PYELONEPHRITIS



As in the textbook examples above, there are many ways to "slice" the knowledge which is relevant to the problem of hematuria. We have allowed the clinician to make these slices in what ever way seems most natural to him. Our emphasis has been on encoding each slice in an orderly and clear way. This is the reason for the graphical form we have chosen - clinicians seem to be able to work with this form comfortably.

We still face the problem of relating all these slices to one another. We plan to do that in the computer. A program for accepting these slices (in some form) and making all the proper associations to link the slices together will be produced. This program will be based on the GOBBLE system we have developed and which is discussed in a later section. The network of concepts which results from the assimilation of these slices by this program will serve as the knowledge base upon which programs for differential diagnosis can be constructed.

We should note here that the construction of even rudimentary programs for diagnosis is an important step in obtaining the clinical knowledge in question. We have found, however, that only part of the knowledge possessed by an expert can be elicited from him in a direct manner. An additional component of this knowledge can be identified only through interaction with a computer program which makes decisions based on the knowledge which he has already cataloged. We found this to be true in our work on decision analysis, and we are finding it true here. After a certain point, the clinician must see someone (in this case a program) do something with the knowledge in order to see whether it is complete, has been understood, etc.

Because of this, we have started to build an interface through which clinicians can interact with a knowledge base of these slices and some rudimentary diagnostic programs. The purpose is to identify places where there are gaps or errors in these slices, and in the process, to learn something about diagnostic process. The interface will permit the clinician to use a subset of English (see the discussion of this in the section on computer science research) to ask questions and to get simple explanations of knowledge in the slices. He will also get explanations of the way in which the diagnostic programs used this knowledge in making decisions. Further, the clinician will be provided with facilities for recording complaints, suggestions, etc.

By making this interface simple and direct, we hope that we can get clinicians other than those working in the project to help us build this knowledge base. Further, such an interaction may encourage some of these clinicians to become more actively involved in the efforts of the Laboratory.

In addition to this work, we are currently analyzing protocols of differential diagnoses of hematuria to see if the slices we have identified are adequate representations of the knowledge employed by the clinicians. This activity is useful, because we can "hand simulate" a diagnostic program which uses the slices, and thereby learn quickly whether our basic concepts are sound. More detailed studies, using computer programs, will, be required in the long run, but these experiments should prove very valuable in the short run.

Model-Based Decision Making Project

Principals
Professor G. Anthony Gorry
Dr. Stephen G. Pauker
Howard Silverman

Introduction

For a number of problems of clinical medicine, there exist formal models upon which decisions can be based. In these cases, it is sometimes true that the best decisions are made through a dependence on the model. The reasons for the superiority of the model-based decision may be several.

First, the relevant physiology or pathophysiology underlying the problem may have been modeled with precision surpassing that which the clinician can maintain in his own, less formal model. In some cases, the clinician's model is inferior because it fails to account for certain details of a process. In other cases, the clinician cannot (or will not) do the computations required to achieve the accuracy of the formal model. In still other cases, the clinician does not know the parameters of the system with sufficient precision to make predictions of system behavior which are as good as those of the formal model.

In any event, there are situations in which models (perhaps coupled with automated decision making procedures) can outperform the average physician, and in certain cases do better than even the best physician in solving particular problems. Examples which come to mind are acid-base chemistry and the administration of antibiotics.

In general, the problem domains in which models such as these have been successful share an important characteristic. This is that the clinical problem can be dealt with in isolation from the most of the other problems which the patient might have. This does not mean that the model (or computer program based on the model) does not consider aspects of the patient's condition other than the particular problem in question, but rather that the number of such considerations is small, and *in toto* these problems can be rather neatly circumscribed. Of course, it is rather obvious that this property greatly increases the likelihood that such a model can be developed.

There are other clinical areas where models exist, but a variety of factors which are not (or perhaps cannot be) incorporated in the model are relevant to the decisions required in the clinical area in question.

Here the clinician wishing to use a program based on the model encounters some difficulty. First he may know certain facts about the current clinical situation which he would like to combine with the program's results. The program cannot accomodate this additional information. This is to be expected; not all models can incorporate *all* potentially relevant factors. The problem is, however, that the physician is not sure *how* to combine his judgments with the results of the program. For example, exactly how did the program arrive at its conclusion? What assumptions was it making? Did it already include consideration of some of the information he is considering?

In some circumstances, the program could produce *packaged* responses to standard questions which would satisfy the clinician. If they do not, then it is not clear what he should do.

Of course, an ideal solution from the clinician's point of view is for him to have access to a consultant who understands the program and the model on which it is based. Then when questions arise, or when the clinician simply wants to learn some more about the model, he can go to the consultant. The consultant will understand the language and the background of the clinician, and he will know how to make his explanations understandable.

Now the reader may easily guess that we would propose that the *program become the consultant*. The program should know much more than how to *compute* the model. It should know what the model is, how it was developed, and what relation it has to the problems facing the users (clinicians). Such a program, of course would have to possess a great deal of knowledge. It would need the knowledge of the consultant described above. Before we discuss this possibility and the research problems involved further, let us offer another argument for trying to build programs which are "knowledgable" about models.

We noted above that various models have been developed which now serve as the basis of decision-making programs. In several instances, these programs are real clinical successes. If we look to the future, we can see the need to bring a (potentially large) number of such programs together in a common system. Such a system will need a great amount of administrative knowledge as we discussed above. One aspect of that knowledge will need to be knowledge about these model-based programs. In general, the administrator of the system will need answers to all the questions posed by the clinician above. (What assumptions have been made in this program? Are its assumptions compatible with the clinical situation? With the assumptions of a second program which will be used?, etc.) If programs such as these are to be marshalled together in some clinical situation, questions such as these become paramount. The major research problem is how to insure that some supervisory system can get answers to these questions when it needs them.

For these reasons, we have undertaken the study of model-based decision making. Specifically we are studying situations in which a model is relevant, even central, but not all-inclusive. In these situations, the best decisions are made by clinicians who are experts in the area and well acquainted with the model in question. We want to build a program which is really an expert in the domain in question (and generally this domain is very limited). With the model as a core, the program would possess a knowledge base which encompassed all the facts and procedures use by the expert in his work with the model.

In addition, the representation of this knowledge would be such as to support an inquiry and explanation facility which was natural and direct for a clinician, and this representation would also facilitate the supervision of the model by some higher level program monitoring the overall clinical strategy. Finally, this representation scheme would be suitable for a variety of different models.

These efforts directed at developing the technology for such programs and models will be discussed below in our section on representation research.

The specific problem we have chosen for our initial project in this area is the administration of digitalis-digoxin. We now turn to a discussion of this problem.

<u>The Digitalis/Digoxin Therapy Advisor</u>

The clinical use of digitalis preparations has been one of the classical skills of the experienced clinician. Although this drug is often life-saving, its proper administration is difficult and requires careful clinical judgment. Digitalis possesses a rather low toxic-therapeutic ratio, and signs of under-digitalization are often very similar to signs of toxicity.

There have been several recent advances in clinical biochemistry and pharmokinetics which have significantly altered the use of this drug, and much of this new technology and knowledge is now available to clinicians throughout the country. However, administration of this class of drugs still remains a significant clinical problem, and we feel that the availability of a knowledge-based system concerning the cardiac glycosides may be of additional <u>clinical</u> use.

<u>Background</u>

Use of the foxglove began several hundred years ago, but until recently techniques of administration have changed very little. Withering's original advice was to administer the drug until signs of improvement or signs of toxicity occurred, and that remains the cornerstone of digitalis therapy today. Problems arise, however, because the signs of toxicity can often be confused with signs of insufficient drug dosage, and mistakes can be costly since the first sign of excess drug administration can be sudden death. The <u>clinical</u> signs of digitalis excess are <u>cardiac</u> (disturbances of cardiac rhythm) and <u>extra-cardiac</u> (nausea, vomiting, anorexia, visual changes), but the dangers of excess drug are by and large cardiac. The extra-cardiac signs are helpful if they occur before the dangerous cardiac manifestations of toxicity and if they are predictive of those more serious toxic problems.

Quite often, however, the first hint of excess drug dosage is a potentially serious disturbance or cardiac rhythm. The interpretation of these arrhytmias is often less than straightforward. The same arrhythmia can often be a sign of either under- or over-digitalization. For example, ventricular premature beats may be caused by digitalis toxicity or by congestive heart failure (by enlarging the heart and stretching its conduction system). In the case of under-digitalization, administration of more drug might suppress these extra beats by decreasing heart size. However, if the ventricular premature beats were indicators of early excess digitalis effect, then the slight increase in drug dosage could easily lead to a fatal arrhythmia.

In addition to this complex problem of recognizing toxicity, there are other <u>complicating factors</u> in using digitalis. A variety of myocardial processes (varying from myocardopathy to acute myocardial infarction) make the heart more sensitive to cardiac glycosides and thus make toxicity more likely to develop. In addition, there are non-cardiac problems which alter sensitivity, including thyroid dysfunction, electrolyte imbalance, hypoxemia, acidosis and the like. The astute clinician is continually aware of these factors and tries to adjust his dosage to what he judges the patients clinical state to be.

<u>Recent Advances</u>

Jelliffe (14) and Doherty (15) have demonstrated a variety of kinetic factors influencing the amount of active glycoside available to the myocardium after a given dose. These factors include variation in absorption, distribution and excretion of the drug. Because the drug is usually given over a relatively short dosage cycle (once or twice daily down to every other day or so) compared to its <u>in vivo</u> half life (for digoxin 1.6 days and up; for digitoxin and digitalis leaf 6.0 days and up), there is an exponential accumulation of body stores. Therefore changes in excretion and absorption can have a marked influence on body stores. For example, administration of digoxin to a man with normal renal function in a dose of 0.25mg daily would give body stores of roughly 0.625 mg at equilibrium, whereas if the patient had moderate renal functional impairment (a stable creatinine of 2.5mg%) his body stores would be approximately 1.25mg. With a drug of such a low toxic therapeutic ratio, variations of this magnitude are potentially dangerous.

Other studies(16)) have shown variation in the bio-availability of the drug from patient to patient and from <u>brand to brand</u>. This naturally limits the usefulness of a model which only deals with distribution and excretion.

Direct measurement of serum drug levels have recently become fairly common. The assumption that these serum levels bear a reasonable correlation to myocardial levels seems to have been borne out clinically, in that these serum measurements can, on the average, predict the occurrence of drug toxicity. However, we have already mentioned that sensitivity and toxic threshold varies from patient to patient in different clinical settings, so serum levels can only serve as a rough guide.

The State of the Art

What, then, is the behavior of the cardiologist today with respect to the administration of digitalis? He first tries to establish that the drug is indicated, and depending on the indications, decides on how rapidly the patient must be digitalized (loaded with the drug to reach equilibrium levels). He then selects a preparation whose kinetics fit these objectives. Most cardiologists next decide on what maintenance dose they would tend to use in this setting (based on those factors which influence sensitivity to the drug), although they might equivalently select a serum or body store level to fit the situation. The loading and maintenance schedules are then determined based on the patient's renal function and fat-free body mass.

This program is then begun, with careful, frequent examination of the patient for signs of beneficial effect and toxicity. Depending on patient response to his initial program, the cardiologist modifies his plan. If the patient demonstrates either early, unexpected signs of toxicity, or fails to demonstrate clinical response at reasonable doses, the physician may then obtain serum drug levels to clarify the situation. For the vast majority of patients on digitalis preparations, serum levels are used either as a guide in confusing situations or as a source of comfort to the physician. It is still ultimately the patient's clinical response to the drug that dictates changes in therapy.

When faced with a patient who requires therapy with digoxin and who is undergoing changes in renal function, the physician uses both the parmacokinetic models and serum drug level measurements. The model is used to prospectively adjust dosage to reasonable ranges, and then this is "fine-tuned" retrospectively by clinical observation and drug level determinations. In this situation, the pharmocokinetic model assumes a central importance. One might imagine the physician selecting arbitrary dosage plans and tuning them by clinical response and serum drug levels. Although this technique might arrive at the same end-point, it would make it more likely that the patient would be exposed to toxic levels for some brief period. Since toxicity can be fatal, a predictive approach, using the model, is preferable.

Current Computer Approaches

Jelliffe and others have developed computer implementations of various kinetic algorithms which modify suggested administration schedules for renal function (stable or changing), body size and route of administration. These programs also allow for the smooth transition from one preparation to another with differing pharmocokinetics. Studies have shown (15) that availability of these programs can make a significant difference in the incidence of digitalis intoxication. Sheiner has added the feature of feedback data based on measured serum level to further adjust dosage for the individual patient. However, a recent study by Peck (17) failed to demonstrate a significant difference in the performance of expert physicians given access to computer-predicted schedules with serum level feedback, when compared to similar physicians not having access to the program. This suggests that the _expert_ physician already uses the gross prediction algorithm, and that a significant part of his "expert" behavior centers about the tuning of his predictions based on _clinical observation of patient response_.

Our Approach

We propose to implement a knowledge-based digoxin dose advisor, which uses the generally available pharmacokinetic models for its initial prediction phase, but which also has the ability to guide the _non-expert_ physician through the feedback loop of adjusting drug dosage based on clinical response. We would hope that this program might better allow the non-expert to model his behavior after that of the cardiologist, and that interaction with such a program would both improve his treatment for the individual patient and teach him the principles of sophisticated drug use. We feel that this goal can be accomplished because the use of this drug constrains us to a fairly circumscribed, well-defined group of clinical settings.

The development of a program to predict dosage based on age, body size and renal function has already been accomplished in many centers, and we have such an implementation currently available. This system will first determine why the drug is being given (arrhythmia, congestive heart failure, prophylactically) and also look for any factors that might predict increased patient sensitivity. Based on these determinations, it will establish a desired speed of approach to equilibrium. With this factor and knowledge of patient size, age, sex and renal function (as estimated by whatever parameters are then available), it will suggest an initial loading and maintenance schedule.

The physician will then be encouraged to interact with the program, prior to administration of each dose at first, and later, at intervals throughout the equilibrium phase. The program will guide his search for cardiac and extra-cardiac signs of toxicity and will collect data about clinical effect. We do not propose that the program will directly interact with the patient's electrocardiogram in search for manifestations of effect or toxicity, but rather will ask the physician about specific features of the EKG. For the marginally experienced physician a set of labeled examples will be provided. Based on this information concerning patient response, the program will suggest modifications of drug schedule.

If the situation becomes confusing or if unexpected effects are observed, the program will have the ability to ask for and use data about serum drug levels. We would also envision this program to be useful in dealing with a patient already receiving digoxin or digitoxin, but whose response is either troublesome or requires confirmation.

Dealing with Discrepant Information

Principals
Prof. G. Anthony Gorry
Dr. Jerome P. Kassirer
Dr. Stephen G. Pauker
Dr. William B. Schwartz

Introduction

In the above discussion, we have emphasized the rapidity of the focusing which clinicians do during their interactions with patients. Our observation of clinicians at work has caused us to view them as rather aggressive with respect to hypothesis construction and testing. Because they assume this aggressive posture in their problem solving activities, they frequently confront situations in which new facts are in conflict with their working hypotheses. An important aspect of expert performance is the facility with which the expert can respond to these instance of discrepant information.

In some cases, the problem is readily apparent: two pieces of information are clearly contradictory. For example, he may be told that the patient has no hematuria but he does have red blood cell casts. Except in the rarest of circumstances, these two statements are contradictory because hematuria is a prerequisite for the formation of red blood cell casts. So the clinician has the obvious choice of assuming that there really are red blood cell casts and the hematuria was overlooked, or there in fact is no hematuria and the red blood cell casts are illusory. In accepting either alternative, he must account for the implied error.

In other, more complex situations, a fact may not directly contradict other facts, but the acceptance of the new fact by the clinician may cast serious doubt on one or more hypotheses he is maintaining. For example, suppose that the findings support the hypothesis that the patient has idiopathic nephrotic syndrome. Assume that the records from the hospital to which the patient was admitted before being transferred to this hospital show that his serum creatinine was 1.0 mg. per. cent. two weeks ago. The same test run today in this hospital yields a value of 7.6 mg. per. cent. Clearly the acceptance of these two values as accurate measures of the patient's renal function requires the conclusion that the patient is suffering rapidly progressing renal failure. On the other hand, patients with idiopathic nephrotic syndrome almost never suffer rapidly progressing renal failure, and so there is a significant discrepancy between these values taken together and the hypothesis concerning the underlying disease. Of course the hypothesis of idiopathic nephrotic syndrome can be rejected, or one or both of the serum creatinine values can be dismissed, but either course will require new hypotheses to be generated and melded into the overall picture the clinician has of the patient.

The problem of dealing with discrepant information is a common and important one for clinicians. The strategies which experts use to solve these problems are not well understood at present. Nonetheless, a number of observations can be made which can serve as a basis for further research and discussion. The importance of this investigation should be underscored, because without the capability to deal with discrepant information, a computer program cannot succeed in the face of the complexities of real clinical situations.

<u>Recognizing Discrepancies</u>

The recognition of a contradiction always is conditioned on some assumed state of knowledge about the world. For example, the fact that the hematuria-red blood cell casts situation mentioned above constitutes a contradiction is based on physiological knowledge about the formation of these casts. In other cases, a contradiction is recognized as such only on the assumption of a hypothesis about the disease state of the patient. The only difference in these two situations is the degree of certainty the clinician possesses about the state of the world. In the first case, he is so certain of the physiological mechanisms involved that he only considers the possibilities that the hematuria has been missed or the red cell casts are spurious. In the second case, he might also consider the possibility that his hypothesis about the underlying disease state is in error.

For convenience, we recognize three types of assumed states of knowledge about the world:

1) physiologic knowledge,
2) hypotheses about the disease state of the patient, and
3) common sense knowledge.

These categories of assumed knowledge are not precisely defined, nor are they exclusive, but they do provide a rough cut at the bases on which contradictions are recognized.

For any of these states of knowledge, different situations can produce contradictions. We have identified a number of these situations. For example, these five situations can occur conditioned on the acceptance of knowledge of one of the three kinds suggested above.

1) More than one of a set of mutually exclusive alternatives are asserted to be true. (for example, a patient is said to have normal renal function, but the radiologist reports that KUB studies

show no kidneys.)

2) A state of the world is asserted, but one or more prerequisites for that state are denied.
(The hematuria-red blood cell cast example above)

3) A "cause" is asserted, but one or more of its certain "effects" are denied. (For example, it is believed that decreased renal function is the cause of observed hyperkalemia, but the patient's serum creatinine is normal.)

4) A measurement exceeds absolute or experiential limits.

5) The rate of change of a physical state exceeds absolute or experiential limits (For example, a patient claims to have gained 40 pounds in one day).

Contradictions are most easily recognized when they violate principles or facts which are known to be always true. When the known principles or facts are conditioned on the acceptance of a hypothesis, the contradiction can be asserted only on the assumption of the underlying hypothesis. For example, in the example of the patient with apparent rapidly progressing renal failure, the discrepancy is not absolute; there are many examples of situations in which such acute renal failure can occur. It is the acceptance of the hypothesis of idiopathic nephrotic syndrome which produces the conditional discrepancy.

A complicating factor in the identification of discrepancies is that they need not be direct. Inferences drawn from one fact may contradict those drawn from another. Here it is required that the contradiction itself be recognized, but in addition the original facts which triggered the contradictory deductions must be identified as discrepant. Further, such <u>indirect</u> discrepancies may arise through chains of deductions conditioned on various hypotheses.

As a small example of this kind of problem, consider a patient whose presenting signs and symptoms suggest a cardiac problem. Further suppose that the patient tells the doctor that when he was a young boy he was treated for a "heart murmur" by his family physician. This latter fact strengthens the physician's belief that the patient's problems are the result of heart disease, in particular heart disease of long duration. Then in passing, the patient mentions that he served in the army during the Korean war. This fact is discrepant with the hypothesis that the patient's current heart disease is a progression of his childhood problem. If he served in the army, then he passed an army physical exam. Such an exam probably would have revealed his heart murmur (especially if it was loud), and he would not have been accepted. Further, it can be presumed that he had a reasonable exercise tolerance, and this too argues against the assumption of long-standing heart disease.

<u>How Experts Deal with Discrepancies</u>

As might be expected, experts use a number of approaches in their attempts to resolve discrepancies during the diagnostic process. Basically these approaches can be divided into three categories: 1) doubting or dismissing one or more of the stated facts; 2) constructing alternative relationships or connections among the discrepant facts which make the discrepancy only apparent, not real; and 3) revising or dismissing an underlying hypothesis about the disease state of the patient. The choice of a method for dealing with discrepancies in many cases is dictated by specific real world knowledge. In other cases, although there is a certain amount of specific knowledge concerning the situation in question, the clinician must fall back on more general problem solving strategies.

One point is worth noting here, because it seems to be characteristic of the approach used by experts. When confronted by a situation in which several facts appear to be discrepant, the expert makes a specific choice of explanations which resolve the discrepancy. If later facts cause him to discard this explanation, he will return to select another explanation if possible. Further, if his explanation appears to be confirmed, he will make at least a cursory check of the alternative explanations to make certain he is correct. He does not, however, attempt to process alternative world views (one in which one fact is assumed to be in error, another in which a second fact is assumed to be incorrect, etc.) in parallel. When discrepancies arise, they are almost always dealt with directly, and a specific explanation is constructed.

In order to indicate some of the richness of the information used to resolve discrepancies, we offer two real medical problems, and we will identify the knowledge used by the clinician to construct an explanation of the way in which the problem arose. The first is relatively easily resolved; the last is considerably more complex.

In many instances, a problem arises because of a simple factual error. An example of such a problem is given above in which it is asserted that there are red blood cell casts but no hematuria. Here, because of the physician's firm belief in his understanding of the pathophysiological mechanisms involved, he must reject one of these facts. The physician clearly would like to have the urine studies repeated in order to resolve the problem; but in certain cases, the facts are historical, and no further information can be gathered. In this case, the clinical's knowledge of the relative likelihoods of error will determine his choice of explanation. Many more mistakes are made in the detection of red blood cell casts than in the detection of hematuria, and so he would proceed on the assumption that the patient had neither hematuria nor red blood cell casts.

The more complex situation is the case of the patient cited above who was thought to have idiopathic nephrotic syndrome. Recall that a problem arose because two measurements of serum creatinine taken two weeks apart indicated rapidly progressing renal failure. Here we have a conditional contradiction, in that the development of renal failure in patients with idiopathic nephrotic syndrome is insidious. Hence, the clinician must resolve the situation, perhaps at the expense of the hypothesis of idiopathic nephrotic syndrome.

If the other evidence favoring the hypothesis of idiopathic nephrotic syndrome is quite strong, then the natural inclination of the clinician will be to doubt the evidence for rapidly progressing renal failure. The simplest way to do this is to attribute the problem to a simple factual error. Either the serum creatinine done at the other hospital or the one done here is in error.

Of course, it is a simple matter to repeat the test in this hospital, and to make the situation interesting, let us assume that repeating the test yields the same result. So the clinician now knows that the patient is in renal failure. The question of the rapidity of its onset remains, however, and the lab test result from the other hospital becomes suspect.

Now in trying to ascertain the validity of a test result from the past, the clinician faces a different problem. Obviously, the test cannot be repeated; the only avenue open to him is to gather other facts about the patient, and to consider whether they are consistent with the result in question. For example, if an x-ray of the kidneys was taken at the first hospital and the physician has access to it, it may cast some light on the problem.

If the x-ray shows that the kidneys are small, then it is reasonable to assume that the serum creatinine measurement from the first hospital was in error, because kidneys of reduced size indicate a renal problem of relatively long duration and severity and atrophy of the kidneys takes a year or more with chronic renal failure (except with renal infarction). This in turn is inconsistent with normal renal function (as indicated by the lab test).

If the x-ray shows normal-sized kidneys, then the validity of the lab test cannot be determined in this way, because although people with kidneys of normal size usually have normal renal function, when disease is present, impaired renal function will precede atrophy of the kidneys. Therefore, the patient could have been in renal failure during his stay in the first hospital (the lab test is in error) and the x-ray of the kidneys would show normal size.

For the purposes of our example, let us assume that attempts such as this to ascertain the validity of the first serum creatinine all fail, and the clinician is left with the two values which are inconsistent with his diagnosis of idiopathic nephrotic syndrome. There is another way he can try to resolve the conflict, namely by retaining the diagnosis, and trying to show that the presence of renal failure is not a direct consequence of severe damage to the kidneys. This requires some rather specialized, expert knowledge on his part.

If the patient is losing enough protein in his urine, he can become hypovolemic. The mechanism for this involves a severe reduction in his serum albumin with an accompanying reduction in blood volume. This reduced blood volume in turn can cause a reduction in the glomerular filtration rate which is sufficient to produce a markedly elevated serum creatinine concentration. Experience indicates that only under special circumstances can this occur, but when it does, it produces elevations of the serum creatinine which can be mistakenly interpreted as the result of severe structural renal damage.

The expert knows the limits of proteinuria, hypoalbuminemia, and serum creatinine which are consistent with this mechanism. He can match the patient's findings to these limits in order to test this hypothesis. Further, he knows that if this mechanism is operative, the patient should manifest low blood pressure (at least posturally), and so he would use blood pressure as evidence for or against this hypothesis.

Of course, the third possibility which the clinician should consider is that his original hypothesis of idiopathic nephrotic syndrome is incorrect. To follow this route, however, probably will require a major reorganization of the facts in his mind in order to fit them into another framework. Whether he is willing to make this reorganization will depend on the success of the approaches described above, and the strength of his belief in his diagnosis based on the totality of the facts in hand.

Reasoning of this complexity is often required in difficult clinical situations. We plan to undertake some studies of the way in which clinicians deal with such complexity. At present, we see aspects of the problem of discrepant information throughout all our work with clinicians, but our work has not produced a single, coherent project. We have raised the problem of discrepant information here however, despite our rather vague plans for dealing with it, because we realize its importance, and we plan to initiate an effort focused on it as soon as possible.

<u>Research on Dealing with Discrepancy</u>

In the absence of a specific research plan, we will suggest a number of goals we hope to achieve with the work we will initiate in this area.

1) <u>How Are Discrepancies Recognized</u>?

A problem which we will face immediately is that of finding a good characterization of discrepancies. What exactly constitutes a problem of this type? How does a clinician recognize such a problem?

This problem is more difficult than it appears at first glance. Consider, for example, the addition of a SINGLE fact to a knowledge base. How should this fact be "tested" to see if it contradicts one or more facts already accepted. Does a clinician test the incoming fact with every fact he knows? With every fact he knows about the patient? If he uses only some of the facts he knows, how are this subset selected?

The "obvious" answer to this last question is that he tests the new knowledge only against existing knowledge which "relates" to it. But of course, this simply avoids the issue; how do we measure "relatedness" in a meaningful way?

This problem of recognizing discrepant information is really a difficult one. A great deal of effort will be required to solve it. Our immediate goal is to first develop a theory of how potential conflicts among facts and hypotheses are recognized. This work will involve not only introspection and protocol analysis, but also it will require some innovations with respect to the ways we have for representing knowledge in a computer. Thus this work will interact with the work on GOBBLE discussed below.

Although we do not know now how this effort will develop, we think that it most likely will involve the detailed study of a number of clinical examples. These studies may be augmented by studies of the way people recognize discrepancies in situations other than clinical ones.

2) How Are Discrepancies Dealt With?

Once a discrepancy has been recognized (at least tentatively), the clinician must deal with it (if only by ignoring it). We will study the way in which clinicians deal with discrepancies using our basic approach of protocol analysis and interview. The result of this effort will be the description of a number of the strategies they use, and the characteristics of the situations in which these strategies are employed.

These strategies will be tested by simulation, and their efficacy will be considered in various clinical situations. As soon as possible, we will begin to integrate the work on conflict identification with this work. It should be noted, however, that both these efforts can proceed in parallel at the outset.

Supporting Computer Science Research

Principals
Professor G. Anthony Gorry
Kenneth Kahn
Peter Miller
Dr. Stephen G. Pauker
Charles Safran
Howard Silverman

Introduction

In the projects discussed above, the present illness project, the formalization of medical knowledge project, and the model based decision-making project, a number of computer science issues were raised (at least implicitly). In some cases, a need for improved technology is more or less clear; further we see ways to produce the required improvements. In other cases, we will need to do more fundamental research to achieve the facilities required by the medical projects.

In this section, we will discuss some computer science problems which arise in the context of the medical projects, and will review our current work on these problems and our plans for the future. Much of this work is in preliminary stages, and so the examples we give show our first prototypical programs. Undoubtedly much will change as we proceed, and so we offer these examples only as that, not for their technical details.

We also want to emphasize the advantage which our close association with the computer science community at M.I.T. offers us with respect to these problems. A considerable amount of research is being pursued by members of that community which is either directly in line with or supportive of our efforts. We plan to draw heavily on the expertise of these workers, and whenever possible, we will incorporate their ideas into our work. On the other hand, we believe that our research will produce ideas and technology which they will find equally interesting and useful. In all, we are anticipating a close and fruitful collaboration.

Computer Representation of Clinical Knowledge

One of the needs of each of the above projects is a means for representing knowledge in the computer. This representational scheme must be capable of accomodating diverse forms of knowledge, and at the same time, it must allow flexible retrieval of knowledge. We have undertaken the development of a program, called GOBBLE (written in LISP), for managing a data base of knowledge. It is our intention that GOBBLE (or some descendant of it) will serve the needs of all or most of the above projects. The advantage of this is that it would greatly facilitate the merging of the efforts of these projects. For example, if the formal representation of clinical knowledge could be expressed in

GOBBLE, and the strategies produced by the studies of the present illness were compatible with GOBBLE, the two efforts could be readily combined. The results of this combination would be a program with <u>both</u> good strategies for dealing with knowledge, and a detailed data structure which it could use for problem solving.

Although such a "knowledge management" program would be very important, our initial aims for GOBBLE were rather pragmatic. We wanted a program for our immediate needs (writing experimental present illness programs and rudimentary simulations of clinical cognitive process), but we did not wish to undertake a major language development effort, especially when our understanding of the clinical decision-making process was as yet unclear and poorly developed. Hence we opted for the implementation of a flexible representation scheme with a small set of primitives for accessing a knowledge base. This, then, is what GOBBLE is, a way of writing down facts, for 'grouping' facts together, and a set of programs for retrieving facts which have been written in this way and 'digested' by the GOBBLE program.

It is fitting to note the strong similarity of GOBBLE to MAPL 2 [17], a formalism developed by Professor William A. Martin at M.I.T. We have found that many of the ideas Martin had for MAPL 2 were well suited for our work in medicine, and so we incorporated them directly into GOBBLE. Because of our close association with Martin and his research project in Automatic Programming, we expect that GOBBLE will continue to be influenced by the work of that group. Another influence on our thinking has been the CONNIVER language [18] developed by Professor Gerald Sussman and Drew McDermott, also of M.I.T. Our understanding of the issues was considerably enhanced by our experiences with CONNIVER.

Our emphasis on the antecedents of GOBBLE is to underscore the close involvement we have with fundamental computer science research at M.I.T. Our initial design of GOBBLE is only one example of the benefit which accrue to us from this association.

<u>The GOBBLE Program</u>

GOBBLE is a data base handling system which we have written in LISP. The principal features of GOBBLE are: 1) the use of <u>contexts</u> to create 'clumps' of associated facts, and 2) the threading of facts in such a way as to permit the retrieval of expressions representing facts through the specification of subexpressions of these expressions.

A context name is associated with a <u>set</u> of ordered doubles or triples called "valid expressions" where the validity of an expression is determined through checks in a user-built, system maintained <u>dictionary</u>. A GOBBLE context has no inherent significance other than that all facts in a context are marked with the same context name. The same fact (e.g. "(STATUS EDEMA PRESENT)") can appear in many contexts, but in each it will have a unique incarnation. Each incarnation, however, will be recognized by the system as corresponding to the basic pattern. Thus the user can refer either to the generic pattern (e.g. "(STATUS EDEMA

PRESENT)") or to a particular realization of the pattern ("the edema which is present in Acute Glomerulonephritis"). This latter reference would be to "(STATUS EDEMA PRESENT)" in the context "Acute Glomerulonephritis".

It should be noted that the system imposes no overall structure on contexts. By mentioning context names in "subcontext" expressions in other contexts, however, the user can organize an explicit hierarchy of contexts. By mentioning the name of a context in a fact expression in another context, the user creates a link in an implicit network of contexts. (We will give some examples of below.) Of course, it is incumbent upon him to make such a network useful.

A context may contain any number of facts, each one represented by a an expression in GOBBLE form. By creating a context, the user represents a <u>theme</u> for the facts, much as the writer of a book selects the theme around which his presentation is organized. For instance, Acute Glomerulonephritis (AGN) might be the context name, and the expressions associated with it could represent the clinical picture of this disease. Thus it would be a simple matter for a diagnostic program to find out what kinds of things (e.g. sodium-retention) complicated the identification of this disease, and how likely this was to happen. There might also be contexts about edema, hematuria, proteinuria, etc. in which AGN is mentioned, but in which the central theme is the finding in question. Thus various points of view about AGN would be found in individual contexts (representing "clumps" or frames). To this extent, GOBBLE represents information much as do the writers of the chapter cited above. There is a major difference, however, in that in GOBBLE, all these clumps are linked by the through extensive cross-referencing. GOBBLE stores information in a complex <u>association</u> network, and provides functions for the flexible retrieval of facts from this network.

<u>**The GOBBLE Formalism**</u>

The general form of expression for GOBBLE is:

(<function> <argument> <value>)

where the value is optional. In our formalism, facts are equivalent to applications of functions to arguments to produce values. In our current work, we use such "functions" as LOCATION, AMOUNT, CAUSE, FINDING, SUGGESTS, ETC. Thus, for example, to represent the fact that the patient has light proteinuria, we could GOBBLE into the "patient" context an expression for this fact.

```
(GOBBLE PATIENT (AND (STATUS PROTEINURIA PRESENT)
                     (AMOUNT PROTEINURIA LIGHT)))
```

Below, we will show how this new fact can be related to other facts about light proteinuria already in the knowledge base.

As another example, consider the structure:

```
(PREREQUISITE (STATUS STREPTOCOCCAL-INFECTION PRESENT)
              (AND (STATUS STREPTOCOCCAL-EXPOSURE PRESENT)
                   (TIME-OF (STATUS STREPTOCOCCAL-EXPOSURE PRESENT)
                   (BEFORE (ONSET STREPTOCOCCAL-INFECTION)
                           (INTERVAL (WEEK 1.) (WEEK 3.)))))))
```

This is an encoding of the fact that one must be exposed to the streptococcal bacteria a few weeks before the disease develops.

More complex structures can be GOBBLE'd by the system, with the context mechanism serving as the key to bind these structures together. A fragment of a context for AGN is shown in the Figure 8. Here facts about the time relationships of symptoms of the preceding streptococcal infection and a few of the symptoms of AGN.

<u>Pattern-Matching and Fact Retrieval</u>

As noted above, our short term interest in GOBBLE is rather pragmatic, and as a result, we have restricted the development of pattern matching and fact retrieval facilities to a few basic functions. After we have gained experience with these functions and the GOBBLE data structure in the medical projects, we will undertake a more extensive development of these facilities. It seems, however, that our short term needs in the other projects will be reasonably well met by the current version of GOBBLE.

The facilities for pattern based retrieval of facts which we have built into GOBBLE allows the specification of a "theme" for the organization of facts at a time <u>after</u> the facts have been stored. Facts can be retrieved either <u>in</u> a context or <u>through</u> all (or some set of) contexts.

Suppose the piece of advice (suitably encoded in GOBBLE) "The presence of light proteinuria and gross hematuria together suggests either a stone, or a tumor, or recent coagulopathy." were stored in the knowledge base. If the program was given the fact "proteinuria is present", it could find hypotheses about the cause of the proteinuria by using one of the pattern matching programs. Among the suggestions returned would be the one above. Then a dialogue could be initiated to "fill" the pattern:

> What is the amount of the proteinuria?
> LIGHT
> Does the patient have hematuria?
> YES
> Is it gross?
> YES
> etc.

[redacted]

```
(TYPICALLY (*STREP-SYMPTOMS FINDING
                            AGN
                            (AND (STATUS PHARYNGITIS PRESENT)
                                 (STATUS FEVER PRESENT)
                                 (STATUS MALAISE PRESENT))))

(USUALLY (TIME-OF *STREP-SYMPTOMS
                  (AFTER (ONSET STREPTOCOCCAL-INFECTION)
                         (INTERVAL (DAYS 1.) (DAYS 5.)))))

(ALMOST-ALWAYS (*AGN-SYMPTOMS FINDING
                              AGN
                              (AND (NOT *STREP-SYMPTOMS)
                                   (STATUS WEAKNESS PRESENT)
                                   (STATUS ANOREXIA PRESENT))))

(TIME-OF *AGN-SYMPTOMS
         (AFTER (ONSET STREPTOCOCCAL-INFECTION)
                (INTERVAL (WEEKS 1.) (WEEKS 2.))))
```

FIGURE 8. FRAGMENT OF THE AGN CONTEXT

NOTE:
For convenience, GOBBLE permits
expressions to be labelled for later reference. Expressions beginning
with starred words are labelled. The starred word is discarded,
but it is remembered as standing for the rest of the expression.
Later mentions of the name are replaced by the full expression.
We have used this convention in this Figure.

and it would have the tentative hypotheses of stone, renal tumor, etc.

We have begun to integrate GOBBLE into our various projects. For example, we are planning to convert the present illness program to this system., and we are experimenting with the conversion of the formal representation of clinical knowledge to this format. Also the digitalis/digoxin advisor project is using GOBBLE in its preliminary programming. Some further examples of the use of GOBBLE will be presented in the next section when we discuss the time specialist.

<u>Building "Specialists"</u>

Any expert system needs specialists in common sense knowledge. A doctor in addition to needing medical knowledge must know rather everyday things about time , location or quantities. During the process of diagnosis the doctor must be able to understand that if a patient is 25 years old and he was told that when the patient was about 22 years old he had a heart murmur, that it occured three years ago or during 1970-1971.

The GOBBLE system also needs specialists. When asked if there is a mention of edema of the face, the system must respond positively if there is periorbital edema mentioned. This requires that the system know that periorbital edema is located around the eyes and the eyes are part of the face. Many such <u>elementary</u> deductions are required for accessing a large knowledge. The question is how best to provide such a facility.

One solution is to distribute the requirement for such deductions through the system. Another solution, which seems much more promising is to concentrate as much special knowledge about such matters as time, location, etc. in isolated <u>specialists</u>, programs which are expert in the rather shallow deductions needed. Our belief is that most of the questions about time can be answered by a <u>time specialist</u>. The same holds true for location, status, amount, etc. Undoubtedly there will be special questions, in certain contexts, which may be beyond the competence of the specialists, but we think that such questions will be rare.

With these considerations in mind, a time specialist for was developed as part of the GOBBLE framework. First a representation of time expressions was developed. Two different time representations were chosen to be as close to everyday usage as possible . One is absolute time where the time is given as a date and a <u>fuzz</u> factor to describe the uncertainty of the time of the event's occurrence. The format is:

```
(TIME-OF <event> (DATE (19NN NN NN)
                 (FUZZ <days,weeks,months,years> NN)))
```

Where event is either an event such as "(STATUS EDEMA PRESENT)" or an event preceded by either "beginning-of" or "end-of". Beginning-of and end-of are used to specify that an event occured over a period of time longer than a day. If only the beginning-of an event is specified it is assumed to be currently true as in "(BEGINNING-OF LIFE)" . The fuzz is simply the length of time from the date given, that one considers it possible that the event occured and is used in the routines that search the data base.

The other representation for the time of an event is more common in everyday speech, that is the time is given relative to some other event whose time is presumably known. Thus "25 years old" translates to "(AFTER (BEGINNING-OF LIFE) (BY-AMOUNT (YEARS 25.) (FUZZ MONTHS 6.)))". "Exactly three weeks ago" becomes "(BEFORE TODAY (BY-AMOUNT (WEEKS 3.) NIL))". To express the fact that edema occured two weeks after a strep infection one would GOBBLE:

```
(TIME-OF (STATUS EDEMA PRESENT)
         (AFTER (STATUS STREP-INFECTION PRESENT)
                (BY-AMOUNT (WEEKS 2.) (FUZZ DAYS 3.))))
```

<u>What the Time Specialist Does</u>

When a fact is GOBBLE'd in the relative time format the corresponding absolute time is computed and GOBBLE'd, leaving the original alone. In addition when an absolute time is GOBBLE'd the event is put on a "time line" which orders the events on a number line as either points or segments. This time line is used by a function called "SEARCH" which takes one or two dates in the form "(19NN NN NN)" and finds all events that were true during that period regardless of whether they began or ended between those dates.

The other main interrogator of the data base is the function "TIME-OF" which when applied to an event, a time specification identical to that of the time specification for general non-fact rules, i.e. interval instead of amount, and a context, returns the internal identifier of the first fact it finds that meets the time specification which in the case of non-fact contexts is found in that context and is matched in the facts context. For example,

```
(TIME-OF '(STATUS EDEMA PRESENT)
         '(AFTER STREP-INFECTION A-FEW-WEEKS)
         'FACTS)
```

would return "nil" if edema was not a few weeks after the strep infection otherwise the identity of the expression whose TIME-OF edema matched. If the context were say, edema, then the time expression would be searched for in the edema context and matched in facts. (See Figure 9.)

```
========================================================================

The following is a sample conversation with the time specialist.
Lower case letters are typed by the user and upper case by the
the computer.  Comments are preceeded by "**** ".

when he was 21 years old he had a heart attack.

IF YOU THINK THE FOLLOWING IS  RIGHT THEN RESPOND  YES
AND IT WILL BE GOBBLED INTO FACTS.
(TIME-OF HEART-ATTACK (AFTER (BEGINNING-OF LIFE)
                             (BY-AMOUNT (YEARS 21.) (FUZZ MONTHS 9.))))
                   **** This is the Gobble form translated from
                   **** the English.  After the present testing
                   **** stage this will automatically be Gobbled.

THE TIME OF HEART-ATTACK IS WHEN THE PATIENT WAS 21. YEARS
GIVE OR TAKE 9. MONTHS  OLD
                   **** This is the English paraphrasing of the
                   **** Gobble form.

yes                **** The fact in now in the data base.

(cp 'facts)        **** This displays the "FACTS" context.

THE TIME OF HEART-ATTACK IS ABOUT JANUARY 25. , 1973.
GIVE OR TAKE 9. MONTHS
                   **** The date was calculated and Gobbled by
                   **** the time specialist.
THE TIME OF HEART-ATTACK IS WHEN THE PATIENT WAS 21. YEARS
GIVE OR TAKE 9. MONTHS
THE TIME OF THE BEGINNING OF LIFE IS EXACTLY JANUARY 25. ,
1952.

three years ago he had edema.
                   **** The system can presently accept a wide
                   **** variety of ways of expressing when
                   **** something happened.

IF YOU THINK THE FOLLOWING IS  RIGHT THEN RESPOND  YES
AND IT WILL BE GOBBLED INTO FACTS.
(TIME-OF EDEMA (BEFORE TODAY
                       (BY-AMOUNT (YEARS 3.) (FUZZ MONTHS 9.))))
THE TIME OF EDEMA IS 3. YEARS GIVE OR TAKE 9. MONTHS AGO

yes
========================================================================
```

FIGURE 9. Session With The TIME SPECIALIST

```
***********************************************************************

(time-of 'heart-attack
                 '(before today (interval (years 2.)
                                          (months 4.))) 'facts)
                        **** Here the computer is asked if
                        **** the heart attack occured between
                        **** two years and four months ago.

YES, THE TIME OF HEART-ATTACK IS ABOUT JANUARY 25. , 1973.
GIVE OR TAKE 9. MONTHS

(*why)                  **** Here the system is asked why it
                        **** answered as it did.

I WAS ASKED IF THE TIME OF HEART-ATTACK WAS BETWEEN 2. YEARS
AND 4. MONTHS AGO
SO FIRST I TRIED TO FIGURE OUT WHAT YOU MEANT BY BETWEEN 2.
YEARS  AND 4. MONTHS AGO
THE EARLIEST POSSIBLE TIME TODAY COULD HAVE BEEN IS 1974. 2.
27.  AND THE LATEST POSSIBLE TIME IT COULD HAVE BEEN IS 1974.
2. 27.
I THEN FIGURED OUT THE DATES WHICH WERE BETWEEN 2. YEARS
AND  4. MONTHS BEFORE 1974. 2. 27.  AND 1974. 2. 27.
WHICH WERE 1972. 2. 27.  AND  1973. 10. 27.
THEN I CHECKED TO SEE IF THE TIME OF HEART-ATTACK FELL INTO
THE RANGE BETWEEN 1972. 2. 27.  AND  1973. 10. 27.
THE EARLIEST TIME HEART-ATTACK COULD HAVE HAPPENED IS 1972. 4.
25. AND THE LATEST IT COULD HAVE HAPPENED IS 1973. 10. 25.
SO I DECIDED THAT
THE TIME OF HEART-ATTACK PROBABLY WAS BETWEEN 2. YEARS  AND 4.
MONTHS AGO

(setq time-of-talk-sw t)    **** This tells the system to
                            **** "think outloud".

(time-of 'edema
         '(after (beginning-of life) (interval (years 20)
                                               (years 18))) 'facts)

I WAS ASKED IF THE TIME OF EDEMA WAS WHEN THE PATIENT WAS
BETWEEN 20.  AND 18. YEARS OLD
SO FIRST I TRIED TO FIGURE OUT WHAT YOU MEANT BY
WHEN THE PATIENT WAS BETWEEN 20.  AND 18. YEARS OLD
THE EARLIEST POSSIBLE TIME THE BEGINNING OF LIFE COULD HAVE
BEEN IS 1952. 1. 25.  AND THE LATEST POSSIBLE TIME IT COULD
HAVE BEEN IS 1952. 1. 25.
I THEN FIGURED OUT THE DATES WHICH WERE BETWEEN 20. YEARS
AND  18. YEARS AFTER 1952. 1. 25.  AND 1952. 1. 25.   WHICH
WERE 1970. 1. 25.  AND  1972. 1. 25.
THEN I CHECKED TO SEE IF THE TIME OF EDEMA FELL INTO THE RANGE
BETWEEN 1970. 1. 25.  AND  1972. 1. 25.
THE EARLIEST TIME EDEMA COULD HAVE HAPPENED IS 1970. 5. 27.
AND THE LATEST IT COULD HAVE HAPPENED IS 1971. 11. 27. SO I
DECIDED THAT
THE TIME OF EDEMA PROBABLY WAS WHEN THE PATIENT WAS BETWEEN
20.  AND 18. YEARS OLD

***********************************************************************
```

FIGURE 9. Continued

Note: Patient is known to have been born on
January 25, 1952, and the discussion is being held
on February 27, 1974.

Research on the Time Specialist and Other Specialists

Although the time specialist deals well with rudimentary questions about time, some additional work is needed to expand its capabilities. One of the most important problems is to incorporate into it some understanding of rates. For example, it should understand such statements as

The onset of the disease is abrupt.
Usually the disease develops insidiously.
The hypertension subsides slowly after the diuresis.
etc.

Now it is clear that in certain circumstances, even doctors would have difficulty saying exactly what these statements mean. So we are not proposing to equip the time specialist with more than human expertise. On the other hand, we can get very good agreement on what these statements do not mean. For example, if the symptoms of the disease mentioned in the first statement appear over a two week interval, then we would not call the onset abrupt. Similarly, we would not call the development of a disease within a few weeks insidious. The time specialist should be aware of these distinctions, too.

It is very important to realize that even rough definitions of these concepts will allow the time specialist to answer a great many questions. People have developed these concepts and have used them successfully because in most instances, their exact definitions do not matter. If someone tells you that an event will occur "within a few days", you may find that acceptable, never ascertaining whether two days, three days, or more is meant. The language of medicine is rich in terms which are understood, but never precisely defined. In certain instances, this lack of precise definition can be troublesome, but for the most part, a rough idea, commonly shared, of the meaning of the concept is sufficient.

We propose to pursue our research on the time specialist and other specialist with such a bias. The goal will be to equip each specialist with just enough knowledge to permit a reasonable discussion with a clinician. The program should answer the questions of the clinician directly even when they contain vague phrases of the type mentioned above. The goal is to have the specialist have trouble only when most people would have trouble in interpretting a question.

In addition to the problems associated with rates, we want to look at another important problem for the time specialist. This is the concept of episodes. In a sense, this problem belongs in the domain of representation work as well as here in the province of the time specialist. In any event, the representation and understanding of episodic disease is very important, and will require considerable research before a good solution can be developed. Basically we need a mechanism to describe the "prototypical" episode and the time intervals between occurrences of episodes. For certain instances, this is quite straightforward, but for other situations, this is quite difficult. Because we have just begun to work on this problem, we cannot discuss it further here, other than to note that it will receive careful attention in the near future.

Inquiry and Explanation

The development of markedly improved facilities for inquiry and explanation is one of the central computer science research projects of the proposed Laboratory. The importance of such facilities should be recognized, because without them, it is doubtful whether a large, knowledge-based program can be built for a complex clinical problem. The construction of such a program will require three things:

1) understanding of the processes of clinical cognition
2) mechanization of a very large amount of knowledge
3) development of new programming concepts and technology

The achievement of the first two goals will require the close collaboration of clinicians and computer scientists. The former must be able to actively work with the computer realizations of the cognitive theories, and they must also be able to explore the knowledge base of the programs in use. Hence, the clinicians will need direct interaction with the developing system. Further as the system grows, computer scientists as well will need such access. As the system grows in complexity, it must be able to answer questions about its knowledge and performance.

Further, if we look to the day in which such systems are introduced into the health care system, we see the additional need for such facilities. It is unreasonable to expect that clinicians will accept advice from such a system about a serious problem without any access to the knowledge or reasoning upon which the advice is based. In addition, this explanation of the reasoning of the system must be in terms which the clinician can understand.

So for our own immediate needs, and for the long run needs of the field, we will actively pursue research in both inquiry and explanation. Of the two, explanation will receive the most attention. The reason for this is that other researchers at M.I.T. are vigorously pursuing natural language research. This research has already led to significantly improved parsers. We plan to adopt one of these parsers when it has reached a satisfactory state of development. We plan to invest only enough time and resources to assure that the special needs of an interface designed for clinicians can be accomodated by the parser we select.

As an example of this policy, consider the English language facility used in the dialogue with the time specialist. The parser used there is called LINGOL (19) and it was developed by Professor Vaughn Pratt of M.I.T. We found that we could easily adapt it to our needs, and that it provides us with a reasonable interface. Certainly, there are problems which it doesn't handle, but we will leave most all of these problems to the language researchers. For the small effort involved in adapting it to our needs, LINGOL has returned considerable benefit. Other language research at M.I.T. may yield even better facilities. If so, we will be able to further improve our interface with the clinicians, and thereby improve our ability to achieve our research objectives.

The matter of explanation, however, is one to which we will put more effort. Because of its importance, and because it appears to be a problem in which we are more interested than other computer science researchers, we feel that we must take more of a lead in research. To this end, we have undertaken the development of an explanation facility to incorporated into GOBBLE.

Now the first issue to be considered is what constitutes an adequate explanation. In certain instances, simply retrieving a fact may suffice. In other cases, the explanation may require the use of a model (e.g. of physiology). In still other cases, a dialogue may be required to accomplish the task.

As a first step, we have developed a rudimentary "explainer" for GOBBLE. It gives simple explanations of relationships in a knowledge base. An example of an interaction with this explanation program is shown in Figure 10a. The program is following back through a causal chain to produce an explanation of the findings of nephrotic syndrome. Although it is not apparent from this particular interaction, this program differs from one which has a "canned" response in that it can enter this chain at any point as requested by the user. A more interesting example of this capability is shown in Figure 10b.

[redacted]

```
==========================================================================

 AT ANY TIME YOU MAY ENTER THE QUESTION MODULE BY TYPING A ?
   FOLLOWED BY A SPACE. A  RETURNS TO PROGRAM.
 I HAVE FOUND 2. POSSIBLE CONTEXTS
 1. NEPHROTIC SYNDROME
 2. GENERAL
 WHICH DO YOU WANT TO REFERENCE ?

>>> 1

 EDEMA CAN OFTEN BE OBSERVED IN NEPHROTIC SYNDROME
 THE BASIC CAUSE OF EDEMA IN NEPHROTIC SYNDROME IS MASSIVE PROTEINURIA
 WOULD YOU LIKE TO KNOW MORE ?

>>> Y

 EDEMA IS DIRECTLY CAUSED BY INCREASED INTERSTITIAL FLUID
 WOULD YOU LIKE TO KNOW MORE ?

>>> Y
 INCREASED INTERSTITIAL FLUID IS DIRECTLY CAUSED BY RENAL SALT AND
 WATER RETENTION
 WOULD YOU LIKE TO KNOW MORE ?

>>> Y

 I THINK THAT THERE ARE 2. POSSIBLE CAUSES OF RENAL SALT AND WATER
   RETENTION
 1. DECREASED GLOMERULAR FILTRATION RATE
 2. INCREASED MINERALOCORTICOID SECRETION
 WHICH WOULD YOU LIKE TO KNOW MORE ABOUT ?

>>> 2

 INCREASED MINERALOCORTICOID SECRETION IS DIRECTLY CAUSED BY
 DECREASED PLASMA VOLUME AND DECREASED CARDIAC OUTPUT
 BY THE WAY DECREASED PLASMA VOLUME AND DECREASED CARDIAC OUTPUT
 ALSO CAUSES DECREASED RENAL BLOOD FLOW
 WOULD YOU LIKE TO KNOW MORE ?

>>> Y

 IN NEPHROTIC SYNDROME
 DECREASED PLASMA VOLUME AND DECREASED CARDIAC OUTPUT IS DIRECTLY
 CAUSED BY DIFFUSION OF WATER AND SOLUTES TO INTERSTITIAL FLUID
==========================================================================
```

FIGURE 10a. An explanation from the knowledge base

```
**************************************************************************
WOULD YOU LIKE TO KNOW MORE ?

>>> Y

DIFFUSION OF WATER AND SOLUTES TO INTERSTITIAL FLUID IS DIRECTLY
CAUSED BY DECREASED ONCOTIC PRESSURE
WOULD YOU LIKE TO KNOW MORE ?

>>> Y

DECREASED ONCOTIC PRESSURE IS DIRECTLY CAUSED BY HYPOPROTEINEMIA
WOULD YOU LIKE TO KNOW MORE ?

>>> Y

HYPOPROTEINEMIA IS DIRECTLY CAUSED BY MASSIVE PROTEINURIA
BY THE WAY MASSIVE PROTEINURIA IS ASSOCIATED WITH
HYPERCHOLESTEROLEMIA AND LIPIDURIA
WOULD YOU LIKE TO KNOW MORE ?

>>> Y

I DO NOT KNOW ANY MORE
**************************************************************************
```

FIGURE 10a. Continued

```
YOU HAVE JUST ENTERED THE QUESTION MODULE. TYPE ANY QUESTION THAT
 YOU LIKE TERMINATED BY A "?" FOLLOWED BY A SPACE.
 PLEASE CONTINUE

>>> what causes decreased renal blood flow ?

   THE MOST IMMEDIATE CAUSE OF DECREASED RENAL BLOOD FLOW IS DECREASED
   PLASMA VOLUME AND DECREASED CARDIAC OUTPUT
   PLEASE CONTINUE

>>> what is the usual result of decreased oncotic pressure ?

   THE RESULT OF DECREASED ONCOTIC PRESSURE IS DIFFUSION OF WATER AND
   SOLUTES TO INTERSTITIAL FLUID
   PLEASE CONTINUE

>>> what could cause decreased renal blood flow ?

   THE MOST IMMEDIATE CAUSE OF DECREASED RENAL BLOOD FLOW IS DECREASED
   PLASMA VOLUME AND DECREASED CARDIAC OUTPUT
   PLEASE CONTINUE

>>> what could cause a decreased glomerular filtration rate ?

   THE MOST IMMEDIATE CAUSE OF DECREASED GLOMERULAR FILTRATION RATE IS
   DECREASED RENAL BLOOD FLOW OR DAMAGE TO GLOMERULI

>>> what is the possible cause of hypoproteinemia ?

   THE MOST IMMEDIATE CAUSE OF HYPOPROTEINEMIA IS MASSIVE PROTEINURIA
```

FIGURE 10b. Explanations in inquiry mode

Methods of Procedure

Introduction

Much of the work reviewed above is already underway. Some activities are more advanced than others, but all the projects discussed are receiving the attention of at least one member of our group. In most cases, most of the members of the group are involved in at least some aspect of each project. We expect that this mode of operation will be common in the Laboratory, and as a result, it is not a simple matter to give a detailed timetable for each project. The researchers in our group will naturally tend to shift their attentions somewhát to those problems which loom most prominantly at any point in time. We believe that this flexibility will prove tremendously beneficial to the Laboratory, but it, coupled with our present uncertainty about the degree of difficulty each project will manifest, makes our current projections only informed guesses.

Nonetheless, we offer here our best guesses as to the course the research of the Laboratory will take. As our work proceeds, we will undoubtedly modify these plans in the light of new problems and developments.

Present Illness Project

Because of its complexity, it is most difficult to chart the course of the present illness project. The broad outlines are clear, but the details are hard to discern at this point in time.

For the next six months or so, we will continue our detailed analysis of the problem-solving behavior of a few renal experts. The procedure we will use will include protocol analysis and close man-machine interaction involving a computer simulation of cognitive process. This approach has been quite successful so far, and we expect it will become one of the major methodological tools of the Laboratory.

The work on the simulation program for the present illness will remain focused on the presenting problem of edema during the next six months. We believe that a very detailed study of the way in which one or two experts deal with this one problem will prove extremely useful and interesting.

Within a year, we will have a simulation of this behavior which is rather complete, in that the program can take a a present illness for edema which will deal with all the major issues outlined in the above discussion (e.g., pattern-matching of signs and symptoms, finding a specific context for the problem, "backing up" in the face of failure, etc.) in at least a preliminary way.

We cannot expect that the program will take a present illness of edema which is fully comparable to that which would be taken by an expert.

The major problem, as we see things now, is not so much a matter of strategy (although some knotty problems are apparent), but rather the amount of real world knowledge which the expert uses. Thus the program might do quite well on one problem, but on a second problem, it might "fail" because it didn't know that "waitresses who stand up all day often get swollen ankles at night".

At this time (approximately July, 1975), we expect to produce a paper aimed a a medical audience which discusses the cognitive theory we have developed, and the implications of this theory with respect to such issues as the assessment of problem solving skill, medical education, etc. This paper will draw on the study of cognitive style which at this point should have produced some new and interesting results. (Of course, this may be best presented in a separate paper.) The second major paper will be focused on the use of computer science methodology in cognitive theory formulation.

At this point, we expect that our experiences of the first six to eight months will prompt us to undertake a re-design of the simulation program, and will help us structure the "knowledge acquisition" problem so that several teams can be set to work on it. During the year 1975-1976, the emphasis should be on the broadening and deepening of the knowledge base for the program. If large areas of knowledge can be dealt with by separate groups, our work should proceed much more rapidly.

Here we expect that the work on the formalization of clinical knowledge will begin to yield great benefits. By this time, a scheme for codifying knowledge should be available, and a "compiler" for knowledge expressed in this scheme will have been developed. This will greatly facilitate the expansion of the knowledge base of the simulation program.

It should also aid in the exploration of another medical area. During this year (1975-1976), we expect to begin a similar project in a different medical speciality (perhaps cardiology). We would be interested in assessing the usefulness of our theories and concepts in a different area. Although we expect that some modifications will be required, we believe the bulk of the theory will apply.

By July, 1976, we expect to have built sufficient knowledge about the present problems of edema, hematuria, etc. into the present illness program that its performance can be meaningfully compared with that of clinicians of various skill levels. Such comparisons will involve detailed studies of the protocols of the clinicians and the trace of the program on the same cases.

Undoubtedly, this study will also point out deficiencies in theory and in the program. The direction of this research beyond this point will be determined in large part by the outcome of tests such as this. At this point in time, we can say little other than that the basic effort will be directed at expanding the theory and developing the program.

As we proceed, however, we will make a concerted effort to publicize successes of the Laboratory and to find ways to make these successes available to researchers in other centers. One way in which we will do this is through publications; another way may be through the ARPA network. A third way is through conferences and research meetings. The point is that our proposed work touches on so many central issues that it will be to our advantage and to the benefit of others for us to maintain close contacts with the existing research community in computer science and medicine.

Digitalis Advisor

It is anticipated that the central mathematical algorithm will be implemented and packaged in simple routines for limited physician use within six months' time. Programming of criteria for speed of administration, interpretation of therapeutic and toxic effects and searches for factors influencing sensitivity should take an additional two to three months, with allowance for an additional two months to create a crude set of programs to facilitate more extensive physician interaction with the model. Thus, by April, 1975, we would hope to have a crude program available for testing by physicians both in our Laboratory and possibly in limited areas of the hospital. We would envision this initial testing phase to encompass about three months time, and then another three months for further program development before a second stage program is available for testing. At that stage, we would hope to be able to begin testing effectiveness among non-expert physicians. We would plam that this trial include some of our surgical colleagues, who deal with patients requiring this drug.

This test of effectiveness will require careful study of the decision-making of clinicians and surgeons both before and after their introduction to the program. This raises the question of how one should measure the effectiveness of clinical decision-making, and we will have to give this question careful thought. The particular problem we have chosen, however, may make this problem somewhat less troublesome, because over a sufficient number of trials, the toxic/therapeutic response of the patient can be taken as the prime indicator of effectiveness of decision-making.

Papers recounting the development of the program and the experience with it in the clinical setting will be prepared at this time. Further, steps will be taken to provide the program to other researchers for their use and evaluation.

If this project is successful, we plan to initiate another "model-based" effort such as the administration of antibiotic therapy or the like to gain more experience, and to test our ability to transfer the technology and understanding we have gained to other problems.

GOBBLE Development

By introducing GOBBLE into the various projects which are underway, we expect to learn a great deal about its limitations. Some are already known to us, because we have made a conscious decision to defer the development of certain features of the system until we have more experience with medical problems. Others will arise in the course of the research in the various projects. Thus at present, we can only give a rough time-table for the development of the system.

The basic development of GOBBLE should be complete within the next six months. That is, by December 1, 1974, we should have the first version in sufficiently de-bugged and polished state that it can be "frozen" and it can be a major tool in the program development activities of the Laboratory. The features of this first version of the system will be:

1) An improved facility for stringing sub-contexts together
2) Semantics for specifying retrieval searches through various contexts and subcontexts
3) Facilities for specifying "a-kind-of" relationships (e.g. pedal edema is a kind of edema) such that the subclasses automatically take on the properties of the main class unless otherwise indicated
4) A rudimentary capability for responding to questions about the knowledge base
5) An improved dictionary facility to automatically check new additions to the knowledge base for obvious errors (misspellings, etc.) and obvious contradicitons

At this time, a small manual will be written on the use of the system, and it will be formally introduced into each of the projects. For a period of three months, we will record problems and failings in the system. After this trial period, several decisions will be made.

First, we will decide whether GOBBLE is a viable and useful concept. At present, we believe that it almost certainly will prove to be one. It may prove more useful for some projects than for others, however, and at this point, we will decide which projects should continue to use the system.

From the recorded problems with the system and from our general understanding of its limitations, we will identify the most important additions to and revisions of the system which are required, and undertake a new design. Into this design, we will incorporate the results of the three projects described below, the specialists project, the explanation and inquiry project, and the interface project. This new implementation should be completed within a month or so, and then GOBBLE will be a basic part of the work of the Laboratory, with revisions being made as necessary by members of the staff.

A detailed description of the system with examples of its applications in the medical project will be issued by the Laboratory about six months after the second implementation of the system.

In addition to further work on the time specialist, the development of other specialists will be undertaken. The current choice for the next project is the <u>location specialist</u>. This program will manage the common sense knowledge about the parts of the body and their locations relative to one another. This specialist will know the difference between the inside and the outside of the body as well. In large, the location specialist will be like the time specialist. Instead of a time-line for organizing facts, the location specialist will maintain a model of the body, and it will organize statements about locations around this model.

We expect that a first version of the location specialist can be developed with eight months, and so by December, 1974, this specialist, and the improved time specialist should be available in the second version of GOBBLE. Although other specialists will be developed, we cannot say at this time how many there will be, or in what order they will be built.

Further developments of GOBBLE or its descendants will flow from the use of this technology in the medical projects. Their needs will determine the efforts in this area.

<u>Significance of the Research</u>

The impediments to the use of computer science and technology to favorably influence the quality and the quantity of health care available to the community are large and complex. These impediments will not fall to simple extensions of past work, rather new, more powerful combinations of resources and people will be required. The most immediate significance of the proposed laboratory is that it can focus the attentions of first rate medical scientists and computer scientists on one of the most important of these problems, the lack of a well-articulated theory of clinical cognition. Further the efforts of these researchers can be built on the base of the most advanced technology and methodology of its kind in existence.

The development of such a theory and the successful application of the technology which will be developed in concert with the theory will radically alter the way in which expert physicians can interact with programs, and the kind of expertise these programs can have. Further the technology which results will allow an attack on many clinical areas by other workers. Thus we see the techniques and facilities which will result from our research as being the vital first step on the road to creating distributable expertise in the form of specialist consultant programs.

In this way the physician dealing with even the most complex problems in a site remote from consultants could be assured of guidance that would allow him to enormously upgrade his performance. The expectation is not that the local physician can perform at a level equal to the best consultant but simply at a level approaching that of the expert, a level far above that generally achieved today.

Beyond the use of programs such as these, and perhaps even more significant in the long run, lies the prospect of analogous programs being prepared for the support of allied health personnel in the delivery of primary medical care. Such support is vital, because even if the current shortage of physicians can be overcome, it is unlikely that the problem of maldistribution of physicians will be resolved. Few physicians wish to practice in the rural areas (consisting of nearly 40 million people without adequate access to physicians) nor in the inner city where tens of millions more face a similar problem. For this reason it seems to be highly likely that new classes of allied health personnel must be trained to fulfill the primary care functions. Such personnel must, if they are to be accepted by the patient, be able to provide care of good quality. Current programs for use of allied health Personnel, such as the MEDEX effort, promise quantity but cannot provide quality and it is here that the computer can make its contribution.

Once the basic problems related to computer-support of the physician have been worked out, as described in the present proposal, it should be possible through utilization of this knowledge and experience to develop programs geared to the needs of the allied health professional in his triage function-making as certain as possible that he does not overlook serious disease and restraining him from taking on complex problems beyond his capability. These programs could also provide him with the assistance necessary for dealing with crises under circumstances in which a transfer of the patient is not feasible.

We realize, that most patients coming to most primary care physicians (or or new kinds of allied health personnel envisioned as delivering primary care) do not have serious diseases and that a wide range of relatively simple algorithms will be necessary to assist in the care of the patient. Nevertheless, these procedures must be organized within the context of a knowledgable system in order to insure their correct application. Our studies and those being pursued at the Massachusetts General and Beth Israel Hospitals and elsewhere should complement each other . Thus in the long term we believe that our work can assist in solving our manpower and quality problem by contributing to an understanding of the use of the computer in serious management problems by both physicians and non-physicians.

A second major benefit of this research is its potential impact on medical education. The development of clearly understood theories of expert knowledge and its application is a major goal of our effort. Although it is undoubtedly true that effective decision-making is one of the central factors in clinical practice, little, if any, attention is directed to this subject in current medical education. Most medical students are forced to infer from their observation and experience the general principles of diagnostic and therapeutic decision-making. At present there exist no well-articulated theories of medical decision-making, and it is very difficult for the average medical student to become a good problem-solver.

We believe that our work will result in extensive new knowledge of the way in which clinical experts solve problems, and further it will suggest many new ways in which students can be introduced to the processes upon which expertise is built. Rather than simply being a collection of facts about the medical problem in question, programs will provide procedures for <u>solving</u> the problem, and students can study and interact with these programs. Such procedures, supported by additional reference material, organized in more <u>associative</u> ways, will allow the student to enlarge his understanding of a given area.

A further benefit which will result from the activities of the Laboratory will be the training of computer science graduate students to work with clinicians on important research questions, and in turn the Laboratory will offer clinicians the opportunity to learn about the methodology of computer science. We believe that the Laboratory will be the basis for a whole new area of collaborative research and education, an area which can greatly benefit society.

The Management of the Laboratory

As Principal Investigator and Director of the Laboratory, Professor Gorry ultimately will be responsible for all activities of the Laboratory, both scientific and administrative. Because of the interdisciplinary nature of the activities of the Laboratory, Professor Gorry will draw on the advice and assistance of key senior people in both medicine and computer science. Dr. Schwartz has accepted the responsibility for overseeing the medical aspects aspects of the research, and he will be the Deputy Director of the Laboratory. His judgments concerning the medical importance and relevance of projects will be a key factor in determining the directions in which our efforts go.

Professors Fredkin and Minsky will help with the development and maintenance of close relations between the Laboratory and Project MAC and the Artificial Intelligence Laboratory.

One of the goals of the Laboratory will be to promote a real community devoted to research on computer science and clinical decision-making. The facilities and research programs of the Laboratory represent on nucleus about which such a community could be centered. Through a concerted effort to publicize these facilities and resources, we will establish relationships with individuals and groups who are already active in this area or who could be fruitfully encouraged to become active. A variety of relationships between the Laboratory and these individuals and groups will be explored. We expect that some relationships will be very close, while others will be quite loose.

We believe that it will be to the advantage of the research programs of the Laboratory to develop such contacts, and in certain cases, to grant the use of some of its resources to researchers who are technically outside it. We would like to accept certain proposals from research outside the Laboratory to use resources of the Laboratory, particularly the computer. If such a proposal were in keeping with the broad aims of the Laboratory, and if the required resources were available, it would be accepted.

As an extension of the above idea, we would consider inviting certain researchers to come to the Laboratory for a period of time ranging from a few days to a few months. These guests would be chosen for the potential of the contribution they could make to the programs of the Laboratory. Such contributions might be lectures or consultations with staff and students. These visitors would also provide a good source of criticism of our activities, either from a medical or from a computer science point of view.

Because we believe that informed criticism is very valuable, we plan to form a small visiting committee composed of three or four respected computer scientists and physicians from other institutions. They would come to the Laboratory for a day or two every six months to review and criticize our activities. We feel that careful consideration of our work by this committee will be extremely valuable.

If it is possible, we would like to hold some form of conference once each year on computer science and clinical decision-making at the Laboratory. Currently, we envision this as a working research conference attended by people who are active in the field. We also will encourage Laboratory staff to prepare papers for conferences and publication as appropriate to help transfer the ideas and technology of the Laboratory to others in the field.

Facilities

The Laboratory computer will be directly linked to 4 large time-sharing computer systems at M.I.T.: the MULTICS system which is owned by M.I.T. and operated by the Information Processing Center, and 3 compatible PDP-10 systems, 2 at Project MAC and one at the Artificial Intelligence Laboratory. Through this connection, we will have direct access to an impressive array of software including an advanced operating system and programming languages such as LISP. These languages will operate on all these systems.

All these machines are linked to the ARPA network, and thus are accessible to researchers and general users at 25 other locations. We plan to connect our machine to this network as well to facilitate use of our technology by selected researchers at other institutions.

In addition to these computers *per se*, we can draw on a large reservoir of computer talent. The Laboratory will be located in the same building with Project MAC and the Artificial Intelligence Laboratory, and many members of these two research efforts have an active interest in our work. Further, we expect to attract some very good graduate students in computer science by virtue of our close proximity to these laboratories and the inherent appeal of our research program.

Further, the Laboratory will have access to a library of computer science publications, a printing and reproduction section, an electronics shop, and a machine shop, all housed in the same building with the Laboratory.

The primary offices of the clinical members of the effort will be located at the New England Medical Center Hospital. The Hospital is a general hospital consisting of about 400 beds. This private, non-profit university hospital has 11,000 admissions per year and 140,000 out-patient visits per year. Approximately 30% of these out-patient visits are handled by the Department of Medicine. The in-patient Medical Service is divided into units of 15 beds each, each of which has a professional staff consisting of an attending physician, an assistant resident, an intern, and two medical students. One or more of these units will serve as a test environment for programs developed in the

Laboratory.

As Physician-in-Chief, Dr. Schwartz has control of the beds in the hospital. In addtion, Dr. Kassirer is the Director of the House Staff Training Program. Both these facts should greatly facilitate the interaction of the research program of the Laboratory with the clinical environment.

Principal Investigator Assurance

The undersigned agrees to accept responsibility for the scientific and technical conduct of the research project and for provision of required progress reports if a grant is awarded as the result of this application.

G. Anthony Gorry 3/22/1974

Principal Investigator

References

1) Schwartz, W.B., "Medicine and the Computer: The Promise and Problems of Change", New England Journal of Medicine, 283: (December 3, 1970), pp. 1257-1264

2) Newell, A., and Simon, H.A., Human Problem Solving, Prentice Hall, Englewood Cliffs, N.J., 1972

3) Gorry, G.A. and Barnett, G.O., "Sequential Diagnosis by Computer", Journal of the American Medical Association, 205: (1968), pp. 849-854

4) Schwartz, W.B., et al, "Decision Analysis and Clinical Judgment", American Journal of Medicine, 55: (October, 1973), pp. 459-472

5) Gorry, G.A., "Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure", American Journal of Medicine, 55: (October, 1973), pp. 473-484

6) Weed, L.L., Medical Records, Medical Education and Patient Care, Case Western Reserve University Press, Cleveland, 1970

7) Slack, W.V., et al, "A Computer-Based Medical History System", New England Journal of Medicine, 274: (1966), pp.194-198

8) Bleich, H.L., "Computer Evaluation of Acid-Base Disorders", Journal of Clinical Investigation, 48: (1969), pp. 1686-1696

9) Ginsburg, A.S., "Decision Analysis in Clinical Patient Management", R-751-RC/nlm, Rand Corporation, Santa Monica, 1970

10) Winograd, T., "A Procedural Model of Language Understanding", in Schank, R.C. and Colby, K.M. (eds.) Computer Models of Thought and Language, Freeman, San Francisco, 1973

11) Sussman, G.J. "A Computational Model of Skill Acquisition" M.I.T. Ph.D. Thesis, 1973.

12) Minsky, M. and Papert, S., Progress Report of the M.I.T. Artificial Intelligence Laboratory, January 1, 1972

13) Minsky, M. "Frame-Systems: A System for the Representation of Knowledge", unpublished paper, 1973.

14) Jelliffe, R.W., Buell, J., and Kalaba, R., "A Computer Program for Digitalis Dosage Regimens" Math Biosci, (1970), pp. 179-193

15) Doherty, J, "Digitalis Glycosides: Pharmacokinetics and their Clinical Implications", Ann Intern Med, (1972), pp. 891-906

16) Jelliffe, R.W., Buell, J., and Kalaba, R., "Reduction of Digitalis Toxicity by Computer-assisted Glycoside Dosage Regimens", Ann Intern Med, (1972), pp. 891-906

17) Peck C.C. and Sheiner L.B., "Computer-assisted Digoxin Therapy", New England Journal of Medicine, (1973), pp. 441-446

18) Martin, W.A., "MAPL 2", Automatic Programming Group Memo., Project MAC, 1973.

19) Sussman, G.J. and McDermott, D.V. "From PLANNER to CONNIVER: A Genetic Approach", Proceedings of the Fall Joint Computer Conference, 1972, pp. 1171-1179

20) Pratt, V., "A Linguistics Oriented Programming Language", Artificial Intelligence Memo No. 277, Febuary 1973.