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 *clinically* 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.

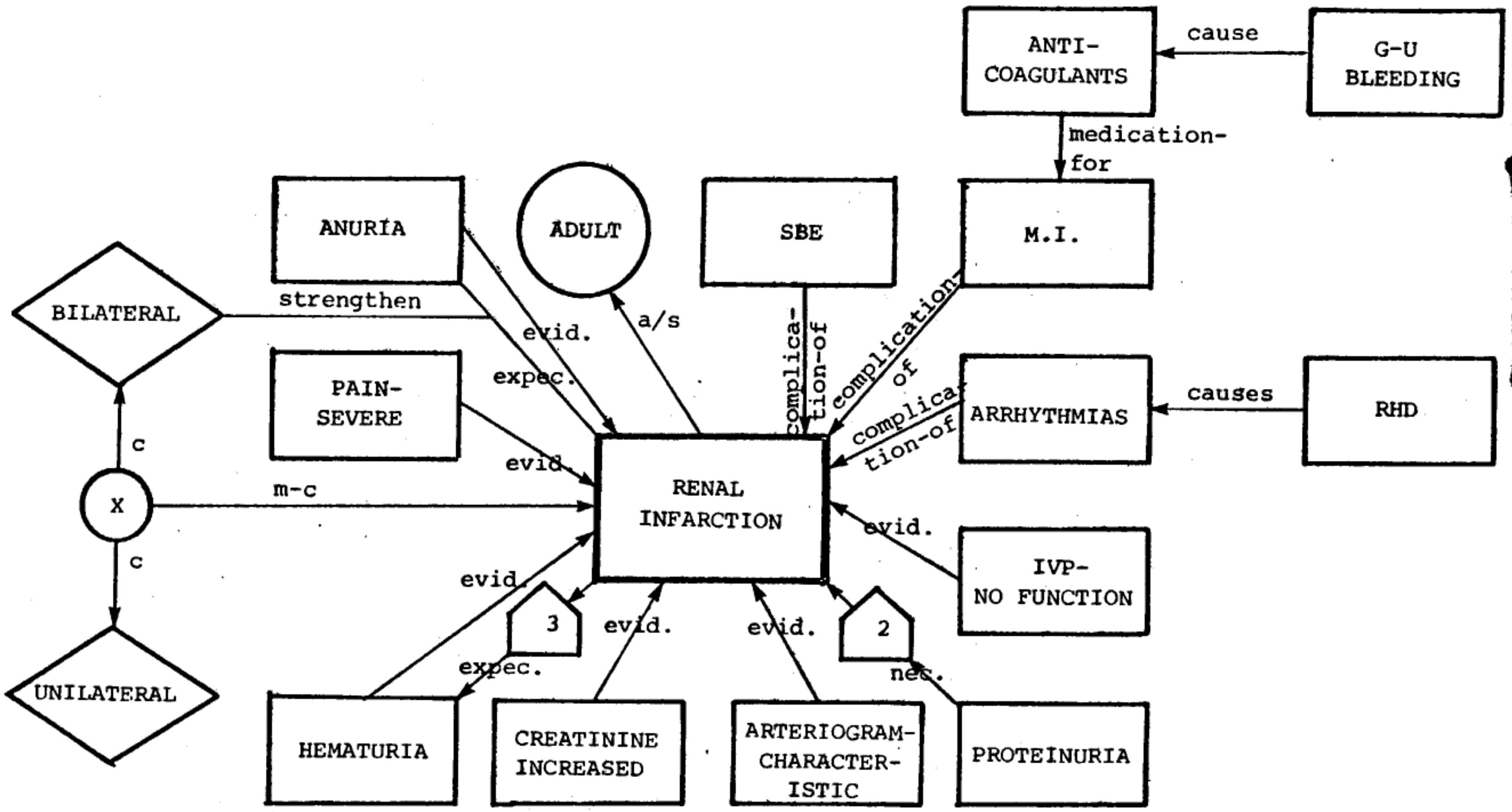
The CPC was presented to him <u>one fact at a time</u>. After each fact was given to him, he was asked to discuss the "meaning" of the fact. The meaning of the fact to him included the immediate conclusions which he could draw from it, its effects on hypotheses currently being considered, its suggestions 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 <u>introspective</u> 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 <u>slices</u> 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.



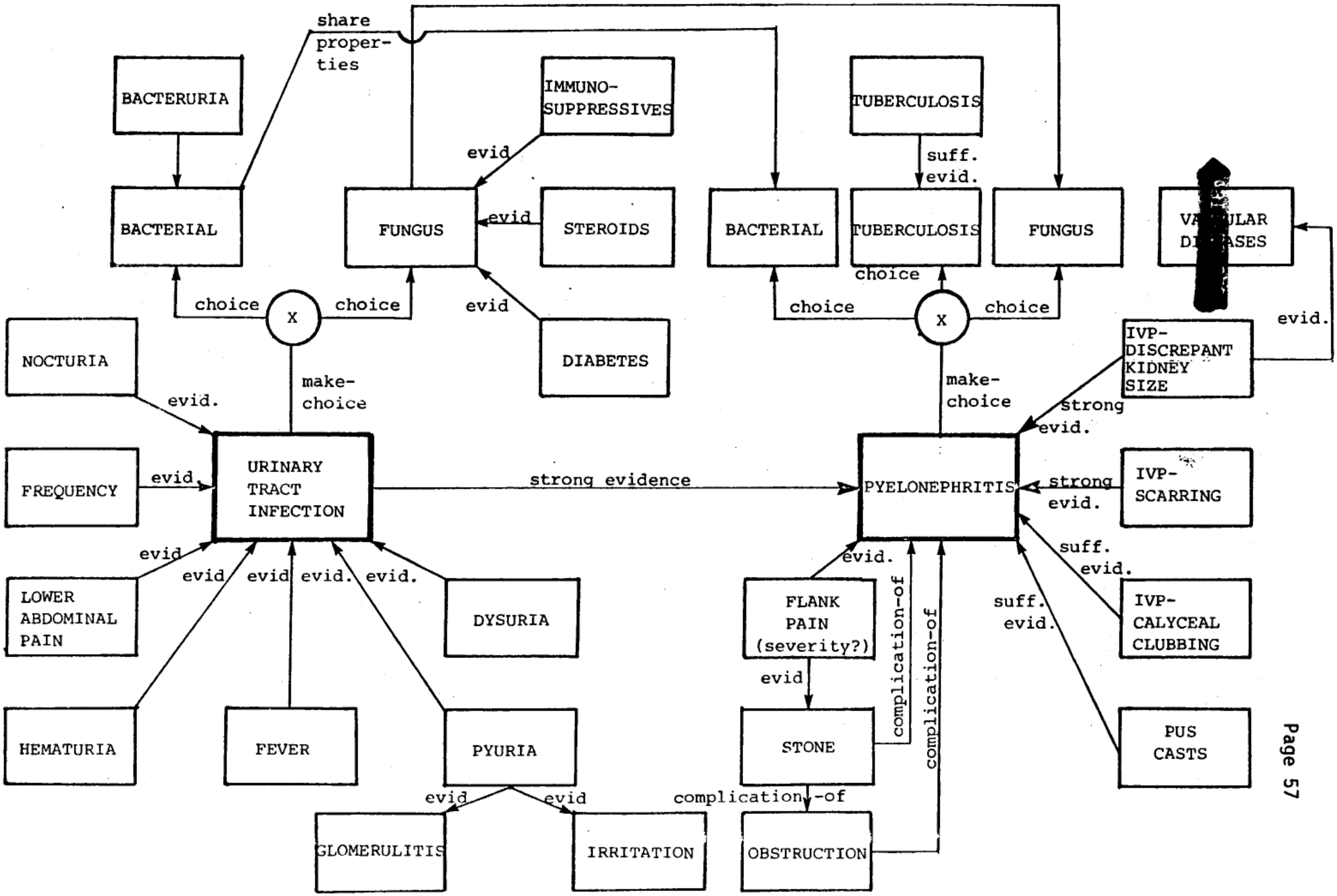
RENAL INFARCTION

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 <u>obtaining</u> 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) <u>do something</u> 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 <u>explanations</u> 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.

The Digitalis/Digoxin Therapy Advisor

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 clinical use.

Background

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 clinical signs of digitalis excess are cardiac (disturbances of cardiac rhythm) and extra-cardiac (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 complicating factors 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.

Recent Advances

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 in vivo 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 brand to brand. 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.

Recognizing Discrepancies

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.

How Experts Deal with Discrepancies

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.

Research on Dealing with Discrepancy

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) How Are Discrepancies Recognized?

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 both 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.

The GOBBLE Program

GOBBLE is a data base handling system which we have written in LISP. The principal features of GOBBLE are: 1) the use of contexts 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 set of ordered doubles or triples called "valid expressions" where the validity of an expression is determined through checks in a user-built, system maintained dictionary. 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.

Pattern-Matching and Fact Retrieval

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 after the facts have been stored. Facts can be retrieved either in a context or through 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.
```