6

Towards the Simulation of Clinical Cognition: Taking a Present Illness by Computer

Stephen G. Pauker, G. Anthony Gorry, Jerome P. Kassirer, and William B. Schwartz

Remarkably little is known about the cognitive processes employed in the solution of clinical problems. This paucity of information is probably accounted for in large part by the lack of suitable analytic tools for the study of the physician's thought processes. In the following early work, which arose from Gorry's observations outlined in Chapter 2, Pauker and his colleagues report on the use of the computer as a laboratory for the study of clinical cognition.

Their experimental approach consisted of several elements. First, cognitive insights gained from the study of clinicians' behavior were used to develop PIP, a computer program designed to take the present illness of a patient with edema. The program was then tested with a series of prototypical cases, and the present illnesses generated by the computer were compared to those taken by the clinicians in their group. Discrepant behavior on the part of the program was taken as a stimulus for further refinement of the evolving cognitive theory of the present illness. Corresponding refinements were made in the program, and the process of testing and revision was continued until the program's behavior closely resembled that of the clinicians.

The advances in computer science that made this kind of effort possible included goal-directed programming, pattern matching, and a large associative memory, all of which were products of research in the AI field.

From *American Journal of Medicine*, 60: 981–996 (1976). Used with permission.

The information used by the program is organized in a highly connected set of associations, which are then used to guide such activities as checking the validity of facts, generating and testing hypotheses, and constructing a coherent picture of the patient. As the program pursues its interrelated goals of information gathering and diagnosis, it uses knowledge of diseases and pathophysiology, as well as limited "common sense," to assemble dynamically many small problem-solving strategies into an integrated history-taking process.

Although the work was preliminary and aimed more at understanding cognitive processes and the related computer science issues than at short-term development of a clinical tool, PIP provided important new insights regarding the links among cognitive psychology, computer science, and the expertise of clinical problem solving. The article is also noteworthy because it represented the first time that the concepts of artificial intelligence appeared in a clinical medical journal. In addition, the research challenges that grew out of the PIP work have to a large extent defined the research directions of the AIM researchers at Tufts–New England Medical Center and M.I.T. in subsequent years.

6.1 Introduction

During the last decade there has been increasing interest in the use of the computer as an aid to both clinical diagnosis and management. Programs have been written that can carry out a review of systems (Slack et al., 1966), guide in the evaluation of acid-base disorders (Bleich, 1969; 1972), recommend the appropriate dose of digitalis (Peck et al., 1973; Jelliffe et al., 1972), and weigh the risks and benefits of alternative modes of treatment (Gorry et al., 1973). Some of these programs have been used to a limited extent in clinical practice, whereas others are prototypes that, although not yet of practical value, offer promise for the future. All, however, have the underlying characteristics that they are highly structured and that they deal with well-defined, sharply constrained problems. In nearly all instances, the use of a formalism, such as a flow chart (Slack et al., 1966; Bleich, 1969; 1972), decision analysis (Gorry et al., 1973), or a mathematical algorithm (Peck et al., 1973; Jelliffe et al., 1972), is the guiding principle used to capture clinical expertise in the computer.

There are, however, aspects of clinical medicine that cannot be reduced to formalisms, that is, situations in which a fixed recipe cannot provide the skilled guidance of the experienced clinician. To deal with this class of problems, new and more flexible strategies are under development, but work on such strategies is still in its embryonic phase (Shortliffe et al., 1973; Kulikowski et al., 1973; Pople and Werner, 1972).

In this paper, we report on the development of a computer program that uses unstructured problem-solving techniques to take the history of the present illness of a patient with edema.[1] We have chosen the problem of present illness for investigation because it is prototypical of clinical problems that demand complex problem-solving strategies. The present illness is, furthermore, the keystone on which a physician builds his or her diagnosis and bases many subsequent management decisions. Although we have examined only a limited range of issues in the present program, we believe that our effort is a first step toward a full understanding of the way in which a physician carries out the history-taking process.

6.2 Computer Science in the Study of Clinical Cognition

Our attempt to simulate the unstructured problem-solving processes of the present illness falls into the domain of computer science known as *artificial intelligence*. Research in this field is concerned with producing computer programs that exhibit behavior that would be termed *intelligent* if such behavior were that of a person. Examples of such work are programs that, to a limited extent, understand English, make sense of certain kinds of visual scenes, and control the operations of robots (Winston, 1974). Such research has been underway for 20 years (Feigenbaum, 1963) and, during this time, some major lessons have been learned. Perhaps the most important discovery has been that formalisms alone, for example, cybernetics (Bell, 1962), mathematical logic (McCarthy, 1968), and information theory (Shannon and Weaver, 1949), cannot produce intelligent behavior in complex, real-world situations. It has become abundantly clear that no single, formal approach can accommodate the knowledge of first principles and the experience, common sense,[2] and guesswork (Minsky, 1975) required for "intelligent" activities.

Because of the obvious competence of people in carrying out activities that formalisms cannot, artificial intelligence researchers have turned more recently to the study of human problem solving (Winston, 1974; Minsky, 1968). The study of natural intelligence, in fact, has become the central activity of artificial intelligence, and the experimental method of the field now emphasizes the use of computer systems as laboratories in which the-

[1]The program described here should be contrasted with the well-bounded "present illness algorithms" (Stead et al., 1972), which rely on flow charts for their implementation and which refer the patient to the physician for further questioning whenever the situation appears to be complex or serious.

[2]By common sense, we mean all the ordinary, rather pedestrian knowledge about everyday occurrences that is possessed by reasonably intelligent people.

ories of human problem solving can be represented and tested (Newell and Simon, 1972).

Conventional computer-programming concepts and structures have proven inadequate to express complex theories of human problem solving; however, new techniques have been developed that ameliorate these technological difficulties. Greatly improved systems have been created for managing very large collections of facts, and new goal-directed programming languages have been designed for utilizing these facts in the solution of difficult problems. Through the appropriate statement of goals, it is possible to construct a program that brings knowledge to bear *when it is required.* As new facts are obtained, such programs can dynamically organize many small problem-solving techniques into a coherent strategy that can respond flexibly to the changing picture of the world.[3] Equally important, as we shall discuss, is that these new languages provide means for giving a program *advice* as to when a particular piece of knowledge may be useful and how that knowledge should be applied to particular situations.

We believe that the ideas and technology now emerging from artificial intelligence research should make possible realistic simulations of human problem-solving strategies. In assessing the feasibility of building an "intelligent" program, however, some vital questions must be answered:

What is expert knowledge?
How much knowledge is required?
How should it be organized and how should it be applied?

The answers to these questions will come only from the careful study of real problem domains, and the success of such studies will be determined in large part by the *boundedness* of the problem domain under consideration. We believe that medicine, with its highly developed taxonomy, its codified knowledge base, the generally repetitive nature of the problem-solving encounters, and the existence of acknowledged experts, constitutes a promising problem domain because of its relatively well-bounded character. We therefore believe that, building on the technology at hand, acceptable progress can be made toward the development of sophisticated

[3]Expressed in technical terms, these languages do not require a detailed, rigid program because of pattern-directed invocation. Each subroutine contains a statement of what it potentially can accomplish, so the programmer need not specify which subroutines (or even that any subroutine) should carry out a desired action. Rather, he or she can specify the desired effect or goal and ask the computer to identify and use those subroutines that appear relevant. This type of program organization has many applications, such as offering heuristic advice and generating hypotheses. As an example of one class of problem that is very difficult to solve with conventional techniques, but that is trivial with this type of language, consider the problem of logical deduction. The program is told "All Greeks are poets" and "Anyone born in Athens is a Greek." We then tell the program that "Constantine was born in Athens" and ask "Is Constantine a poet?" The program automatically deduces the answer, basically using the same process we would. That is, it sets about to find out if Constantine is a poet. It realizes that the way to answer this question is to determine if he is a Greek, and therefore it asks if Constantine was born in Athens. When it discovers that he was, the original question is answered.

systems that can deal competently with complex clinical problems. To achieve such progress, however, the essential first step is to examine in depth the nature of the clinician's cognitive processes.

6.3 Methods of Procedure

Our first efforts were directed toward elucidating a number of the problem-solving strategies that physicians use in taking the history of the present illness of a patient with edema. This analytic effort was carried out through introspection and through direct observations of clinicians' problem-solving behavior. The insights gained in this way were represented as a computer program [using the CONNIVER programming system (Sussman and McDermott, 1972)] that incorporates the goal-directed techniques described in Section 6.2. The program was then tested with a series of prototypical cases in which edema was the presenting problem, and the questioning strategy followed by the program was compared to that of the physicians whom it was intended to simulate.

It immediately became apparent that the program's behavior differed markedly from that of the physicians, but, by examining specific discrepancies, we were able to recognize components of the clinicians' reasoning process that had been misunderstood or neglected in our initial analysis. With these new insights, we revised the program and evaluated its history-taking performance again. With each iteration of this process, the performance of the program improved and our insights into the cognitive process deepened. The study was terminated when the program closely simulated the manner in which the physician members of the team take the present illness of a patient with edema.[4]

6.4 Examples of Computer-Generated Analyses of Present Illnesses

Figure 6-1 presents a portion of a typical dialogue between a user (a physician) and the program. The language of both the questions and the computer-generated summaries (Figures 6-2 through 6-5) is rather stilted because most of our effort has been devoted to examining the history-taking process, not to producing a polished output. Each case demonstrates the program's use of somewhat different overall strategies of history taking.

[4]Our program simulates the behavior of the particular physicians in our group. The question of differences in problem-solving behavior among physicians is one we intend to pursue further.

```
PRESENTING PROBLEM: A YOUNG MAN WITH MASSIVE PEDAL EDEMA

1A IS PEDAL EDEMA, WHICH IS MASSIVE,
   1. FOR-DAYS?
   2. FOR-WEEKS?
   3. FOR-MONTHS?
   4. FOR-YEARS?                                        >>1
1B IS IT,
   1. FIRST-TIME?
   2. INFREQUENT?
   3. OCCASIONAL?
   4. FREQUENT?                                         >>1
1C IS IT,
   1. WITHOUT-TEMPORAL-PATTERN?
   2. WORSE-IN-MORNING?
   3. WORSE-IN-EVENING?                                 >>3
1D IS IT,
   1. SYMMETRICAL?
   2. ASYMMETRICAL?                                     >>1
1E IS IT,
   1. PITTING?
   2. NON-PITTING?                                      >>1
IF IS IT,
   1. ERYTHEMATOUS?
   2. NOT ERYTHEMATOUS?                                 >>2
1G IS IT,
   1. PAINFUL?
   2. NOT PAINFUL?                                      >>2
2 WELL, IS THERE DYSPNEA?                               >>NO
3 BY THE WAY, IS THERE HEAVY ALCOHOL CONSUMPTION?       >>NO
3A IS ALCOHOL CONSUMPTION,
   1. NONE?
   2. RARE?
   3. SOCIAL?                                           >>3
4 TELL ME, IS THERE JAUNDICE?                           >>NO
5 TELL ME, WAS THERE PROTEINURIA IN THE PAST?           >>UNKNOWN
5A IS THERE A LARGE-POLICY LIFE INSURANCE?              >>NO
5A1 IS LIFE INSURANCE,
   1. NONE?
   2. UNABLE-TO-OBTAIN?
   3. SMALL-POLICY?                                     >>3
5B IS THERE SERVED-IN ARMED FORCES?                     >>YES
6 WELL, IS THERE VARICOSE VEINS?                        >>NO
7 WAS THERE EDEMA IN THE PAST?                          >>NO
8 WELL, IS THERE ANY HEMATURIA?                         >>NO
8A IS THERE . . .                                       >>NO
```

FIGURE 6-1 Case 1. Initial portion of a representative present illness dialogue. The questions were asked by the computer, and the responses (preceded by >>) are those of the user. The complete present illness analysis, as generated by the computer, is shown in Figure 6-2.

PRESENTING PROBLEM: A YOUNG MAN WITH MASSIVE PEDAL EDEMA

THIS CASE CAN BE SUMMARIZED AS FOLLOWS:

THIS IS A YOUNG MAN WHO HAS PEDAL EDEMA WHICH IS NOT-PAINFUL, NOT-ERYTHEMATOUS, PITTING, SYMMETRICAL, WORSE-IN-EVENING, FIRST-TIME, FOR-DAYS AND MASSIVE. HE DOES NOT HAVE DYSPNEA. HE HAS SOCIAL ALCOHOL CONSUMPTION. HE DOES NOT HAVE JAUNDICE. IT IS NOT EXPLICITLY KNOWN WHETHER IN THE PAST HE HAD PROTEINURIA, BUT HE HAS SMALL-POLICY LIFE INSURANCE, AND HE HAS SERVED IN ARMED FORCES. HE DOES NOT HAVE VARICOSE VEINS. IN THE PAST HE DID NOT HAVE EDEMA. HE DOES NOT HAVE HEMATURIA. HE HAS NORMAL BUN. HE HAS NORMAL CREATININE. HE HAS PERI-ORBITAL EDEMA, WHICH IS WORSE-IN-MORNING, FIRST-TIME, FOR-DAYS AND SYMMETRICAL. HE HAS LOW ALBUMIN CONCENTRATION. HE HAS HEAVY PROTEINURIA, WHICH IS >5GRAMS/24HRS. HE HAS MODERATELY-ELEVATED, RISING WEIGHT. IN THE RECENT PAST HE DID NOT HAVE PHARYNGITIS. IN THE RECENT PAST HE HAD NOT-ATTENDED SCHOOL. IN THE RECENT PAST HE HAD NOT-ATTENDED SUMMER CAMP. IN THE RECENT PAST HE HAD NOT BEEN EXPOSED TO STREPTOCOCCI. IN THE RECENT PAST HE DID NOT HAVE FEVER. IT IS SAID, BUT HAS BEEN DISREGARDED, THAT HE HAS RED-CELL-CASTS-IN URINARY SEDIMENT. HE DOES NOT HAVE JOINT PAIN. HE DOES NOT HAVE RASH. HE HAS NEGATIVE ANA. HE DOES NOT HAVE FEVER. HE HAS NOT-RECEIVED ANTIBIOTIC. HE DOES NOT HAVE ANEMIA. IN THE PAST HE DID NOT HAVE HEMATURIA.

DIAGNOSES THAT HAVE BEEN ACCEPTED ARE: NEPHROTIC SYNDROME AND SODIUM RETENTION.

THE LEADING HYPOTHESIS IS IDIOPATHIC NEPHROTIC SYNDROME.

HYPOTHESES BEING CONSIDERED:

	A. fit of case to hypothesis	B. fraction of findings explained by hypothesis	average of A and B
IDIOPATHIC NEPHROTIC SYNDROME	0.80	0.37	0.58
ACUTE GLOMERULONEPHRITIS	0.22	0.27	0.24
HENOCH-SCHOENLEIN PURPURA	0.07	0.10	0.09

FIGURE 6-2 Case 1. Computer-generated summary of present illness of a patient with idiopathic nephrotic syndrome. Note that the diagnosis was not available to the computer; the program was provided only with the description of the presenting problem. The ranking at the bottom of the figure is based on the average of the "fit" of the case of the hypothesis (column A) and the fraction of the findings explained by the hypothesis (column B). For details of the evaluation (scoring) procedure, see text.

Case 1. Figure 6-2 shows the computer-generated summary of Case 1, a patient with idiopathic nephrotic syndrome. The computer was given as the chief complaint "a young man with massive pedal edema." The behavior of the program can be briefly summarized as follows. The computer characterized the edema in detail and, in light of the specific findings, turned to questions designed to elucidate etiology. After quickly determining that there was no history suggestive of congestive heart failure, alcoholic cirrhosis, varicosities, or renal failure, it noted that the patient had several findings strongly suggestive of nephrotic syndrome. The program then initiated a search for causes of the nephrotic syndrome, first exploring the possibility that the patient was suffering from poststrepto-

PRESENTING PROBLEM: A MIDDLE-AGED WOMAN WITH PEDAL EDEMA

THE CASE CAN BE SUMMARIZED AS FOLLOWS:

THIS IS A MIDDLE-AGED WOMAN, WHO HAS PEDAL EDEMA, WHICH IS NOT-PAINFUL, NOT-ERYTHEMATOUS, PITTING, SYMMETRICAL, 4+, WITHOUT-TEMPORAL-PATTERN, OCCASIONAL AND FOR-WEEKS. SHE DOES NOT HAVE DYSPNEA. SHE HAS HEAVY ALCOHOL CONSUMPTION. SHE HAS JAUNDICE. SHE HAS PAINFUL HEPATOMEGALY. SHE HAS SPLENOMEGALY. SHE HAS ASCITES. SHE HAS PALMAR ERYTHEMA. SHE HAS SPIDER ANGIOMATA. SHE DOES NOT HAVE PAROTID ENLARGEMENT. SHE HAS PROLONGED PROTHROMBIN TIME. SHE HAS MODERATELY-ELEVATED SGPT. SHE HAS MODERATELY-ELEVATED SGOT. SHE HAS MODERATELY-ELEVATED LDH. SHE HAS NOT RECEIVED BLOOD TRANSFUSIONS. SHE HAS NOT EATEN CLAMS. SHE DOES NOT HAVE ANOREXIA. SHE HAS MELENA. SHE DOES NOT HAVE HEMATEMESIS. SHE HAS LOW SERUM IRON. SHE HAS ESOPHAGEAL VARICES.

DIAGNOSES THAT HAVE BEEN ACCEPTED ARE: ALCOHOLISM AND GI BLEEDING.
THE LEADING HYPOTHESIS IS CIRRHOSIS.

HYPOTHESES BEING CONSIDERED:

	A. fit of case to hypothesis	B. fraction of findings explained by hypothesis	average of A and B
CIRRHOSIS	0.72	0.78	0.75
HEPATITIS	0.75	0.30	0.53
PORTAL HYPERTENSION	0.72	0.17	0.45
CONSTRICTIVE PERICARDITIS	0.17	0.13	0.15

FIGURE 6-3 Case 2. Computer-generated summary of the present illness of a patient with cirrhosis of the liver. The format is identical to that of Figure 6-2.

coccal glomerulonephritis and then looking for evidence of a systemic disease such as lupus erythematosus. Finding no evidence of a systemic disorder, the program made the diagnosis of nephrotic syndrome, probably idiopathic in character, but indicated that acute glomerulonephritis remained as a second, albeit much less likely, possibility. Note, incidentally, that the program disregarded the statement that red cell casts had been seen because it concluded that in the absence of hematuria the report of red cell casts was almost certainly in error. Also note that the questions about life insurance and military service were utilized because normal earlier physical examinations can suggest that proteinuria had not been present in the past.

Case 2. Figure 6-3 summarizes the present illness of a patient with Laennec's cirrhosis. The computer was given as the chief complaint "a middle-aged woman with pedal edema." In response, it obtained a detailed description of the character of the edema and then undertook an exploration of possible etiologies. On finding that the patient drank large quantities of alcohol, it turned to cirrhosis as a working hypothesis and quickly uncovered many stigmata of liver disease. The program also briefly ex-

PRESENTING PROBLEM: A YOUNG MAN WITH PEDAL EDEMA AND OLIGURIA

THE CASE CAN BE SUMMARIZED AS FOLLOWS:

THIS IS A YOUNG MAN, WHO HAS OLIGURIA. HE HAS PEDAL EDEMA, WHICH IS NOT-PAINFUL, NOT-ERYTHEMATOUS, PITTING, SYMMETRICAL, WITHOUT-TEMPORAL-PATTERN, FIRST-TIME AND FOR-DAYS. IT HAS BEEN DENIED THAT HE HAS RECENT SCARLET FEVER. IN THE RECENT PAST HE DID NOT HAVE PHARYNGITIS. IN THE RECENT PAST HE HAD NOT-ATTENDED SUMMER CAMP. IN THE RECENT PAST HE HAD NOT-BEEN-EXPOSED-TO STREPTOCOCCI. HE HAS NOT-RECEIVED RADIOGRAPHIC CONTRAST MATERIAL. HE HAS NOT-RECEIVED NEPHROTOXIC DRUGS. IN THE RECENT PAST HE DID NOT HAVE HYPOTENSION. HE HAS MODERATELY-ELEVATED URINE SODIUM. HE HAS URINE SPECIFIC GRAVITY WHICH IS ISOSTHENURIC. HE HAS NO-RED-CELLS-IN, NO-WHITE-CELLS-IN, RENAL-CELLS-IN, NO-RENAL-CELL-CASTS-IN, HYALINE-CASTS-IN URINARY SEDIMENT. IT IS NOT EXPLICITLY KNOWN WHETHER HE HAS BEEN-EXPOSED-TO A CLEANING FLUID. HE DOES NOT HAVE HYPOTENSION. HE DOES NOT HAVE TACHYCARDIA. HE HAS NORMAL TURGOR-AND-PERFUSION-OF SKIN. HE HAS MODERATELY-ELEVATED, RISING WEIGHT.

DIAGNOSES THAT HAVE BEEN ACCEPTED ARE: SODIUM RETENTION, EXPOSURE TO NEPHROTOXINS, EXPOSURE TO HEPATOTOXINS AND ACUTE RENAL FAILURE.

THE LEADING HYPOTHESIS IS ACUTE TUBULAR NECROSIS.

HYPOTHESES BEING CONSIDERED:

	A. fit of case to hypothesis	B. fraction of findings explained by hypothesis	average of A and B
ACUTE TUBULAR NECROSIS	0.50	0.37	0.43
ACUTE GLOMERULONEPHRITIS	0.20	0.21	0.20
IDIOPATHIC NEPHROTIC SYNDROME	0.18	0.16	0.17
CHRONIC GLOMERULONEPHRITIS	0.19	0.11	0.15

FIGURE 6-4 Case 3. Computer-generated summary of the present illness of a patient with acute tubular necrosis. The format is identical to that of Figure 6-2.

plored other etiologies of liver disease, such as the hepatitis induced by transfusions or by the ingestion of raw shellfish, but could find no evidence in support of these diagnoses. It then returned to the primary hypothesis of cirrhosis and, in searching for possible complications, noted the presence of both esophageal varices and chronic gastrointestinal bleeding. It concluded that the patient had alcoholic cirrhosis and that hepatitis was an alternative, but much less likely, possibility.

Case 3. Figure 6-4 shows the computer-generated summary of a patient with acute tubular necrosis produced by carbon tetrachloride exposure. The computer was given as the chief complaint "a young man with edema and oliguria." The program immediately undertook a search for causes of acute renal failure. It first focused on the diagnosis of acute glomerulonephritis but could find no evidence of streptococcal exposure. It next explored the possibility of acute tubular necrosis but was unable to find an etiological factor. When the program later assessed the characteristics of the urine sediment, however, it noted many hallmarks of tubular

PRESENTING PROBLEM: A MIDDLE-AGED MAN WITH ASCITES AND PEDAL EDEMA

THE CASE CAN BE SUMMARIZED AS FOLLOWS:

THIS IS A MIDDLE-AGED MAN WHO HAS ASCITES. HE HAS PEDAL EDEMA, WHICH IS NOT-PAINFUL, NOT-ERYTHEMATOUS, PITTING, SYMMETRICAL, WORSE-IN-EVENING, OCCASIONAL AND FOR-MONTHS. HE HAS SOCIAL ALCOHOL CONSUMPTION. HE HAS HEPATOMEGALY. HE DOES NOT HAVE JAUNDICE. HE DOES NOT HAVE PALMAR ERYTHEMA. HE DOES NOT HAVE SPIDER ANGIOMATA. HE DOES NOT HAVE PAROTID ENLARGEMENT. HE DOES NOT HAVE GYNECOMASTIA. HE DOES NOT HAVE TESTICULAR ATROPHY. HE HAS NORMAL BILIRUBIN. HE HAS NORMAL PROTHROMBIN TIME. HE HAS NORMAL SGPT. HE HAS NORMAL SGOT. HE HAS CHEST PAIN WHICH IS RELIEVED-BY-SITTING-UP, WITHOUT-RADIATION, MODERATE, OCCASIONAL, FOR-SECONDS AND SHARP. HE HAS EXERTIONAL DYSPNEA. HE HAS ORTHOPNEA. HE DOES NOT HAVE PAROXYSMAL NOCTURNAL DYSPNEA. HE HAS ELEVATED NECK VEINS. HE HAS KUSSMAUL'S SIGN. HE HAS PERICARDIAL KNOCK. HE HAS DISTANT HEART SOUNDS. HE HAS PERICARDIAL-CALCIFICATION-ON, NORMAL-HEART-SIZE-ON, CLEAR-LUNG-FIELDS-ON CHEST XRAY.

THE LEADING HYPOTHESIS IS CONSTRICTIVE PERICARDITIS.

HYPOTHESES BEING CONSIDERED:

	A. fit of case to hypothesis	B. fraction of findings explained by hypothesis	average of A and B
CONSTRICTIVE PERICARDITIS	0.78	0.50	0.64
CONGESTIVE HEART FAILURE	0.44	0.21	0.32

FIGURE 6-5 Case 4. Computer-generated summary of the present illness of a patient with constrictive pericarditis. The format is identical to that of Figure 6-2.

injury. Pursuing this lead, it soon uncovered an exposure to a cleaning fluid that it presumed contained carbon tetrachloride. It then explored the possibility that acute hypotension had also contributed to the development of the oliguria but could obtain no evidence in support of this hypothesis. Finally, it determined that body weight was increasing, and from this fact concluded that the patient was retaining sodium. Because the data base does not currently include the distinction between the retention of salt and the retention of free water, the program could not arrive at the correct interpretation of the weight gain, namely, that the overhydration was due to water retention *per se*.

Case 4. Figure 6-5 gives the summary of the present illness of a man with constrictive pericarditis secondary to tuberculosis. The program was given as the chief complaint "a middle-aged man with ascites and pedal edema." After further characterizing the edema, the computer focused on a hepatic etiology and found that the patient had an enlarged liver. Although subsequent questioning revealed only social alcohol consumption, the program persisted in its search for stigmata of cirrhosis. When none was found, it turned to a possible cardiac etiology and noted that chest pain was a prominent complaint; the pain was, however, more characteristic of pleural or pericardial than of myocardial disease. It next found that there was both neck vein distention and orthopnea, but that there was no paroxysmal nocturnal dyspnea. These clinical findings, in combination with the ascites, suggested the diagnosis of pericardial disease. Further questioning then revealed many of the stigmata of constrictive pericarditis.

Because even the experienced clinician often confuses constrictive pericarditis with cirrhosis, it is understandable why the diagnosis of cirrhosis was pursued with such vigor. Note, however, that the program was deficient in that it failed to explore other etiologies of predominantly right-sided cardiac failure, such as cor pulmonale and multiple pulmonary emboli; this shortcoming is explained by the fact that the current knowledge base does not include information about these latter diagnoses.

6.5 Nature of the Underlying Computer Programs

In this section we shall first discuss the overall behavior of the program in terms of its major components and the way that these components interact. We shall then consider in detail the underlying processes used by the program.

6.5.1 An Overview of the Present Illness Program

In taking a history of the present illness, the program, much like the physician, tries to develop a sufficient "understanding" of the patient's complaints to form a reasonable basis on which to evaluate the clinical problem and to lay the groundwork for subsequent management decisions. It accomplishes this goal by undertaking two processes: information gathering and diagnosis. Although these two threads of the problem-solving process are interwoven, for clarity of exposition we shall consider them separately.

By *information gathering,* we mean the accumulation of a profile of data concerning the patient. Because there are innumerable facts that could be gathered, one needs a sharp focus for this activity. This focus is obtained through the pursuit of a small set of diagnostic hypotheses that are suggested by the presenting complaints.

The process of *diagnosis,* in contrast, is an attempt to infer the meaning of a constellation of given findings and does not involve the acquisition of additional information about the patient; rather, it is concerned with the processing of the available facts. When additional findings are required, the diagnostic process turns again to the information-gathering process. Thus the history-taking process is directed both at establishing *what the facts are* and at establishing *what the facts mean* (Feinstein, 1967).

In taking a present illness, our program uses the chief complaint to generate hypotheses about the patient's condition. It also actively seeks additional clinical information to accomplish a number of different tasks, including testing hypotheses and eliminating unlikely ones. Any of these activities may spawn further tasks, such as checking the validity of a newly discovered fact or asking about related findings. As will become evident, however, this brief description understates both the complexity of the program's behavior and the differences between this program and others previously reported.

6.5.2 The Basic Components of the Program

The complexity of the program's behavior is the result of the interaction of the four factors schematically shown in Figure 6-6: (1) the patient-specific data, (2) the supervisory program, (3) the short-term memory, and (4) the long-term (associative) memory.

1. *The patient-specific data.* These are the facts provided by the user either spontaneously or in response to questions asked by the program. These data comprise the computer's knowledge about the patient.

2. *The supervisory program.* The supervisory program guides the computer in taking the present illness and oversees the operation of various subprocesses, such as selecting questions, seeking and applying relevant advice, and processing algorithms (such as flow charts). The principal goal of this supervisor is to arrive at a coherent formulation of the case, by quickly generating and testing hypotheses and by excluding competing hypotheses. At the present time, there are about 300 potential questions that relate to over 150 different concepts that the program can employ in its information-gathering activities.

3. *The short-term memory.* The short-term memory is the site in which data about the patient interact with general medical knowledge that is kept in long-term memory (see below). The supervisory program determines which aspects of this general knowledge enter the short-term memory and how such knowledge is melded with the patient-specific data that are under consideration. The amount of information in short-term memory is quite variable, depending on the complexity of the case and the number of active hypotheses. For a simple case, the short-term memory might contain only two or three hypotheses and the knowledge and deductions associated with them. In a complex or puzzling case, it might contain five or ten hypotheses.

4. *The long-term (associative) memory.* The long-term memory contains a rich collection of knowledge, organized into packages of closely related

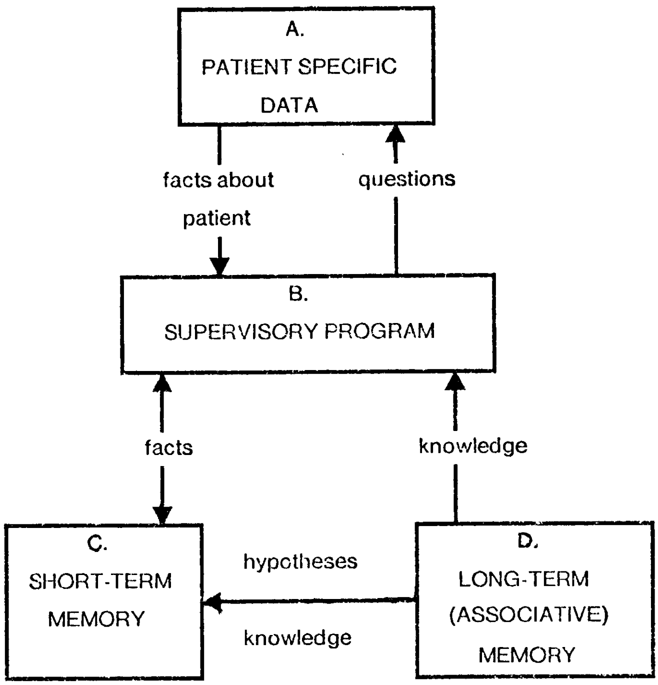


FIGURE 6-6 Overview of program organization. Clinical data (A) are presented to the supervisory program (B), which places them in short-term memory (C). The supervisory program, after consulting both short-term (C) and long-term memories (D), generates hypotheses and moves the information associated with these hypotheses from long-term to short-term memory. The supervisory program then asks for additional patient-specific data relevant to its hypotheses. At every stage, each hypothesis is evaluated (scored) by the program to determine whether it should be rejected, accepted, or considered further.

facts called *frames* (Minsky, 1975). Frames are centered around diseases (such as acute glomerulonephritis), clinical states (such as nephrotic syndrome), or physiologic states (such as sodium retention). Within each frame is a rich knowledge structure that includes prototypical findings (signs, symptoms, laboratory data), the time course of a given illness, and rules for judging how closely a given patient might match the disease or state that the frame describes. A typical example of a frame (nephrotic syndrome) is shown in Figure 6-7.

As shown in Figure 6-8, the frames are linked into a complex *network*. In the figure each frame is represented as a shaded sphere (diseases are

```
NAME: NEPHROTIC SYNDROME
IS-A-TYPE-OF: CLINICAL STATE
FINDING: LOW SERUM ALBUMIN CONCENTRATION
FINDING: HEAVY PROTEINURIA
FINDING: >5GRAMS/24HRS PROTEINURIA
FINDING: MASSIVE SYMMETRICAL EDEMA
FINDING: EITHER FACIAL OR PERI-ORBITAL AND SYMMETRICAL EDEMA
FINDING: HIGH SERUM CHOLESTEROL CONCENTRATION
FINDING: URINE LIPIDS PRESENT
MUST-NOT-HAVE: PROTEINURIA ABSENT
IS-SUFFICIENT: BOTH MASSIVE PEDAL EDEMA AND >5GRAMS/24HRS PROTEINURIA
MAJOR SCORING:
    SERUM ALBUMIN CONCENTRATIONS
        LO: 1.0
        HIGH: -1.0
    PROTEINURIA:
        >5GRAMS/24HRS: 1.0
        HEAVY: 0.5
        EITHER ABSENT OR LIGHT: -1.0
    EDEMA:
        MASSIVE AND SYMMETRICAL: 1.0
        NOT MASSIVE BUT SYMMETRICAL: 0.5
        ERYTHEMATOUS: -0.2
        ASYMMETRICAL: -0.5
        ABSENT: -1.0
MINOR SCORING:
    SERUM CHOLESTEROL CONCENTRATION:
        HIGH: 1.0
        NOT HIGH: -1.0
    URINE LIPIDS:
        PRESENT: 1.0
        ABSENT: -0.5
MAY-BE-CAUSED-BY:
    ACUTE GLOMERULONEPHRITIS,
    CHRONIC GLOMERULONEPHRITIS,
    NEPHROTOXIC DRUGS,
    INSECT BITE,
    IDIOPATHIC NEPHROTIC SYNDROME,
    SYSTEMATIC LUPUS ERYTHEMATOUS, OR
    DIABETES MELLITUS
MAY-BE-COMPLICATED-BY:
    HYPOVOLEMIA
    CELLULITIS
MAY-BE-CAUSE-OF: SODIUM RETENTION
DIFFERENTIAL DIAGNOSIS:
    IF NECK VEINS ELEVATED, CONSIDER: CONSTRICTIVE PERICARDITIS
    IF ASCITES PRESENT, CONSIDER: CIRRHOSIS
    IF PULMONARY EMBOLI PRESENT, CONSIDER: RENAL VEIN THROMBOSIS
```

FIGURE 6-7 A typical frame. Information about a disease, a physiologic state, etc., is stored in the form of a frame within the long-term memory. Included in a typical frame, as shown here for nephrotic syndrome, are descriptions of typical findings, numerical factors to be used in scoring, and links to other frames (e.g., MAY-BE-CAUSED-BY, MAY-BE-COMPLICATED-BY). There are also rules for excluding (MUST-NOT-HAVE) and satisfying (IS-SUFFICIENT) the fit of the frame to the case at hand. For further details, see text.

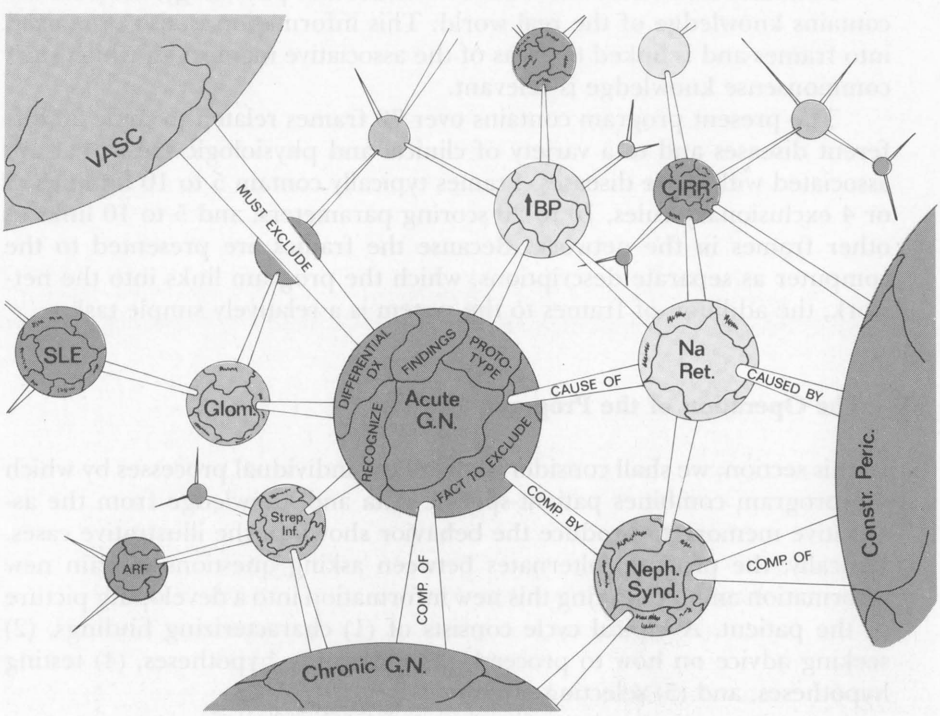


FIGURE 6-8 The long-term (associative) memory. The long-term memory consists of a rich collection of knowledge about diseases, signs, symptoms, pathologic states, real-world situations, etc. Each point of entry into the memory allows access to many related concepts through a variety of associative links shown as rods. Each rod is labeled to indicate the kind of association it represents. Note that the dark gray spheres denote disease states, medium gray spheres denote clinical states (e.g., nephrotic syndrome) and light gray spheres denote physiologic states (e.g., sodium retention). Abbreviations used in this figure are Acute G.N. = acute glomerulonephritis, Chronic G.N. = chronic glomerulonephritis, VASC = vasculitis, CIRR = cirrhosis, Constr. Peric. = constrictive pericarditis, ARF = acute rheumatic fever, Na Ret. = sodium retention, SLE = systematic lupus erythematosus, ↑BP = acute hypertension, Glom. = glomerulitis, Strep. Inf. = streptococcal infection, Neph. Synd. = nephrotic syndrome.

dark gray, clinical states are medium gray, and physiologic states are light gray), and the links between the frames are represented as labeled rods. These links depict a variety of relations, such as MAY-BE-CAUSED-BY and MAY-BE-COMPLICATED-BY.

In addition to information about diseases and physiology, the network contains knowledge of the real world. This information is also organized into frames and is linked to areas of the associative memory in which such commonsense knowledge is relevant.

The present program contains over 70 frames related to some 20 different diseases and to a variety of clinical and physiologic states that are associated with these diseases. Frames typically contain 5 to 10 findings, 3 or 4 exclusionary rules, 10 to 20 scoring parameters, and 5 to 10 links to other frames in the network. Because the frames are presented to the computer as separate descriptions, which the program links into the network, the addition of frames to the system is a relatively simple task.

6.5.3 The Operation of the Program

In this section, we shall consider in detail the individual processes by which the program combines patient-specific data and knowledge from the associative memory to produce the behavior shown in the illustrative cases. Basically, the program alternates between asking questions to gain new information and integrating this new information into a developing picture of the patient. A typical cycle consists of (1) characterizing findings, (2) seeking advice on how to proceed, (3) generating hypotheses, (4) testing hypotheses, and (5) selecting questions.

Characterizing Findings

After being presented with the chief complaint, the supervisor retrieves from the associative memory a procedure that characterizes that complaint in detail. This procedure is a flow chart that follows a set pattern in eliciting such features as the location, severity, and duration of the complaint. The program uses this detailed description of the complaint to limit the number of hypotheses that it will later have to consider.

Seeking Advice on How to Proceed

One of the most important features of our program is its ability to assemble small history-taking strategies into an overall approach that is tailored to the case at hand. This ability is critically dependent on the availability of appropriate advice about efficient methods for the exploration and organization of the case. Here we shall present three examples of the program's use of this facility:

1. Advice can be given that alerts the supervisor to ask one or more questions that will "zero in" on the presenting problem and thus, at the stage

of hypothesis generation (see below), limit the number of diagnostic possibilities that must be evaluated.

2. Advice can be given that guides the supervisor in its evaluation of information that is being presented. Such validity checks can be of several types. First, the program might point out that a finding itself is clearly in error, e.g., a weight gain of 50 pounds in 48 hours. Second, it might note that new information is inconsistent with other facts known about the patient, e.g., the presence of red cell casts in the absence of hematuria. Finally, it might indicate that a new finding contradicts a conclusion already drawn about the case.[5]
3. Advice can be given that alerts the supervisor to errors that might stem from a patient's misinterpretation of a particular sign or symptom. For example, if a patient complains of "blood in the urine," the supervisor is told that dark urine, which is attributed by the patient to blood, may be caused by the presence of bile, myoglobin, or anthocyanins (from beets).

Hypothesis Generation

After the complaint has been characterized and all relevant advice has been acted upon, the supervisory program proceeds to generate working hypotheses. Hypothesis generation consists of moving frames from long-term memory to short-term memory, where each frame plays a special role in guiding further exploration of the patient's problem. Frames can exist in one of four states: dormant, semiactive, active, and accepted. Initially, the short-term memory contains no frames; all frames are in the long-term memory and are said to be in the *dormant state.* In this nascent condition, however, some of the findings in the frames are associated with small, independent computer programs called *daemons.* A few of these daemons extend like tentacles from the frame into the short-term memory (see Figure 6-9, BEFORE); these are primarily the daemons of those findings that are strongly suggestive of their associated frames. When the matching fact for a daemon is added to the short-term memory, the entire frame attached to the daemon is added to the short-term memory (see Figure 6-9, AFTER). As pointed out, this process is synonymous with forming a hypothesis. Those frames that have entered short-term memory as hypotheses are called *active.* As is reflected in the AFTER half of Figure 6-9, frames one link away from an active frame are also affected in that during the activation process they are pulled closer to short-term memory. Consequently, more tentacles from such frames can reach into memory where they can now watch for their matching facts. These related frames,

[5]The latter two kinds of advice would not be provided in the initial cycle, which deals with the chief complaint, because, at such an early stage, the short-term memory would not contain any detailed information about the patient.

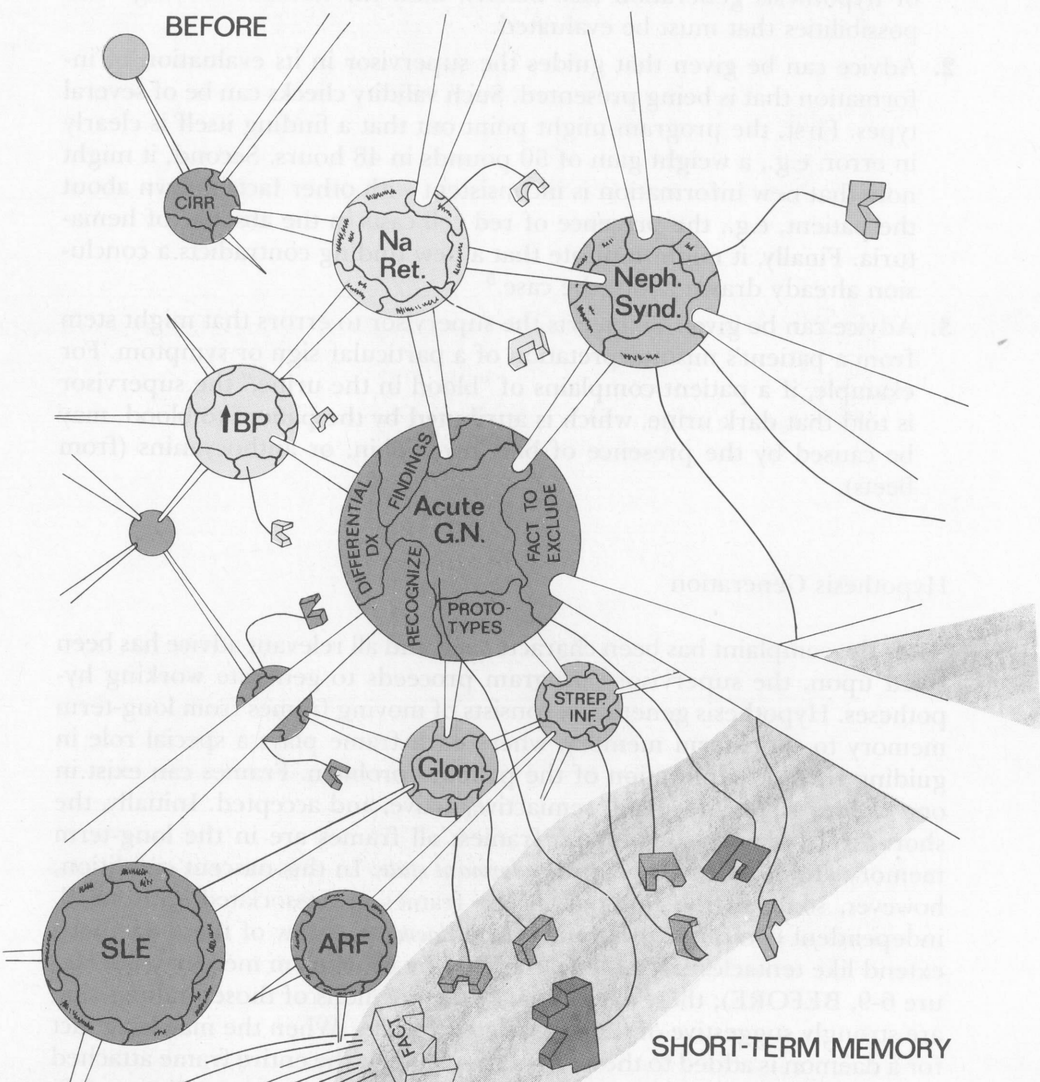
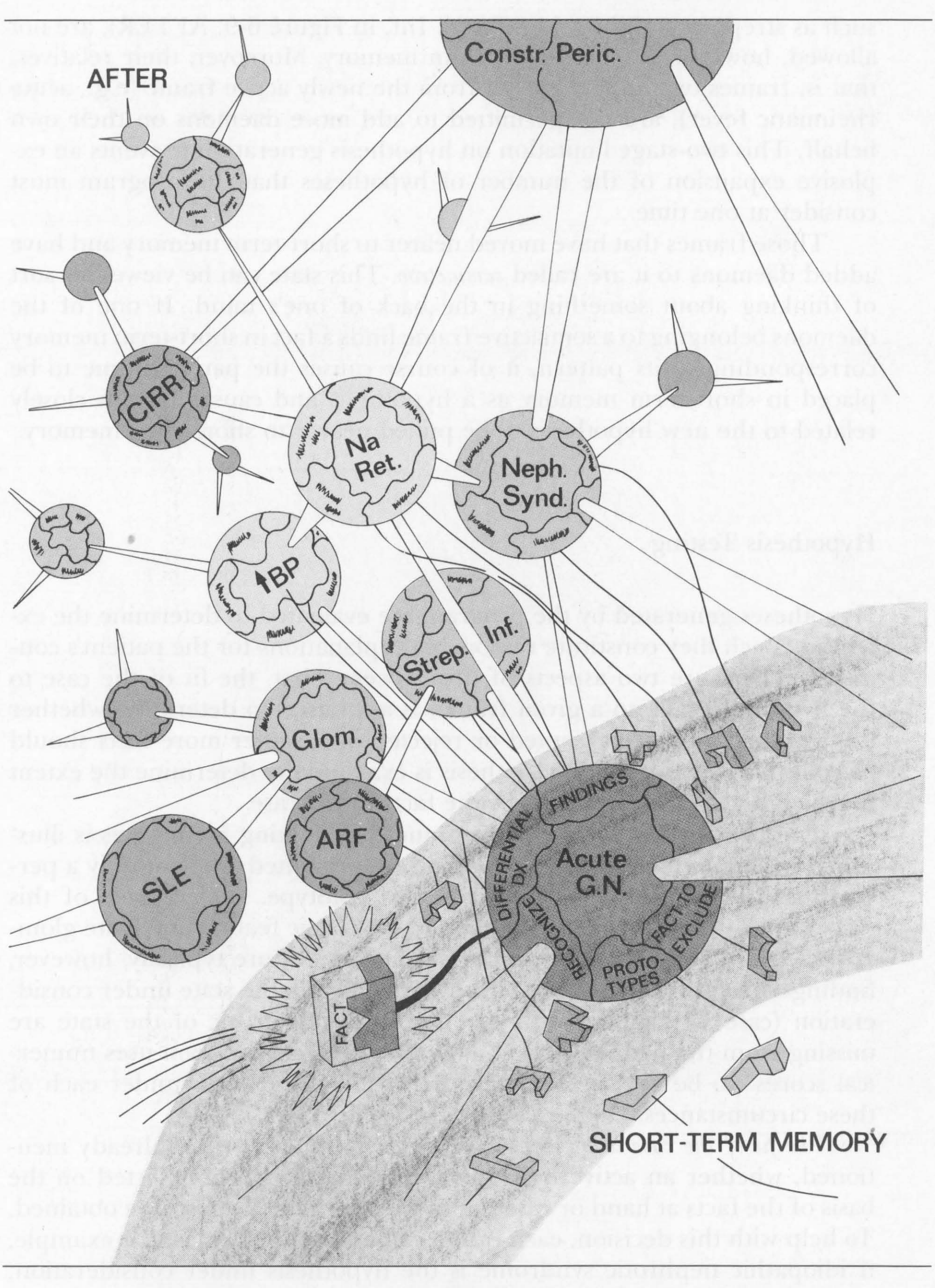


FIGURE 6-9 Hypothesis generation. (The abbreviations are the same as those used in Figure 6-8.) BEFORE: in the nascent condition (when there are no hypotheses in short-term memory), tentacles (daemons) from some frames in long-term memory extend into the short-term memory, where each constantly searches for a matching fact.



AFTER: the matching of fact and daemon causes the movement of the full frame (in this case, acute glomerulonephritis) into short-term memory. As a secondary effect, frames immediately adjacent to the activated frame move closer to short-term memory and are able to place additional daemons therein. Note that, to avoid complexity, the daemons on many of the frames are not shown.

such as streptococcal infection (Strep. Inf. in Figure 6-9, AFTER), are not allowed, however, to enter short-term memory. Moreover, their relatives, that is, frames two links removed from the newly active frame (e.g., acute rheumatic fever), are not permitted to add more daemons on their own behalf. This two-stage limitation on hypothesis generation prevents an explosive expansion of the number of hypotheses that the program must consider at one time.

Those frames that have moved nearer to short-term memory and have added daemons to it are called *semiactive.* This state can be viewed as sort of thinking about something in the back of one's mind. If one of the daemons belonging to a semiactive frame finds a fact in short-term memory corresponding to its pattern, it of course causes the parent frame to be placed in short-term memory as a hypothesis and causes frames closely related to the new hypothesis to be pulled nearer to short-term memory.

Hypothesis Testing

Hypotheses generated by the program are evaluated to determine the extent to which they constitute reasonable explanations for the patient's condition. There are two aspects of this process. First, the fit of the case to the hypothesis (i.e., to a given frame) is appraised to determine whether the hypothesis can be accepted or rejected or whether more facts should be collected. Second, each hypothesis is examined to determine the extent to which it can account for all of the facts in the case.

The problem faced by the program in evaluating hypotheses is illustrated in Figure 6-10. In case A, we have represented schematically a perfect match between patient and disease prototype. An example of this situation would be a patient who has all the classic features of acute glomerulonephritis and no other abnormal findings. More typically, however, findings are present that are not ordinarily seen in the state under consideration (case B, Figure 6-10), or findings characteristic of the state are missing from the patient (case C, Figure 6-10). The program uses numerical scores (to be discussed) to measure the degree of fit under each of these circumstances.

The fit of the case to the hypothesis serves to determine, as already mentioned, whether an active hypothesis can be accepted or rejected on the basis of the facts at hand or whether more information should be obtained. To help with this decision, each frame contains specific rules. For example, if idiopathic nephrotic syndrome is the hypothesis under consideration, and the program then learns that the patient has had gross hematuria, an *exclusionary rule* rejects the hypothesis and permanently removes the nephrotic syndrome from short-term memory. On the other hand, if the patient has both edema and massive proteinuria (protein excretion of

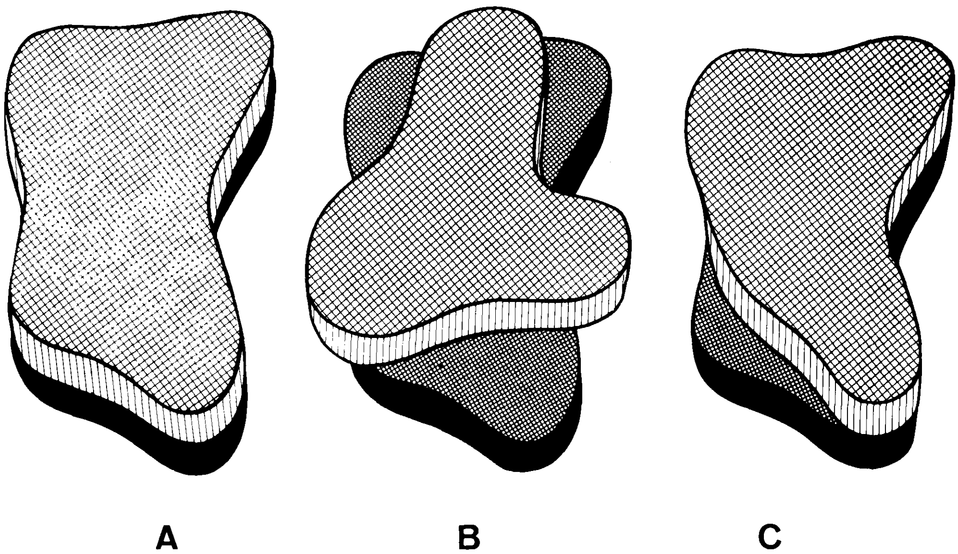


FIGURE 6-10 Schematic representation of pattern matching. Two wafers are shown in each instance, the lower one denoting the prototype being sought and the upper one denoting the case being tested. Case A: an exact match. Every important feature of the prototype is found in the case, and there are no features of the case that are not explained by the prototype. Case B: there are features of the case that are not explained by the prototype. Case C: there are features in the prototype that are not found in the case.

greater than 5 g/24 hours), a *sufficiency rule* immediately accepts the nephrotic syndrome hypothesis.[6]

In this accepted state, the hypothesis is asserted as if it were a fact. This new "fact" then is added to the short-term memory where it, in turn, can be found by daemons belonging to other frames. We should emphasize, however, that if later facts contradict the original conclusion, the acceptance is revoked.

In many instances, of course, there is no simple rule that can serve either to exclude or to establish a given hypothesis, and a *scoring process* is required. This scoring process uses numerical values (contained in the frame) that reflect the likelihood that various clinical findings will occur in

[6]Not only can disease frames be accepted or rejected, but frames corresponding to physiologic and clinical states can be similarly accepted or rejected.

the given disorder.[7] Major features are given more weight in the final scoring process than are the minor features.

Consider, for example, the nephrotic syndrome frame shown in Figure 6-7. Those features of the frame that can be readily identified as present or absent (e.g., low serum albumin concentration or heavy proteinuria) are given a numerical value. The remaining features are not initially assigned values; instead their contribution to the scoring process is determined by affiliated frames, e.g., hypovolemia is evaluated by means of a specific hypovolemia frame. Once such an affiliated frame has carried out its scoring function, the resulting value is passed on to the central frame. For example, acute glomerulonephritis receives a score for "elevated blood pressure in the absence of signs of chronic hypertension" from the acute hypertension frame to which it is related by a "may be complicated by" link. Similarly, the acute hypertension frame itself depends on affiliated frames (e.g., hypertensive encephalopathy) for its own score. In some instances, such propagation of scoring proceeds through several levels.

If the score for a hypothesis exceeds a defined threshold, the frame is accepted by the supervisor. Similarly, if the score falls below a given threshold (i.e., if the hypothesis no longer fits the patient "well enough"), the supervisor forces the hypothesis into a semiactive state. *The ability of the hypothesis to account for the findings of the case* is the extent to which all the facts of the case are explained by the hypothesis and its affiliated frames. The hypothesis of acute glomerulonephritis can explain, for example, both "low serum complement" and "oliguria"; the former finding is a part of the acute glomerulonephritis frame itself, and the latter is a part of a closely linked frame, "acute renal failure." It cannot, however, account for the finding of long-standing hypertension. The program computes for each frame a value equal to the fraction of all findings in the patient profile that are explained by the hypothesis. This value and the measure of the fit of the case to the hypothesis are averaged, and the hypotheses are assigned a rank order based on the average.

Selecting Questions

After the supervisor has ranked the hypotheses, it seeks to gather more information about the patient in order to improve its understanding of the clinical problem. The hypothesis that has received the highest overall score is explored first, with the initial inquiries directed to the classic findings of the disorder. The answer to each question that is posed causes the reevaluation of all hypotheses; as new information is obtained, the supervisor determines whether the leading hypothesis being pursued is still plausible, should now be accepted, or should be discarded from active consideration.

---

[7]The weights associated with those features of the frame known to be present in the patient are summed and then normalized by the maximum attainable score.

After the program has gathered information on prototypical features of the frame, it turns to questions about minor features of the disorder and then to inquiries about complications, etiological factors, and differential diagnoses. A change in the train of questioning usually indicates that, as the result of the continuous process of reevaluation, a new hypothesis has moved into the leading position.[8]

Repeating the Cycle

Any new finding obtained in the course of the questioning process sets into motion a cycle that is the same as that just described for the chief complaint: each new sign or symptom is characterized, advice is sought, hypotheses are generated and tested, and additional questions are asked. In its cycle of response to new findings, the program will, however, make use of the information acquired earlier (and hypotheses already generated) and will thus focus its questioning more sharply than if such a context were not available.

Controlling the Proliferation of Hypotheses

As discussed, the information gathering and the diagnostic competence of the present illness program depend critically on its ability to quickly generate hypotheses to account for the patient's condition. For this reason, "aggressive" hypothesis generation occurs even when only a few rather isolated facts are available. To avoid the excessive computational burden that is often produced by such an aggressive strategy, the program employs several methods to restrict the number of hypotheses under active consideration.

Two of these methods have already been mentioned—the "zeroing-in" on a complaint and the two-stage process of hypothesis generation. A third method is the application of the *principle of parsimony.* Let us take, as an example, a patient with edema and massive proteinuria who is hypothesized to have nephrotic syndrome with sodium retention. If the program discovers that the patient has a positive test for antinuclear antibody, it does not simply add a new hypothesis, "systematic lupus erythematosus." Instead, it incorporates the hypothesis of "nephrotic syndrome with sodium retention" into a new, overall hypothesis of "lupus erythematosus

[8]Sometimes the program considers a new hypothesis because of advice stored in the frame currently under consideration. For example, if nephrotic syndrome is the current hypothesis and symptoms suggestive of pulmonary emboli are reported, advice in the nephrotic syndrome frame will suggest that attention be shifted to renal vein thrombosis. The supervisor will then call up the questions designed to explore this latter possibility.

with nephrotic syndrome." If at a later time the more parsimonious hypothesis is rejected, the subhypotheses are again given independent status as active frames or, alternatively, are returned to a dormant state.

6.6 Comments

The present report demonstrates that insight derived from the study of clinical cognition can be combined with advanced techniques for computer simulation to create computer programs that possess a powerful problem-solving capability. The major technological advance embodied in our program is the capacity to retrieve and apply knowledge *when that knowledge is required,* thus freeing the programmer from the virtually impossible task of specifying all contingencies in advance [as would be necessary, for example, in a branching flow chart (Slack et al., 1966; Bleich, 1969; 1972; Stead et al., 1972)]. The key to implementation of the present system lies in the goal-directed nature of its operation. It is this goal-directed character that permits the supervisory program to select pertinent medical and real-world knowledge from the computer's memory and to dynamically assemble many small problem-solving techniques that efficiently guide the acquisition of additional clinical information. Another central feature of the program is the organization of its data base into an associative memory in which clusters of closely related facts about diseases and clinical states are stored in a fashion analogous to a richly cross-referenced encyclopedia. These groups of facts, called *frames* (Figure 6-7), are further organized into a *network* (Figure 6-8) that facilitates efficient retrieval of closely related blocks of information. When the supervisory program is presented with a clinical problem (that is, a chief complaint), it generates hypotheses about the case by moving frames from the long-term (associative) memory into short-term memory, where the frames interact with a profile of the patient's clinical data. When a hypothesis is generated, the supervisory program becomes "aware" of all the etiologies, complications, and other features of the hypothesized condition because the frames describing such related facts are drawn into close proximity to short-term memory.

The goal of the program is to arrive at the best possible diagnostic appraisal by evaluating these hypotheses. To accomplish this purpose, the computer characterizes each finding in detail, seeks relevant advice from the associative memory, and tests the hypotheses. The set of hypotheses under consideration also provides a framework within which additional information, both medical and real-world, is sought and interpreted. Throughout the questioning process, the supervisory program searches for inconsistencies in the information that it has obtained. When such inconsistencies arise, the program consults the memory for specific advice on how to deal with the conflicting information.

As new facts are obtained, additional hypotheses may be generated. From time to time, the supervisory program may also combine several hypotheses to form a more coherent diagnostic picture. As the questioning proceeds, all hypotheses under consideration are repetitively tested and scored to measure the "goodness of fit" between the description of the disease or physiologic state and the profile of facts about the patient. This testing provides the basis for either the acceptance or rejection of each hypothesis. At the termination of the questioning, the accepted hypotheses are listed, and the other hypotheses are rank ordered on the basis of the final score calculated for each.

The fundamental principles embodied in the present illness program are, we believe, applicable not solely to the problem of edema but are broadly relevant to the history-taking process. It is obvious, however, that many strategies other than those we have employed must be uncovered if a system such as we have described is to deal effectively with a wide range of clinical problems. In addition, numerous medical and real-world facts must be added to the program, and ingenious new techniques must be devised that can deal with multiple coexisting diseases, that can draw appropriate inferences about the temporal aspects of a patient's history, and that can choose the point at which the questioning process should be terminated. The solutions to such problems obviously will require many years of intensive work.

6.6.1 The Problem of Scale

We believe that, over the long term, computer programs can be developed that should be capable not only of taking the present illness but also of assisting in virtually all aspects of patient management. If this view is correct, one must then ask whether the existing technology will be able to cope with the volume of information that might be required by such a system; that is, will it be possible to store the requisite number of facts at a reasonable cost and to retrieve them in an efficient and effective manner? To answer this question we must first ask how much a computer program must "know" before it knows all of general internal medicine. Obviously, any calculation of this sort must be highly speculative, but it seems certain that the program must have available at least that body of information that is contained in a standard textbook of medicine. As shown in section A of Table 6-1, it appears that each of the two most widely used textbooks of medicine contains on the order of 200,000 facts.[9] This estimate far un-

---

[9]This estimate was arrived at by the crude technique of estimating the number of facts on several pages (not only basic facts, but the relationships between them) and multiplying this average by the total number of pages in the book. The major source of variability in such an estimate is the definition of what constitutes a single fact, because such a definition is to a certain extent arbitrary. In our calculations, we have used as a yardstick the amount of information that is treated as a single fact by our program; however, the choice of any other reasonable yardstick would not have changed our results appreciably.

TABLE 6-1 Estimate of total number of facts contained in standard textbooks of medicine and in representative subspecialty texts*

Title	*Pages (approx.)*	*Facts per page (approx.)*	*Total facts*
A. GENERAL INTERNAL MEDICINE			
Principles of Internal Medicine (Wintrobe, 1974b)	2,035	100	200,000
Textbook of Internal Medicine (Beeson and McDermott, 1975)	1,892	100	190,000
B. SPECIALTY TEXTS			
Diseases of the Kidney (Strauss, 1971)	1,456	40	60,000
The Heart (Hurst, 1974)	1,755	50	90,000
Clinical Hematology (Wintrobe, 1974a)	1,788	40	70,000

*Estimated as described in Footnote 9.

derstates, however, the total amount of information that is relevant to the practice of internal medicine. It is clear, for example, that there is a fund of basic science information used by the clinician that does not appear in such a textbook of medicine. To account for this body of data, we will double our estimate to a total of 400,000 facts. Finally, there is a considerable body of information about the real world (life insurance examinations, army physicals, time of day, seasons of the year), which, we will estimate, requires knowledge of still another 100,000 facts.[10] This brings us to a total of 500,000 facts. If we now double this value to take cognizance of possible underestimates, we arrive at an upper bound of approximately 1 million facts as the core body of information in general internal medicine.

The core knowledge embodied in the approximately ten separate subspecialties of internal medicine is, of course, considerably larger. To estimate the volume of clinical information basic to the entire domain of subspecialty medicine, we first have estimated the number of facts in textbooks of nephrology, cardiology, and hematology. As shown in section B of Table 6-1, each of the subspecialty treatises contains on the order of 60,000 facts. From this we estimate that the core body of information in all medical subspecialty texts combined is about 600,000 facts. If we assume that approximately one-third of this information represents duplications among the specialty fields, we arrive at a total body of 400,000 facts, a value approximately twice that estimated for general internal medicine. Using the same ratio between facts and other kinds of relevant information as we used in the case of general medicine, we calculate, correcting again for any possible underestimate, that the core of information in the subspecialties of internal medicine does not exceed 2 million facts.

[10]Note that we are only concerned with real-world knowledge that is relevant to medicine, not with all such knowledge possessed by the average person.

6.6.2 Can the Knowledge Base of Internal Medicine Be Stored at a Reasonable Cost?

Can 2 million facts be stored in a computer system at a reasonable cost? If we assume that each fact requires for its representation an average of 10 words of computer memory,[11] a computer storage capacity of 20 million words would be necessary. A memory of this size is certainly large, and, if core storage were required, the cost would at present be prohibitive. Because only a small part of the data is used at any one time, an inexpensive mass storage device (such as a magnetic disc or drum) would be a practical alternative storage medium; even at present prices, the cost would probably be no more than $20,000.[12]

We should note, furthermore, that even with rapid progress in the development of "expert" consulting programs, it is unlikely that a system could reach the size we have envisioned in a period of less than ten years. By that time, given the rapid evolution of computer technology, it is almost certain that a memory capable of storing 2 million facts could be purchased at a very low cost. From these considerations, it can reasonably be concluded that data storage will not be the limiting factor in the development of consulting capability within the computer. Indeed, even the storage of an additional large body of specialized information drawn from the literature, should, some years hence, pose no great technological difficulties.

6.6.3 Can the Data Base of Internal Medicine Be Efficiently Managed?

The problems of organizing, retrieving, and applying the relevant data are far more formidable than is the problem of data storage. We believe, however, that the task is probably not insuperable, because in any given case only a very small fraction of the available knowledge needs to be retrieved. Furthermore, the retrieval of whatever information is required will be greatly eased by the fact that pertinent information can be dealt with in the highly organized clusters known as frames. Assuming that the average frame will contain on the order of 100 facts, only 20,000 frames would be required for the postulated data base of 2 million facts.

Probably the most difficult aspect of data management will be the problem of *coding*, the process of ensuring that each fact is properly associated with other facts. Only if the large data base of internal medicine

---

[11] We are considering this representation in a computer language such as LISP, which is quite efficient at storing and retrieving the type of symbolic data that we envision will be used.

[12] *Ed. note:* This cost estimate, accurate in 1976 when this article appeared, is excessive in 1984 due to technological advances.

can be transferred automatically from English to the appropriate representation within the computer is there hope that serious errors, omissions, and contradictions can be avoided. Current efforts to develop computer programs that understand English give promise that this fundamental problem will eventually be solved (Schank and Colby, 1973).

The arguments we have considered here have led us to the conclusion that, over the long term, there are not likely to be intrinsic technological constraints on the realization of a system capable of coping with all of internal medicine. In fact, the availability of increasingly powerful technology suggests a future in which computer programs may well "know" far more than any individual physician. For the short term, however, we look toward the development of programs that know a great deal, but not all, of internal medicine.

6.6.4 Some Reflections on the Cognitive Process

As discussed earlier, the present illness simulation described here is based on insight derived from introspection and from observation of the problem-solving behavior of experienced clinicians. Here we offer a brief discussion of certain key ideas that we believe merit further study by investigators interested either in computer-aided decision making or in clinical cognition.

Our study clearly illuminates an important difference between the expert in practice and the expert as often pictured in literature or folklore. The epitome of the expert in fiction is the detective who, through superior deductive powers and by sheer force of logic, organizes the facts at hand in such a way that they lead to a single, inevitable conclusion. By contrast, the real-world clinician seems to rely much more heavily upon "guessing," the initial hypothesis typically being based on precious little data. These guesses are apparently prompted by patterns of clinical findings or by specific complaints that bring to mind particular diseases. The physician then tries to demonstrate the correctness of his or her guesses, moving to new hypotheses only if the initial impressions prove untenable.

The rapidity with which the initial hypotheses are generated and the ostensibly fragile basis of the guessing process together constitute the most striking feature of the behavior of experienced clinicians. Often with only the age, sex, and presenting complaint of the patient, the clinician unhesitatingly selects a single working hypothesis. Even in ambiguous situations, he or she rarely begins with more than a few hypotheses.

Another characteristic of the experienced physician is the fashion in which he or she continually pares the list of diagnostic possibilities. The physician discards some, accepts others, and often combines individual possibilities into a single, new, integrated hypothesis. In this way, he or she is generally able to limit sharply the number of diagnoses which must actively be considered. We can understand the value of such a sharp focus when we consider that, in taking a present illness, the physician can gather only a small fraction of the potential set of facts concerning the patient and must therefore seek information very selectively. In consequence, the clinician must find a context within which to properly focus his or her questioning and to organize the information that is obtained.

Because the initial hypotheses are usually generated on the basis of relatively few facts, they will often later prove to be incorrect. In such cases, how does the experienced clinician proceed to undo any "damage" done by aggressive hypothesis generation? Our observations suggest that he or she often employs the rather efficient strategy of associating one hypothesis with others with which it may be readily confused (e.g., "multiple pulmonary emboli are often confused with cardiomyopathy"). By explicitly remembering such situations, the physician can move directly from a hypothesis that has become suspect to one that offers another plausible explanation for the presenting findings.

Unlike the seasoned clinician, the medical student or young physician does not have an extensive knowledge of such relations and so is unlikely to move from one hypothesis to another in such a skillful fashion. Therefore, the novice who acts aggressively in hypothesis generation risks making serious errors. We have observed that the student or house officer, apparently to counter this problem, often approaches the diagnostic process in a highly structured, methodical fashion. Similarly we have noted that the experienced physician performing outside his or her area of expertise uses a far more structured approach than is his or her usual custom. The seasoned clinician's expertise in taking a present illness thus appears to derive in considerable part from a complex set of associations and from a familiarity with many alternative scenarios within that individual's "frames."

We believe that the experimental methods utilized in the present study, if extensively employed, will provide important new insights into the process of clinical problem-solving. Furthermore, as our understanding of problem-solving processes grows, it seems likely that the study of clinical cognition will assume a significant place in the medical curriculum. Such increased attention to this neglected aspect of medical education should eventually make an important contribution to improving the quality of physician performance.

ACKNOWLEDGMENTS

This research was supported in part by the Health Resources Administration, U.S. Public Health Service, under grant 1 R01 MB 00107-01 from the Bureau of Health Manpower and under grant HS 00911-01 from the National Center for Health Services Research.