TOWARDS A COMPUTER MODEL OF PSYCHIATRIC REASONING *

Janet L. Kolodner

School of Info. and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332

Robert M. Kolodner

Veterans Administration Medical Center
Department of Mental Health
1670 Clairmont Road
Decatur, GA

ABSTRACT

Human experts are able to introspect about their knowledge and learn from past experience. It is this view of expertise that we are exploring. This paper will present the basis for this view, the reasoning model it implies, and a computer program which begins to implement the theory. The program, called SHRINK, simulates a psychiatrist.

INTRODUCTION

In the past few years, computers have entered more and more into the process of clinical decision making by doctors. Within the field of artificial intelligence, the programs that have been developed are called "expert systems". Expert systems research to date has concentrated on extracting the rules that experts use (e.g., [3]), developing problem solving methodologies for dealing with those rules (e.g., [6]), and developing consultation protocols so that diagnostic questions will be presented in coherent order to doctors using a system (e.g., [1]). The premise behind this work has been that the knowledge used for diagnosis and treatment is of primary importance. Furthermore, the claim is that representing that knowledge in the form of rules makes use and explanation of a system easy. Thus, the stress has been on extracting rules from experts and building systems which from the beginning contain all of the compiled knowledge an expert has.

We are considering an alternate approach to expert reasoning. In this approach, knowledge is built up incrementally on the basis of experience. Reasoning processes, the knowledge they use, and incremental changes in the organizational structure of the knowledge and the reasoning processes are equally important. Experience is organized in diagnostic episodes in a long term memory, and guides reasoning processes. Failures in reasoning processes guide incremental change to the memory organization and thus to the reasoning processes themselves. Successful diagnosis and treatment reinforces knowledge and processes already in use.

The evolution from novice to expert requires introspection and examination of the knowledge used in solving problems. This results in learning from experience. A human expert can interpret a new case in terms of a previous case or set of cases he is already familiar with. This implies that as an expert is having new experiences, he is evaluating and understanding them in terms of previous ones. In the process, the new experiences gets integrated into his memory so that it too will be accessible to use in understanding a later case.

Our long term goal is to address the following problems:

1. What makes an expert expert?
2. What processing capabilities are implied when we say somebody is an expert?
3. What understanding capability does an expert have that a novice does not?
4. Which processes comprise expert reasoning?
5. What is the relationship between experience and expertise?

This requires discovery of the processes used in expert reasoning, how memory must be organized to access that knowledge easily, and how memory's organization changes over time. We are drawing on recent research into the organization of experience in long term memory ([4], [5], [7]), observing physicians as they make decisions, and building a computer program, called SHRINK, which implements and tests our theory of expertise in the domain of psychiatry. When complete, the program will analyze new psychiatric cases based on previous cases it has seen. It will integrate new cases into its memory as it is processing them, and will build up its expertise based on generalizations it has made concerning the similarities between cases it has seen. Its expertise will also be heightened through analysis of failures in diagnosis and treatment.

EPISODIC MEMORY

To explain our approach to expert reasoning, it will be necessary to explain episodic memory. Episodic memory is memory for experience. It records and organizes individual episodes or events in a person's life. It also creates and records generalized episodes. To see what that means, consider the fact that some episodes are reminiscent of others. When one is "reminded" [7] of a previous episode, the similarities between the two can be extracted to form a generalized episode. From two experiences at restaurants, for example, a generalized restaurant episode is formed which holds knowledge about events that usually happen in restaurants. If a doctor sees a particular set of symptoms a number of times, and if he finds that a

* The work of the first author is supported in part by NIH-BRSG Grant #5 S07 RR 07024-16 and in part by NSF Grant #IST-8116892.

particular drug seems to work in each of those cases, a similar type of generalized episode can be formed.

This generalization capability is important for a number of reasons. First, it is economical in terms of storage. Once a generalized episode is formed, full details of the individual episodes which support it don't have to be stored. Instead, only differentiating features of each individual episode need be recorded. Generalized episodes are also useful as organizational structures. The generalized episode serves an an organizing point for storing similar episodes with respect to each other. If memory is organized in this way, memory search (i.e., remembering) can be directed to only the most relevant items. Generalized episodes also aid understanding and reasoning processes. When a new experience is reminiscent of a generalized episode, the generalizations formed previously can be used to analyze and act on the new experience.

Memory Organization for Psychiatric Diagnosis

We call these generalizes structures for episodes Episodic Memory Organization Packets (E-MOPs) [5]. In developing a medical system, we must determine what types of episodes the system should be aware of. Doctors see patients, diagnose their illnesses and treat them. Thus, our system must be aware of at least "DIAGNOSIS" and "TREATMENT" as generalized types of episodes. The "DIAGNOSIS" E-MOP used in SHRINK involves the following steps [2]:

Psychiatric DIAGNOSIS

patient data
||
input
||
\/
establish primary indices
(i.e., predominant clinical features)
||
next
||
\/
differential diagnosis
identifies all possible disorders
||
next
||
\/
diagnostic evaluation
(check each in more detail, narrow list)
||
output
||
\/
short list of probable disorders

Each of these steps in turn is made up of a number of steps. This process is a textbook process, and one appropriate for "naive reasoning". SHRINK starts with this naive reasoning process. It currently uses it to diagnose normal cases of major depression. As its learning processes are developed, it will refine this process to make it a more expert one.

An individual diagnostic experience happens each time a patient is diagnosed. If the diagnostic process and data are similar for a group of patients, then generalizations are made about that group of patients and compiled paths through the process will be established. Such paths allow uncomplicated cases of particular disorders to be easily identified. Thus, there will be a compiled path with particular signs, symptoms, primary indices, differential diagnoses, etc. filled in for each commonly seen disorder. A path through the process above for diagnosing a first episode of major depression consists of establishing that the predominant clinical feature is a mood disturbance, doing a differential diagnosis outputing "major depression" as the most probable disorder, and in the diagnostic evaluation finding that none of the exclusion criteria for major depression are present and that there have been no previous depressive or manic episodes. The output of the diagnostic evaluation will be "Major Depression, Single Episode".

At the end of each compiled path through DIAGNOSIS will be pointers to the identified disorder and to relevant treatment information. Thus, there is a second type of episodic information that the memory must have -- disorder knowledge. Knowledge about a particular disorder includes its identifying signs and symptoms and how to treat it. It must include knowledge about how other diagnostic categories relate to it. In addition, each disorder category must organize individual cases of the disorder. Knowledge about individual cases is necessary in developing new treatment and diagnostic rules and in refining existing ones.

Below is some of the information SHRINK has about "Depressive Episodes" [2]. Presence of a "depressive episode" is necessary to diagnose all "major affective disorders". In the program, this information is represented in a language-free form. We are showing it in English to make it easily readable.

A third type of knowledge the system must have is symptom knowledge -- knowledge of how particular symptoms and their combinations tend to predict particular diagnostic categories. Attempted suicide, for example, suggests a possibility of depression.

Each step of the diagnostic process is guided by either disorder or symptom knowledge. After predominant clinical features (major symptoms) are chosen, the knowledge associated with them is used to choose starting points for differential diagnosis. Only those categories implied by the primary symptoms are considered in initial differential diagnosis. Differential diagnosis is guided by knowledge about the relationships between diagnostic categories. Disorder knowledge (in particular, knowledge about the normal symptoms and exclusion criteria associated with particular

disorders) also guides diagnostic evaluation once possible disorders have been established through differential diagnosis.

DEPRESSIVE EPISODE

ESSENTIAL FEATURES:
  at least 1 of:
    (1) dysphoric mood
    (2) pervasive loss of interest or pleasure in usual pastimes and activities.
SUPPORTING FEATURES:
  at least 4 of:
    (1) a significant change in appetite.
    (2) a disturbance in sleep habits.
    (3) psychomotor agitation or retardation.
    (4) decrease in sexual drive.
    (5) loss of energy or fatigue.
    (6) feelings of worthlessness, self-reproach, or excessive or inappropriate guilt
    (7) complaints or evidence of diminished ability to think or concentrate.
    (8) recurrent thoughts of death, suicide, death wishes, or suicide attempt.
  time constraint: symptoms must have been present simultaneously and for at least 2 weeks
EXCLUSION CRITERIA:
 (1) preoccupation with a mood-incongruent delusion or hallucination (=> PSYCHOTIC DISORDER)
 (2) bizarre behavior (=> poss PSYCHOTIC DISORDER)
ADDITIONAL CRITERIA:
   negate diagnoses of:
     (1) SCHIZOPHRENIA
     (2) SCHIZOPHRENIFORM DISORDER
     (3) ORGANIC MENTAL DISORDER
     (4) UNCOMPLICATED BEREAVEMENT

INCREMENTAL LEARNING

In order for failures in procedure to refine previously-held diagnostic rules, memory's organization must be updated with each new experience. For that to happen in the memory structures we have defined, two things must happen when an experience deviates from the expected. First, the deviant path through "DIAGNOSIS" must be recorded. Second, the differences between the well-known diagnostic categories and the deviant case must be recorded.

There are two reasons for recording these differences. First, they should be recorded so that when a similar deviation occurs later, the original case can be remembered. The two cases can then be examined for similarities, and generalizations based on their similarities can be used to evolve a new diagnostic category. When later similar cases occur, the knowledge necessary to deal with them will already be in memory. Second, recording deviations acts as a signal that additional reasoning must be attempted. Extra effort can then be applied to "explain" the deviation. If an explanation is found, diagnostic procedures are updated, and that case is maintained as support of the new procedure. If no explanation is found, the deviation marks a problem that must be resolved later. Later, when a similar problem case is encountered, the marker signals that both the old case and the new one should be examined to see if between them they provide enough evidence to explain the violation.

An Example

Consider the following example (adapted from [9], Case #125), which illustrates some of the processes alluded to above:

> Dr. X sees a patient who seems to show classic signs of major depression. She is 38 and complains of depression. In the last month, she has been unhappy, suffering from insomnia and crying. She reports poor concentration and diminished interest in activities, and that she has been depressed since childhood when her father deserted the family. She has previously been diagnosed as depressive, and was treated in a mental hospital with antidepressants. She was sickly as a child, has had a drinking problem, and has had a number of physical illnesses which doctors have not been able to find causes for.
>
> Seeing that she has been treated previously for depression, that her chief complaint is depression, and that she has insomnia, poor concentration, and diminished interest, Dr. X concludes that this patient is suffering from Major Depression, Recurrent, without Melancholia. He treats her with antidepressants. The antidepressants seem to work, but the woman keeps coming back complaining of additional major physical disorders. Dr. X begins to think that there may be some other problem which he had not accounted for.
>
> Taking a further history, he finds out more about the medical problems she has had. They have been numerous, and doctors have been unable to find organic reasons for them. He realizes that the large number of medical problems is important to consider. Going through the diagnostic process again using that symptom as the predominant clinical feature, he realizes that he should have diagnosed her for Somatization Disorder in addition to the diagnosis of Depression.

As a result of this case, Dr. X should learn the following:

1. it is important to take medical history into account in choosing predominant clinical features,
2. depression can camouflage somatization disorder, and
3. a patient who is highly depressed but who complains about medical problems may be suffering from somatization disorder in addition to depression.

Using the first fact, he should be able to refine his rules for choosing predominant clinical features. This case should help him conclude that

medical history as a primary index may be more important than he had thought. In addition, the next time he finds that a treatment has failed, he may be biased towards finding out facts about the patient's medical history that he hadn't known. Using the second fact, he should be skeptical of diagnoses of depression coming from other doctors, and will want to find out more about the medical history of a new patient before taking a previous diagnosis seriously. The third fact gives the relationship between Depression and Somatization Disorder, which could be helpful in diagnosing and treating later cases.

Furthermore, this case should enable the doctor to hypothesize that there is often more than one symptom that is primary, and that the patient's chief complaint may not be the most important symptom to look at. Current problems must be separated from long-standing problems and both must be taken into account. Later cases treated by this doctor should enable him to recognize and separate current from long-standing problems, to confirm these two hypotheses, and to learn specifically which long-standing problems and which chief complaints are likely to be significant.

The Process in More Detail

The process of incremental learning can be divided into six parts:

1. initial decision
2. noticing the failure
3. blame assignment
4. correcting the failure
5. explanation of the problem
6. memory update

Initial decision -- The first step in incremental learning is to follow the current procedure to come to a conclusion. In this case, that involves following the normal DIAGNOSIS procedure to make a diagnosis. Noticing the patient's reported depression and using his knowledge about unhappiness and mood disturbance, the doctor makes a diagnosis of Major Depression, Recurrent, without Melancholia. He prescribes anti-depressants, the common treatment for depression.

Noticing the failure -- Any unexpected result must be noticed and explained [8]. In this case, the woman exhibited the expected result plus additional unexpected physical symptoms.

Assigning blame -- After noticing a violation, blame must be assigned for the failure. In this case, the task is to explain what caused the woman's new symptom. There are three possibilities:

a. the medicine is producing side effects,
b. she has developed a new physical ailment independent of the medication, or
c. she is imagining her disorders.

Each must be checked. We assume the first doesn't hold. To check the second, appropriate diagnostic tests, a further history, etc. must be done. If no organic reason for the illness is found, then the doctor should be reminded of the patient's previous history. She has a history of physical illnesses for which doctors have been unable to find causes. This should also remind the doctor of the following rule, causing him to check the possibility of Somatization.

> Excessive medical history in females implies a Somatization Disorder.

Note that this process requires many types of "reminding". The reasoner must be reminded of previous cases similar to the current one so that knowledge gleaned from those cases can be used to reason about the new case. In addition, patterns of behavior associated with the current patient must be available for recall. Reminding must also allow symptom knowledge to be remembered.

Our claim is that knowledge becomes available only when there is a way to direct memory processes toward it. Initially, the doctor was focussing on mood disturbance, the then current problem. As a result, only knowledge associated with mood disturbances was accessible for reasoning. Only when attention was directed to the patient's physical disorders and medical history could the knowledge associated with those symptoms become available. The memory organization described above supports such processing.

Correcting the failure -- The doctor can now hypothesize that the patient has a Somatization Disorder in addition to depression. Repeating the diagnosis, he would find that the hypothesis is supported.

Explaining the failure -- To learn from this episode, the doctor must figure out whether and why he made an initial mistake. There are four possibilities:

a. the diagnosis was wrong, and therefore the treatment is unsatisfactory,
b. the diagnosis was right, but the treatment was not appropriate
c. the diagnosis was right, but the treatment didn't work, or
d. the treatment and diagnosis were right, but something new has come up.

To distinguish between (a) and (d), the possibilities in this case, the doctor must decide whether he had the necessary information in the beginning to make the correct diagnosis. The key to making that decision is figuring out where in the diagnosis there might have been a problem and how it could have been corrected. Because the failure was tracked to a rule associated with establishing primary indices, the failure can be assumed to have occurred in that step of the process. He must thus determine whether he had enough information initially to include Somatization Disorder in his original differential. He did have the necessary knowledge. His conclusion must be that he should have paid attention to that initially, and marks "medical history" as a patient

feature which he should pay more attention to in the future.

Memory update -- In addition to concluding that medical history is an important clinical feature to consider in diagnosis, the relationship between Somatization Disorder and Depression is also learned. This will be represented in a number of ways in episodic memory.

IMPLICATIONS AND DISCUSSION

We take human expert reasoning as a model and attempt to both understand the human model and implement a computer program which copies it. Our long-range goal is to understand expertise. Our shorter-range goal is to understand expertise in the domain of medical reasoning, particularly psychiatry.

This research is relevant to medicine for two reasons — education and decision support. First, an understanding of the reasoning processes involved in medical diagnosis and treatment and of the relationship between experience and expertise, can ultimately aid in educating new experts. Second, if we understand expert reasoning, then we can build expert computer systems to aid medical practitioners in their decision making.

There are a number of advantages and disadvantages to the approach we are taking in terms of building expert consultation systems. The first major advantage of our approach is that it is geared toward exceptional cases or novel situations, exactly those kinds of situations which a clinician in need of further advice cannot deal with himself.

Another advantage this approach has is that it deals with both experiential knowledge and facts in the same way. Both are stored in the same structures and organized identically. This means that both are equally accessible and both can be used in reasoning. An implication of this organization is that it is amenable to new information from any source.

If we are aiming towards building expert consultation systems, this approach has the disadvantage of being highly complex. Rule-based expert systems (e.g., MYCIN) have the advantage of having a fairly simple algorithm. This makes them easy to understand by non-computer scientists. Our approach, based largely on learning, is not as simple. It may be difficult to verify a new generalization the system has made, perhaps making the system's whole operation suspect. Another challenge associated with making a system such as this acceptable as a diagnostic consultant is giving it the capability of explaining itself. This is important since acceptance by clinicians is dependent on being able to follow a program's rationale.

In the long run, we think that these will not be limitations. In developing a system based on human reasoning, we hope to be developing a system which can explain itself easily. Its process of diagnosis will be similar to what the diagnostician would be doing if he were an expert in the field. Explanation problems should be equivalent to those found when an expert in one field attempts to explain his conclusions to an expert in another field.

This paper has pointed out how experience aids in developing the expertise necessary for expert reasoning. The research and the program are still in a state of infancy. Nevertheless, we see this approach as having a great deal of promise both in terms of implementing expert computer systems and in helping us to understand the cognitive processes underlying expertise. The system we are proposing, which learns from its own experiences and from the experiences of others, should ultimately be able to diagnose and treat illnesses, and also to keep up with new practices and treatments as a good human expert does.

ACKNOWLEDGEMENTS

Keith McGreggor has been the programmer for this effort. He has also been helpful in formulating the ideas presented in this paper.

REFERENCES

[1] Aikens, J. S. (1980). Prototypes and production rules: A knowledge representation for computer consultations. Technical Report. Dept. of Computer Science, Stanford University, Palo Alto, CA.

[2] American Psychiatric Association (1980). Diagnostic and Statistical Manual of Mental Disorders (Third Edition). APA, Washington, D.C.

[3] Davis, R., Buchanan, B. G., and Shortliffe, E. H. (1977). Production rules as a representation for a knowledge-based consultation program. Artificial Intelligence, Vol. 8, pp. 15-45.

[4] Kolodner, J. L. (1981a). Organization and retrieval in a conceptual memory for events. In Proceedings of the International Joint Conference on Artificial Intelligence, Vancouver, B.C., Canada.

[5] Kolodner, J. L. (1981b). Knowledge-Based Self-Organizing Memory. In Proceedings of the IEEE International Conference on Cybernetics and Society, Atlanta, GA.

[6] Pople, H. E. (1977). The Formation of Composite Hypotheses in Diagnostic Problem Solving. In Proceedings of ICJAI-77.

[7] Schank, R. C. (1980). Language and memory. Cognitive Science, Vol. 14, No. 3.

[8] Schank, R. C. (1982). DYNAMIC MEMORY: A Theory of Learning in People and Computers. (forthcoming).

[9] Spitzer, R. L., Skodol, A. E., Gibbon, M., and Williams, J. B. W. (1980). DSM-III Case Book. APA, Washington, D.C.