Reprinted from PROCEEDINGS MEDCOMP '82 - FIRST IEEE COMPUTER SOCIETY INTERNATIONAL CONFERENCE ON MEDICAL COMPUTER SCIENCE/COMPUTATIONAL MEDICINE, September 1982

MINIMAL SET COVERS AS A MODEL FOR DIAGNOSTIC PROBLEM SOLVING

James A. Reggia,* Pearl Y. Wang, and Dana S. Nau
University of Maryland

Abstract: This paper suggests that the minimal set covering problem provides a reasonable model for diagnostic problem solving. Such a model is useful because it directly addresses the problem of multiple simultaneous disorders, and because it offers a framework for a formal theory of diagnostic inference. In addition, this approach provides a new basis for computer-aided diagnostic decision making. We present the model here using intuitive rather than formal terms and illustrate its application to medical diagnosis.

In this paper we introduce a theoretical model of diagnostic reasoning based on the concept of minimal set covers. This "set cover model" is of interest because it addresses the problem of multiple simultaneous diseases, it provides the basis for a formal theory of diagnostic inference, and it can be used as a unifying framework in which to view decision support systems (DSS's) that model diagnostic reasoning.

At the present time, statistical pattern classification and rule-based deduction are the predominant methods used to implement diagnostic DSS's. While these two methods have achieved some notable successes, both face significant limitations to their general applicability [12]. Furthermore, efforts to overcome these limitations by developing DSS's which model the diagnostic reasoning process of the physician have been only partially successful. Diagnostic problems where multiple diseases are present simultaneously have proven extremely difficult to handle [10]. In addition, these models of diagnostic reasoning are often criticized as being "ad hoc" because of the absence of a formal theoretical foundation [e.g., 2]. This is in contrast to statistical pattern classification (based on probability theory) and rule-based deduction (founded in the first-order predicate calculus and formal language theory).

Our set cover model addresses each of these issues. We describe the model here in intuitive rather than formal terms and illustrate it with an implemented DSS based on the model. A subsequent paper will present the model in a more formal fashion and in greater detail.

*Corresponding Author: James A. Reggia, Department of Neurology, University of Maryland Hospital, Baltimore, Maryland 21201.

DIAGNOSTIC PROBLEM SOLVING: EMPIRICAL RESULTS

Several studies of the physician's diagnostic reasoning process have been published during the last few years (see [12] for a review). This work has shown that diagnostic reasoning involves a sequential hypothesize-and-test process during which the physician conceptually constructs a model of the patient. This model, subsequently referred to as the hypothesis, is based largely on what disease manifestations (signs or symptoms) are known to be present. It postulates the presence of one or more diseases that could explain these manifestations. To construct and modify the hypothesis, the physician relies on his medical "knowledge base," which ideally includes the set of all possible causative diseases for each manifestation, and the set of all possible manifestations for each disease.

The physician's hypothesis may at times be relatively complex. Not only may it contain a great deal of uncertainty about which of several diagnoses account for a certain manifestation, but it might also presume the simultaneous presence of multiple diseases. The empirical evidence suggests that the hypothesis can best be viewed as a resolution of two conflicting goals:

Coverage Goal: the goal of explaining all of the patient's manifestations;

Minimality Goal: the goal of minimizing the complexity of the explanation.

The second goal is sometimes referred to as the "Principle of Parsimony" or "Occam's Razor."

It is important to appreciate the sequential nature of diagnostic reasoning. As the physician gradually learns information about a patient, his or her hypothesis changes to reflect this new information. For example, if a patient complained of sudden onset of chest pain, the physician's initial hypothesis might be something like:

"heart attack, or pulmonary embolus, or ...".

Acknowledgement: Research supported by NINCDS, grants 5 K07 NS 00348 and 1 P01 NS 16332. Computer time provided in part by Computer Science Center, University of Maryland.

As further details became available, some of these initially possible diseases might be eliminated. If it was then learned that the patient also had a chronic cough and was a heavy smoker, the hypothesis might change to

"heart attack, or pulmonary embolus, or ..." and "bronchitis, or asthma, or ...",

reflecting the physician's belief that at least two diseases must be present to account for this patient's symptoms. Note that at this point, the hypothesis contains both uncertainty (indicated by "or") and the presumption that multiple simultaneous diseases are present (indicated by "and").

A FRAMEWORK FOR DIAGNOSTIC PROBLEM SOLVING

To provide a theoretical framework for viewing diagnostic problem solving, at least two points need to be addressed. First, an abstract representation must be specified for the underlying medical knowledge and for the hypothesis that explains a patient's manifestations. Second, a model must be stated for the sequential diagnostic problem solving process itself in terms of the chosen abstract representation.

Representation. A simple but reasonable representation for the knowledge used to solve a diagnostic problem involves specifying two sets (Figure 1). The first set $D = \{d_1,\ldots,d_n\}$ represents all possible diseases d_i, while the second set $M = \{m_1,\ldots,m_k\}$ represents all manifestations that can be caused by any disease or disorder in D. The key concept here is that associated with each d_i in D there is a corresponding subset of M, designated $man(d_i)$, which represents all possible manifestations caused by disorder d_i (Figure 1). When d_i occurs, some but not necessarily all of its manifestations may be observed. Similarly, for each m_j in M, there is a corresponding subset of D, designated $causes(m_j)$, which represents all possible causes of manifestation m_j (Figure 1). We assume that the definitions of these sets of manifestations and causes are consistent: if m_j is in $man(d_i)$, then d_i must be in $causes(m_j)$, and vice versa. We also assume that $man(d_i)$ and $causes(m_j)$ are never empty. Finally, if $D = \{d_i,\ldots,d_j\}$ is a subset of D, then $man(D)$ is defined to be the union of $man(d_i)$ for all $d_i \in D$. Similarly, if $M = \{m_1,\ldots,m_n\}$ is a subset of M, then $causes(M)$ is the union of $causes(m_i)$ for all $m_i \in M$.

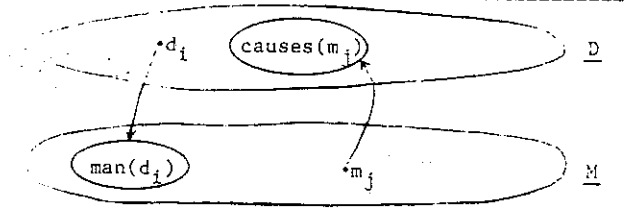

Figure 1: Modelling the knowledge base.

Given this representation for medical knowledge, a specific patient's manifestations can be represented as a subset M^+ of M (Figure 2). The goal in solving a diagnostic problem centers on explaining the presence of M^+ by postulating the presence of one or more disorders in D. We can therefore proceed to develop a framework for solving a diagnostic problem as follows. A set of disorders E in D is defined to be an explanation for M^+ if:

i) $M^+ \subseteq man(E)$
(The disorders in E together can account for all manifestations present in M^+. We say E covers M^+.); and

ii) $|E| \leq |D|$ for any other cover D of M^+
(No other set of disorders D which satisfies part (i) is of smaller cardinality).

These two requirements are intuitively plausible, and correspond to the Coverage Goal and Minimality Goal stated earlier. As a simple example, suppose M^+ is given while d_1 and d_7 are diseases with manifestation sets as shown in Figure 2. Clearly $D = \{d_1\ d_7\}$ covers all of M^+, so D would be an explanation if no other single disorder in D could cover M^+.

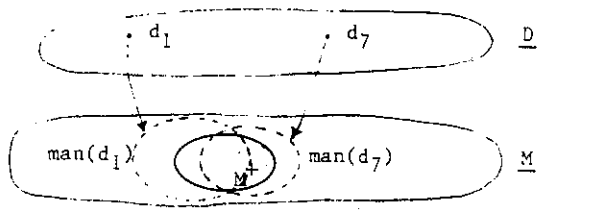

Figure 2: Explaining a set of manifestations.

The above definition of "explanation" is not only intuitively plausible, but it also implies that finding an explanation for a set of manifestations M^+ is analogous to the well-known minimal set cover problem, and that an explanation corresponds to the concept of a minimal set cover [4]. The only difference from the traditional set cover problem in mathematics is that we do not require an explanation as defined here to cover M^+ exactly. The set man(E) for any explanation E must contain M^+, but it might also contain manifestations not present in M^+ (as illustrated in Figure 2). This difference reflects the fact noted earlier that sometimes when a disease is present not all of its manifestations occur.

Having defined an explanation, we can now proceed to define the solution of a diagnostic problem to be the set of all explanations for M^+. This is analogous to finding all minimal set covers. For example, suppose $D = \{d_1,d_2,\ldots,d_9\}$ and $M = \{m_1,m_2,\ldots,m_6\}$, and the relevant relationships between elements in these two sets are as illustrated in Figure 3. Let a specific patient's manifestations consist of $M^+ = \{m_1,m_4,m_5\}$. Since no single disorder can cover all of M^+, the solution to this problem would consist of the following eight possible

explanations: $\{d_1\ d_7\}$, $\{d_1\ d_8\}$, $\{d_1\ d_9\}$, $\{d_2\ d_7\}$, $\{d_2\ d_8\}$, $\{d_2\ d_9\}$, $\{d_3\ d_8\}$, and $\{d_4\ d_8\}$. This solution can be interpreted as saying that any one of these eight sets satisfies both the Coverage Goal and the Minimality Goal and is therefore a plausible explanation for the patient's manifestations.

d_i	man(d_i)	d_i	man(d_i)
d_1	$\{m_1\ m_4\}$	d_6	$\{m_2\ m_3\}$
d_2	$\{m_1\ m_3\ m_4\}$	d_7	$\{m_2\ m_5\}$
d_3	$\{m_1\ m_3\}$	d_8	$\{m_4\ m_5\ m_6\}$
d_4	$\{m_1\ m_6\}$	d_9	$\{m_2\ m_5\}$
d_5	$\{m_2\ m_3\ m_4\}$		

m_j	causes(m_j)	m_j	causes(m_j)
m_1	$\{d_1\ d_2\ d_3\ d_4\}$	m_4	$\{d_1\ d_2\ d_5\ d_8\}$
m_2	$\{d_5\ d_6\ d_7\ d_9\}$	m_5	$\{d_7\ d_8\ d_9\}$
m_3	$\{d_2\ d_3\ d_5\ d_6\}$	m_6	$\{d_4\ d_8\}$

Figure 3: A simple abstract knowledge base.

Rather than representing the solution to a diagnostic problem as an explicit list of all possible explanations for M^+, it is advantageous to represent the disorders involved as a collection of explanation generators. An explanation generator is a collection of sets of "competing" disorders that implicitly represents a set of explanations in the solution and can be used to generate them. A generator is analogous to a Cartesian set product, the difference being that the generator produces unordered sets rather than ordered tuples. To illustrate this idea, consider the example diagnostic problem in the preceding paragraph. Two generators are sufficient to represent the solution to that problem: $\{d_1\ d_2\} \times \{d_7\ d_8\ d_9\}$ and $\{d_3\ d_4\} \times \{d_8\}$. The second generator here implicitly represents the two explanations $\{d_3\ d_8\}$ and $\{d_4\ d_8\}$, while the first generator represents the other six explanations in the solution.

There are at least three advantages to representing the solution to a diagnostic problem as a set of generators. First, this is usually a more compact form of the explanations present in the solution. Second, generators are a very convenient representation for developing algorithms to process explanations sequentially (see below). Finally, and perhaps most important, generators are closer to the way the physician organizes the possibilities during problem solving (i.e., the "differential diagnosis"). For example, compare the form of the first generator immediately above with the differential diagnosis for the patient with acute chest pain given previously.

Sequential Problem Solving. As noted earlier, diagnostic problems in the real world are highly sequential in nature. Rather than simply being given M^+, the physician is typically given only one or more manifestations in M^+, and must actively seek the rest by questioning, examining and testing the patient. It is during this process that the physician constructs and subsequently revises a tentative hypothesis about what is occurring in the patient.

This sequential process can be modelled in terms of the representational framework presented above. The tentative hypothesis at any point during problem solving is defined to be the solution for those manifestations already known to be present, assuming, perhaps falsely, that no additional manifestations will be subsequently discovered. To construct and maintain a tentative hypothesis like this, three simple data structures prove useful:

MANIFS: the set of manifestations known to be present so far;

SCOPE: causes(MANIFS), the set of all diseases d_i for which at least one manifestation is already known to be present; and

FOCUS: the tentative solution for just those manifestations already in MANIFS; FOCUS is represented as a collection of generators.

Using these three data structures, a hypothesize-and-test algorithm can simulate diagnostic problem solving. As each new manifestation m_i is discovered to be present, the three data structures are adjusted to accomodate this new information. The working hypothesis is then examined to select further information that would be helpful, and the entire process is repeated.

The central issue here is how to maintain the three data structures MANIFS, SCOPE and FOCUS in a fashion consistent with the properties of a solution as described above. This is done as follows. As each manifestation m_i that is present is discovered, MANIFS is updated simply by adding m_i to it. SCOPE is augmented to include any possible causes d_j of m_i which are not already contained in it (set union of causes(m_i) and SCOPE). Finally, FOCUS is adjusted to accomodate m_i based on intersecting causes(m_i) with the sets of diseases in the existing generators. These latter operations are done such that any explanation implicitly represented by the FOCUS which can no longer account for the augmented MANIFS (which now includes m_i) are eliminated.

Perhaps the best way to understand this process is to follow a simple example. Recall the abstract knowledge base illustrated in Figure 3, and consider the same diagnostic problem $M^+ = \{m_1\ m_4\ m_5\}$ that was used earlier. Suppose that the sequence of events occurring during problem solving were ordered as listed in the first column of Figure 4. Initially, MANIFS, SCOPE and FOCUS

are all empty. When m_1 is discovered to be present, m_1 is added to MANIFS, and the new SCOPE is the union of the old SCOPE with causes(m_1). Since there are no generators in FOCUS, the intersection of causes(m_1) with them is trivially empty. In such situations, a new generator is created, in this case consisting of causes(m_1). In the terms defined earlier, this generator represents a solution for $M^+ = \{m_1\}$. It tentatively postulates that there are four possible explanations for M^+, any one of which consists of a single disease. The FOCUS thus asserts that "d_1 or d_2 or d_3 or d_4 is present."

Events in order of their discovery	MANIFS, SCOPE and Generators in FOCUS
Initially	all empty
m_1 present	
MANIFS:	$\{m_1\}$
SCOPE:	$\{d_1\ d_2\ d_3\ d_4\}$
FOCUS:	$\{d_1\ d_2\ d_3\ d_4\}$
m_2 absent	all unchanged
m_3 absent	all unchanged
m_4 present	
MANIFS:	$\{m_1\ m_4\}$
SCOPE:	$\{d_1\ d_2\ d_3\ d_4\ d_5\ d_8\}$
FOCUS:	$\{d_1\ d_2\}$
m_5 present	
MANIFS:	$\{m_1\ m_4\ m_5\}$
SCOPE:	$\{d_1\ d_2\ d_3\ d_4\ d_5\ d_7\ d_8\ d_9\}$
FOCUS:	$\{d_1\ d_2\} \times \{d_7\ d_8\ d_9\}$, $\{d_8\} \times \{d_3\ d_4\}$
m_6 absent	all unchanged

Figure 4: A sequential diagnostic problem.

The absence of m_2 and m_3 do not change this initial hypothesis. However, when m_4 is discovered to be present, MANIFS and SCOPE are augmented appropriately. A new FOCUS is developed, representing the intersection of causes(m_4) with the single set in the only pre-existing generator in FOCUS. Note that the new generator $\{d_1\ d_2\}$ in the FOCUS that results from this intersection operation represents precisely all explanations for the augmented MANIFS. This new FOCUS also illustrates another important point. In general, as information about each possible manifestation becomes available, the FOCUS changes incrementally with a monotonic decrease in the number of explanations it represents.

When m_5 is noted to be present, MANIFS and SCOPE are again adjusted appropriately. However, in this case the intersection of causes(m_5) with the single generator set in the FOCUS is empty (none of the previous explanations represented by the old FOCUS can now cover all known manifestations). The occurrence of an empty FOCUS like this again triggers a restructuring of the FOCUS: a procedure is called that produces a new set of generators from the now augmented MANIFS and SCOPE. These new generators are based on the assumption that the cardinality of any new explanation now contained in the FOCUS must be exactly one greater than the cardinality of its old explanations. Thus, when m_5 is found to be present, the new generators represent explanations consisting of two diseases.

Since m_6 is absent, the final solution to the problem is given by these same two generators (last line in Figure 4). Note that these two generators implicitly represent the eight explanations for M^+ that were listed earlier. Had this been a larger knowledge base with additional manifestations, the FOCUS would have continued to evolve using similar but more complex set intersection operations.

PROBLEM DECOMPOSITION

Since the problem of finding a minimal set cover is known to be NP-complete [4], and since the set cover problem can clearly be reduced to a diagnostic problem in polynomial time, it follows that the task of constructing the solution for a diagnostic problem is potentially combinatorially expensive. Thus an important question to address is when, within the context of the model presented here, a diagnostic problem can be reduced or decomposed into smaller independent subproblems.

One example of when this can be done is best presented by introducing the concept of "connected" manifestations. Two manifestations m_a and m_b are said to be connected if either causes(m_a) and causes(m_b) have a non-empty intersection, or there exists a finite set of manifestations $\{m_1, m_2, \ldots, m_n\}$ such that $m_1 = m_a$, $m_n = m_b$, and each m_j is connected to m_{j+1}. All of the manifestations appearing in Figure 3, for example, are connected to one another. It can be shown that if M^+ can be partitioned into N subsets of connected manifestations, each subset of which contains no manifestation connected to another manifestation in a different subset, then the original diagnostic problem can be partitioned into N independent subproblems. The generators for the solution to the original problem are then easily constructed by appending in an appropriate fashion the generators for the solutions to the subproblems.

Furthermore, sequentially constructing and maintaining independent subproblems in this way, each with its own SCOPE, FOCUS and MANIFS, is relatively easy. When a new manifestation m_i is found to be present, the set causes(m_i) is intersected with the SCOPE of each pre-existing subproblem. When this intersection is non-empty,

m_i is said to be related to the corresponding subproblem. There are three possible results of identifying the subproblems to which m_i is related. First, m_i may not be related to any pre-existing subproblems. In this case, a new subproblem is created, with MANIFS = $\{m_i\}$, SCOPE = causes(m_i), and FOCUS = a single generator consisting of the single set of competing disorders found in causes(m_i). This is what always occurs when the first manifestation becomes known, as was illustrated in Figure 4. Second, m_i may be related to exactly one subproblem, in which case m_i is assimilated into that subproblem as described earlier and illustrated with m_4 and m_5 in Figure 4. Finally, m_i may be related to multiple existing subproblems. In this situation, these subproblems are "joined" together to form a new subproblem, and m_i is then assimilated into this new subproblem (not illustrated in Figure 4).

APPLYING THE MODEL

In the interest of clarity, we have so far ignored several aspects of real-world diagnostic problem solving in presenting the set cover model. Rather than examining all of these details here, we will briefly present an implemented DSS which is based on our model. Our example implementation of the set cover model is a DSS designed to aid physicians with the diagnosis of dizziness in adults. Dizziness is a difficult diagnostic problem because of the numerous possible causes that exist. Furthermore, a patient's dizziness might be due to multiple disorders, and the DSS therefore serves to illustrate how the minimal cover approach functions in such situations. The dizziness DSS illustrated here was built using KMS, a domain-independent programming environment for constructing and evaluating DSS's [11].

The knowledge base for the dizziness DSS consists of a DESCRIPTION for each possible cause of dizziness (at present the system knows of 50 different causes). Figure 5a illustrates a DESCRIPTION for basilar migraine as it is actually encoded in the knowledge base. The letters inside angular brackets are "symbolic probabilities," rough estimates of how frequently an event occurs. Their interpretation is: A = always, H = high likelihood, M = moderate likelihood, L = low likelihood, and N = never (M is the default). Thus the "<L>" immediately after the name BASILAR MIGRAINE indicates that in general, this disorder is an uncommon cause of dizziness.

The bulk of the DESCRIPTION in Figure 5a consists of a series of statements. With the exception of the first statement (AGE = . . .), these statements represent a list of all the manifestations that can be caused by basilar migraine. For example, the third statement

```
HEAD PAIN = PRESENT <A>
  [LOCATION = OCCIPITAL <H>, RE$T <L>]
```

indicates that when present, basilar migraine always causes headache which is usually located in

---

a)
```
BASILAR MIGRAINE <L>
  [DESCRIPTION:
AGE = FROM 20 THROUGH 30 <H>, 30
      THRU 50 <L>, 50 THRU 110 <N>;
DIZZINESS = PRESENT
  [TYPE = VERTIGO <H>, RE$T <L>;
  COURSE = EPISODIC
    [EPISODE DURATION = MINUTES <L>,
      HOURS <H>, DAYS <L> ],
    ACUTE AND PERSISTENT];
HEAD PAIN = PRESENT <A>
    [LOCATION = OCCIPITAL <H>, RE$T <L>];
NEUROLOGICAL SYMPTOMS =
  TINNITUS <M>,
  DIPLOPIA
    [DURATION=TRANSIENT DURING DIZZINESS <A>],
                . . .
  SYNCOPE;
NEUROLOGICAL EXAM FINDINGS =
  HOMONYMOUS FIELD CUT
    [DURATION = TRANSIENT DURING DIZZINESS],
                . . .
  CNS FINDINGS
    [TYPE = NON-SPECIFIC <H>, RE$T <L>;
    DURATION = TRANSIENT DURING DIZZINESS]]
```

b) "Basilar migraine is an uncommon disorder that usually occurs in individuals from 20 to 30 years old, but may occur up to age 50. If a person is over 50, basilar migraine can be categorically discarded as a possible etiological factor. Basilar migraine causes dizziness which is usually of a vertiginous nature and occurs either in an episodic or an acute, persistent fashion. When episodic, the dizziness usually lasts for hours but may last for minutes or days. Headache, usually in an occipital location, is always present. Neurological symptoms caused by basilar migraine are . . ., and examination findings include . . .".

Figure 5: a) A DESCRIPTION of a cause d_i of dizziness as it is symbolically written by a physician; b) natural language form of the same information.

---

the occipital (back) part of the head (the symbol REST here indicates "any other part of the head"). A complete natural language translation for the DESCRIPTION of basilar migraine is given in Figure 5b.

The key point here is that the DESCRIPTION of each disorder d_i in the knowledge base specifies, among other things, all possible manifestations $m_1,\ldots,m_n$ of that disorder. It therefore provides man(d_i) as defined earlier in the set cover model. Since man(d_i) is explicitly encoded in the symbolic knowledge base for each and every d_i, the corresponding sets causes(m_i) are implicitly specified. Thus all the information needed to use the set cover model is present.

The inference process used by the dizziness DSS is based on the set cover model presented earlier. After being told that dizziness is present, the system asks questions of the user to

obtain further information in a sequential fashion. An example conversation with the dizziness DSS is given in the Appendix. Questions produced by the dizziness DSS are not preprogrammed but are heuristically generated based on the current FOCUS of the system. Ranking of competing disorders is done after the final FOCUS has been constructed. This scoring uses the symbolic probabilities in the knowledge base and is context-sensitive [11].

DISCUSSION

This paper has proposed the construction and maintenance of minimal set covers ("explanations") as a general model of diagnostic reasoning and as a method for diagnostic decision support systems. The set cover model is attractive in that it directly handles multiple simultaneous disorders, it can be formalized, it is intuitively plausible, and it is justifiable in terms of past empirical studies of diagnostic reasoning (e.g., [3], [5]). To our knowledge the analogy between the classic set covering problem and general diagnostic reasoning has not previously been examined in detail, although some related work has been done (e.g., assignment of HLA specificities to antisera, see [7], [14]).

The set cover model provides a useful context in which to view past work on diagnostic DSS's. In contrast to the set cover model, most diagnostic DSS's that use hypothesize-and-test inference mechanisms or which might reasonably be considered as models of diagnostic reasoning depend heavily upon the use of production rules [e.g., 1,6,8]. These systems use a hypothesis-driven approach to guide the invocation of rules which in turn modify the hypothesis. Researchers building rule-dependent systems like this have not significantly addressed the general problem of multiple simultaneous disorders. It is therefore difficult to believe this approach will be readily adaptable to more general diagnostic problems than the limited domains to which it has been applied so far. Futhermore, a rule-based hypothesize-and-test process does not provide a convincing model of human diagnostic reasoning. Rules have long been criticized as a representation of diagnostic knowledge [e.g., 13], and their invocation to make deductions or perform actions does not capture in a general sense such concepts as the Coverage Goal and Minimality Goal.

Perhaps the diagnostic DSS closest to the set cover model is INTERNIST [9]. INTERNIST represents diagnostic knowledge in a DESCRIPTION-like fashion and does not rely on production rules to guide its hypothesize-and-test process. In contrast to the set cover model, however, it uses a heuristic scoring procedure to guide the construction and modification of its hypothesis. This process is essentially depth-first, unlike the breadth-first approach we described above. INTERNIST first tries to establish one disorder and then proceeds to establish others. This roughly corresponds to constructing and completing a single generator set in the set cover model, and then later returning to construct the additional sets for the generator. Reportedly, this depth-first approach led to less than optimal performance [10]. It is also unclear that the INTERNIST inference mechanism is guaranteed to find all possible explanations for a set of manifestations. Recent enhancements in INTERNIST-II or CADUCEUS attempt to overcome thse limitations through the use of "constrictors" to delineate the top-level structure of a problem [10]. These changes are quite distinct from the approach taken in our set cover model, but do add a breadth-first component to hypothesis construction.

The set cover model presented here is still evolving both theoretically and in terms of its evaluation in practice. Our intent is to continue the formal development of the concepts presented here, and to analyze the algorithmic solution to diagnostic problems in further detail.

REFERENCES

1. Aikins, J: Prototypes and Production Rules - A Knowledge Representation for Computer Consultations, Memo HPP-80-17, Stanford Heuristic Programming Project, 1980.
2. Ben-Bassat M, et al: Pattern-Based Interactive Diagnosis of Multiple Disorders - The MEDAS System, IEEE Trans. Pat. Anal. Machine Intell., 2, 1980, 148-160.
3. Elstein A, Shulman L, and Sprafka S: Medical Problem Solving - An Analysis of Clinical Reasoning, Harvard University Press, 1978.
4. Karp R: Reducibility Among Combinatorial Problems, in R. Miller and J. Thatcher (eds.), Complexity of Computer Computations, Plenum Press, New York, 1972, 85-103.
5. Kassirer J and Gorry G: Clinical Problem Solving - A Behavioral Analysis, Ann. Int. Med., 89, 1978, 245-255.
6. Mittal S, Chandrasekaran B, and Smith J: Overview of MDX - A System for Medical Diagnosis, Proc. Third Symposium on Computer Applications in Medical Care, IEEE, 1979, 34-46.
7. Nau D, Markowsky G, Woodbury M, and Amos D: A Mathematical Analysis of Human Leukocyte Antigen Serology, Math. Biosci., 40, 1978, 243-270.
8. Pauker S, et al: Towards the Simulation of Clinical Cognition, Am. J. Med., 60, 1976, 981-996.
9. Pople H, Myers J, and Miller R: DIALOG: A Model of Diagnostic Logic for Internal Medicine, IJCAI, 4, 1975.
10. Pople H: The Formation of Composite Hypotheses in Diagnostic Problem-Solving - An Exercise in Synthetic Reasoning, IJCAI, 5, 1977, 1030-1037.
11. Reggia J: Knowledge-Based Decision Support Systems - Development Through KMS, TR-1121, Department of Computer Science, University of Maryland, Oct. 1981.

12. Reggia J: Computer-Assisted Medical Decision Making, in Applications of Computers in Medicine, M. Schwartz, editor, IEEE Press, 1982, 198-213.
13. Reggia J: A Production Rule System for Neurological Localization, Proc. Second Ann. Symp. Comp. Applic. Med. Care, 1978, 254-260.
14. Woodbury M, Ciftan E, and Amos D: HLA Serum Screening Based on an Heuristic Solution of the Set Cover Problem, Comp. Pgm. Biomed., 9, 1979, 263-273.

*** APPENDIX ***

(All output from the dizziness DSS is in upper case, while the user's typing is in lower case and underlined. The conversation is annotated parenthetically. Periodically a "snapshot" of the system's FOCUS is displayed (enclosed in box). This is produced by the tracing mechanism of the DSS for purposes of illustration and normally would not be seen by a user. Ellipses indicate where material has been excised for brevity.)

. . .

READY FOR COMMAND:
assert dizziness.

FOR DIZZINESS = PRESENT:

TYPE:
(1) DEFINITE ROTATIONAL SENSATION
(2) SENSATION OF IMPENDING FAINT
(3) ILL-DEFINED LIGHT-HEADEDNESS
(4) LOSS OF BALANCE WITHOUT HEAD SENSATION
= ? 2 / 3.

COURSE:
(1) ACUTE OR SUBACUTE PERSISTENT DEFICIT LASTING MORE THAN ONE HOUR
(2) ONE OR MORE CLEAR-CUT EPISODES
(3) GRADUAL ONSET OR SLOW PROGRESSION WITHOUT CLEAR-CUT EPISODES
= ? 2.

(The user indicates that the patient has episodic dizziness described as faintness or light-headedness.)

. . .

READY FOR COMMAND:
assert blood pressure examination = significant orthostatic hypotension.

(The user indicates that the patient's blood pressure falls when he stands up, i.e., there is orthostatic hypotension.)

READY FOR COMMAND:
obtain diagnosis.

```
FOCUS:
 GENERATOR:
  MULTIPLE SYSTEM ATROPHY
  IDIOPATHIC ORTHOSTATIC HYPOTENSION
  SYMPATHOTONIC ORTHOSTATIC HYPOTENSION
  AUTONOMIC NEUROPATHY
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    PHENOTHIAZINES
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    ANTIDEPRESSANTS
  ORTHOSTATIC HYPOTENSION SECONDARY TO L-DOPA
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    ANTIHYPERTENSIVE MEDICATIONS
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    DIURETICS
  HYPERBRADYKINISM
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    PARKINSONISM
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    PREVIOUS SYMPATHECTOMY
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    PROLONGED RECUMBENCY
```

(The initial FOCUS for the two known manifestations, dizziness and orthostatic hypotension, consists of a single generator which in turn consists of a single set of competing diseases. Each disease represents a minimal cover for the two manifestations.)

NEURO-OTOLOGICAL SYMPTOMS ASSOCIATED WITH DIZZINESS:
(1) DIPLOPIA
(2) LOSS OF OR BLURRED VISION
(3) SCINTILLATING SCOTOMAS
(4) FOCAL SYMPTOMS REFERRABLE TO CNS
(5) SYNCOPE
(6) HEARING LOSS BY HISTORY
(7) TINNITUS
= ? 3 & 5.

(Based on the active hypothesis, the DSS generates its first hypothesis-driven question and discovers two new manifestations, syncope and scintillating scotomas.)

```
FOCUS:
 GENERATOR:
  BASILAR MIGRAINE
        &
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    PROLONGED RECUMBENCY
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    SYMPATHECTOMY
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    PARKINSONISM
  HYPERBRADYKINISM
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    DIURETICS
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    ANTIHYPERTENSIVE MEDICATIONS
  ORTHOSTATIC HYPOTENSION SECONDARY TO L-DOPA
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    ANTIDEPRESSANTS
  ORTHOSTATIC HYPOTENSION SECONDARY TO
    PHENOTHIAZINES
```

```
AUTONOMIC NEUROPATHY
SYMPATHOTONIC ORTHOSTATIC HYPOTENSION
IDIOPATHIC ORTHOSTATIC HYPOTENSION
MULTIPLE SYSTEM ATROPHY
```

(None of the previous possible explanations of the patient's manifestations can now account for the new manifestations. In particular, SCINTILLATING SCOTOMAS is not explained. A new FOCUS is therefore created consisting of a single generator representing minimal covers containing two disorders.)

. . .

```
CURRENT MEDICATIONS:
    (1)   ANTICOAGULANTS
    (2)   LARGE AMOUNTS OF QUININE
    (3)   PHENOTHIAZINES
    (4)   ANTIHYPERTENSIVE AGENTS
    . . .
    (12)  BARBITURATES
= ?  2 & 3 & 4.
```

```
ELEMENTARY DISORDERS NOW CATEGORICALLY REJECTED:
    DIZZINESS SECONDARY TO BARBITURATES
    OTOTOXICITY SECONDARY TO AMNIOGLYCOSIDES
    OTOTOXICITY SECONDARY TO SALICYLATES
    ORTHOSTATIC HYPOTENSION SECONDARY TO
     ANTIDEPRESSANTS
    ORTHOSTATIC HYPOTENSION SECONDARY TO L-DOPA
    ORTHOSTATIC HYPOTENSION SECONDARY TO DIURETICS
```

(After learning the patient's medications, the DSS is able to narrow down the number of possibilities.)

```
ABNORMALITIES ON NEUROLOGICAL EXAMINATION:
    (1) OPTIC ATROPHY
    (2) PAPILLEDEMA
    (3) HOMONYMOUS FIELD CUT
    (4) NYSTAGMUS
    (5) IMPAIRED HEARING
    (6) PERIPHERAL NEUROPATHY
    (7) PARKINSONISM
    (8) FOCAL CNS FINDINGS
= ?  3 & 4 & 5.
```

(The user indicates the presence of three additional manifestations: homonymous hemianopsia, nystagmus, and impaired hearing)

. . .

```
FOCUS:
 GENERATOR:
    BASILAR MIGRAINE
          &
    COGAN'S SYNDROME
    OTOTOXICITY SECONDARY TO QUININE
    OTOSCLEROSIS
    LABYRINTHINE FISTULA
    MENIERE'S DISEASE
          &
    ORTHOSTATIC HYPOTENSION SECONDARY TO
      PROLONGED RECUMBENCY
    ORTHOSTATIC HYPOTENSION SECONDARY TO
      PREVIOUS SYMPATHECTOMY
    ORTHOSTATIC HYPOTENSION SECONDARY TO
      PARKINSONISM
    HYPERBRADYKINISM
    ORTHOSTATIC HYPOTENSION SECONDARY TO
      ANTIHYPERTENSIVE MEDICATIONS
    ORTHOSTATIC HYPOTENSION SECONDARY TO
      PHENOTHIAZINES
    AUTONOMIC NEUROPATHY
    IDIOPATHIC ORTHOSTATIC HYPOTENSION
    MULTIPLE SYSTEM ATROPHY
```

(The new FOCUS resulting from the additional information is illustrated here. Each explanation now consists of three disorders. The previous explanations could not account for the impaired hearing or nystagmus, and these new manifestations are now assumed to be due to one of five causes: COGAN'S SYNDROME, etc. Note that this FOCUS represents 1 * 5 * 9 = 45 potential explanations in a compact fashion.)

. . .

```
READY FOR COMMAND:
display value(diagnosis).

    BASILAR MIGRAINE <A>
      &
    OTOTOXICITY SECONDARY TO QUININE <H>
    OTOSCLEROSIS <M>
    LABYRINTHINE FISTULA <L>
    MENIERE'S DISEASE <L>
      &
    ORTHOSTATIC HYPOTENSION SECONDARY TO
     ANTIHYPERTENSIVE MEDICATIONS <H>
    ORTHOSTATIC HYPOTENSION SECONDARY TO
     PHENOTHIAZINES <H>
    IDIOPATHIC ORTHOSTATIC HYPOTENSION <M>
    AUTONOMIC NEUROPATHY <M>
    MULTIPLE SYSTEM ATROPHY <L>
    ORTHOSTATIC HYPOTENSION SECONDARY TO
     PARKINSONISM <L>
```

(This final diagnosis means: "The patient has basilar migraine. In addition, the patient also probably has ototoxicity secondary to the quinine he is taking, although he could have otosclerosis or even one of the other unlikely inner ear disorders listed. Finally, the patient also has orthostatic hypotension which is probably due to his medications, but might be due to one of the other listed causes." This final diagnostic account of the patient's complex set of signs and symptoms is very plausible.)

```
READY FOR COMMAND:
    . . .
```
