Int. J. Man–Machine Studies (1983) **19**, 437–460

Diagnostic expert systems based on a set covering model

JAMES A. REGGIA,*† DANA S. NAU† AND PEARL Y. WANG§

*Departments of *Neurology, †Computer Science and §Mathematics, University of Maryland*

This paper proposes that a generalization of the set covering problem can be used as an intuitively plausible model for diagnostic problem solving. Such a model is potentially useful as a basis for expert systems in that it provides a solution to the difficult problem of multiple simultaneous disorders. We briefly introduce the theoretical model and then illustrate its application in diagnostic expert systems. Several challenging issues arise in adopting the set covering model to real-world problems, and these are also discussed along with the solutions we have adopted.

1. Introduction

A diagnostic problem can be defined to be a problem in which one is given a set of abnormal findings (manifestations) for some system, and must explain why those findings are present. Problems of this kind are very common: they include diagnosing a patient's signs and symptoms, determining why a computer program failed, deciding why an automobile will not start, finding the cause of noise in a plumbing system, localizing a fault in an electronic circuit, etc. Because of this ubiquity, developing general methods for expert systems which support the decision making of human diagnosticians is an important issue at present.

This paper introduces a new model for diagnostic expert systems based on the concept of minimal set covers. This model is of interest because it captures several intuitively plausible features of human diagnostic inference, it directly addresses the problem of multiple simultaneous causative disorders, and it provides a basis for a theory of diagnostic inference.

In the following, section 2 discusses the set covering model, and section 3 explains how the model can be adopted for use in expert diagnostic systems. Section 4 and Appendix B give examples of operational expert systems based on set covering, and section 5 presents in a more detailed fashion some of the issues involved in implementing these systems. Section 6 contains some concluding remarks.

2. The set covering model

In the set covering model the underlying knowledge for a diagnostic problem is organized as pictured in Fig. 1(a) (a table of symbols is given in Appendix A). There are two discrete finite sets which define the scope of diagnostic problems: **D**, representing all possible *disorders* d_i that can occur, and **M**, representing all possible *manifestations* m_j that may occur when one or more disorders are present. For example, in medicine, **D** might represent all known diseases (or some relevant subset of all diseases,

All correspondence to: Dr James Reggia, Department of Neurology, University of Maryland Hospital, Baltimore, Maryland 21201, U.S.A.

0020-7373/83/110437+24$03.00/0

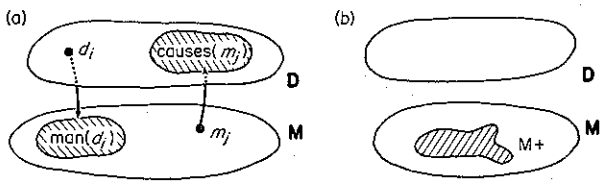

FIG. 1. Organization of diagnostic knowledge (a) and problems (b).

see below), and **M** would then represent all possible symptoms, examination findings, and abnormal laboratory results that can be caused by diseases in **D**. We will assume that $\mathbf{D} \cap \mathbf{M} = \varnothing$.

To capture the intuitive notion of causation, we assume knowledge of a relation $\mathbf{C} \subseteq \mathbf{D} \times \mathbf{M}$, where $\langle d_i, m_j \rangle \in \mathbf{C}$ represents "d_i can cause m_j". Note that $\langle d_i, m_j \rangle \in \mathbf{C}$ does not imply that m_j always occurs when d_i is present, but only that m_j may occur. For example, a patient with a heart attack may have chest pain, numbness in the left arm, loss of consciousness, or any of several other symptoms, but none of these symptoms are necessarily present.

Given **D**, **M**, and **C**, the following sets can be defined:

$$\text{man}(d_i) = \{m_j | \langle d_i, m_j \rangle \in \mathbf{C}\} \qquad \forall d_i \in \mathbf{D}, \text{ and}$$

$$\text{causes}(m_j) = \{d_i | \langle d_i, m_j \rangle \in \mathbf{C}\} \qquad \forall m_j \in \mathbf{M}.$$

These sets are depicted in Fig. 1(a), and represent all possible manifestations caused by d_i, and all possible disorders that cause m_j, respectively. These concepts are intuitively familiar to the human diagnostician. For example, medical textbooks frequently have descriptions of diseases which include, among other facts, the set man(d_i) for each disease d_i. Physicians often refer to the "differential diagnosis" of a symptom, which corresponds to the set causes(m_j). Clearly, if man(d_i) is known for every disorder d_i, or if causes(m_j) is known for every manifestation m_j, then the causal relation **C** is completely determined. We will use $\text{man}(\mathrm{D}) = \bigcup_{d_i \in \mathrm{D}} \text{man}(d_i)$ and $\text{causes}(\mathrm{M}) = \bigcup_{m_j \in \mathrm{M}} \text{causes}(m_j)$ to indicate all possible manifestations of a set of disorders D and all possible causes of any manifestation in M, respectively.

Finally, there is a distinguished set $\mathrm{M}^+ \subseteq \mathbf{M}$ which represents those manifestations which are known to be present (see Fig. 1(b)). Whereas **D**, **M**, and **C** are general knowledge about a class of diagnostic problems, M^+ represents the manifestations occurring in a specific case.

Using this terminology, we can now make the following definition.

Definition. A *diagnostic problem P* is a 4-tuple $\langle \mathbf{D}, \mathbf{M}, \mathbf{C}, \mathrm{M}^+ \rangle$ where these components are as described above.

We will assume in what follows that diagnostic problems are well-formed in the sense that man(d_i) and causes(m_j) are always non-empty sets.

Having characterized a diagnostic problem in these terms, we now turn to defining the solution to a diagnostic problem by first introducing the concept of explanation.

Definition. For any diagnostic problem P, $\mathrm{E} \subseteq \mathbf{D}$ is an *explanation* for M^+ if; (i) $\mathrm{M}^+ \subseteq \text{man}(\mathrm{E})$, or in words: E *covers* M^+; and (ii) $|\mathrm{E}| \leq |\mathrm{D}|$ for any other cover D of M^+, i.e. E is *minimal.*

This definition captures what one intuitively means by "explaining" the presence of a set of manifestations. Part (i) specifies the reasonable constraint that a set of disorders E must be able to cause all known manifestations M^+ in order to be considered an explanation for those manifestations. However, that is not enough: part (ii) specifies that E must also be one of the smallest sets to do so. Part (ii) reflects the Principle of Parsimony or Ockham's Razor: the simplest explanation is the preferable one. This principle is generally accepted as valid by human diagnosticians. Here, we have equated "simplicity" with minimal cardinality, reflecting an underlying assumption that the occurrence of one disorder d_i is independent of the occurrence of another.

With these concepts in mind, we can now define the solution to a diagnostic problem.

Definition. The *solution* to a diagnostic problem P, designated Sol(P), is the set of all explanations for M^+.

The concepts defined above are illustrated in the following example.

Example. Let $P = \langle \mathbf{D}, \mathbf{M}, \mathbf{C}, M^+ \rangle$ where $\mathbf{D} = \{d_1, d_2, \ldots, d_9\}$, $\mathbf{M} = \{m_1, \ldots, m_6\}$, and man($d_i$) and causes($m_j$) are as specified in Table 1. Note that the top (or bottom) half of Table 1 implicitly defines the relation **C**, because $\mathbf{C} = \{\langle d_i, m_j \rangle | m_j \in \text{man}(d_i)$ for some $d_i\}$. Let $M^+ = \{m_1, m_4, m_5\}$. Note that no single disorder can cover (account for) all of M^+, but that some pairs of disorders do cover M^+. For instance, if $D = \{d_1, d_7\}$ then $M^+ \subseteq \text{man}(D)$. Since there are no covers for M^+ of smaller cardinality than D, it follows that D is an explanation for M^+. Careful examination of Table 1 should convince the

TABLE 1
Knowledge about a class of diagnostic problems. The relation C is implicitly defined by either the top or bottom half of this table

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

reader that

$$\text{Sol(P)} = \{\{d_1\ d_7\}\ \{d_1\ d_8\}\ \{d_1\ d_9\}\ \{d_2\ d_7\}\ \{d_2\ d_8\}\ \{d_2\ d_9\}\ \{d_3\ d_8\}\ \{d_4\ d_8\}\}$$

is the set of all explanations for M^+.

It is of interest to compare the model of diagnostic problems presented here with the classic set covering problem. The set covering problem is typically stated along the following lines (Edwards, 1962):

> For a finite set S of elements and a family F of subsets of S, a cover K of S from F is a subfamily $K \subseteq F$ such that $\bigcup(K) = S$. A cover K is called minimum if its cardinality is as small as possible.

In this definition, S corresponds to M^+ and F corresponds to **D** in the sense that each $d_i \in \mathbf{D}$ labels a subset of M^+ (the intersection of man(d_i) with M^+). A minimum cover K corresponds roughly to the idea of an explanation E except man(E) is required only to contain M^+ rather than be equal to M^+.

3. Expert systems using the set covering model

We now turn to the description of expert systems for diagnostic problem solving based on the set covering model presented above. Such systems are organized as shown in Fig. 2 and consist of three parts.

1. A *database*, which is divided into case-specific information and general knowledge about some domain of diagnostic problems. We will use the term *knowledge base* for the latter.
2. An *inference mechanism* which is a hypothesize-and-test process that mimics diagnostic reasoning by using the set covering model.
3. A *user interface* which accepts assertions and queries from the user and translates them into internal data structures.

We now present a specific example of an expert system called System D for diagnostic problem solving. This implemented system illustrates how the set cover model can be adopted to the demands of real world problems. While System D is medically oriented, it should be remembered that the set cover model is domain-independent and not restricted to problems of medical diagnosis.

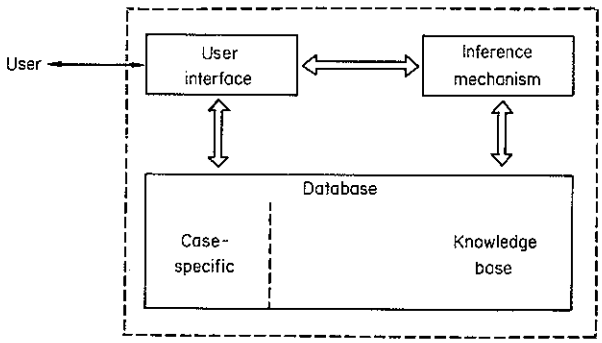

FIG. 2. Architecture of an expert system based on the set covering model.

System D is a relatively large expert system for diagnosing patients with dizziness. Dizziness is in general a very difficult diagnostic problem for the physician because there are numerous potential causes that are distributed across multiple medical specialties. Examples of possible diagnoses include:

orthostatic hypotension secondary to drugs (orthostatic hypotension is a fall in blood pressure upon standing up, and can be a side effect of certain medications);
heart disease, such as an irregular heart beat or an abnormal heart valve;
basilar migraine: headache due to painfully dilated blood vessels which supply blood to the balance centers of the brain;
inner ear diseases: these interfere with the balance mechanisms of the inner ear, and include viral labyrinthitis, Meniere's disease, and otosclerosis; and
hyperventilation: overbreathing, typically secondary to anxiety.

It is entirely possible that more than one cause of dizziness could be present simultaneously.

The knowledge base for System D is derived from numerous references and currently contains information about 50 causes of dizziness. It was built using KMS, a domain-independent software facility for constructing expert systems (Reggia, 1981). We now describe its components in detail.

THE CASE-SPECIFIC DATABASE

The case-specific database for System D contains a collection of assertions that describe a specific diagnostic problem. For example,

```
AGE=50;

DIZZINESS=PRESENT
    [TYPE=VERTIGO; COURSE=EPISODIC]; and

NEUROLOGICAL SYMPTOMS=DIPLOPIA
```

represent three assertions that might appear in the database. Each assertion is of the form

attribute relation value [elaboration],

so the three statements here mean: "A 50-year-old individual with episodic vertigo (a type of dizziness where one feels a sensation of motion) and double vision (diplopia)". During a problem solving session this case-specific information is acquired in a sequential fashion, generally in response to questions generated by the expert system. The legal attributes and their possible values are predefined in a database schema by the creator of the expert system (the "knowledge base author").

THE KNOWLEDGE BASE: REPRESENTING DIAGNOSTIC KNOWLEDGE

One of the attractive features of the set cover model is that it permits the organization of diagnostic knowledge in a form familiar to the human diagnostician. Information in the knowledge base is organized into frame-like entities called DESCRIPTIONs. Each DESCRIPTION provides a textbook-like summary of the disorder with which it is associated. An example of a DESCRIPTION from System D is illustrated in Fig. 3. To understand this descriptive knowledge representation more fully, it is necessary

```
MENIERE'S DISEASE  <L>
  [DESCRIPTION:
     AGE = FROM 20 TO 30 <L>;
     DIZZINESS = PRESENT
       [TYPE = VERTIGO;
        COURSE = ACUTE AND PERSISTENT,
        EPISODIC [EPISODE DURATION = MINUTES <L>, HOURS <H>;
                  OCCURRENCE = POSITIONAL <H>, ORTHOSTATIC <M>,
                    NON-SPECIFIC <L>]      ];
     HEAD PAIN = PRESENT <L> [PREDOMINANT LOCATION = PERIAURAL];
     NEUROLOGIC SYMPTOMS = HEARING LOSS BY HISTORY <H>, TINNITUS <H>;
     PULSE DURING DIZZINESS = MARKED TACHYCARDIA <L>;
     NEUROLOGIC SIGNS = NYSTAGMUS  [TYPE = HORIZONTAL, ROTATORY],
                        IMPAIRED HEARING <H>                        ]
```

FIG. 3. The knowledge base of System D currently consists of a set of data structures called DESCRIPTIONs such as that shown here for Meniere's Disease.

to know about three conventions being used: symbolic probabilities, separation of causal and non-causal associations, and elaboration.

Symbolic probabilities, indicated in angular brackets in Fig. 3, are subjective, non-numeric estimates of how frequently an event occurs. While exact probabilities of diagnostic associations are usually not available, a great deal of descriptive information about diagnosis exists in the form

x *frequently* causes y,
x *can* cause y,
x is *never* associated with y,
x is *commonly* associated with y,
x is *rare* (*common*, *very common*, . . .), and
x *only* occurs if y,

where x is some disorder and y is some fact about a case. Symbolic probabilities capture this coarse but useful information. The five† possible estimates we use are:

A = always,
H = high likelihood,
M = medium likelihood,
L = low likelihood, and
N = never.

Thus, the "⟨L⟩" following MENIERE'S DISEASE in Fig. 3 indicates that this disorder is relatively uncommon, and the "⟨H⟩" on the last line of the DESCRIPTION indicates that Meniere's disease often causes impaired hearing.

The second convention used in DESCRIPTIONs is the separation of causal and non-causal associations. Certain features of a disorder can be viewed as being caused by the disorder being described. For example, in medicine loss of vision, chest pain, dizziness and confusion are all abnormalities that conceptually are caused by some underlying problem. We have been using the term *manifestations* for these causally-related features. In contrast, other features of a disorder are not causally associated with it. For example, a patient's age and sex may provide very significant information about the likelihood of a certain disease being present, but they are not caused by

† Five is obviously somewhat arbitrary, but it has proven sufficient for our applications so far.

that disease. Features such as these will be referred to as *setting factors*. Which features in a knowledge base are manifestations and which are setting factors are indicated in the database schema specified by the knowledge base author (see Reggia, 1981). In the DESCRIPTION in Fig. 3, only the first assertion concerning age specifies a setting factor, while each of the other assertions specify manifestations.

Finally, elaboration provides further details about a manifestation and is indicated as part of an assertion inside of square brackets. For example, in Fig. 3 the assertion

```
DIZZINESS=PRESENT
 [TYPE=VERTIGO;
  COURSE=ACUTE AND PERSISTENT,
         EPISODIC . . .]
```

elaborates on the type of dizziness manifested by Meniere's Disease by indicating that it is vertiginous in nature and that it occurs either in an acute, persistent fashion or in episodes.

With the above conventions in mind, the DESCRIPTION in Fig. 3 should now be relatively understandable. It indicates that Meniere's disease is a relatively uncommon cause of dizziness (because of "⟨L⟩" immediately following the name of the disease). The dizziness it causes is vertiginous in nature and either acute and persistent or episodic. When episodic, the episodes usually last for hours and are especially produced by positional changes of the head. Meniere's Disease occasionally causes periaural headache, frequently causes hearing loss and tinnitus (ringing in the ears), and so forth.

What is most important here in the context of the set covering model is that the DESCRIPTION associated with any disorder d_i specifies, among other things, the set man(d_i) of all manifestations caused by d_i. Thus, a knowledge base containing a set of disorders along with all of their DESCRIPTIONs completely specifies the information needed to solve diagnostic problems as they were defined earlier. Returning to our example, the knowledge base of System D consists of a listing of 50 causes of dizziness and their DESCRIPTIONs, each similar to that illustrated in Fig. 3. The knowledge base thus explicitly specifies the set **D** of all causative disorders of dizziness as well as the set man(d_i) for each disorder. Furthermore, the set **M** is implicitly specified, as it consists of every manifestation listed in any of the DESCRIPTIONs. The relationship **C** and the sets causes(m_j) are also implicitly specified by the collective information in the DESCRIPTIONs. As explained earlier, the DESCRIPTIONs in this knowledge base contain additional information about setting factors and estimates of relevant probabilities.

THE INFERENCE MECHANISM: A SEQUENTIAL HYPOTHESIZE-AND-TEST PROCESS

In adapting the set covering model for use in a real-world expert system several issues must be addressed and resolved. Perhaps the most obvious of these issues is the fact that diagnostic problem-solving is inherently sequential in nature. Rather than knowing all of the manifestations which are present in a specific case at the start, the human diagnostician usually begins knowing that one or a few manifestations are present, and must actively seek further information about others. In medicine, for example, the physician is typically confronted with a patient complaining of some symptom (the "chief complaint"), and must uncover other manifestations through questions, examining the patient, and laboratory-testing.

Empirical studies done over the last decade have provided convincing evidence that this sequential diagnostic reasoning is guided by a hypothesize-and-test process (Elstein, Schulman & Sprafka, 1978; Kassirer & Gorry, 1978; see Reggia, 1982, for a review). Given a few initial manifestations, the human diagnostician constructs a tentative hypothesis about the cause of those manifestations. Further information is then sought for generally two reasons: either for completeness (so-called "protocol-driven" questions), or to uncover facts specifically needed to modify the evolving hypothesis (so-called "hypothesis-driven" questions). These latter questions "test" the validity of the hypothesis, possibly confirming or eliminating part of it.

This sequential diagnostic process can be captured in terms of the set covering model presented earlier. 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 disorders 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 presented as a collection of generators.

The term "generator" used here needs further definition. Rather than representing the solution to a diagnostic problem as an explicit list of all possible explanations for M^+ or MANIFS, 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 represent 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 presented earlier (Table 1). 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 human diagnostician organizes the possibilities during problem solving (i.e. the "differential diagnosis").

Using the three data structures MANIFS, SCOPE and FOCUS, a hypothesize-and-test algorithm based on the set covering model can perform diagnostic problem solving. The FOCUS represents the tentative or working hypothesis at any point during problem-solving. The algorithm, described informally, is:

(1) Get the next manifestation m_j.
(2) Retrieve causes(m_j) from the knowledge base.

(3) MANIFS ← MANIFS ∪ $\{m_j\}$.
(4) SCOPE ← SCOPE ∪ causes(m_j).
(5) Adjust FOCUS to accommodate m_j.
(6) Repeat this process until no further manifestations remain.

Thus, as each manifestation m_j that is present is discovered, MANIFS is updated simply by adding m_j to it. SCOPE is augmented to include any possible causes d_i of m_j which are not already contained in it (derived by taking the union of causes(m_j) and SCOPE). Finally, FOCUS is adjusted to accommodate m_j based on intersecting causes(m_j) with the sets of diseases in the existing generators. These latter operations are done such that any explanation which can no longer account for the augmented MANIFS (which now includes m_j) are eliminated.

The key step in this process is Step 5, the adjustment of the FOCUS or working hypothesis. Perhaps the best way to understand this step is to follow a simple example. Recall the abstract knowledge base illustrated in Table 1, and consider the same diagnostic problem $M^+ = \{m_1\ m_4\ m_5\}$ that was used earlier. The order in which information about manifestations is discovered is determined by question generation heuristics, as described later in section 5. For now, suppose that the sequence of events occurring during problem-solving were ordered as listed in Fig. 4. What happens during problem-solving is as follows.

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

FIG. 4. Sequential problem-solving using the set covering model.

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 previously there were no generators in the 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".

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 set 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. As information about each possible manifestation becomes available, the FOCUS changes incrementally with a monotonic decrease in the number of explanations it represents (with the exception of situations where the FOCUS becomes empty).

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 fact that the cardinality of any 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 Fig. 4). Note that these two generators implicitly represent the eight explanations for M^+ that were listed earlier. It is also interesting to note that d_3 and d_4, eliminated from the FOCUS when m_4 was found to be present, are once again viable possibilities. Had this been a larger knowledge base with additional manifestations, the FOCUS would have continued to evolve using similar set intersection operations.

4. An example application of 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 immediately examining these details, we will first briefly illustrate an interactive session with System D.

Expert System D must handle situations where multiple causes of dizziness are present simultaneously. The following conversation illustrates just such a situation. User typing is underlined and ellipsis indicates where material has been excised for brevity. The system's tracing mechanism has been turned on producing the normally unseen output highlighted in boxes which represents periodic "snapshots" of the FOCUS.

.
.
.

```
DIZZINESS DIAGNOSIS DECISION SUPPORT SYSTEM
TYPE 'CONTINUE' TO BEGIN

READY FOR COMMAND:
assert dizziness.
```

(Rather than waiting to be asked questions, the user elects to volunteer some initial information.)

```
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. The answer "2/3" means "2 *or* 3".)

```
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.
```

(The user activates System D's inference mechanism, which is based on the set covering process explained above.)

```
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, System D 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 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. The "&" here represents the "×" used in generators as described in section 3.)

```
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, System D is able to narrow down the number of possibilities. How this is done is described in the next section.)

```
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>

READY FOR COMMAND:
```

This final diagnosis offered by System D, including a ranking of competing alternatives which will be explained below, 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.

5. From model to functioning expert system

As noted earlier, the implementation of functioning expert systems like System D based on the set covering model requires that several issues be addressed and resolved. We have already discussed adopting the model to sequential problem solving, so we now turn to several other aspects of real-world diagnostic problem-solving. Further details about these issues can be found in the references (Reggia, 1981).

QUESTION GENERATION AND TERMINATION CRITERIA

The vast majority of questions generated by Expert System D, representing Step 1 in the informal sequential algorithm presented earlier, fall into the category of hypothesis-driven questions. In other words, each question is based solely on the disorders in the FOCUS at that point during problem-solving. Let us say that a disorder

is *active* if it is currently in the FOCUS. Then to select its next question the expert system extracts from the DESCRIPTION of each active disorder the first attribute in an assertion whose current value is not yet known (recall that assertions, attributes and values were defined in section 3, "Case-Specific Database"). From these candidate attributes, the one in the largest number of DESCRIPTIONs of active disorders is selected to form the basis of the next question.

This simple, heuristic approach to question generation makes no claim to optimality. However, it does have certain properties that make it a useful strategy to follow. Since it selects one of the most commonly referred to attributes of active disorders, it usually produces questions that help to discriminate among the competing explanations in the FOCUS. In addition, since it selects candidate questions from the *first* unknown attributes remaining in these DESCRIPTIONs, it allows the knowledge base author to exert partial control over the order in which questions are generated (i.e. by consistently ordering the assertions in DESCRIPTIONs in a similar fashion). Finally, this approach to question generation has the advantage of being computationally inexpensive when compared with more elaborate optimization schemes that might be used.

Once a new question has been asked and answered by the user, another hypothesize-and-test cycle begins. This continues until no further questions can be generated because no assertions in the DESCRIPTIONs of active disorders contain attributes whose values have not been acquired from the user. This termination condition is a somewhat arbitrary approach to deciding when sufficient information is known. While it asks about all attributes relevant to ranking the competing explanations involved at termination time, it might leave some information unsought. To permit the knowledge base author to insure the level of completeness of information collection that is desired from an expert system, protocol-driven questions that should always be asked may optionally be included as explicit instructions to an expert system at the time it is constructed.

SETTING FACTORS AND THE RANKING OF COMPETING DISORDERS

Once the termination condition is satisfied, expert systems like System D enter a final scoring phase during which competing disorders are ranked relative to one another for the first time. In other words, the hypothesize-and-test control cycles have previously only been concerned with the construction of all possible explanations (a differential diagnosis) without regard to their relative likelihood. For each active disorder at termination time two numeric scores are calculated: a *setting score* and a *match score*. These scores are calculated using the symbolic probabilities in the knowledge base as well as any symbolic probabilities incorporated in a user's response to questions. A simple weighting scheme ($A = 4$, $H = 3, \ldots, N = 0$) is used for these calculations.

The setting score for an active disorder is initialized to the numerical equivalent of its symbolic probability originally specified following its name in the knowledge base. This initial score is then incrementally adjusted upwards or downwards based only on assertions about setting factors in its DESCRIPTION. The setting score is intended to provide a generalization of the concept of prior probability in that it reflects the general likelihood of a disorder in the context of the specific setting in which it is occurring.

The match score of an active disorder is based only on M^+ and the assertions about manifestations in its DESCRIPTION. The match score is also derived using a simple weighting scheme. At termination time, an expert system conceptually has in the FOCUS all of the possible competing explanations for M^+. It can therefore derive a match score for any disorder based on the "best" explanation which contains that disorder (i.e. the explanation that as a whole would be most likely to cause M^+). For example, if d_1 is in two explanations, then the match score for d_1 would be based on its role in the "better" of these two possible explanations. Furthermore, a manifestation can be assigned to the disorder in an explanation which is most likely to be producing it in situations where that manifestation can be caused by more than one of the disorders in the explanation. The match score is intended as a measure of how closely a disorder fits the manifestations of a case, irrespective of the setting in which they are occurring.

A final score is calculated for each active disorder based on both its setting score and its match score. Since this final numerical weighting is intended to provide only a "ballpark" indication of how likely the disorder is, it is subsequently converted back into a symbolic probability to emphasize its imprecise and heuristic nature. This was illustrated in the conversation with System D when the final diagnostic possibilities were listed in order of likelihood.

It should be appreciated that the set covering model used in this fashion permits scoring which can be considered to be truly context-dependent. Not only does a disorder's likelihood depend on the specific environment in which it is occurring (setting score), but it also depends on what other disorders are postulated to be simultaneously present in an explanation, and on which of several competing explanations contain it (match score).

One final point about the use of symbolic probabilities needs to be made. While the ranking of competing disorders is done *after* the termination condition is satisfied, the symbolic probabilities A and N are used in one other way *during* the hypothesize-and-test control cycles. They are used to determine when any disorder d_i should be *categorically rejected* by the inference mechanism. For example, the DESCRIPTION of ORTHOSTATIC HYPOTENSION SECONDARY TO L-DOPA in System D's knowledge base contains the categorical assertion

CURRENT MEDICATIONS=L-DOPA <A>.

Thus, if System D discovered that a patient was not taking L-DOPA, ORTHOSTATIC HYPOTENSION SECONDARY TO L-DOPA would be immediately discarded from any further consideration by the inference mechanism (as occurred after the question on CURRENT MEDICATIONS in the conversation with System D earlier). In effect, what occurs is that the set D is changed: the set of all possible disorders is modified by removing any disorders discovered to be categorically rejected during problem-solving. All subsequent development of the SCOPE and FOCUS by the inference mechanism reflects this change in the very framework of the problem.

PROBLEM DECOMPOSITION

Since finding a minimal set cover is known to be NP complete (Karp, 1972), the task of constructing the solution to a diagnostic problem is potentially combinatorally expensive as the size of an explanation increases. This difficulty is only academic for

some classes of diagnostic problems. For example, it is not uncommon for a patient seen by a physician to have more than one disease simultaneously, but it would be exceedingly rare for someone to have more than 50 diseases simultaneously. However, since the potential for combinatorial explosion exists, it is still important to address the question of when 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 Table 1, 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 (Reggia, 1981).

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 Fig. 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 Fig. 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 Fig. 4 nor in the conversation with System D, both of which involved only a single subproblem).

OTHER CONSIDERATIONS

Many other considerations go into expanding the generality and robustness of the set covering model for use in real world expert systems. We will mention just three of these issues here: unexplainable manifestations, the single-disorder constraint, and non-independent disorders.

Assuming that an expert system's knowledge base and a relevant case are both correct and complete, the set cover model as described above can handle a broad range of diagnostic problems. Unfortunately, in the real world, this ideal situation is sometimes not present. A knowledge base might be incomplete or contain errors, especially during system development, and a user might enter incorrect information about a problem.

One example of such a situation is the *unexplainable manifestation*: a manifestation m_j whose associated set causes(m_j) is empty. If undetected, such a manifestation would

result in repeated futile attempts by the inference mechanism to create progressively larger and larger explanations to account for all known manifestations. The important point here is that the inference mechanism must continuously monitor for unexplainable manifestations at run time. This is because an initially non-empty set causes(m_i) could potentially become empty during problem-solving if all of the disorders in it were discovered to be categorically rejected. Our expert systems currently handle this anomaly by informing the user of the situation, discarding the unexplainable manifestation, and offering the user the option of continuing with the understanding that all is not well.

Another issue deserving special attention is situations where only one disorder is expected to occur at a time. Even though such a *single-disorder constraint* may not be strictly correct in a theoretical sense, there are situations where such an assumption is justified by practical considerations. An example of an expert system called System P which uses the set covering model with the single disorder constraint is given in Appendix B. System P uses this constraint because of the exceedingly low likelihood that two of the individually very rare disorders in its knowledge base would occur in a single individual. The advantage of using the single-disorder constraint when appropriate is that it permits the automatic recognition by the inference mechanism of potential errors. This is illustrated in the conversation with System P in Appendix B when it indicates that no single disorder can account for all of the facts in the case under consideration. Such a situation might have been due to (i) user error in describing the case, (ii) an incomplete or incorrect knowledge base, or (iii) a patient with a previously unknown form of peroneal muscular atrophy.

When the single-disorder constraint is employed, two adjustments are made to the inference mechanism of expert systems using the set covering model. First, at the start of a case the FOCUS is initialized to a single generator whose single set includes all possible disorders within the domain of the expert system. This initial FOCUS represents the initial hypothesis that exactly one possible disorder is present. Second, the inference mechanism as usual monitors for the occurrence of an empty FOCUS, but interprets such an occurrence as an anomaly. It does not try to construct explanations containing two disorders, but indicates to the user that it cannot explain the current case findings with a single disorder (see conversation with System P).

Finally, we have assumed so far that the disorders in D are independent of one another, an assumption that may not be valid in some domains. One possible approach to this non-independence would be to award a "bonus" during scoring to explanations where associated disorders were involved [this was the approach used in INTERNIST; see Pople, Myers & Miller (1975)]. We have elected to study instead those situations where disorders can be partitioned into classes, with disorders in one class causing disorders in another. For example, one expert system currently being constructed involves both localization of damage in the nervous system and diagnosis.

6. Discussion

This paper has proposed the construction and maintenance of minimal set covers ("explanations") as a general model of diagnostic reasoning and has illustrated its use as an inference method for diagnostic expert systems. The set cover model is attractive in that it directly handles multiple simultaneous disorders, it can be formal-

ized, it is intuitively plausible, and it is justifiable in terms of past empirical studies of diagnostic reasoning (e.g. Elstein *et al.*, 1978; Kassirer & Gorry, 1978). 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 [for example, assignment of HLA specificities to antisera; Nau, Markowsky, Woodbury & Amos (1978) and Woodbury, Ciftan & Amos (1979)].

The set cover model provides a useful context in which to view past work on diagnostic expert systems. In contrast to the set cover model, most diagnostic expert systems that use hypothesize-and-test inference mechanisms or which might reasonably be considered as models of human diagnostic reasoning depend heavily upon the use of production rules (e.g. Aikins, 1979: Mittal, Chandrasekaran & Smith, 1979; Pauker, Gorry, Kassirer & Schwartz, 1976). These systems use a hypothesis-driven approach to guide the invocation of rules which in turn modify the hypothesis. A rule-based hypothesize-and-test process does not provide a convincing model of what has been learned about human diagnostic reasoning in the empirical studies cited earlier. Furthermore, rules have long been criticized as a representation of diagnostic knowledge (e.g. Reggia, 1978), and their invocation to make deductions or perform actions does not capture in a general sense such intuitively attractive concepts as coverage, minimality, or explanation.

Perhaps the previous diagnostic expert system whose inference method is closest to the set cover model is INTERNIST (Pople *et al.*, 1975). INTERNIST is a large and well-known expert system that 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, INTERNIST's inference mechanism uses a heuristic scoring procedure to guide the construction and modification of its hypothesis. This process is essentially serial or depth-first, unlike the more parallel or breadth-first approach implied in the set cover model. In other words, 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. The criteria used by INTERNIST to group together competing disorders (i.e. a set in a generator) is based on a simple heuristic: "Two diseases are competitors if the items not explained by one disease are a subset of the items not explained by the other; otherwise, they are alternates (and may possibly coexist in the patient)" (Miller, Pople & Myers, 1982). In the terms of our model, this corresponds to stating that d_1 and d_2 are competitors if $M^+ - \text{man}(d_1)$ contains or is contained in $M^+ - \text{man}(d_2)$. It can be proven that while this simple heuristic may generally work in constructing a differential diagnosis, there are clearly situations for which it will fail to correctly group competing disorders together.† Reportedly, the serial or depth-first approach used in INTERNIST resulted in less than optimal performance (Pople, 1977; Miller *et al.*, 1982), and it has been criticized as "*ad hoc*" by some individuals working in statistical pattern classification because of the lack of a formal

† For example, suppose $M^+ = \{m_1 \cdots m_8\}$ and only d_1, d_2, and d_3 have been evoked where $M^+ \cap \text{man}(d_1) = \{m_2\ m_4\ m_5\ m_6\ m_7\ m_8\}$, $M^+ \cap \text{man}(d_2) = \{m_3\ m_4\ m_5\ m_6\ m_7\ m_8\}$, and $M^+ \cap \text{man}(d_3) = \{m_1\ m_2\ m_3\}$. In the set cover model, $\text{Sol(P)} = \{\{d_1\ d_3\}\ \{d_2\ d_3\}\}$ which can be represented by the single generator $\{d_1\ d_2\} \times \{d_3\}$ where d_1 and d_2 are grouped together as competitors. Suppose that d_1 was ranked highest by the INTERNIST heuristic scoring procedure. Then $M^+ - \text{man}(d_1) = \{m_1\ m_3\}$ and $M^+ - \text{man}(d_2) = \{m_1\ m_2\}$, so INTERNIST would apparently fail to group d_1 and d_2 together as competitors.

underlying model (e.g. Ben-Bessat *et al.*, 1980). 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's successor CADUCEUS attempt to overcome some of these limitations through the use of "constrictors" to delineate the top-level structure of a problem (Pople, 1977). These changes are quite distinct from the approach taken in the 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. Work is clearly needed in at least three directions: further theoretical development of the model, assessment of its application in expert systems involving a broad range of real-world diagnostic problems, and assessment of its adequacy as a cognitive model. We intend to pursue these issues in the future.

This research was supported by NIH grants 5 K07 NS 00348 and 1 P01 NS 16332 and in part by NSF grant MCS81-17391. Computer time was provided by the Computer Science Center of the University of Maryland.

References

AIKINS, J. (1979). Prototypes and production rules: an approach to knowledge representation for hypothesis formation. *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, Tokyo, Japan, pp. 1–3.

BEN-BASSAT, M., CARLSON, R., PURI, V., DAVENPORT, M., SCHRIVER, J., LATIF, M., SMITH, R., PORTIGAL, L., LIPNICK, E. & WEIL, M. (1980). Pattern-based interactive diagnosis of multiple disorders: the MEDAS system. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2**, 148–160.

EDWARDS, J. (1962). Covers and packings in a family of sets. *Bulletin of the American Mathematics Society*, **68**, 494–499.

ELSTEIN, A., SHULMAN, L. & SPRAFKA, S. (1978). *Medical Problem Solving—An Analysis of Clinical Reasoning*. Cambridge, Massachusetts: Harvard University Press.

KARP, R. (1972). Reducibility among combinatorial problems. In MILLER, R. & THATCHER, J., Eds, *Complexity of Computer Computations*, pp. 85–103. New York: Plenum Press.

KASSIRER, J. & GORRY, G. (1978). Clinical problem solving: a behavioral analysis. *Annals of Internal Medicine*, **89**, 245–255.

MILLER, R., POPLE, H. & MYERS, J. (1982). INTERNIST-1, An experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine*, **307**, 468–476.

MITTAL, S., CHANDRASEKARAN, B. & SMITH, J. (1979). Overview of MDX: a system for medical diagnosis. *Proceedings of the Third Annual Symposium on Computer Applications in Medical Care*, pp. 34–46. Piscataway, New Jersey: IEEE Press.

NAU, D., MARKOWSKY, G., WOODBURY, M. & AMOS, D. (1978). A mathematical analysis of human leukocyte antigen serology. *Mathematical Biosciences*, **40**, 243–270.

PAUKER, S., GORRY, G., KASSIRER, J. & SCHWARTZ, W. (1976). Towards the simulation of clinical cognition. *American Journal of Medicine*, **60**, 981–996.

POPLE, H. (1977). The formation of composite hypotheses in diagnostic problem solving: an exercise in synthetic reasoning. *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Carnegie–Mellon University, Pittsburgh, Pennsylvania, pp. 1030–1037.

POPLE, H., MYERS, J. & MILLER, R. (1975). DIALOG: a model of diagnostic logic for internal medicine. *Proceedings of the Fourth International Joint Conference on Artificial Intelligence*, pp. 848–855. Cambridge, Massachusetts: M.I.T. Artificial Intelligence Laboratory Publications.

REGGIA, J. (1978). A production rule system for neurological localization. *Proceedings of the Second Annual Symposium on Computer Applications in Medical Care*, pp. 254–260. Piscataway, New Jersey: IEEE Press.
REGGIA, J. (1981). Knowledge-based decision support systems: development through KMS. *TR-1121*, Department of Computer Science, University of Maryland, College Park.
REGGIA, J. (1982). Computer-assisted medical decision making. In SCHWARTZ, M., Ed., *Applications of Computers in Medicine*, pp. 198–213. Piscataway, New Jersey: IEEE Press.
WOODBURY, M., CIFTAN, E. & AMOS, D. (1979). HLA serum screening based on an heuristic solution of the set cover problem. *Computer Programs in Biomedicine*, **9**, 263–273.

Appendix A: Table of symbols

Symbol	*Meaning*
$\in$	element of
$\subseteq$	subset of
$\leq$	less than or equal to
$\varnothing$	empty set
$\forall$	for all
$\cup$	set union
$\cap$	set intersection
$\lvert\ \rvert$	set cardinality
M	universe of manifestations (effects)
D	universe of disorders (causes)
C	causative relationship

Appendix B

Expert System P addresses the problem of diagnosing the cause of wasting of the muscles of the lower legs (peroneal muscular atrophy). The causes of this problem are very rare, and include entities such as hypertrophic neuropathy of Charcot–Marie–Tooth and Refsum's Syndrome. Because of the rarity of each of these disorders, it is extremely unlikely that more than one would occur in an individual, so System P uses the set covering model with the additional constraint that only one possible cause can be present in any case. This leads to the following interesting result (user typing is underlined in the conversation, and the system's tracing mechanism is turned on producing the output highlighted in boxes which represent normally unseen periodic "snapshots" of the FOCUS):

```
.
.
.
PERONEAL MUSCULAR ATROPHY DECISION SUPPORT SYSTEM FOR
DIAGNOSTIC CLASSIFICATION. TYPE 'CONTINUE' TO BEGIN.

READY FOR COMMAND:
continue
OK–PLEASE ANSWER THE FOLLOWING . . .
```

```
AGE OF ONSET:
 (1) INFANCY
 (2) CHILDHOOD OR ADOLESCENCE
 (3) ADULT LIFE
=?2.
```

(The manner in which the system selects questions to ask was addressed in section 5.)

```
FOCUS:
GENERATOR:
  HEREDITARY SPASTIC PARAPLEGIA WITH PERONEAL MUSCULAR
   ATROPHY
  FRIEDRICH'S ATAXIA
  REFSUM'S SYNDROME
  PROGRESSIVE SPINAL MUSCULAR ATROPHY OF CHARCOT-MARIE-
   TOOTH
  NEURONAL TYPE OF CHARCOT-MARIE-TOOTH DISEASE
  NON-DOMINANT HYPERTROPHIC NEUROPATHY OF CHARCOT-MARIE-
   TOOTH
  HYPERTROPHIC NEUROPATHY OF CHARCOT-MARIE-TOOTH
```

(Knowing that exactly one cause of peroneal muscular atrophy is present, System P starts off with an initial hypothesis consisting of a single generator containing the eight possibilities. One cause is immediately discarded because it always starts in infancy, leaving the seven potential explanations listed here.)

```
INHERITANCE:
 (1) SPORADIC
 (2) RECESSIVE
 (3) DOMINANT
=?unknown.
SEVERITY OF MOTOR DEFICIT:
 (1) MILD
 (2) MODERATE
 (3) SEVERE
=?3.
```

```
FOCUS:
 GENERATOR:
  FRIEDRICH'S ATAXIA
  REFSUM'S SYNDROME
```

(The three possible causes of a severe motor deficit are FRIEDRICH'S ATAXIA, REFSUM'S SYNDROME, and HYPERTROPHIC NEUROPATHY OF DEJERINE-SOTTAS. Intersecting this set with the previous generator decreases the alternative explanations to the two listed here.)

```
SENSORY DEFICIT:
 (1) PRESENT
 (2) ABSENT
=?1.
FOR SENSORY DEFICIT=
  PRESENT . . .
  SEVERITY:
  (1) MILD
  (2) MODERATE
  (3) SEVERE
=?2.
```

(When told that the patient has a sensory deficit, System P automatically asks for relevant elaboration about its severity. This request for elaboration is the only example of a protocol-driven question in this conversation. The discovery of a sensory deficit does not change the FOCUS, shown immediately below, because this manifestation can be caused by both active disorders.)

```
FOCUS:
 GENERATOR:
  FRIEDRICH'S ATAXIA
  REFSUM'S SYNDROME
```

```
PES CAVUS:
 (1) PRESENT
 (2) ABSENT
=?1.
```

```
FOCUS:
 GENERATOR:
  FRIEDRICH'S ATAXIA
```

(A third manifestation, PES CAVUS or "high arches", is uncovered, and intersecting its set of causes with the FOCUS results in a single remaining possible explanation. Rather than accepting this hypothesis as the diagnosis, System P continues to generate questions to confirm it.)

```
OTHER NEUROLOGICAL FINDINGS:
 (1) RETINITIS PIGMENTOSA
 (2) DEAFNESS
 (3) NYSTAGMUS
 (4) DYSARTHRIA
 (5) CEREBELLAR ATAXIA
 (6) PYRAMIDAL SIGNS
=? (MULTIPLE ANSWERS PERMITTED)
1.
```

```
A SINGLE VALUE OF
TYPE OF PERONEAL MUSCULAR ATROPHY CANNOT EXPLAIN
ALL OF THE FEATURES OF THIS CASE
SHOULD PROCESSING CONTINUE USING THOSE VALUES THAT
ARE NOT CATEGORICALLY REJECTED
=?(YES/NO)
no.

TYPE OF PERONEAL MUSCULAR ATROPHY=
 UNKNOWN

READY FOR COMMAND: . . . .
```

(When System P learns that RETINITIS PIGMENTOSA is present, intersection of its causes with the generator results in an empty FOCUS. System P gives up and classifies this patient as having an unknown disease.)

What is striking here is that System P automatically detects that it does not know the diagnosis in this case. This is because of the special constraint that only a single disorder be present, imposed by the creator of System P, contradicts the definition of adequacy required of an explanation in the context of this specific case.