7

A Model-Based Method for Computer-Aided Medical Decision Making

Sholom M. Weiss, Casimir A. Kulikowski, Saul Amarel, and Aran Safir

While MYCIN and PIP were under development at Stanford and Tufts/M.I.T., a group of computer scientists at Rutgers University was developing a system to aid in the evaluation and treatment of patients with glaucoma. The group was led by Professor Casimir Kulikowski, a researcher with extensive background in mathematical and pattern-recognition approaches to computer-based medical decision making (Nordyke et al., 1971), working within the Rutgers Research Resource on Computers in Biomedicine headed by Professor Saul Amarel. Working collaboratively with Dr. Arin Safir, Professor of Ophthalmology, who was then based at the Mt. Sinai School of Medicine in New York City, Kulikowski and Sholom Weiss (a graduate student at Rutgers who went on to become a research scientist there) developed a method of computer-assisted medical decision making that was based on causal-associational network (CASNET) models of disease. Although the work was inspired by the glaucoma domain, the approach had general features that were later refined in the development of the EXPERT system-building tool (see Chapters 18 and 20).

A CASNET model consists of three main components: observations of a patient, pathophysiological states, and disease classifications. As observations are recorded, they are associated with the appropriate intermediate states. These states, in turn, are typically causally related, thereby forming a network that summarizes the mechanisms of disease. It is these patterns of states in the network that are linked to individual disease classes. Strat-

From *Artificial Intelligence*, 11: 145–172 (1978).

egies of specific treatment selection are guided as much by the individual pattern of observations and diagnostic conclusions as they are by the disease classification itself.

Unlike mathematical models of disease processes, a CASNET model is inherently symbolic and focuses on causality and temporal sequences of events. Although not all medical topics are well understood at this level, CASNET demonstrated that there are areas of medicine in which explicit model representations permit powerful reasoning strategies that go beyond simple matching of treatments with diseases. It is this ability to match treatment plans with the patient's current stage in the progression of a disease process and with expectations of future events that set CASNET apart from the other early AIM systems. More recently ABEL (Chapter 14) and VM (Chapter 10) have extensively studied similar issues, and Pople has discussed at length the need to incorporate causal reasoning and a sense of temporal progression into future versions of INTERNIST (Pople, 1982).

7.1 Introduction

In the present paper, a general approach to structuring medical knowledge for computer-aided diagnosis and therapy is presented. We have developed a representation that models disease processes as a causal-associational network (CASNET). This model-based method has been used successfully in designing a consultation program for the diagnosis and long-term treatment of the glaucomas. The consultation program uses a set of general decision-making strategies in conjunction with a class of causal-associational models (Kulikowski and Weiss, 1971; Weiss, 1974). In this paper, examples will be given from a CASNET model of glaucoma. However, the model representation and decision-making procedures are generalizable to other medical domains.

Diagnostic problems have often been cast into a pattern-recognition or statistical decision-theory framework. Computer representation is not difficult, and as a result many well-known methods such as those based on Bayes' Theorem have been used (Brodman et al., 1959; Warner et al., 1964; Gorry and Barnett, 1968a). The difficulties with applying these methods (such as scarcity of statistics and the use of invalid approximations) are also sufficiently persistent that alternative approaches have been sought. In many medical areas, existing knowledge could enhance the decision-making capabilities of a diagnostic system. There are many useful decision rules specific to a given medical application that the physician directly applies in his or her reasoning.

In the past few years, there has been increased interest in the application of artificial intelligence (AI) techniques to medical decision making.

AI techniques attempt to capture decision-making rules explicitly, while statistical methods may extract them implicitly from accumulated sample experience. The AI approaches intend to overcome some of the limitations of purely statistical methods by developing a more structured representation of the diagnostic and therapy selection problems. A program that uses decision strategies based on explicit representations of medical knowledge can more easily incorporate evolving changes in its knowledge base, independently of the reasoning strategies. It can also incorporate the results of clinical experience by matching the more explicit patterns of reasoning to the decisions and opinions of physicians. Such systems are more likely to be accepted because they are expressed in a decision-making context familiar to the clinician. A structured representation can also permit the formulation of complex hypotheses that express progression and severity of disease. Some researchers have attempted to increase the scope, accuracy, and explanation capabilities of their systems by increasing structure, while still preserving a statistical framework (Patrick et al., 1974). Others have relied on logical and semantic encodings of contextual knowledge within an artificial intelligence framework (Pople et al., 1975; Shortliffe et al., 1973; Wortman, 1972) (see also Chapter 6).

Several fundamental AI issues are raised by medical decision-making problems. One important issue concerns the development of representations that are powerful enough to capture a complex and changing knowledge base in a realistic task domain. There has been an increased interest in recent years in developing AI systems that use expert knowledge in a variety of application areas (Buchanan et al., 1969; Duda et al., 1977; Reddy, 1977; Sridharan and Schmidt, 1977). Methods of acquiring knowledge from experts, the choice of appropriate levels of abstraction and resolution for describing a given problem, and the choice of computer representation of the knowledge base are all problems that immediately arise in developing such systems. They are closely linked to the control strategies or methods used to produce interpretations for individual cases. Fundamental to most such control strategies is the capability of approximate reasoning. This is needed to manage the multiple hypotheses that can be generated from a large and complex knowledge base, which includes statements at different levels of uncertainty. Once decisions are reached, producing explanations becomes an important task if the acceptability of the system is to be enhanced. Practical issues of implementation for these large knowledge-based systems include ease of knowledge management (updating), efficiency, choice of languages, and transferability into practical use in both the original domain and other similar ones.

The present paper describes the methods of representation and interpretation developed while building a knowledge-based system for medical consultation. In the course of describing these methods, specific solutions to some of the issues raised above are offered.

7.2 Causal-Associational Network (CASNET) Models

A causal-associational network is a particular type of semantic network (Woods, 1975) designed to:

a. describe dynamic processes in terms of (loop-free) causal relationships among a set of internal variables;
b. relate this description to external variables that are considered to be manifestations of the internal processes; and
c. describe various classifications imposed on the dynamic processes.

CASNET models can be used to describe many different complex processes, but we have developed them to describe pathophysiological processes of disease (Weiss, 1974). Knowledge, in our scheme, is represented by three types of data elements, corresponding to the three kinds of description outlined above: observations of the patient; pathophysiological states; and diagnostic, prognostic, and therapeutic categories. Observations are the direct evidence obtained about a patient. Pathophysiological states are intermediate constructs that describe internal conditions assumed to take place in the patient; they summarize results from many different observations. Categories of disease are conceptually at the highest level of abstraction, summarizing patterns of states and observations. In Figure 7-1 we summarize this three-level description of disease processes. Considerations of all three levels enter into the recommendation of therapy. Bonner et al. (1964) developed a single-level model with causal and associational relations intermingled. When diagnosis is to be modeled in a domain of knowledge where mechanisms of disease are understood, the cause-and-effect model can be used to significantly improve the basis on which decisions are made. When, however, less information is available, associations between findings must be relied on to a greater extent, and the goals of reaching structured and well-explained conclusions and recommendations may not be fully satisfied.

7.2.1 Causal Network of States

In our model of disease, the pathogenesis and mechanisms of a disease process are described in terms of cause-and-effect relationships between pathophysiological states. States are summary descriptions of events that are deviations from normality. Strict causality (Bunge, 1963) is not as-

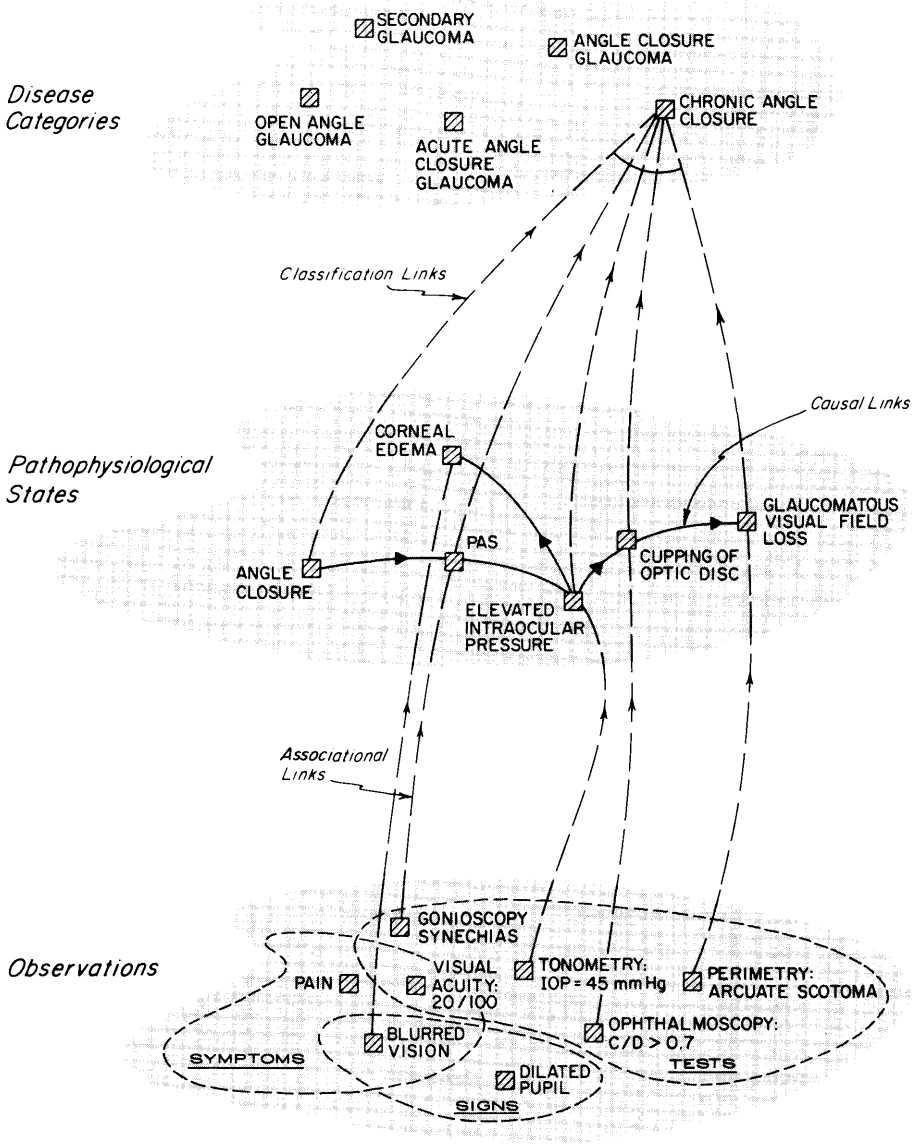

FIGURE 7-1 Three-level description of a disease process.

sumed—there may be multiple causes and effects, and in a given patient, a cause may be present without any of its effects occurring at the same time. Various effects can follow from a given cause, each produced with a different strength of causation. Examples of states would be "increased

intraocular pressure" or "glaucomatous visual field loss." Many such states may occur simultaneously in any disease process. A state thus defined may be viewed as a set of values of a state variable as used in control systems theory. It does not correspond to one of the mutually exclusive states that could be used to describe a probabilistic system. This definition was chosen to correspond to the basic entities physicians use when they describe disease mechanisms. A somewhat simplified graph model of glaucoma is illustrated in Figure 7-2, where each node, n_i, is a pathophysiological state, and each edge is a causal connection. Disease processes may be characterized by pathways through the network. A complete pathway from a starting to a terminal node usually represents a complete disease process, while partial pathways, from starting to nonterminal nodes, represent various degrees of evolution within the disease process. Progression along a causal pathway is usually associated with increasing seriousness of the disease. For example, in Figure 7-2 a complete pathway is traversed from n_{35} (PRIMARY OPEN ANGLE MECHANISM) to n_{31} (GLAUCOMATOUS VISUAL FIELD LOSS): (n_{35} n_{25} n_{26} n_{27} n_{28} n_{29} n_{30} n_{31}). A partial pathway is traversed from n_{35} (PRIMARY OPEN ANGLE MECHANISM) to n_{26} (ELEVATED INTRAOCULAR PRESSURE): (n_{35} n_{25} n_{26}).

When a set of cause-and-effect relationships between states is specified, the resulting structure is a network, or directed acyclic graph of states. The state network is defined by (S, F, N, X), where S is the set of starting states, those states with no antecedent causes; F is the set of final states, those states with no effects; N is the total set of states; X is the set of mappings between states indicating causal relationships.

The mappings are of the form

$$n_i \xrightarrow{a_{ij}} n_j$$

where a_{ij} is the strength of causation (interpreted in terms of frequency of occurrence) and n_i and n_j are states. This rule is interpreted as follows: state n_i causes state n_j, independently of other events, with frequency a_{ij}. Starting states are also assigned a frequency measure indicating a prior or starting frequency. The strengths of causation are represented by numerical values, fractions between 0 and 1 that correspond to qualitative ranges such as sometimes, often, usually, or always.

States are summary statements. Many events and many complex relationships may be summarized by a single state. For example, in Figure 7-2, "neural tissue loss and cupping of the nerve head" is a summary of a much more complex situation. If a higher-resolution description is desired, several different types of nerve loss and cupping could be specified. The resolution of states should be maintained at a level consistent with the objective of efficient decision making. A state network can be thought of as a streamlined model of disease that unifies several important concepts and guides us in our goal of diagnosis. It is not meant to be a complete model of disease.

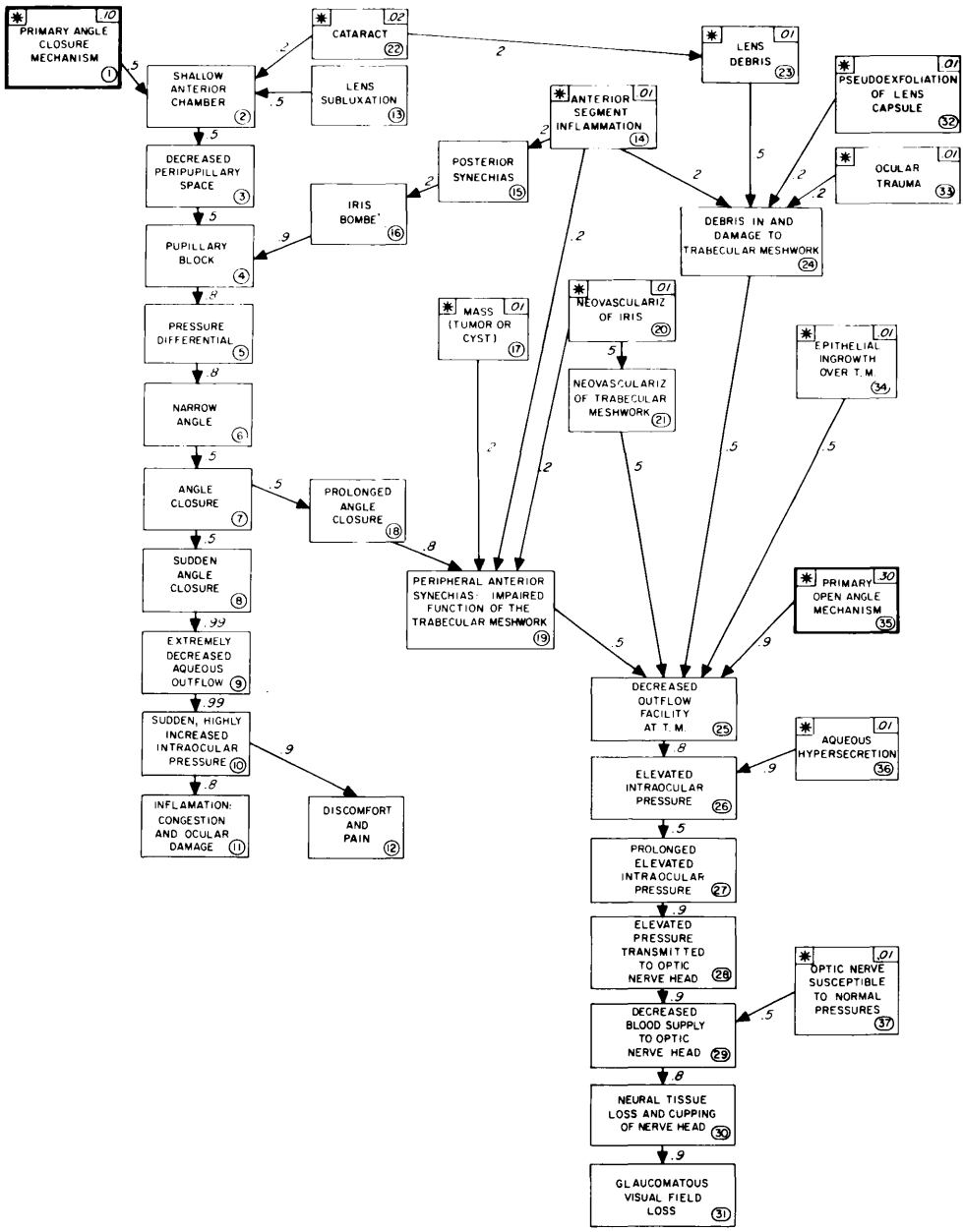

FIGURE 7-2 Partial causal network for glaucoma. States with no antecedent causes are indicated by asterisks (*). The circled numbers correspond to the state labels (n_i) used in examples in the text.

7.2.2 Rules for Associating States with Observations

Observations (tests)—the history, signs, symptoms, and laboratory tests—are the form in which information about a patient is presented. These clinical features, however, must be unified into some coherent framework for explanation and diagnosis. Observations about a patient are used to confirm or deny certain states in the network that describe the disease process. A single state may be associated with many observations. These states can then be related by causal pathways that explain the mechanisms of disease in a patient. The relationship between tests and states is noncausal; it is associational. For a given observation, confidence measures are used to indicate a degree of belief in the presence of specific states.

The rules for associating tests with states are represented as

$$t_i \xrightarrow{Q_{ij}} n_j$$

where t_i is the ith observation (or Boolean combination of observations), n_j is the jth node, and Q_{ij} $(-1 \leq Q_{ij} \leq +1)$ is the confidence in n_j given that t_i is observed to be true. Positive values of Q_{ij} correspond to an increased confidence in n_j, and negative values correspond to a decreased confidence in n_j when t_i is observed. Associated with each observation are costs $C(t_i)$ that reflect the cost of obtaining the result t_i.

Example 1. Two different instruments may be used to measure intraocular pressure (tension): a Schiotz tonometer and an applanation tonometer. A high Schiotz tension reading may indicate an elevated intraocular pressure with a confidence of 0.5. A high applanation tension reading, which is usually more reliable, may be assigned a higher confidence, such as 0.7. If by ophthalmoscopy it is further demonstrated that the appearance of the optic disc indicates damage to the optic nerve (with a confidence of 0.3), these results may be combined and assigned a confidence of 0.8 that the pressure has been and is truly elevated. Figure 7-3 illustrates these relationships, with circular nodes standing for states and square nodes for observations. The number on the link that connects a test to a state is the confidence with which a test supports a state.

7.2.3 Rules for Associating Disease Categories with States

Diagnostic and prognostic categories of disease are defined in terms of ordered patterns of rules, which we refer to as classification tables. The tables contain rules of the form

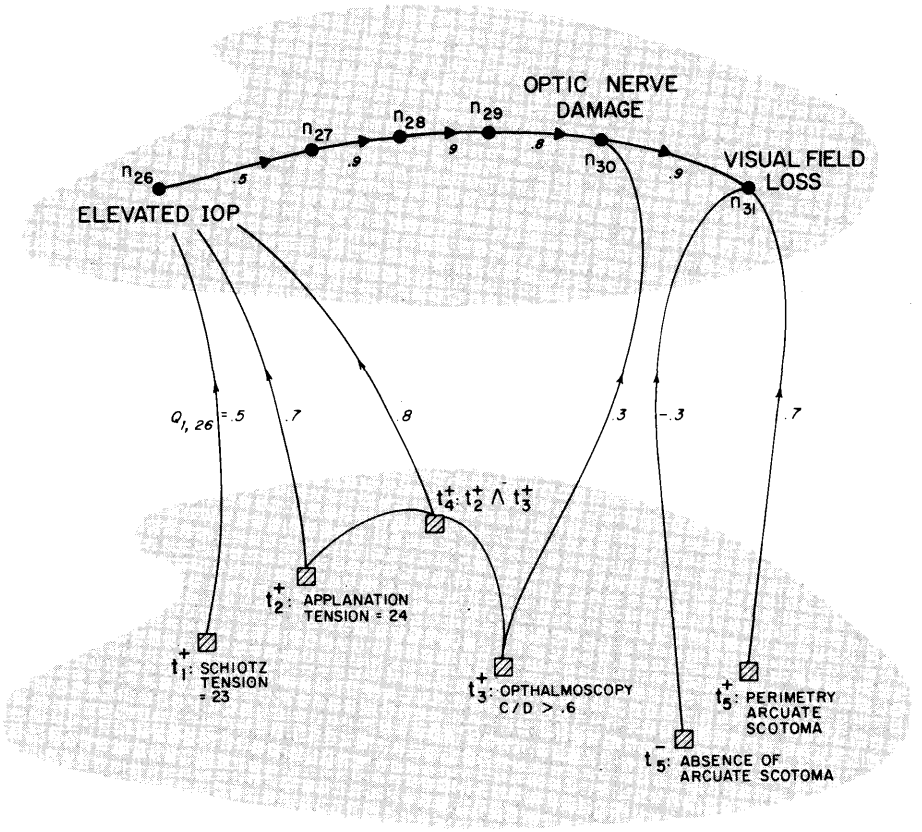

FIGURE 7-3 Examples of rules for associating states with observations.

$$
\begin{array}{llllll}
n_1 & \wedge & \neg n_2 & & \rightarrow & D_1 \\
n_1 & \wedge & n_2 & \wedge \quad \neg n_3 & \rightarrow & D_2 \\
 & & & \vdots & & \\
n_1 & \wedge & n_2 & \wedge \ldots \quad \neg n_i & \rightarrow & D_{i-1} \\
n_1 & \wedge & n_2 & \wedge \ldots \quad n_i & \rightarrow & D_i
\end{array}
$$

where D_i is the ith diagnostic and prognostic category, which is implied by the given ordered pattern of states, $n_1 \wedge n_2 \ldots \wedge n_i$. For this chosen form and ordering of rules, the tables can be referred to by using an abbreviated notation of ordered pairs:

$$(n_1, D_1), (n_2, D_2), \ldots (n_{i-1}, D_{i-1}), (n_i, D_i)$$

The classification tables can be augmented to include therapy recommendations. These tables are ordered triples of the form

$$(n_1, D_1, T_1), (n_2, D_2, T_2), \ldots (n_{i-1}, D_{i-1}, T_{i-1}), (n_i, D_i, T_i)$$

where T_i are treatments (or treatment plans) for patients falling into particular diagnostic categories.

In the following sections, clinical decision making will be considered as a problem of using a CASNET model for (a) selecting and interpreting observations, (b) analyzing and resolving conflicts and contradictions in the observations, (c) selecting diagnostic and prognostic categories, and (d) recommending treatments.

7.3 Test Result Interpretation

A test result has the following form: observation t_i is true, false, or uncertain. Based on a given result for t_i, a measure of confidence, Q_{ij}, may be assigned to state n_j. More than one test may confirm or deny a single state with varying degrees of confidence. The total confidence in the presence or absence of a state is derived from all local mappings from tests to states occurring for a given patient. Each node, n_j, in the state network is assigned a measure, $\mathrm{Cf}(n_j)$. Initially, the Cf of all nodes is undetermined; i.e., $\mathrm{Cf}(n_j) = 0$.

Rule 1. When a test result is received and a rule $t_i \overset{Q_{ij}}{\rightarrow} n_j$ is found applicable, the $\mathrm{Cf}(n_j)$ is affected as follows:

a. If $|\mathrm{Cf}(n_j)| < |Q_{ij}|$, then $\mathrm{Cf}(n_j)$ is reset to Q_{ij}.

b. If $Cf(n_j) = -Q_{ij}$, then $Cf(n_j)$ is set to 0 (and the conflict is noted) until another result t_k is received such that $|Q_{kj}| > |Q_{ij}|$.

c. Otherwise, $Cf(n_j)$ is unchanged.

Thus, of all the test results that are evidence for a given state, we choose the result in which we have the greatest confidence. When a new test result is received with a confidence measure equal but counter to the previously accumulated evidence, the conflict is noted, and the status of the node is reset to be undetermined.

A Cf measure is used to evaluate whether the status of a node is assumed to be confirmed, denied, or undetermined. Let Θ be a nonnegative integer that serves as a threshold fixed in advance for a specific model. (The threshold for test selection may be fixed at a level different from that used for classification.)

Rule 2.

a. If $Cf(n_j) > \Theta$, then n_j is assumed confirmed.

b. If $Cf(n_j) < -\Theta$, then n_j is assumed denied.

c. Otherwise, the status of n_j is assumed undetermined.

In this way, the designer of a model can assign confidence to the test results. Whenever the status of a node n_j exceeds (or is less than) a uniform and consistent threshold, node n_j is assumed confirmed (or denied). At some point there is enough confidence in these findings to draw at least tentative conclusions about some specific aspects of the disease, which are summarized in the states. These conclusions can change when other test results, in which we have greater confidence, are received.

The initial state network graph is a static structure. However, based on a series of observations, a configuration, or labeled subnet, of the state network can be generated that is applicable to a given patient. For a given patient, a configuration of the state network is described by assigning each node either a confirmed, denied, or undetermined status. The state network dynamically evolves into different configurations, each determined by the interpretation of the test results. Tentative diagnostic conclusions and decisions can be reached for each configuration of the state network.

7.4 Strategies for Test Selection

A configuration of the state network can be used not only to reach conclusions, but also to select questions. An interactive sequential questioning procedure that is guided by the results of previously asked questions can usually reduce the number of questions that must be asked, often elimi-

nating irrelevant and redundant questions. Asking the right questions in an intelligent order is an important aspect of the diagnostic process.

The strategies for test selection that have been developed for CASNET-type models can be categorized as those emphasizing (a) local logical constraints among questions, (b) categories of causal pathways, and (c) likelihood measures over the states.

These strategies are not mutually exclusive, and all three may be combined into a single overall strategy. The simplest strategy, yet perhaps the most effective for a well-circumscribed domain of application, is the strategy that emphasizes local logical constraints among the questions. For this strategy, questions on related topics are organized into small local tree structures. Each group of questions is asked only when a fixed set of logical conditions is satisfied.

The second strategy depends on isolating the causal pathways that potentially explain the observations that have already been recorded. The strategy would then pursue observations that are related to the states found on these pathways. The identification of pathways that may explain the current observations and related processes of disease is discussed in Section 7.5.

A likelihood strategy for the CASNET model is based on the assignment of weights to each of the nodes in the state network. Tests that can produce results having greater measures of confidence than are currently held for the states are considered possible candidates for further testing. Of these tests, the one that relates to the highest-weighted node is selected.

A number of characteristics of the state network are important for the specification of inference strategies:

a. No loops may exist in the network because all transitions between nodes are unidirectional under the assumption of causal production.

b. Starting nodes have no antecedent causes (or predecessors in the network) and represent events taken as the starting events in the causal chains. These nodes are assigned (prior) weights, a_i, based on their relative frequency of occurrence.

c. Each transition weight has a maximum value of unity. The sum of transition weights leaving node n_i is not necessarily unity, because the successors of n_i are not necessarily mutually exclusive. In addition, the model incorporates only consequences of events that are of interest to the process being described, leaving unspecified any other possible outcomes.

d. The transition weight, a_{ij}, in a link $n_i \xrightarrow{a_{ij}} n_j$ *is assigned on the assumption that* n_j is caused by n_i with frequency a_{ij}, independently of the way in which n_j was entered from other nodes of the network. For a given model, a consistent interpretation must be given to the transition weights throughout the network.

During a procedure of sequential test selection, a given node of the network can be in one of three status conditions: confirmed, denied, or undetermined. Initially, all nodes are undetermined, but as tests are selected and results obtained, some of the nodes will be confirmed and others denied.

At every stage of the selection procedure, each node in the network is assigned a weight that is determined by the current configuration of confirmed and denied nodes of the network. The derivation of these weights is given below. The weights serve as an index for the selection of further nodes to be queried. In a model of disease, where the nodes represent states of the disease process, the weights are used to choose the sequence of states to be tested. The assignment of weights in a causal network has a superficial similarity to a Markov chain (Gheorghe et al., 1976). The important differences are found in the lack of mutual exclusivity between successors of a node and in the assumption of causal production between successor nodes. A model could be designed as a Markov network, but it would require the specification of a much larger number of nodes and transitions for the many possible combinations of events that can occur in a complex process.

7.4.1 Forward Weights for Test Selection

An *admissible pathway* is said to exist from node n_i to node n_j when none of the intermediate nodes in the pathway are denied. For the remainder of the discussion on the calculation of weights, a reference to a pathway refers to an admissible pathway. Also, it is assumed that successive nodes on a pathway are numbered consecutively.

The weight of entering node n_j from a single admissible pathway starting at node n_i is defined as the product of the transition weights between all pairs of successive nodes (n_k, n_{k+1}) in the pathway:

$$w_F(j|i) = \prod_{k=1}^{j-1} a_{k,k+1}$$

The total forward weight of node n_j is computed as the sum of the weights $w_F(j|i)$ for those admissible pathways entering n_j, starting at the nearest confirmed or starting nodes, n_i, of the network. A nearest confirmed node within a pathway is a node such that there are no other confirmed nodes in the pathway between it and n_j. In the case that n_i is an unconfirmed starting node, the weight of this pathway is multiplied by the starting weight.

$$\text{Let } \mu_i = a_i \text{ when } n_i \text{ is an unconfirmed starting state,}$$
$$= 1 \text{ otherwise.}$$

The total weight of n_j is then

$$w_F(j) = \sum_i w_F(j|i) \cdot \mu_i$$
$$= a_j \text{ when } n_j \text{ is a starting node}$$

where i ranges over the set of nearest confirmed or starting nodes.

Example 2. (See Figure 7-2.) Assume that n_7 is confirmed, n_{17} and n_{20} are denied, and the other nodes are of undetermined status. The forward weight of $n_{19} = (0.5)(0.8) + (0.01)(0.20) = 0.402$. The weight is calculated from pathways beginning at n_7 (the nearest confirmed node) and at n_{14} (the only undenied starting node that leads to n_{19}).

The rationale for choosing the product of successive transition weights in a pathway lies in the assumption that each transition weight, a_{ij}, is independent of all preceding transitions. If n_{j-1} is a confirmed node, the weight $w_F(j-1|i) = 1$, so all previous transition weights within the pathway need not be computed. Hence the weight of the end node of a pathway is calculated only from the nearest confirmed or starting node.

When there is only one admissible pathway leading to a node, ignoring the possibility of overlapping causal events introduces no error in the computation of weights. If the network is defined with mutual exclusivity between all successors of any given node, the problem can be operationally treated as a Markov chain calculation. That is, all confirmed nodes do in reality lie on a unique pathway. In our representation, when overlap between pathways does occur, some nodes in the network may be given a greater weight by the above computations than would result from exact frequency assignments over disjointly defined pathways. Yet, because a pathway to a node, n_j, represents the manner in which n_j is produced, this greater weighting is acceptable and even helpful. The tendency toward overweighting is related to the number of pathways that lead to n_j and the strength of transitions between the nodes in these pathways. But it is precisely those nodes that have many possible ways of occurring and that have strong causal and frequency connections that are the most likely to occur for the patient. Since a product of fractions not greater than unity is employed for the computation of weights, weights computed on the basis of few observations will result in relatively small weight assignments to the nodes. For some nodes, this weight assignment may be an accurate measure of frequency. Even without exact frequency assignments, the manner in which confirmation or denial of nodes is included in the weight calculation can be quite effective in guiding the selection of tests. The confirmation of a given node, n_i, will usually greatly increase the weight of all of its effects. Many fewer fractional multiplications will be used in computing the necessary pathways from n_i, since for the successors of n_i the weight of n_i can be assumed to be unity. Similarly, the weights resulting from a

denied node, n_i, will decrease, since n_i cannot lie on any admissible pathway to another node. This is precisely the response needed to guide the search for information to topics suggested by the accumulated evidence.

In general, the overall effect of forward weight calculation is to increase the weights of those nodes resulting from confirmed nodes while decreasing the weights of those from denied nodes.

7.4.2 Inverse Weights for Test Selection

The forward weight of a node, n_i, summarizes the weight of evidence carried from the causes of n_i. The weight of n_i can also take into account the confirmed nodes that are effects of n_i. For this we must define some inverse weight of confirmed effect n_j on the cause n_i. In analogy to Bayes' formula for inverse probability, an inverse weight can be defined as

$$w_I(i|j) = [w_F(j|i) \cdot w_F(i)]/w_F(j)$$

where confirmed nodes are ignored in the pathways (and forward weights are, therefore, computed from starting states). Because an admissible pathway cannot contain a denied node, an inverse weight is proportional to the weight of pathways passing through n_j that also pass through n_i divided by the weight of all currently possible pathways to n_j.

Example 3. (See Figure 7-2.) Assume that all pathways are denied, except those beginning with n_{35} and n_{37}. Let n_{31} be confirmed and the remaining nodes undetermined. The inverse weights for n_{35} and n_{37} are calculated as follows:

$$w_F(35) = 0.30,\ w_F(37) = 0.01 \text{ (starting weights)}$$

$$\begin{aligned} w_F(31) &= (0.3)(0.9)(0.8)(.05)(0.9)(0.9)(0.8)(0.9) + (0.01)(0.5)(0.8)(0.9) \\ &= 0.067 \end{aligned}$$

$$w_F(31|35) = (0.9)(0.8)(0.5)(0.9)(0.9)(0.8)(0.9) = 0.210$$

$$w_F(31|37) = (0.5)(0.8)(0.9) = 0.360$$

$$\begin{aligned} w_I(35|31) &= w_F(35) \cdot w_F(31|35)/w_F(31) \\ &= (0.30)(0.210)/0.067 = 0.940 \end{aligned}$$

$$\begin{aligned} w_I(37|31) &= w_F(37) \cdot w_F(31|37)/w_F(31) \\ &= (0.01)(0.36)/0.067 = 0.054 \end{aligned}$$

Since several effects may follow from a single cause, it is desirable to choose some function of all the inverse weights to represent the overall

inverse weight of a node, n_i. A reasonable choice is the maximum of the inverse weights for each n_i:

$$w_I(i) = \max_j \{w_i(i|j)\}$$

This function was selected because we are searching the network for strong evidence that n_i is present. For the important situations where nodes lie on a single pathway to a confirmed node or nodes lie on every pathway to a confirmed node, the inverse weight for those nodes will be correctly assigned as unity. Where there are two or more mutually exclusive pathways to a confirmed node, the inverse weight remains a relatively accurate frequency measure. However, the pathways to a confirmed node need not be mutually exclusive. Therefore, the maximum of the inverse weights is used as an overall measure of inverse weight. The inverse weight for a confirmed node assigns weight on a fractional basis to the node's potential causes, even when more than one cause is strongly indicated. The maximum weight compensates for the lack of mutual exclusivity by considering evidence other than a single confirmed node. Since several confirmed nodes may in fact be unrelated, an average or sum would appear to be less effective. The maximum is effective because it preserves the weight of a strong piece of confirmatory evidence without dilution from other nonexclusive causes, because it recognizes the possible multiplicity of confirmed effects from a single cause, and because it generally provides a reasonable basis of comparison with the forward weights.

The calculation of inverse weights is strongly influenced by evidence for the confirmation or denial of nodes. The weight of a node may be increased when its effects are confirmed. Initially, a pathway may be an unlikely alternative, but after some testing it may become the only feasible pathway to a particular confirmed node. This results in increased weight assignments to the remaining causes of the confirmed node.

7.4.3 Overall Weight for Test Selection

In order to choose a node for testing, a single function of the forward and inverse weights of the node is needed as an overall measure. The maximum of these two weights has been chosen:

$$W_i = \max \{w_F(i), w_I(i)\}$$

This choice reflects the need to have a measure of strong confirmatory evidence for the potential presence of a node, n_i. Evidence of the denial of n_i is included in both $w_F(i)$ and $w_I(i)$. These forward and inverse weights represent the contribution from different parts of the network toward the likelihood of confirmation of n_i. The maximum is thus a measure of strong

confirmatory evidence toward n_i throughout the network. It should usually provide good testing candidates. Relatively efficient algorithms can be specified for the computation of weights (Weiss, 1974). These algorithms take advantage of the acyclic nature of the state network so that the states may be topologically sorted.

7.4.4 Test Selection with Cost Assignment

A weight is a measure of likelihood, based on the evidence gathered for the possible causes of a node. The weight does not take into account the cost of performing a test that may confirm or deny the node. Let t_i be a test for node n_j and C_i be the cost of t_i. W_j is the currently assigned (nonzero) weight of node n_j.

Two cost strategies have been used for test selection:

a. *Maximum weight-to-cost ratio:* select t_i such that $W_j/C_i = \max_{m,n}(W_m/C_n)$.

b. *Maximum weight within a certain range of costs:* select t_i such that $W_j = \max_m(W_m)$ for all t_n with $C_n < C$.

A strategy of maximum weight selection is a special case of strategy a when the costs are equal or are ignored. A minimum cost strategy is a special case of strategy b where C_n is taken as the minimum cost (for the remaining tests) and W_j is any nonzero weight.

The stopping rule for the likelihood strategy consists of terminating test selection when no weight exceeds a fixed threshold. For an in-depth consultation in an application such as glaucoma, where all topics must be covered thoroughly, all questions are asked that have not been logically excluded by prior responses. This corresponds to setting the threshold to zero.

A form of hypothesis-driven test selection has also been formulated. A hypothesis corresponds to a class of likely causal pathways that explain the patient's observations and related but as yet unknown findings. The strategy then selects tests that support the hypotheses. The following section describes methods of identifying the likely hypotheses (classes of pathways) for a patient.

7.5 Interpretation of Disease Processes Within the State Network

The state network is a general structure that implicitly contains large numbers of both complete and partial causal state pathways, representing processes of disease. Several general classes of pathways can be described that

are useful for decision making and explanation. These classes of pathways are characterized by (a) their starting nodes, and (b) their terminal nodes.

Starting nodes or states are those states in the network for which no causes have been defined. The starting nodes are explicitly determined by the structure of the state network; the complete set of possible starting states is independent of any configuration of confirmed states. In Figure 7-2, n_{14} is a starting state and n_{19} is not; n_{19} will never be a starting state, even when all of its causes (n_{18}, n_{17}, n_{14}, n_{20}) are denied. Within the model, a starting state is the most antecedent cause of further progression of disease in a patient. It represents a basic causal mechanism that characterizes a disease process. Any causal pathway that explains the disease process involved in a particular patient can be characterized by its starting state. When a nonstarting state has all of its antecedent clauses denied, this state will not appear on any pathway that attempts to explain the manifestation of disease in a patient. The nonstarting states represent events that should be explained by the events that cause them.

The clinician is usually most concerned with the most likely causes of disease found in a patient. The most likely starting node is taken as the node that explains the greatest number of states of disease. This is the starting state from which pathways (containing no denied nodes) are generated that traverse the greatest number of confirmed nodes. If two or more starting states are found, a likelihood measure is computed for the states, and the starting node with the greatest weight is selected. If a single starting state does not explain all of the confirmed nodes, then another starting state is found that explains the greatest number of remaining states. The procedure is continued until all of the confirmed nodes are explained, and the complete set of most likely starting nodes is identified. The pathways generated from these nodes represent the most likely explanations of the disease processes manifested in the patient.

The physician may also wish to discover alternative though less likely causes that potentially explain the disease mechanisms present in a patient. Potential explanations of the disease processes for a patient can be found by generating all pathways that reach confirmed states, without traversing any denied states. In addition, since a state network is usually designed for a restricted domain of diseases, the clinician may wish to determine those causes of disease that have not yet been eliminated. These may be observed by generating all undenied pathways in the state network.

Observations of a patient are often gathered sequentially. History questions are asked before the physical examination, which precedes the laboratory tests. For a given configuration of the state network, pathways may be generated that, by necessity, are based on an incomplete set of observations. For a specific patient the physician is often interested in determining those disease processes that have not yet been ruled out and may be uncovered by additional observations. Pathways that explain disease processes for a specific patient are usually terminated at a confirmed node. This provides the direct explanation of the events that have been observed. By continuing causal pathways beyond this usual termination point of a con-

firmed node and extending them to include all nodes with an undetermined status, those aspects of the disease process that remain possible can be indicated.

Many diseases are (irreversibly) progressive. Once the particular processes are determined, the physician is concerned with identifying the stages of the disease to which the patient may subsequently proceed. The particular pathways that have been generated to explain observations for a patient may be continued to the terminal nodes of the state network, even if they traverse currently denied nodes. These pathways will give an indication of possible future events, and provide the basis for prognostic assessment.

Example 4. (See Figure 7-2.) Assume that nodes n_{15} and n_{19} are confirmed and the remaining nodes are undetermined; n_{14} will be selected as the most likely mechanism, because it explains both n_{15} and n_{19}. The pathways emanating from n_{14} ($n_{14}\,n_{15}\,n_{16}\,n_4\,n_5\,n_6\,n_7\,n_{18}\,n_{19}$ and $n_{14}\,n_{19}$) directly explain the current observations of the patient. However, for future examinations, more observations may be recorded, and one will probably be interested in continuing the pathways, to check for elevations of intraocular pressure (n_{10} or n_{26}). There are also other mechanisms that are less likely, but that may potentially explain n_{19} (e.g., n_{20}).

7.6 Conflicts and Contradictions

The diagnostician is sometimes faced with the task of interpreting test results that are seemingly conflicting and in some cases contradictory. It is possible to recognize and resolve many conflicts and contradictions because the test results for a patient are interpreted through a model of disease that expresses the meaning of these observations. The model may be viewed as containing an implicit set of consistency conditions that must be satisfied for each patient.

The procedures for interpreting test results have been designed to resolve explicit conflicts in these results. As described earlier, the test result that is held with greatest confidence is taken as the accepted result. If conflicting results are received with equal confidence, then the conflict is noted, and the status of the state of disease remains undetermined until additional results, with greater confidence, resolve the conflict.

A typical contradictory situation occurs when a state is confirmed, yet all of its potential causes in the network are denied. For example, in Figure 7-2, a contradiction would result if n_{19} is confirmed, and n_{18}, n_{17}, n_{14}, and n_{20} are all denied. There is not an admissible pathway to confirmed node n_{19}, because all of the pathways contain a denied node. One potential explanation for this difficulty is that the model of disease may be incom-

plete and some causes (of confirmed node n_{19}) are missing from the network. For example, although it is not indicated in Figure 7-2, n_{33} (OCULAR TRAUMA) may in fact cause n_{19} (PERIPHERAL ANTERIOR SYNECHIAS). The model designer may intentionally not specify all potential causes; instead, he or she may indicate that for some nodes no contradiction should be assumed because the model of causes for these nodes is incomplete. Either the model is incomplete or a contradiction has been found.

Based on a configuration of confirmed and denied nodes in the state network, pathways of disease are generated to explain the processes of disease found in a given patient. Some of the nodes in these pathways may have an undetermined status, with $|\mathrm{Cf}(n_i)| < \Theta$. When the Cf of a node generated in a pathway is undetermined but in the direction of denial, i.e., $-\Theta < \mathrm{Cf}(n_i) < 0$, then the explanation of disease is inconsistent. The explanation provided by the model may be valid, but it indicates that further, more conclusive evidence is needed. If any inconsistencies are found in these pathways, it is important to check for any alternative explanation that, while not the most likely, is entirely consistent with the states that are explained. This can be accomplished by changing the threshold Θ to zero and then finding the most likely starting node. Now all nodes that have been tested with any degree of confidence will be assumed either confirmed or denied. Either the same most likely starting nodes will be selected or alternative mechanisms will be found.

Example 5. (See Figure 7-2.) Assume that n_{35} is the most likely starting node and the pathway $n_{35}\ n_{25}\ n_{26}\ n_{27}\ n_{28}\ n_{29}\ n_{30}\ n_{31}$ is generated. If, however, the status of n_{25} is undetermined in the direction of denial, an inconsistency is indicated. If a search is made for alternative but consistent explanations and n_{25} is assumed denied, then n_{36} is selected as the most likely starting node, and the consistent pathway $n_{36}\ n_{26}\ n_{27}\ n_{28}\ n_{29}\ n_{30}\ n_{31}$ is generated.

7.7 Classification of Diseases

Recognition of the basic mechanisms of disease for a patient often is insufficient for diagnostic classification. An evaluation of the status of a patient must also determine the degree of progression and severity of disease. Patients with the same disease may exhibit different degrees of dysfunction. For example, glaucoma may lead to total blindness, but many cases will be encountered with little or no loss of vision, and these cases must be treated quite differently.

The CASNET system differentiates between two important categories of classification: (a) the mechanism of disease, and (b) the severity and the

degree of progression of disease. The cause or mechanism of disease is described in terms of the state network by the starting nodes. For a given patient, a set of most likely starting nodes will be found that identifies the underlying causal mechanism of disease. Implicit in the most likely pathways that follow from these starting nodes is a description of the progression of the disease. Statements are needed to summarize significant findings that take into account such factors as the current severity of disease and the prognosis for the patient. Additionally, specific and well-established disease labels often exist to give diagnostic descriptions. While each name may directly correspond to a specific mechanism of disease, several mechanisms of disease are frequently summarized by a single name.

The classification tables, (n_1, D_1), (n_2, D_2), ... (n_i, D_i), enable us to produce such descriptions of the status of the patient. These tables contain ordered sets of diagnostic statements interpreting the significance of the various findings and pathways of disease. When the processes of disease found in the patient are known, as displayed by the most likely pathways generated for the patient, classification tables will be searched to determine the appropriate statements.

Each starting state has pointers to the particular classification tables that contain diagnostic statements that evaluate this disease mechanism. Several starting states may refer to the same table, since several causal mechanisms may be included in the same diagnostic category. For a given patient, the most likely starting states point to the appropriate tables. The classification tables contain a series of rules ordered by seriousness of disease. The appropriate diagnostic statement corresponds to the single rule that is satisfied in the table. This rule will correspond to the deepest confirmed state in the pairs (n_k, D_k) that is reached from any of the most likely pathways that refer to this table. In most instances, an additional constraint will be added to the search of the classification table: when a state, n_i, within a table is confirmed, it must be traversed by a pathway generated from a most likely starting node that refers to this table. Otherwise, the statement for n_i is inappropriate; other pathways may refer to n_i in a different table. The deepest state is appropriate since any statement that is found earlier in the table is for a less serious stage of disease and can be ignored.

Example 6. A classification table for the primary open angle mechanism, n_{35}, from Figure 7-2, is given as

$$(n_{25}, D_1), (n_{26}, D_2), (n_{30}, D_3), (n_{31}, D_4)$$

where

D_1 = mild risk of open angle glaucoma
D_2 = high risk of open angle glaucoma
D_3 = very high risk of open angle glaucoma; significant risk of visual field loss
D_4 = open angle glaucoma

If n_{35} is selected as a most likely starting state, and n_{25} and n_{26} are confirmed but not n_{30}, then D_2 is appropriate. If n_{25}, n_{26}, n_{30}, and n_{31} are confirmed, then D_4 is appropriate.

Within a table, differing intensities of a disease process can be determined by differences in the magnitude or intensity of the states. In glaucoma, different intensities of pressure may be distinguished by defining states of moderately elevated pressure or extremely elevated pressure. These states, when found in classification tables, may then lead to different conclusions.

In some instances, it is necessary to have classification rules that indicate that specific states are denied. The same notation and interpretation for a classification table is used, where each entry in the table is not a confirmed state, but rather the required truth value (confirmed, denied, or undetermined) for that state. Multiple causes for a particular patient's disease may be either independent or related. If they are independent, separate classification tables are required. If they are related, the same classification tables are referenced for each of the multiple causes. Rules that are based on truth values (and not confirmation alone) are used to distinguish situations where multiple causes cannot be classified independently.

We can now summarize our diagnostic method as a series of transformations. As test results are received, they are related to individual states. These states are then organized into pathways inferred from configurations of a state network. The generated pathways are then related to classification tables containing the detailed diagnostic categories.

7.8 Treatment Recommendations

In some cases a therapy recommendation can be explicitly linked to a specific diagnostic conclusion. There may be a unique treatment for a given condition. In other instances, a category of treatments may be described (for example, the class of miotic medications) without an indication of a specific medication. For these simplified situations, the recommendation of a treatment is a continuation of the diagnostic statement found in a classification table.

Example 7. The classification table of Example 6 may be augmented to include treatment recommendations as follows:

$$(n_{25}, D_1, T_1), (n_{26}, D_2, T_2), (n_{30}, D_3, T_3), (n_{31}, D_4, T_4)$$

where the treatment recommendations are

T_1 = return visit in 6 months
T_2 = careful follow-up with repeated tension readings
T_3 = careful follow-up or a therapeutic trial with pilocarpine 1% QID
T_4 = miotic therapy (or, if medically uncontrolled, surgery)

The pairs (D_i, T_i) are linked together for this table; they may not always be found together for other tables and other mechanisms of disease.

A recommendation for therapy is usually a more complex problem than is described above. While the number of potential treatments that are applicable to a patient may be greatly reduced by the precision of the diagnosis, many treatments may still remain feasible. One of these treatments must then be selected. In addition, once a treatment is recommended and given to the patient, it is important to evaluate and monitor the effectiveness of that treatment.

While the purpose of therapy is to control and if possible to cure disease, the recommendation of treatment often introduces factors that are external to the original diagnosis. Specific treatments may be contraindicated because of particular conditions of the patient that do not relate directly to the diseases that are modeled. These factors must be considered before a treatment is recommended. For example, age, allergies, and history of other illnesses may all play an important role in the recommendation of a medication. A treatment for disease may in itself cause new processes of dysfunction that are unrelated to the original diagnosis. Many medications are known to cause side effects, and unwanted complications may ensue from surgical procedures.

A plan of action can be designed to select treatments for patients who fall into a particular diagnostic category. A strategy of treatment selection adapts the general treatment plan to the specific circumstances of a patient. The treatment plan must take into account (a) the effectiveness of the current treatment, and (b) indications or contraindications for various therapies.

Diagnostic conclusions for a patient are found by interpreting the specific observations within the model of disease. These diagnostic conclusions will consider severity and progression of disease. The diagnostic statements may then point to one or more treatment plans. These are shown in Figure 7-4. Each plan consists of an ordered list of treatments, $T_{i1}, T_{i2}, \ldots, T_{in}$. The list is ordered by preference: treatment 1 is tried before treatment 2, which will be tried before treatment 3, etc. This plan represents a prototypical sequence of treatments for patients in the appropriate diagnostic categories, as agreed to in advance by the experts in the domain. A strategy for recommending treatment for an individual will usually follow the order

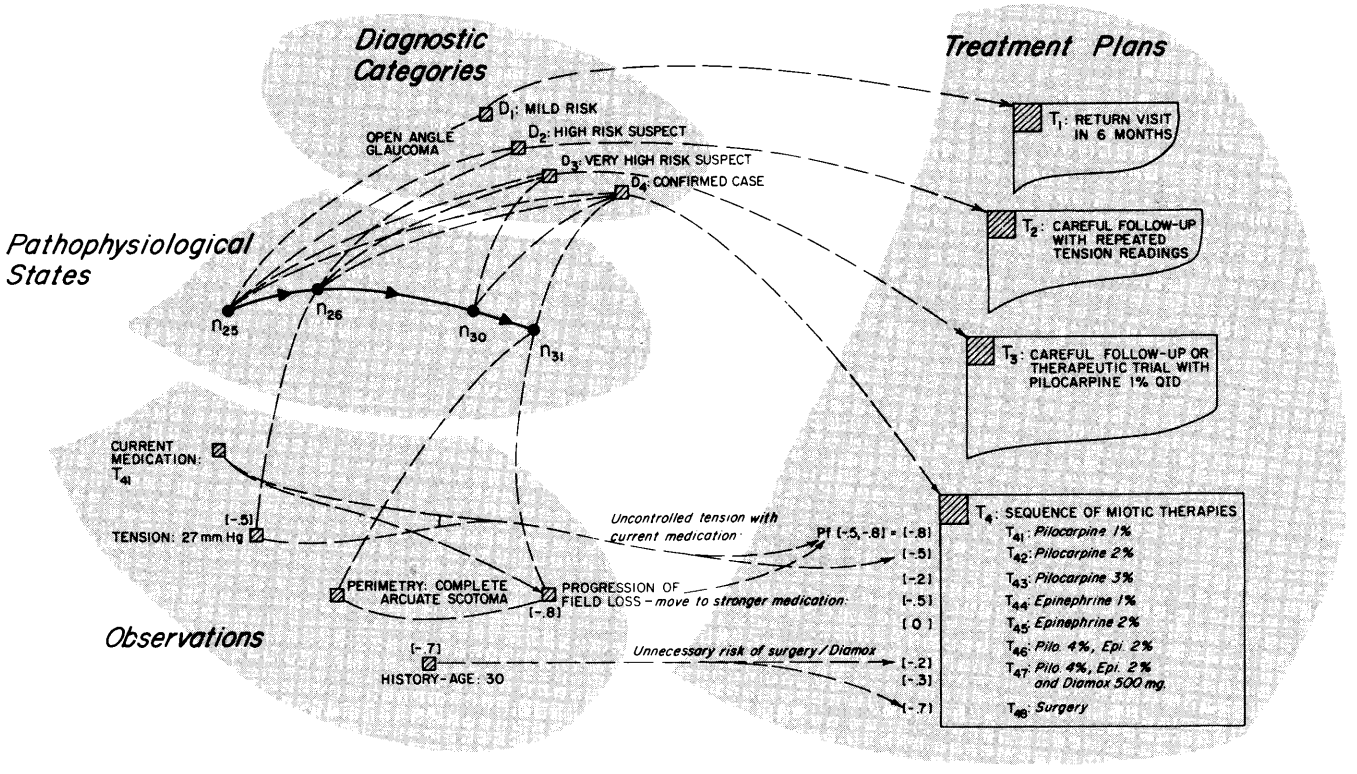

FIGURE 7-4 Examples of treatment plans.

of treatment preferences. However, the ordering may be changed in response to particular observations noted in the individual patient. Within a treatment plan, deviations from the prototype result from changes in the degrees of preference or contraindications for specific treatments. In certain situations, no well-established set of preferences exists, and the selection of a treatment from within the general plan is almost completely determined by the pattern of observations for each patient.

The strategy for treatment selection is described as follows: within a treatment plan, T_k, each specific treatment, T_{kj} has associated with it a preference measure, $\mathrm{Pf}(T_{kj})$, which is assigned from direct observations of the patient. Each observation, t_i, that affects the preference of T_{kj} contributes a measure, $\mathrm{Pf}_{ij}(-1 \leq \mathrm{Pf}_{ij} \leq +1)$, which is assigned in a manner similar to the Q_{im} for relating observations to the disease states, n_m. For example, a drug intolerance may associate a negative preference with a particular treatment. For glaucoma, a very high tension reading after treatment would indicate ineffective control of the disease and contribute a negative preference to the current treatment. Being in a particular age group may increase the preference measure of one treatment over another. The overall preference measure, Pf, is computed by the same rules used to compute the confidence measure, Cf, for the disease states. Once the Pf values have been computed, the rule for selecting a specific treatment, T_{kj}, from within its plan, T_k, can be summarized as follows:

a. Select T_{kj} such that $\mathrm{Pf}(T_{kj}) = \max_i [\mathrm{Pf}(T_{ki})]$.

b. If there is more than a single treatment with maximum Pf, select the one with smallest index j in the *a priori* prototypical ordering.

Example 8. A treatment plan T_4 corresponding to a confirmed case of open angle glaucoma (D_4), as indicated in Example 7, is shown in Figure 7-4. The $\mathrm{Pf}(T_{4j})$ are computed from the observations of a patient, some of which are illustrated in the figure. The patient shown is currently under treatment T_{41}, yet the observed tension of 27 mm of Hg indicates an uncontrolled intraocular pressure. This assigns a decreased preference measure of -0.5 to the current treatment T_{41} and the related treatment T_{42}. The patient also showed progression of field loss, which decreases the preference (-0.8) for the current medication T_{41} even more strongly. Because the patient is under 30, a systemic medication such as Diamox is less preferred (-0.3). The relatively higher risk of surgery versus medication results in the assignment to T_{48} of a decreased preference, -0.7. As a result of comparing these and other Pf's derived from the observations, the treatment with the maximum preference for this patient is T_{45}, which is recommended.

7.9 Results and Discussion

A general method for solving a class of diagnostic and therapy selection problems has been presented. These ideas on model-based interpretation have been put into practice through the implementation of a computer system for medical decision making. Much experience has been gained in the development of a model for the diagnosis and treatment of glaucoma, which has led to the design of a system with a high level of "expertise." This has influenced our general approach toward knowledge representation and reasoning procedures. The consultation system, however, is not specific to glaucoma. Other models of disease have been developed for the anemias, thyroid dysfunction, diabetes, and hypertension. Glaucoma, however, is the one application that has been pursued in depth and has undergone clinical testing.

The design of a consultation system can be broken down into two important tasks. These are the design and representation of models and the design of general problem-solving algorithms that use a suitably defined model for decision making. In the general CASNET system, a medical expert describes or modifies a model, but does not alter the reasoning procedures that select diagnostic interpretation and treatment plans. Two separate computer programs have been developed: the modeling program for designing application models (Kulikowski and Weiss, 1973b), and the consultation program that uses models for reaching diagnoses and recommending therapies (Weiss et al., 1978).

The current glaucoma consultation system has more than 100 states, 400 tests, 75 classification tables, and 200 diagnostic and treatment statements. Results must be interpreted for each eye, so that, in effect, twice the number of rules are involved in any ophthalmological model. There are also many special rules for binocular comparisons of states, tests, and diagnostic and treatment statements. A set of the program's conclusions for a sample case is given in Figure 7-5. This session illustrates the level of performance that the program has attained in reasoning about complex cases of glaucoma.

The consultation program has been designed for efficient performance. Human-engineering aspects of program design have also been emphasized. The program has been developed primarily as a tool for the research of medical decision making by computer. However, our approach to program development involved the collaboration of a network of physicians with minimal prior experience in the use of computers. Their active participation in the project required careful attention to programming details that would allow our collaborators and other ophthalmologists to use the programs with little difficulty. This implies that only limited typing

VISIT 1:

RIGHT EYE:

(1) PRESENT DIAGNOSTIC STATUS:

PIGMENTARY GLAUCOMA. OPEN ANGLE GLAUCOMA. CHARACTERISTIC VISUAL FIELD LOSS WITH CORRESPONDING DISC CHANGES. EARLY FIELD LOSS.

(2) TREATMENT RECOMMENDATIONS:

PILOCARPINE 2% QID.

RESEARCH STUDIES

ALTERNATIVE INTERPRETATIONS OF PIGMENTARY GLAUCOMA:
. SECONDARY GLAUCOMA
. PRIMARY OPEN ANGLE GLAUCOMA

REFERENCES:

1. "WHEN PIGMENTARY GLAUCOMA WAS FIRST DESCRIBED IT WAS THOUGHT TO BE A FORM OF SECONDARY GLAUCOMA CAUSED BY PLUGGING OF THE TRABECULAR MESHWORK BY THE SAME PIGMENT THAT FORMED THE KRUKENBERG'S SPINDLES. HOWEVER, AN INCREASING NUMBER OF OBSERVERS NOW BELIEVE THAT IT IS A VARIANT OF PRIMARY OPEN ANGLE GLAUCOMA..." (WILENSKY, PODOS 1975, TRANSACTIONS NEW ORLEANS ACAD. OPTH.)
2. "MORE RECENT EVIDENCE SUGGESTS THAT PIGMENTARY GLAUCOMA IS A SEPARATE ENTITY..." (ZINK, PALMBERG, ET AL, A.J.O., SEPT. 1975)

VISIT 7:

RIGHT EYE:

(1) PRESENT DIAGNOSTIC STATUS:

PIGMENTARY GLAUCOMA. OPEN ANGLE GLAUCOMA. CHARACTERISTIC VISUAL FIELD LOSS WITH CORRESPONDING DISC CHANGES. ADVANCED FIELD LOSS. CURRENT MEDICATION HAS NOT CONTROLLED IOP IN THE EYE. (AS INDICATED BY PROGRESSION OF CUPPING) (AS INDICATED BY VISUAL FIELD LOSS PROGRESSION)

(2) TREATMENT RECOMMENDATION:

FILTERING SURGERY IS INDICATED. AS AN ALTERNATIVE, PHOSPHOLINE MAY BE TRIED (BUT NOT USED 2 WEEKS BEFORE SURGERY).

FIGURE 7-5 Examples of program-generated decisions for a case of pigmentary glaucoma, abstracted from a sequence of seven visits.

would be required and that quick response time, even for complex diagnostic interpretations, would be essential.

Initially, we designed and built a prototype model that was demonstrated to a select audience of ophthalmologists. At this point, the program was far from being expert. However, rapid progress in the development of a decision-making system can be made by building a small simplified prototype and modifying and improving the prototype.

A very significant event in the development of the program has been the formation of ONET—the Ophthalmological Network. Using the SUMEX-AIM computer,[1] we have put together a nationwide group of ophthalmological clinician-researchers who have participated in the development of the program's knowledge base. They enter cases and suggest improvements. Their suggestions have not been based on a comprehensive review of the logical rules contained in the program. Rather, we have concentrated on entering realistic cases and comparing the program's questioning sequence and conclusions with those of the experts.

Within a period of approximately a year and a half of ONET collaboration, the program achieved an expert level in the long-term diagnosis and treatment of many types of glaucoma. The program's performance has been validated by our group of experts and by the system's participation in panel discussions of glaucoma cases at ophthalmological symposia. In November 1976 a scientific exhibit of the program was presented at the annual meeting of the American Academy of Ophthalmology and Otolaryngology. Ophthalmologists were invited to present difficult cases to the computer. The program did well, with 77% of the ophthalmologists who entered cases describing the program as performing at an expert or very competent level (Weiss et al., 1978).

In comparing the experiences of modeling glaucoma and other diseases, we have obtained some insight into the advantages and limitations of the CASNET representation and its associated decision-making methods. When an understanding of the mechanisms of disease serves as a basis for decision making, the CASNET approach is most valuable. When reasoning is mostly judgmental and based more on empirical information than knowledge of the disease mechanisms, other decision models may prove more appropriate (Patrick et al., 1974; Shortliffe et al., 1973).

In the MYCIN system (see Chapter 5), descriptive domain knowledge is implicitly contained within the system of production rules that encode the clinical judgment of an expert consultant. Therapy selection for infectious diseases is a medical domain in which empirical knowledge plays a predominant part in the problem-solving process, and it is not surprising that this domain has been successfully modeled in terms of judgmental rules alone. In glaucoma, as in other diseases where mechanisms of dysfunction are reasonably well known and have an important effect on the selection of treatments, we have developed a more structured representation for causal knowledge. And yet, since strict Aristotelian causality is hardly applicable in medicine, the causal representation is embedded within an associational structure of observations that accounts for the uncertainties of clinical findings.

In questions of hypothesis generation and approximate reasoning, the

[1]This computer was established at Stanford University with NIH support to provide a national shared resource for research in AIM (AI in medicine).

CASNET approach is quite distinctive in its use of the causal-associational structuring of knowledge. An overall diagnostic hypothesis for a patient is usually a composite of several hypotheses. It is not uncommon to find five or six hypotheses included in the final diagnostic statement. Many of these hypotheses include statements of uncertainty within them, as, for example, "very high risk of glaucoma" or "mild risk of glaucoma."

Approximate reasoning takes place at several levels. Measures of uncertainty are used to interpret observations in terms of the most elementary subhypotheses: the pathophysiological states of the causal network. A thresholding of the measures of uncertainty for all observations that are relevant to a given state determines whether that state is to be considered a "confirmed," "denied," or "undetermined" subhypothesis for the patient. At this level the method corresponds to the usual approach of assigning a likelihood or degree of belief to a hypothesis. At a higher level of abstraction these subhypotheses of states are grouped together in a more deterministic fashion. Measures of uncertainty are less important at this stage because the hypotheses themselves include qualifying statements as described above. Thus the greatest reduction of uncertainty takes place between the observations and the states, which serve as local and relatively simple summaries of events in the course of a disease. The detailed structural relationships among states allows a fine-resolution encoding of the possible patterns of the disease. Because statements of uncertainty are associated with these patterns, they can be related to final hypotheses in a deterministic logical manner without losing the soundness of the outcome. It is often advantageous to do this, in so far as it corresponds more closely to the conclusions expressed by an expert physician.

The explanations produced by the CASNET/Glaucoma system also appear to correspond more closely to those of the physician. Instead of tracing all the rules involved in arriving at the final diagnosis, the composite hypothesis includes certain key subhypotheses that the physician recognizes as necessary elements in justifying the conclusions or recommendations. For example, in glaucoma, a typical subhypothesis would be "corresponding disc and visual field changes," which is both explanatory and supportive of a higher-level hypothesis of "open angle glaucoma." The subhypothesis is itself the summary of many different observations. In building the CASNET system, we have found that exhaustive tracing of rules is much more valuable as a debugging tool than as an explanation for the physician.

The CASNET/Glaucoma system has proved to be highly efficient and sufficiently expert to be accepted as such by many ophthalmologists. Its solutions to many of the representational and strategy questions have been shown to be effective in a realistic problem domain. Nevertheless, the role of such large knowledge-based consultation systems in routine clinical practice remains an open question.

ACKNOWLEDGMENTS

Supported in part by grant RR-643 from the Biotechnology Resources Program, Division of Research Resources, NIH, and in part by grant R01-MB-00161 of the Health Resources Administration.