

To appear in Applied Artificial Intelligence, Spring 1989.

Medical Diagnosis Using a Probabilistic Causal Network

William Long

Principal Research Associate

M.I.T. Laboratory for Computer Science

545 Technology Square, Room 368

Cambridge, MA 02139, U.S.A.

December 12, 1988

Abstract

This paper relates our experience in developing a mechanism for **reasoning** about the differential diagnosis of cases involving the symptoms of heart failure using a causal model of the cardiovascular hemodynamics with probabilities relating cause to effect. **Since** the problem requires the determination of causal mechanism as well as primary cause, the model has many **intermediate** nodes as well as causal **circularities** requiring a heuristic approach to **evaluating** probabilities. The method we have developed builds hypotheses **incrementally** by adding the highest **probability** path to each finding to the hypothesis. With a number of enhancements and computational tactics, this method has proven effective for **generating** good hypotheses for typical cases in less than a minute.

1 Acknowledgements

The research reported here was done with my medical collaborators at New England Medical Center, Shapur **Naimi**, M. G. Criscitiello, Stephen Pauker, Greg Larsen, and in earlier years of the project, Robert Jayes and Steven **Kurzrok**. They are responsible for the medical content of the program, although any mistakes in this paper are my own. The work reported here has been supported by National Institutes of Health grants **R01** HL 33041 from the National Heart, Lung, and Blood Institute and **R01** LM 04493 from the National Library of Medicine.

2 Introduction

Over the past few years we have been developing a system to assist the physician in the diagnosis and management of patients with diseases that may cause or resemble heart failure. The program will be used by interns, residents, house staff, and other physicians who are managing patients with complex disorders over a period of days in a setting such as an intensive care unit. It will assist the physician in reasoning about the patient for both diagnosis and management. The physician can enter information about the history, physical examination, and laboratory tests (i.e., the findings) and the program will provide a differential diagnosis list with a graphical explanation of how each set of causes could produce the findings, suggestions about what other information would help to differentiate among the possibilities, suggestions about therapies that could correct the causal paths leading to undesirable states, and provide predictions about the overall effects of various therapies given the patient's pathophysiological state.

The focus of this paper is the problem of providing useful information about the diagnosis. The cardiovascular domain, as well as many others in medicine, is full of uncertain causal mechanisms. This has led to a representation of the domain knowledge as a network of causal probabilistic links between physiologic parameter states. However, this network contains multiple paths between nodes and forward cycles, necessitating the development of heuristic methods for evaluating the state of the network when findings are known. These methods are used to produce multiple hypotheses representing possible explanations for the findings, which are put together as a differential diagnosis. We have found that these methods are effective for interactive use in a model that contains about 150 internal nodes and about 300 possible findings.

The following sections will discuss the special problems of the medical domain, the nature of a differential diagnosis, the kinds of causal relationships and the way they are modeled as probabilities in the program, the approach to producing diagnostic hypotheses, and an example of its use. The emphasis throughout is on the lessons we have learned about the nature of the problems and the practical concerns in building tools to meet the needs of the users.

3 The Domain Context

The first issue in designing an expert system is to determine an appropriate view of the domain and of the reasoning required in problem solving. It may be easier to take a solution and look for a suitable problem, but the potential for addressing the needs of the user are greater if we investigate the problem first and find or develop methods to address the problems that are there.

The medical domain for our program is the diagnosis of patients with heart failure. Heart failure is the state of the patient when some disease process makes cardiac output inadequate for the demands of the body. There are many possible causes of heart failure including diseases of or damage to the myocardium (heart muscle), dysfunction of a heart valve, restriction of the heart by the pericardium (enclosing membrane), electrical dysfunction, and so forth. The cardiovascular system has strong mechanisms that attempt to compensate for the inadequate output. One mechanism alters the capacitance and volume of fluid compartments to increase the heart input pressure (preload) in an attempt to increase cardiac output. A longer term mechanism also starts retaining fluid to increase the blood volume to also increase preload. If the patient has lost blood, these are appropriate responses. When the problem is not low blood volume, the mechanisms can produce high pressures in the lungs, stiffening them, and causing fluid to leak into the air space, producing the common symptoms of pulmonary congestion. When the pressures increase on the right side of the heart, the pressure balance changes in the systemic venous circulation and the patient gets edema in the ankles and up the legs. Low cardiac output also causes the systemic blood vessels to contract to maintain the blood pressure to the important organs. While this is appropriate when blood volume is low, it often puts excessive stress on an already weakened heart. Since many different causes produce some combination of the same overt symptoms, they are grouped together as the syndrome of heart failure.

The problem for the physician is not simply to determine what disease the patient has and order an appropriate therapy. Since many of the most important diseases are incurable short of replacing the heart, the physician must determine how the disease is causing the findings in order to identify points where therapy can reduce the undesirable effects. If the disease is correctable, it is important to recognize that — indeed to rule out all correctable diseases. Fortunately, with most of the incurable diseases the patient is usually able to live an active and relatively normal life if the treatable problems and compensatory mechanisms are appropriately controlled. Since

these tend to be chronic problems, many of the cases admitted to the hospital have multiple diseases — often a new problem added to a chronic problem. In addition, the patients are usually on therapies which may change the symptoms of the diseases or cause symptoms of their own. The physician needs to determine what diseases the patient has and the causal mechanisms by which they are producing the observed symptoms in order to determine whether there are treatable causes and what therapies might limit or correct the undesirable effects of the other causes. Thus, the goal of diagnosis is to determine the pathophysiological state of the patient in enough detail to guide therapy.

These diagnostic requirements imply that we need a disease model with enough intermediate states to distinguish among situations where different therapies are appropriate. For example, a myocardial infarction (heart attack) can produce both diastolic (ventricular filling) and systolic (ventricular emptying) dysfunction, each of which has different implications for therapy. To make such distinctions, a simple associational scheme such as those used by DXplain[Barnett87], or QMR[First85] would not be sufficient.

4 Differential Diagnosis

While the goal of diagnosis is to determine the pathophysiology in sufficient detail to suggest a rational approach to therapy, this goal usually can not be achieved from the initial history, presenting complaints, and physical examination. Tests or reexamination may be needed to determine the appropriate diagnosis. The tool for directing the diagnostic process is the differential diagnosis. The differential diagnosis identifies possible explanations for the known findings, identifies states that have been ruled out, and provides the data needed to determine what information can be used to refine the diagnosis.

Over the course of this research our view of differential diagnoses has evolved. Initially, we used the model and findings to identify parameter states that were well supported. For example, pulmonary findings would be used to decide whether pulmonary congestion was true or false in the patient. Then these states (e.g., pulmonary congestion) could be used as evidence for other parameter states. In this way the diagnosis could be built state by state. At any point in the process, the differential diagnosis was the union of the possible values of the states not yet known with the states known to be true. This is essentially the approach that was

taken by CASNET[Weiss78] to reason from a causal network. There are two problems with this approach. First, it assumes there will always be enough evidence local to a parameter state (easily computable causes or consequences) to determine its truth, either immediately or given more test results. In fact, what often happens is that there are parameters with states whose probabilities are not overwhelmingly true or false, but any decision about one parameter state will change the probability of the others with possible far reaching consequences. For example, a minor decision of whether the right ventricle has decreased compliance versus fixed capacity in the model could determine whether the best diagnosis is restrictive cardiomyopathy or constrictive pericarditis — a minor distinction on which to make the diagnosis. Second, the view of the differential as all unknown states does not give a good picture of the likelihood of different possibilities, especially the significant combinations of states that would have consequences in patient management.

Because of the restrictions of the parameter by parameter view of the diagnosis problem, we started considering complete hypotheses — hypotheses that include causal chains from the primary causes (causes that do not require a further cause) to the findings and that cover **all** of the patient findings. With these it is possible to get a complete picture of what might be happening in the patient. This approach also allows the user to evaluate the reasoning of the computer. Since the hypotheses are complete, there are no hidden assumptions or implications. The user can decide whether the scenario proposed by the computer is reasonable and if not, why not.

Given that a hypothesis is complete from primary cause to findings, we are still left with the problem of how to form a differential. The set of **all** possible hypotheses is much too large to list and the hypothesis with the highest probability gives no indication of the alternatives. One could generate alternative hypotheses by taking the best hypothesis and proposing alternate paths for single parameter states or short chains of states. However, many of the alternatives that might be generated would differ from the starting hypothesis in inconsequential ways, while major alternatives would be overlooked. We have addressed the issue of what constitutes important alternatives by designating a subset of the parameter states (nodes) in the model as *diagnostic nodes*. Hypotheses are significantly different if they contain a different set of diagnostic nodes. This heuristic has proven effective in limiting the differential diagnosis to a reasonable number

of alternatives that have significant differences.

5 The Causal Disease Model

Developing an appropriate disease model requires two major steps, deciding what diseases need to be included and determining the causal structure among the diseases, intermediate states, and findings. The domain is management of the hemodynamic consequences of heart failure. Therefore we need to include the diseases that alter the hemodynamics. However, there are many definitions of a disease. One important disease of the myocardium is congestive cardiomyopathy. This disease can be further classified by causes (alcohol, sarcoidosis, idiopathic, etc.) but no matter what the cause, the effect on the cardiovascular hemodynamics is the same. That is not to say that the entire treatment is the same — if there is a treatable cause weakening the heart, that must be treated — but those aspects involved in correcting the hemodynamic dysfunction remain the same. Therefore, we have one node for congestive cardiomyopathy and it is considered a primary cause. Since there is also a node for ischemic heart disease in the model which can cause congestive cardiomyopathy, that causal link must be included. Thus, congestive cardiomyopathy may be primary or caused by other nodes. The decision about what to include is not always so simple. There is an uncommon disease, left atrial myxoma, that partially obstructs the mitral valve and produces the same hemodynamic dysfunction as mitral stenosis. The physical examination clues to distinguish the two diseases are quite subtle, but if either disease were suspected, an echocardiogram or other imaging test would be done which would clarify the situation. Therefore, there is no advantage to including the less common disease and the common disease node is used to cover both. There are also a number of congenital diseases that have been left out of the program because such diseases are not initially found in the adult (our target population). Thus the model has in it the cardiovascular causes of the hemodynamic dysfunctions at the level needed to distinguish the different hemodynamic manifestations.

There are also a number of diseases with findings similar to those produced by hemodynamic dysfunction. For example, the pulmonary findings of pneumonia are often very difficult to distinguish from those of cardiac pulmonary congestion. Since such diseases are often the most important differential diagnoses, they must be included. The problem again is the level of detail. Both low cardiac output and renal *insufficiency* cause retention of fluids, but there are many

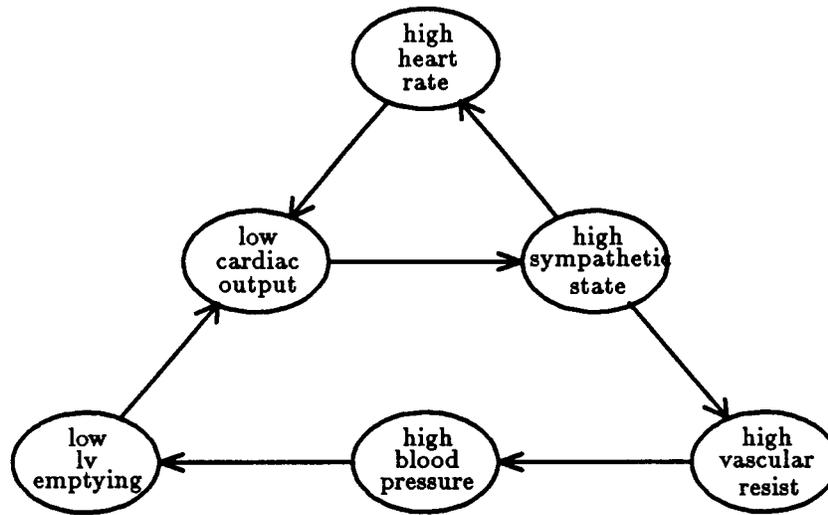


Figure 1: Circularities in the Causal Structure

different kinds of renal insufficiency because there are many kinds of kidney disease. Our solution has been to only cover such diseases to the extent that they produce findings similar to cardiac findings. Thus, we have entities such as renal insufficiency and primary liver disease. The result is a model with decreasing specificity as one gets away from the central focus of hemodynamic disorders.

Developing the model requires a change of perspective for the cardiologist expert. Usually the cardiologist thinks in terms of diseases and findings and not in terms of the intermediate causal mechanisms. To develop a model that includes the intermediate nodes, we started with lists of associations between diseases and findings, identified findings generated by the same mechanisms, and added intermediate nodes representing those mechanisms. For example, many diseases cause shortness of breath, rales (noises in the lungs), and certain X-ray findings. These are all produced by fluid buildup in the lungs, so an intermediate node was added for pulmonary congestion. Identifying intermediate nodes was not done blindly. We started with the model used in the program to predict the effects of **therapy**[Long88]. This model has equations relating the constraints among parameters and dependencies on physiologic state. Since diagnosis of the patient provides much of the information needed to establish the pathophysiologic state for this model, the intermediate nodes must be compatible.

The use of the predictive model also pointed out situations requiring circularities in the

causal structure. Two circularities are shown in Figure 1. In both cases low cardiac output causes a high sympathetic state. The high sympathetic state increases the heart rate, but a high heart rate can decrease cardiac output, especially in situations such as aortic stenosis, when ejection time is prolonged. The high sympathetic state can also increase the systemic vascular resistance, which increases blood pressure and the load on the heart, which decreases left ventricular emptying and cardiac output. Both of these situations are positive feedback loops. If any of the nodes is triggered by some pathophysiological situation, there is a significant probability that the succeeding nodes in the loops will result. Indeed, one reason that therapies to decrease heart rate or decrease vascular resistance are often effective in patients having low cardiac output is that these mechanisms often overdo their jobs.

Using the therapy model and clustering the nodes by physiologic mechanism, we built a model that covers the pathophysiology of cardiovascular hemodynamics. The resulting model has as many as a dozen intermediate nodes between a primary node and a terminal node. It makes explicit the mechanisms that produce the findings and allows the representation to distinguish the variations and combinations of diseases and alterations due to therapies. Because we started with the list of associations, we also have that as a check of the expanded model. That is, we can compare the expected associations in the expanded model to our initial list of associations and use the differences to refine the model.

6 Probability Model

Given the form of the causal disease model, the next step is to represent the relation between each cause and effect. To simplify reasoning with the causal relations, we have imposed a uniform view of *probabilistic causality* on the model. That is, each causal relation is summarized as a probability of the cause producing the effect. Furthermore, causes precede effects and nodes are either true or false. By imposing this view, we are able to handle the circularities in the model in a consistent manner. The causation can be viewed as growing an acyclic **subgraph** from primary causes within the model graph. Anytime a circularity in the causation would be encountered, the effect node is already true and further causation is blocked. This causal view eliminates anomalous conditions such as circular causation with no initiating primary cause.

The probabilities on the links summarize several kinds of causality in the model. Some of

the relations are randomly triggered causal events. For example, there is a probability that a patient with obstructed coronary arteries will have a myocardial infarction. The probability is partially dependent on other factors, such as the heart rate, blood pressure, and so forth, but the event is essentially random and binary.

Many other relations can be expressed as equations. For example, blood pressure is a function of cardiac output, systemic vascular resistance, and right atrial pressure. Usually however, not enough of the parameter values are known to constrain the equation and the easiest way to represent the potential for low cardiac output or low systemic vascular resistance to produce low blood pressure is to approximate the relationships as causal probabilities.

Other causal relations are dependent on duration or severity of the cause to produce the effect. For example, cardiac dilatation usually takes weeks to occur but a severe myocardial infarct can cause it in days or even hours. In such cases, the probability on the link is approximating the more complex relation of duration and severity. The probability could be defined as a function of duration and severity, but even without information about either, there must be an average time after an acute infarct when the diagnosis is done and an average severity and the probability reflects those averages.

Representing the states of cardiac output, blood pressure, and other measurable parameters with qualitative descriptors such as low, normal, and high is a compromise to constrain the diagnostic problem. The ranges of the parameters could be divided into much finer intervals, approximating their probability distribution functions. Such a scheme requires a high computational cost without much benefit because the distinctions between states are harder to recognize from the findings, and the differences between hypotheses become insignificant. The method used is to define the qualitative values by causal distinctions and let the relationship between the qualitative values and the quantitative measures be probabilistic. The criteria for distinct states is a difference in the causal structure. Thus, low cardiac output is different from normal cardiac output because low cardiac output can produce fatigue, inadequate renal perfusion, sympathetic response, or other effects. If there are different causes or effects, it is a different state. For the most part, simple designations such as low, normal, and high have been appropriate for the parameters. Since the parameter state is determined by the existence of causes and effects, the relationship to measured values is not always easily defined. The level of cardiac output at which

the effects of low cardiac output start happening varies from patient to patient. The program handles this by specifying the probability that each qualitative state produces a measured range of the parameter:

low: (range cardiac-index 0.7 2.3 0.2 2.5 0.1 2.7 0.0)

normal: (range cardiac-index 0.0 2.3 0.05 2.5 0.05 2.7 0.9)

Thus, 70% of low cardiac outputs have a cardiac index (cardiac output normalized for body size) below 2.3, 20% between 2.3 and 2.5, and so forth. Similarly, no normal cardiac outputs are below 2.3, but 5% are between 2.3 and 2.5, and so forth. These distributions allow the program to handle the cardiac index as evidence for the qualitative values of cardiac output in the same way as it would a categorical measurement value.

The probabilities on the links may be fixed, dependent on patient parameters, or dependent on the diagnostic hypothesis. In most cases there is not enough data or experience to recognize situations in which the probability of a causal link will change. Therefore the probability on the link is just a number. In other situations the probability varies with measurable patient parameters, usually age and sex. For example, pneumonia is less likely to produce a fever in the elderly than in a younger person and the probability on the link is:

pneumonia: $p(\text{fever}) = (0.9 \text{ (range age } 0.95 \text{ } 70 \text{ } 0.9 \text{ } 80 \text{ } 0.8 \text{ } 90 \text{ } 0.7))$

If nothing is known about the age, 0.9 is used. If the patient is less than 70, the probability of fever is 0.95. If the age is between 70 and 80, 0.9 is used, and so forth.

Since the qualitative parameter state can represent a wide range of values, the probability of the state having an effect can be dependent on the actual parameter value. For example, the probability that high heart rate will cause low cardiac output is strongly dependent on the actual heart rate (as well as age). Many of the primary cause probabilities also vary with age and sex. The dependencies on patient parameters are handled by adjusting the probabilities on links before attempting to compute a differential diagnosis.

The probability of an effect being produced also varies with the number of causes. The approach we have taken is to assume independence of causes and use the "noisy-or" combining function unless there is specific knowledge about how a combination of causes changes the probability. That is, the probability of the effect is $1 - \prod(1 - p_i)$, where p_i is the probability of

the cause producing the effect by itself. There are in addition, two kinds of combinations that are treated differently. First, there are factors that increase the probability of a cause producing an effect, but can not produce that effect themselves. For example, a high heart rate will make a myocardial infarct more likely but can not produce it without another cause being present. The probabilistic contributions of these worsening or precipitating factors are combined as if they were causes when another cause is present, otherwise they are treated as zero. Similarly, there are factors that decrease the probability of a cause producing an effect. Usually these are therapies but there are also pathophysiological states that prevent or make other states less likely. For example, mitral stenosis usually produces high pressure in the left atrium (the chamber before the constricted valve). However, if there is tricuspid regurgitation, the right side of the heart is less likely to be able to maintain an elevated pressure. Such factors are combined to existing causes and multiplicatively decrease the causal probability. For example, if the the causes imply a 0.5 probability of producing the effect and there is a correcting factor that prevents the effect 80% of the time, the probability of the effect is 0.1.

With these mechanisms we are able to model the physiologic relationships in the medical domain as a probabilistic network. The network is the knowledge base from which diagnostic reasoning is done.

7 Computing the Differential Diagnosis

The knowledge base, with its binary nodes and probabilities, resembles a Bayesian belief network as defined by Pearl[Pearl86]. Such networks have been the subject of considerable research by Pearl, Lauritzen and Spiegelhalter[Lauritzen88], Cooper[Cooper86], and others. However, the network of our model does not conform to the definition of a Bayesian belief network because it has forward cycles. Even if we removed all of the forward cycles, there are still multiple paths between nodes. These are not forbidden in Bayesian belief networks, but they greatly increase the amount of computation required to compute the probabilities. The algorithms that have been developed to handle multiple paths are exponential and Cooper has shown that the problem is NP-hard[Cooper87]. We did an analysis of an earlier version of the model and determined that to use the method described by Pearl[Pearl86] and interpolate from a set of nodes whose values cut the multiple paths would require more than 40 nodes in the cut set. Since the probabilities

in the network for all combinations of values of this set would have to be evaluated, an exact solution is infeasible.

These observations led us to develop heuristic methods for approximating the probabilities of nodes and finding good hypotheses. The approach we developed is based on an observation about the nature of hypotheses. A hypothesis is a subset of the model such that every finding and node in the hypothesis is accounted for by another node or is primary. This subset can be viewed as the union of one or more paths through the model from a primary node to each finding. Thus, the best hypothesis is the set of paths covering all of the findings which taken together has the highest probability. The problem is difficult because paths that share links have higher probability than the product of the individual path probabilities. Thus, the best explanation for two findings may not include either of the best paths for the findings taken separately. Still, in a reasonably behaved network, a good heuristic approach is to *take the findings one at a time and build up a hypothesis using the best path to each finding*.

While this is the basic approach we are using, there are a number of enhancements that can be made, most of which have already been incorporated into our algorithm. Each of these is a heuristic for restricting the search and increasing the probability of the hypotheses found. The test of whether a subset of the nodes in a hypothesis is better than an alternative set of nodes is to compare the total probability of the hypotheses.

1. Some nodes may be definitely true or definitely false from the findings. These are set and used as if they were findings. For example, if there is a finding with only one cause and that requires a cause (i.e., it can not exist by itself), the cause must be true. Similarly, if there is a node that always produces a finding and that finding is absent, the node is false. Often there are several nodes definitely known, significantly constraining the search for hypotheses.
2. The search for a path for a finding can make use of the paths already selected for other findings. Often the best explanation of a finding is a short path from a node in the partially generated hypothesis rather than a path from a primary node.
3. Since sharing of paths decreases the number of primary nodes in the hypothesis, the algorithm starts the search by selecting a cover set of primary nodes to account for all of

the findings and restricts the search to these. Currently, the program uses the primary nodes that account for a large number of findings and for each of these generates all minimal additional sets of primary nodes needed to account for the findings. This way all minimal cover sets are used plus cover sets that include each primary node that accounts for a large number of findings. Care must be taken in generating cover sets because there are times when an explanation with two causes has a higher probability than an explanation with only one of the causes even though the single cause is sufficient.

4. When there are multiple pieces of evidence (findings or true nodes) for a node, it is often possible to show that the node is the best explanation for the evidence under fairly conservative assumptions by comparing the probability of the findings with the node in the hypothesis to an estimated maximum probability with the node false. If this comparison is done later in the hypothesis generation process, the probability estimates are better but there is less savings in search. This process takes considerable computation and we have yet to establish a definite benefit.
5. The probabilities of the paths can be adjusted for the negative findings and nodes that are direct consequences of the nodes in the path. Since the probability that a node is false is $\prod(1 - p_i)$, where p_i is the probability of true causes, $(1 - p_i)$ is the probability that the node will remain false when the node i in the path is true. This gives a better estimate of the probability that the path is part of the diagnosis.
6. The diagnostic nodes provide a natural place to break the paths so that fewer paths have to be considered. Thus, if the probability of the diagnostic nodes can be approximated, they can be used as if they were primary and divide the search. When this approach was added to the algorithm, the number of causal paths was reduced by a factor of about four.
7. Ordering the findings by the difference in probability between the best path and the second best path, improves the chances that the best combinations of paths will be found. Thus, it is selecting first the paths that are most clearly indicated as best paths. We have tried several methods of ordering the findings and while it is clear that the ordering makes a difference, it is difficult to tell how much better this method is than others.

8. Pruning a hypothesis can sometimes improve it — essentially hill climbing to a local maximum probability. Sometimes a primary cause is unneeded because some findings may be better left unexplained or other primary causes cover them. Sometimes sections of paths are unneeded because paths added later account for **all** of the findings. Care must be taken in removing sections to insure that all nodes remain reachable from a primary node, because circularities in the causality appear to be self supporting.

There are also a number of computational tactics that can be employed to make computing the differential diagnosis reasonably efficient.

1. Precompute **all** of the causal paths in the model when the system is initialized. This can be done efficiently by generating a tree structure through the effects from each primary node and diagnostic node. Thus, the paths share structure and the only link unique to each path is the last link. Before the diagnostic nodes were used to divide the paths and unusable paths were eliminated, there were about 70,000 paths. There are now about 7,000.
2. Some paths can be eliminated if there is another path with a subset of its nodes and the minimum probability of the shorter path is always higher than the maximum probability of the longer path. (The longer path has a detour around one link in the shorter path.) Eliminating such paths more than halved the number of paths.
3. Use bit vectors to do comparisons between paths and sets of nodes. For example, examining paths that come from the primary nodes in the cover set is done by checking for a non-zero intersection between the bit vector of the cover set and the bit vector of each path. Also, a path is discarded if its bit vector intersects the bit vector of the set of nodes known to be false.
4. The probabilities along the paths can be computed once the user enters the findings. Since paths are examined multiple times and share structure, this is much more efficient than computing the probabilities as needed.
5. Once consideration has been restricted to a small number of primary nodes in generating a hypothesis, there are often intermediate nodes that must be true because **all** causal paths

from the selected primary nodes to the findings go through them. These can be found by intersecting bit vectors.

6. Hypotheses can be compared by their total probability — essentially the probability that a patient with that mechanism producing that set of findings would exist in the general population. It is unnecessary to normalize by the probability of the findings because **all** of the hypotheses to be compared have the same findings.

With these enhancements and computational tactics, the algorithm has proven effective for generating differential diagnoses in our domain. Even with the large size of the probabilistic model, the algorithm runs in less than a minute on a typical case with about a dozen abnormal findings on a Symbolics 3650. This makes it useful as an interactive tool for physicians analyzing a case. The method is still heuristic. Indeed we have found cases where there was a better hypothesis than any of those that were generated. These have **all** been hypotheses that differed from generated hypotheses in relatively minor ways. From the viewpoint of the domain, the performance is still adequate. It also appears that we should be able to extend the final pruning process to identify many of these situations and modify the hypothesis — in effect a local hill climbing process to make sure there is no better hypothesis in the region of the candidate.

7.1 The Algorithm for the Differential Diagnosis

With the approach and enhancements outlined above we will describe the implementation of the algorithm for the differential diagnosis.

1. The facts about the patient are entered via a menu system that supports both categorical and numeric values. These facts are turned into finding objects that become the terminal nodes for the diagnostic network.
2. The probabilities of the paths are computed, using the patient information to adjust the probabilities on the links.
3. The input is searched for definite consequences. If any node is definitely true or definitely false because of a finding, it is asserted along with any definite implications of these nodes. These definite nodes provide the first constraint on the hypotheses that can be generated.

4. All of the findings from the input list that potentially reflect abnormal states are selected to guide the hypothesis generation process. This list usually contains about ten to thirty items and is the list of findings that will need some kind of explanation. Each is either caused by a causal chain in the model or has some probability of existing independently (reflecting possible causes outside of the medical domain or occurrence of the finding in normal individuals).
5. For each of the abnormal findings and any true nodes that need causes, a search is made to find all of the diagnostic nodes with causal paths to the finding.
6. The diagnostic nodes and primary nodes that could account for findings are ordered by the number of findings they account for. The top node and any others that account for almost as many findings (almost is four fewer) are used as a seeds for generating hypotheses.
7. For each of the seed nodes, the list of findings and unaccounted nodes (which may include the seed node, if it is a diagnostic node that is not primary) is checked for ones without a causal path from anything in the hypothesis. These cause additional diagnostic or primary nodes to be added to the hypothesis until all of the findings have at least one possible cause. The additional nodes are selected from the list of possible diagnostic or primary nodes that could cause the finding.
8. The algorithm tries each explanation for the finding and adds each new covering set of causes that is not a superset of another generated from the same seed node.

For each of the cover sets, a hypothesis is built. The hypothesis starts with the cover set and the known true nodes.

9. First a search is made for states that would be true no matter which causal paths were chosen for the findings. These are added to the hypothesis.
10. The findings and unaccounted states are sorted in decreasing order of the difference in probability between the first and second best path, where those paths start at one of the cover set nodes and include no false nodes.

11. For each finding the best path is found from any state in the accumulating hypothesis and the states on that path are added to the hypothesis. If the prevalence of the finding is higher than the additional probability of the path, the path is not added. However, that finding will be tried again later to make sure there is not a good explanation for it in the hypothesis.
12. When the hypothesis is complete, the links among the nodes in the hypothesis are checked to make sure there are no circularities in the hypothesis. If there are, the last link before the loop is completed is marked as having probability zero.
13. The hypothesis is then pruned of unnecessary causes and unnecessary paths that decrease the probability.
14. The probability of the total hypothesis is determined by multiplying the probabilities of each node in the hypothesis, each node in the false set, and each finding given the truth of their causes in the hypothesis.
15. The hypotheses are ordered by their probability. The differential diagnosis list consists of the best hypothesis and any others with a probability greater than one percent of the best.

8 An Example

Figure 2 is the computer display of one of the hypotheses for an actual case as generated by the program. The figure shows the causal explanation that the program is proposing as the most likely explanation of the patient's findings. The findings entered by the user are shown in lower case and the nodes of the hypothesis are in upper case. The four nodes and one finding that are considered primary or diagnostic are printed in bold. The numbers in parentheses following the names indicate the probability that the item is true without being caused by anything else — the probability of being primary. The arrows indicate the most important causal relationships. (Ones with low probability are left out of the display if there is a higher probability link in the hypothesis. For instance, digitalis can also cause nausea/vomiting, but renal insufficiency is a more likely cause.) The numbers on the arrows are the probability that the cause will produce the effect. The probabilities from nodes to findings have been left out to keep the

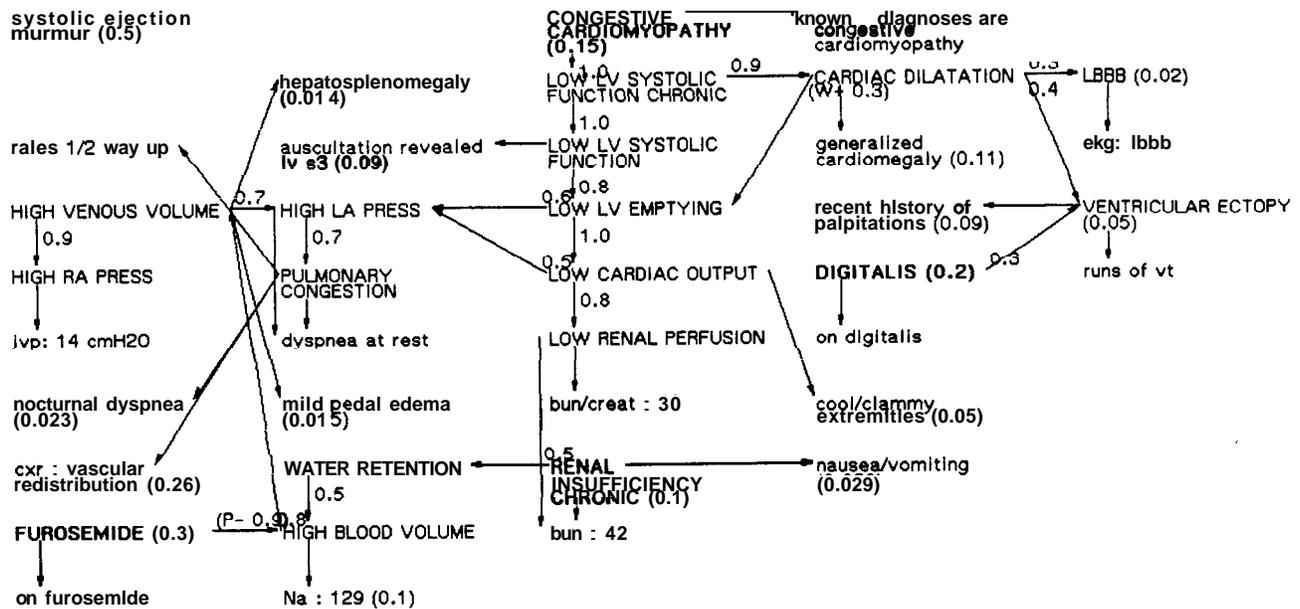


Figure 2: Congestive Cardiomyopathy and Renal Insufficiency Hypothesis

display readable. The links with a *W+* are worsening factors. For example, cardiac dilatation increases the probability of low left ventricular emptying, but can not cause it alone. Links with *P-* decrease the probability of the effect. For example, the use of furosemide (a diuretic) decreases the probability of high blood volume, but does not eliminate it. In this case there is direct and indirect evidence that the blood volume is still high.

There are several points to notice about the hypothesis. First, there are a number of findings that can only be explained by one node. For example, the finding that the patient is on digitalis or furosemide (therapies), causes the corresponding nodes to be true. Also, some test results always indicate the condition (at least for our purposes). In this case, left bundle branch block (LBBB) on electrocardiogram always indicates that LBBB exists. These findings cause the corresponding nodes to be asserted true. Second, the program will generate hypotheses with multiple causes. In this case renal insufficiency and congestive cardiomyopathy are used to account for the findings, even though it is actually possible to account for the findings with just congestive cardiomyopathy. Third, the program will find multiple paths to a node. There are paths to dyspnea at rest (shortness of breath) from both pulmonary congestion and high left atrial (LA) pressure. Pulmonary congestion has independent evidence and high LA pressure is needed to explain the pulmonary congestion.

One of the advantages of this method of generating the hypothesis is the detailed information

it provides about the internal nodes of the hypothesis — the mechanism by which the causes are producing the findings. In the medical domain, this information is useful both as a way for the user to check that the hypothesis is reasonable and as a way of identifying therapies that may be beneficial.

9 Summary

Because of the requirements of our medical domain, we have developed a causal model of cardiovascular hemodynamics that includes enough intermediate nodes to represent the mechanisms in sufficient detail to account for the actions of therapies. The causal nature of several different kinds of mechanisms in the model have been represented as probabilities and to uniformly account for potential circularities in the causal structure, we have introduced the notion of probabilistic causality. The requirements for a differential diagnosis in this domain led to the introduction of diagnostic nodes to capture the idea of significant differences between hypotheses. Finally, we have developed a heuristic method for generating differential diagnosis hypotheses based on causal paths from primary causes to the findings and the heuristic of building the hypothesis finding by finding. The mechanism has been tested on more than 40 patient cases and provides good hypotheses in less than a minute using the model which has about 150 internal nodes and about **300** terminal nodes.

References

- [Barnett87] G. Octo Barnett, James J. Cimino, J. A. Hupp, and E. P. Hoffer, "DXplain: An Evolving Diagnostic Decision-Support System," *Journal of the American Medical Association*, **285**: 67-74, 1987.
- [Cooper86] G. F. Cooper, "A Diagnostic Method That Uses Causal Knowledge and Linear Programming in the Application of Bayes' Formula," *Computer Methods and Programs in Biomedicine*, **22**: 223-237, 1986.

- [Cooper87] G. F. Cooper, "Probabilistic Inference Using Belief Networks is NP-Hard," Technical Report KSL-87-27, Medical Computer Science Group, Knowledge Systems Laboratory, Stanford University, May 1987.
- [First85] Michael B. First, Lynn J. Soffer, and Randolph A. Miller, "QUICK (QUick Index to Caduceus Knowledge): Using the Internist-1/Caduceus Knowledge Base as an Electronic Textbook of Medicine," *Computers and Biomedical Research*, **18**: 137-165, 1985.
- [Lauritzen88] S. L. Lauritzen and D. J. Spiegelhalter, "Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems," *Journal of the Royal Statistical Society B* 50: 157-224, 1988.
- [Long88] W. J. Long, S. Naimi, M. G. Criscitiello, R. Jayes "Using a Physiological Model for Prediction of Therapy Effects in Heart Disease" *Proceedings of Computers in Cardiology Conference*, pp. 15-20, October 1986.
- [Pearl86] Judea Pearl, "Fusion, Propagation, and Structuring in Bayesian Networks," *Artificial Intelligence*, 29: 241-288, 1986.
- [Weiss78] Sholom M. Weiss, Casimir A. Kulikowski, Saul Amarel, Aran Safir, "A Model-Based Method for Computer-Aided Medical Decision-Making," *Artificial Intelligence*, **11**: 145-172, 1978.