Reprinted from *Artificial Intelligence in Medicine*, **8:** 193-215, 1996

Temporal Reasoning for Diagnosis in a Causal Probabilistic Knowledge Base

William Long
MIT Lab for Computer Science
545 Technology Square, Room 420A
Cambridge, MA 02139, USA
wjl@mit.edu
FAX: 617-258-8682
Phone: 617-253-3508

Abstract

We have added temporal reasoning to the Heart Disease Program (HDP) to take advantage of the temporal constraints inherent in cardiovascular reasoning. Some processes take place over minutes while others take place over months or years and a strictly probabilistic formalism can generate hypotheses that are impossible given the temporal relationships involved. The HDP has temporal constraints on the causal relations specified in the knowledge base and temporal properties on the patient input provided by the user. These are used in two ways. First, they are used to constrain the generation of the pre-computed causal pathways through the model that speed the generation of hypotheses. Second, they are used to generate time intervals for the instantiated nodes in the hypotheses, which are matched and adjusted as nodes are added to each evolving hypothesis.

This domain offers a number of challenges for temporal reasoning. Since the nature of diagnostic reasoning is inferring a causal explanation from the evidence, many of the temporal intervals have few constraints and the reasoning has to make maximum use of those that exist. Thus, the HDP uses a temporal interval representation that includes the earliest and latest beginning and ending specified by the constraints. Some of the disease states can be corrected but some of the manifestations may remain. For example, a valve disease such as aortic stenosis produces hypertrophy that remains long after the valve has been replaced. This requires multiple time intervals to account for the existing findings.

This paper discusses the issues and solutions that have been developed for temporal reasoning integrated with a pseudo-Bayesian probabilistic network in this challenging domain for diagnosis.

Keywords: temporal reasoning, causality, Bayesian probability networks, physiologic causality, constraint reasoning, diagnosis, heart disease

1 Introduction

Temporal reasoning is a tool to enhance other types of reasoning. Many reasoning tasks such as planning, understanding, or diagnosis, have an aspect of time. The temporal reasoning

enhances the performance of the task by applying the particular characteristics of time to the problem. The most prominent characteristic of time is the one-way flow with *now* occupying a single point. Such tools as Allen's formalization of temporal reasoning[1] give a functional understanding of such temporal properties as *before* and *after*. This paper will take a closer look at temporal reasoning in the service of medical diagnosis. Others have also incorporated temporal reasoning into medical programs to a varying degree. Kahn recently provided a review of such efforts[7].

1.1 Diagnosis

Diagnosis is determining the cause of a pathological state. Often medical diagnostic programs simply generate a list of diseases that might account for the given findings[2, 5, 12]. Whether this is an adequate diagnosis depends on the purpose of the diagnosis. Most diagnosis is done in order to treat the patient. If the name of the disease is sufficient guidance for treatment, the job is done. However, in many domains more detailed information is needed about the mechanisms involved in producing the observed manifestations, the complications, severity, and so forth.

In such complex domains with multiple interacting mechanisms leading to the observed state, diagnosis can be better characterized as the reconstruction of the likely scenario that produced the observed state. The essential features of the scenario are that it adequately and consistently explain the observations in terms of the mechanisms involved at a level of detail sufficient to guide the further management of the patient. Indeed, it is sometimes true that only knowing the immediate mechanisms without the ultimate etiology is sufficient to guide therapy and the diagnostic process need not determine the name of the disease. Determining the scenario has the additional benefit of providing justification that the disease or diseases postulated do indeed account for the findings. That is, the scenario is a detailed argument for the consistency and completeness of the hypothesis and can be critiqued by the physician user. This solves one of the problems of associational diagnostic reasoners in that they tend to find diseases with some matching findings for which the overall combination of findings is inconsistent with the disease[3].

In a domain with a significant degree of uncertainty, diagnosis should be further characterized as the process of determining a set of one or more possible scenarios covering the range of likely explanations for the observed findings. In such a domain, diagnostic reasoning is a step in an iterative process of further measurement and testing, terminating when there is sufficient information on which to base treatment. This process of generating a set of hypotheses, called the differential diagnosis, each element of which is a possible explanation for a given set of findings is the task of the Heart Disease Program.

2 Heart Disease Program

The Heart Disease Program (HDP)[1] is a computer program to assist the physician in the management of patients with complex cardiovascular disorders. The diagnostic aspect of the management process is the most developed part of the program and the only part of concern

[1] In earlier papers the program has been referred to as the Heart Failure Program, but the domain is much broader than heart failure, so Heart Disease Program is a more appropriate name.

in this paper. The program's diagnostic reasoning task is to take the findings reported by the physician and put together a differential consisting of the most likely hypotheses accounting for the findings. Each hypothesis consists of a complete causal explanation detailing how the diseases and mechanisms in the hypothesis provide a consistent accounting for the findings. The program has gone through two evaluations, one when the reasoning was based entirely on a probability network[10] and a more recent evaluation after temporal and severity constraints were incorporated into the reasoning[11].

2.1 Problem Domain

The domain of the HDP is particularly challenging because it involves multiple interacting mechanisms operating over a variety of time periods. Furthermore, the available observations are limited, requiring significant reasoning to ferret out what is taking place. The domain consists of those disorders that cause or complicate hemodynamic dysfunction in the patient. When for some reason the heart is not able to pump as much blood as the body requires, a set of compensatory mechanisms are set in motion which tend to maintain the blood pressure and increase the blood volume. To accomplish this the body constricts blood vessels, selectively maintaining blood pressure to the heart and brain and decreasing blood supply to the kidneys and less critical organs. While these mechanisms are very effective in the patient with a normal heart who has lost blood, they can be counterproductive when the heart muscle has been weakened by disease. For example, the blood pressure filling the ventricles (the heart's primary pumping chambers) from the atria increases to help the ventricles maintain cardiac output. However, the increased left atrial pressure causes an increase of back pressure in the lungs and ultimately fluid in the lungs called pulmonary congestion. The lung congestion, fluid accumulation throughout the system, and increased stress on the heart presents a characteristic pattern called congestive heart failure.

The diseases that cause such hemodynamic dysfunction include diseases of the heart muscle such as myocardial infarction (heart attack) and several kinds of cardiomyopathy, valvular dysfunction, and restriction of the heart by the pericardium. There are also a number of diseases that decrease the effectiveness of a healthy heart, including hypertension, pulmonary hypertension, anemia, pulmonary disease, and renal disease. Finally, some diseases simulate the effects of hemodynamic compromise, such as liver disease. Each of the diseases has particular characteristics and findings that differentiate it from the others even when most of the findings are similar. These diseases and the mechanisms by which they produce hemodynamic compromise are the domain of the HDP.

Most of the cardiovascular disorders of concern in this program are chronic, progressive, and many can not be corrected short of a cardiac transplant. As a result, patients typically arrive with existing diseases and existing therapies. The problem is to determine what new diseases or complications are now present and their relationship to the known diseases. Thus, the therapies with both their beneficial effects and side effects are an important part of the domain.

2.2 Diagnosing Heart Disease

Having characterized the diagnostic problem as generating causal physiologic explanations for the given findings, the computational mechanism with the best fit to generate those

explanations is a Bayesian probability network. To first approximation, the best hypothesis or explanation is the subset of the network that is true in the maximum likelihood state of the network. This computational characterization of the problem assumes that the links in the network, which represent the conditional probabilities in the domain, can also represent the causal relations needed to characterize a scenario accounting for the findings.

In the heart disease domain, this fit has some problems. First, there are situations in which the appropriate causal characterization is that A can cause B and (perhaps with intervening links) B can cause A. Keeping the Bayesian network faithful to the sense of causality results in forward loops in the network, which are inconsistent with the mathematics of a Bayesian network. Because of this, the HDP uses a pseudo-Bayesian probability network. The knowledge base has forward loops in it, although any particular hypothesis does not. To reason with such a network, heuristic methods are necessary[9].

A probability network assumes that nodes are completely characterized by their truth (the conditional independence assumption that gives the network its power). That is, if the node is true (or has one of a small fixed set of values), it isolates its causes from its effects unless there are other paths between them. Thus, a node only needs to know about its immediate causes. Unfortunately, this assumption is false if links are intended to represent causality. For example, if low cardiac output has only been true for a few hours, its effects, whether immediate or reached through a number of causal steps, can only have been true for a few hours. In other causal relationships, it takes time for effects to develop and the duration of the cause may rule out effects further down the causal chain. This problem could be solved by having multiple low cardiac output nodes representing different periods of time. However, in the heart disease domain this proliferation of nodes would have to happen over the whole model, increasing the size and complexity of the model enormously.

The strategy of duplicating nodes to represent different times has been successfully applied in the domain of diabetes therapy. In that domain the combination of diagnostic and temporal reasoning has been handled by having a copy of the Bayesian network for each hour over a 24 hour period[6]. In the heart disease domain, there are no well-defined convenient time periods to divide up the past, since minutes, hours, days, and years are often pertinent to the reasoning. Even if 20 or 30 suitable time periods could be devised, a model with a couple hundred nodes for each time period would make the reasoning intractable at current computational speeds.

For this reason we have added temporal relationships as constraints on the probabilistic network. For example, if the cause for low cardiac output is an acute MI that took place over the past four hours, the temporal constraints also determine what effects low cardiac output and anything further down the causal chain can have. Since we are already using heuristic methods to reason with the probability network, the addition of temporal reasoning does not compromise the mathematical integrity of formal methods.

3 Example Problem

The kind of situation that calls for temporal reasoning is illustrated by the following example, which will be used as a running example for the rest of the paper. In an actual (and quite typical) case diagrammed in figure 1, a patient was admitted to the Emergency Department with chest pain of an hour duration beginning two hours prior and was given nitroglycerin.

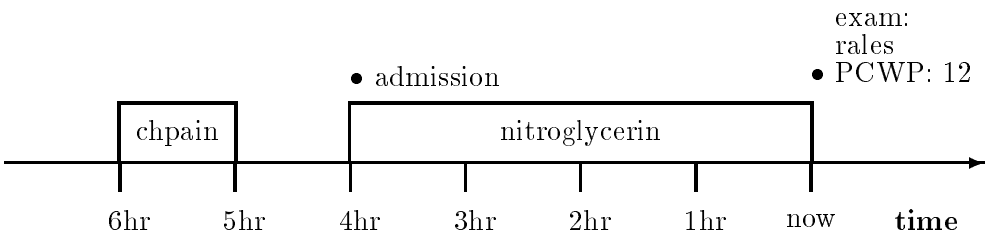

Figure 1: Scenario of Patient Presenting with Chest Pain

Four hours later, when the data was entered into the computer, the patient had a Swan-line in place providing cardiac pressure information. At that time, the patient had rales in the chest examination (fluid in the lungs) and a pulmonary capillary wedge pressure (PCWP) of 12, indicating a normal left atrial pressure.

In this patient, there is adequate evidence from the chest pain (and other findings which we will ignore for simplicity) indicating a myocardial infarction (MI). It can also be concluded from the rales that the patient has pulmonary congestion. An MI can cause pulmonary congestion (figure 2) by poor left ventricular function (LVF), which elevates the left atrial pressure (LAP), causing fluid to accumulate in the lungs. Prior to temporal reasoning, the HDP had difficulty accounting for the rales because the PCWP of 12 indicated that the LAP was normal, breaking the causal pathway from MI to pulmonary congestion. As a result, the program proposed pneumonia as an explanation for the rales.

Because pneumonia can also cause pulmonary congestion, this is a reasonable hypothesis, but it is not the best hypothesis. The nitroglycerin decreases the LAP as well as improving blood flow in the myocardium. Since the rales can take many hours to go away after the LAP has returned to normal, a better explanation is that the rales were caused by the MI, but the nitroglycerin has now decreased the LAP and the rales have not had time to clear. A third explanation is that the MI only transiently decreased the LVF. While the rales have not had time to clear, the causal mechanism is no longer present. Thus, there are three reasonable mechanisms to account for the findings, each with different implications for treatment, but unless the reasoning considers the time relationships only one explanation can be generated.

4 Desiderata for a Temporal Representation

To handle such problems, we have developed a mechanism for temporal reasoning. The essential functions of the temporal reasoning are to deduce and maintain the causal temporal constraints in hypotheses and to support the possibility of nodes having different values over different time intervals. In the example, the temporal reasoning must allow the LAP to have a value of *high* for some period after the MI but also be *normal* by the time of the examination. In addition, it must enforce the restriction that the time interval of pulmonary congestion is after the MI if it is to be explained by the MI. Thus, the temporal reasoning must support the logical constraints of the causal theory allowing some situations that would be ruled out in a purely probabilistic network and disallowing some situations that would satisfy such a

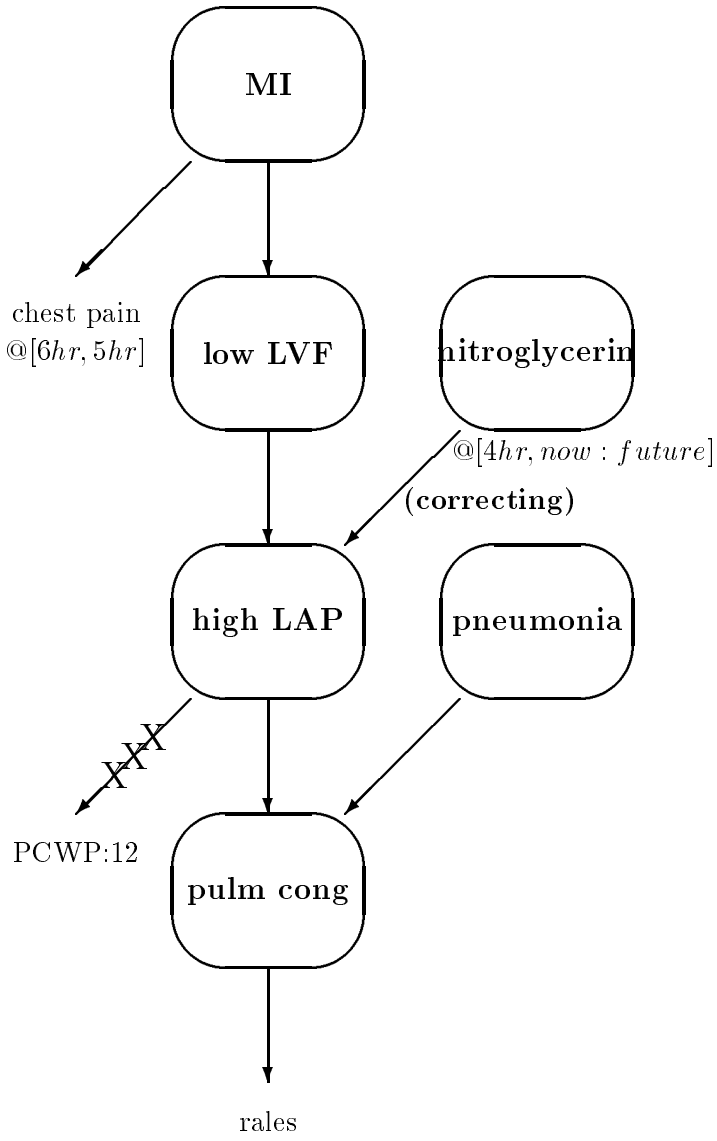

Figure 2: Causal Structure for Example

network.

Besides the logical constraints provided by the temporal model, the temporal characteristics provide an opportunity to adjust the probabilities when there are different causal situations with different probabilities. For example, if the depression of LV function is sometimes transient but often continues for days, it is appropriate to use a different probability for a causal link involving LV dysfunction a day after an MI than for a link involving LV dysfunction during an MI.

4.1 Causality in the Heart Domain

There are a small number of patterns of causation that the program needs to capture. These can be characterized as follows:

Immediate This takes place immediately if it happens at all. For example, most valvular lesions cause a murmur and if so, it is there when the lesion is present.

Event-like This happens as a result of pathophysiological states that predispose the patient to the entity. For example, a patient with coronary artery disease can have an MI any time with no particular time relationship. Perhaps some additional stress actually triggers the MI, but in practical terms the immediate cause is rarely known and not important.

Delayed This is like immediate causation except that there is some definable time interval between cause and effect. For example, a patient with an MI may develop pericarditis within the first two weeks after the MI, usually more than two days after the event.

Progressive This is causation that, once it takes place, continues. Progressive states persist chronically and just get worse unless radical correction takes place. For example, rheumatic heart disease often causes mitral stenosis (restriction). The mitral stenosis will gradually get worse until valve repair or replacement is done.

Accumulative This requires the cause to exist over a period of time. For example, the fluid accumulation that produces pedal edema (swelling of the feet) is caused (indirectly) by low cardiac output. The cardiac output does not have to be consistently low, but if it is low on average for long enough there will be enough water retention to produce the pedal edema.

Intermittent When a cause is present some effects and findings occur intermittently. For example, paroxysmal atrial fibrillation is a rhythm disturbance that may persist for seconds to days and may happen again at random intervals. Often it gradually increases in frequency and eventually becomes chronic atrial fibrillation.

Corrective Some states return other states to normal. These are usually therapies, but in some cases they are other pathophysiologic states that have influences in opposition to the effect, such as dehydration 'correcting' high blood volume.

Besides these characteristics of causation, there are some characteristics of the states that have implications for temporal reasoning. States produced by accumulative causation and

sometimes by delayed causation tend to remain after the cause has ended since the same mechanism tends to delay the return to a normal state. The progressive and accumulative descriptions could also be viewed characteristics of the state, since once they start, their continuation is independent of considerations of what caused them. Some states have a maximum time period. For example, stress of many kinds including an MI cause an increased sympathetic nervous state with sweating, rapid heart rate, and so forth, but this only lasts for a few days at most, even though the cause may continue. Other states only last for a period of time, not because they are self-limiting but because they are either corrected or the patient does not survive. These characteristics constrain the time bounds of causes and effects and what conclusions can be drawn from findings.

It should be noted that these characterizations are not mutually exclusive. Delayed states can be progressive or intermittent and so forth. Probably the most complicated causal pattern in the domain is that of an MI. An MI is usually caused by coronary artery disease but the time is random, without any triggering event in the majority of cases, and having an acute onset. Thus, it is the archetype for event-like causation. Over the first few hours, the MI causes chest pain, signs of acute ventricular dysfunction, and signs of sympathetic nervous response. Over the first two days, the enzyme changes are evident. Over the first two weeks the electrocardiographic patterns change into ones characterized as evolving and the patient may experience pericarditis. Often the damage to the heart is permanent, producing chronic ventricular dysfunction and the electrocardiographic pattern of an old MI.

This temporal characterization of the heart domain may not be adequate for other medical domains, but it will probably be nearly so. One obvious temporal pattern that has been left out is cyclic phenomena. Those do happen in this domain. For example, Cheyne-Stokes respiration is an oscillation between slow and rapid breathing that sometimes happens with older patients in heart failure. However, all such phenomena in this domain are recognized as entities in themselves and are treated in the same way as other findings. That is, the cyclic nature of the phenomena is not something the program needs to reason about.

5 Representation of the Temporal Relationships

To represent these causal patterns, we need a number of additional properties on node and link descriptions in the knowledge base. Three of these are time ranges: *onset, delay,* and *persistence.* Persistence is the range of time the state might remain true after the cause ceases. Thus, progressive states can be represented as having infinite persistence. Accumulative states have a *persist* property that gives the range of time needed for the state to resolve. For example, pedal edema can remain for days to weeks after the cause has been removed, dilatation of the heart takes weeks to many months to return to normal. Immediate causes have a persistence of zero.

Onset and delay refer to the initiation of a state. Accumulative states require time to start. This is captured by the *onset* property. The value is either zero for immediate states, or the range of time it takes for the state to become apparent. One might argue that even though it takes a day before one could observe pedal edema, the swelling was actually there earlier. However, the concern is on the relationship between states and observed findings, so requiring a time for onset makes sense even in such a case. Because the onset of the state is often a time in which the effects of the state may start to develop as well, the onset period

is treated as part of the time of the state for reasoning about effects. The delay is the onset plus any additional time between the beginning of the cause and the beginning of the effect. Thus, if an effect has both a *delay* property and an *onset* property, the time between the cause and the effect is the delay. When the effect is observed it is considered to have been in existence for the onset time.

The most common type of delay is reflected in event-like causation. The delay of events is zero to infinity relative to their cause. Some of these also have an onset. For example, anemia can be caused by renal insufficiency. The delay is arbitrary but the onset is sufficiently gradual that it should be reasoned with as if it had been there for a week when it is observed. Other than events, it is unusual for a causation to have a delay in addition to an onset that makes any difference for diagnostic reasoning. The clearest instance is constrictive pericarditis. Pericardial calcification takes place over many months and does not have any effect until it starts to restrict the heart's filling capacity. At that point there are many effects that are observable.

The remaining properties are the *max-exist* (maximum existence time) and two binary properties, *intermittent* and *self-limiting*. The maximum existence is the maximum length of time a patient would stay in that state, even though the cause continued. The *self-limiting* property says that the state will return to normal even if no correcting state is present. For example, high sympathetic states only last a short time and return to normal (normal with respect to clinically important manifestations) without any therapy being directed at the state. The *max-exist* property is also needed for states that are not self-limiting when the continuation of the state is not compatible with life. For example, there are no effects that would be caused by months of septic shock because no patient would survive that long. Either the shock is successfully treated or the patient dies within a few days. The property *intermittent* implies that the state or finding does not always have to be observable. For example, many arrhythmias are intermittent and not observing them during a particular examination does not rule them out or mean that they can not have any effects.

In summary, the representation of causality in the knowledge base requires the following properties:

Onset The range of time that can be assumed for the effect when it is observed.

Delay The range of time the cause must be true before the effect can start. This includes the onset time.

Persist The range of time that the effect will remain if the cause ceases to be true.

Max-exist The maximum time the effect will remain, even though the cause continues.

Self-limiting Whether the state will cease by *max-exist* without a corrective node being true.

Intermittent Whether the state can be absent over subintervals of the interval in which it is true.

The relationships among these times are diagrammed in figure 3. The rules for applying these properties are as follows:

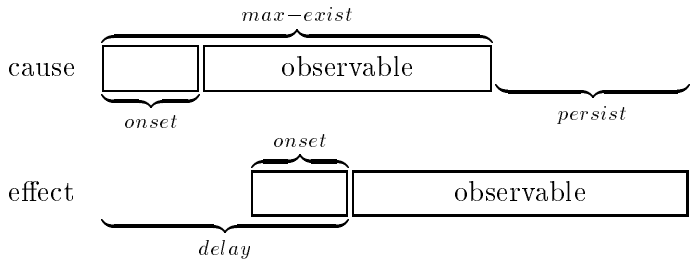

Figure 3: Relationship of Causality Times

1. When a node is observed, it is assumed to already be producing effects for *onset* time.
2. Effects are observable at a time after the cause given by the *delay*, if it exists, otherwise by the *onset*.
3. Effects are observable after the cause is observable and overlap the cause.
4. Effects continue until the cause ceases, unless the *max-exist* is exceeded or the effect is *intermittent*.
5. Effects continue after the cause ceases, in accordance with the *persist*.

These rules follow from the discussion above except for the third rule. This additional rule encompasses two practical considerations. First, causes have manifestations before (or contemporaneous with) their effects. From the representation it would be possible for a cause to have a longer onset than its effect making the effect observable before the cause. However, what it means for a state to be observable is that it has manifestations, so such a situation would not make sense. As a practical consideration, if a cause develops slowly, the effect will also develop slowly. Computationally, this is a useful constraint because it provides tighter constraints on what causes can produce a particular effect. Secondly, the cause and effect must overlap. In other words there are no remote causes. Computationally, this is also a useful constraint because it means that the effect must start before the cause ends regardless of how long a delay could otherwise exist. Physiologically, it is a matter of perspective. In situations where there appear to be remote causes, there is some underlying mechanism that covers the time period, although it may have fewer observable findings. For example, when an MI causes pericarditis a week later, there is an underlying process of myocardial modification going on that may continue for several weeks.

Besides causal links to states, there are also correcting links to many of the states. These are primarily the therapies that can counter the states, The model used in this program for the effects of correcting influences is the same as that for causes. That is, a therapy may take a period of time (onset) before it produces the desired effect. We have not come across sufficient reason to use persistence or other aspects of the representation to model the effects of the correcting influences. This also is a matter of perspective. One could model the effect of a surgical procedure, such as a valve replacement, as a therapy taking a short time with an

infinite persistence, but it is just as easy to handle it as a therapy that always remains true. Most drugs have some persistence (the pharmacokinetic half-life), but that can be considered part of the treatment time.

With this information it is possible to compute the time limits either for a cause when the effect is known or an effect when the cause is known. When a node is instantiated for a case, it is given a temporal interval, representing the observable time of the node. These temporal intervals have earliest and latest beginning and ending times, similar to the representation used in the CHECK system[4]. These intervals are used in a number of ways. The findings attributed to the node and the causes contributing to it are used to refine the limits of the interval as a hypothesis is built. To determine the causal pathways to be added to the hypothesis, the intervals are used to determine consistency. To speed this process, the overall temporal constraints on causal pathways are pre-computed in the model. This alone eliminates as inconsistent about 20% of the pathways that were computed in the purely probabilistic model. In the context of a specific node or finding, the probability along a pathway from a known node or primary node is recomputed using the more specific temporal information from the case. Thus, all of the nodes added to a hypothesis are consistent with the temporal constraints of causation. This time mechanism combines the ideas about reasoning with time outlined in our earlier paper[8] with the probabilistic reasoning described more recently[9].

5.1 Representing the Example

The following is a knowledge base fragment sufficient to cover the example MI and nitroglycerin problem. Not all of the temporal parameters are needed in this example, but it will give the flavor of the temporal reasoning. The definitions have been simplified to keep the example understandable:

```
(defnode MI
   full-name myocardial-infarction
   caused-by (primary prob 0.001)
   persist (time 0 1hr 0.5 1day)
   measure (chest-pain prob 0.9))
```

For purposes of this example, there is no cause for an MI. It just has a probability of 0.001 of being true in a patient. The persistence time is the time from the end of a cause being true to the end of the node being true. Since in this case the cause is random chance, the persistence is the same as the length of time the condition is true. The *time* clause indicates that the MI is never over in less than an hour (probability 0) and always over if more than a day has gone by. If the time is between, there is a 0.5 probability that the MI is over. This is clearly an approximation since the length of an MI is actually a smooth curve, but this is sufficient. For simplicity, this ignores everything but the most acute effects of an MI. The *measure* clause has the possible findings for the node. In this case an MI has a 0.9 probability of producing chest pain. There are a number of other possible findings, but this is sufficient for the example.

```
(defnode (low LVF)
```

```
   full-name (low left-ventricular-systolic-function)
   caused-by (MI prob 0.6)
   persist (time 0.3 6hr 0.8 inf))
```

An MI causes low left ventricular systolic function with a probability of 0.6. 30% of the time the low LVF returns to normal immediately or within six hours after the end of the MI. After six hours 80% of the cases of low LVF have returned to normal and the rest are permanently impaired. There is no direct measure of LVF for the example.

```
(defnode nitroglycerin
   caused-by (therapy))
```

Nitroglycerin is a therapy. Since therapies are under the control of the physician there is an implicit measure in the therapy history.

```
(defnode (high LAP)
   full-name (high left-atrial-pressure)
   caused-by ((low LVF) prob 0.8)
   corrected-by (nitroglycerin prob 0.8)
   measure (PCWP prob (range 0 15 .1 18 .9)))
```

High LAP is caused by low LVF with a probability of 0.8. It can be corrected by nitroglycerin and the probability that a LAP that is high will be brought down to normal by nitroglycerin is 0.8. The measure for LAP is the PCWP. Since that is a continuous variable, the *range* statement breaks it into ranges in which the probability can be specified. A PCWP less than 15 never indicates a high LAP. 10% of high LAPs will produce a PCWP between 15 and 18 and the rest will produce a PCWP above 18.

```
(defnode pneumonia
   caused-by (primary prob 0.01)
   persist (time 0 2day 0.5 2week))
```

Pneumonia is another primary node, with a probability of 0.01, that takes 2 days to 2 weeks to return to normal.

```
(defnode PC
   full-name pulmonary-congestion
   caused-by ((high LAP) prob (onset now 0.2 1hr 0.5 6hr 0.8))
             (pneumonia prob (onset 2day 0.5))
   persist (time 0 6hr 0.5 1day)
   measure (rales prob 0.9))
```

Pulmonary congestion can be caused by either high LAP or pneumonia. The high LAP may take up to 6 hours to produce the congestion if it is going to produce it at all. In 20% of cases the congestion occurs in less than an hour. Pneumonia takes 2 days before half the cases produce pulmonary congestion. Persistence on the other hand is always between 6 hours and a day.

There must also be knowledge about the findings including how often they might be falsely positive, but we will assume for now that they are perfect.

5.2 Findings

The source of temporal information about a case is the user input. From the viewpoint of temporal properties, findings can be grouped into four classes: observations, symptoms, history, and tests. The observations include the results of the physical exam. These only provide information about what is true at the time of the exam and say nothing about how long the findings might have been true. The symptoms are typically reported by the patient and include time information. For example, the patient may say that the chest pain started two hours ago and lasted for an hour. In practice descriptions can be much more complex, such as complaining of shortness of breath that only occurs at night. However, these more complex descriptions can either be handled as associations of findings or as specialized findings with names of their own. Important associations include descriptions such as having palpitations with shortness of breath. The shortness of breath at night is called paroxysmal nocturnal dyspnea or PND. History information has essentially the same temporal properties as symptoms. Patients could have coronary artery disease for five years or have had an episode of endocarditis a year ago. Tests occur at a specific time. They are essentially observations except that often they were done at a time in the past and, unlike physical exam observations, will not be repeated unless there is some specific need. For example, an echocardiogram done five years ago may provide useful information about the patient's present condition. Observations of the past are either summarized as patient history or are considered irrelevant to the current situation.

Thus, the user can provide the system with the appropriate temporal information with a small number of additional capabilities in the input interface. Observations are assumed to refer to the current time. Symptoms and history need to have durations, event times, and pertinent associations. Tests need to have event times. With these attributes it is possible in the heart disease domain for the user to provide the pertinent temporal details for diagnosis.

In the example, the user provides the following input:

> chest pain: anginal at rest 6hr ago for 1hr, therapy: nitroglycerin for 4hrs, chest: rales, PCWP: 12 now

Given this input, the case provides the following facts: (In the following discussion we will use the convention that times such as $6hr$ will mean 6 hours in the past and *now* refers to the current time.)

- chest pain @$[6hr, 5hr]$ (chest pain 6 hours ago, lasting an hour)
- nitroglycerin @$[4hr, now]$ (nitroglycerin from 4 hours ago until now)
- PCWP = 12 @now (PCWP is currently 12)
- rales @now (patient currently has rales)

The causal relations and the initial facts are shown graphically in figure 4. The nodes are shown with initial probabilities at the top and the persistence time at the bottom. The links have the causal probabilities.

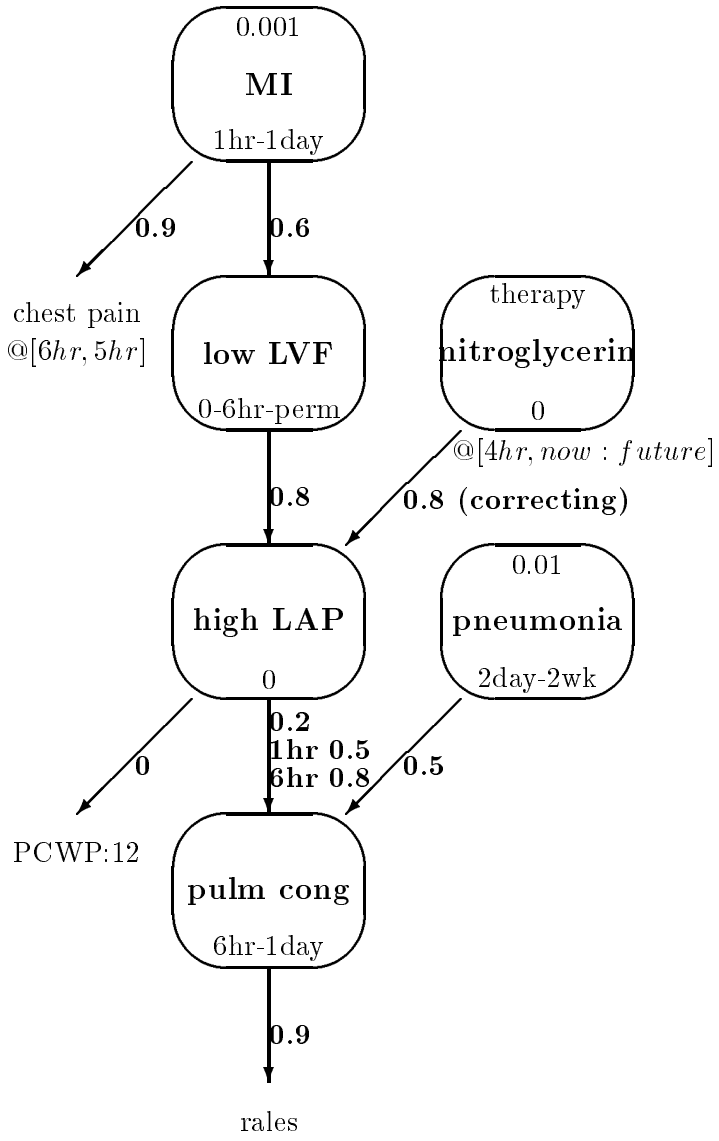

Figure 4: Causal Relations in Nitroglycerin Case

6 Time Intervals

To represent this input information and the diagnostic conclusions we need a representation for the temporal properties of the instantiated nodes. This is accomplished by representing the truth of a finding or node over a temporal interval. For example, the chest pain is true over $[6hr, 5hr]$. Unfortunately, with varying delays, onsets, and persistences the diagnostic conclusions from the findings are not as definite and require indefinite time bounds to specify the ends of the intervals. For example, if pneumonia is responsible for the pulmonary congestion, the pneumonia could be true now or could have ended in the last day, since it may take up to a day for the rales to clear. Furthermore, the pneumonia must have started within the last two weeks, otherwise it would be over by now. It must have started a couple of days ago for the effects to be present now. Deductions such as these can be captured by time intervals that have four time parameters: earliest beginning, latest beginning, earliest ending, and latest ending. For the pneumonia this can be represented as: $[2wk : 2day, 1day : future]$.

This representation of time is not completely sufficient for reasoning because it loses some information. In general, it is not possible to determine the minimum and maximum extent of the node from the interval. In the case of pneumonia, it lasts two days to two weeks. This information is needed to rule out certain findings as effects, or in the case of other nodes to rule out possible causes. Since the interval refers to a node or finding and that information is already in the knowledge base, including it in the temporal interval structure is a matter of computational convenience rather than necessity. Another such piece of information is the causal relationship between nodes. If the interval only places the node in the last week and another node has similar bounds, it is not possible to tell whether one can be the cause of the other. Such questions must be answered by deduction from the causal networks of the two nodes. Thus, the four parameter time interval representation provides the information needed to carry out the temporal reasoning.

In diagnosis the most common reasoning step is to infer a cause from an effect. Given the time intervals and the node representations, the determination of the time interval of the cause proceeds as follows:

$$\begin{aligned} be_c &= be_e + \lceil d_e \rceil - \lceil o_c \rceil \\ bl_c &= bl_e + \lfloor d_e \rfloor - \lfloor o_c \rfloor \\ ee_c &= ee_e + \lceil p_e \rceil \\ el_c &= el_e + \lfloor p_e \rfloor \end{aligned}$$

Where be is the earliest begin time, bl the latest begin time, ee the earliest end time, el the latest end time, d the delay (which includes the onset time), o the onset, and p the persist time. The subscripts are c for the cause and e for the effect. Since delay, onset, and persist are ranges, the maximum and minimum delays are used as appropriate. The reason the maximum and minimum onsets of the causes are used, instead of the opposite, is the observation that slow causes produce slow effects. The times are all temporal distances before the current time, so $1day + 1day$ is two days prior to the reference time.

These time interval values must be modified to account for the max-exist of the cause, if there is one. This is a further constraint that can make the earliest begin time later (shorter time) or the latest end time earlier. If x is the max-exist, then $be - ee \leq x$ and $bl - el \leq x$. Therefore, $be_c \leq ee_c + x_c$ and $el_c \geq bl_c - x_c$. The additional constraint that observation of

cause precedes observation of effect implies that $\lceil d_e \rceil \geq \lceil o_c \rceil$ and $\lfloor d_e \rfloor \geq \lfloor o_c \rfloor$. That they are overlapping implies $ee_c \leq be_e$ and $el_c \leq bl_e$. Once the parameters of the time interval have been adjusted for these constraints, it is as exactly specified as is possible.

7 Drawing Conclusions in the Example

The first step in the diagnostic process is to draw any definite conclusions from the input. In the example, these are:

- chest pain @$[6hr, 5hr]$ ⇒MI @$[6hr, 5hr]$
- PCWP = 12 @now ⇒(high LAP)=false @$[past : now, now : future]$
- rales @now ⇒PC @$[past : now, now : future]$
- nitroglycerin @$[4hr, now : future]$

Since the chest pain is a symptom with immediate effect and no persistence, the time of the chest pain determines the time of the MI. Not all MI's have chest pain, so this is only true when the MI causes chest pain. The rales are an observation on physical exam, so it does not determine the time extent of the PC. All we know from the rales is that the PC started before *now* and will end after *now*. The PCWP is a test which happened now, and thus is just like an observation.

7.1 Generating A Hypothesis

Once the definite conclusions have been drawn, the HDP looks for findings and nodes that need explanations and considers causal pathways that might explain them. In the example, there are four nodes with known states. The MI node is true over a time interval, but it is primary (in this limited knowledge base) and needs no explanation. Similarly, the nitroglycerin is a therapy and needs no explanation. The high LAP node is known to be false now, which is its normal state, so no explanation is necessary. The only node that needs an explanation is the pulmonary congestion (PC) node.

The HDP uses the unexplained nodes to generate a list of possible hypotheses. In this case the two primary nodes that could cause PC are MI and pneumonia. The pneumonia causes PC directly and the MI causes it through the causal chain of low LVF and high LAP. The HDP tries to generate each of these hypotheses in turn.

Pneumonia Hypothesis The pneumonia hypothesis requires a time interval for causing PC, determined from the properties on PC and pneumonia. The PC onset is 2day from the causal statement. PC maximum persist is 1day from the persist clause. Since the max-exist for a primary node is the same as the persist, the max-exist for pneumonia is 2wk. Using these facts and the rules specified above determines the time interval for pneumonia:

- PC @$[past : now, now : future]$ ⇒
- $be = past + 2day = past$

- $bl = now + 2day = 2day$
- $ee = now + 1day = 1day$
- $el = future + 1day = future$
- max-exist: $be \leq 1day + 2wk \Rightarrow be = 2wk$ (within the time resolution)
- $\Rightarrow$pneumonia $@[2wk : 2day, 1day : future]$

Since pneumonia is primary there are no further causes. This corresponds to the explanation that is generated without a time representation. The additional provision that the pneumonia could actually have ended within the last day was not captured before.

MI Causing PC Hypothesis For the MI to cause the PC would have been ruled out by the normal LAP without the time relations. Temporal reasoning makes this hypothesis possible. Since the onset for high LAP is zero and the max-exist is unspecified, the only causal constraint is the persistence of PC. Applying the rules as above:

- PC $@[past : now, now : future] \Rightarrow$
- high LAP $[past : now, 1day : future]$

However, this time interval is constrained by the existing time interval in which high LAP is false. Multiple time intervals for a node have to have a strict time ordering. By default they are assumed to be abutting unless there is some reason to infer an additional change in between. These time intervals are reconciled by adjusting the early begin and late end times:

- high LAP $[past : now, 1day : future]$ then normal LAP $[past : now, now : future] \Rightarrow$
- high LAP $[past : now, 1day : now]$ then normal LAP $[1day : now, now : future]$

The result captures the fact that the LAP is now normal but was high within the last day.

The high LAP needs further explanation, since it is not a primary node. There are two possible explanations: either the low LVF ended before now, or the high LAP was corrected by the nitroglycerin.

Low LVF Ended Hypothesis Since the high LAP is an immediate effect and low LVF has no *max-exist* clause, the time interval of low LVF is the same as that of the high LAP.

- low LVF $@[past : now, 1day : now]$ (low LVF starting before now and ending before now but not more than 24 hours before now)

This in turn is caused by the MI, which already is known to be true over the time interval $[6hr, 5hr]$. Since effects can not precede causes, this further constrains the low LVF and high LAP.

- MI $[6hr, 5hr] \Rightarrow$
- low LVF $@[6hr, 5hr : now] \Rightarrow$

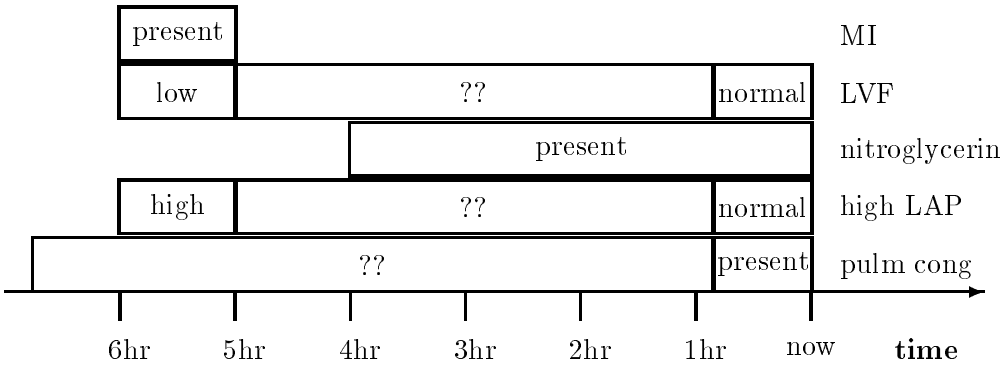

Figure 5: Hypothesis with low LVF ending

- high LAP $@[6hr, 5hr : now] \Rightarrow$
- PC $@[past : now, now : future]$

Low LVF has a definite begin time from the MI and the onset time of zero for low LVF. It definitely continues until 5hr because it is not intermittent nor is its maximum existence exceeded. In applying the new time interval for high LAP to PC, there is another consideration. Since there is another possible cause for PC (pneumonia) which is still unknown, the beginning could still be prior to 6 hours ago. The latest begin time can not be constrained further either since the onset can be up to 6 hours. Thus, the time interval for PC remains as it was before. The resulting hypothesis is shown in figure 5.

Nitroglycerin Correcting High LAP Hypothesis In the final hypothesis the nitroglycerin corrects the high LAP for the time the nitroglycerin is present. This implies the following time intervals:

- nitroglycerin $@[4hr, now] \Rightarrow$
- normal LAP $@[past : 4hr, now]$ (normal over the 4 hours) $\Rightarrow$
- high LAP $@[past : 4hr, 1day : 4hr]$ (as constrained by PC) $\Rightarrow$
- normal LAP $@[1day : 4hr, now]$ (as constrained by high LAP)

Now the high LAP is accounted for by low LVF, except that the end of the high LAP does not determine the end of the high LVF, because the nitroglycerin removes the causal relationship once it has begun.

- high LAP $@[past : 4hr, 1day : 4hr] \Rightarrow$
- low LVF $@[past : 4hr, 1day : future]$

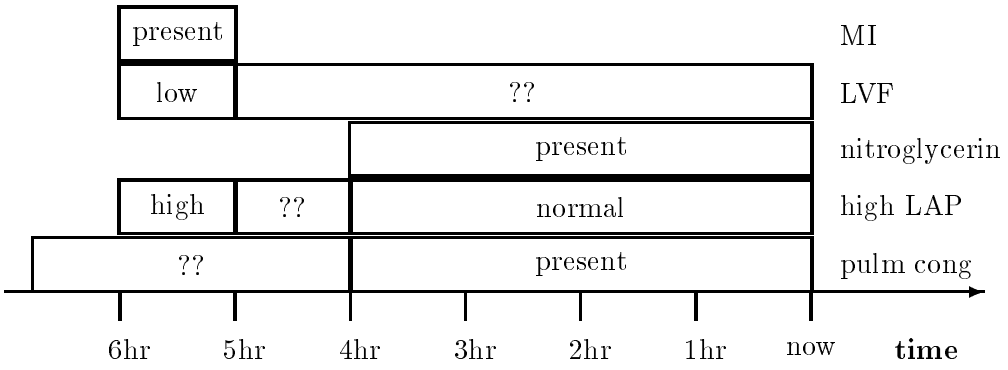

Figure 6: Hypothesis with Nitroglycerin Correcting LAP

Again, this is caused by the MI, which already is known to be true over the time interval $[6hr, 5hr]$. Since effects can not precede causes, this constrains the low LVF, high LAP, and PC.

- MI $[6hr, 5hr] \Rightarrow$
- low LVF $@[6hr, 5hr : future] \Rightarrow$
- high LAP $@[6hr, 5hr : 4hr] \Rightarrow$
- PC $@[past : 4hr, now : future]$

The PC must have started before the high LAP was corrected to overlap with its cause, but the earliest beginning is undetermined because there is another possible cause. This hypothesis is shown in figure 6.

8 Probability of Hypotheses

The reasoning described thus far is sufficient to determine hypotheses that are temporally consistent. The next step is to compute the probabilities of the hypotheses. This is complicated because the hypotheses are actually constrained patterns of possible scenarios. The different scenarios within each pattern will have different probabilities. The issue is how precisely a hypothesis needs to be defined. For example, the pneumonia hypothesis does not specify whether the LVF was ever low and neither high LAP hypothesis specifies whether the PC came on immediately or after an hour or two. The first of these situations results in different nodes in the hypothesis. The second has different time intervals on the nodes. There is enough knowledge in the model to distinguish between either of these situations, but there is no reason to distinguish beyond what is clinically relevant. In the following we will require the nodes to be fully specified, but not the time intervals. Thus the problem is to compute the maximum probability represented by a path through the nodes of a hypothesis.

In all of the hypotheses, MI and nitroglycerin are true and not dependent on any other nodes. Their probabilities will be ignored in the following analysis since the probabilities are only used to rank order the hypotheses. We will first consider the two hypotheses with high LAP.

LVF Normalized Hypothesis Probability In this hypothesis, pneumonia was unspecified. Specifying that pneumonia is false further determines the time of PC.

- high LAP @$[6hr, 5hr : now]$ & pneumonia false $\Rightarrow$
- PC @$[6hr : now, now : future]$

The probability of pneumonia being false is 0.99, but since this and the probability of all other primary nodes being false is so close to 1.0, they can be ignored. The heuristic strategy for computing the probability of the best path through a hypothesis is to use the segment of the causal range of each relation with the highest probability. Thus, the computation of the probability is as follows:

MI $\rightarrow$low LVF	=	0.6
low LVF end $<$ 6hr	=	0.3
low LVF $\rightarrow$high LAP	=	0.8
high LAP $<$ 6hr $\rightarrow$PC	=	0.5
PC persist $>$ zero	=	1.0
$\Rightarrow$ *total probability*	=	0.072

The probability of the MI producing low LVF is 0.6. Since the low LVF ended before now, it normalized in less than 6 hours, for which the probability is 0.3. The probability of low LVF producing high LAP is 0.8. There are two regions in which to consider the high LAP. Either it ended before the nitroglycerin started or it continued for some time after the nitroglycerin started but ended before now. This decision does not determine which nodes are in the hypothesis, so it is left undetermined. Finally, the probability of high LAP causing PC depends on how long the high LAP has been present. We know it happened in less than 6 hours, but not whether it happened in less than one hour. Again, the decision also does not determine which nodes are in the hypothesis, so the time interval is left unchanged and the higher probability is picked. The probability that the PC normalized in less than 6 hours is zero, so no further adjustment is necessary and the total estimated probability is the product of the probabilities.

High LAP Corrected Hypothesis Probability The computation of the probability for this hypothesis is similar. Again, pneumonia is assumed false with a probability of one. The probability of the MI producing low LVF is still 0.6, but now there is no need to decide when the low LVF ends. The probability of high LAP is still 0.8. The high LAP explicitly ends with the addition of the nitroglycerin, for which the probability is 0.8. Since the high LAP therefore lasts for 2 hours, the probability of PC is 0.5. The PC must persist for the 4 hours until the exam time, but PC always lasts at least 6 hours. The estimated probability is summarized as follows:

MI $\rightarrow$low LVF	=	0.6
low LVF $\rightarrow$high LAP	=	0.8
nitroglycerin $\rightarrow$normal LAP	=	0.8
high LAP $<$ 2hr $\rightarrow$PC	=	0.5
PC persist $>$ 4hr	=	1.0
$\Rightarrow$ *total probability*	=	0.192

Pneumonia Hypothesis Probability The third hypothesis has the intermediate nodes for low LVF and high LAP false. The probability of this is also calculated by following the causal pathways. The probability of the pneumonia is 0.01 and the probability that it causes PC is 0.5. The probability that the MI did not produce low LVF is 0.4.

pneumonia	=	0.01
pneumonia $\rightarrow$PC	=	0.5
MI $\not\rightarrow$low LVF	=	0.4
$\Rightarrow$ *total probability*	=	0.002

There is also opportunity for local optimization of this hypothesis by adding nodes. Since the probability of low LVF is higher than the probability of it being false, the program explores whether this leads to a better hypothesis. The highest probability scenario corresponds to the time intervals in the hypothesis with high LAP corrected by nitroglycerin. The primary difference is that there are now two causes for PC and the probability of PC is $0.5 + 0.5 - 0.5 \times 0.5 = 0.75$, using the default combination rule. Thus, if those two nodes are added to the pneumonia hypothesis, the probability becomes 0.0029.

pneumonia	=	0.01
MI $\rightarrow$low LVF	=	0.6
low LVF $\rightarrow$high LAP	=	0.8
nitroglycerin $\rightarrow$normal LAP	=	0.8
high LAP $<$ 2hr & pneumonia $\rightarrow$PC	=	0.75
$\Rightarrow$ *total probability*	=	0.0029

Thus, the HDP is able to rank order the hypotheses and say that the most likely hypothesis is that the nitroglycerin corrected the high LAP. About a third as likely is the hypothesis that the low LVF normalized. About 1.5% as likely is that pneumonia was at least partially the cause for the pulmonary congestion.

9 Pre-Computing Temporal Constraints

With the HDP knowledge base there are thousands of paths that lead to a node such as pulmonary congestion. Most of these paths are never used, but without further information they must be investigated. In the version of the HDP without temporal reasoning, we pre-computed all of the paths to provide a fast way of picking the most likely candidates for hypotheses and picking the best causal paths. Temporal reasoning provides a way to extent this mechanism. Even though the times of nodes are not known before a case is entered, the minimum and maximum extent are known from the temporal properties in the knowledge base. These can be transferred down the causal chain by applying each new constraint to the one above it. When there is an inconsistency, that causal path can be eliminated.

Therefore, we added properties *min-exist* and *max-exist* to each node in the computed causal path. The node constraint min-exist is determined by the onset and the max-exist is determined by the max-exist and persist properties. In the causal path, if the max-exist of the cause is less than the delay or onset or if the min-exist of the cause is greater than the max-exist of the effect, the causal path is stopped. In this way we were able to eliminate about 20% of the causal pathways that were generated in the older version of the HDP.

This also allows the computation of time bounds on nodes before all of the causality has been determined. In the example case, it is possible to determine that the earliest possible begin time for PC is two weeks, because that is the earliest time for any causal path (in the very limited model) that causes PC. This in turn assures that any hypothesis that would require PC longer than two weeks would be eliminated without further computation. Thus, pre-computing the implications of the temporal constraints allows the program to make optimal use of them.

10 Discussion

Our approach to temporal reasoning for the HDP raises several issues: Why temporal constraints? What does a hypothesis represent? What hypotheses belong in a differential? What do the probabilities represent?

One possible alternative to temporal constraints is to use probability density functions (PDFs). Probability statements such as that for high LAP causing PC with different probabilities for times less than an hour, one to six hours, and greater than six hours are essentially approximations of PDFs. There are two problems with using PDFs. The obvious one is the increased computational burden imposed on an already computation intensive task. The second problem is how to break up a structure consisting of PDFs into hypotheses. The explicit time bounds provide a natural way to generate and compare different hypotheses. Finally, it is difficult to estimate the time bounds and probabilities for this model and the task of estimating PDFs for each causal relation in a 200 node model would be nearly impossible.

Given a hypothesis consisting of a network of causally linked nodes with temporal intervals, what does it represent? If one thinks in terms of possible scenarios producing the observed findings, the hypothesis is a finite region in the space of possible scenarios. That is, it is all scenarios meeting the constraints on the hypothesis nodes. The bounds of the region are defined by the clinically significant distinctions that determine the time bounds in the model. Thus, each region defined by a hypothesis network should differ from every other region in some detail of potential clinical significance. The question then is what differences might have clinical significance. The most extreme position would be to use each region defined by a distinction in the model as an indication of clinical significance. In the analysis of the example, we took a less restrictive position and once the temporal constraints of the data was accounted for no other distinctions were enforced unless they involved nodes being included in the hypothesis. A third possible position would be to only enforce the temporal distinctions of the data and leave nodes for which there is no evidence other than a possible cause as unknown. The appropriate strategy depends on the purpose of the diagnosis, since diagnosis is a tool for patient management and not an end in itself.

What hypotheses belong in the differential again depends to some extent on the purpose of the differential. If the user is interested in the overall diagnosis of the patient, only

hypotheses that differ in nodes of diagnostic significance should be included. In the example, the distinction of with and without pneumonia is significant but the distinction of continuing or ended low LV function may be too small a detail. However, if the user is considering what changes to make in the therapy for the patient, the fact that the low LV function may have ended and therefore the nitroglycerin may no longer be necessary is a useful consideration. So far, the principle use of the HDP has been to explain the overall diagnosis, so the program only presents hypotheses with diagnostic distinctions.

If the hypothesis represents all possible combinations of times of causation and persistence that are consistent with the pattern of nodes in the hypothesis within the time constraints on the nodes, then the probability should be the sum of the probabilities of all of the mutually exclusive allowable combinations of times through the hypothesis — essentially a multiple integration of the possible probabilities over time. Because of the computational difficulties associated with this strategy, we have chosen a heuristic for estimating the probability. The probability for each time interval is determined locally from the constraints on the causes. Thus, in the 'corrected high LAP' hypothesis, there was no need to decide how long the low LVF continued and no reason to decide how long the high LAP continued in the 'normalized LVF' hypothesis. It would be possible to make a model in which there were situations where unlikely hypotheses would be attributed significant probabilities, but in practice the relative probabilities for the hypotheses produced are consistent with our expectations.

11 Conclusion

The temporal reasoning of the HDP was implemented a couple of years ago and has gone through one evaluation and soon will go through another. The addition of temporal reasoning has eliminated a major class of the errors that were being made before, such as findings with chronicity longer than the proposed cause or ignoring possible causes because some part of the causal chain was no longer true. Examination of the hypotheses produced by the HDP make sense to cardiologists and do not reveal any shortcomings in the temporal reasoning process.

The domain of heart disease provides a good test bed for developing temporal reasoning in a diagnostic context. There are a wide variety of temporal situations that require explicit reasoning to handle properly. Thus, our expectation is that the lessons learned in this domain will be transferable to a number of other domains.

Acknowledgments The development of the HDP knowledge base has been the result of a long-term collaboration with Drs. Shapur Naimi and M. G. Criscitiello of the Tufts-New England Medical Center. However, they are not responsible for the pedagogical changes made for this paper. This research was supported by National Institutes of Health Grant No. R01 HL33041 from the National Heart, Lung, and Blood Institute.

References

[1] J. F. Allen, Towards a General Theory of Action and Time, Artificial Intelligence 23 (1984) 123-154.

[2] G. O. Barnett, J. J. Cimino, J. A. Hupp, and E. P. Hoffer, DXplain: An Evolving Diagnostic Decision-Support System, Journal of the American Medical Association 285 (1987) 67-74.

[3] E. S. Berner, Performance of four computer-based diagnostic systems, New England Journal of Medicine 330 (1994) 1792-1796.

[4] L. Console, A. J. Rivolin, and P. Torasso, Fuzzy Temporal Reasoning on Causal Models, Intl J. Intelligent Systems, 6 (1991) 107-133.

[5] M. B. First and R. 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 (1985) 137-165.

[6] O. K. Hejlesen, S. Andreassen, and S. K. Andersen, Implementation of a Learning Procedure for Multiple Observations in a Diabetes Advisory System Based on Causal Probabilistic Networks, in: S. Andreassen et al., eds., Artificial Intelligence in Medicine, (IOS Press, Amsterdam, 1993) 63-74.

[7] M. G. Kahn, Modeling Time in Medical Decision-support Programs, Medical Decision Making 11 (1991) 249-264.

[8] W. Long, Reasoning About State from Causation and Time in a Medical Domain, American Association for Artificial Intelligence 1983 Conference, (1983) 251-254.

[9] W. Long, Medical Diagnosis Using a Probabilistic Causal Network, Applied Artificial Intelligence 3 (1989) 367-383.

[10] W. J. Long, S. Naimi and M. G. Criscitiello, Development of a Knowledge Base for Diagnostic Reasoning in Cardiology, Computers in Biomedical Research 25 (1992) 292-311.

[11] W. J. Long, S. Naimi, and M. G. Criscitiello, Evaluation of a New Method for Cardiovascular Reasoning, Journal of the American Medical Informatics Association 1 (1994) 127-141.

[12] R. A. Miller, H. E. Pople, and J. D. Myers, INTERNIST-I, An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine, New England Journal of Medicine 307 (1982) 468-476.