Computación y Sistemas Vol. 5 No. 2 pp. 109 - 119
© 2001, CIC - IPN. ISSN 1405-5546 Impreso en México

Medical Diagnosis Using Temporal Probabilistic Networks

Red Probabilística Temporal Aplicada al Diagnóstico Médico

Gustavo Arroyo-Figueroa[1] **and Luis Enrique Sucar Succar**[2]
[1]Gerencia de Sistemas Informáticos, Instituto de Investigaciones Eléctricas
Av. Reforma # 113, Col. Palmira, 62490 Temixco, Mor. México
[2]Instituto Tecnológico de Estudios Superiores de Monterrey
Cuernavaca, Morelos, México
e-mail : garroyo@iie.org.mx, esucar@campus.mor.itesm.mx

Article received on October 18, 2000; accepted on October 20, 2001

Abstract

Diagnosis and prediction in some domains, like medicine, require an adequate representation that combines uncertainty management and temporal reasoning. This paper presents a novel representation called Temporal Nodes Bayesian Network (TNBN). A TNBN is a Bayesian network in which each node represents an event or state change of a variable, and an arc corresponds to a causal-temporal relationship. A temporal node represents the time that a variable changes state, including an option of no-change. The temporal intervals can differ in number and size for each temporal node, so this allows multiple granularity. The proposed approach is applied to medical diagnosis through a case study when a car accident occurs. The results of this study for different cases are presented and compared to a simple Bayesian network and a dynamic Bayesian network.

Keywords: Bayesian networks, Temporal reasoning, medical diagnosis, prediction, intelligent system.

Resumen

En este trabajo se presenta el desarrollo y la aplicación de un nuevo formalismo que combina el manejo de incertidumbre con el manejo de relaciones temporales. El formalismo es denominado como Red Bayesiana con Nodos Temporales (RBNT). Una RBNT puede verse como una estructura dinámica compacta que maneja cambios de estados (eventos), en vez de estados de las variables, asociados a un intervalo de tiempo donde es factible su ocurrencia. Estos intervalos de tiempo pueden ser definidos de manera diferente en número y en magnitud para cada nodo temporal, lo cual hace que el formalismo sea de múltiple granularidad y con alta capacidad expresiva. El modelo es aplicado al diagnóstico médico de las lesiones en un paciente cuando ha ocurrido un accidente automovilístico.

Palabras clave: Redes Bayesianas, razonamiento temporal, diagnóstico médico, sistemas inteligentes.

1 Introduction

Representing and reasoning about time is an essential part of medical problem solving. Many of the main medical tasks, such as prevention, dia gnosis, therapeutic planning/ intervention, prognosis, and medical discovery involve time modeling and inference [Hanks et al., 1995; Aliferis and Cooper, 1996]. Similarly, in fields such as industrial diagnosis [Arroyo and Sucar, 1999], planning [Santos, 1996], scheduling [Brusoni et al., 1994], forecasting [Dagum et al., 1992], speech recognition [Zweig and Rusell, 1998] among others, representing the dynamic aspects of the domain is essential for the success of the model. In particular, the evolutionary nature of these domains requires a representation that takes into account temporal aspects. The exact timing information for things like lab-test results, occurrence of symptoms, observations, measures, as well as faults, can be crucial in this kind of applications.

In medicine, temporal information can be very important in settings like "blood test B was taken 5 hours after drug D was administered to the patient" [Santos, 1996]; "what is the patient's state 10 minutes after a collision had occurred" [Hanks et al., 1995]; or "the feedback between glucose and insulin levels in the human body" [Aliferis and Cooper, 1996]. In spite of the need of time modeling, there are a few number of medical knowledge-based systems developed throughout the years that have incorporated explicitly temporal aspects. In general, modeling and reasoning with time is considered to be one of the greatest challenges in developing knowledge-based systems for medical diagnosis.

To model temporal relations is a complex task. Temporal models are much more complex than atemporal ones [Haddawy, 1996]. In a temporal model each variable and its relationships with other variables must be examined over multiple points of time. These tasks often entail an inordinate amount of computation, due to the size and the complexity of the resulting model. In the context of intelligent systems, a temporal model must be capable of reasoning about the present, past and future state of the domain. Moreover, an ideal temporal knowledge representation should satisfy the criteria of *temporal expressiveness* (sound and flexible time model); *computational tractability* (adequate temporal inference that allows to examine the variables interactions over multiple points of time); *and temporal knowledge acquisition* (amenable to machine-learning methods for temporal concepts acquisition).

Aside from temporal considerations, real world information is usually imprecise, incomplete and noisy. The temporal model must be able to deal with uncertainty. Among the formalism proposed for dealing with uncertainty, one of the most used techniques for the development of intelligent systems are Bayesian networks (BN). Bayesian networks [Pearl, 2000], also know as probabilistic networks, causal networks or influence diagrams, are graphical structures used for representing expert knowledge. A BN is a directed acyclic graph (DAG) whose structure corresponds to the dependency relations of variables represented in the networks (nodes), and which is parameterized by the conditional probabilities (links) required to specify the underlying distribution. In this graphical representation of dependencies and independencies, each node represents a random variable and each arc a probabilistic dependency. Given a knowledge base represented as a Bayesian network, it can be used to reason about the consequences of specific input data, by what is called probabilistic reasoning. This consists of instantiating the input variables, and propagating their effect through the network to update the probability of the hypothesis variables. The updating of the probabilities is consistent with probability theory, based on the application of Bayesian calculus and the independencies represented in the network. Probability propagation in a general network is a complex problem, but there are efficient algorithm for certain restricted structures: singly connected networks (tree and polytrees structured) and alternative approaches for more complex networks.

Bayesian networks usually represent a static causal model of certain domain. That is, the nodes represent the values of the variables at a specific time point and temporal relations are not considered. Diagnosis and prediction in some domains, like medicine, require an adequate representation that combines uncertainty management and temporal reasoning. In some medical applications, there is typically enough time to gather information about the patient's state and consider alternative diagnoses and treatments, but the temporal interaction between the timing of the tests, treatments, and the course of the disease must also be considered. For such domain a static Bayesian network is not very useful: the estimation of probability distributions of domains variables based on appropriate prior knowledge and observation of other variables is reliable only for a limited period of time. Thus, to cope with such dynamic systems using probabilistic networks we need to consider temporal reasoning under uncertainty [Boyen and Koller., 1999]. The extension of Bayesian networks semantics to deal with temporal relationships can be complicated. The main problem is to represent each node with its dependence relationships over multiple points of time.

Significant research has been done exploring probabilistic networks which are evaluated at each point in time. For modeling dynamic systems using probabilistic networks we need interconnect multiple instances of static networks. Obviously, as time evolves, new "slices" must be added to the model and old ones cut off. This introduces the notion of dynamic Bayesian networks [Kjaerulff, 1992]. A DBN may be defined as a sequence of submodels each representing the state of the dynamic domain at a particular point in time. Hence, a DBN consists of a series of, most often structurally identical, subnetworks (time slice) interconnected by temporal relations. Time slices are duplicated over a predetermined time grid representing the temporal range of interest. Temporal relations are represented with links between nodes of different time slices. The conditional probability tables for a DBN include a state evolution model, which describes the transition probabilities between states, and a sensor model, which describes the observations that can result from a given state. Typically, one assumes that the conditional probability tables in each slice do not vary over time. The exact belief propagation in DBN is a complex problem. Clustering algorithms for DBN's are too expensive and impractical for realistic applications problems. Recently; approximate simulation algorithms had been proposed [Kanazawa et al., 1995] based on the idea that exact probabilities are not needed. The justification is based on the intuition that, if processes interact only weakly, the error can not be too large.

In this paper we present an alternative representation called Temporal Nodes Bayesian Network (TNBN), a probabilistic network of events in discrete time. A TNBN is a Bayesian network in which each node represents an event or state change of a variable, and an arc corresponds to a causal-temporal relation. A temporal node represents the time that a variable changes state, including an option of no-change. Including more temporal nodes can represent more than one change for the same variable. The temporal intervals can differ in number and size for each temporal node, so this allows multiple granularity. Temporal information is relative, that is, there is not an absolute temporal reference. We developed a mechanism for transforming the relative times to absolute, based on the timing of the observations. With this representation we can model complex real-world systems with a simple network, and use standard probability propagation techniques for diagnosis and prediction. The proposed approach is applied to the diagnosis and prediction of events that occur after an automobile accident.

The rest of this paper is organized as follows. Section 2 describes a medical case study of an accident and its effects, presents a simple Bayesian network and a dynamic Bayesian network representation, and introduces our approach, which is contrasted with both models. Section 3 presents a formal definition of the proposed model and the section 4 describes the inference mechanism for evaluation of a TNBN. In section 5, an empirical evaluation is presented for a medical example. Section 6 presents a brief discussion of several extensions of BNs for time modeling. Finally, section 7 presents the conclusions and future research.

2 A Medical Example

To illustrate the proposed temporal probabilistic model, we present the hypothetical example of the consequences of an automobile accident based on [Hanks et al., 1995]. The example expresses the necessity for representing temporal relations for medical diagnosis.

Assume that at time t=0 an automobile accident occurs. The driver is a healthy 45 years old man and contact with steering wheel is noted. This kind of accidents can be classified as *severe*, *moderate* or *mild*. The immediate consequences in this sort of accident are injuries to the *head*, *abdominal cavity* and *internal organs*, *chest* and *extremities*. For demonstration purpose we only consider *head* and *chest* injuries. Injury of the head can bruise the brain, which will cause it to begin swelling. Chest injuries can include a fractured sternum, one or both punctured lungs, and bleeding in the chest cavity. These instantaneous state changes can initiate a set of internal changes that will generate subsequent changes. For example, brain trauma will cause the brain to begin swelling. This increase of the brain volume tends to increase intracranial pressure, which in turn eventually cause *dilated pupils*, *destabilized vital signs* (pulse and blood pressure) and loss of consciousness. Bleeding into the chest cavity decreases blood volume over time, which also tends to destabilize vital signs. Internal bleeding will also eventually increase pressure on the heart, decreasing its efficiency, further destabilizing vital signs. The collision itself can be modeled as an external event, which can immediately cause certain changes in the patient's state: trauma to the brain, broken sternum, punctured lung, and bleeding in the chest cavity. These changes cause internal changes, which are not immediate: dilated pupils, vital signs unstable, and loss of consciousness, and depend on the severity of the accident.

Suppose that we gathered the following statistics about the accidents that occurred in a specific zone of a city:

- 36.80% of the collisions (**C**) are severe, 39.20% are moderate and 24% are mild.
- If the accident is mild, then the probability that head injury occurs is 0.1 and the probability of an injury resulting in slight internal bleeding (**IB**) is 0.6 and gross is 0.05.
- If the accident is moderate, then the probability that head injury occurs is 0.4 and the probability of an injury resulting in slight internal bleeding is 0.15 and gross is 0.65.
- If the accident is severe, then the probability that head injury occurs is 0.9 and the probability of an injury resulting in slight interval bleeding is 0.4 and gross is 0.5.

This information indicates that there is a strong causal relationship between the severity of the accident and the immediate effect in the patient's state. Additionally, there are some important temporal information about the relation between the instantaneous consequences (head injury and internal bleeding) and the symptoms (pupils dilated and vitals signs unstable).

- If a head injury (**HI**) occurs, the brain will start to swell, and if left unchecked, the swelling will cause the pupils to dilate (**PD**) within 0 to 10 minutes.
- If internal bleeding (**IB**) begins, the blood volume will start to fall, which will tend to destabilize vital signs (**VS**). The time required to destabilize signs will depend on the severity of bleeding:
- If the bleeding is gross, it will take from 10 to 30 minutes.
- If the bleeding is slight, it will take between 30 to 60 minutes.
- A head injury (**HI**) also tends to destabilize vital signs, taking between 0 to 10 minutes to make them unstable.

Figure 1 shows the temporal occurrence of the symptoms in relation with the time of the immediate effects.

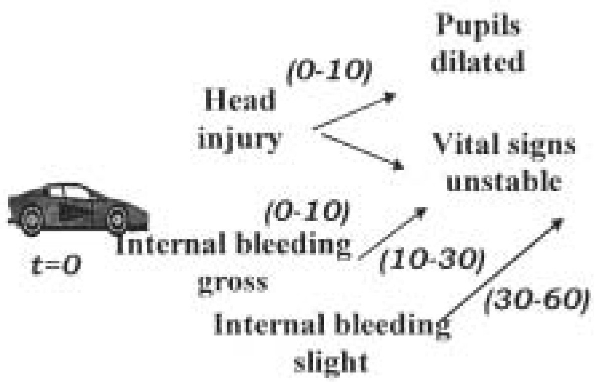


Figure 1. Temporal relations between immediate effects and symptoms

3 Probabilistic Models for the Medical Example

3.1 Static Bayesian Network

In this example there is an external event: the collision (**C**); that generates two immediate effects in the patient's state: head injury (**HI**) and internal bleeding (**IB**). These internal events produce certain posterior endogenous changes in the patient. These changes are not immediate and are manifest through two observations: *pupils dilated* (**PD**) and *vital signs unstable* (**VS**). The severity of the collision has a direct causal relation with the variables head injury and internal bleeding. Figure 2 shows a static Bayesian network for the collision event, and table 1 shows the required conditional probabilities.

A simple Bayesian network can not represent the temporal information of the dynamic domain. That is, the temporal relationships between the occurrence of the immediate effects and the symptoms.

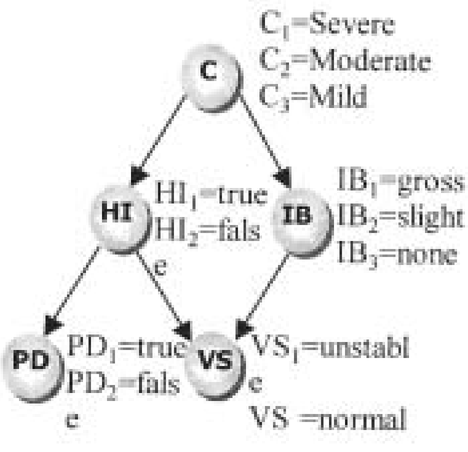


Figure 2. Static Bayesian network for the accident example

Node "accident" (C)						
Severe	0.37					
Moderate	0.39					
Mild	0.24					
Node "head injury" (HI)						
Accident	Severe	Moderate	Mild			
True	0.9	0.4	0.1			
False	0.1	0.6	0.9			
Node "internal bleeding" (IB)						
Accident	Severe	Moderate	Mild			
Gross	0.65	0.5	0.05			
Slight	0.25	0.4	0.60			
None	0.1	0.1	0.35			
Node "vital signs" (VS)						
Head injury	True			False		
Internal bleeding	Gross	Slight	None	Gross	slight	none
Normal	0.05	0.08	0.2	0.01	0.1	0.95
Unstable	0.95	0.92	0.8	0.99	0.9	0.05
Node "pupils dilated" (PD)						
Head injury	True	False				
True	0.97	0.1				
False	0.03	0.9				

Table 1. Conditional probabilities for the static Bayesian network model.

3.2 Dynamic Bayesian Network

The most common representation of dynamic relations are dynamic Bayesian networks [Kjaerulff, 1992]. Figure 3 shows a simple DBN for the "accident" example, with a time slice each 10 minutes, the maximum common divisor of the time intervals.

This is a simple DBN for the example, which considers the following assumptions: (i) a state depends only on the previous one (Markovian assumption), (ii) there are links only between the same variable at different slices. Even with these simplifications, it is a complex model in terms of storage requirements and computation time for probability propagation. If we consider a more complex model, relaxing the previous assumptions, it could become prohibitive for realistic applications. Also, the acquisition of the model (structure and parameters) could become a problem.

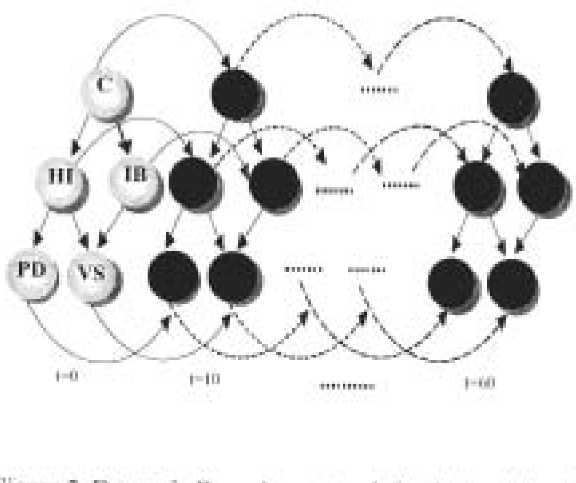


Figure 3. Dynamic Bayesian network for the "accident" example

Unfortunately DBN's present the following problems for realistic applications [Aliferis and Cooper, 1996; Santos, 1996]: (1) DBN is a model of high complexity, hundreds or even thousands of nodes and conditional probability distributions may have to be defined; (2) DBN's handle a predefined temporal range of interest, the DBN's do not allow to vary the temporal range as a model parameter; and (3) DBN's do not have an integrated temporal/causal semantics, the knowledge about time can not be exploited easily to prevent serious inconsistencies.

3.3 Temporal Nodes Bayesian Network

We propose an alternative representation of temporal aspects in Bayesian networks. In many cases, there are few states changes (events) in the temporal interval of interest in the domain. The timing of these events is usually important for diagnosis and prediction. For instance, in the medical example, the time when "vital signs unstable" and "pupils dilated" occur is crucial for the accident diagnosis.

To model these changes, we require a representation of events. We propose a temporal representation based on events and its time interval of occurrence, called *Temporal Node Bayesian Networks* (TNBN) [Arroyo and Sucar, 1999]. A TNBN is a Bayesian network in which each node represents an event or state change of a variable, and an arc corresponds to a causal-temporal relation. A *temporal node* represents a possible state change of a variable and the time when it happens. Each value of a temporal node is defined by an ordered pair: the value of the variable to which it changes and the time interval of its occurrence. Time intervals represent relative times between the parent events and the corresponding state change. A temporal node has an initial or default state, so a value is associated to this state with a maximum time interval (*temporal range*) and it indicates the condition of no change. Figure 4 shows a TNBN model for the "accident" example, and table 2 shows the conditional probability distribution for each temporal node.

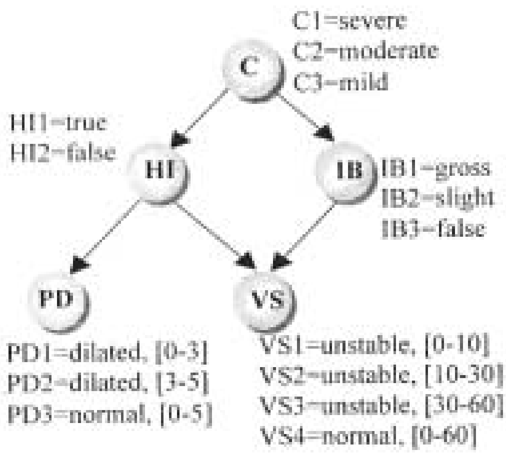


Figure 4. TNBN for the accident example

Node "vital signs unstable"						
Head injury	False			True		
Internal bleeding	gross	Slight	None	gross	Slight	None
Normal [0-60]	0.05	0.08	0.20	0.05	0.10	[illegible]
Unstable [0-10]	0.80	0.70	0.35	0.80	0.05	[illegible]
Unstable [10-30]	0.15	0.15	0.15	[illegible]	0.25	[illegible]
Unstable [30-60]	0.05	0.07	0.12	0.05	0.70	0.05

Node "pupils dilated"		
Head injury	True	False
False [illegible]	0.05	0.90
True [illegible]	0.80	0.05
True [illegible]	0.15	0.05

Table 2. Table of conditional probabilities for the temporal nodes

In the accident example, there are 3 instantaneous events: *collision, head injury* and *internal bleeding*; and two events that can be represented by nodes with temporal intervals: *pupils dilated* and *vital signs unstable*. ***PD*** has *normal* as initial state, and can change to *dilated* in 2 temporal intervals ([0,3],[3,5]); while ***VS*** has *normal* as initial state, and can change to *unstable* in 3 different time intervals ([0-10],[10-30],[30-60]). Both variables have the default state associated to the overall time interval, ([0-5] and [0-60]) which correspond to the no change condition. The time intervals were defined based on the temporal information of the accident example.

The TNBN model can be used, for example, to predict the consequences of an accident or to diagnose its severity. The resulting network is less complex than the corresponding DBN. The main difference with a DBN is that the representation is based on state changes at different times represent by temporal nodes instead of state values at different times, represent by networks. A TNBN can be seen as natural extension of a Bayesian network and its properties are parallel. The temporal intervals can differ in number and size for each temporal node, so this allows multiple granularity. Temporal information is relative, that is, there is not an absolute temporal reference. We also developed a mechanism for transforming the relative times to absolute, based on the timing of the observations. A formal definition of a TNBN is presented in the next section.

4 Formal TNBN Definition

A TNBN is a Bayesian network in which each node represents an event or state change of a variable, and an arc corresponds to a causal-temporal relation. This representation is based on the definition of a temporal node. A temporal node is defined by a set of states. Each state is defined by an ordered pair: the value of the variable (to which it changes) and a time interval associated to the change of value of the variable. A temporal node is defined as follows:

Definition 1. A temporal node (***TN***) is defined by a set of states, each defined by an ordered pair (σ,τ), where σ is a value of a random variable and τ is the time interval associate to the change of variable value. There is a default state of not change that corresponds to the initial value (generally the "normal" value), associated to the temporal range of the node.

The values of each **TN** can be seen as the "cross product" between the set of values (Σ) and the set of time intervals (T), except for the default state, which is associated only to the temporal of interest (**TR**). The definition of the **TNs** for the accident example was presented in table 2.

TNs are connected by edges. Each edge represents a causal-temporal relationship between TNs. The conditional probability distribution for each node is defined as the probability of each ordered pair (σ_i,τ_i) given the ordered pairs of its parents (σ_j,τ_j).

As a TN is defined by a set of time intervals, we can relate these time intervals based on Allen's temporal algebra [Allen, 1983]. A temporal relationship between the time intervals of a TN is defined as:

Definition 2. In a **TN** the definition of the default state is associated to temporal range of interest, TR. The possible temporal relationships between TR with the time intervals, Ti, of a node are: *start* (**si**), *during* (**di**) and *finish* (**fi**). The temporal relationship between each pair of time intervals is *meet* (**m**): Ti {**m**} Tj.

For example, for the temporal node *vitals signs* the relationships between its time intervals are represented in the figure 5.

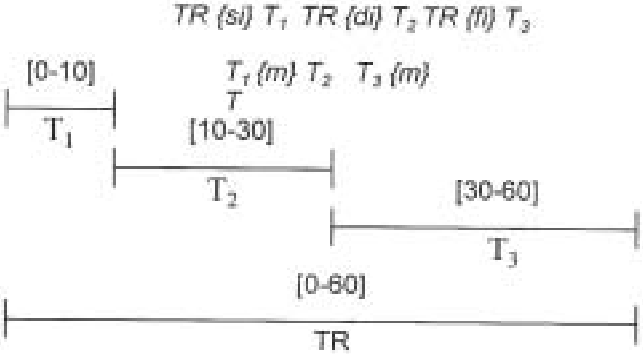


Figure 5. Temporal relationships between the time intervals of node vitals signs

Finally, a Temporal Node Bayesian Network (TNBN) is defined as:

Definition 3. A TNBN is defined as **TNBN=(V, E)**, where V is the set of temporal nodes and E is the set of edges. Each temporal node is defined by an ordered pair (σ, τ) and the conditional probability matrix that specifies the probability of each ordered pair given its parents.

In each temporal node, the temporal intervals are relative to the parent nodes, that is, there is not an absolute temporal reference. This makes the representation more general; but, for its application, we need to associate these relative times to the actual or absolute times of the observed events. We developed a mechanism for transforming the relative times to absolute, based on the timing of the observations. In the next section, we present the definition of the inference mechanism.

5 Inference Mechanism

As we mentioned before, the temporal intervals in each node are relative to its parents, that is, there is not absolute timing of the events until one is observed. When an initial event is detected, its time of occurrence "fixes" temporally the network. The timing of the observation is used as temporal reference for the other events. This means that the actual timing of the events represented in the network is dynamic. For definition of the inference mechanism, we need to define some additional parameters:

Tc (*real time of occurrence*): is the actual time when an event is detected. As the net does not have any temporal reference, the time of occurrence of the initial event fixes temporally the network.

α (*real time occurrence function*): is the absolute value of the difference between the time of occurrence of a pair of events, $\alpha = |tc - tc^{'}|$, where tc is the time of occurrence of the first event and tc^{N} is the time of occurrence of the second event.

These parameters are used by the inference mechanism for determining the actual time intervals of occurrence of each event. The mechanism consists of 3 basic steps, which are as follows.

Step 1. Event detection and time interval definition

When an event initial is detected, its time of occurrence, "$\mathbf{tc}_{initial}$", is utilized as temporal reference for all the network. There are two possible cases, depending on the position of the *initial* node in the network: (a) the initial event corresponds to a root node, and (b) the initial event corresponds to an intermediate or leaf node.

1-(a). In the first case, the actual value of the node can be determined (root nodes are always instantaneous events).

1-(b). For the second case, it is not possible to determine the value of the variable, because the event could be associated to any time interval for the state. It is necessary to wait for a second observation to determine the interval. When the next event is detected, its time of occurrence, $\mathbf{tc}_{posterior}$, is utilized for definition of the time interval associated with the real time occurrence function, $\alpha = |tc_{initial} - tc_{posterior}|$. The value of α is used to set the time interval of the child node considering the parent node as the initial event. This step is applied recursively to subsequent events.

Step 2. Propagation of the evidences

Once the value of a node is obtained (time interval and associated state), the next step is to propagate the effect of this value through the network to update the probability of other temporal nodes. It can be use any standard algorithm for probability propagation.

Step 3. Determination of the past and future events

With the posterior probabilities, we can estimate the potentially past and future events based on the probability distribution of the each temporal node.

If there is not enough information, for instance there is only one observed event which corresponds to an intermediate node, the mechanism handles different *scenarios*. The node is instantiated to all the intervals corresponding to the observed state, and the posterior probabilities of the other nodes are obtained for each scenario. These scenarios could be used as a set of possible alternatives, which will be reduced when another event occurs.

The TNBN representation and inference mechanism were applied for diagnosis and prediction of events in a medical example, the next section presents these experiments.

6 Experimental Results

Suppose that at time t=10 minutes the patient is observed by paramedics. They observe a probable broken sternum, the patient is complaining of shortness of breath and dizziness, vital signs are unstable, but pupils are not dilated. From this symptoms the probable chest injury and unstable vital signs suggest internal bleeding, which will soon cause serious problems if left unattended. Intravenous fluids should probably be administered immediately to increase the blood volume, and if the transportation to the hospital is expected to take more than 20 minutes it might be best to insert a chest tube to drain blood from the chest cavity and reduce pressure on the heart. Finally, the collision and shortness of breath suggest a collapsed lung and decreased oxygen transfer, which should be treated immediately by administering oxygen.

6.1 First Experiment: Static Bayesian Network and TNBN

We contrast the results of a static Bayesian network with the results of a TNBN. According to the accident event example [Hanks et al., 1995], the paramedics arrived ten minutes after that the collision occurred. The paramedics realize a clinical analysis of the patient's state. They determine that the *vitals signs* are *unstable* but the *pupils* are *not dilated*. With this information of the domain, we can update the probabilities of the events of the network using standard probability propagation techniques. Applying this evidence to the static Bayesian model, table 3 shows the update of the posterior marginal probabilities given the evidence.

Accident	Probability
Severe	0.20
Moderate	0.50
Mild	0.30
Head injury	Probability
True	0.20
False	0.80
Internal bleeding	Probability
Gross	0.45
Slight	0.53
None	0.02

Table 3. Marginal posterior probabilities for the static Bayesian network

Given the previous information, we can conclude the following: (i) the symptoms "vital signs unstable" and "pupils normal" are independent of the hour in which they were observed and (ii) it seems that the information is not sufficient for an adequate diagnosis. This shows that the information generated by a static Bayesian model is limited and that we require additional information of the domain to determine the probable effects in the patient's state. The static Bayesian model does not take into account the arrival time of the paramedics to the collision scene.

Now, we use the same example for demonstration of the importance of modeling the temporal aspects. In this case, it is very important the arriving timing of the paramedics and the time of occurrence of the events and the observations. We consider only two temporal nodes *pupils dilated* and *vitals signs unstables*. Given the time of arrival of the paramedics, the time interval of occurrence of both temporal nodes is 10 to 30 minutes. Table 4 shows the results of probability propagation with the TNBN model.

Accident	Probability
Severe	0.15
Moderate	0.68
Mild	0.17
Head injury	Probability
True	0.03
False	0.97
Internal bleeding	Probability
Gross	0.77
Slight	0.22
None	0.01

Table 4. Marginal posterior probabilities with the TNBN model

The results show the potential benefits of the use of the temporal information of the domain. The temporal causal model reduces the uncertainty and generates a diagnosis task of better quality. The conclusions generated by temporal causal model are consistent with the diagnosis generated by the paramedics (see section medical example). Additionally, the temporal causal model provides the severity of the internal bleeding and the accident event, these conclusions are difficult to obtain by the paramedics when they arrive to the collision scene.

Another comparison between the static and TNBN models is given by the entropy of the marginal posterior probabilities given by the following expression:

$$W = \sum_{i=1}^{N} - P_i \log_2(P_i)$$

Where N is the number of variable values and $\boldsymbol{P}_i$ is the marginal posterior probability of each value. Maximal entropy means equal probabilities, in this case the information is not sufficient for an adequate diagnosis. Table 5 shows the entropy for the static and temporal models.

Node	Static Bayesian	Temporal Node Bayesian
Accident	0.4417	0.3683
Head injury	0.2173	0.0585
Internal bleeding	0.3362	0.2521

Table 5. Entropy of marginal posterior probabilities

6.2 Second Experiment: Diagnosis with Scenarios

In this section we present an example of application of the diagnosis mechanism. Figure 6 shows a TNBN that represents knowledge about the injuries that occur in a patient caused by a "severe blow" and an "accident". The model is an extension of the accident example presented in the previous section. In this case it is considered that "head injuries" can be caused by a "severe blow" of a heavy object or by an "accident".

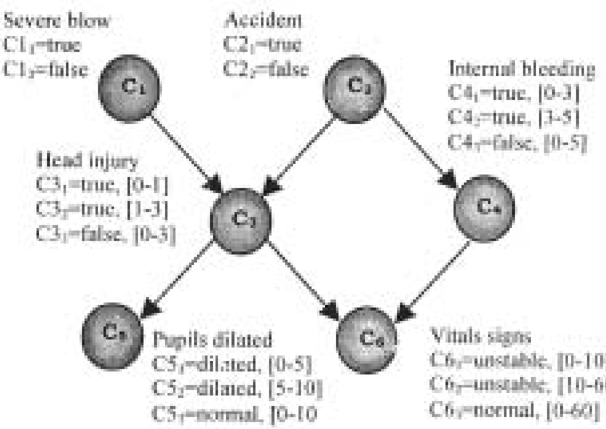


Figure 6. TNBN for the diagnosis example

1. We suppose that "vital signs unstable" (node **C6**) was detected at 10:45:00. This time is defined as the real time of occurrence t_{C6}. Since the initial node is a leaf node, it is not feasible to determine the interval of occurrence. Therefore, the evidence is propagated in the possible scenarios, in this case for two time intervals. Table 6 shows the marginal probabilities for both scenarios: **C6** changes in the time interval [0-10] (**$C6_1$**) and **C6** changes in the time interval [10-60] (**$C6_2$**).

It is required to check the previous events to determine which were presented before C_4 and in this way to determine the actual time inferred. The table 7 shows the marginal probabilities for the two scenarios.

Scenario 1		Scenario 2	
Node	Probability	Node	Probability
$C1_1$ {true, [10:34-10:45]}	0.80	$C1_1$ {true, [9:40-10:35]}	0.53
$C1_2$ {false}	0.20	$C1_2$ {false}	0.47
$C2_1$ {true, [10:30-10:45]}	0.57	$C2_1$ {true, [9:44-10:35]}	0.83
$C2_2$ {false}	0.43	$C2_2$ {false}	0.17
$C3_1$ {true, [10:35-45]}	0.63	$C3_1$ {true, [9:45-10:35]}	0.53
$C3_2$ {true, [10:35-45]}	0.33	$C3_2$ {true, [9:45-10:35]}	0.43
$C3_3$ {false, [10:35-45]}	0.04	$C3_3$ {false, [9:45-10:35]}	0.04
$C4_1$ {true, [10:35-45]}	0.42	$C4_1$ {true, [9:45-10:35]}	0.70
$C4_2$ {true, [10:35-45]}	0.38	$C4_2$ {true, [9:45-10:35]}	0.21
$C4_3$ {false, [10:35-45]}	0.20	$C4_3$ {false, [9:45-10:35]}	0.09
$C5_1$ {dilated, [10:35-55]}	0.38	$C5_1$ {dilated, [9:45-10:45]}	0.13
$C5_2$ {dilated, [10:35-55]}	0.49	$C5_2$ {dilated, [9:45-10:45]}	0.50
$C5_3$ {normal, [10:35-55]}	0.13	$C5_3$ {normal, [9:45-10:45]}	0.37
$C6_1$ {unstable, [10:45:00]}	1.0	$C6_1$ {unstable, []}	0
$C6_2$ {unstable, []}	0	$C6_2$ {unstable, [10:45:00]}	1.0
$C6_3$ {normal}	0	$C6_3$ {normal}	0

Table 6. Marginal probabilities for the scenarios $C6_1$ and $C6_2$

2. We find that the event **C4**, "internal bleeding" occurred at 10:34:00. Therefore we define the real time of occurrence as t_{C4} =10:34:00. The determination of the associated time interval will be given by the difference between the real time of occurrence for the events **C4** and **C6**: t_{C6} - t_{C4} = 10:45:00 - 10:34:00 = 11:00. This difference is compared with the time intervals defined for the node **C6**. The difference "t_{C6} - t_{C4} "corresponds to the time interval [10:00 to 60:00] and the evidence is associated with state **$C6_2$** {unstable, [10-60]}. This evidence **$C6_2$** is propagated and the marginal probabilities are updated. Now, table 7 shows the marginal probabilities of the nodes for the two possible scenarios: **$C4_1$** and **$C4_2$** since that the time interval of occurrence for the node **C4** is unknown. Table 8 shows the marginal probabilities for the two scenarios.

Scenario 1		Scenario 2	
Node	Probability	Node	Probability
$C1_1$ {true, [10:34-10:45]}	0.53	$C1_1$ {true, [9:40-10:35]}	0.53
$C1_2$ {false}	0.47	$C1_2$ {false}	0.47
$C2_1$ {true, [10:31-10:34]}	0.93	$C2_1$ {true, [10:29-10:31]}	0.87
$C2_2$ {false}	0.07	$C2_2$ {false}	0.13
$C4_1$ {true, [10:34:00]}	1.00	$C4_1$ {true, []}	0
$C4_2$ {true, []}	0	$C4_2$ {true, [10:34:00]}	1.0
$C4_3$ {false}	0	$C4_3$ {false}	0
$C6_1$ {unstable, []}	0	$C6_1$ {unstable, []}	0
$C6_2$ {unstable, [10:45:00]}	1.0	$C6_2$ {unstable, [10:45:00]}	1.0
$C6_3$ {normal}	0	$C6_3$ {normal}	0

Table 7. Marginal probabilities for the scenarios $C4_1$ and $C4_2$

It is required to check the previous events to determine which event occurred before the event **C4**, and in this way obtain the time interval for **C4**.

3. Checking the event history, we find that the event "accident" (**C2**) was detected at 10:30:00. Therefore, we defined the occurrence time as t_{C2} =10:30:00. The determination of the associated time interval will be given by the difference of the events **C4** and **C_2**: tc_4 - tc_2 = 10:34:00 - 10:30:00 = 04:00. This

difference is compared with the time intervals defined for **C6**. The difference te_4 -te_1 corresponds to the time interval [3:00 to 5:00] and the evidence is associated with state **$C4_2$** {true, [3-5]}. So the value **$C4_3$** is propagated and the marginal probabilities are updated.

As **C2** is a root node and by definition a root node does have temporal intervals, it is not necessary to obtain its temporal interval and the mechanism stops. Table 8 shows a summary of the events and their associated times.

Node	Time	Interval	Scenario
C6	10:45:00	Initial node	
C4	10:34:00	$t_{C6}-t_{C4}$ = 10:45:00 - 10:34:00 = 11:00	$C6_1$
C2	10:30:00	$t_{C4}-t_{C2}$ = 10:34:00 - 10:30:00 = 04:00	$C4_1$

Table 8. Summary of the events

7 Related Work

In this section, we review related work in temporal Bayesian networks and contrast it with our approach. Significant research has been done exploring probabilistic networks which are evaluated at each point in time. The network is arranged into "time slices" representing the system's complete state at a single point in time. Time slices are duplicated over a predetermined time grid representing the temporal range of interest. The "time net" of Kanazawa [Kanazawa, 1991] is a kind of Bayesian network with a formal declarative language of random variables for making inferences. The events are considered to occur at an instant of time while facts are considered to occur over a series of time points. The "Dynamic Belief Networks" (DBN) [Boyen and Koller, 1999] considers a dynamic Bayesian network where the network has certain structure at time *ti* and a different structure at time *ti+1*. The DBN is built dynamically, reflecting the dynamic changes in the environment. The "Network of exogenous events and endogenous changes" [Hanks et al., 1992] is a probabilistic model for reasoning about the system as it changes over time, both due to exogenous events and endogenous changes. An exogenous event generally refers to an instantaneous change in the process state. Endogenous changes are modeled using a local inference model, a simple arbitrary linear model. All the previous approaches are based on point models of time, and as such require that events occur instantaneously. It is difficult to consider that events take place at time points, often it is more natural to consider events taking place over time intervals.

Others authors propose the use of time intervals as the primitive temporal notion. Santos [1996] proposed a model based on the Temporal Abduction Problem (TAP). In TAP, each event has an associated interval during which the event occurs. Relationships between events are expressed as directed edges from cause to effect within a weighted directed acyclic graph structure. The TAP has strong interval-based temporal semantics, but lacks strong probabilistic semantics.

Later, Santos and Young [1996] proposed a new model, the probabilistic temporal network (PTN). Bayesian networks provide the probabilistic basis for the management of uncertainty. Allen's interval algebra and its thirteen relations provide the temporal basis [Allen, 1983]. The nodes of the network are temporal aggregates and the edges are the causal/temporal relationships between aggregates. Each aggregate represents a process changing over time. The temporal aggregates are temporal random variables, defined by an ordered pair (random variable plus Allen's intervals). These approaches are based on time-intervals and consider the temporal relationships that occur between the events. The time-constrained models provide a trade off between both strong uncertainty and temporal semantics.

Aliferis and Cooper [1996] proposed an extension of Bayesian networks called "Modificable Temporal Bayesian Networks with Single-granularity (MTBN-SG). A MTBN-SG is an extended time-sliced Bayesian network defined over a range of time points. The temporal graph is a directed graph (possibly cyclic) composed of nodes and arcs corresponding to 3 types of variables: ordinary, mechanism and time-lag quantifier variable. As indicated in the name, the MTBN-SG model only supports a single granularity for the size of the time step in any given network. The resulting graph can have cycles to allow expressions of recurrence and feedback. This model has great representation capacity of temporal and atemporal information. The problems with this model are that joint probability distribution is not compatible with the Bayesian model and that it only supports a single granularity for the size of the time step. Extending the model to support multiple granularity appears problematic, it is difficult to combine two events with different time ranges. Also, the acquisition of quantitative information appears a big problem because the excessive number of probabilities required.

In summary, previous probabilistic temporal models are, in general, based on variable states that are repeated at different times. A static probabilistic model of the system is built and repeated at various time points, and directed temporal links are drawn between nodes of the different "static" slices. The resulting models are quite complex for realistic applications, so they do not satisfy the knowledge acquisition and computational tractability criteria (see section 1). These models support a single granularity and it is difficult to extend them for multiple granularity. In contrast, the TNBN model is based on representing changes of state in each node. If the number of possible state changes for each variable in the temporal range is small, as it is in many practical problems, the resulting model is much simpler. This facilitates temporal knowledge acquisition and allows efficient inference using standard probability propagation techniques. Also, the model supports in a natural way multiple granularity, with different number of temporal intervals for each node, and different duration for each interval within a node.

8 Conclusions and Future Work

This paper presented the definition and application of an approach for dealing with uncertainty and time called Temporal Nodes Bayesian Network (TNBN). A TNBN generates a formal and systematic structure used to model the temporal evolution of dynamic process. TNBN model is an extension of Bayesian networks for dealing with temporal information. Each event or state change of a variable is associated with a time interval. The definition of the number of time intervals and their duration for each node is free (multiple granularity) and can be see as a trade off between the complexity and the accuracy needed for depicting the knowledge of the temporal domain.

Although many BN variants have been introduced for temporal modeling, we believe that the TNBN is a good candidate for diagnosis and prediction of events in real complex environments, such as medical diagnosis. The formalism satisfies the requirements of temporal knowledge acquisition, low computational cost and temporal expressiveness. The main difference with previous probabilistic temporal models is that the representation is based on state changes at different times instead of state values at different times. This makes the model much simpler in many applications in which there are few changes for each variable in the temporal range of interest.

The temporal information in a TNBN is relative, that is, there is not absolute temporal reference. We developed a mechanism for transforming the relative times to absolute. The temporal reasoning mechanism is based on the propagation of probabilities and gives the time of occurrence of events or state changes with some probability value. The mechanism has three main steps: (1) event detection and time interval definition; (2) evidence propagation through the net; and (3) determination of past and future events. If there is not enough information, the mechanism handles scenarios. These scenarios could be use as a set of possible alternatives, which will be reduced when another event occurs.

In order to demonstrate the ideas present in the article, the formalism was applied to the diagnosis and prediction of the consequences of the patient's state after an accident occurs, a case study presented by Hanks in 1995. The results are consistent with the medical diagnosis generated by the paramedics that arrive to the collision scene and they show the necessity to model temporal concepts in dynamic domains. The TNBN model is also simpler than chain state network of Hanks [Hanks et al., 1995].

The TNBN model can be used for the diagnosis of a cascade of anomalies arising with certain delays; this situation is typical in medicine or in the diagnosis of industrial processes. In contrast, dynamic bayesian networks, using time slices, seem more adequate for monitoring the evolution of a system that fluctuates around its normal state, specially if there is a cyclic pattern.

Our future work will focus on developing and validating our approach with additional experiments on prediction and diagnosis of events in other real-world domains, such as industrial process diagnosis. Also, we will try incorporate qualitative temporal constraints for knowledge acquisition and validation.

References

Aliferis C. F. and **Cooper G. F.**, "A Structurally and Temporally Extended Bayesian Belief Network Model: Definitions, Properties, and Modeling Techniques", in *Proc. 12th Conf. on Uncertainty in Artificial Intelligence*, 1996, pp. 28-39.

Allen J.F., "Maintaining Knowledge about Temporal Intervals", in *Communications of the ACM*, Vol. 26, No. 11, 1983, pp. 832-843.

Arroyo G., and **Sucar E.**, "A temporal Bayesian Network for diagnosis and prediction", in *Proc. 15th Conf. On Uncertainty in Artificial Intelligence*, 1999, pp. 13-20.

Boyen X. and **Koller D.**, "Exploiting the Architecture of Dynamic Systems", in *Proc. 16th National Conference on Artificial Intelligence*, Orlando Florida, 1999, pp. 18-22.

Brusoni V., **Console L.**, **Pernici B.**, and **Terenziani P.**, "Later: a general purpose manager of temporal information", in *Proc. 8th Intl. Symposium ISIM 94, Methodologies for Intelligent Systems*, 1994, pp. 255-364.

Dagum P., **Galper A.**, and **Horvitz E.**, "Dynamic Network Models for Forecasting", in *Proc. 8th Conf. on Uncertainty in Artificial Intelligence*, 1992, pp. 41-48.

Haddawy P., "A Logic of Time, chance, and action for representing plans", in *Artificial Intelligence*, Vol. 80, No. 2, 1996, pp. 243-308.

Hanks S., **Madigan D.** and **Gavrin J.**, "Probabilistic Temporal Reasoning with Endogenous Change", in *Proc. 11th Conf. on Uncertainty in Artificial Intelligence*, 1995, pp. 245-254.

Kanazawa K., "A Logic and time nets for probabilistic inference", in *Proc 10th National Conf. on Artificial Intelligence AAAI*, 1991, pp. 360-365.

Kanazawa K., **Koller D.** and **Russell S.**, "Stochastic Simulation algorithms for dynamic probabilistic networks", in *Proc. 11th Conf. On Uncertainty in Artificial Intelligence*, 1995, pp. 346-351.

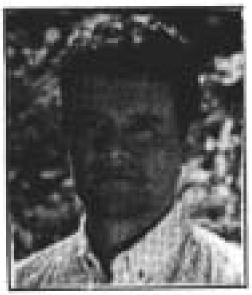
Kjaerulff U., "A Computational Scheme for Reasoning in Dynamic Probabilistic Networks", in *Proc. 8th Conf. on Uncertainty in Artificial Intelligence*, 1992, pp. 121-129.

Pearl Judea, *Causality Models, Reasoning, and Inference*, Cambridge University Press, 2000.

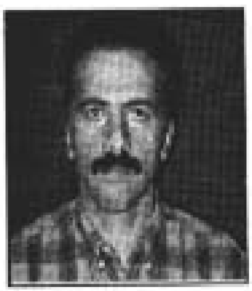
Santos E. Jr., "Unifying Time and Uncertainty for Diagnosis", in *J. Expt Theor. Artif. Intell.* Vol. 8, 1996, pp. 75-94.

Santos E. and Young J., "Probabilistic Temporal Networks", in report AFIT/EN/TR96-06 *Air Force Institute of Technology* USA, 1996.

Zweig G. and **Russell S.**, "Speech Recognition with Dynamic Bayesian Networks", in *Proc. 15th Conference on Artificial Intelligence*, Madison Wisconsin, 1998, pp. 26-30.



Gustavo Arroyo-Figueroa *received his B.S. and M.S. degrees in Chemical Engineering from the Celaya Institute of Technology, Mexico, and a Ph.D degree in Computer Science (Artificial Intelligence) from the Monterrey Institute of Technology (ITESM) Campus Morelos, Mexico. He is currently the Head of the Department of Information Systems at the Electric Research Institute, Mexico. His research interest lies in the fields of uncertainty management, knowledge-based systems, and process control, and their applications to the electric power industry. He is a member of the National Researchers Systems of Mexico.*



Luis Enrique Sucar Succar *received his B.S. in Electronics and Communications from the Monterrey Institute of Technology, in Monterrey, Mexico, an M.Sc. in Electrical Engineering from Stanford University, and a Ph.D. degree in Computer Science from the Imperial College of Science Technology and Medicine, University of London, England. He is currently the Head of Department of Computer Science at the Monterrey Institute of Technology, Campus Cuernavaca, Mexico. His major research interests include uncertainty management, knowledge-based systems, machine learning, mobile robotics and computer vision. He is a member of the National Researchers Systems of Mexico and president of the Mexican AI Society.*