Comparison of two kinds of event Bayesian networks: a case study

S. F. Galán
Dpto. de Inteligencia Artificial
E.T.S.I. Informática (UNED)
Juan del Rosal, 16
28040 Madrid, Spain
seve@dia.uned.es

G. Arroyo-Figueroa
IIE - GSI
Av. Reforma 113, Cuernavaca, Mor.
62490 Mexico
garroyo@iie.org.mx

F. J. Díez
Dpto. de Inteligencia Artificial
E.T.S.I. Informática (UNED)
Juan del Rosal, 16
28040 Madrid, Spain
fjdiez@dia.uned.es

L. E. Sucar
INAOE
Luis Enrique Erro, 1
62020 Tonantzintla, Puebla, Mexico
esucar@itesm.mx

January 18, 2007

Abstract

Temporal Nodes Bayesian Networks (TNBNs) and *Networks of Probabilistic Events in Discrete Time* (NPEDTs) are two different types of *Event Bayesian Networks* (EBNs). Both are based on the representation of uncertain events, alternatively to *Dynamic Bayesian Networks*, which deal with real-world dynamic properties. In a previous work, Arroyo-Figueroa and Sucar applied TNBNs to the diagnosis and prediction of the temporal faults that may occur in the steam generator of a fossil power plant. We present an NPEDT for the same domain, along with a comparative evaluation of the two networks. We examine different methods suggested in the literature for the evaluation of Bayesian networks, analyze their limitations when applied to this temporal domain, and suggest a new evaluation method appropriate for EBNs. In general, the results show that, in this domain, NPEDTs perform better than TNBNs, possibly due to be the finer time granularity used in the NPEDT.

Keywords: Bayesian networks, temporal reasoning, fault diagnosis and prediction, evaluation, case study.

1 Introduction

Bayesian networks (BNs) [21] have been usually applied without considering an explicit representation of time. However, important efforts have also been made to model temporal processes by means of BNs.

The usual way to apply BNs to dynamic domains consists in discretizing time and creating an instance of each random variable for each point in time. In the formalism of *Dynamic Bayesian Networks* (DBNs) [8, 9, 18, 19], initially a static causal model is built. Then, a copy of this model is generated for each instant in the temporal range of interest. Finally, links between nodes in adjacent static networks are established. In this way, a DBN obeys the Markov property: The future is conditionally independent of the past given the present.

While in a DBN the value of a variable represents the state of a real-world property at a particular time, in other types of temporal BNs[1], like *Temporal Nodes Bayesian Networks* (TNBNs) [3] or *Networks of Probabilistic Events in Discrete Time* (NPEDTs) [13], each value of a variable represents the time at which a certain event may occur. Both formalisms can be grouped under the term "Event Bayesian Network" (EBN). For domains involving temporal fault diagnosis or prediction, EBNs present some advantages over DBNs, since, in the former, faults can be easily represented through events.

In a previous paper [13], we discussed the similarities and differences between TNBNs and NPEDTs, but no empirical study was carried out in order to contrast their performance in a real-world domain. In this paper, we present a comparative evaluation of TNBNs and NPEDTs by means of a case study: the diagnosis and prediction of the temporal faults that may occur in the steam generator of a fossil power plant. Arroyo-Figueroa and Sucar applied TNBNs to this domain [4]. The application of NPEDTs to the same domain is presented in this paper.

The rest of the paper is organized as follows. Sec. 2 offers an overview of TNBNs and NPEDTs. Sec. 3 reviews the main methodologies used for the evaluation of BNs, and proposes a new evaluation method applicable to TNBNs and NPEDTs. Sec. 4 describes the application of TNBNs and NPEDTs to a case study: the diagnosis and prediction of the possible temporal faults taking place in the steam generator of a fossil power plant. Sec. 5 presents the empirical results for the evaluation of the two networks. Finally, Sec. 6 includes some remarks, and summarizes the main results of this work.

2 Temporal Bayesian networks

This section outlines the characteristics of the two types of temporal BNs that we empirically compare in this work: TNBNs and NPEDTs. These two approaches could form a new subgroup of temporal BNs that we call *Event Bayesian Net-*

[1] For historical reasons, some authors use the term "temporal BN" as a synonym of "dynamic BN", despite the fact that there exist other types of BNs for temporal reasoning (see [13] for a review of the different types of temporal BNs).

works (EBNs). We will also give some details regarding the main aspects that differentiate EBNs from DBNs; these differences make the former more appropriate than the latter for modeling domains like the industrial one considered in this paper.

2.1 Bayesian networks

BNs have been successfully applied to the modeling of problems involving uncertain knowledge. A BN is an acyclic directed graph whose nodes represent random variables, and whose links define probabilistic dependencies between variables. BNs specify dependence and independence relations in a natural way through the network topology. Those relations are quantified by associating a conditional probability table (CPT) to each node. A CPT defines the probability of a node given each possible configuration of its parents. Probability trees [6] allow for a compact representation of CPTs. Diagnosis or prediction with BNs consists in fixing the values of the observed variables and computing the posterior probabilities of some of the unobserved variables.

In the general case, it is necessary to assign each node in a BN a set of conditional probabilities that grows exponentially with the number of its parents. This complicates the acquisition of the parameters, their storage, and the propagation of evidence. For these reasons, causal interaction models —called canonical models [10]— were developed in order to simplify both BN construction and probability computation. A well-known example is the noisy OR-gate [21, 10], which requires just one independent parameter per parent.

In some domains, like medicine or industry, diagnosis and prediction require a representation combining uncertainty and time. Temporal information between observations and manipulations is usually critical for a correct diagnosis. For example, in medicine, representing and reasoning about time is crucial for many tasks like prevention, diagnosis, therapeutic management, or prognosis. In industrial domains, it is also critical for diagnosis and prediction of events and disturbances.

The usual method of applying BNs to the modeling of temporal processes is based on the use of DBNs [8, 9, 18, 19]. In a DBN, time is discretized, and an instance of each random variable is created for each point in time. While in a DBN the value of a variable V_i represents the state of a real-world property at time t_i, in an EBN —either a TNBN or an NPEDT (see below)— each value of a variable represents a possible occurrence time for a certain event. Therefore, EBNs are more appropriate for temporal fault diagnosis or prediction, because only one variable is necessary for representing the occurrence of a fault and, consequently, the networks involved are much simpler than those obtained by using DBNs (see [13, Sec. 4]). However, DBNs are more appropriate for monitoring tasks, since they explicitly represent the state of the system at each moment.

2.2 Temporal Nodes Bayesian Networks

Arroyo-Figueroa and Sucar developed TNBNs [3] as a type of temporal BN, and applied this formalism to fault diagnosis and prediction for the steam generator of a fossil power plant [4].

A TNBN is a BN in which each node represents a temporal event or change of state of a variable. There is at most one state change for each variable in the temporal range of interest. The value taken on by the variable represents the interval in which the event occurs. Time is discretized in a finite number of intervals, allowing a different number and duration of intervals for each node (multiple granularity). Each interval defined for a child node represents the possible delays between the occurrence of one of its parent events (cause) and the corresponding child event (effect). Therefore, this model makes use of **relative time** in the definition of the values associated to each temporal node with parents.

Due to the use of relative time, there is an asymmetry in the way evidence is introduced in the network: The occurrence of an event represented by a node without parents constitutes direct evidence, while evidence about a node with parents is introduced by considering several scenarios. When an initial event is detected, its occurrence time fixes the network temporally. A TNBN permits reasoning about the probability of occurrence of certain events, for diagnosis or prediction, using standard probability propagation techniques for BNs.

TNBNs lack a formalization of canonical models for temporal processes; consequently, all conditional probabilities must be given explicitly. Another shortcoming of TNBNs is that each value defined for an effect node, which is associated to a determined time interval, means that the effect has been caused during that interval by only one of its parent events. However, this assumption is not appropriate in some domains where a child event can be simultaneously caused by several of its parents.

2.3 Networks of Probabilistic Events in Discrete Time

In NPEDTs [13], each variable represents an event that can occur at most once. However, they differ from TNBNs in that time is discretized by adopting the same temporal unit (seconds, minutes, etc.) for all the variables. The value taken on by a variable indicates the **absolute time** at which the event occurs.

Formally speaking, a temporal random variable V in the network can take on a set of values $v[i]$, $i \in \{a, \ldots, b, never\}$, where a and b are instants — or intervals— defining the limits of the temporal range of interest for V. The links in the network represent temporal causal mechanisms between neighboring nodes. Therefore, each CPT represents the most probable delays between the parent events and the corresponding child event. For the case of general dynamic interaction in a family of nodes, giving the CPT involves assessing the probability of occurrence of the child node over time, for each temporal configuration of the parent events. In a family of n parents $X_1, \ldots, X_n$ and one child Y, the CPT is given by $P(y[t_Y] \mid x_1[t_1], \ldots, x_n[t_n])$ with $t_Y \in \{0, \ldots, n_Y, never\}$

and $t_i \in \{0, \ldots, n_i, never\}$. The joint probability is given by the product of all the CPTs in the network. Any marginal or conditional probability can be derived from the joint probability.

In many domains, the dynamic causal relations have the property of *time invariance*:

$$P(y[t_Y + \Delta t] \mid x_1[t_1 + \Delta t], \ldots, x_n[t_n + \Delta t]) = P(y[t_Y] \mid x_1[t_1], \ldots, x_n[t_n]).$$

If we consider a family of nodes with n parents, and divide the temporal range of interest into i instants, in the general case the CPT associated to the child node requires $O(i^{n+1})$ independent conditional probabilities. In real-world applications, it is difficult to find a human expert or a database that allows us to create such a table, due to the exponential growth of the set of required parameters with the number of parents. For this reason, temporal canonical models were developed as an extension of traditional canonical models. In this fault-diagnosis domain, we only need to consider the *temporal noisy OR-gate* [13]. We have no *temporal noisy AND-gate* (see also [13]).

2.4 Comparison of DBNs, TNBNs, and NPEDTs

DBNs were the first formalism that used BNs to carry out temporal reasoning, and nowadays constitute the usual method of applying BNs to the modeling of temporal processes. EBNs are relatively recent methods, and have been applied in domains like industry [4] or medicine [14].

In DBNs there is an instance of each variable V for each point in time, and each random variable V_i represents the state of a property at time t_i. In EBNs, variables represent events, and each value of a variable is associated to a possible occurrence time for a certain event.

In DBNs, initially a static model is built, and then a copy of this model is generated for each instant within a certain time range. Links between nodes in adjacent static networks can be established, so that the DBN obeys the Markov property: The future is conditionally independent of the past given the present. EBNs use just one variable to represent an event, no copies of the initial causal model are needed, and no assumption about the Markovian nature of the processes involved needs to be made.

In DBNs, the representation of irreversible processes, such as fault propagation, requires introducing new nodes called *memory nodes* (cf. [17, Sec. 4.3.1]), which results in models with a high complexity. However, faults can be easily represented through events in EBNs. On the contrary, DBNs are more appropriate for dealing with monitoring tasks, since they explicitly represent the state of the system at each moment.

Inference in DBNs consists in estimating the state of the past, the present, or the future, given all the evidence up to the current time; these tasks are called *smoothing*, *monitoring*, and *prediction*, respectively. If there are long observation sequences, smoothing in DBNs becomes computationally complex (see [5]). Unlike in DBNs, in EBNs the presence of additional evidence from the future does not imply an increase in the complexity of the inference process.

The most significant differences between TNBNs and NPEDTs are:

- In a TNBN, intervals may have different durations, even within the same variable. In an NPEDT, usually a unique interval duration is defined for the whole network.
- In a TNBN, each value associated to a node is defined as a time interval relative to the occurrence of one of its parents. In an NPEDT, time is absolute.
- NPEDTs use temporal canonical models, while in a TNBN all the conditional probabilities must be explicitly given.
- The occurrence of an event at a particular instant does not constitute direct evidence in a TNBN and, therefore, several scenarios need to be considered. NPEDTs do not need to analyze different scenarios.

3 Evaluation method

The expression "evaluation of a BN" could in short be defined as "estimation of performance of a BN" or "estimation of quality of recommendations obtained by using a tool based on a BN". Evaluation constitutes a requisite for the practical application of BNs. Conventional BN evaluation consists of obtaining a set of cases from records or from experts, querying the network for a diagnostic or predictive recommendation for each case, and determining how well the recommendations agree with the actual results known for the cases. There are two important issues with regard to the process of evaluation of a BN: on the one hand, the selection of the cases and, on the other hand, the method for measuring the performance. The cases can be obtained in two different ways:

- from the BN itself, or
- from a database or with the help of an expert in the domain.

The assessment of performance can be addressed following two distinct strategies:

- by relying on expert opinion to judge the results produced by the BN, or
- by executing a mathematical method whose entries are the cases available and the inferential results.

3.1 Previous work on evaluation of BNs

Przytula et al. [22] propose an approach that automatically generates its own cases in a way that guarantees a complete evaluation of the model. Their approach uses Monte Carlo simulation to automatically generate diagnostic cases that uniformly cover all the parts of the BN model. Certain visualization tools

allow inference results to be easily analyzed and interpreted by the experts. They apply this approach to the evaluation of BNs used for diagnosis of component defects in complex systems. In general, this approach is a good alternative in situations where the available cases are incomplete or erroneous, or when experts are biased in their selection of cases.

The most widespread method for BN evaluation consists in running real cases (or cases constructed by an expert) through the BN, presenting those cases to an expert, and asking the expert to analyze the performance of the BN for each case. The expert opinion may then be used to make modifications in the network, so that it reflects the behavior of the real system considered. Some examples of this approach can be found in [1, 12, 15, 20].

Other evaluation methods do not need the help of any expert in order to be applied. Let us examine some significant examples:

1. One simple method is based on comparing the value of a variable in a case with the value that maximizes its posterior probability distribution. Given a set of cases, a score can be obtained, namely the *accuracy*. A similar method involves calculating the *total square error* (*Brier score*) for each variable V:

   $$BS(V) = \sum_{i=1}^{N} \left(1 - P\left(v_i \mid \mathbf{e}\right)\right)^2$$

   where N is the number of cases, v_i is the value taken on by V in case i, and $\mathbf{e}$ is the evidence. Examples of the application of these methods can be found in [23, 2].

2. MammoNet [16] is a BN for the diagnosis of breast cancer. The evaluation of MammoNet is based on the existence in the network of a diagnosis variable called "Breast cancer", which has two possible states: *present* and *absent*. MammoNet was tested from 77 cases by following the next steps:

   - posterior probability computation of "Breast cancer" for each case, and
   - analysis of the results through the receiver operating characteristic (ROC) curve.

   An ROC curve depicts the true-positive fraction of diagnosed cases against the false-positive fraction of diagnosed cases. The area under the ROC curve determines the usefulness of MammoNet at discriminating between the patients with cancer and those without cancer in the set of 77 cases studied. The closer the area is to 1, the better the discrimination is. The area under the ROC curve reported for MammoNet is 0.881 ± 0.045 .

3. Dagum and Galper [7] describe a system for forecasting sleep apnea by means of *Dynamic Network Models* (DNMs) [8]. DNMs are a probabilistic

	Results analyzed by experts	*Results analyzed mathematically*
Cases generated from the BN itself	**Przytula et al. [22]**	
Real cases or cases generated by experts	**MUNIN [1], DIAVAL [12], DIABNET [15], HEPAR [20]...**	**accuracy, Brier, MammoNet [16], Dagum & Galper [7], and our method (Sec. 3.2)**

Table 1: Classification of methods for BN evaluation.

forecasting methodology that generalizes DBNs (see Sec. 2.4) by allowing structure and conditional probabilities to be updated as new evidence becomes available, and dependencies between nonadjacent time points to be included in the model. The continuous variables present in the sleep apnea system are conveniently discretized, and the data for the construction (27.000 recordings) and evaluation (7.000 recordings) of the system were collected from a patient suffering from sleep apnea. The inference results are analyzed by using several measures from statistical forecasting:

Prediction error at time t: difference between the observation at time t, o_t, and the *expected forecast value* for time t, p_t. For variable Z_t, this quantity can be expressed as

$$PE(Z_t) = \hat{z} - \sum_{Z_t=z} z \cdot P(Z_t = z \mid \mathbf{e}_{t-1}) \tag{1}$$

where $\hat{z}$ represents the observed value for Z_t, and $\mathbf{e}_{t-1}$ are all the observations made up to time $t-1$.

Mean prediction error:

$$\frac{1}{N} \sum_{t=1}^{N} \frac{o_t - p_t}{o_t}$$

where N is the number of time instants considered.

Table 1 shows a classification of the evaluation methods examined so far. We have organized them according to two concepts: the way the cases are generated, and the way the testing results are analyzed.

The new method for evaluating EBNs belongs to the group that uses real cases and analyzes inferential results through a certain mathematical measure. The rest of the methods included in that group cannot be applied to EBNs for the following reasons:

- Both accuracy and Brier score only consider one of the values of a variable in each step of the testing process: the value that maximizes the posterior

probability distribution, or the value taken on by the variable in a case, respectively. We have observed that this fact tends to deteriorate the test results in our dynamic domain, where posterior probabilities for a variable can be distributed among a considerable number of states.

- The existence of a binary variable in MammoNet whose posterior probability is used to decide the diagnostic result makes it possible to use ROC curves as a method for evaluating the network. This is not the case in our systems because several possible initial faults have to be considered in order to elaborate a complete diagnosis.

- The system for sleep apnea by Dagum and Galper deals with variables like "Heart rate", "Chest volume", or "Blood oxygen concentration". These continuous variables, once discretized, are analogous to the temporal nodes of a TNBN, or to the temporal events of an NPEDT. However, the equivalence is not complete because the variables enumerated above lack a state analogous to *never*. This fact has to be taken into account in the evaluation of EBNs.

3.2 Evaluation method for EBNs

In this section, we introduce a new domain-independent evaluation method for EBNs. The method is based on the calculation of a mathematical measure that quantifies how well an EBN performs.

In general, we suppose that a total of C cases have been collected for evaluation purposes. Each case consists of a list

$$((event_1, t_1), (event_2, t_2), \ldots, (event_K, t_K))$$

where t_i is the occurrence time for $event_i$. There are K possible events. If $event_i$ did not occur then $t_i = never$.

Our first attempt to quantify the performance of an EBN was carried out as follows. For each node or event X not included in the evidence:

1. Calculate $P(X \mid \mathbf{e})$, the posterior probability of node X, given evidence $\mathbf{e}$.

2. For each case whose evidence is $\mathbf{e}$, obtain a *measure of error*, $ME(P(X \mid \mathbf{e}), \hat{t}_X)$ —see below—, which quantifies the difference between the posterior probability and $\hat{t}_X$, the real (or simulated) occurrence time for X.

3. Calculate the mean and variance of the measures of error obtained in the previous step.

Given a *probability density function* for a variable V, $f_V(t)$, if we know that V took place at $\hat{t}_V$, a possible measure of error is

$$ME(f_V(t), \hat{t}_V) = \int_0^{+\infty} f_V(t) \cdot \left| t - \hat{t}_V \right| \, dt. \tag{2}$$

This measure represents the average time distance between an event occurring at $\hat{t}_V$ and another one that follows distribution $f_V(t)$. For example, if $f_V(t)$ is a constant distribution between t_i and t_f (with $t_i < t_f$):

$$f_V(t) = \begin{cases} 0 & \text{if} \quad t < t_i \\ \frac{p}{t_f - t_i} & \text{if} \quad t_i \leq t \leq t_f \\ 0 & \text{if} \quad t > t_f \end{cases} \tag{3}$$

then

$$ME(f_V(t), \hat{t}_V) = \begin{cases} p \cdot \left(\frac{t_i + t_f}{2} - \hat{t}_V \right) & \text{if} \quad \hat{t}_V \leq t_i \\ \frac{p}{t_f - t_i} \cdot \left[\left(\hat{t}_V - \frac{t_i + t_f}{2} \right)^2 + \left(\frac{t_f - t_i}{2} \right)^2 \right] & \text{if} \quad t_i \leq \hat{t}_V \leq t_f \\ p \cdot \left(\hat{t}_V - \frac{t_i + t_f}{2} \right) & \text{if} \quad \hat{t}_V \geq t_f \end{cases} .$$

Note that *ME* is equivalent to the *prediction error* defined in Eq. 1 if time is the variable considered. The probability distribution $f_V(t)$ can be directly obtained from $P(V \mid \mathbf{e})$ in an NPEDT, while in a TNBN it is necessary to know which parent node is really causing V, which can be deduced from the information contained in the corresponding case.

Two problems arise when we try to apply Eq. 2 to a node of either a TNBN or an NPEDT:

- If, given a certain case, event V does not occur, we can only assign $\hat{t}_V$ the value $+\infty$; as a consequence, the integral in Eq. 2 will be equal to $+\infty$. However, if $P(V = v[never] \mid \mathbf{e}) > 0$, we would expect to obtain a finite measure of error, for example, $1 - P(V = v[never] \mid \mathbf{e})$.
- If $P(V = v[never] \mid \mathbf{e}) > 0$, the value t in Eq. 2 cannot be precisely defined for $V = v[never]$; if we supposed that $t = +\infty$, the integral would be equal to $+\infty$ if $\hat{t}_V$ takes a finite value. However, if $\hat{t}_V$ is finite, we would expect to obtain an infinite measure of error only in the case that $P(V = v[never] \mid \mathbf{e}) = 1$, but not in the case that $P(V = v[never] \mid \mathbf{e}) < 1$.

In order to avoid these two problems, we adopted an alternative point of view: Instead of a *measure of error*, we used a *measure of proximity* between $P(V \mid \mathbf{e})$ and $\hat{t}_V$ for evaluating the networks. Given a probability density function for a variable V, $f_V(t)$, if V took place at $\hat{t}_V$, the measure of proximity we used is

$$MP(f_V(t), \hat{t}_V) = \int_0^{+\infty} \frac{f_V(t)}{1 + \left(\frac{t - \hat{t}_V}{c} \right)^2} \, dt \tag{4}$$

where c is an arbitrary constant. We have selected this function because it has the following desirable properties:

1. As $\int_0^{+\infty} f_V(t) \, dt = 1$, $0 \leq MP \leq 1$. $MP = 1$ if and only if $f_V(t)$ is a Dirac delta function at $\hat{t}_V$.

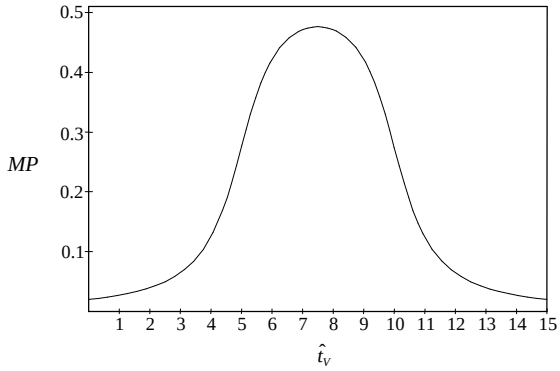

Figure 1: *MP*s (with $c = 1$) for a constant distribution between $t_i = 5$ and $t_f = 10$, with $p = 1$.

2. When $t = \hat{t}_V$, the value of the integrand is $f_V(\hat{t}_V)$; however, as $\left|t - \hat{t}_V\right| \longrightarrow +\infty$, the integrand approaches 0 regardless of the value of f_V.

3. If, given a case, event V does not occur ($\hat{t}_V = +\infty$), the integrand is zero when $t \neq never$, and we consider that $MP = P(V = v[never] \mid \mathbf{e})$.

4. If $\hat{t}_V$ takes on a finite value and $P(V = v[never] \mid \mathbf{e}) > 0$, we consider that the contribution of $V = v[never]$ to MP is 0.

5. When the density function is constant inside an interval, MP can be easily calculated. Given the constant distribution defined in Eq. 3,

$$MP\left(f_V(t), \hat{t}_V\right) = \frac{p \cdot c}{t_f - t_i} \cdot \left(\arctan\left(\frac{t_f - \hat{t}_V}{c}\right) - \arctan\left(\frac{t_i - \hat{t}_V}{c}\right)\right). \quad (5)$$

   Fig. 1 depicts MP (with $c = 1$) against $\hat{t}_V$ for $p = 1$, $t_i = 5$ and $t_f = 10$, i.e., all the probability is uniformly distributed in the interval $[5, 10]$. As expected, the maximum measure of proximity appears when $\hat{t}_V = \frac{t_i + t_f}{2}$, 7.5 in this case.

Note that properties 3 and 4 of this measure of proximity solve the two problems of the measure of error mentioned above. Since TNBNs and NPEDTs are discrete-time models, we calculate MP (given by Eq. 4) by adding the contributions of each interval associated to the values of node V and the contribution of value *never*. $P(V \mid \mathbf{e})$ defines a constant probability distribution over each of the intervals defined for V. Apart from EBNs, the evaluation method introduced in this section can be applied to any formalism that:

- represents time explicitly (although EBNs use discrete time, the evaluation method can also be applied to formalisms using continuous time), and
- is based on events whose uncertainty is expressed through probabilities.

4 Case study: a fossil power plant

4.1 The domain

Steam generators of fossil power plants are exposed to disturbances that may provoke faults. The appearance and propagation of these faults is a non-deterministic dynamic process whose modeling requires representing both uncertainty and time.

We are interested in studying the disturbances produced in the drum level control system of a fossil power plant. As shown in Fig. 2, the drum provides steam to the superheater, and water to the water wall of a steam generator. The drum is a tank with a steam valve at the top, a feedwater valve at the bottom, and a feedwater pump which provides water to the drum.

There are four potential disturbances that may occur in the drum level control system: a power load increase (LI), a feedwater pump failure ($FWPF$), a feedwater valve failure ($FWVF$), and a spray water valve failure ($SWVF$). These disturbances may provoke the events shown in Fig. 3.

Note that node DRL represents two possible events: either an increase of the drum water level as a consequence of a feedwater flow increase (FWF), or a reduction of the level due to a steam flow increase (STF). In this domain, we consider that an event occurs when a signal exceeds its specified limit of normal functioning.

4.2 A TNBN for this domain

The causal model used by Arroyo-Figueroa and Sucar in the construction of their model is shown in Fig. 3. The network structure was defined based on the knowledge of an expert operator.

The definition of the time intervals for each temporal node was obtained from knowledge about the process dynamics, combined with data from a simulator of a 350 MW fossil power plant [4]. The parameters of the TNBN were estimated from data generated by the simulator. A total of more than 900 cases were simulated. Approximately 85% of the data were devoted to estimate the parameters, while the remaining 15% was used in the evaluation of the model, which we discuss later in the paper.

A complete description of the application of TNBNs to this industrial domain can be found in [4].

4.3 An NPEDT for this domain

This section presents the application of NPEDTs to the same industrial domain of fossil power plants, which constitutes one of the contributions of this paper.

4.3.1 New causal model

The causal model depicted in Fig. 3, which was used by Arroyo-Figueroa and Sucar in the construction of the TNBN, needs to be transformed in order to

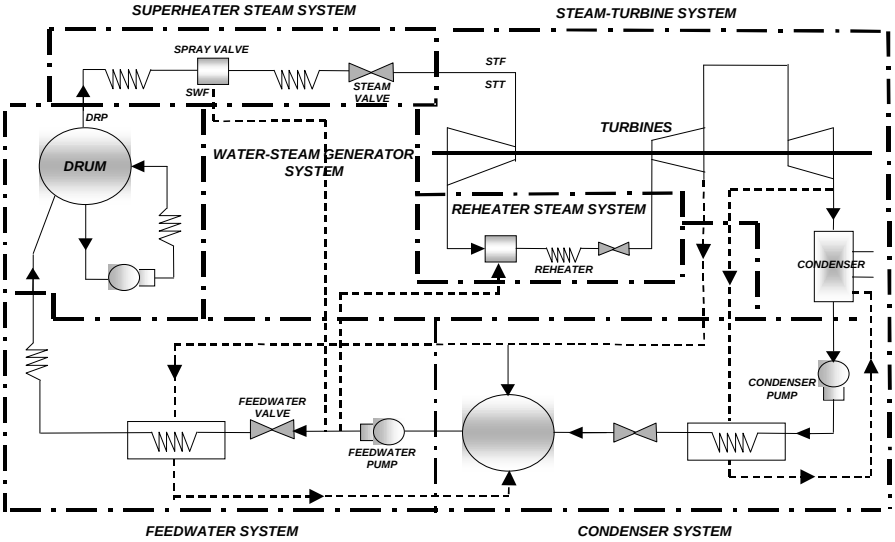

Figure 2: Steam generator system (taken from [3]).

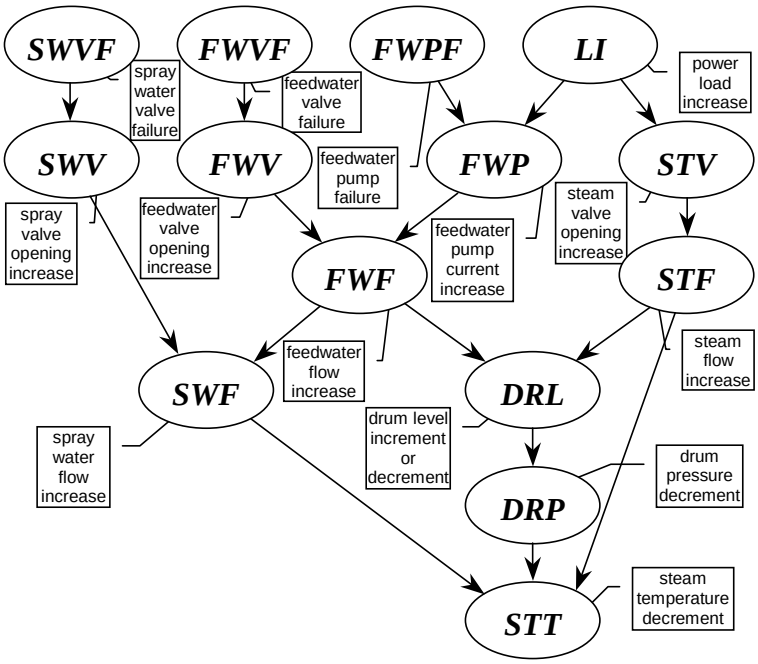

Figure 3: Causal graph of the steam generator (taken from [3]).

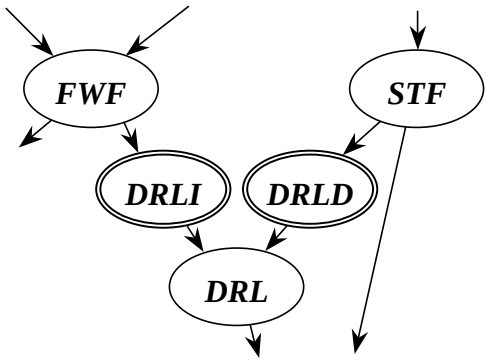

Figure 4: New graph with additional simple events.

generate an NPEDT. As mentioned above, node *DRL* (drum level increment or decrement) is not associated with a unique event. In the new causal model, we introduce two parent nodes for *DRL*, *DRLI* and *DRLD* (see Fig. 4), representing an a priori increment and decrement in the drum level, respectively. Whether the drum level really increases or decreases is established a posteriori in node *DRL*. The way to define the CPT for node *DRL* is explained below. The variables associated to nodes *DRLI* and *DRLD* are not directly observable. Consequently, the evidence on the drum level is introduced in node *DRL*.

4.3.2 Numerical parameters of the NPEDT

For the NPEDT, we consider a time range of 12 minutes, and divide this period into 20-second intervals. Therefore, there are 36 different intervals in which any event in Fig. 3 may occur. Given a node E, its associated random variable can take on values $\{e[1], \ldots, e[36], e[never]\}$, where $e[i]$ means that event E takes place in interval i, and $e[never]$ means that E does not occur during the time range selected. For example, $SWF = swf[3]$ means "spray water flow increase occurred between seconds 41 and 60". As the values of any random variable in the network are exclusive, its associated events can only occur once over time. This condition is satisfied in the domain, since the processes involved are irreversible. Without the intervention of a human operator, any disturbance would provoke a shutdown of the fossil power plant.

We use the temporal noisy OR-gate as the model of causal interaction in the network. In this model, each cause acts independently of the rest of the causes to produce the effect. This property of independence of causal interaction is satisfied in the domain, according to the experts' opinion. For example, a feedwater flow increment (FWF) can be produced by either a feedwater valve opening increase (FWV) or by a feedwater pump current augmentation (FWP); both processes act independently of each other.

For nodes without parents, each value is assigned a prior probability close to

10	10	10	10	11	11	11	11	12
12	12	12	12	12	12	12	13	13
13	13	15	15	15	15	18	18	18
18	20	20	20	20	20	20	20	20
21	21	21	21	21	21	21	21	22
22	22	22	27	27	27	27	29	29
29	29	29	29	29	29	29	29	29
29	31	31	31	31	32	32	32	32

Table 2: Delays between $FWPF$ and FWP for 72 simulations.

zero. Computing the CPT for a node Y in the network (except for node DRL) requires specifying

$$c_{y[k]}^{x_i[j_i]} \equiv P(y[k] \mid x_i[j_i], x_l[never], l \neq i)$$

for each possible delay, $k - j_i$, between cause X_i and Y, when the rest of the causes are absent. Therefore, given that X_i takes place during a certain 20-second **interval**, it is necessary to specify the probability of its effect Y taking place in the same interval —if the rest of the causes are absent—, the probability of Y taking place in the next interval, and so on. These parameters were estimated from the same dataset used by Arroyo-Figueroa and Sucar in the construction of their TNBN. The simulator provided us with delays between the occurrence of X_i and that of Y. Table 2 shows the delays —in seconds— obtained for arc $FWPF \rightarrow FWP$ for 72 different simulations.

The numerical data in Table 2 allow us to obtain $\tilde{c}_Y^{X_i}(\Delta t)$. Parameters $\tilde{c}_Y^{X_i}(\Delta t)$ represent, given that X_i takes place at a certain **instant**, the probability of its effect Y taking place in the next 20-second interval ($\Delta t = 1$) if the rest of the causes are absent, the probability of Y taking place in the interval after the next interval ($\Delta t = 2$), and so on. From the data in Table 2, where double lines separate delays belonging to different 20-second intervals, $\tilde{c}_{FWP}^{FWPF}(\Delta t = 1) = \tilde{c}_{FWP}^{FWPF}(\Delta t = 2) = 0.5$. It can be proved (see [14, Sec. 3.2]) that

$$c_{y[j_i+\Delta t]}^{x_i[j_i]} = \frac{\tilde{c}_Y^{X_i}(\Delta t) + \tilde{c}_Y^{X_i}(\Delta t + 1)}{2}. \tag{6}$$

In this way, any CPT can be computed from the available data. According with the experts, we suppose that our domain satisfies the property of time invariance:

$$c_{y[k]}^{x_i[j_i]} = c_{y[k+\Delta t]}^{x_i[j_i+\Delta t]}.$$

Besides, the effect cannot occur if none of its causes are present:

$$c_{y[k=never]}^{x_i[j_i=never]} = 1.$$

Finally, the effect cannot precede the cause:

$$c_{y[k]}^{x_i[j_i>k]} = 0.$$

$P(DRL \mid drli[j], drld[k])$	
0	$if \;\; DRL = \{d[l] \; or \; i[l]\} \; and \; l < j \; and \; l < k$
1	$if \;\; DRL = i[l] \; and \; j = l \; and \; j < k$
1	$if \;\; DRL = d[l] \; and \; k = l \; and \; k < j$
0.5	$if \;\; DRL = \{d[l] \; or \; i[l]\} \; and \; (j = k = l)$

Table 3: Conditional probabilities for node *DRL*.

Only node *DRL* does not interact with its parents through the temporal noisy OR model. This node can take on values $\{d[1], \ldots, d[36], i[1], \ldots, i[36], unchanged\}$, where d and i stand for "decrement" and "increment", respectively. Table 3 shows how the CPT for node *DRL* was established. The value 0.5 in Table 3 corresponds to the case in which there are —a priori— a drum level increment and decrement during the same 20-second interval. As we ignore which event happens first, both of them are assigned the same probability.

4.3.3 An example

Consider the possible faults or events occurring in the steam generator of a fossil power plant from 10:00:00 a.m. to 10:12:00 a.m. . As shown in Fig. 5, we divide this period into 20-second intervals. We are interested in determining, from the available evidence, the most probable initial fault (*SWVF*, *FWVF*, *FWPF*, or *LI*), and when it occurred.

At 10:06:50 a.m., a feedwater pump current increase (*FWP*) is detected. This event takes place during interval I_{21}, hence $FWP = fwp[21]$. This finding produces an important increase in the probabilities of occurrence of *FWPF* and *LI*, as shown in Table 4. Note that when there was no evidence, the prior probabilities of these events were close to zero for each interval.

At 10:07:45 a.m., a second event is detected corresponding to a steam temperature decrement (*STT*); therefore, $STT = stt[24]$. This new finding gives rise to the new posterior probabilities shown in Table 5. The posterior probabilities for node *FWPF* have now fallen to almost zero for any of the intervals. Additionally, *LI* has become the only initial event that is able to explain the evidence (see Table 5). Note that $P(li[never] \mid fwp[21], stt[24]) \approx 0$.

In this NPEDT, evidence propagation through exact algorithms takes, in general, a few seconds by using Elvira[2] and the factorization described in [11]. (If that factorization is not used, evidence propagation takes almost one minute.) Consequently, this network could be used in a fossil power plant to assist human operators in real-time fault diagnosis and prediction.

[2]Elvira is a software package for the construction and evaluation of BNs and influence diagrams, which is publicly available at `http://www.ia.uned.es/~elvira` .

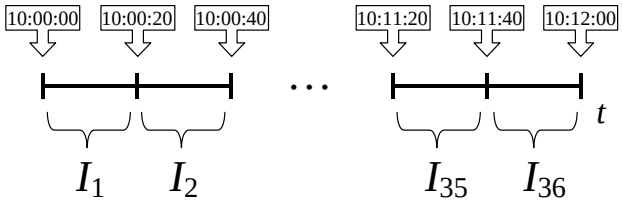

Figure 5: Temporal range for the example.

intervals	$P(fwpf \mid fwp[21])$	$P(li \mid fwp[21])$
...	...	...
I_{15}	0	0.029
I_{16}	0	0.066
I_{17}	0	0.163
I_{18}	0	0.198
I_{19}	0.098	0.110
I_{20}	0.197	0.037
I_{21}	0.098	≈ 0
...	...	...
never	0.605	0.394

Table 4: Posterior probabilities for $FWP = fwp[21]$.

intervals	$P(li \mid fwp[21], stt[24])$
...	...
I_{15}	0.002
I_{16}	0.023
I_{17}	0.182
I_{18}	0.448
I_{19}	0.283
I_{20}	0.058
I_{21}	≈ 0
...	...
never	≈ 0

Table 5: Posterior probabilities for $FWP = fwp[21]$ and $STT = stt[24]$.

4.4 Comparison of the two EBNs for this domain

The graph of the TNBN for the fossil power plant is more simple than that of the NPEDT. This fact is a consequence of the presence of one multiple event, *DRL*. This multiple event is more easily represented in the NPEDT if some additional nodes associated to simple events, *DRLD* and *DRLI*, are introduced.

The definition of the 36 values (or intervals) per variable was immediate and systematic in the NPEDT. However, the simulated cases available and the dynamics of the steam generator had to be studied in order to define the intervals for each temporal node in the TNBN.

Although the NPEDT for this domain needs a high number of probabilities to be assessed, the assumption of time invariance and the use of temporal canonical models greatly facilitates their acquisition. In the case of the TNBN, both the general model of interaction for each family of nodes and the use of relative time complicate the estimation of the parameters. As a consequence, additional configurations of faults need to be run in the simulator in order to obtain the probabilities for the TNBN.

Sec. 4.3.3 demonstrates that the introduction of evidence in the NPEDT is straightforward, as soon as some disturbances are observed. In contrast, due to the use of relative time in the TNBN, several scenarios have to be considered if the parents of the evidence nodes are not observed. As far as the time spent in inference is concerned, while it is negligible for the TNBN, for the NPEDT it takes a few seconds to propagate evidence (see Sec. 4.3.3). However, inference results lead to finer —more informative— probability distributions in the case of the NPEDT: it uses a time unit of 20 seconds, while the mean length of the intervals in the TNBN is 45 seconds.

5 Empirical results

A total of 127 simulation cases were generated for evaluation purposes by means of a simulator of a 350 MW fossil power plant [4]. Each case consists of a list

$$((event_1, t_1), (event_2, t_2), \ldots, (event_{14}, t_{14}))$$

where t_i is the occurrence time for $event_i$. There are 14 possible events, as Fig. 3 shows. If $event_i$ did not occur then $t_i = never$. In general, among the 14 pairs included in each case, some of them correspond to evidence about the state of the steam generator. Table 6 shows an example of a case generated by the simulator, using seconds as time unit.

Due to the intrinsic complexity of this industrial domain, where complicated dynamic processes can happen, it would be rather difficult for any expert in fossil power plants to establish accurate diagnosis or predictions for any partial case presented to him/her. Note that the complete description of a process taking place in the steam generator not only requires specifying which events really occur, but also the time at which they occur. Consequently, in this domain it is impossible for an expert to judge precisely the performance of the

FWVF	*FWPF*	*LI*	*SWVF*	*FWV*	*FWP*	*STV*
never	*never*	0	*never*	*never*	46	18

SWV	*FWF*	*STF*	*SWF*	*DRL*	*DRP*	*STT*
never	142	90	250	127 (decrement)	170	119

Table 6: Example of a case generated by the simulator.

networks evaluated; therefore, an authomatic approach needs to be applied for this purpose.

By using the measure of proximity proposed in Sec. 3.2, we have performed tests for prediction and for diagnosis from the 127 simulation cases available for evaluation. We have used Eq. 5 with the value $c = 360$ for the TNBN and the NPEDT; while $t_f - t_i$ is always 20 seconds in the NPEDT, it is specific of each interval in the case of the TNBN.

5.1 Predictive accuracy

In order to analyze the predictive performance of the networks, we have carried out four different types of tests. In each of them there was only an initial fault event present: *spray water valve failure* ($SWVF$), *feedwater valve failure* ($FWVF$), *feedwater pump failure* ($FWPF$), and *power load increase* (LI), respectively. The states of the rest of the nodes in the networks were unknown. The time at which the corresponding initial fault event occurred defines the beginning of the global time range considered. Among the 127 simulated cases, 64 are associated to the presence of LI, and the rest of the initial fault events are simulated by means of 21 cases each. Tables 7 through 10 contain the means and variances of the measures of proximity obtained separately for both the TNBN and the NPEDT in the predictive tests. The average of the values shown in the last file of each table are: μ(TNBN) = 0.789003, σ^2(TNBN) = 2.603E-4, μ(NPEDT) = 0.945778, and σ^2(NPEDT) = 1.509E-3. Please note that the measure of proximity is, by definition, between 0 and 1.

These results show that the NPEDT predicts more accurately than the TNBN for most of the nodes. In general, the difference between the accuracy of the predictions from the two networks grows as we go down in the graph, i.e., as the distance between the evidence node and the predicted node increases. Both networks predict correctly that some events do not occur. Such events have been omitted in the tables.

5.2 Diagnostic accuracy

The diagnostic performance of the TNBN and the NPEDT was studied on one type of test consisting in occurrence of the bottom fault event, *steam temperature decrement* (STT) (see Fig. 3). The analysis was carried out on 127 simulated cases. Since in a TNBN the introduction of evidence for a node with parents

Node	μ (TNBN)	σ^2 (TNBN)	μ (NPEDT)	σ^2 (NPEDT)
SWV	0.99786	2.185E-6	0.996066	3.456E-6
SWF	0.85793	4.267E-6	0.987375	4.138E-5
STT	0.55219	2.515E-3	0.874228	0.011258
Average	0.80266	8.404E-4	0.952556	3.767E-3

Table 7: Means and variances of *MP* when *SWVF* is present.

Node	μ (TNBN)	σ^2 (TNBN)	μ (NPEDT)	σ^2 (NPEDT)
FWV	0.99844	2.053E-6	0.995963	1.286E-6
FWF	0.88154	8.99E-5	0.957003	2.39E-4
SWF	0.71559	1.069E-6	0.818828	4.704E-4
DRL	0.85165	1.035E-3	0.914457	0.003583
DRP	0.93127	6.317E-6	0.895225	0.006293
STT	0.14576	5.724E-8	0.832621	0.001401
Average	0.754041	1.89E-4	0.902349	1.998E-3

Table 8: Means and variances of *MP* when *FWVF* is present.

Node	μ (TNBN)	σ^2 (TNBN)	μ (NPEDT)	σ^2 (NPEDT)
FWP	0.87001	3.522E-8	0.996962	7.425E-7
FWF	0.90496	1.478E-5	0.989113	2.018E-5
SWF	0.89533	1.537E-5	0.972282	1.598E-4
DRL	0.88665	5.776E-8	0.976043	9.658E-5
DRP	0.93262	9.382E-7	0.975946	6.43E-5
STT	0.14463	4.961E-8	0.954822	4.348E-4
Average	0.772366	5.205E-6	0.977528	1.294E-4

Table 9: Means and variances of *MP* when *FWPF* is present.

Node	μ (TNBN)	σ^2 (TNBN)	μ (NPEDT)	σ^2 (NPEDT)
FWP	0.93043	9.377E-6	0.988255	2.594E-5
STV	0.99858	1.455E-6	0.997886	8.289E-7
FWF	0.83527	1.682E-5	0.97729	2.795E-4
STF	0.99694	8.829E-6	0.992117	1.35E-5
SWF	0.69901	2.722E-6	0.967232	6.314E-4
DRL	0.62306	5.715E-8	0.71745	2.147E-5
DRP	0.99204	1.433E-5	0.978515	1.271E-4
STT	0.540239	1.094E-7	0.986699	5.011E-5
Average	0.826946	6.712E-6	0.95068	1.437E-4

Table 10: Means and variances of *MP* when *LI* is present.

Node	μ (TNBN)	σ^2 (TNBN)	μ (NPEDT)	σ^2 (NPEDT)
SWVF	0.701059	0.041	0.802444	0.118323
FWVF	0.449978	3.147E-3	0.742676	0.071108
FWPF	0.450636	3.064E-3	0.754429	0.062699
LI	0.655584	0.114252	0.995538	3.649E-5
SWV	0.762227	0.01292	0.801968	0.118188
FWV	0.446817	3.518E-3	0.742875	0.070954
FWP	0.428903	0.044449	0.688344	0.041785
STV	0.6529	0.115369	0.995781	1.091E-5
FWF	0.488658	0.04751	0.5607	0.075691
STF	0.630933	0.100964	0.999926	2.671E-9
SWF	0.609587	0.025558	0.999525	2.228E-7
DRL	0.562921	0.049913	0.510202	0.109071
DRP	0.809094	0.117564	0.662385	0.164828
Average	0.588448	0.052248	0.788984	0.064053

Table 11: Means and variances of *MP* when *STT* and one of its parents are present.

requires knowing which of them is causing the appearance of the child event, in this type of test it was necessary to consider information from two nodes: *STT* and its causing parent. Table 11 includes the means and variances of the measures of proximity obtained in this test. Again, the NPEDT performs better than the TNBN for most of the nodes: 0.789 vs. 0.588 .

Although in general the measures of proximity for diagnosis are lower than those for prediction, that does not mean that our EBNs perform in diagnosis worse than in prediction. There is another reason that explains this result: If we had

- two different probability density functions, $f_V(t)$ and $f_W(t)$, the former more spread out than the latter, and
- two infinite sets of cases, C_V and C_W, following distributions $f_V(t)$ and $f_W(t)$, respectively,

then, from Eq. 4, *MP* would be lower on average for V than for W, since $t - \hat{t}_V$ is on average greater than $t - \hat{t}_W$. Therefore, even if a BN yielded satisfactory inference results both for variable V and variable W, Eq. 4 would in general produce different average *MP*s for V and W. This is taking place in our tests. For example, in the NPEDT we calculated the mean number of states per node whose posterior probability was greater than 0.001 . While in the prediction tests this number was approximately 5, in diagnosis it rose to nearly 9. Anyhow, the measure of proximity defined in Eq. 4 allows us to carry out a comparative evaluation of the TNBN and the NPEDT.

6 Discussion and conclusions

The following issues make EBNs more appropriate than DBNs to model the domain considered in this work:

- The propagation of faults in the steam generator of a fossil power plant is not a Markovian process: A causal mechanism between two faults has in general several possible associated delays, each having a different probability.

- Irreversible events, like the faults that may take place in a steam generator, are more easily represented in EBNs than in DBNs. The representation of irreversible processes in DBNs requires defining memory nodes, which add complexity to the model.

The main reason why the NPEDT for the steam generator performs better than the TNBN seems to be the way of defining the intervals. There are two aspects to be considered in this regard: the **number** of intervals defined for each node and their **duration**.

A major advantage of an NPEDT is that it allows us to make use of different temporal noisy gates that facilitate knowledge acquisition and inference. This is the reason why an NPEDT allows a finer granularity (a greater **number** of intervals) than a TNBN, and hence an improvement in the temporal precision of the inference results. For example, in our domain the NPEDT uses a time unit of 20 seconds, while the mean length of the intervals in the TNBN is 45 seconds.

From the discussion in last paragraph, the use of a greater number of intervals or values for nodes in NPEDTs with respect to nodes in TNBNs, results in inference processes for TNBNs faster than for NPEDTs. In our case study, both networks led to short computation times, but this difference might become relevant in the case of more complex networks.

Once the number of relative intervals has been set for a temporal node in a TNBN, the **duration** of each of them needs to be defined. In general, different durations are allowed in the same node or variable. Establishing the right durations for the intervals of each temporal node is not straightforward. The dynamic causal mechanisms taking place in the domain need to be thoroughly studied in order to guarantee that the exclusivity of the values associated to each temporal node is not violated. On the contrary, NPEDTs are not subject to the previous disadvantage, since absolute time is used for all of the variables. This is another possible factor explaining the results obtained in Sec. 5.

In this work, we have carried out a comparative evaluation of two different types of BNs for temporal reasoning. An option for obtaining an absolute evaluation of the two networks separately would be to define a metric between the two following probability distributions: that produced for a node by each network in a certain test, and that obtained for the same node from the available cases. For this ideal method to be applicable, a high number of cases is necessary. This number is even higher in our case, due to the unusual number of values present in each variable.

The main results obtained from the work described in this paper are the following:

- After reviewing the main methodologies for evaluation of BNs, we have

designed **a new method for carrying out a comparative evaluation of EBNs**. The new evaluation method is based on a proximity measure between the posterior probabilities obtained from the networks and each of the cases available for evaluation, rather than on an error measure. Since the events represented in EBNs form dynamic processes, the proximity measure takes into account time as an important variable.

- **We have applied the previous evaluation method** to the TNBN and the NPEDT generated from a case study: the faults that may occur in the steam generator of a fossil power plant. To this end, we have performed different tests in order to compare the predictive as well as the diagnostic accuracy of the TNBN and the NPEDT. The results show that, in general, the NPEDT yields better predictions and diagnoses than the TNBN. There are two main reasons for that: Firstly, the use of temporal canonical models in the NPEDT allows for a finer granularity than in the case of the TNBN and, secondly, the definition of the intervals in a TNBN is not so systematic as in an NPEDT, and depends strongly on the domain.

Acknowledgements

This research was supported by the Spanish CICYT, under grant TIC2001-2973-C05-04. The first author was supported by a grant from the Mexican *Secretaría de Relaciones Exteriores*.

References

[1] S. Andreassen, F. V. Jensen, and K. G. Olesen. Medical expert systems based on causal probabilistic networks. *International Journal of Biomedical Computing*, 28:1–30, 1991.

[2] G. Arroyo-Figueroa. *Razonamiento Probabilístico con Nodos Temporales y su Aplicación al Diagnóstico y Predicción de Eventos*. PhD thesis, División de Ingeniería y Ciencias del ITESM, Cuernavaca, Mexico, 1999. In Spanish.

[3] G. Arroyo-Figueroa and L. E. Sucar. A temporal Bayesian network for diagnosis and prediction. In *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence (UAI'99)*, pages 13–20, Stockholm, Sweden, 1999. Morgan Kaufmann, San Francisco, CA.

[4] G. Arroyo-Figueroa, L. E. Sucar, and A. Villavicencio. Probabilistic temporal reasoning and its application to fossil power plant operation. *Expert Systems with Applications*, 15:317–324, 1998.

[5] J. Binder, K. Murphy, and S. Russell. Space-efficient inference in dynamic probabilistic networks. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI–97)*, pages 1292–1296, Nagoya, Japan, 1997.

[6] C. Boutilier, N. Friedman, M. Goldszmidt, and D. Koller. Context–specific independence in Bayesian networks. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence (UAI'96)*, pages 115–123, Portland, OR, 1996. Morgan Kaufmann, San Francisco, CA.

[7] P. Dagum and A. Galper. Forecasting sleep apnea with dynamic network models. In *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence (UAI'93)*, pages 64–71, Washington D.C., 1993. Morgan Kaufmann, San Francisco, CA.

[8] P. Dagum, A. Galper, and E. Horvitz. Dynamic network models for forecasting. In *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence (UAI'92)*, pages 41–48, Stanford University, CA, 1992. Morgan Kaufmann, San Francisco, CA.

[9] T. Dean and K. Kanazawa. A model for reasoning about persistence and causation. *Computational Intelligence*, 5:142–150, 1989.

[10] F. J. Díez and M. Druzdzel. Canonical probabilistic models for knowledge engineering. Technical Report, Decision Systems Laboratory, University of Pittsburgh, PA, 2003. In preparation.

[11] F. J. Díez and S. F. Galán. Efficient computation for the noisy MAX. *International Journal of Intelligent Systems*, 18:165–177, 2003.

[12] F. J. Díez, J. Mira, E. Iturralde, and S. Zubillaga. DIAVAL, a Bayesian expert system for echocardiography. *Artificial Intelligence in Medicine*, 10:59–73, 1997.

[13] S. F. Galán and F. J. Díez. Networks of probabilistic events in discrete time. *International Journal of Approximate Reasoning*, 30:181–202, 2002.

[14] S. F. Galán, F. J. Díez, F. Aguado, and J. Mira. Nasonet, modeling the spread of nasopharyngeal cancer with networks of probabilistic events in discrete time. *Artificial Intelligence in Medicine*, 25:247–264, 2002.

[15] M. E. Hernando, E. J. Gómez, R. Corcoy, and F. del Pozo. Evaluation of DIABNET, a decision support system for therapy planning in gestational diabetes. *Computer Methods and Programs in Biomedicine*, 62:235–248, 2000.

[16] C. E. Kahn Jr., L. M. Roberts, K. A. Shaffer, and P. Haddawy. Construction of a Bayesian network for mammographic diagnosis of breast cancer. *Computers in Biology and Medicine*, 27:19–29, 1997.

[17] K. Kanazawa. *Reasoning about Time and Probability*. PhD thesis, Dept. of Computer Science, Brown University, RI, 1992.

[18] U. Kjærulff. A computational scheme for reasoning in dynamic probabilistic networks. In *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence (UAI'92)*, pages 121–129, Stanford University, CA, 1992. Morgan Kaufmann, San Francisco, CA.

[19] A. E. Nicholson and J. M. Brady. Sensor validation using dynamic belief networks. In *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence (UAI'92)*, pages 207–214, Stanford University, CA, 1992. Morgan Kaufmann, San Francisco, CA.

[20] A. Oniśko. Evaluation of the HEPAR II system for diagnosis of liver disorders. In *Working notes of the AIME'01 Workshop on Bayesian Models in Medicine*, pages 59–64, Cascais, Portugal, 2001.

[21] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, CA, 1988. Revised second printing, 1991.

[22] K. W. Przytula, D. Dash, and D. Thompson. Evaluation of Bayesian networks used for diagnostics. In *Proceedings of the 24th Annual IEEE Aerospace Conference*, Big Sky, MT, 2003.

[23] L. E. Sucar. *Probabilistic Reasoning in Knowledge–based Vision Systems*. PhD thesis, Imperial College, London, UK, 1992.