MODELLING NETWORKS OF DISCRETE AND CONTINUOUS VARIABLES WITH AN APPLICATION TO DAMAGE ASSESSMENT

Enrique Castillo*, **José Manuel Gutiérrez*** and **Ali S. Hadi****

* *Department of Applied Mathematics and Computational Sciences, University of Cantabria, SPAIN*

** *Department of Statistics, Cornell University, USA*

ABSTRACT

In this paper we show how discrete and continuous variables can be combined using parametric conditional families of distributions and how the likelihood weighting method can be used for propagating uncertainty through the network in an efficient manner. To illustrate the method we use, as an example, the damage assessment of reinforced concrete structures of buildings and we formalize the steps to be followed when modelling probabilistic networks. We start with one set of conditional probabilities. Then, we examine this set for uniqueness, consistency and parsimony. We also show that cycles can be removed because they lead to redundant probability information. This redundancy may cause inconsistency, hence the probabilities must be checked for consistency. This examination may require a reduction to an equivalent set in *standard canonical* form from which one can always construct a Bayesian network, which is the most convenient model. Finally, a sensitivity analysis is performed showing that the model is robust.

Key Words: Bayesian networks, Compatibility, Existence, Simulation, Uniqueness.

1 Introduction

In recent years much attention has been focussed on the use of probability models in expert systems. Today, probability models, especially those associated with Bayesian networks, are successfully replacing other uncertainty measures. Bayesian networks are effective and efficient instruments for dealing with uncertainties in expert systems; see, for example, Pearl (1986a), Lauritzen and Spiegelhalter (1988), Castillo and Alvarez (1991) and Díez and Mira (1994). Existing methods for propagating uncertainty deal with discrete or special cases of continuous random variables (see Spiegelhalter (1993) and Norman and Tritchler (1992)), but no method exists for propagating uncertainty in general mixed networks (networks with discrete and continuous variables). In this paper we show how continuous variables, belonging to parametric families, can be combined with discrete variables and how uncertainty can be propagated in the resulting networks by the likelihood weighting method.

In addition, when modelling probabilistic networks, one of the key problems is the specification of the joint probability distribution of the nodes. When the number of nodes is large, direct specification of the joint probability distribution is practically impossible. It is possible, however, to specify the joint probability distribution indirectly by specifying a set of conditional distributions. For this set to produce a bona fide and unique joint probability distribution, it must satisfy certain compatibility and uniqueness conditions. Gelman and Speed (1993), Castillo, Gutiérrez and Hadi (1993) and Arnold, Castillo, and Sarabia (1995) consider this problem and give simple conditions for the given set of conditional to be compatible and lead to a unique joint probability distribution. These conditions have many practical implications in modelling probabilistic networks.

In this paper, we also formalize and discuss the steps to build a Bayesian network. We start in Section 2 by a formulation of the model. As an illustrative practical example, we use the damage assessment of reinforced concrete structures of buildings. In Section 3, methods by which the network and its associated set of conditional probabilities can be checked for consistency, compatibility, and uniqueness are summarized. In Sections 4, 5 and 6 the network is set up for propagation of uncertainties using the likelihood weighting method and certain questions regarding the assessment of the damage of reinforced concrete structures of buildings are answered. A sensitivity analysis is performed in Section 7 to analyze the robustness of the selected model.

2 Formulation of the Model

In the example we use in this paper, the objective is to assess the damage of reinforced concrete beams of buildings. The example is taken from Liu and Li (1994), but slightly modified for illustrative purposes. The first stage of model formulation involves two steps: variable selection and identification of dependencies.

2.1 Variable Selection

Model formulation process usually starts with the selection or specification of a set of variables of interest. This specification is dictated by the subject matter specialists. In our example, the goal variable (the damage of a reinforced concrete beam) is denoted by X_1. A civil engineer initially identifies 16 variables $(X_9, X_{10}, \ldots, X_{24})$ as the main variables influencing the damage of reinforced concrete structures. In addition, the engineer identifies seven intermediate conceptual variables $(X_2, X_3, \ldots, X_8)$ which define some partial states of the structure and are known functions of some of the above variables (see Section 4). Table 1 shows the list of variables and their physical meanings. The table also shows whether each variable is continuous or discrete and the possible values that each variable can take. The variables are measured using a scale that is directly related to the goal variable, that is, the higher the value of the variable the more the possibility for damage. To generalize, let the set of variables be denoted by $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$. In this example, $n = 24$.

	Variable	Type	Possible Values	Description
Goal	X_1	discrete	0,1,2,3,4	Damage assessment
Intermediate	X_2	mixed	(0,1)	Cracking state
	X_3	mixed	(0,1)	Cracking state in shear domain
	X_4	mixed	(0,1)	Steel corrosion
	X_5	mixed	(0,1)	Cracking state in flexure domain
	X_6	mixed	(0,1)	Shrinkage cracking
	X_7	mixed	(0,1)	Worst cracking state in flexure domain
	X_8	mixed	(0,1)	Corrosion state
Main	X_9	continuous	(0,1)	Weakness of the beam
	X_{10}	continuous	(0,1)	Deflection of the beam
	X_{11}	continuous	(0,1)	Position of the worst shear crack
	X_{12}	continuous	(0,1)	Breadth of the worst shear crack
	X_{13}	continuous	(0,1)	Position of the worst flexure crack
	X_{14}	continuous	(0,1)	Breadth of the worst flexure crack
	X_{15}	continuous	(0,1)	Length of the worst flexure cracks
	X_{16}	continuous	(0,1)	Cover
	X_{17}	continuous	(0,1)	Structure age
	X_{18}	continuous	(0,1)	Humidity
	X_{19}	discrete	0,1,2	PH level in the air
	X_{20}	discrete	0,1,2	Chlorine content level in the air
	X_{21}	discrete	0,1,2,3	Number of shear cracks
	X_{22}	discrete	0,1,2,3	Number of flexure cracks
	X_{23}	discrete	0,1,2,3	Shrinkage level
	X_{24}	discrete	0,1,2,3	Corrosion level

Table 1: Definitions of the variables related to damage assessment of reinforced concrete structures.

X_i	$\mathcal{N}(X_i)$	$Par(X_i)$
X_1	$\{X_9, X_{10}, X_2\}$	$\{X_9, X_{10}, X_2\}$
X_2	$\{X_3, X_6, X_5, X_4, X_1\}$	$\{X_3, X_6, X_5, X_4\}$
X_3	$\{X_{11}, X_{12}, X_{21}, X_8, X_2\}$	$\{X_{11}, X_{12}, X_{21}, X_8\}$
X_4	$\{X_{24}, X_8, X_5, X_2, X_{13}\}$	$\{X_{24}, X_8, X_5\}$
X_5	$\{X_{13}, X_{22}, X_7, X_2, X_4\}$	$\{X_{13}, X_{22}, X_7\}$
X_6	$\{X_{23}, X_8, X_2\}$	$\{X_{23}, X_8\}$
X_7	$\{X_{14}, X_{15}, X_{16}, X_{17}, X_8, X_5\}$	$\{X_{14}, X_{15}, X_{16}, X_{17}, X_8\}$
X_8	$\{X_{18}, X_{19}, X_{20}, X_7, X_4, X_6, X_3\}$	$\{X_{18}, X_{19}, X_{20}\}$
X_9	$\{X_1\}$	ϕ
X_{10}	$\{X_1\}$	ϕ
X_{11}	$\{X_3\}$	ϕ
X_{12}	$\{X_3\}$	ϕ
X_{13}	$\{X_5, X_4\}$	$\{X_4\}$
X_{14}	$\{X_7\}$	ϕ
X_{15}	$\{X_7\}$	ϕ
X_{16}	$\{X_7\}$	ϕ
X_{17}	$\{X_7\}$	ϕ
X_{18}	$\{X_8\}$	ϕ
X_{19}	$\{X_8\}$	ϕ
X_{20}	$\{X_8\}$	ϕ
X_{21}	$\{X_3\}$	ϕ
X_{22}	$\{X_5\}$	ϕ
X_{23}	$\{X_6\}$	ϕ
X_{24}	$\{X_4\}$	ϕ

Table 2: Variables and their corresponding neighbors, $\mathcal{N}(X_i)$, and parents, $Par(X_i)$ as initially seen by the engineer.

2.2 Identification of Dependencies

The next step in model formulation is the identification of the dependency structure among the selected variables. This identification is also given by the subject matter specialists and is usually done by identifying the minimum set of variables, $\mathcal{N}(X_i)$, for each variable X_i such that

$$P(X_i|\mathcal{X} - X_i) = P(X_i|\mathcal{N}(X_i)), \tag{1}$$

that is, given $\mathcal{N}(X_i)$, X_i is conditionally independent of $\mathcal{X} - \mathcal{N}(X_i) - X_i$. The set $\mathcal{N}(X_i)$ is referred to as the neighbors of X_i. The variables and their corresponding neighbors are shown in the first two columns of Table 2. It follows that if $X_j \in \mathcal{N}(X_i)$, then $X_i \in \mathcal{N}(X_j)$.

Additionally, but optionally, the engineer can impose certain cause-effect relationships among the variables, that is, specifying which variables among the set $\mathcal{N}(X_i)$ are direct causes of X_i and which are direct effects of X_i. The set of direct causes of X_i is referred to as the parents of X_i and is denoted by $Par(X_i)$. Similarly, the set of direct effects of X_i is referred to as the Children of X_i and is denoted by $C(X_i)$.

In our example, the engineer imposes the following cause-effect relationships. The goal variable, X_1, depends primarily on three factors, X_9, the weakness of the beam available in the form of a damage factor, X_{10}, the deflection of the beam, and X_2, its cracking state. The cracking state, X_2, in turn is characterized by four variables: X_3, the cracking state in the shear domain; X_6, the evaluation of the shrinkage cracking; X_4, the evaluation of the steel corrosion; and X_5, the cracking state in the flexure domain. Shrinkage cracking, X_6, depends on shrinkage, X_{23}, and the corrosion state, X_8. Steel corrosion, X_4, is defined by X_8, X_{24}, and X_5. The cracking state in the shear domain, X_3, depends on X_{11}, the position of the worst shear crack; X_{12}, the breadth of the worst shear crack, X_{21}, the number of shear cracks, and X_8. The cracking state in the flexure domain, X_5 is determined by X_{13}, the position of the worst flexure crack, the worst cracking state in the flexure domain without considering the position, X_{22}, the number of flexure cracks, and X_7, the worst cracking state in the flexure domain. The variable X_{13} is influenced by X_4. The variable X_7 is a function of X_{14}, the breadth of the worst flexure crack, X_{15}, the length of the worst flexure crack, X_{16}, the cover, X_{17} the structure age, and X_8, the corrosion state. Node X_8 is determined by X_{18}, the humidity, X_{19}, the PH value in the air, and X_{20}, the content of chlorine in the air.

These causal-effect relationships among the variables are depicted in Figure 1. Each node in this diagram represents a variable. The relationships are represented by links (a directed line emanating from one node and pointing to another). For example, there are three arrows emanating from the nodes X_9, X_{10}, and X_2 and pointing to X_1 indicating that X_1 depends on the three variables. Thus, the nodes X_2, X_9, and X_{10} are the parents of X_1 and X_1 is the child of each of X_2, X_9, and X_{10}. The sets

$Par(X_i)$ are shown in the third column of Table 2. The sets of the children of X_i is $C(X_i) = \mathcal{N}(X_i) - Par(X_i)$.

3 Diagnosing the Model

Once the set of conditional probabilities as in (1) are given, the task of the statistical expert begins. Before propagation of uncertainty can start, the given set of conditional probabilities have to be checked for consistency, compatibility, and uniqueness, i.e. the statistical expert determines whether the set of conditional probabilities corresponds to a well defined joint probability distribution of the variables in the network. To this purpose we use some results given by Gelman and Speed(1993) and Arnold, Castillo, and Sarabia (1995). We also show that cycles imply redundant information and can always be removed.

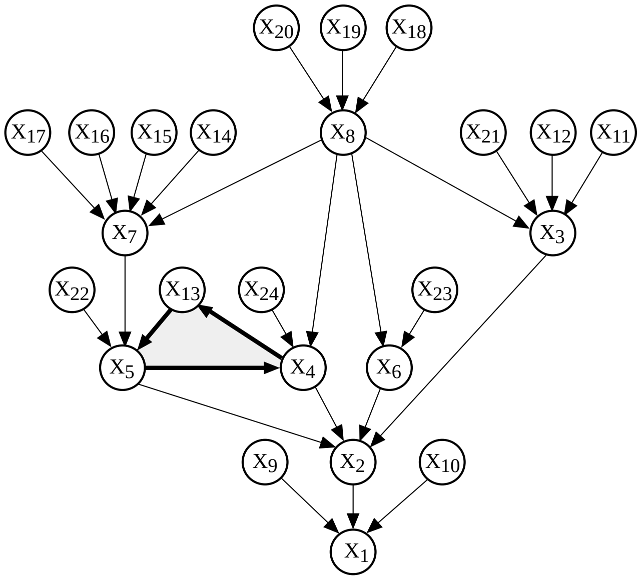

Figure 1: Diagram of the damage assessment of reinforced concrete structure example as initially seen by an engineer. The arrows define local dependencies between variables.

THEOREM 1 Canonical Representation (Gelman and Speed, 1993): *Suppose that*

$$\mathcal{P} \equiv \{P_1(A_1|B_1), \ldots, P_m(A_m|B_m)\}$$

is a given collection of conditional probabilities, where A_i, B_i are subsets of $\mathcal{X}$, such that $A_i \cap B_i = \phi$. Then from the above collection we can obtain an equivalent representation such that all new sets A_i contain a single element of $\mathcal{X}$.

Note that when $B_i = \phi$, the conditional probability $P_i(A_i|B_i)$ is simply $P(A_i)$, the marginal probability of A_i.

The resulting set of conditionals and marginals is known as the *canonical representation* of the probability distribution of $\mathcal{X}$. As a consequence of this theorem, and without loss of generality, we assume in the following that the set $\mathcal{P}$ is given in a canonical form. An examination of the set of probabilities $\mathcal{P}$ in Table 3 shows that the set is already given in a canonical form.

Given a set $\mathcal{P}$ in a canonical form, the set may or may not uniquely define a joint probability distribution of all variables. The conditions under which a given canonical representation uniquely defines a joint probability distribution are given below.

THEOREM 2 Uniqueness (Gelman and Speed, 1993): *Given a collection of conditional distributions in canonical form, and assuming that it is compatible with at least one joint distribution for $\mathcal{X}$, this collection uniquely determines the joint distribution of $\mathcal{X}$ if and only if it, after possible permutation of the variables, contains a nested sequence of probability functions of the form*

$$P_i(X_i|S_i) \;\; \forall X_i \in \mathcal{X} \text{ and } S_i \supset H_i = \{X_{i+1}, X_{i+2}, \ldots, X_n\}. \tag{2}$$

If $S_i = H_i$, for all i, then the consistency is guaranteed, otherwise, the set of conditionals must be checked for consistency. When $S_i = H_i$, for all i, a canonical form is referred to a *standard canonical form* and the term $P_i(X_i|H_i)$ is referred to as a *standard canonical component.*

A practical and important consequence of this theorem is that a minimum set of conditionals is required to have uniqueness, that is, a well defined (unambiguous) joint distribution of $\mathcal{X}$. We therefore, make use of this theorem as follows. If we can find a given order of the nodes, say $X_1, X_2, \ldots, X_n$, and, for each node X_i, we define, with respect to this order, a set of parents of X_i as

$$Par(X_i) = \{X_j \in \mathcal{N}(X_i) : j > i\}, i = 1, 2, \ldots, n, \tag{3}$$

then we have $H_i \supset Par(X_i)$ and, by theorem 2, the sequence $P_i(X_i|Par(X_i))$ is consistent with a joint distribution of X. This can be demonstrated as follows. Using the decomposition axiom of independence, equation (1) can be written as

$$\begin{aligned} P(X_i|Par(X_i), C(X_i), H_i - Par(X_i)) &= P(X_i|Par(X_i), C(X_i), \mathcal{X} - \mathcal{N}(X_i) - X_i) \\ &= P(X_i|Par(X_i), C(X_i)). \end{aligned} \tag{4}$$

Multiplying by $P(C(X_i))$ and integrating with respect to $C(X_i)$ we get

$$P(X_i|H_i) = P(X_i|Par(X_i), H_i - Par(X_i)) = P(X_i|Par(X_i)); \forall i. \tag{5}$$

Equation (5) is an important condition implied by Equation (1) that leads to uniqueness of the joint distribution. Note that there are many joint distributions that are compatible with the set of conditionals $\{P(X_i|Par(X_i)), i = 1, 2, \ldots, n\}$ but at most one satisfies (5). From (5), and by the chain rule, the joint probability distribution of $\mathcal{X}$ can be written as

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^{n} P(X_i|H_i) = \prod_{i=1}^{n} P(X_i|Par(X_i)). \tag{6}$$

Thus, the engineer need to provide a set of conditional probabilities

$$\mathcal{P} = \{P(X_i|Par(X_i)) : i = 1, 2, \dots, n\}.$$

This set is given in Table 3, where the continuous variables are assumed to have a $Beta(a, b)$ distribution with the specified parameters. The reason for this choice is that the beta distribution has finite bounds and also has a variety of shapes depending on the choice of the parameters. The discrete variables are assumed to be Binomial $B(n, p)$. The intermediate variables $X_j; j = 2, \dots, 8$ are assumed to have a $Dirac(h(Par(X_j)))$ function, where $h(x_1, x_2, \dots, x_n)$ is given by

$$h(x_1, x_2, \dots, x_n) = \sum_{j=1}^{n} \frac{x_j/u_j}{n}, \tag{7}$$

where u_j is an upper-bound (e.g., the maximum value) of the random variable X_j.

Note that if no cause-effect relationship are given, an ordering satisfying (3) can always be found because we have no restriction on the choice of parents, that is, any subsets of $\mathcal{N}(x_i)$ can serve as $Par(X_i)$. On the other hand, if some cause-effect relationships are given, the ordering in (3) must satisfy these relationships, that is, a child must receive a lower number than, all of its parents. It follows then that if the given cause-effect relationships contain cycles, there exists no ordering which satisfies (3). Therefore, cycles have to be removed.

We have arranged the variables in Table 3 in the order required to check uniqueness so that a permutation of the variables is not necessary. It can be seen that the condition for consistency, $S_i = H_i$, for all i, is satisfied for all nodes except for X_{13} where X_{13} depends on X_4 which is a variable preceding it in the list. Thus, we have $H_{13} = \{X_{14}, X_{15}, \dots, X_{24}\}$ which is not equal to $S_{13} = X_4 \cup H_{13}$. Therefore, the set of conditional distributions in Table 3 needs to be checked for consistency.

Existence or Compatibility. The representation in (2) ensures only uniqueness but it does not guarantee the existence of a joint probability distribution for the set $\mathcal{X}$. In fact, one can give contradictory conditional probabilities that have no joint probability distribution. Arnold et al. (1995) provide a theorem by which one can determine whether a given set of conditionals define a feasible joint probability distribution for $\mathcal{X}$. They also give an algorithm for checking the compatibility, one step at a time, and for constructing a canonical form of the above type with $S_i = H_i$, $\forall i = 1, 2, \dots, n$. Let H_i be as defined in (2) and $\bar{H}_i = \{X_1, X_2, \dots, X_i\}$.

Node	Probabilities $P(X_i \vert H_i)$	Family
X_1	$P(X_1 \vert X_9, X_{10}, X_2)$	$B(4, 0.3x_9 + 0.1x_{10} + 0.6x_2))$
X_2	$f(X_2 \vert X_3, X_6, X_5, X_4)$	$Dirac(h(x_3, x_6, x_5, x_4))$
X_3	$f(X_3 \vert X_{11}, X_{12}, X_{21}, X_8)$	$Dirac(h(x_{11}, x_{12}, x_{21}, x_8))$
X_4	$f(X_4 \vert X_{24}, X_8, X_5)$	$Dirac(h(x_{24}, x_8, x_5))$
X_5	$f(X_5 \vert X_{13}, X_{22}, X_7)$	$Dirac(h(x_{13}, x_{22}, x_7))$
X_6	$f(X_6 \vert X_{23}, X_8)$	$Dirac(h(x_{23}, x_8))$
X_7	$f(X_7 \vert X_{14}, X_{15}, X_{16}, X_{17}, X_8)$	$Dirac(h(x_{14}, x_{15}, x_{16}, x_{17}, x_8))$
X_8	$f(X_8 \vert X_{18}, X_{19}, X_{20})$	$Dirac(h(x_{18}, x_{19}, x_{20}))$
X_9	$f(X_9)$	$Beta(2, 6)$
X_{10}	$P(X_{10})$	$Beta(1, 1)$
X_{11}	$P(X_{11})$	$Beta(0.9, 0.9)$
X_{12}	$P(X_{12})$	$Beta(1, 4)$
X_{13}	$P(X_{13})$	$Beta(2, 2)$
X_{14}	$P(X_{14})$	$Beta(1, 4)$
X_{15}	$f(X_{15})$	$Beta(1, 4)$
X_{16}	$P(X_{16})$	$Beta(1, 4)$
X_{17}	$f(X_{17})$	$Beta(6, 2)$
X_{18}	$f(X_{18})$	$Beta(6, 2)$
X_{19}	$P(X_{19})$	$B(2, 0.2)$
X_{20}	$P(X_{20})$	$B(2, 0.2)$
X_{21}	$P(X_{21})$	$B(3, 0.2)$
X_{22}	$P(X_{22})$	$B(3, 0.2)$
X_{23}	$P(X_{23})$	$B(3, 0.1)$
X_{24}	$P(X_{24})$	$B(3, 0.1)$

Table 3: Beta and Binomial marginal and conditional probability distributions for variables X_1 and X_9 to X_{24}.

The compatibility and uniqueness results together imply that cycles can be removed. Even more, if they are not removed, that is, if the corresponding conditional probabilities are given, then consistency must first be checked. In addition, removing the cycles allows specifying the conditional probabilities with no restrictions (apart from the probability axioms that each individual distribution must satisfy).

Thus, once the network and the associated set of conditional probabilities are given, the statistical expert can perform the following tasks:

1. Check that there are enough links (in the sense of satisfying the uniqueness theorem, that is, that the minimal required nested sequence (2) is included in the network.).

2. Remove all cycles, if present. Note that given a cycle one can remove it by just removing one of its links or changing the direction of some arrows. When removing a link, however, one must be careful not to remove any of the dependence relationships stated by the human specialists. For example, if the link $X_4 - X_{13}$ is removed, this would imply that $P(X_4|X_5, X_{13}) = P(X_4|X_5)$, that is, X_4 is independent of X_{13}, given X_5, which is in contradiction with the engineer's specification.

3. Remove redundant links (in the sense explained by the compatibility theorem).

4. Propagate uncertainties.

5. Answer queries posed by the engineer regarding the probabilities of the goal variable given the data (the evidence set).

As can be observed, the diagram in Figure 1 contains one cycle, which is indicated as a shadowed region and thick arrows. It involves nodes X_5, X_4, and X_{13}. This implies that in the engineer's mind, the cracking state in the flexure domain influences the steel corrosion, the steel corrosion influences position of the worst flexure crack and the worst flexure crack influences the cracking state in the flexure domain.

However, as it has been indicated above, cycles represent redundant conditional probability information which can lead to incompatibility. Thus, we can remove cycles without affecting the joint probability assignment and avoiding compatibility checks. We have reversed the direction of the link $X_4 \to X_{13}$ thus obtaining another graph without cycles (Figure 2) which allow us to define the joint distribution of all nodes without restrictions in the selection of the conditional probabilities. Note that reversing this link requires changing the conditional probabilities for X_4 and X_{13} in Table 3 from $h(X_{24}, X_8, X_5)$ to $h(X_{24}, X_8, X_5, X_{13})$, and from $P(X_{13}|X_4)$ to $P(X_{13}) = B(2, 2)$, respectively. Thus, we arrive at a set of conditionals in a standard canonical form and the probability assignment does not cause any compatibility problems, i.e. we obtain a Bayesian network model which arises in a natural way.

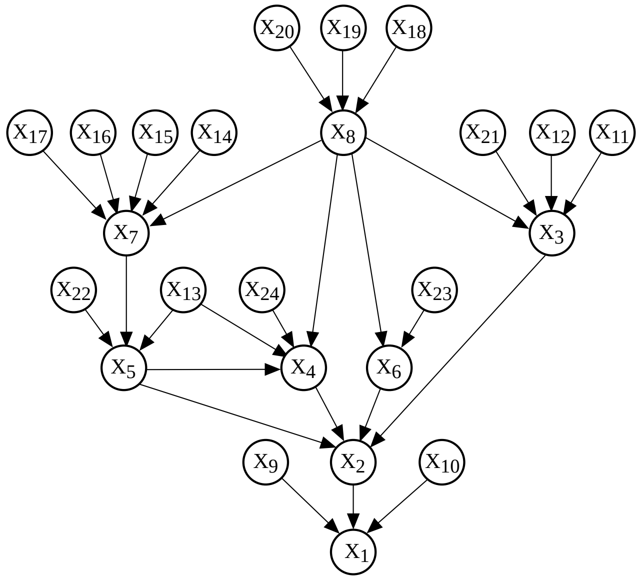

Figure 2: The network in Figure 2 after reversing the link from X_4 to X_{13}. Now it becomes a directed acyclic graph.

4 Specification of Conditional Distributions

To simplify the probability assignment, the engineer assumes that the conditional probabilities belong to some parametric families (e.g., Binomial, Beta, etc.). Table 3 specifies a parametric family for each of the nodes in the network. The variable X_1 can assume only one of five values (states): $0, 1, 2, 3, 4$, with 0 meaning the building is free of damage and 4 meaning the building is seriously damaged. The values in between are intermediate states of damage. All other variables are defined similarly using a scale that is directly related to the goal variable, that is, the higher the value of the variable the more possibility for damage.

All discrete variables are assumed to have a binomial distribution with parameters N and p, with $N+1$ being the number of possible states of each variable. These distributions, however, can be replaced by any other suitable distributions. The parameter $0 \leq p \leq 1$, associated with node X_1, is specified as follows.

$$p(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} \alpha_j \frac{x_j}{u_j}, with \sum_{j=1}^{n} \alpha_j = 1;\ \alpha_j \geq 0, \tag{8}$$

where α_j is a weight associated with X_j. Thus $p(x_1, x_2, \ldots, x_n)$ is a weighted function of $x_1, x_2, \ldots, x_n$ Figure 3 shows the probability density functions (pdf) of some of the Beta functions used in the example. A Beta(0.9,0.9) and a Beta(2,2) have been chosen for the positions of the worst shear and flexure cracks, respectively, to reflect the fact that the largest shear forces and bending moments occur at the end points and the center of the beam, respectively. A Beta(1,4) has been selected for both the breadth of the worst shear and flexure cracks to reproduce the fact that large cracks are less frequent than small cracks. Similarly, a Beta(1,1) (uniform) has been used for the deflection of the beam because all deflections occur in reality, more or less with the same frequency. Finally, a Beta(6,2) is used for the age and the humidity because there are more old structures than young ones and high humidity is more common than low humidity.

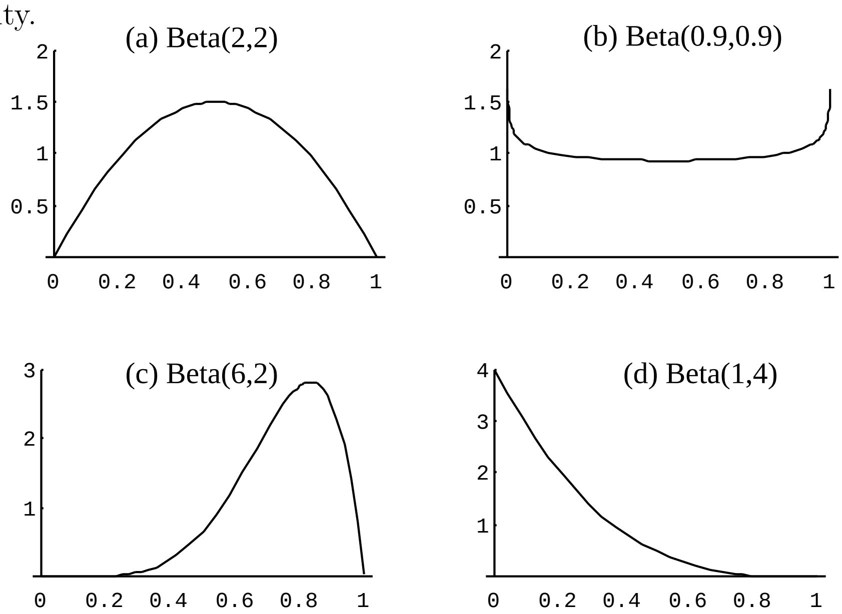

Figure 3: pdfs of several Beta models used in the example.

On the other hand, the beta random variable has been used by its smoothness and flexibility to concentrate a maximum probability on a given zone. The parameter values associated with variables X_{19} to X_{24} have been selected according to the accumulated practical experience.

The mixed graphical association model given by Lauritzen (1992) could also be used here but we consider that the basic assumption of this model stating the conditional distribution of the quantitative variables given the qualitative is multivariate Gaussian is too strong for this case, where the beta family seems more adequate.

5 Propagating Uncertainty

In this paper we deal with both continuous and discrete random variables that are combined in the same network. Also, the network in Figure 2 contains many loops. For example, there is a loop involving the variables X_2, X_6, X_8, X_4, X_2 and another loop involving $X_2, X_3, X_8, X_7, X_5, X_2$. Thus, we need an uncertainty propagation mechanism to allow for this general type of networks.

Exact propagation mechanisms are available for some special topologies of the network when the variables are discrete or belong to simple families, as the normal, (see, for example, Lauritzen and Spiegelhalter, 1988, Pearl, 1991) but no exact efficient and general methods exist when all the variables are continuous or they are combined with discrete variables (for one particular case see Wermouth and Lauritzen, 1990 or Lauritzen (1992)).

The case of continuous variables complicates things because sums need to be replaced by integrals. Thus, several simulation alternative methods have been proposed, e.g., probabilistic logic sampling, Gibbs sampling, likelihood weighting, etc. (see Pearl, 1986b, Henrion, 1988, Shachter and Peot, 1990 or Fung and Chang, 1990).

Because of its computational efficiency, the likelihood weighting method has been selected in this paper as the propagation mechanism. Note that the probabilistic logic sampling and the Gibbs sampling are much less efficient because the first leads to a very high rejection rate and the second needs some iterations before convergence, and it requires the specification of full conditionals which is not possible in this case.

Suppose that the joint probability, $P_{\mathcal{X}}(x)$, can be represented in the form

$$P_{\mathcal{X}}(x) = Cg(x)h(x), \tag{9}$$

where $C > 1$ and. Let

$$\begin{aligned} g(x) &= \prod_{j \in E} P_j(x_{j0}|H_j^*), \\ h(x) &= \prod_{i \in V-E} P_i(x_i|H_i^*), \end{aligned} \tag{10}$$

where E is the evidence set and x_{j0} is the known evidence associated with the variable x_j and H_i^* is the set H_i with all the evidences instantiated. Note that this is equivalent to assuming $P_j(x_{j0}|H_j^*) = 1$. Clearly, $0 < g(x) < 1$, and $h(x)$ is a probability density function (pdf).

In the light of the above, we may use the following algorithm to simulate the pseudo-random variables corresponding to the density $P_{\mathcal{X}}(x)$:

1. Using the corresponding conditional distribution, that is, $P_i(x_i|H_i^*)$ for x_i, generate the random variables not in the evidence set E, one by one and in the reversed order.

2. Calculate the associated sample score $\prod_{j \in E} P_j(x_{j0}|H_j^*)$ and accumulate it over the samples.

3. Repeat the above two steps for a specified number of replications.

4. Calculate the marginals by adding the scores associated with the feasible values and normalizing by the sum of all scores.

It is worthwhile mentioning that this likelihood weighting procedure allows not only the univariate but the multivariate marginals to be obtained. In fact, in the above algorithm, in addition to the scores, we store the frequencies of all feasible values of the discrete variables and all simulated values of the continuous variables. Both can be used to plot estimates of the resulting marginal distributions given the evidence.

6 Answering Queries

To illustrate the uncertainty propagation and to answer certain queries prompted by the engineer, we assume that the engineer examines a given concrete beam and obtain the values $x_9, x_{10}, \ldots, x_{24}$ corresponding to the observable variables $X_9, X_{10}, \ldots, X_{24}$. Note that these values can be measured sequentially. In this case, the inference can also be made sequentially. As illustrative examples, suppose we wish to assess the damage (the goal variable, X_1) in each of the following hypothetical situations:

Q1: Before Observing Evidence. We are given only the conditional and marginal probabilities in Table 3 without any evidence (i.e., without knowledge of the values $x_9, x_{10}, \ldots, x_{24}$).

A1: Tables 4 to 6 show the probabilities of the damage X_1 of a given beam for various types of evidence ranging from no knowledge at all to the knowledge of all the observed values $x_9, x_{10}, \ldots, x_{24}$. Thus, the answer to *Q1* is given in the row corresponding to the cumulative evidence "None". Thus, for example, the probability that a randomly selected building has no damage ($X_1 = 0$) is 0.2313 and the probability that the building is seriously damaged ($X_1 = 4$) is 0.0129. These probabilities can be interpreted as 23% of the buildings in the area are save and 1.29% are seriously, damaged. Other values in Table 4 are explained and interpreted below.

Q2: Evidence of High Damage. Now, suppose that we have the data for all the observable variables as given in Table 4, but the data are measured sequentially in the order given in the table.

A2: The answer is given in Table 4, where the probabilities in the ith row is computed using $x_9, x_{10}, \ldots, x_i$, that is, they are based on accumulated evidence. Except for the key variables X_9 and X_{10}, the values of all other variables attain high values resulting in high probabilities of damage. For example, as can be seen in the last row of the table, when all the evidences are considered, $P(X_1 = 4) \simeq 1$ an indication that the building is seriously damaged.

Q3: Evidence of Low Damage. The same as Q2 but with the evidences given in Table 5.

Known variables	Damage of the beam				
	0	1	2	3	4
None	0.2313	0.3977	0.2665	0.0916	0.0129
$X_9 = 1.00$	0.0468	0.2153	0.3663	0.2806	0.0910
$X_{10} = 1.00$	0.0321	0.1630	0.3501	0.3325	0.1223
$X_{11} = 1.00$	0.0246	0.1540	0.3352	0.3466	0.1396
$X_{12} = 1.00$	0.0178	0.1292	0.3149	0.3673	0.1708
$X_{13} = 1.00$	0.0107	0.0858	0.2769	0.4012	0.2254
$X_{14} = 1.00$	0.0081	0.0779	0.2697	0.4152	0.2291
$X_{15} = 1.00$	0.0085	0.0681	0.2532	0.4134	0.2568
$X_{16} = 1.00$	0.0070	0.0673	0.2473	0.4152	0.2632
$X_{17} = 1.00$	0.0070	0.0639	0.2487	0.4165	0.2639
$X_{18} = 1.00$	0.0064	0.0607	0.2312	0.4078	0.2939
$X_{19} = 2.00$	0.0026	0.0355	0.1812	0.4231	0.3576
$X_{20} = 2.00$	0.0007	0.0198	0.1380	0.3970	0.4445
$X_{21} = 3.00$	0.0006	0.0122	0.0976	0.3668	0.5228
$X_{22} = 3.00$	0.0001	0.0030	0.0486	0.2953	0.6530
$X_{23} = 3.00$	0.0000	0.0002	0.0056	0.1203	0.8739
$X_{24} = 3.00$	0.0000	0.0000	0.0000	0.0003	0.9997

Table 4: The probability distribution of the damage, X_1, given the accumulated evidence of $x_9, x_{10}, \ldots, x_{24}$ as indicated in the table. The results are based on 10,000 replications.

A3: In this case all the variables attain low values indicating that the building is in good condition. When considering all evidence, for example, the probability of no damage is as high as 1.

Q4: **Observing Partial Evidence.** Finally, suppose that the data we have available is only for a subset of the observable variables as given in Table 6.

A4: The probabilities are reported in Table 6 and can be interpreted in a similar way.

It can be seen from the above examples that any query posed by the engineer can be answered simply by propagating uncertainties using the evidence given. Note also that it is possible for the inference to be made sequentially. An advantage of the sequential inference is that we may be able to make a decision concerning the state of damage of a given building immediately after observing only a subset of the variables. Thus, for example, once a very high value of X_9 or X_{10} is observed, the inspection can stop at this point and the building is declared to be seriously damaged.

Known variables	Damage of the beam				
	0	1	2	3	4
None	0.2313	0.3977	0.2665	0.0916	0.0129
$X_9 = 0.00$	0.3538	0.4058	0.1915	0.0458	0.0031
$X_{10} = 0.00$	0.4383	0.3950	0.1410	0.0243	0.0014
$X_{11} = 0.00$	0.4891	0.3727	0.1206	0.0168	0.0008
$X_{12} = 0.00$	0.5020	0.3754	0.1051	0.0164	0.0011
$X_{13} = 0.00$	0.6373	0.2931	0.0627	0.0068	0.0001
$X_{14} = 0.00$	0.6362	0.3004	0.0588	0.0045	0.0001
$X_{15} = 0.00$	0.6408	0.2943	0.0598	0.0046	0.0005
$X_{16} = 0.00$	0.6466	0.2928	0.0562	0.0044	0.0000
$X_{17} = 0.00$	0.6843	0.2648	0.0467	0.0042	0.0000
$X_{18} = 0.00$	0.8164	0.1635	0.0189	0.0012	0.0000
$X_{19} = 0.00$	0.8498	0.1376	0.0121	0.0005	0.0000
$X_{20} = 0.00$	0.8791	0.1117	0.0087	0.0005	0.0000
$X_{21} = 0.00$	0.9156	0.0787	0.0056	0.0001	0.0000
$X_{22} = 0.00$	0.9544	0.0434	0.0022	0.0000	0.0000
$X_{23} = 0.00$	0.9851	0.0149	0.0000	0.0000	0.0000
$X_{24} = 0.00$	1.0000	0.0000	0.0000	0.0000	0.0000

Table 5: The probability distribution of the damage, X_1, given the accumulated evidence of $x_9, x_{10}, \ldots, x_{24}$ as indicated in the table. The results are based on 10,000 replications.

Known variables	Damage of the beam				
	0	1	2	3	4
None	0.2313	0.3977	0.2665	0.0916	0.0129
$X_9 = 0.70$	0.1006	0.2997	0.3560	0.1977	0.0460
$X_{10} = 0.80$	0.0768	0.2721	0.3684	0.2317	0.0510
$X_{17} = 0.80$	0.0797	0.2674	0.3670	0.2308	0.0551
$X_{13} = 0.90$	0.0569	0.2307	0.3755	0.2653	0.0716
$X_{20} = 1.00$	0.0513	0.2208	0.3677	0.2813	0.0789
$X_{21} = 2.00$	0.0386	0.1963	0.3685	0.3059	0.0907

Table 6: The probability distribution of the damage, X_1, given the accumulated evidence of the variables as indicated in the table. The results are based on 10,000 replications.

Node	Parameter	Known Variables							
		No Evidence		Evidence 1		Evidence 2		Evidence 3	
		Mean	Var	Mean	Var	Mean	Var	Mean	Var
	None	**1.26**	**0.91**	**0.94**	**0.75**	**0.76**	**0.63**	**0.00**	**0.00**
X_9	$Beta(\mathbf{p}, q)$	1.22	0.91	0.94	0.75	0.74	0.62	0.00	0.00
X_{10}	$Beta(\mathbf{p}, q)$	1.21	0.88	0.94	0.75	0.73	0.61	0.00	0.00
X_{11}	$Beta(\mathbf{p}, q)$	1.24	0.91	0.95	0.75	0.74	0.63	0.00	0.00
X_{12}	$Beta(\mathbf{p}, q)$	1.24	0.91	0.95	0.75	0.73	0.61	0.00	0.00
X_{13}	$Beta(\mathbf{p}, q)$	1.23	0.91	0.95	0.77	0.75	0.64	0.00	0.00
X_{14}	$Beta(\mathbf{p}, q)$	1.22	0.90	0.94	0.75	0.74	0.62	0.00	0.00
X_{15}	$Beta(\mathbf{p}, q)$	1.25	0.91	0.94	0.73	0.74	0.63	0.00	0.00
X_{16}	$Beta(\mathbf{p}, q)$	1.25	0.92	0.93	0.75	0.74	0.62	0.00	0.00
X_{17}	$Beta(\mathbf{p}, q)$	1.25	0.92	0.94	0.75	0.74	0.63	0.00	0.00
X_{18}	$Beta(\mathbf{p}, q)$	1.25	0.92	0.93	0.76	0.74	0.63	0.00	0.00
X_9	$Beta(p, \mathbf{q})$	1.25	0.90	0.94	0.77	0.75	0.62	0.00	0.00
X_{10}	$Beta(p, \mathbf{q})$	1.26	0.92	0.97	0.77	0.75	0.62	0.00	0.00
X_{11}	$Beta(p, \mathbf{q})$	1.26	0.94	0.95	0.75	0.76	0.65	0.00	0.00
X_{12}	$Beta(p, \mathbf{q})$	1.22	0.91	0.95	0.77	0.74	0.63	0.00	0.00
X_{13}	$Beta(p, \mathbf{q})$	1.25	0.91	0.96	0.77	0.76	0.63	0.00	0.00
X_{14}	$Beta(p, \mathbf{q})$	1.27	0.93	0.95	0.77	0.75	0.62	0.00	0.00
X_{15}	$Beta(p, \mathbf{q})$	1.23	0.90	0.94	0.76	0.75	0.62	0.00	0.00
X_{16}	$Beta(p, \mathbf{q})$	1.23	0.88	0.95	0.76	0.74	0.63	0.00	0.00
X_{17}	$Beta(p, \mathbf{q})$	1.24	0.91	0.96	0.77	0.75	0.63	0.00	0.00
X_{18}	$Beta(p, \mathbf{q})$	1.27	0.93	0.94	0.75	0.75	0.65	0.00	0.00
X_{19}	$B(n, \mathbf{p})$	1.27	0.92	0.97	0.78	0.78	0.65	0.00	0.00
X_{20}	$B(n, \mathbf{p})$	1.27	0.95	0.97	0.76	0.77	0.65	0.00	0.00
X_{21}	$B(n, \mathbf{p})$	1.27	0.93	0.95	0.76	0.76	0.63	0.00	0.00
X_{22}	$B(n, \mathbf{p})$	1.28	0.93	0.96	0.77	0.77	0.63	0.00	0.00
X_{23}	$B(n, \mathbf{p})$	1.27	0.90	0.98	0.78	0.79	0.67	0.00	0.00
X_{24}	$B(n, \mathbf{p})$	1.26	0.93	0.96	0.75	0.75	0.62	0.00	0.00

Table 7: Means and variances of node X_1 for 4 different evidence situations for different parameter modifications. The modified parameters are written in boldface. Binomial p parameters are increased by 0.1 and Beta parameters are decreased by 0.2 with respect to the model parameters. Evidence 1 corresponds to $X_9 = 0$. Evidence 2 corresponds to $X_9 = X_{10} = 0$. Evidence 3 corresponds to $X_9 = \ldots = X_{24} = 0$.

7 Sensitivity Analysis

In this section we discuss the sensitivity of the model to the specified parameter values. We repeated the simulations by changing the parameters one parameter at a time. The Binomial p parameters are increased by 0.1 and the Beta parameters are decreased by -0.2. Note that changing the parameters implies changing the shape of the distribution. This is specially so for the beta distribution. The resulting means and variances of the X_1 variable are shown in Table 7, where one can see that although the changes in the parameters are large, the changes in the means and variances are small. This is an indication that the method is robust with respect to the parameter values.

Acknowledgements

The authors are grateful for having received information and useful comments from Francisco Javier Díez Vegas. This paper was written while Professor Hadi was visiting the Department of Applied Mathematics and Computing Sciences of the University of Cantabria (Spain). The authors are also grateful to the DGICYT of the Spanish Ministry of Education for a partial support Grant PB94-1056.

REFERENCES

Arnold, B., Castillo, E. and Sarabia, J. M. (1992), "Conditionally Specified Distributions," *Lecture Notes in Statistics*, Vol. 73, Springer Verlag, Berlin.

Arnold, B., Castillo, E. and Sarabia, J. M. (1995), "Specification of Distributions by Combinations of Marginal and Conditional Distributions," *Statistics and Probability Letters*, to appear.

Castillo, E. and Alvarez, E. (1991), *Experts Systems: Uncertainty and Learning*, Computational Mechanics Publications and Elsevier Applied Science, London.

Castillo, E., Gutiérrez, J. M., and Hadi, A. S. (1993), *Modelling Probabilistic Networks Based on Conditional/Marginal Probabilities*," Technical report No. 93-102, Department of Applied Mathematics and Computing Sciences, University of Cantabria.

Díez, F. J. and Mira, J. (1994), "Distributed Reasoning and Learning in Bayesian Expert Systems," in *Advances in Fault-Diagnosis Problem Solving*, (C. A. Newton, Ed.), CRC Press, Boca Raton, FL.

Fung, R. and Chang, K-C. (1990), "Weighing and Integrating Evidence for Stochastic Simulation in Bayesian Networks," in *Uncertainty in Artificial Intelligence 5*, Machine Intelligence and Pattern Recognition Series, Vol 10, (Henrion et al. Eds.), North Holland, Amsterdam, 209−219.

Gelman, A. and Speed, T. P. (1993), "Characterizing a Joint Probability Distribution by Conditionals," *Journal of the Royal Statistical Society*, B 55, No. 1, 185–188.

Henrion, M. (1988), "Propagating Uncertainty in Bayesian Networks by Probabilistic Logic Sampling," in *Uncertainty in Artificial Intelligence 2*, (J.F. Lemmer and L. N. Kanal, Eds.), North Holland, Amsterdam, 317–324.

Lauritzen, S. L. and Spiegelhalter, D. J. (1988), "Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems," *Journal of the Royal Statistical Society (B)*, 50, 157–224.

Lauritzen, S. L. (1992), "Propagation of Probabilities, Means, and Variances in Mixed Graphical Association Models", *Journal of the American Statistical Association*, Vol 87, No. 420, 1098–1108.

Liu, X. and Li, Z. (1994), A Reasoning Method in Damage Assessment of Buildings, Microcomputers in Civil Engineering, Special Issue on "Uncertainty in Expert Systems". To appear.

Norman, S. L. and Tritchler, D. (1992), "Parameter Updating in a Bayes Network", *Journal of the American Statistical Association*, Vol 87, No. 420, 1109–1115.

Pearl, J. (1986a), "Fusion, Propagation and Structuring in Belief Networks," *Artificial Intelligence*, 29, 241–288.

Pearl, J. (1986b), "Evidential Reasoning Using Stochastic Simulation of Causal Models," it Artificial Intelligence, 32, 245–287.

Pearl, J. (1991), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, Second printing, San Mateo, California.

Shachter, R. D. and Peot, M. A. (1990), "Simulation Approaches to General Probabilistic Inference on Belief Networks," in *Uncertainty in Artificial Intelligence 5*, Machine Intelligence and Pattern Recognition Series, Vol 10 (Henrion et al. Eds.), North Holland, Amsterdam, 221–231.

Spiegelhalter, D. J., Dawid, A. P., Lauritzen, S. L. and Cowell, R. G. (1993), "Bayesian Analysis in Expert Systems", *Statistical Science*, Vol 8, No. 3, 219–283.

Wermouth, N. and Lauritzen, S. L. (1990), "On Substantive Research Hypothesis, Conditional Independence Graphs, and Graphical Chain Models," *Journal of the Royal Statistical Society (B)*, 52, No. 1, 21–50.