Propagation of Linguistic Labels in Causal Networks*

Jorge Bolaños, Luis M. de Campos, Serafín Moral
Departamento de Ciencias de la Computación e I.A.
Universidad de Granada
18071 - Granada - Spain

***Abstract*— In this paper the problem of the propagation of linguistic labels in polytrees is considered. The approach has been purely simbolic, and does not consider semantic translations of the terms. It starts from a general axiomatic framework to propagate information in graphs, which is later particularized to the case of linguistic labels. It is observed how it is very difficult to define operations with linguistic labels verifying all the required properties, mainly because of the lack of granularity of finite sets of terms. Finally, to cope with these problems, modifications of the propagation algorithms are proposed.**

I. Introduction

The propagation of probabilities in graphical structures (see [9, 12]) is an efficient and very well founded procedure of handling uncertainty in Artificial Intelligence. Propagation algorithms take advantage of the independence relationships among the variables of a problem in order to calculate 'a posteriori' conditional probabilities, without explicitely giving a global probability for all the variables in the problem.

The weakest point of this methodology is that it requires a complete specification of some conditional probabilities. However, in some cases, they are unknown or only vaguely known. Several authors have studied propagation algorithms for the case in which we only have a set of possible conditional probabilities instead of knowing the specific values of conditional probabilities (see [3, 1, 16]). The resulting algorithms are, in general, more complex than their probabilistic counterpart.

The aim of this paper will be to devise propagation algorithms for linguistic probabilities. In most real situations, when we ask the expert for the value of a probability, he is unable to give a precise numerical value, however, in general, he will give a linguistic term, expressing his knowledge about this probability. One approach to make use of this kind of knowledge is to transform the linguistic terms into fuzzy numbers and then to use fuzzy arithmetic to generalize probabilistic propagation algorithms. This is the method used by Jain and Agogino, [7]. One characteristic of this method is that the resulting procedures involve exact calculations with real numbers, in fact, more than in the case of probabilistic propagation.

A completely different way of approaching this problem is by considering that a calculus with linguistic values does not need precise calculations, and that it is better to define symbolic operations directly in linguistic terms (see [11]). This is closer to the way we manage uncertainty in everyday reasoning. We do not rely on complicated operations, but on some qualitative relationships among the terms we use to express the degree of occurrence. This will be the perspective adopted in this work. We shall try to generalize probabilistic propagation algorithms to linguistic probabilities.

In the second section we shall give a general axiomatic framework to propagate uncertainty in directed acyclic graphs that was established in [4]. Next (section 3) we shall consider the desirable properties of the operations defined on a set of linguistic labels, which will allow to verify the general framework. These properties will hardly be verifiable by most common sets of linguistic labels. The problem will be the lack of precision available when using small sets of labels. In section 4 we shall consider how propagation algorithms can be modified to minimize the effect of this lack of precision.

II. The Valuation Framework

In [4] we have proposed an axiomatic framework for the propagation of uncertainty in Directed Acyclic Graphs. It is based on Shenoy and Shafer, [13] valuation based systems. Shenoy and Shafer, [13], introduce the primitive concept of valuation, which can be considered as the mathematical representation of a piece of information. A valuation may be particularized to a possibility distribution, a probability distribution, a belief functions, etc.. Then

*This work has been supported by the Commission of the European Communities under ESPRIT BRA 3085: DRUMS.

0-7803-0614-7/93$03.00

they develop and express propagation algorithms in terms of operations with valuations. These algorithms may be particularized to any concrete theory by translating valuations and operations to their special interpretation in this theory.

To fix the notation, let us assume that we have an n dimensional variable, $(X_1, \ldots, X_n)$, each X_i, taking values on a finite set U_i. Then the following conventions will be followed:

- if $I \subseteq \{1, \ldots, n\}$, by X_I we shall denote the family of variables $(X_i)_{i \in I}$, and by U_I the cartesian product $\prod_{i \in I} U_i$, that is the set in which X_I takes its values.
- For each $I \subseteq \{1, \ldots, n\}$, the set of valuations defined on U_I will be denoted by $\mathcal{V}_I$. An element of $\mathcal{V}_I$ will be the representation of a piece of information about the variable X_I.
- if $I \subseteq J$ and $u \in U_J$ then $u^{\downarrow I}$ will denote the point from U_I obtained from u eliminating the coordinates in U_J that are not in U_I. For example, if

$$I = \{1, 2\}, J = \{1, 2, 3\}, u = (u_1, u_2, u_3)$$

then $u^{\downarrow I}$ will be (u_1, u_2).
- $\mathcal{V}$ will be the set of all valuations $\mathcal{V} = \cup_{I \subseteq \{1,\ldots,n\}} \mathcal{V}_I$.

The following two operations are defined on the set of valuations, [13]:

- *Marginalization.-* If $J \subseteq I$ and $V_1 \in \mathcal{V}_I$ then the marginalization of V_1 to J is a valuation $V_1^{\downarrow J}$ in $\mathcal{V}_J$.
- *Combination.-* If $V_1 \in \mathcal{V}_I$ and $V_2 \in \mathcal{V}_J$, then their combination is a valuation $V_1 \otimes V_2$ in $\mathcal{V}_{I \cup J}$

The propagation algorithms are based on the verification of the following six axioms.

Axiom 1 $V_1 \otimes V_2 = V_2 \otimes V_1, \quad (V_1 \otimes V_2) \otimes V_3 = V_1 \otimes (V_2 \otimes V_3)$.

Axiom 2 If $I \subseteq J \subseteq K$, and $V \in \mathcal{V}_K$, then $(V^{\downarrow J})^{\downarrow I} = V^{\downarrow I}$.

Axiom 3 If $V_1 \in \mathcal{V}_I$, $V_2 \in \mathcal{V}_J$, then $(V_1 \otimes V_2)^{\downarrow I} = V_1 \otimes V_2^{\downarrow (J \cap I)}$.

Axiom 4 *Neutral Element.-* There is one and only one valuation V_0 defined on $U_1 \times \ldots \times U_n$ such that $\forall V \in \mathcal{V}^I, \forall J \subseteq I$, we have $V_0^{\downarrow J} \otimes V = V$.

Axiom 5 *Contradiction.-* There is one and only one valuation, V_c, defined on $U_1 \times \ldots \times U_n$, such that $\forall V \in \mathcal{V}$, $V_c \otimes V = V_c$.

Axiom 6 $\forall V \in \mathcal{V}_\emptyset$, if $V \neq V_c^{\downarrow \emptyset}$, then $V = V_0^{\downarrow \emptyset}$.

The first three axioms were introduced by Shafer and Shenoy, [13]. Axioms 4-6 were introduced by Cano, Delgado and Moral, [4].

The following definitions are fundamental to our work.

Definition 1 *A valuation $V \in \mathcal{V}_I$ is said to be absorbent if and only if it is not the contradiction in $\mathcal{V}_I$ and $(\forall V' \in \mathcal{V}_I)((V \otimes V' = V)$ or $(V \otimes V' = V_c))$.*

Definition 2 *If $V \in \mathcal{V}_{I \cup J}$, it is said that V is a valuation on U_I conditioned to U_J, if and only if $V^{\downarrow J} = V_0 \in \mathcal{V}_J$, the neutral element on $\mathcal{V}_J$. The subset of $\mathcal{V}_{I \cup J}$ given by the valuations on U_I conditioned to U_J will be denoted by $\mathcal{V}_{I|J}$.*

Details about the meaning of operations and the above definitions can be found in [4]. The following example describes valuations and operations in the probabilistic case.

Example 1 *In Probability Theory a valuation is the representation of a probabilistic piece of information about some of the variables, $X_I, I \subseteq \{1, \ldots, n\}$. More concretely, if we have three variables (X_1, X_2, X_3) taking values on $U_1 \times U_2, \times U_3$, where $U_i = \{u_{i1}, u_{i2}\}, i = 1, 2, 3$, then a valuation may be a probability distribution about X_1,*

$$p(u_{11}) = 0.8, \qquad p(u_{12}) = 0.2$$

It may also be a conditional probability distribution about X_3 given X_2,

$p(u_{31}|u_{21}) = 0.9 \quad p(u_{32}|u_{21}) = 0.1$
$p(u_{31}|u_{22}) = 0.6 \quad p(u_{32}|u_{22}) = 0.4$

From a mathematical point of view, a probabilistic valuation about variables X_I is a non-negative mapping,

$$p : U_I \longrightarrow \Re_0^+$$

where $\Re_0^+$ denotes the non-negative real numbers.

These mappings are not considered normalized, but are considered equivalent upon multiplication by a positive constant. That is, two valuations p_1, p_2 defined on the same frame U_I are considered equivalent if there is a constant $\alpha > 0$, such that

$$\forall u \in U_I, p_1(u) = \alpha . p_2(u)$$

From a strict mathematical point of view, a valuation should be considered as an equivalence class on the set of non-negative mappings from U_I on $\Re_0^+$, under the above equivalence relation. However, to simplify the language and notation, we shall consider that a valuation is a mapping, but that two mappings are considered identical if they are equivalent.

Combination is defined by point-wise multiplication. If p_1 and p_2 are non-negative functions defined on U_I and U_J respectively, then $p_1 \otimes p_2$ is a mapping defined on $U_{I\cup J}$ to $\Re_0^+$ given by,

$$p_1 \otimes p_2(u) = p_1(u^{\downarrow I}).p_2(u^{\downarrow J}), \forall u \in U_{I\cup J}$$

This operation is used in probability to combine a marginal distribution with a conditional one to produce a bidimensional distribution, or used to calculate conditional information. Remember that as we are not concerned about normalization, conditioning to a set A may be considered as multiplication by the likelihood associated to A (its characteristic function: $l_A(u) = 1$, if $u \in A$; $l_A(u) = 0$, otherwise).

Marginalization is defined in the usual way: If p is a valuation defined on U_I and $J \subseteq I$, then

$$p^{\downarrow J}(v) = \sum_{u^{\downarrow J}=v} p(u), \qquad \forall v \in U_J$$

The neutral element is a constant (non-zero) valuation: $p_0(u) = 1, \quad \forall u \in U_1 \times \ldots \times U_n$

The contradiction is given by the zero valued function: $p_c(u) = 0, \quad \forall u \in U_1 \times \ldots \times U_n$

An absorbent valuation is a probability degenerated on a point. It represents perfect knowledge about the true value of the variable.

The above conditional distribution about X_3 given X_2 verifies the definition of conditional valuations: If we marginalize to X_2 we obtain a constant valuation (1 at every point).

In general, when we have a set of pieces of information (valuations), $\mathcal{H} = \{V_j\}_{j\in J}$ about an n-dimensional variable $(X_1, \ldots, X_n)$ each one of them will only be related to some of the variables. For example, we may have an 'a priori' valuation about (X_2, X_3) or a conditional valuation about X_4 given X_1. In [4] rules to obtain a global (unconditional) valuation for the n-dimensional variable $(X_1, \ldots, X_n)$ are studied. These rules take advantage of the independence relationships among the variables. For example, if V_1 is an 'a priori' valuation about X_1, V_2 is a conditional valuation about X_2 given X_1, V_3 is a conditional valuation about X_3 given X_2, and X_3 and X_1 are conditionally independent given X_2 then we can deduce that $V_1 \otimes V_2 \otimes V_3$ is an unconditional valuation about (X_1, X_2, X_3).

An observation about a variable will be an absorbent valuation defined on it. If we have a global valuation V about variables $(X_1, \ldots, X_n)$ and $\{O_k\}_{k\in K}$ is a family of observations for some of the variables, $\{X_k\}_{k\in K}$,then the conditioning of this valuation to these observations is defined as

$$V|\{O_k\}_{k\in K} = \left[\bigotimes_{k\in K} O_k\right] \otimes V \qquad (1)$$

In probability theory we carry out a division, but here this division is avoided because valuations are equivalent under multiplication by a constant factor.

The problem of the calculus with valuations can be expressed in the following way:

1. We start with a set of pieces of information (valuations), $\mathcal{H} = \{V_j\}_{j\in J}$, representing our general knowledge about the problem, and a family of observations $\{O_k\}_{k\in K}$ for a particular case which we are studying.

2. In an intermediate step, we calculate an unconditional valuation V for all the variables. This valuation is conditioned to the observations, calculating $V|\{O_k\}_{k\in K}$.

3. $V|\{O_k\}_{k\in K}$ is marginalized to the variables in which we are interested.

The main problem of the above procedure is that step 2 is very expensive if the number of variables is high. An unconditional valuation would be defined on the cartesian product $\prod_{i\in\{1,\ldots,n\}} U_i$ and the number of elements of this set is the product of the number of elements of all the sets U_i. Probabilistic propagation algorithms in directed acyclic graphs developed by Pearl, [12], take advantage of the independence relationships among variables in order to avoid the second step, carrying out step 3 directly from step 1, using only initial local pieces of information. In [4] these algorithms have been generalized to the case of valuations. The main results of this work can be summarized in the following points:

1. In a directed acyclic graph (nodes represent variables and the arcs express dependence relationships) if we have a conditional valuation, V_i, for each variable, X_i, given the parent variables in the graph then we can determine one and only one unconditional global valuation which is equal to $V_1 \otimes \ldots \otimes V_n$.

2. If the graph does not have loops (undirected cycles) and we want to calculate $PS_j = \{V|\{O_k\}_{k\in K}\}^{\downarrow\{j\}}$ (step 3), then this can be carried out by using algorithms that are very similar to the Pearl's, in which each variable receives (send messages) from (to) his neigbourhood variables. If

   - $O'_k = V_0 \in \mathcal{V}_{\{j\}}$, if $j \notin K$
   - $O'_k = O_k$, if $k \in K$

- $P(j)$ is the set of parents of node X_j
- $C(j)$, is the set of children of node X_j.

then $\forall j \in \{1, \ldots, n\}, PS_j = \pi_j \otimes \lambda_j$, where

$$\pi_j = \left[V_j \otimes \left(\bigotimes_{k \in P(j)} \pi_j^k \right) \right]^{\downarrow \{j\}} \tag{2}$$

$$\lambda_j = \left(\bigotimes_{k \in C(j)} \lambda_j^k \right) \otimes O'_j \tag{3}$$

$\forall k \in P(j)$,

$$\pi_j^k = \left[\pi_k \otimes O'_k \otimes \left(\bigotimes_{i \in C(k),\ i \neq j} \lambda_k^i \right) \right] \tag{4}$$

$\forall k \in C(j)$,

$$\lambda_j^k = \left[\lambda_k \otimes O'_k \otimes V_k \left(\bigotimes_{i \in P(k),\ i \neq j} \pi_k^i \right) \right]^{\downarrow \{j\}} \tag{5}$$

π_j^k is called the message that node j sends to his child k. λ_j^k is the message that node j receives from his child k. The fundamental fact for the propagation is:

> An outgoing message from node X_i to node X_j can be calculated from all the incoming messages to X_i, except the incoming message coming from X_j.

We can consider propagation algorithms as an ordered way of calculating all the messages among nodes taking into account the above considerations. Equations (4) and (5) are the ones used in the calculations of the messages going out of a node from the incoming messages to it. More details can be found in [12, 4, 3].

As in [12] directed acyclic graphs without loops will be called polytrees.

III. Valuations Taking Values on Sets of Labels

In [8] the calculus with linguistic probabilities taking values on a set of labels was studied. The idea was to determine the necessary operations and properties which we should define on a set of labels in order to reproduce probability calculus. In this section we shall propose the properties which will allow us to define linguistic valuations and to propagate them in directed acyclic graphs.

Definition 3 *A probabilistic set of labels will be a set, L, in which two operations (addition , $\oplus$, and multiplicacion, $\odot$) and a total order relation, $\preceq$, are defined, verifying the following properties:*

1. Commutative.- $\forall l_1, l_2 \in L,\ \ l_1 \oplus l_2 = l_2 \oplus l_1,\ \ l_1 \odot l_2 = l_2 \odot l_1$
2. Associative.- $\forall l_1, l_2, l_3 \in L,\ \ (l_1 \oplus l_2) \oplus l_3 = l_1 \oplus (l_2 \oplus l_3),\ \ (l_1 \odot l_2) \odot l_3 = l_1 \odot (l_2 \odot l_3)$
3. Neutral Element.- *There is one and only one element $0 \in L$, such that $0 \oplus l = l, \forall l \in L$.*

   There exists one and only one element $1 \in L$, such that $1 \odot l = l, \forall l \in L$.
4. Distributive.- $\forall l_1, l_2, l_3 \in L,\ \ l_1 \odot (l_2 \oplus l_3) = (l_1 \odot l_2) \oplus (l_1 \odot l_3)$
5. *$a \preceq b$ if and only if $\exists c \in L$ such that $a \oplus c = b$.*
6. Inverse element.- *If $a \preceq b$, then $\exists d \in L$ such that $b \odot d = a$, d being unique if $a \neq 0$.*

An example of a probabilistic set of labels is the set of non-negative real numbers.

The valuations defined in a probabilistic set of labels are very similar to the probabilistic valuations. A valuation on U_I will be a mapping:

$$V : U_I \longrightarrow L \tag{6}$$

Two valuations, V_1, V_2, defined on the frame U_I are considered equivalent if a label $l \neq 0$ exists, such that

$$\forall u \in U_I, V_i(u) = l \odot V_j(u) \tag{7}$$

where $i = 1, j = 2$ or $i = 2, j = 1$.

With the same reasoning as in the probabilistic case, two valuations will be considered identical if they are equivalent.

Combination is defined in the following way: If $V_1 \in \mathcal{V}_I, V_2 \in \mathcal{V}_J$,

$$V_1 \otimes V_2(u) = V_1(u^{\downarrow I}) \odot V_2(u^{\downarrow J}), \forall u \in U_{I \cup J} \tag{8}$$

Marginalization is defined as follows: If V is a valuation defined on U_I and $J \subseteq I$, then

$$V^{\downarrow J}(v) = \bigoplus_{u^{\downarrow J} = v} V(u), \qquad \forall v \in U_J \tag{9}$$

The neutral element is a constant (non-zero) valuation: $V_0(u) = 1,\ \ \forall u \in U_1 \times \ldots \times U_n$

The contradiction is given by the zero valued function:

$$V_c(u) = 0, \qquad \forall u \in U_1 \times \ldots \times U_n$$

It can be shown that valuations defined in this way verify Axioms 1-6, then these valuations can be represented and propagated by using dependence graphs.

Example 2 *Let $L_1 = \{0\} \cup \{l_i \mid i \geq 0\}$, where addition is defined as*

- $0 + l = l, \quad \forall l \in L_1$
- $l_i + l_j = l_{\mathrm{Min}\{i,j\}}, \forall i, j \geq 0$

and multiplication is

- $l \times 0 = 0, \forall l \in L_1$
- $l_i \times l_j = l_{i+j}, \forall i, j \geq 0$

The minimum element is 0 and $l_i \preceq l_j$ if and only if $i \geq j$.

When assigned to an event, a value of 1 (l_0) means that there is no limitation of the probability of this event. A value of 0 means that this event is impossible. A value of l_i means that the probability of this event is of the order ρ^i, where ρ is a 'small' number. This semantic establishes a direct relationship with Spohn conditionals, [14, 15], which have the same interpretation.

In this case, all the properties are satisfied. The neutral element for multiplication is l_0. As a conclusion, this kind of labels can be propagated in polytrees.

IV. Finite Linguistic Sets of Labels

Unfortunately things are not so easy when we are working with labels which are more similar to the ones used in everyday human speech. In general these labels are generated by a finite set of terms, [2, 5, 6, 8, 10]. The following set is from [6]:

$$L = \{0, AN, F, AH, M, 1\}$$

where the intuitive meaning of the labels is

0	None	AN	Almost None
F	Few	AH	About Half
M	Most	1	Almost All or All

The order of the labels is immediate:

$$0 \preceq AN \preceq F \preceq AH \preceq M \preceq 1$$

To define the operations among labels, first let us note that the addition of some labels is not defined. For example, there is no label that could be assigned to $1 \oplus 1$. The reason is that we are only working with the labels between 0 and 1. The complete set of labels, L', would include the elements greater than one. In general, to work with valuations is enough to consider the interval $[0, 1]$ of labels. The operations will only be defined when the operators and the result are in this interval.

Table 1: Operations with a finite set of labels

$\oplus$	0	AN	F	AH	M	1
0	0	AN	F	AH	M	1
AN	AN	AN	F	AH	M	1
F	F	F	AH	M	1	–
AH	AH	AH	M	1	–	–
M	M	M	1	–	–	–
1	1	1	–	–	–	–

$\odot$	0	AN	F	AH	M	1
0	0	0	0	0	0	0
AN	0	AN	AN	AN	AN	AN
F	0	AN	AN	AN	AN	F
AH	0	AN	AN	AN	F	AH
M	0	AN	AN	F	AH	M
1	0	AN	F	AH	M	1

Another important point is how to determine the result of one operation, for example, $F \oplus AH$. For this we may adopt a semantic approach which consists in translating the labels to numerical intervals or fuzzy numbers, [2, 6], carrying out the operation at this level and then coming back to the labels by means of an approximation procedure. It is also possible [5, 10] to consider a direct elicitation of these operations from the experts. Here, we shall follow a purely symbolic point of view in which operations are given by means of tables, without worrying about the procedure of the construction of these tables. As an example of the operations see table 1.

As was noticed in [8], some of the required properties are very difficult to verify, so we have to discard them from the beginning. These are:

- Inverse element.- Consider the multiplication of a label, for example AH, by a label different from 0. The result should be less than or equal to AH and different from 0. In our case, by multiplying AH by all the non 0 labels we get:

  $$AN - AN - AN - F - AH$$

  As we have a finite granularity, and we multiply more labels than the number of possible results, then some of them have to be repeated. The inverse of one repeated label with respect to AH is not well defined: There are several labels, that multiplied by AH, give AN as the result.

- *Distributive*.- If we assume that $AN \oplus AN = AN$ and there is a label l such that $l \odot l = AN$, and

$\overbrace{l \oplus \cdots \oplus l}^{i\ times} = 1$ (F verifies these conditions in our case), then the distributive property can not be satisfied. In effect, we have

$$l \odot (\overbrace{l \oplus \cdots \oplus l}^{i\ times}) = l \odot 1 = l$$

and

$$\overbrace{(l \odot l) + \cdots + (l \odot l)}^{i\ times} = \overbrace{AN \oplus \cdots \oplus AN}^{i\ times} = AN.$$

These properties can not be verified because of the lack of granularity in the set of terms. These properties are desirable but they are impossible to verify with the finite sets we are considering. Anyway, we can establish that these properties are verified in most of the cases. In our example, the main problem is with the AN value. The other cases verify inverse property with respect to a greater value and distributive property is verified if AN is not one of the used labels. The question is: How to cope with these limitations of a finite set of terms? First, we shall consider the consequences of not verifying these properties. Valuations can be defined and propagation formulas can be used. But there are two axioms that are not verified. The first is Axiom 3 (this axiom is based on distributive property). As a consequence the result of applying propagation formulas is not equal to $PS_j = \{V|\{O_k\}_{k\in K}\}^{\downarrow\{j\}}$, this is the result of building a global piece of information V, combining it with the observations, and marginalizing the result to U_j. One could view the result of propagation as being incorrect. However, our point of view is that the result of propagation is a better assesment of the 'a posteriori' information than PS_j. As we do not have enough granularity, building a global valuation and marginalizing it after can degrade the information, while it can be more appropriate to use only local valuations. The following example is an extreme case of this idea.

Example 3 *Consider two variables, X_1 and X_2, taking values on $U_1 = \{u_{11}, u_{12}\}, U_2 = \{u_{21}, u_{22}\}$, respectively. Assume that we know that these two variables are independent and that we have the 'a priori' valuations V_1 and V_2 about them, where*

- $V_1(u_{11}) = 1, \quad V_1(u_{12}) = AN$
- $V_2(u_{21}) = AN, \quad V_2(u_{22}) = 1$

Assume that we observe the value u_{21} for X_2. The induced information in X_1 is:

a) By using propagation formulas we have a graph consisting of two nodes without links (the variables are independent), then the result of propagation is the same valuation V_1.

b) By calculating PS_1 through a global valuation we get:

$$PS_1(u_{11}) = AN, \quad PS_1(u_{12}) = AN$$

In the second case the information has been degenerated. However, in the first case we have used the independence relationship to do less calculations and therefore the result is more appropriate.

The other axiom that is not fulfilled is Axiom 6. If inverse property is not verified, we can not consider two valuations as equivalent if condition (7) is given. As a consequence a constant valuation is not longer a neutral element and multiplying by a value may change a valuation. The consequences of using propagation algorithms are analyzed in the following example.

Example 4 *Let us use the example of a machine that is prepared to work in very tough environmental conditions. The performance of the machine is very good in bad conditions and excellent in good conditions. This machine, almost always works in bad conditions. This situation can be modeled with two variables: X_1 (conditions) and X_2 (state), where X_1 takes the values: G (good) and B (bad), and X_2 the values W (working) and NW (not working). The graph consists of 2 nodes, X_1 and X_2, with an arc from X_1 to X_2. As 'a priori' information about X_1 we have:*

$$V_1(G) = AN, V_1(B) = 1$$

As conditional valuation about X_2 given X_1 we have,

$$V_{2|1}(W|G) = 1, V_{2|1}(NW|G) = AN$$

$$V_{2|1}(W|B) = M, V_{2|1}(NW|B) = F$$

Let us now assume that we know that in a particular situation the machine is working in good conditions. The message that X_1 sends to X_2 is π_1^2 with

$$\pi_1^2(G) = AN, \pi_1^2(B) = 0$$

Calculating the value of PS_2 according to propagation formulas we get,

$$PS_2(W) = AN, PS_2(NW) = AN$$

We get the same value for working *and* not working. *However, we should obtain the value 1 forW and AN for NW. The problem is that with the lack of granularity it is impossible to discriminate between $1 \odot AN$ and $AN \odot AN$. The value AN contaminates the rest of the valuations.*

To solve the problem of propagation of the lack of granularity we propose a solution consisting in normalizing the messages that nodes send each other. In the general framework, when all properties are verified this normalization is senseless: we obtain an equivalent valuation. But here it is important to maintain the maximum precision in the calculations. First, we give a definition of the normalization of valuations defined on a set of labels.

Definition 4 *Let V be a valuation on U_I conditioned to U_J we say that this valuation is normalized if and only if*

$$Sup\ \{V^{\downarrow J}(u) \mid u \in U_J\} = 1 \qquad (10)$$

For an unconditional valuation ($J = \emptyset$) the meaning of (10) is that the addition of all the values is 1. π_j^i messages are unconditional valuations in U_j, thus the normalization is a transformation to a valuation in which all the values add up to 1. The λ_j^k messages are valuations in the empty set conditioned to U_j. Then the normalization condition is,

$$\text{Sup}\ \{\lambda_j^k(u) \mid u \in U_j\} = 1$$

That is, that the supremum of the values is equal to 1.

Definition 5 *A normalization function N is a mapping applying a valuation, V, in a set of normalized valuations $N(V)$ such that for every $V' \in N(V)$ there exists a label $l \in L$ such that*

$$\forall u, \quad V(u) = l \odot V'(u) \qquad (11)$$

The result of normalizing a valuation is a set of valuations in the general case. The reason is that as the property of the inverse element is not verified we may find several normalized valuations, V', fulfilling (11). We shall assume that this set is never empty. This happens with the labels we are considering.

Example 5 *Consider the set $U = \{u_1, u_2\}$ and the valuation V given by*

$$V(u_1) = AN, V(u_2) = AN$$

For all normalized valuations, V', in U with all the values different from 0, there exists a label $l \in L$ (this label can be AN) such that (11) is verified.

In the case of the valuation π_1^2 of the above example, $N(\pi_1^2)$ has only a valuation, V', being

$$V'(u_{11}) = 1, V'(u_{12}) = AN$$

The propagation problems of this example could be solved after transforming π_1^2 into a normalized valuation.

To cope with the problem of normalization in the general case, we have two extreme approaches:

- *The strong approach.-* In this method the normalization of a valuation is the set of all the valuations verifying (11). It consists in considering a new concept of valuation. A new valuation in U_I, $\mathbf{V}$, will be a set of mappings from U_I to L. The combination is carried out by making the pointwise combination of all the mappings in the valuations:

$$\mathbf{V_1} \otimes \mathbf{V_2} = \{V \mid V = V_1 \otimes V_2, \quad V_1 \in \mathbf{V_1}, V_2 \in \mathbf{V_2}\} \qquad (12)$$

The marginalization is defined on an analogous way:

$$\mathbf{V}^{\downarrow I} = \{V' \mid V' = V^{\downarrow I}, \quad V \in \mathbf{V}\} \qquad (13)$$

The normalization of a new valuation, $\mathbf{V}$, will be the set of all the mappings obtained by normalization of the mappings in $\mathcal{V}$. In this way we may use propagation formulas (4,5) taking into account that valuations are now sets of mappings and that normalization is applied after their calculation.

This method is very similar to the one used in [6] in which the result of the propagation is a set of labels. This imprecision is a consequence of the imprecision of the set of labels we are using.

- *The weak approach.-* This method consists in considering that the result of normalization is a unique valuation. In this case we ignore some of the imprecision of the set of labels we are using but the calculations are simpler. Valuations are composed of an unique mapping. This can be considered as an approximate approach. We find that it is difficult to deduce any particular rule to select a normalized valuation. However we propose the following rule based on the heuristic principle of avoiding extreme probability assignments as much as possible:

  1. Let l be AN.
  2. While the set of normalized valuations has more than one element do
     3. Remove from the set of normalized valuations those with a number of values equal to l greater than the number of values equal to l of other valuation in the set.
     4. Let l be the following label in L.

With our set of labels, this algorithm ends with an unique valuation. In the case of

$$V(u_1) = AN, \quad V(u_2) = AN$$

we get

$$N(V)(u_1) = AH, \quad N(V)(u_2) = AH$$

This is a kind of maximum entropy principle.

References

[1] Amarger S., D. Dubois, H. Prade (1991) Constraint propagation with imprecise conditional probabilities. In: *Uncertainty in Artificial Intelligence. Proceedings of the 7th Conference* Morgan & Kaufmann, San Mateo, 26-34.

[2] Bonissone P.P. and K.S. Decker (1986) Selecting uncertainty calculi and granularity: an experiment in trading-off precision and complexity. In: *Uncertainty in Artificial Intelligence* (L.N. Kanal and J.F. Lemmer, eds.) Amsterdam, North-Holland, 217-247.

[3] Cano J.E., M. Delgado, S. Moral (1991) Propagation of Uncertainty in Dependence Graphs. En: *Symbolic and Quantitative Approaches to Uncertainty* (R. Kruse, P. Siegel, eds.) Springer Verlag. Lecures Notes in Computer Science 548, 42-47.

[4] Cano J.E., M. Delgado, S. Moral (1992) An axiomatic system for the propagation of uncertainty in directed acyclic networks. To appear in: *International Journal of Approximate Reasoning.*

[5] Delgado M., J.L. Verdegay and M.A. Vila (1988) Ranking linguistic outcomes under fuzziness and randomness. *Proceedings of the Eighteenth International Symposium on Multiple-Valued Logic*, Computer Society Press, 352-356.

[6] Dubois D., H. Prade, L. Godó, R, López de Mántaras (1992) A Symbolic Approach to Reasoning with Linguistic Quantifiers. *Uncertainty in Artificial Intelligence. Proceedings of the Eighth Conference*, Morgan & Kaufmann, San Mateo, 74-82.

[7] Jain P., A.M. Agogino (1990) Stochastic sensitivity analysis using fuzzy influence diagrams. In: *Uncertainty in Artificial Intelligence, 4* (R.D. Shacter et al., eds.) North-Holland, Amsterdam, 79-92.

[8] Lamata M.T., S. Moral (1992) Calculus with Linguistic Probabilities and Beliefs. To appear: *Advances in the Dempster-Shafer Theory of Evidence* (M. Fedrizzi et al., eds.) John Wiley, New York.

[9] Lauritzen S.L., D.J. Spiegelharter (1988) Local computation with probabilities on graphical structures and their application to expert systems. *J. of the Royal Statistical Society*, B 50, 157-224.

[10] López de Mántaras R., L. Godo and R. Sanguesa (1990) Connective operators elicitation for linguistic term sets. *Proc. Inter. Conference on Fuzzy Logic and Neural Networks*, Iizuka, Japan, 729-733.

[11] Parsons S., J. Fox (1991) Qualitative and Interval Algebras for Robust Decision Making under Uncertainty. In: *Decision Support Systems and Qualitative Reasoning* (M. Singh, L Travé-Massuyès, eds.)

[12] Pearl J. (1989) *Probabilistic Reasoning with Intelligent Systems.* Morgan & Kaufman, San Mateo.

[13] Shenoy P.P., G. Shafer (1990) Axioms for probability and belief-functions propagation. In: *Uncertainty in Artificial Intelligence, 4* (R.D. Shachter et al., eds.) North-Holland, Amsterdam, 169-198.

[14] Spohn W. (1988) Ordinal conditional functions: a dynamic theory of epistemic states. In: *Causation in Decision, Belief Change, and Statistics* (W.L. Harper and B. Skyrms, eds.) Dordrecht, Kluwer, 105-134.

[15] Spohn W. (1990) A general non-probabilistic theory of inductive reasoning. In: *Uncertainty in Artificial Intelligence 4* (R.D. Shacter et al., eds.), Amsterdam, North-Holland, 149-158.

[16] Tessen B. (1989) Interval Representation of Uncertainty in Artificial Intelligence. Ph. D. Thesis, Department of Informatics, University of Bergen, Norway.