Proceedings of the 4th WSEAS Int. Conf. on Information Security, Communications and Computers, Tenerife, Spain, December 16-18, 2005 (pp472-477)

Knowledge Representation with Possibilistic and Certain Bayesian Networks

ABDELKADER HENI
Preparatory Institute for Engineering Studies
Department of Technology
Kairouan Road, 5019, Monastir
TUNISIA

MOHAMED-NAZIH OMRI
Preparatory Institute for Engineering Studies
Department of Technology
Kairouan Road, 5019, Monastir
TUNISIA

ADEL M ALIMI
National School of Engineers
Department of Electrical Engineering
Soukra Road, 3000 Sfax,
TUNISIA

Abstract: -Possibilistic logic and Bayesian networks have provided advantageous methodologies and techniques for computer-based knowledge representation. This paper proposes a framework that combines these two disciplines to exploit their own advantages in uncertain and imprecise knowledge representation problems. The framework proposed is a possibilistic logic based one in which Bayesian nodes and their properties are represented by local necessity-valued knowledge base. Data in properties are interpreted as set of valuated formulas. In our contribution possibilistic Bayesian networks have a qualitative part and a quantitative part, represented by local knowledge bases. The general idea is to study how a fusion of these two formalisms would permit representing compact way to solve efficiently problems for knowledge representation. We show how to apply possibility and necessity measures to the problem of knowledge representation with large scale data.

Key-Words: - Possibilistic logic, Bayesian networks, Certain Bayesian networks, Local knowledge bases

1 Introduction

Bayesian networks have attracted much attention recently as a possible solution to complex problems related to decision support under uncertainty. These networks are systems for uncertain knowledge representation and have a big number of applications with efficient algorithms and have strong theoretical foundations [1],[2],[3],[4],[5] and [11]. They use graphs capturing causality notion between variables, and probability theory (statistic data) to express the causality power.

Although the underlying theory has been around for a long time, the possibility of building and executing realistic models has only been made possible because of recent improvements on algorithms and the availability of fast electronic computers. On the other hand, one of the main limits of Bayesian networks is necessity to provide a large number of numeric data; a constraint often difficult to satisfy when the number of random variables grows up. The goal of this paper is to develop a qualitative framework where the uncertainty is represented in possibility theory; an ordinal theory for uncertainty developed since more than ten years [6], [7], and [8]. Our framework propose to define a qualitative notion of independence (alternative to the probability theory), to propose techniques of decomposition of joined possibility distributions, and to develop some efficient algorithms for the revision of beliefs. Thus, on the first hand limitations of quantitative structure in Bayesian networks that use simple random variables have been noted by many researches. These limitations have motivated a variety of recent research in hierarchical and composable Bayesian models.

On the other hand, another limitation of the use of probabilistic Bayesian networks in expert systems is difficulty of obtaining realistic probabilities. So to solve these problems we use a new modified possibilistic Bayesian method. Our new modified possibilistic Bayesian networks simultaneously make use of both possibilistic measures: necessity measure and possibility measure.

Our work extends and refines these proposed frameworks in a number of crucial ways. The language defined in [12] [13] and in [15] has been modified to enhance usability and to support a more powerful system. We are trying in this paper to describe a language that provides the important capability of uncertainty modeling. We have also combined different element from works cited above to describe our possibilistic networks based on local necessity-valued knowledge bases.

In this paper we consider a type of possibilistic network that is based on the context model interpretation of a degree of possibility and focused on imprecision [14]. The first section presents an overview of standard possibilistic networks and their extensions. The following section describes our contribution with the use of necessity-valued knowledge bases as quantitative representation for uncertainty in nodes. And eventually we will talk about the transformations between average certain-possibilistic Bayesian networks and average knowledge bases.

2 Basic notions of necessity-possibility measures and possibilistic networks

In order to be able to discuss our framework for possibilistic networks we shall in this section give a few preliminary definitions and notational conventions. At the same time, we present a brief outline of few important notations and ideas in possibility theory and possibilistic networks relevant to the subject of this paper.

2.1 Possibilistic logic

Let L be a finite propositional language. p; q; r; . . . denote propositional formulae.

$\top$ and $\bot$, respectively, denote tautologies and contradictions. $\vdash$ denotes the classical syntactic inference relation. Ω is the set of classical interpretations ω of L, and $[p]$ is the set of classical models of p (i.e, interpretations where p is true $\{\omega \mid \omega \models p\}$) [13].

2.1.1 Possibility distributions and possibility measures

The basic element of possibility theory is the possibility distribution $\prod$ which is a mapping from Ω to the interval [0 1]. The degree $\pi(\omega)$ represents the compatibility of ω with the available information (or beliefs) about the real world. By convention, $\pi(\omega)= 0$ means that the interpretation ω is impossible, and $\pi(\omega) = 1$ means that nothing prevents ω from being the real world [13].

Given a possibility distribution π, two different ways of rank ordering formulae of the language are defined from this possibility distribution. This is obtained using two mappings grading, respectively, the possibility and the certainty of a formula p:

• The possibility (or consistency) degree:

$$\prod(p) = max\ (\ \pi(\omega) : \omega \in [p]) \quad (1)$$

Which evaluates the extent to which p is consistent with the available beliefs expressed by p [16]. It satisfies:

$$\forall p,\ \forall q \quad \prod(p \vee q) = \max(\prod(p), \prod(q)) \quad (2)$$

• The necessity (or certainty, entailment) degree

$$N(P) = 1 - \prod(\neg p) \quad (3)$$

Which evaluates the extent to which p is entailed by the available beliefs. We have [17]:

$$\forall p,\ \forall q \quad N(p \wedge q) = \max(N(p), N(q)) \quad (4)$$

To note here that in our case, we consider that both necessity degree and possibility degrees for a given formulae should be given by an expert. On the other hand, when a data is required (a possibility degree or necessity degree) one should deduce it by applying equation (3).

2.1.2 Possibilistic knowledge base

A possibilistic formula is a pair (φ, α) where φ is a classical first-order closed formula and $\alpha \in [0,1]$ is a positive number. (φ, α) expresses that φ is certain at least to the degree α, i.e. $N(\varphi) \geq \alpha$, where N is a necessity measure modelling our possibly incomplete state of knowledge. The right part of a possibilistic formula, i.e. α, is called the *weight* of the formula.

A possibilistic knowledge base $\sum$ is defined as the set of weighted formulae [18]. More formally $\sum = \{(\varphi_i, \alpha_i),\ i = 1....m\}$ where φ_i is a propositional formula and α_i is the lower bound of necessity accorded to this formula (certainty degree).

3 Certain-possibilistic Bayesian networks

A standard possibilistic network is a decomposition of a multivariate possibility distribution according to:

$$\pi(A_1,\ldots,A_n) = \min_{i=1..n} \pi(A_i \mid \text{parents}(A_i)) \quad (5)$$

where *parents*(A_i) is the set of parents of variable Ai, which is made as small as possible by exploiting conditional independencies of the type indicated above [9] and [10]. Such a network is usually represented as a directed graph in which there is an edge from each of the parents to the conditioned variable.

In our work an average certain-possibilistic Bayesian networks is considered as a graphical representation of uncertain information. It offers an alternative to probabilistic causal network when numerical data are not available.

Let $V= \{A_1,A_2,..A_n\}$ be a set of variables (i.e attributes or proprieties). The set of interpretations is the Cartesian product of all domains of attributes in V. When each attribute is binary, domains are denoted by $D_i=\{a_i, \neg a_i\}$.

An *average certain-possibilistic graph* denoted by $\prod G^A$ is an acyclic graph where nodes represents attributes i.e. a patient temperature and edges represent causal links between them. Uncertainty is represented by

Proceedings of the 4th WSEAS Int. Conf. on Information Security, Communications and Computers, Tenerife, Spain, December 16-18, 2005 (pp472-477)

possibilities distribution, certainties distribution and conditional possibilities and necessities for each attribute explaining the link force between them.
The conditional possibilities and necessities distributions are associated to the graph as follow:

For each root attribute A_i, we specify prior possibility distribution $\Pi(a_i), \Pi(\neg a_i)$ and the prior normalization) and the prior necessity distribution $N(a_i)$, $N(\neg a_i)$ with the constraint that :

$$\begin{cases} N(a_i) = 1 \vdash N(\neg a_i) = 0 \\ N(\neg a_i) = 1 \vdash N(a_i) = 0 \end{cases} \quad (6)$$

- For other attributes A_j, we specify the conditional possibilities distribution $\Pi(a_j|u_j)$, $\Pi(\neg a_j|u_j)$ with $\max(\Pi(a_i|u_j), \Pi(\neg a_i|u_j)) = 1$ where u_j is an instance of a_j parents and the conditional necessity distribution $N(a_i)$, $N(\neg a_i)$ with the constraint that :

$$\begin{cases} N(a_i|u_j) = 1 \vdash N(\neg a_i)|u_j) = 0 \\ N(\neg a_i)|u_j) = 1 \vdash N(a_i)|u_j) = 0 \end{cases} \quad (7)$$

Example: the next figure gives an example of possibilistic Bayesian networks with four nodes and their conditional possibilities.

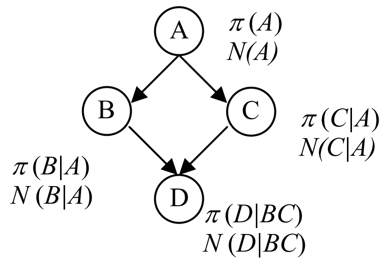

Fig. 1: example of an certain-possibilistic Bayesian network

The *joint average distribution* is obtained then by applying the chain rule:

$$A(A_1,.,A_n) = \min(\pi(A_i|U(A_i)) * \min(N(A_i|U(A_i)) \quad (8)$$

Where:

- $A(A_1,.,A_n)$ is The joint average distribution.
- $\min(\pi(A_i \mid U(A_i))$ is the lower bound of the possibilities degrees associated to $(A_i|U(A_i))$.
- $\min(N(A_i|U(A_i))$ is the lower bound of the necessities degrees associated to $(A_i|U(A_i)$

Example: let the prior possibilities-necessities and the conditional possibilities-necessities be as described in table 1:

	π	N	A
a	1	0.6	0.6
¬a	0.5	0.1	0.05

	π	N	π	N
B|A	a	a	¬a	¬a
b	1	0.5	0.75	0.2
¬b	0.5	0.25	0.3	0

	π	N	π	N
C|A	a	a	¬a	¬a
C	1	0.3	0.7	0.2
¬c	0.6	0.1	0.4	0.1

	π	N	π	N	π	N
D|BC	Bc	bc	b¬c	b¬c	else	Else
d	1	0.2	0.5	0.1	1	0.3
¬d	0.5	0.4	0.3	0.1	0.7	0.2

Table 1: possibilities-necessities distribution

By the use of the chain rule defined by equation (8) we obtain the average distribution associated with the average certain-possibilistic Bayesian network cited above as described in table.

A	B	C	D	minΠ	minN	𝒜
a	b	c	d	1	0.2	0.2
a	b	c	¬d	0.5	0.3	0.15
a	b	¬c	d	0.5	0.1	0.05
a	b	¬c	¬d	0.3	0.3	0.09
a	¬b	c	d	0.5	0.2	0.1
a	¬b	c	¬d	0.5	0.1	0.05
a	¬b	¬c	d	0.5	0.1	0.05
a	¬b	¬c	¬d	0.5	0.1	0.05
¬a	b	c	d	0.5	0.1	0.05
¬a	b	c	¬d	0.5	0.1	0.05
¬a	b	¬c	d	0.4	0.1	0.04
¬a	b	¬c	¬d	0.3	0.1	0.03
¬a	¬b	c	d	0.3	0	0
¬a	¬b	c	¬d	0.3	0	0
¬a	¬b	¬c	d	0.3	0	0
¬a	¬b	¬c	¬d	0.3	0	0

Table 2: joint average possibility-necessity distribution

4 Average certain-possibilistic valued knowledge base

We would like to represent a class of possibilistic Bayesian networks using a local average certain-possibilistic valued knowledge base consisting of a collection of possibilistic logic sentences (formulae) in such a way that a network generated on the basis of the information contained in the knowledge base is isomorphic to a set of ground instances of the formulae. As the formal representation of the knowledge base, we use a set of possibilistic formulae. We represent random variables with necessities and possibilities weights and restrict ourselves to using only the average of these two measures.

Formally an average necessity-possibility valued knowledge base is defined as the set :

$$\Sigma = \{(\varphi_i, \alpha_i, \beta_i), i = 1 \ldots m\} \quad (9)$$

Where φ_i denotes a classical propositional formula, α_i and β_i denote respectively the lower bound of certainty (i.e necessity) and the lower bound of possibility.

We can represent the information contained in each node of a Bayesian network, as well as the quantitative information contained in the link matrices, if we can represent all the direct parent/child relations. We express the relation between each random variable and its parents over a class of networks with a collection of quantified formulae. The collection of formulae represents the relation between the random variable and its parents for any ground instantiation of the quantified variables. The network fragment consisting of a random variable and its parents with a set of formulae of the form (φ, α, β).

We give next some definitions inspired from [12] and [13].

Definition 1:
Two average knowledge bases Σ^A_1 and and Σ^A_2 are said to be equivalent if their associated possibility distributions (respectively necessity distributions) are equal, namely:

$$\begin{cases} \forall\omega \in \Omega,\ \pi\Sigma^A_1(\omega) = \pi\Sigma^A_2(\omega) \\ \text{and} \\ \forall\omega \in \Omega,\ N\Sigma^A_1(\omega) = N\Sigma^A_2(\omega) \end{cases} \quad (10)$$

Definition 2:
Let (φ, α, β) a formula in Σ^A Then (φ, α, β) is said to be subsumed by Σ^A if Σ^A and $\Sigma^A \backslash \{(\varphi, \alpha, \beta)\}$ are equivalent knowledge bases.

This is means that each redundant formula should be removed from the average valued knowledge base since it can be deduced from the rest of formulae.

5 From average certain possibilistic network to certain possibilistic valued knowledge base

In this section, we describe the process that permit to deduce an average valued knowledge base from an average network.

Let ΠG^A be an average certain-possibilistic Bayesian network consisting of a set of labeled variables $V= \{A_1,A_2,..A_n\}$. Now let A be a binary variable and let $(a\ \neg a)$ be its instances.
Given the two measures $\pi\ (a_i|u_i)$ and $N(a_i|u_i)$ witch represent respectively the local possibility degree and the local necessity degree associated with the variable A where $u_i \in U_A$ is an instance of parents(a_i). the local average knowledge base associated with A should be defined using the next equation :

$$\Sigma^A_A = \{(\neg a_i \vee u_i, \alpha_i, \beta_i), \alpha_i = 1 - \pi(a_i|u_i) \neq 0 \text{ and } \beta_i = 1 - N(a_i|u_i) \neq 0\ \} \quad (11)$$

To note here that in [15] the authors prove the possibility to recover conditional possibilities from Σ_A where Σ_A is a possibilistic knowledge base.

Based o the results obtained in [15] , we can check in our case that it is possible to recover both conditional necessities from Σ^A_A according to equations (12) and (13).

$$\Pi\Sigma^A(\omega) = \begin{cases} 1 \text{ if } \forall\ (\varphi_i, \alpha_i) \in \Sigma\ \ \omega \vDash \varphi_i \\ 1 - \max\{\alpha_i : \omega \nvDash \varphi_i\} \text{ otherwise} \end{cases} \quad (12)$$

and

$$N\Sigma^A(\omega) = \begin{cases} 1 \text{ if } \forall\ (\varphi_i, \alpha_i) \in \Sigma\ \ \omega \vDash \varphi_i \\ 0 \text{ otherwise} \end{cases} \quad (13)$$

Proceedings of the 4th WSEAS Int. Conf. on Information Security, Communications and Computers, Tenerife, Spain, December 16-18, 2005 (pp472-477)

Example: by applying equation (11), we get the average knowledge base associated to the average certain-possibilistic Bayesian network described in section 3.

Σ^{A}_{A} = {(a, 0.5, 0.9)} = {(a, 0.45)}
Σ^{A}_{B} = {(b∨a,0.7),(b∨¬a,0.5,0.75)(¬b∨a,0.25, 0.8)}
= {(b∨a, 0.7), (b∨¬a, 0.375) (¬b∨a, 0.2)}

Σ^{A}_{C}={(c∨a,0.6,0.9),(c∨¬a,0.4,0.9) (¬c∨a,0.3, 0.8)}
= {(c∨a, 0.54), (c∨¬a, 0.36 0.9) (¬c∨a, 0.24)}

Σ^{A}_{D} = {{(d∨b∨c, 0.3, 0.8), (d∨b∨¬c, 0.3, 0.8), (d∨¬b∨c, 0.7, 0.9), (d∨¬b∨¬c, 0.5, 0.6), (¬d∨¬b∨c, 0.5, 0.9)}
= {{(d∨b∨c, 0.24), (d∨b∨¬c, 0.24), (d∨¬b∨c, 0.63), (d∨¬b∨¬c, 0.3), (¬d∨¬b∨c, 0.45)}

Remark: for each knowledge base the first equality represents the initial knowledge base weighted by possibilities and necessities when the other represents the average based knowledge base (namely average necessity-possibility valued knowledge base).

Next section shows the other face of transformation between average valued knowledge base and average certain-possibilistic Bayesian network.

6 From Average valued knowledge base to average certain-possibilistic Bayesian network

In [15] the authors describe a process permitting to deduce a possibilistic network from a possibilistic knowledge base. In this section we follow the same way to transform our average necessity-valued knowledge bases into an average certain-possibilistic Bayesian network.

To note here that the average certain-possibilistic Bayesian network deduced from an average necessity-valued knowledge bases will have the same graphical structure as the starting network

The conditional average distributions are simply the ones associated with the average knowledge bases. More precisely, let A_i be variable and u_i be an element of parents(A_i). Let Σ^{A}_{Ai} be the local average knowledge base associated with the node A_i. Then, the conditional average degree $A(a_i|u_i)$ is defined by $\pi(a_i|u_i) = \pi(a_i \wedge u_i) = \pi\Sigma^{A}_{Ai}(a_i \wedge u_i)$ and $\Sigma^{A}_{Ai}(a_i \wedge u_i)$ is defined using equation (12) and equation (13).

Respectively the conditional necessity degree $N(a_i|u_i)$ is defined by $N(a_i|u_i) = N(a_i \wedge u_i) = N\Sigma^{A}_{Ai}(a_i \wedge u_i)$.
Example:

From the average knowledge base associated to the node *A* and by the use of equations 11 and 12
Σ^{A}_{A} = {(a, 0.5, 0.9)}
= {(a, 0.45)}
We can deduce the conditional average table for node *A* by the use of equations 11 and 12

	π	N	A
a	1	0.6	0.6
¬a	0.5	0.1	0.05

Same to rest of nodes we can deduce the rest of conditional averages associated to other nodes and so we can recover the average distribution presented in table 2.

7 Conclusion

This paper has presented a definition of certain-possibilistic Bayesian networks and how to use them to deduce average knowledge bases and vice versa. Uncertainty in nodes in our models is represented by local knowledge bases.

The key benefits of this representation to the practitioner are that both knowledge declaration and possibilistic inference are modular. Individual knowledge bases should be separately compilable and query complete. Also this representation specifies an organized structure for elicitation of the graph structure. We only defined the transformation process for knowledge bases.

A future work is to extend this representation by definition of efficient algorithms for locally inferences.

References

[1] Pearl Judea. Bayesian networks : *a model of self-activated memory for evidential reasoning*. Cognitive Science Society, UC Irvine, pages 329–334, (1985).
[2] Pearl Judea. *A constraint-propagation approach to probabilistic reasoning*. In : L.N. Kanal and J. F. Lemmer (eds), Uncertainty in Artificial Intelligence, Amsterdam, NorthHolland, pages 3718–1986, (1986).
[3] Pearl Judea. *Fusion, propagation and structuring in belief networks*. UCLA Computer Science Department Technical Report 850022 (R-42) ; Artificial Intelligence,, 29 :241– 288, (1986).
[4] Pearl Judea and A. Paz. *Graphoids : a graph-based logic for reasoning about relevance relations*. UCLA Computer Science Department Technical Report

850038; In B. Du Boulay, et.al. (Eds.), Advances in Artificial Intelligence-II, North-Holland Publishing Co., (1987).
[5] Pearl Judea and T. Verma. *Influence diagrams and d-separation.* UCLA Cognitive Systems Laboratory, Technical Report 880052, (1988).
[6] D. Dubois, S. Moral, H. Prade, *A semantics for possibility theory based on likelihoods*, J. Math. Anal. Appl. 205, 359-380. (1997)
[7] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.
[8] D. Dubois, H. Prade, *Théorie des Possibilités: Applications à la Représentation des Connaissances en Informatique*, Masson, Paris,1985.
[9] N.Ben Amor. *Qualitative possibilistic graphical models from independance to propagation algorithms*. Thèse de doctorat, (2002).
[10] J.Gebhardt and R.Kruse, *Background and perspectives of possibilistic graphical models*, in proceeding of European C onference of Symbolic and Quantitative Approches to Rreasoning and Uncertainty, pages 108-121, Bad Honnef Germany, (1997).
[11] J. Pearl. Probabilistic *Reasoning in Intelligent Systems: Networks of Plausible Inference*.Morgan Kaufmann Publishers, Palo Alto, (1988).
[12] S. Benferhat, S. Smaoui. *Hybrid possibilistic networks*. in proceeding of the Twentieth National Conference on Artificial Intelligence (AAAI-05),AAAI Press. Pittsburgh, 2005.
[13] S. Benferhat, D. Dubois, L. Garcia, H. Prade. *On the transformation between possibilistic logic bases and possibilistic causal networks*. International Journal of Approximate Reasoning, Vol. 29, N. 21, 35-173, 2002.
[14] J. Gebhardt and R. Kruse. *The context model an integrating view of vagueness and uncertainty* Int. Journal of Approximate Reasoning 9 283–314, 1993
[15] Salem Benferhat and Salma Smaoui *Possibilistic networks with locally weighted knowledge bases* 4th International Symposium on Imprecise Probabilities and Their Applications, Pittsburgh, Pennsylvania, 2005
[16] L.A. Zadeh, *Fuzzy sets as a basis for a theory of possibility*, Fuzzy Sets and Systems 1 (1978) pp 3-28.
[17] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and applications*, Academic Press, New York, 1980.
[18] Dubois D. (1986) *Belief structures, possibility theory and decomposable confidence measure. on finite sets*. Computers and Artificial Intelligence, 5(5), 403-416.