Inference in possibilistic hypergraphs

Didier Dubois, Henri Prade

5. NON-STANDARD LOGICS

INFERENCE IN POSSIBILISTIC HYPERGRAPHS

Didier DUBOIS - Henri PRADE

Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier, 118 route de Narbonne
31062 TOULOUSE Cedex (FRANCE)

Abstract

In order to obviate soundness problems in the local treatment of uncertainty in knowledge-based systems, it has been recently proposed to represent dependencies by means of hypergraphs and Markov trees. It has been shown that a unified algorithmic treatment of uncertainties via local propagation is possible on such structures, both for belief functions and Bayesian probabilities, while preserving the soundness and the completeness of the obtained results. This paper points out that the same analysis applies to approximate reasoning based on possibility theory, and discusses the usefulness of the idempotence property for combining possibility distributions, a property not satisfied in probabilistic reasoning. The second part analyzes a previously proposed technique for handling dependencies, by relating it to the hypergraph approach.

Keywords

Hypergraph, Markov trees, possibility theory, approximate reasoning.

1. INTRODUCTION

Starting with a knowledge base made of pieces of information which may be pervaded with uncertainty or vagueness, an ideal inference mechanism should be able to provide the user with conclusions which are i) sound, ii) as certain and precise as it is possible, iii) obtained by a computation procedure which is as efficient as possible. Obviously, this aim may be more or less easily reached, depending on the modeling of uncertainty which is used and the format of the pieces of knowledge allowed in the framework of this modeling. We may distinguish between local and global computation methods (e.g. Dubois and Prade, 1987). A typical local approach to the treatment of uncertainty is the one used in rule-based expert systems where rules are evaluated and triggered one after the other according to the control procedure of the inference engine. The procedure in that case is computationally efficient but no guarantee exists usually about the optimality in terms of certainty and precision of the conclusion and even in some cases its soundness may be questioned. This is mainly due to problems raised by the necessity of combining partial uncertain conclusions pertaining to the same matter, without being able to take into account implicit dependencies or redundancies in the knowledge base. The probabilistic reasoning method advocated by Nilsson (1986) is an example of global approach where the best lower and upper bounds on the probability of a conclusion are obtained by solving a (linear) programming problem whose constraints are directly obtained from the available bounds on the probability of the different granules in the whole knowledge base. In that case the soundness and the optimality of the conclusion can be guaranteed, but the computational cost may be heavy in practice and moreover it is then difficult to provide such an inference system with explanation capabilities.

Local propagation methods have been recently proposed in Bayesian networks (Pearl, 1986), and Markov networks (Lauritzen and Spielgelhalter, 1988). This kind of approach assigns precise probability values to any proposition in a given context, making an extensive use of conditional independence assumptions. Then the soundness and optimality issues coincide in the sense that in these approaches one is interested in getting estimations of the probability values which are as close as possible to the actual values. Both Bayesian and Markov networks aim at representing conditional independence assumptions using oriented graphs and non- oriented graphs respectively. Similarly, Shafer et al.(1987) have proposed to use hypergraphs and Markov trees for representing dependence relations between variables, the uncertainty being modelled in terms of belief functions (Shafer, 1976) and have developed local propagation methods. In a more general framework Shafer and Shenoy (1988, 1990) have studied axioms which propagation and combination operations should satisfy to ensure that local algorithms correctly work in hypergraph representations ; see also (Williams, 1990).

Independently and quite at the same time an approach has been proposed for handling dependencies in approximate reasoning by Chatalic, Dubois and Prade (1987). It is applicable to Shafer's belief functions as well as to Zadeh's (1978) possibility measures. The purpose of this paper is to give an improved presentation of this approach, in the framework of possibility theory, to connect it with the approaches mentioned at the end of the preceding paragraph, and also to relate it to recent extensions of the constraint propagation paradigm for the handling of fuzzy values (Dubois and Prade, 1989a, b ; Yager, 1989). Indeed the updating of the possible ranges of variables linked by a set of constraints has been studied for a long time in Artificial Intelligence and the general procedure originally proposed by Waltz has been refined for many particular cases corresponding to different kinds of relations between the variables ; see Davis (1987) for instance. Let us also mention the work recently developed by Kruse and Schwecke (1988, 1989, 1990) in the possibilistic framework for handling dependencies in (causal) networks.They propose a local propagation algorithm for rule-based systems where rules are expressed as dependence relations, and show that this approach carries over to fuzzy rules following Zadeh(1979).

2. POSSIBILISTIC REASONING AND HYPERGRAPHS

In possibility theory (Zadeh, 1978), what is known about the value of variables or about the existing relations between variables is represented by means of possibility distributions. Namely, a possibility distribution π_X attached to the variable X, is a mapping from the domain $\mathcal{X}$ of X to the interval [0,1] ; $\forall\ x \in \mathcal{X}$, $\pi_X(x)$ reflects to what extent it is possible that X = x according to the available information. If two possibility distributions π_X and π'_X such that $\pi_X \leq \pi'_X$ are available from two different sources, π_X is said to be more restrictive than π'_X since each possible value for X receives a smaller possibility degree according to π_X than according to π'_X ; π_X then expresses a less uncertain and/or imprecise information. The normalization condition which is usually applied to π_X is $\exists\ x \in \mathcal{X}$, $\pi_X(x) = 1$ which expresses that at least one value in $\mathcal{X}$ is considered as completely possible for X (exhaustiveness of the domain) and which allows that distinct values in $\mathcal{X}$ be simultaneously regarded as completely possible. More generally a possibility distribution $\pi_{X_1,\ldots,X_n}$, defined on a Cartesian product $X_1 \times \ldots \times X_n$, expresses a dependency relation between the variables $X_1, \ldots, X_n$; for instance $\pi_{X_1,\ldots,X_n}(x_1, \cdot, \ldots, \cdot)$ characterizes the fuzzy set of the more or less possible values of the tuple $(X_2,\ldots,X_n)$ when $X_1 = x_1$. Normalization of $\pi_{X_1,\ldots,X_n}$ guarantees the existence of a completely possible interpretation $(x_1,\ldots,x_n)$ compatible with the piece of knowledge expressed by $\pi_{X_1, \ldots, X_n}$.

In this framework a general procedure known as the conjunction/projection method (Zadeh, 1979) can be applied to a knowledge base in order to deduce what can be said about the value of a variable of interest, or the relationship which jointly constrains a set of variables. Namely, let $A_1, \ldots, A_n$ be the fuzzy sets of the possible values of the variables $X_1, \ldots, X_n$ (i.e. $\pi_{X_i} = \mu_{A_i}$) ; if nothing is known about the value of X_i, then $\mu_{A_i}(x_i) = 1,\ \forall\ x_i \in \mathcal{X}_i$. Let $R_1,\ldots, R_m$ be the (fuzzy) relations stated in the knowledge base between variables. R_j is then defined on the Cartesian product of the domains $\mathcal{X}_k$'s of the variables X_k involved in the relationship represented by R_j. Then the conjunction/projection method consists in

i) performing the combination

$$\pi^*_{X_1,\ldots,X_n} = \min(\min_{i=1,n} \mu_{A_i}, \min_{j=1,m} \mu_{R_j}) \qquad (1)$$

which is the *least restrictive* possibility distribution for the tuple $(X_1, \ldots, X_n)$ compatible with all the constraints.

ii) projecting the result $\pi^*_{X_1,\ldots,X_n}$ on the domain(s) of the variable(s) we are interested in.For instance we get for X_i

$$\forall\ x_i \in \mathcal{X}_i,\ \pi_{X_i}(x_i) = \sup_{x_j,\ j=1,n,\ j\neq i} \pi^*_{X_1,\ldots,X_n}(x_1, \ldots, x_n) \qquad (2)$$

$\pi^*_{X_1,\ldots,X_n}$ is generally supposed to be normalized.The lack of normalization of $\pi^*_{X_1,\ldots,X_n}$ would express the inconsistency of the knowledge base ; namely, the quantity $1 - \sup_{x_1,\ldots,x_n} \pi^*_{X_1,\ldots,X_n}$ estimates the degree of inconsistency of the knowledge base. The projection preserves the normalization. So, there always exists an interpretation fully compatible with any conclusion obtained by the combination/projection approach when the knowledge base is consistent.

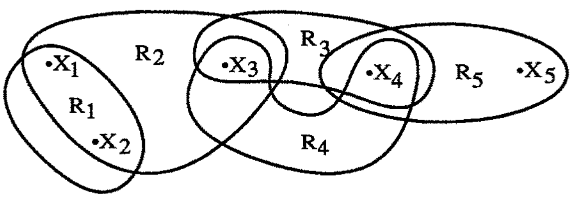


Figure 1

An example of knowledge base is illustrated by the hypergraph on Fig. 1, where variables are nodes and relations are pictured by hyperedges ; an hyperedge, denoted S, corresponds to a subset of variables (nodes) ; however there may be several hyperedges between the same subset of nodes. Then for instance, computing the possibility distribution attached to X_3 from the knowledge of A_2 and A_3 (the other A_i's being such that $A_i = \mathcal{X}_i$) leads to the expression, $\forall x_3 \in \mathcal{X}_3$

$$\pi_{X_3}(x_3) = \sup_{x_1,x_2,x_4,x_5} \min(\mu_{R_1}(x_1,x_2), \mu_{R_2}(x_1,x_2,x_3), \mu_{R_3}(x_3,x_4), \mu_{R_4}(x_3,x_4), \mu_{R_5}(x_4,x_5), \mu_{A_2}(x_2), \mu_{A_3}(x_3)).$$

Here the hypergraph visualises the decomposition of a possibility distribution into non-interactive components; this feature make approximate reasoning similar to Bayesian networks (decomposition of a joint probability into a product of marginals or conditional probabilities) and belief function techniques (decomposition of a belief function into independent pieces of evidence). From the point of view of constraint propagation, the updating of the variable X_i, from A_i to A'_i, is expressed in this framework by

$$\forall i, \forall x_i,\ \mu_{A'_i}(x_i) = \sup_{\substack{x_j \\ j=1,n\ ;\ j\neq i}} \min(\min_{k=1,m} \mu_{R_k}(x_1,\ldots,x_n), \min_{j=1,n} \mu_{A_j}(x_j)) \quad (3)$$

$$\leq (\min(\mu_{A_i}(x_i), \min_{k=1,m}[\sup_{\substack{x_j \\ j=1,n\ ;\ j\neq i}} \min(\mu_{R_k}(x_1,\ldots,x_n), \min_{j=1,n\ ;\ j\neq i} \mu_{A_j}(x_j))]) \quad (4)$$

The inequality (4) expresses that if we take into account each R_k separately in the updating process, we are not sure, even if we iterate the procedure as in the Waltz algorithm, of obtaining the most restrictive possibility distribution for X_i ; however the result provided by the right-hand part of (4) is more easy to compute in general and obviously sound (in the sense that what is obtained is not as restrictive and informative as what could be obtained, but, as such, cannot be arbitrarily restrictive). Note that in case of binary relations, it can be shown that the Waltz procedure (i.e. the separate processing of the R_k's) yields the most accurate result given by (3), provided there is at most one relation R_k between any pair of variables and that there is no cycle in the non-oriented graph whose nodes correspond to variables and edges to binary relations.

More generally, it is not surprizing that the general procedure which consists first in performing the general combination of all the representations of the pieces of information in the knowledge base and then projecting the result, is generally computationally untractable in practice. Then it is natural to try to take advantage of the fact that each relation usually relates only a (small) subset of the variables and to exploit the structural properties enjoyed by the combination and projection operations. Indeed it can be easily checked that the four axioms which are at the basis of the local computation scheme proposed by Shafer and Shenoy (1988a) (see also Williams, 1990), are satisfied in the possibilistic framework. Namely let $I \subset \{1, \ldots, n\}$, R be defined on the Cartesian product of $\mathcal{X}_i$'s where $i \in I$, we have

A0 ("Identity") $\sup_{x_i,\ i\notin I} \mu_R(\ldots, x_i, \ldots) = \mu_R$ (assimilating μ_R with its cylindrical extension on $\mathcal{X}_1 \times \ldots \times \mathcal{X}_n$).
A1 ("Consonance of marginalization") Let $K \subseteq J \subseteq I$, then

$$\sup_{x_i,\ i\notin K} \mu_R(\ldots, x_i, \ldots) = \sup_{x_i, i\notin K} (\sup_{x_i, i\notin J} \mu_R(\ldots, x_i, \ldots))$$

A2 Commutativity and associativity of the combination (obvious with 'min')
A3 ("Distributivity of marginalization over combination") Let S be a relation defined on the Cartesian product of $\mathcal{X}_i$'s where $i \in L \subset \{1, \ldots, n\}$, then

$$\sup_{x_i,\ i\notin I} \min(\mu_R, \mu_S) = \min(\mu_R, \sup_{x_i \notin I\cap L} \mu_S)$$

(for instance, let R and R' be defined on $X_1 \times X_2$ and $X_2 \times X_3$ respectively, then

$$\sup_{x_3} \min(\mu_R(x_1,x_2), \mu_{R'}(x_2,x_3)) = \min(\mu_R(x_1,x_2), \sup_{x_3} \mu_{R'}(x_2,x_3)))$$

A2 and A3 also hold with conjunctive operations other than min (e.g. product, which is used in the Bayesian setting and in evidence theory). These properties ensure that the local treatment of uncertainty is equivalent to the global treatment, provided that the hypergraph can be translated into a Markov tree. A Markov tree is a tree $G = (\mathcal{N},\mathcal{A})$ such that its nodes correspond to hyperedges of the hypergraph, or unions thereof ; there is an edge $(S,S') \in \mathcal{A}$ whenever $S \cap S' \neq \emptyset$, and moreover if S, S' are any two nodes, $S \cap S'$ is contained in all nodes on the (unique) path from S to S' (Markov property). When the Markov tree conditions are fulfilled, $\mathcal{N}$ is called an hypertree. Given a general hypergraph, it is possible to cover it with a hypertree by merging suitable hyperedges. The bigger the hyperedges, the greater the spaces in which global combination projection operations must be done. Hence the problem is to cover hypergraphs with hypertrees having as small hyperedges as possible (see Shafer et al., 1987 ; Shafer and Shenoy, 1990).

In the possibilistic framework the combination is also idempotent. An interesting question is then : how does this property simplify a local treatment of uncertainty ? First, it must be noticed that in possibilistic reasoning the projection operation is the same as for general belief functions. Now, as a consequence of the idempotence of min, performing local inference on a possibilistic hypergraph will never produce wrong results. The reason why this is not so with belief functions is that Dempster rule used on related pieces of evidence can take into account twice the same information, thus leading to arbitrary reinforcement effects. Nothing of the like occurs in possibilistic reasoning. The tree structure turns out to be useful for possibilistic reasoning because a local treatment of cyclic structures usually leads to a loss of information through projection steps.

3. AN HYPERTREE GENERATION AND QUESTION ANSWERING PROCEDURE

A procedure for simultaneous treatment of dependencies and question answering has already been proposed by Chatalic, Dubois and Prade (1987) and Chatalic (1986). The procedure makes use of two basic transformations which are to be applied the dependency hypergraph, namely

The (hyper)edge fusion

It aims at merging dependence relations pertaining to the same variables, so that combination may take place before any projection step. A simple example is on Figure 2 : knowing $\pi_X = \mu_{A_1}$, if we want to compute π_{X_2} we cannot "propagate along" R_1 and R_2 separately and combine the results afterwards, without loosing information, as a consequence of (4) (see Dubois, Martin-Clouaire and Prade, 1988) ; namely :

$$\sup_{x_1} \min(\mu_{A_1}(x_1), \mu_{R_1}(x_1,x_2), \mu_{R_2}(x_1,x_2)) \\ \leq \min(\sup_{x_1} \min(\mu_A(x_1), \mu_{R_1}(x_1,x_2)), \sup_{x_1} \min(\mu_A(x_1), \mu_{R_2}(x_1,x_2))$$

(which is a particular case of (3)-(4)). This operation also applies to any family of hyperedges such that one of them contains the other ones.

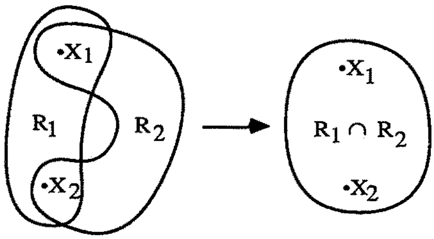


Figure 2

The variable fusion

It aims at destroying cyclic structures by considering two or more variables jointly as a vector-valued variable. It results in the fusion of hyperedges in the original hypergraph and eventually produces a Markov tree structure. A simple example of this operation is on Figure 3. If we want to evaluate X_1, we must evaluate X_2 and X_3 due to R_2 and R_3 ; in turn evaluating X_2 requires the evaluation of X_3 and evaluating X_3 requires the evaluation of X_2 due to R_1. The way out of this tricky situation is to conjointly evaluate X_2 and X_3. Then the hyperedges corresponding to $R_2 \cap R_3$ and R_1 can be fused if needed. Clearly this still applies when there is more than one variable in common between two relations. It is also used when longer cycles are present. Consider the example on Figure 4

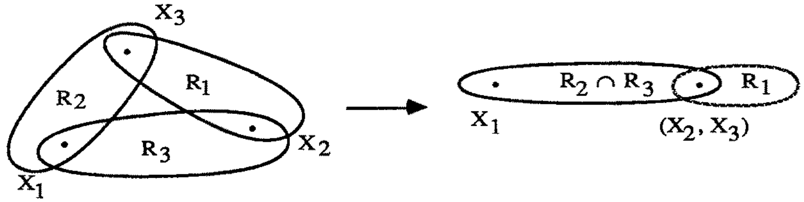


Figure 3 : Evaluating X_1

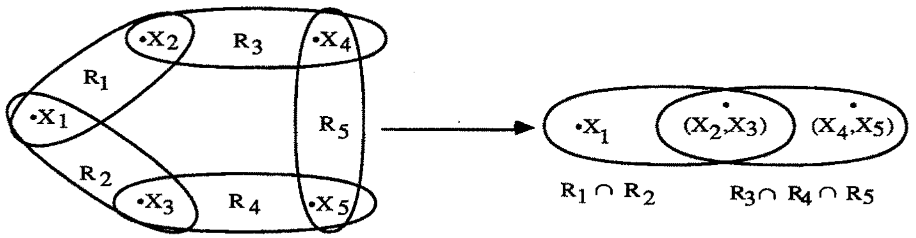


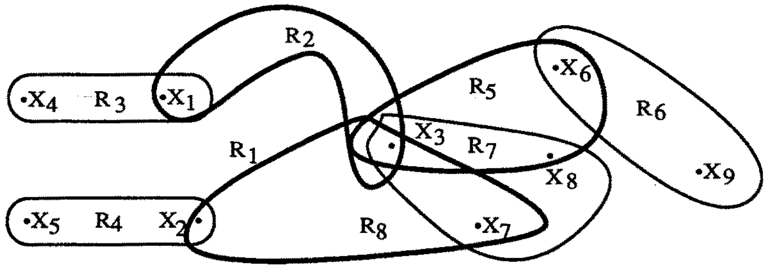
Figure 4

Starting with X_1, X_2 and X_3 are reached via R_1 and R_2, and then X_4 and X_5. The discovery of the dependence between X_4 and X_5 leads to fusing X_2 and X_3, and also X_4 and X_5.

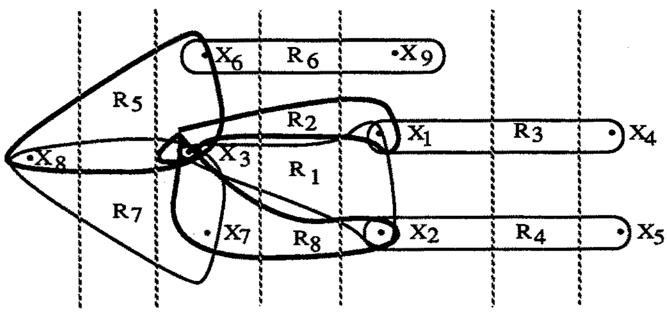
One basic idea of the procedure presented in the following is that the treatment of the "dependency hypergraphs", induced by the relationships stated in the knowledge base, can depend on the query which is addressed to the inference system. Indeed if we consider the hypergraph in the left part of Fig. 4, in order to evaluate X_1 we can perform a variable fusion as explained above, while for instance for computing π_{X_2} we can perform another variable fusion (i.e. combining R_1 and R_3). The procedure consists in

i) starting with variable(s) to evaluate, identifying the relations in which the variable(s) take(s) part, then finding out the other variables involved in these relations and then iterating this process until we have considered all the dependency relations between variables ;
ii) recursively applying the hyperedge fusion and the variable fusion to the dependency graph thus oriented at step i).

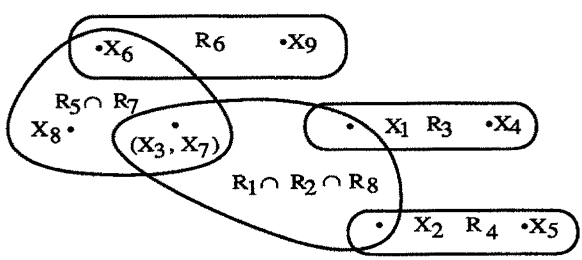
An example of this procedure is shown in Fig. 4, starting with the "dependency hypergraph" shown in Fig. 4.a, the result of step i) is shown in Fig. 4.b and the hypertree finally obtained is in Fig. 4.c. It is easy to see that the obtained hypertree is a covering of the original hyperstructure generally. Details can be found in (Chatalic, 1986 ; Chatalic, Dubois and Prade, 1987).



(a)



(b)



(c)
Figure 5

This procedure produces an hypertree which can be directly exploited in order to obtain a sound, optimal (in the sense that it yields the most restrictive possibility distribution) evaluation of the variable(s) we are interested in. Moreover the hypertree shows what are the sequence(s) of computations to be performed and what computations can be done separately taking advantage as far as possible of the property expressed by axiom A_2. For instance, in Figure 5.c, we see that when looking for evaluation of X_8, we can evaluate X_6 only on the basis of i) R_6 and X_9 and also of ii) the information we may already have on X_6 (updating process) ; we also see that X_3 and X_7 have to be evaluated conjointly, R_1, R_2 and R_8 have to be used in combination for evaluating this pair of variables and so on. Note also that if we have no information on some variable (e.g. X_4), it is allowed to forget the corresponding branch (here X_1 R_3 X_4) in the evaluation, provided that the involved relation is normalized (i.e. $\sup_{x_4} \mu_{R_3}(x_1,x_4) = 1, \forall\ x_1$). The procedure puts the different computation steps at a level which is as local as possible if we want to be sure to have a result which is both sound and optimal, since in the hyperedge and variable fusions, and thus in the proposed procedure, we are only taking advantage of the properties A0-A3 (but not of idempotence).

Chatalic's procedure mixes up the problem of finding an hyperedge cover and the problem of evaluating a variable. Namely the hypertree in Figure 5 is obtained assuming that X_8 is to be evaluated. When evaluating another variable, another hypertree is produced from 5(a). However it is worth noticing that the hypertree on 5(c) can serve for a sound and complete evaluation of any variable. The above procedure can produce as many hypertrees as involved variables in the network. The best hypertree is again the one that has the smallest hyperedges. However the above procedure is not optimal in that respect. Figure 6 gives a better hypertree cover for the example on Figure 4.

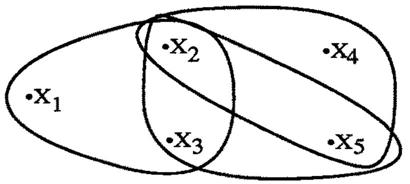


Figure 6

Indeed, the biggest hyperedge of any hypertree cover for the hypergraph, computed via Chatalic's procedure will contain four variables, due to obvious symmetry reasons, while on Figure 6, hyperedges contain only 3 variables.

N.B. : Let R be a fuzzy relation defined on $\mathcal{X}_1$ x...x $\mathcal{X}_k$ x $\mathcal{X}_{k+1}$ x...x $\mathcal{X}_n$ such that $\forall x_1,\ldots, \forall x_k$, $\exists x_{k+1},\ldots, \exists x_n$ such that $\mu_R(x_1,\ldots, x_k, x_{k+1},\ldots, x_n) = 1$. $X_1,\ldots,X_k$ are called input variables of R. This condition guarantees that whatever the available information on $X_1,\ldots,X_k$ is, provided that this information is expressed in terms of *normalized* possibility distributions, the result of any projection of the combination of R with this information, will be always normalized. In other words any information on $X_1,\ldots,X_k$ will be consistent with R. Thus if all the relations obtained in the hyperedge fusion process (e.g., R_3, R_4, R_6, $R_1 \cap R_2 \cap R_8$, $R_5 \cap R_7$ in Fig. 5.c) enjoy this property with respect to their input variables, then whatever the available information on the input variables of the hypertree we have built, we are sure to obtained a normalized possibility distribution on the variable(s) we are evaluating (i.e. X_8 in the example of Fig. 5). When some relations do not satisfy the above condition with respect to their input variables, we are in the situation recently considered by Yager and Larsen (1990) of a *potentially* inconsistent knowledge base.

4. CONCLUDING REMARKS AND OTHER RELATED ISSUES

In this paper we have considered that what is available in the knowledge base, are possibility distributions restricting tuples of variables. In practice, the expert knowledge is rather given under the form of rules from which it is possible to build the associated possibility distributions, taking into account the intended meaning of the rules (see Dubois and Prade, 1989c). So, the procedure described here could be applied to fuzzy expert systems, especially the hyperedge fusion for combining parallel rules relating the same variables, rather than combining their conclusions. However it can be shown that the most efficient way of computing with rules in parallel is not to combine their associated possibility distributions explicitly as formally suggested by the procedure we have presented, but generally to consider them jointly and to build from them a new rule adapted to each input value to consider. It gives the optimal result directly (see Dubois, Martin-Clouaire and Prade, 1988). But this point concerns the local computation within a node of a Markov tree representation. It does not question the interest of representing a knowledge base of fuzzy rules via a Markov tree structure.

Based on this framework, it is interesting to investigate computational problems similar to the ones addressed by Pearl (1986) in Bayesian networks : updating the value of X_i from a new piece of data about X_j elsewhere in the network ; re-computing the most possible tuple(s) of values for the variables $X_1 \ldots X_n$. These problems should be solved via local propagation, provided that decomposability assumptions hold for the global fuzzy relation relating the set of variables. This achievement would be a significant step ahead with regard to current fuzzy logic applications (e.g. in fuzzy control), where only one layer of parallel rules is considered. A first attempt has been recently made by Fonck (1990) who proposes a possibilistic counterpart of Pearl's local computation approach. It would correspond here to the particular case where all the variables take their values on binary universes of discourse. Lastly, another worth-considering issue would use the proposed framework for analyzing the consequences of the *approximation* of complex hyperedges by simpler (hyper)edges, by fixing the values of "secondary" variables by means of fuzzy default values, as already discussed in (Dubois and Prade, 1988).

ACKNOWLEDGEMENTS : This work has been partially supported by the P.R.C.-G.D.R. Intelligence Artificielle (CNRS and French Ministry of Research) and by the ESPRIT-BRA project DRUMS. Thanks also to Sandra Sandri for useful comments on a first draft of this paper.

REFERENCES

Chatalic P. (1986) Raisonnement déductif en présence de connaissances imprécises et incertaines : un système basé sur la théorie de Dempster-Shafer. Thèse de Doctorat, Univ. P. Sabatier, Toulouse, France.

Chatalic P., Dubois D., Prade H. (1987) A system for handling relational dependencies in approximate reasoning. Proc. of the 3rd Inter. Expert Systems Conf., London, June 2-4, 495-502.

Davis E. (1987) Constraint propagation with interval labels. Artificial Intelligence, 32, 281-331.

Dubois D., Martin-Clouaire R., Prade H. (1988) Practical computing in fuzzy logic. In : Fuzzy Computing (M.M. Gupta, T. Yamakawa, eds.), North-Holland, 11-34.

Dubois D., Prade H. (1987) A tentative comparison of numerical approximate reasoning methodologies.

Int. J. of Man-Machine Studies, 27, 717-728.

Dubois D., Prade H. (1988) On the combination of uncertain or imprecise pieces of information in rule-based systems – A discussion in the framework of possibility theory. Int. J. of Approximate Reasoning, 2, 65-87.

Dubois D., Prade H. (1989a) Fuzzy arithmetic in qualitative reasoning. In : Modeling and Control of Systems in Engineering, Quantum Mechanics, Economics, and Biosciences (Proc. of the Bellman Continuum Workshop 1988, June 13-14, Sophia Antipolis, France) (A. Blaquière, ed.), Springer Verlag, 457-467.

Dubois D., Prade H. (1989b) Processing fuzzy temporal knowledge. IEEE Trans. on Systems, Man and Cybernetics, 19(4), 729-744.

Dubois D., Prade H. (1989c) A typology of fuzzy "if... then..." rules. Proc. of the 3rd IFSA Congress, University of Washington, Seattle, Aug. 6-11, 782-785.

Fonck P. (1990) Building influence networks in the framework of possibility theory. Proc. of the RP2 First Workshop DRUMS (Defeasible Reasoning and Uncertainty Management Systems), Albi, France, April 26-28, 1990.

Kruse R., Schwecke E. (1988) Fuzzy reasoning in a multi-dimensional space of hypotheses. Proc. of 7th Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS'88), San Francisco State University, June 8-10, 147-151.

Kruse R., Schwecke E. (1989) On the treatment of cyclic dependencies in causal networks. Proc. of the 3rd IFSA Congress, University of Washington, Seattle, USA, Aug. 6-11, 416-419.

Kruse R., Schwecke E. (1990) Fuzzy reasoning in a multidimensional space of hypotheses. Int. J. of Approximate Reasoning, 4, 47-68.

Lauritzen S.L., Spiegelhalter D.J. (1988) Local computations with probabilities on graphical structures and their application to expert systems (with discussion). J. of the Royal Statistical Society, B-50, 157-224.

Nilsson N.J. (1986) Probabilistic logic. Artificial Intelligence, 28, 71-87.

Pearl J. (1986) Fusion, propagation and structuring in belief networks. Artificial Intelligence, 29, 241-288.

Shafer G. (1976) A Mathematical Theory of Evidence. Princeton University Press.

Shafer G., Shenoy P.P. (1988) Local computation in hypertrees. Working Paper n° 201, School of Business, Univ. of Kansas, USA.

Shafer G., Shenoy P.P. (1990) An axiomatic framework for Bayesian and belief-function propagation. In : Uncertainty in Artificial Intelligence, Vol. 4 (R.D. Shachter et al., eds.), North-Holland, Amsterdam, 119-198.

Shafer G., Shenoy P.P., Mellouli K. (1987) Propagating belief functions in qualitative Markov trees. Int. J. Approximate Reasoning, 1, 349-400.

Williams P.M. (1990) An interpretation of Shenoy and Shafer's axioms for local computation. Int. J. of Approximate Reasoning, 4, 225-232.

Yager R.R. (1989) Some extensions of constraint propagation of label sets. Int. J. of Approximate Reasoning, 3, 417-435.

Yager R.R., Larsen H.L. (1990) On discovering potential inconsistencies in validating uncertain knowledge bases by reflecting on the input. In Tech. Report #MII-1001, Machine Intelligence Institute, Iona College, New Rochelle, N.Y..

Zadeh L.A. (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3-28.

Zadeh L.A. (1979) A theory of approximate reasoning. In : Machine Intelligence, Vol. 9 (J.E. Hayes, D. Mikulich, eds.), Elsevier, Amsterdam, 149-194.