Handling Uncertain Knowledge in an ATMS using Possibilistic Logic [1]

Didier DUBOIS, Jérôme LANG, Henri PRADE
Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier,
118 route de Narbonne - 31062 TOULOUSE Cedex (FRANCE)

ABSTRACT

Possibilistic ATMS are Truth Maintenance Systems oriented towards hypothetical reasoning where *both* assumptions and justifications can receive an uncertainty weight. Uncertainty is represented in the framework of possibility theory. In possibilistic logic uncertain clauses are handled as such and then in possibilistic ATMS the management of uncertainty is not separated from the other capabilities of the ATMS. The main interest of a possibilistic ATMS is to take advantage of the uncertainty pervading the available knowledge and rank-order environments in which a given statement is true. Possibilistic ATMS offers a solution to the revision of uncertain knowledge bases by deleting the least certain piece(s) of knowledge causing the inconsistency.

1. INTRODUCTION

ATMS ([3,4] ; see also [15]) are truth maintenance systems oriented towards hypothetical reasoning since they are able to determine under which set of assumption(s) a given proposition is true (this set is called the "label" of the proposition). In this paper we introduce an extension of the ATMS, called "possibilistic ATMS" (or Π-ATMS for short), where the management of uncertainty is integrated inside the basic capabilities of the ATMS. Uncertainty pervading justifications or grading assumptions is represented in the framework of possibility and necessity measures [16], [7]. These measures agree with the ordinal nature of what we wish to represent (it enables us to distinguish between what is plausible and what is less plausible). The certainty of each granule in the knowledge base (represented by a clause in possibilistic logic [5], [6]) is estimated under the form of a lower bound of a necessity measure. This uncertainty in the deduction process is propagated by means of an extended resolution principle. Uncertainty degrees are then naturally attached to the configurations of assumptions in which a given proposition is true ; one can also evaluate to what degree a given configuration of assumptions is inconsistent or compute the more or less certain consequences of a configuration of assumptions. The approach enables us to handle disjunctions and negations of assumptions without particular problem. Moreover, by rank-ordering configurations according to the degrees attached to them, Π-ATMS provides a way of limiting combinatorial explosion when using ATMS in practice.

We present the basic definitions and results of possibilistic logic first. In Section 3 we give the basic definitions and functionalities of the Π-ATMS and illustrate them on an elementary example. Then Section 4 extends the notions of interpretation and extension to the possibilistic case, and show their possible use for the revision of uncertain knowledge bases. Lastly in Section 5 the proposed approach is compared with existing uncertainty-handling ATMS's.

[1] This work has been supported by the European ESPRIT Basic Research Action number 3085 entitled Defeasible Reasoning and Uncertainty Management Systems (DRUMS). Support from THOMSON-SIMSA-CINTRA is also gratefully acknowledged.

2. POSSIBILISTIC LOGIC

Possibilistic logic [5,6,8] is an extension of classical logic where one manipulates propositional or first-order closed formulas weighted by lower bounds of possibility or necessity degrees which belong to [0,1]. A *necessity measure* N is the dual of a *possibility measure* Π, i.e. $N(\varphi) = 1 - \Pi(\neg\varphi)$. Here only a fragment of possibilistic logic is considered, namely formulas weighted by lower bounds of necessity degrees. A necessity measure N satisfies the following axioms :

(i) $N(\perp) = 0$; (ii) $N(T) = 1$; (iii) $\forall\varphi, \psi$ propositional formulas, $N(\varphi \wedge \psi) = \min(N(\varphi), N(\psi))$ where $\perp$ and T denote respectively contradiction and tautology. We emphasize that we only have $N(\varphi \vee \psi) \geq \max(N(\varphi),N(\psi))$ in the general case. The dual Π is a possibility measure in the sense of Zadeh ([16], [7]). We adopt the following conventions :

- $N(\varphi) = 1$ means that, given the available knowledge, φ is certainly true ; conversely, if φ is said to be true we can consider φ as certain.
- $1 > N(\varphi) > 0$ that φ is somewhat certain and $\neg\varphi$ not certain at all (since the axioms imply that $\forall\varphi, \min(N(\varphi),N(\neg\varphi)) = 0$.
- $N(\varphi) = N(\neg\varphi) = 0$ models total ignorance ; it expresses that, from the available knowledge, nothing enables us to say if p is more certainly true than false (or the converse).

Possibilistic logic is well-adapted to the representation of states of incomplete knowledge, since we can distinguish between the complete lack of certainty in the falsity of a proposition φ ($N(\neg\varphi) = 0$) and the total certainty that p is true ($N(\varphi) = 1$). $N(\varphi) = 1$ entails $N(\neg\varphi) = 0$ but the converse is false. It contrasts with probability measures which are such that $\mathrm{Prob}(\varphi) = 1 \Leftrightarrow \mathrm{Prob}(\neg\varphi) = 0$. Possibilistic logic also contrasts with usual multiple-valued logics which are fully truth-functional and deal with fuzzy propositions (i.e. with vague predicates), while necessity degrees apply to standard propositions and are compositional only with respect to conjunction.

A *possibilistic formula* is a couple $(\varphi\ \alpha)$ where φ is a propositional formula and α a valuation, taken as a lower bound of its necessity measure. Thus, in the following we write $(\varphi\ \alpha)$ as soon as the inequality $N(\varphi) \geq \alpha$ is known. If $\varphi = \wedge_{i\in I} (c_i)$ then $(\varphi\ \alpha)$ is equivalent to $\{(c_i\ \alpha), i \in I\}$, thus it is sufficient to work with possibilistic clauses $(c\ \alpha)$.

Resolution has been extended to possibilistic logic [5,6]. The classical rule for propositional clauses is generalized by

$$(c\ \alpha), (c'\ \beta) \vdash (\mathrm{Res}(c,c')\ \min(\alpha,\beta))$$

where Res(c,c') is a classical resolvent of the classical clauses c and c'. See [5,6] for theorem proving issues with resolution (using extended refutation).

Semantics has been defined for necessity-valued formulas [8], by grading the compatibility between (classical) interpretations and the uncertain knowledge base. This leads to define the *inconsistency degree* Inc ($\mathcal{H}$) of an uncertain knowledge base, being the complement to 1 of the maximal compatibility degree of an interpretation with $\mathcal{H}$; it is shown [8] that Inc($\mathcal{H}$) is equal to the weight of the maximal empty clause deducible from $\mathcal{H}$ by resolution. Finally we say that $\Sigma \models \mathcal{F}$ if and only if any interpretation is at least as compatible with $\mathcal{F}$ as with Σ.

As it was pointed out in [8], the weighted clause $(\neg p \vee q\ \alpha)$ is semantically equivalent to the weighted clause $(q\ \min(\alpha,v(p)))$ where v(p) is the truth value of p, i.e. $v(p) = 1$ if p is true and $v(p) = 0$ if p is false. The equivalence between $(\neg p \vee \neg s \vee q\ \alpha)$ and $(\neg p \vee q\ \min(\alpha,v(s)))$ expresses that the rule "if p and s are true then q is certain (at least) to the degree α" means that "in an environment where s is true, if p is true then q is certain (at least) to the degree α" if we decide to consider s as an assumption. This equivalence leads to the following modified presentation of the resolution rule :

$$(\neg p \vee q\ \ \min(\alpha,v(s))),\ (p \vee r\ \ \min(\beta,v(t))) \vdash (q \vee r\ \ \min(\alpha,v(s),\beta,v(t))).$$

This enables us to express that if the clause $\neg p \vee q$ is certain to the degree α in an environment where s is true and if the clause $p \vee r$ is certain to the degree β in an environment where t is true, then the resolvent clause $q \vee r$ is certain to the degree $\min(\alpha,\beta)$ in an environment where s *and* t are true (since $v(s \wedge t) = \min(v(s),v(t))$). This way of giving to some literals the assumption status by transferring them into the weight part of the clause leads us to define in Sections 3 and 4 a possibilistic ATMS and to develop the associated basic procedures.

Let us give now an example : let Σ be the following set of possibilistic clauses, concerning a given meeting about which we would like to make some predictions. The proposition Albert means that Albert will come to a given meeting (and so on for Betty, etc.).

C1 (¬Albert ∨ ¬Betty ∨ ¬Chris 0.7)

"it is rather certain that Albert, Betty and Chris will not come together "

C2 (¬David ∨ Chris 0.4)

C3 (¬Betty ∨ begins-late 0.3)

C4 (¬begins-late∨ ¬Chris 0.3)

C5 (¬begins-late∨ ¬quiet 0.5)

C6 (¬Albert ∨ ¬David ∨ ¬quiet 0.6)

C7 (¬Albert ∨ ¬Betty ∨ ¬David ∨ ¬quiet 0.9)

C8 (Betty 0.2)

C9 (¬Albert ∨ ¬Eva ∨ productive 0.5)

C10 (¬Albert ∨ ¬Betty ∨ Chris ∨ ¬Eva ∨ productive 0.8)

By resolution we can for example infer (¬Albert ∨ ¬Betty ∨ ¬David 0.4) from C1 and C2. It can be checked that Σ is consistent ; from Σ we can infer that N(¬David) ≥ 0.2, i.e. it is weakly certain that David will not come to the meeting. Clearly, the weights in the above example reflect an ordering of the propositions, in terms of certainty, and are used only as landmarks.

3. BASIC PRINCIPLES AND DEFINITIONS OF A POSSIBILISTIC A.T.M.S.

Classical ATMS require that the clauses contained inside the knowledge base are certain. We may wish to handle more or less uncertain information without losing the capacities of the ATMS. The basic principle of the Π-ATMS is to associate to each clause a weight α which is a lower bound of its necessity degree. Assumptions may also be valued, i.e. the user or the inference engine may decide at any time to believe an assumption with a certainty degree that he/she will give. A Π-ATMS is able to answer the following questions :

(i) Under what configuration of the assumptions (i.e., what assumptions shall we consider as true, and with what certainty degrees) is the datum d certain to the degree α ?

(ii) What is the inconsistency degree of a given configuration of assumptions ?

(iii) In a given configuration of assumptions, to what degree is each datum certain ?

The kind of classical ATMS we extend here is Cayrol and Tayrac's [1] generalized ATMS, where each piece of information is represented by a general propositional clause, which enables :

- a uniform representation for all pieces of knowledge (no differenciated storage and treatment between justifications and disjunctions of assumptions).
- the capability of handling negated assumptions as assumptions, i.e. environments and nogoods may contain negations of assumptions.
- a simple and uniform algorithm for the computation of labels, based on a restricted resolution .

The classical notions of environments, labels, nogoods, contexts for ATMS are now extended to Π-ATMS.

Environments

Let Σ be a set of uncertain clauses. Let E be a set of assumptions. Then :

-) [E α] is an *environment* of the datum d if and only if $N(d) \geq \alpha$ is a logical consequence of E ∪ Σ, where the assumptions of E are considered as certainly true (the certainty degree of the associated clauses is 1) ;

-) [E α] is an *α-environment* of d if and only if [E α] is an environment of d and if $\forall \alpha' > \alpha$, [E α'] is not an environment of d (α is maximal) ;

-) [E α] is an α-contradictory environment, or an *α-nogood* if and only if E ∪ Σ is α-inconsistent (i.e. E ∪ $\Sigma \models (\bot\ \alpha)$), with α maximal. We shall use the notation *$nogood_\alpha E$*. The α-nogood [E α] is said to be minimal if there is no β-nogood [E' β]

such that $E' \supset E$ and $\alpha \leq \beta$.

Labels

In order to define the label of a datum d, we consider only non-weighted assumptions (i.e. they will have the implicit weight 1). It can be shown that it is useless to weight the assumptions inside the labels (which also holds for the nogood base). The label of d, $L(d) = \{[E_i\ \alpha_i], i \in I\}$ is the only fuzzy subset of the set of the environments for which the following properties hold :

- *(weak) consistency* : $\forall [E_i\ \alpha_i] \in L(d)$, *$E_i \cup \Sigma$ is β-inconsistent, with $\beta < \alpha_i$* (i.e. $Inc(E_i \cup \Sigma) < \alpha_i$ in the sense of Section 2, the certainty degree associated to the E_i's being 1) ; it guarantees that either E_i is consistent (i.e. $\beta = 0$), or its inconsistency degree is anyway strictly less than the certainty with which d can be deduced from $E_i \cup \Sigma$ (i.e. we are sure to use a consistent sub-base of $E_i \cup \Sigma$ to deduce d, see [8]).
- *soundness* : L(d) is sound if and only if $\forall\ [E_i\ \alpha_i] \in L$ we have $E_i \cup \Sigma \models (d\ \alpha_i)$ where $\models$ has been defined in Section 2 ; i.e., L(d) contains only environments of d.
- *completeness* : L(d) is complete if and only if for every E' such that $E' \cup \Sigma \models (d\ \alpha')$ then $\exists$ i $\in$ I such that $E_i \subset E'$ and $\alpha_i \geq \alpha'$. I.e., all minimal α-environments of d are present in L(d).
- *minimality* : L(d) is minimal if and only if it does not contain two environments $(E_1\ \alpha_1)$ and $(E_2\ \alpha_2)$ such that $E_1 \subset E_2$ and $\alpha_1 \geq \alpha_2$. It means that L(d) only contains the most specific α-environments of d (i.e. all their assumptions are useful).

We emphasize that ranking environments according to their weight in the label of each datum provides a way for limiting the consequences of combinatorial explosion (problem already pointed out by Provan [13] and Raiman [14]) : indeed when a label contains too many environments, the Π-ATMS can help the user by giving the environments with the greatest weight(s) only.

Example (continued) : we may consider the literals Albert, Betty, Chris, David and Eva as assumptions (abbreviated respectively by A, B, C, D, E), i.e. we wish to be able to fix the set of persons attending the meeting, in order to get it quiet, or productive, etc. In this case, the minimal nogoods are : $nogood_{0.7}\{A,B,C\}$; $nogood_{0.4}\{A,B,D\}$; $nogood_{0.4}\{D,\neg C\}$; $nogood_{0.3}\{B,C\}$; $nogood_{0.3}\{B,D\}$; $nogood_{0.2}\{\neg B\}$; $nogood_{0.2}\{C\}$; $nogood_{0.2}\{D\}$.

We now give examples of labels.

The label of *productive* is $\{\{A, E\}_{0.5}, \{A, B, \neg C, E\}_{0.8}\}$. The notation means that {A,E} is a 0.5-environment, etc. The label of *quiet* is { }, i.e. no configuration of assumptions enables to deduce that the meeting will be quiet with a strictly positive certainty degree.

The label of *¬quiet* is $\{\{\}_{0.2}, \{B\}_{0.3}, \{A,D\}_{0.6}, \{A,B,D\}_{0.9}\}$; let us make a few remarks :

- the presence of the empty 0.2 environment $\{\}_{0.2}$ in the label expresses that if no assumption is made, then we can already deduce *(¬quiet 0.2)*.
- the 0.9-environment $\{A,B,D\}_{0.9}$ contains the 0.3-nogood $nogood_{0.3}\{B,D\}$; in a classical ATMS, such an environment would have been inhibited, since it would have been inconsistent; however the label is weakly consistent. This example show us that there is no equivalence between consistency and weak consistency (we have only consistency $\Rightarrow$ weak consistency).
- the environment {A,B,D} contains the environment {A,D} and nevertheless $\{A,D\}_{0.6}$ and $\{A,B,D\}_{0.9}$ are both in the label of *¬quiet* without violating the minimality of the label, because $\{A,D\}_{0.6}$ does not subsume $\{A,B,D\}_{0.9}$ whereas {A,D} subsumes {A,B,D} in classical logic and thus only {A,D} would be kept in the label of *¬quiet* in a classical ATMS. Besides, the empty environment { } subsuming any other one, it would remain alone in the label of *¬quiet* in a classical ATMS.

Consequently, due to the presence of uncertainty weights, the ATMS becomes more selective, and at the same time supplies richer symbolic information.

We may wish to add the clause C11 : (Chris $\vee$ David 1), i.e. we are absolutely certain that either Betty or David will come ; then the system Σ becomes 0.2-inconsistent, since ($\perp$ 0.2) is derived by resolution from {C2, C3, C4, C8, C11}. In consequence the new nogood base is : $nogood_1\{\neg C,\neg D\}$; $nogood_{0.7}\{A,B,C\}$; $nogood_{0.7}\{A,B,\neg D\}$; $nogood_{0.4}\{A,B\}$; $nogood_{0.4}\{\neg C\}$; $nogood_{0.3}\{B\}$; $nogood_{0.2}\{\}$.

The fact that the empty environment is a 0.2-nogood expresses that Σ is 0.2-inconsistent. Consequently, all environments of a label having a membership degree less or equal to 0.2 are inhibited. In fact, the labels do not take into account the clause (Betty 0.2) which is "responsible" of the 0.2-inconsistency, since it is the least reliable clause involved in it. The label of *¬quiet* is now $\{\{A,D\}_{0.6}, \{A,\neg C\}_{0.6}, \{A,B,D\}_{0.9}, \{A,B,\neg C\}_{0.9}\}$.

Contexts

To extend the notion of context, we now consider *weighted assumptions*, which are pairs (H α) where H is an assumption and $\alpha \in [0,1]$ is the certainty degree assigned to H.

The *context* associated with the set of weighted assumptions $\mathcal{E}$ is the set of all pairs (d, $val_{\mathcal{E}}(d)$), where d is a datum or an assumption, and $val_{\mathcal{E}}(d) = \sup\{\alpha, \mathcal{E} \cup \Sigma \models (d\ \alpha)\}$. Let us now give the following theorem :

Let $\mathcal{E}$ be a set of valued assumptions. Let d be a datum ; it can be shown that $\mathcal{E} \cup \Sigma \models (d\ \alpha)$ in possibilistic logic if and only if $\exists[E_i\ \alpha_i] \in$ Label (d) , $E_i = \{H_{i,1}, H_{i,2}, \ldots, H_{i,n}\}$ such that

(i) $\mathcal{E}^* \supset E_i$ where $\mathcal{E}^*$ is the classical set of assumptions obtained from $\mathcal{E}$ by ignoring the weights.

(ii) $\alpha \leq \min(\alpha_i,\beta_1,\beta_2,\ldots,\beta_n)$ where $\beta_1,\beta_2,\ldots,\beta_n$ are the weights attached to $H_{i,1}$, $H_{i,2},\ldots, H_{i,n}$ in $\mathcal{E}$.

(iii) $\alpha > \mathrm{Inc}(\mathcal{E} \cup \Sigma)$.

The proof of the theorem is obvious if using the results of classical ATMS. This theorem gives an immediate algorithm to compute contexts, given the label of every datum :

```
α_max ← 0;
For every [Ei αi] ∈ Label (d) do
    if ℰ* ⊃ Ei
    then α_max ← max(α_max,min(αi,β1,β2,....,βn) where
    β1,....,βn are the weights attached to Hi,1, ... ,Hi,n in ℰ.
end
```

Intuitively, for each environment of the label of d included in $\mathcal{E}^*$, the algorithm computes the degree of certainty with which this environment entails d. This degree depends upon the justifications used in deriving d (via α_i) and the weights of the assumptions in this environment.

Example (continued) : let us consider the initial system Σ = {C1, ..., C10} and let $\mathcal{E}$ be the set of weighted assumptions $\mathcal{E}$ = {(A 0.8), (¬C 0.9), (D 0.5), (E 0.7)} ; then the inconsistency degree of $\mathcal{E} \cup \Sigma$ is 0.4 and the context of $\mathcal{E}$ is {(A 0.8), (¬C 0.9), (D 0.5), (E 0.7), (¬quiet 0.5), (profitable 0.5)}.

4. INTERPRETATIONS AND EXTENSIONS IN A POSSIBILISTIC ATMS

In a Π-ATMS, like in a classical ATMS, an interpretation is a maximal consistent set of assumptions. The complementary (in the whole set of assumptions) of an interpretation is called a candidate. A candidate is a (minimal) set of assumptions whose deletion makes the knowledge base consistent. An extension is the context associated with an interpretation.

The presence of certainty weights inside the possibilistic ATMS will enable us to rank interpretations, which is useful in the case we get a lot of interpretations and extensions. When the user asks the ATMS for a datum d, the ATMS will search for the truth or the falsity of d in the best extension(s), i.e. the least inconsistent one(s). We now show how we can use possibilistic interpretations for the revision of uncertain knowledge bases.

The problem of the revision of uncertain knowledge bases consists in finding, from a partially inconsistent knowledge base, the "closest" consistent sub-base by taking out some formulae. Since less certain formulae are less "entrenched" in the knowledge base than more certain ones, a formula must never be taken out if a less certain one may be taken out instead. This work can be done with a possibilistic ATMS, by controlling each formula by means of a specific assumption. We say that a justification is controlled by an assumption if the justification contains the assumption in its antecedent part. Let us take an example : let $\mathcal{F}$ be the set of uncertain formulae $\{(p\ 0.6), (\neg p \vee q\ 0.9), (\neg p \vee \neg q\ 0.4), (\neg q \vee \neg r\ 0.3), (\neg p \vee \neg r\ 0.2), (r \vee (\neg p \wedge s)\ 0.1)\}$. Let H_1, H_2,..., H_6 be the assumptions controlling each formula. Note that H_6 controls two clauses. The obtained set of justifications Σ is :

$H_1 \rightarrow p$	(0.6)	$H_2 \rightarrow \neg p \vee q$	(0.9)	$H_3 \rightarrow \neg p \vee \neg q$	(0.4)
$H_4 \rightarrow \neg q \vee \neg r$	(0.3)	$H_5 \rightarrow \neg p \vee \neg r$	(0.2)		
$H_6 \rightarrow r \vee s$	(0.1)	$H_6 \rightarrow r \vee \neg p$	(0.1)		

When H_i is true (resp. false), the formula (i.e. one or more clauses) controlled by H_i is taken as valid (resp. inhibited). The formulae taken out of the base will be those controlled by the assumptions of the "best" candidate, and the remaining formulae will be those controlled by the assumptions of the "best" interpretation. To compute the interpretations, we compute their complements, the candidates, by selecting one assumption out of each minimal nogood [4]. Since the minimal nogoods are $\{H_1,H_2,H_3\}_{0.4}$, $\{H_1,H_2,H_4,H_6\}_{0.1}$ and $\{H_1,H_5,H_6\}_{0.1}$, we have five candidates in our example, i.e. we have 5 solutions for taking out formulae from the knowledge base to make it consistent. These candidates are $\{H_1\}$, $\{H_3,H_4,H_5\}$, $\{H_2,H_5\}$, $\{H_2,H_6\}$ and $\{H_3,H_6\}$. To make a choice among these candidates we must introduce a total ordering among them. Namely we associate to a candidate C the multiset, denoted by list (C), of the weights of the justifications controlled by the elements of C. In our example, to $\{H_1\}$ we associate $\{0.6\}$, to $\{H_3,H_4,H_5\}$ we associate $\{0.4,0.3,0.2\}$, etc. Then, if C_1 and C_2 are two candidates, we say that $C_1 \leq C_2$ if and only if

(i) $C_1 = C_2 = \emptyset$ or (ii) $\sup(\text{list}(C_1)) < \sup(\text{list}(C_2))$

or (iii) $\sup(\text{list}(C_1)) = \sup(\text{list}(C_2))$ and $C_1 - \{\sup(\text{list}(C_1))\} \leq C_2 - \{\sup(\text{list}(C_2))\}$;

$\leq$ is a total ordering among the candidates ; the minimal elements relatively to $\leq$ are called lexicographically optimal candidates, and their complement in the whole set of assumptions lexicographically optimal interpretations. If C_1 and C_2 are two candidates such that $C_1 \leq C_2$ then C_2 is a better candidate to deletion than C_1, since either the most reliable justification controlled by C_2 is strictly more reliable than the most reliable justification controlled by C_1, or these two most reliable justifications controlled by C_1 and C_2 have the same weight, but the second most reliable justification controlled by C_2 is strictly more reliable than C_1's, etc. In our example, there is only one lexicographically optimal candidate, which is $\{H_3,H_6\}$. The lexicographically optimal candidates can easily be computed from the nogoods. Thus the optimal revision of the knowledge base comes down to deleting the justifications controlled by H_3 and H_6.

5. POSSIBILISTIC A.T.M.S. VS. OTHER EXISTING "UNCERTAINTY - HANDLING" A.T.M.S.

De Kleer and Williams [4], in a multiple faults diagnosis problem, compute the probability of a candidate (i.e. a configuration of faults) given the outcomes of the system and assuming that faults are mutually independent. Falkenheiner [10] has incorporated Dempster-Shafer theory into Doyle's TMS ; Provan [13], D'Ambrosio [2], Laskey and Lehner [11] have independently incorporated belief functions into the ATMS ; their approaches are somewhat different but share the same basic features : first, a mass function is associated with each assumption ; then, for each datum x we compute the belief Bel(x) of x by first computing the label of x symbolically (which can be done in the classical way), then we compute Bel(x) from Label(x) ; Provan computes Bel(x) by expressing logically Label(x) under the form of a formula containing only independent sub-formulas, assuming mutual independence of assumptions. D'Ambrosio's approach does not use the full Dempster-Shafer theory ; besides, his way of propagating uncertainty is closer to a rule-based system than to a pure logical system (as he uses conditional beliefs and not beliefs of logical implications). Laskey-Lehner's and Provan's approaches both incorporate the full Dempster-Shafer theory into the ATMS, and differ essentially by their algorithm for the computation of Bel(x). See also Pearl [12] whose proposal is close to Provan's (without normalizing Dempster rule).

Our approach differs from the preceding ones on several points :

First, the ATMS with embedded belief functions enable us to rank interpretations, as they intend to do, but they cannot handle the uncertainty pervading the knowledge base if any, i.e. the justifications. They hold the justifications as classical knowledge (i.e. certain) and, putting weights on the assumptions, they deduce the weights attached to the other data. Our approach does not attach any a priori weight to each datum (assumptions or non-assumptions) but assigns a certainty degree to all clauses of the knowledge base.

Secondly, in the other approaches, truth maintenance and uncertainty management are completely separated, i.e. the label of a datum x is first computed, then Bel(x) is computed (of course, taking into account Label(x)). Our approach completely integrates the uncertainty management into the truth maintenance system, by assigning a weight to each environment of a datum. Besides, the handling of weights will be done during the computation of labels, since certainty weights are present inside the label.

Thirdly, the main purpose of introducing uncertainty into ATMS is the ability of ranking solutions (i.e. interpretations) and eliminating solutions which are too uncertain. Hence the precise values of certainty degrees are not as important as their ordering. Possibility theory offers a framework where the ordering of the uncertainty degrees is more important than their precise values, and which requires easier computations (only min and max operations are used) than the general Dempster-Shafer approach. Furthermore, the weights assigned to nogoods are always available and no extra computations are required in order to renormalize results.

Fourthly, in Provan's model, the belief of the datum x is first computed formally by rewriting the label of x into an "independent form" and then computed numerically ; if we decide to change the weight of an assumption, the formal expression of the label of x remains unchanged and the re-computation of Bel(x) is immediate ; however this is not the case if the label of x changes (which may happen each time we add a new justification) : then we have to transform again the expression of the label of x ; in the possibilistic model, both changes of weights or labels lead to almost no more computations than in De Kleer's original model.

Lastly, the other models cannot handle disjunctions of assumptions ; besides, Provan's model needs to consider the assumptions as independent as long as they are not mutually exclusive (i.e. not containing any nogood). Possibilistic A.T.M.S. does not require this assumption, and can handle disjunctions of assumptions, as well as negated assumptions.

6. CONCLUSION

A possibilistic ATMS, based on a straightforward generalization to possibilistic logic of the works of Cayrol and Tayrac on CAT-correct resolution [1], is currently being implemented.

Possibilistic ATMSs enable a joint handling of assumptions and uncertainty relative to a knowledge base, in the framework of possibilistic logic. The fact that the fragment of possibilistic logic, which is used here remains on many points close to classical logic, facilitates the extension of efficient procedures, such as those based on CAT-correct resolution for the computation of labels, nogoods and contexts, by ordering the environments of a datum according to the certainty with which it can be deduced from each of them.

Possibilistic ATMS can be applied to truth maintenance problems in presence of uncertainty, to default reasoning, to fault diagnosis [9], or to generation of (the most plausible) explanations.

REFERENCES

[1] Cayrol M., Tayrac P. ARC : un ATMS basé sur la résolution CAT-correcte. Revue d'Intelligence Artificielle (Hermès, Paris), 3(3), 1989, 19–39.

[2] D'Ambrosio B.D. A hybrid approach to reasoning under uncertainty. In : Uncertainty in Artificial Intelligence 3 (L.N. Kanal, T.S. Levitt, J.F. Lemmer, eds.), North-Holland, 1989, 267–283.

[3] De Kleer J. "An assumption-based TMS" & "Extending the ATMS". Artificial Intelligence, 28, 1986, 127–162 & 163–196.

[4] De Kleer J., Williams B. Diagnosing multiple faults. Artif. Intell., 32, 1987, 97–130.

[5] Dubois D., Prade H. Necessity measures and the resolution principle. IEEE Trans. on Systems, Man and Cybernetics, 17, 1987, 474–478.

[6] Dubois D., Lang J., Prade H. Theorem proving under uncertainty : a possibility theory-based approach. Proc. of 10th Inter. Joint Conf. on Artificial Intelligence, Milano, Italy, 1987, 984–986.

[7] Dubois D., Prade H. (with the collaboration of Farreny H., Martin-Clouaire R., Testemale C.) Possibility Theory – An Approach to Computerized Processing of Uncertainty. Plenum Press, New York, 1988.

[8] Dubois D., Lang J., Prade H. Automated reasoning using possibilistic logic : semantics, belief revision and variable certainty weights. Proc. of the 5th Workshop on Uncertainty in Artificial Intelligence, Windsor, Ontario, 1989, 81–87.

[9] Dubois D., Lang J., Prade H. Gestion d'hypothèses en logique possibiliste – Un exemple d'application au diagnostic. Proc of 10th Inter. Workshop Expert Systems & Their Applications, Avignon, France, 1990, 299–313.

[10] Falkenheiner B. Towards a general-purpose belief maintenance system, Uncertainty in Artificial Intelligence 2 (J.F. Lemmer, L.N. Kanal, eds.), North-Holland, 1988, 125–132.

[11] Laskey K.B., Lehner P.E. Belief maintenance ; an integrated approach to uncertainty mana-gement. Proc. 7th Nat. Conf. of Amer. for Artificial Intelligence, Saint Paul, Minn., 1988, 210–214.

[12] Pearl J. Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference. Morgan Kaufmann Publ. Inc., San Mateo, Ca., 1988.

[13] Provan G.M. An analysis of ATMS-based techniques for computing Dempster-Shafer belief functions. Proc of the 9th Inter. Joint Conf. Artificial Intelligence, Detroit, 1989, 1115–1120.

[14] Raiman O. Two heuristic integrating probabilities and logic : a preliminary report on parsimonious search. Proc. of the Inter. Conf. on Computational Intelligence, Milano, Italy, 1989.

[15] Smith B., Kelleher G. (Eds) Reason Maintenance Systems and Their Applications. Ellis Horwood Ltd, Chichester, 1988.

[16] Zadeh L.A. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 1978, 3–28.