DEFAULT AND INEXACT REASONING WITH POSSIBILITY DEGREES

Henri Farreny Henri Prade

Laboratoire Langages & Systèmes Informatiques
ENSEEIHT & Université Paul Sabatier
118 route de Narbonne
31062 Toulouse Cedex. France

Abstract : This paper proposes a new approach, based on possibility theory, for dealing with uncertain facts and default rules. The chaining of certain or uncertain facts and rules is discussed in detail. The modus ponens, the modus tollens and the resolution principle in the propositional case, are generalized and patterns of plausible reasoning are recovered. Only min and max operations are used for computing the possibility degrees corresponding to the different alternatives. The end of the paper is devoted to the problem of the combination of results obtained from different sources. Particularly it is shown that the combination may be completely unsuitable when some uncertain conclusions can be obtained directly through a specific rule and indirectly through a chain of inferences. This observation has important consequences when building inference systems for the exploitation of uncertain knowledge bases.

Key Words : default reasoning ; inexact reasoning; uncertainty ; combination of pieces of information ; possibility theory.

1 - Introduction

Rich [9] has suggested to see default reasoning as a likelihood calculus where confidence in a rule is expressed by means of a MYCIN-like certainty factor [11]. More recently Ginsberg [3] has proposed an approach to non-monotonic reasoning in the same spirit, applying available statistical methods (especially Dempster's rule of combination [10]) to ranges of possible values of probabilities representing what is known about these confidences.

In this paper, we propose an approach to inexact and default reasoning based on possibility theory [12]. Possibility theory seems to be well-suited for representing uncertain knowledge ; moreover because we are mainly using max and min operations, the numbers used for the estimation of uncertainty are regarded only as rough tendency indications ; what is meaningful in practice is the ordering among these numbers, not their exact value. This is in agreement with the fact that precise estimates of the amounts of uncertainty are almost impossible to obtain in practice. The proposed approach is quite simple and not sensitive to slight variations of the estimates. Its basic features have been recently given in [2], [6]. In the following the approach is presented in greater detail. Moreover some control issues which are peculiar to the management of uncertain knowledge are discussed. Particularly, in case of conflicting results relative to a same matter, obtained from different chains of inferences, we may have to be more confident in some of these results than in the others, rather than combining together positive and negative conclusions in a blind manner.

We first deal with knowledge representation issues, then with the chaining of rules before discussing combination problems.

2 - Knowledge representation

2.1 - Facts :

Let p be a proposition ; provided that p is non-fuzzy (i.e. p does not contain any vague predicate), the excluded-middle and the non-contradiction laws hold, thus p and $\neg p$ (the negation of p) can be regarded as mutually exclusive alternatives. Then a so-called possibility distribution [12] π can be attached to the set $\{p, \neg p\}$; namely, two numbers, belonging to the real interval [0,1], $\pi(p)$ and $\pi(\neg p)$ are assessed, which supposedly grade the possibility that p is true and the possibility that $\neg p$ is true (i.e. p is false) respectively. The normalization condition

$$\max(\pi(p), \pi(\neg p)) = 1 \qquad (1)$$

must hold ; the constraint (1) departs from probability theory where we should have $\text{prob}(p) + \text{prob}(\neg p) = 1$ (see [12] or [5] for a presentation of the differences between possibility and probability) ; (1) expresses that at least one of the alternatives must be completely possible, since the alternatives are mutually exclusive and cover all the possibilities. If $\pi(p) = 1$ and $\pi(\neg p) = 0$, p is regarded as certainly true since it is impossible that p is false ; similarly $\pi(p) = 0$ and $\pi(\neg p) = 1$ corresponds to p false. When $\pi(p) = \pi(\neg p) = 1$, we don't know if p is true or false, both hypotheses being equally possible. If $\pi(p) = 1$ and $\pi(\neg p) = \lambda$, with $0 < \lambda < 1$ we are not sure that p is true, but it is more possible that p is true than p is false ; the smaller λ, the greater our certainty (or if we prefer the stronger our belief) that p is true ; total certainty would correspond to $\lambda = 0$.

The quantity

$$n(p) = 1 - \pi(\neg p) \qquad (2)$$

can be viewed as a measure of necessity since definition (2) expresses that the necessity of p corresponds to the impossibility of $\neg p$, which is in agreement with our intuition. As recalled in [5], possibility and necessity measures are respectively particular cases of plausibility and belief functions studied by Shafer [10].

Moreover, rather than precisely knowing the values of $\pi(p)$ and of $\pi(\neg p)$, we can only know that they are restricted to some subinterval of [0,1] ; for instance $\pi(p) \in [0,\lambda]$ means that p is true is regarded as possible at most at the degree λ ; in case of total ignorance, we only have $\pi(p) \in [0,1]$; however in any case, the constraint (1) must hold. Thus in the proposed approach a difference is made between total ignorance where $\pi(p)$ and $\pi(\neg p)$ remain completely unknown (this situation will be denoted in the following by $\pi(p) = ? = \pi(\neg p)$), and the indetermination of the truth or the falsity of p due to the fact that $\pi(p) = 1 = \pi(\neg p)$, where it is <u>known</u> that p and $\neg p$ are equally possible. Thus, to each proposition p is attached two numerical or interval-valued quantities (which are independent up to the constraint (1)) rather than only one as in a probabilistic approach. In this framework hard facts as well as uncertain facts can be modeled.

2.2 - Rules :

A rule 'if p, then q' will be represented by means of a conditional possibility distribution defined by the quantities $\pi(q|p)$ and $\pi(\neg q|p)$ which respectively estimate the possibility of having q true when p is true and the possibility of having q false when p is true. We must have

$$\max(\pi(q|p), \pi(\neg q|p)) = 1 \qquad (3)$$

When the rule is certain, we have $\pi(q|p) = 1$ and $\pi(\neg q|p) = 0$; when the rule is uncertain, i.e. the rule is such that "generally if p, then q" or "if p, probably q", we have $\pi(\neg q|p) > 0$. The smaller $\pi(\neg q|p)$, the more certain the rule "if p, then q". A rule of the form "if p, then $\neg$q" will be represented by $\pi(\neg q|p) = 1$ and $\pi(q|p) \geq 0$ which assesses the degree of possibility that the rule fails. Note that $\pi(q|\neg p)$ and $\pi(\neg q|\neg p)$ remain completely unknown since, viewing the rule as the incomplete specification of a mapping, the rule only states that the image by this mapping of an argument which satisfies p should rather satisfy q than $\neg$q (if $\pi(q|p) = 1$ and $\pi(\neg q|p) < 1$) and gives no information at all on the images of the arguments which satisfy $\neg$p.

2.3 - Relation with MYCIN

As pointed out in [5], the knowledge representation technique, based on possibility distributions, we use here, is similar to the MYCIN one [11] based on measures of belief (MB) and on measures of disbelief (MD), graded on [0,1]. More precisely, if we state the following correspondence

$$\pi(q|p) = 1 - MD\ (q,p) \qquad (4)$$

$$\pi(\neg q|p) = 1 - MB\ (q,p) \qquad (5)$$

(4) and (5) are in agreement with the relation $MB(\neg q,p) = MD(q,p)$ which always holds in MYCIN ; then the normalization constraint (3) is satisfied in MYCIN under the form $MB(q,p) > 0 \Rightarrow MD(q,p) = 0$. Moreover, the formulae which are proposed in the following for the treatment of compound conditions are analogous, in the sense of (4)-(5), to MYCIN ones. However, the chaining and the combining operations we use, are different from the MYCIN ones, and are consistent with the possibility theory-based approach we have chosen.

3 - Chaining of a fact and a rule

3.1 - Basic machinery

The characteristic axiom[1] of a possibility measure [12]

(1) The other axioms are $\pi(\mathbb{1})=1$ and $\pi(\mathbb{0})=0$ where $\mathbb{1}$ and $\mathbb{0}$ respectively represent the ever-true and the ever-false proposition.

$$\forall\, p,\ \forall\, q,\ \pi(p \vee q) = \max(\pi(p),\ \pi(q)) \tag{6}$$

together with the relation [1]

$$\pi(p \wedge q) = \min(\pi(q|p),\ \pi(p)) \tag{7}$$

enables us to obtain from $q = (q \wedge p) \vee (q \wedge \neg p)$

$$\pi(q) = \max[\min(\pi(q|p),\ \pi(p)),\ \min(\pi(q|\neg p),\ \pi(\neg p))] \tag{8}$$

$$\pi(\neg q) = \max[\min(\pi(\neg q|p),\ \pi(p)),\ \min(\pi(\neg q|\neg p),\ \pi(\neg p))] \tag{9}$$

(8) and (9) can be written in a matrix form

$$\begin{bmatrix} \pi(q) \\ \pi(\neg q) \end{bmatrix} = \begin{bmatrix} \pi(q|p) & \pi(q|\neg p) \\ \pi(\neg q|p) & \pi(\neg q|\neg p) \end{bmatrix} \begin{bmatrix} \pi(p) \\ \pi(\neg p) \end{bmatrix} \tag{10}$$

where the matrix product is defined by analogy with the usual one, changing the sum into max operation and the product into min operation. The result of (10) is normalized as soon as (1) holds and $\pi(q|.)$ and $\pi(\neg q|.)$ satisfy (3). The analogous of (8) in probability theory is

$$\text{prob}(q) = \text{prob}(q|p)\cdot\text{prob}(p) + \text{prob}(q|\neg p)\cdot\text{prob}(\neg p).$$

(10) enables us to chain the conditional rule "if p, then q" with the fact "p"; rule and fact may be certain or uncertain.

When the possibility degrees are only known to be restricted to subintervals, max and min operations are extended in the following way

$$\max([a,b],\ [c,d]) = [\max(a,c),\ \max(b,d)] \tag{11}$$

$$\min([a,b],\ [c,d]) = [\min(a,c),\ \min(b,d)] \tag{12}$$

Note that a precise value is a particular case of subinterval : $a = [a,a]$.

3.2 - Extending the modus ponens

The matricial product (10) extends the modus ponens since it can be checked that

$$\begin{bmatrix} \pi(q) \\ \pi(\neg q) \end{bmatrix} = \begin{bmatrix} \pi(q|p) & \pi(q|\neg p) \\ \pi(\neg q|p) & \pi(\neg q|\neg p) \end{bmatrix} \begin{bmatrix} \pi(p) \\ \pi(\neg p) \end{bmatrix} = \begin{bmatrix} 1 & ? \\ 0 & ? \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

where '?' stands for any number belonging to [0,1]. Conversely the system of equation $\begin{bmatrix} 1 & ? \\ 0 & ? \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

has for unique solution $x = 1$, $y = 0$, which means that if the rule "if p, then q" is certainly true, only the information that p is true enables us to derive that q is true with this machinery. Similarly, the equation $\begin{bmatrix} 1 & a \\ 0 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ has for unique $x=0=a$ and $y=1=b$, which means that q false can be only obtained from p false, provided that we have the rule "if $\neg$p, then $\neg$q".

In case of the fact p is uncertain, i.e. $\begin{bmatrix} \pi(p) \\ \pi(\neg p) \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$ and the rule "if p, then q" is certain, we get

$$\begin{bmatrix} 1 & ? \\ 0 & ? \end{bmatrix} \begin{bmatrix} 1 \\ \lambda \end{bmatrix} = \begin{bmatrix} 1 \\ [0,\lambda] \end{bmatrix}$$

which is natural since $\pi(q|\neg p)$ and $\pi(\neg q|\neg p)$ remain unknown and depending on their values we may have $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ \lambda \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ \lambda \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$ (in the first case q is true whatever p is, in the second case p is true if and only if q is true). When the rule "if p, then q" is uncertain and the fact "p" is certain, we have

$$\begin{bmatrix} 1 & ? \\ \lambda & ? \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$$

which is agreement with our intuition.

Let us consider a very simple illustrative example. Suppose our knowledge base contains two pieces of information pertaining to the usual meeting behaviour of people :

(1) "if Bob comes, Mary comes", represented by

$$\begin{bmatrix} \pi_1(M|B) & \pi_1(M|\neg B) \\ \pi_1(\neg M|B) & \pi_1(\neg M|\neg B) \end{bmatrix} = \begin{bmatrix} 1 & ? \\ 0 & ? \end{bmatrix}$$

(2) "if Mary comes, then generally Tom comes", represented by

$$\begin{bmatrix} \pi_2(T|M) & \pi_2(T|\neg M) \\ \pi_2(\neg T|M) & \pi_2(\neg T|\neg M) \end{bmatrix} = \begin{bmatrix} 1 & ? \\ \lambda & ? \end{bmatrix}$$

and the instantiated fact

(3) "Bob comes at the next meeting", represented by

$$\begin{bmatrix} \pi_3(B) \\ \pi_3(\neg B) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

By a cascade of two matrix products we get $\begin{bmatrix} \pi(T) \\ \pi(\neg T) \end{bmatrix} = \begin{bmatrix} 1 & ? \\ \lambda & ? \end{bmatrix}\begin{bmatrix} 1 & ? \\ 0 & ? \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$ which expresses that there is only a possibility equal to λ that Tom does not come at the next meeting.

3.3 - Extending the modus tollens

Since the rule "if p, then q" is equivalent to the rule "if $\neg$q, then $\neg$p" when the rules are certain, we may postulate[2] that this equivalence still holds with uncertain rules provided that $\pi(q|p) = 1$ (i.e. when it is the rule "if p, then q" that is somewhat certain rather than the rule "if p, then $\neg$q", which itself would be equivalent to "if q, then $\neg$ p"). Then the contraposition applied to $\pi(.|.)$ yields <u>when</u> $\pi(q|p) = 1$

$$\begin{aligned} &\pi(q|p) = \pi(\neg p|\neg q) \quad \pi(q|\neg p) = \pi(\neg p|q) \\ &\pi(\neg q|p) = \pi(p|\neg q) \quad \pi(\neg q|\neg p) = \pi(p|q) \end{aligned} \tag{13}$$

(2) This may be done only in the absence of contrary evidence, since strictly speaking it may happen that the rule "if p, then q" generally holds, while the rule "if $\neg$q, then $\neg$p" does not hold generally (this situation is feasible if q is rarely false).

From $\begin{bmatrix}\pi(q|p) & \pi(q|\neg p)\\ \pi(\neg q|p) & \pi(\neg q|\neg p)\end{bmatrix} = \begin{bmatrix}1 & ?\\ 0 & ?\end{bmatrix}$, we obtain $\begin{bmatrix}\pi(p|q) & \pi(p|\neg q)\\ \pi(\neg p|q) & \pi(\neg p|\neg q)\end{bmatrix} = \begin{bmatrix}? & 0\\ ? & 1\end{bmatrix}$ whose matrix product with $\begin{bmatrix}\pi(q)\\ \pi(\neg q)\end{bmatrix} = \begin{bmatrix}0\\ 1\end{bmatrix}$ yields $\begin{bmatrix}\pi(p)\\ \pi(\neg p)\end{bmatrix} = \begin{bmatrix}0\\ 1\end{bmatrix}$, which corresponds to the modus tollens. The equivalence "if p, then q" and "if q, then p" (viewed as "if ¬p, then ¬q") will be represented by $\begin{bmatrix}\pi(q|p) & \pi(q|\neg p)\\ \pi(\neg q|p) & \pi(\neg q|\neg p)\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}\pi(p|q) & \pi(p|\neg q)\\ \pi(\neg p|q) & \pi(\neg p|\neg q)\end{bmatrix}$. A rule "if p, then q" almost reversible will be represented by $\begin{bmatrix}\pi(q|p) & \pi(q|\neg p)\\ \pi(\neg q|p) & \pi(\neg q|\neg p)\end{bmatrix} = \begin{bmatrix}1 & \lambda\\ 0 & 1\end{bmatrix}$.

3.4 - Recovering a pattern of plausible inference

Our framework enables us to capture some of the patterns of plausible reasoning considered by Pólya [4]. From the certain rule "if p, then q" and "q true" nothing can be said concerning p. Indeed, using (10) and (13) we get $\begin{bmatrix}\pi(p)\\ \pi(\neg p)\end{bmatrix} = \begin{bmatrix}? & 0\\ ? & 1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} = \begin{bmatrix}[0,1]\\ [0,1]\end{bmatrix}$. If we have the additional knowledge that "q without p is hardly credible, i.e. almost impossible", which can be translated into $\pi(q|\neg p) = \pi(\neg p|q) = \varepsilon < 1$ (which implies using (3) $\pi(p|q) = 1$), we get $\begin{bmatrix}\pi(p)\\ \pi(\neg p)\end{bmatrix} = \begin{bmatrix}1 & 0\\ \varepsilon & 1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} = \begin{bmatrix}1\\ \varepsilon\end{bmatrix}$, i.e. p is very credible since ¬p is almost impossible ($\pi(\neg p) = \varepsilon$).

3.5 - Compound conditions

In order to state the rules more precisely we have to be able to consider rules of the form "if p and q, then r" ; such a rule can be represented by means of the matrix $\begin{bmatrix}\pi(r|p\wedge q) & \pi(r|\neg p\vee\neg q)\\ \pi(\neg r|p\wedge q) & \pi(\neg r|\neg p\vee\neg q)\end{bmatrix}$, which can be chained with $\begin{bmatrix}\pi(p\wedge q)\\ \pi(\neg p\vee\neg q)\end{bmatrix}$, computable from $\begin{bmatrix}\pi(p)\\ \pi(\neg p)\end{bmatrix}$ and $\begin{bmatrix}\pi(q)\\ \pi(\neg q)\end{bmatrix}$ by the formulas

$$\pi(p\wedge q) = \min(\pi(p), \pi(q)) \tag{14}$$

$$\pi(\neg p\vee\neg q) = \max(\pi(\neg p), \pi(\neg q)) \tag{15}$$

Formula (14) holds provided that p and q are logically independent, (see [5]). Note that if $(\pi(p), \pi(\neg p))$ and $(\pi(q), \pi(\neg q))$ are normalized, $(\pi(p\wedge q), \pi(\neg p\vee\neg q))$ is also normalized. When $\pi(q)$ and $\pi(\neg q)$ are completely unknown, (14) and (15) respectively give

$$\pi(p\wedge q) \in [0,\pi(p)] \tag{16}$$

$$\pi(\neg p\vee\neg q) \in [\pi(\neg p),1] \tag{17}$$

it points out the approximation we make when we ignore q in the rule "if p and q, then r".

3.6 - Resolution principle

In proposition logic, the inference rule, named "resolution principle" produces a clause of the form $l \vee l'$ from two clauses of the form $p \vee l$ and $\neg p \vee l'$, where p, l, l' stand for any literals.

In our approach these two latter clauses can be respectively represented by the matrices

$$\begin{bmatrix} \pi(l|p) & \pi(l|\neg p) \\ \pi(\neg l|p) & \pi(\neg l|\neg p) \end{bmatrix} = \begin{bmatrix} ? & 1 \\ ? & 0 \end{bmatrix} \text{ et } \begin{bmatrix} \pi(l'|p) & \pi(l'|\neg p) \\ \pi(\neg l'|p) & \pi(\neg l'|\neg p) \end{bmatrix} = \begin{bmatrix} 1 & ? \\ 0 & ? \end{bmatrix}$$

Applying the formulas (14)-(15), we get $\begin{bmatrix} \pi(l \vee l'|p) & \pi(l \vee l'|\neg p) \\ \pi(\neg l \wedge \neg l'|p) & \pi(\neg l \wedge \neg l'|\neg p) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. Using (10), i.e. $\begin{bmatrix} \pi(l \vee l') \\ \pi(\neg l \wedge \neg l') \end{bmatrix} = \begin{bmatrix} \pi(l \vee l'|p) & \pi(l \vee l'|p) \\ \pi(\neg l \wedge \neg l'|p) & \pi(\neg l \wedge \neg l'|\neg p) \end{bmatrix} \begin{bmatrix} \pi(p) \\ \pi(\neg p) \end{bmatrix}$, we check that we obtain $\begin{bmatrix} \pi(l \vee l') \\ \pi(\neg l \wedge \neg l') \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with $\begin{bmatrix} \pi(p) \\ \pi(\neg p) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as well with $\begin{bmatrix} \pi(p) \\ \pi(\neg p) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

N.B. The application of formula (14) requires the logical independence of l and l'. However when $l = l'$, (14) trivially holds : $\pi(\neg l \wedge \neg l|p) = \pi(\neg l|p) = \min(\pi(\neg l|p), \pi(\neg l|p))$. When $l' = \neg l$ we directly get $\pi(l \vee l') = 1$ and $\pi(\neg l \wedge \neg l') = 0$ since $l \vee l' = \mathbb{1}$ and $l \wedge l' = \mathbb{0}$.

Thus the resolution principle is preserved.

4 - Combining or not combining

4.1 - Combining

Given two possibility distributions $(\pi'(p), \pi'(\neg p))$ and $(\pi''(p), \pi''(\neg p))$ pertaining to a same proposition p and issued from different sources, the simplest and perhaps the most natural way for combining them into a new possibility distribution $(\pi(p), \pi(\neg p))$ is the intersection (see [5] for instance) ; namely

$$\pi(p) = \min(\pi'(p), \pi''(p)) \tag{18}$$

$$\pi(\neg p) = \min(\pi'(\neg p), \pi''(\neg p)) \tag{19}$$

In case of a conflict between the sources (which means that for one source p is more possible than $\neg p$ and it is the contrary for the other source), the result obtained from (18)-(19) is no longer normalized in the sense of (1), which expresses the conflict. We may renormalize[3] the result by dividing it by $\max(\pi(p), \pi(\neg p))$, if we want to use the result of the combination in a new chain of inferences in spite of the conflict. However this is impossible in case of a

(3) See [5] for a discussion on the renormalization in case of conflict in the more general framework of Shafer's approach.

total contradiction, i.e. when $\begin{bmatrix}\pi'(p)\\ \pi'(\neg p)\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}\pi''(p)\\ \pi''(\neg p)\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$, which is natural. But then we have $\text{comb}(\begin{bmatrix}1\\ \varepsilon\end{bmatrix}, \begin{bmatrix}\varepsilon\\ 1\end{bmatrix}) = \begin{bmatrix}1\\1\end{bmatrix}$, for any positive value ε. Thus it makes a difference between assessing the value 0 or the value ε to π(p) ; indeed in the first case we are sure of the falsity of p while in the other we are not. In any case 1-max(π(p), π(¬p)) where π(p) and π(¬p) are given by (18)-(19) evaluates the extent of the conflict. Note that there is no reinforcement since $\text{comb}(\begin{bmatrix}1\\ \lambda\end{bmatrix}, \begin{bmatrix}1\\ \lambda\end{bmatrix}) = \begin{bmatrix}1\\ \lambda\end{bmatrix}$; in case we want to reinforce (but it is risky to do it systematically), we may use the product instead of min operation in (18) and (19). We have $\text{comb}(\begin{bmatrix}1\\ \lambda\end{bmatrix} \begin{bmatrix}1\\ \mu\end{bmatrix}) = \begin{bmatrix}1\\ \min(\lambda,\mu)\end{bmatrix}$; thus the least uncertain information is kept. Particularly,

$$\text{comb}(\begin{bmatrix}1\\ \lambda\end{bmatrix}, \begin{bmatrix}1\\ 1\end{bmatrix}) = \begin{bmatrix}1\\ \lambda\end{bmatrix}.$$

In case of interval values in the combination operation, we may use (11)-(12). Then, due to (1) $\text{comb}(\begin{bmatrix}1\\ \lambda\end{bmatrix}, \begin{bmatrix}?\\ ?\end{bmatrix}) = \text{comb}(\text{comb}(\begin{bmatrix}1\\ \lambda\end{bmatrix}, \begin{bmatrix}1\\ ?\end{bmatrix}), \text{comb}(\begin{bmatrix}1\\ \lambda\end{bmatrix}, \begin{bmatrix}?\\ 1\end{bmatrix}))$ and finally we get $\text{comb}(\begin{bmatrix}1\\ \lambda\end{bmatrix}, \begin{bmatrix}?\\ ?\end{bmatrix}) = \begin{bmatrix}1\\ [0,\lambda]\end{bmatrix}$. In fact here we have taken into account the information that anything was possible concerning p according to one of the sources (it would be different if the second source is ignored or does not exist), it is why we get π(¬p) ≤ λ rather than π(¬p) = λ. The basic point is that when we assess a precise value to a possibility degree, we tacitly admit that this value may be decreased in the light of a least uncertain information, without there is any confict between the pieces of information.

However as we shall see in the next section, there are situations where not combining is better.

4.2 - Not combining :

Let us consider the small example of section 3.2 again. Now we add the following default rule to our knowledge base :

(4) "Usually, if Bob comes, Tom does not come", represented by

$$\begin{bmatrix}\pi(T|B) & \pi(T|\neg B)\\ \pi(\neg T|B) & \pi(\neg T|\neg B)\end{bmatrix} = \begin{bmatrix}\mu & ?\\ 1 & ?\end{bmatrix}$$

The situation is pictured on Figure 1.

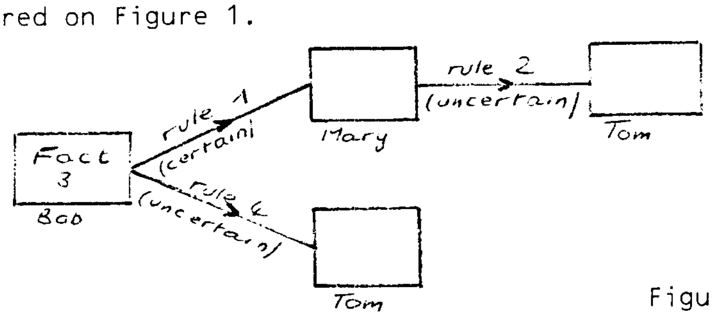

Figure 1

From fact 3 and rules 1 and 2 we got in section 3.2. $\begin{bmatrix}\pi(T)\\ \pi(\neg T)\end{bmatrix} = \begin{bmatrix}1\\ \lambda\end{bmatrix}$.

Applying rule 4 to fact 3 gives

$$\begin{bmatrix}\pi(T)\\ \pi(\neg T)\end{bmatrix} = \begin{bmatrix}\mu & ?\\ 1 & ?\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} = \begin{bmatrix}\mu\\ 1\end{bmatrix}$$

Thus from two different derivations we get two conflicting albeit uncertain conclusions. We might think of combining them. It is not necessarily a good idea for reasons we give in the following. Another related question is to wonder if the conflict between the two partial conclusions is due to some inconsistency in the knowledge base or not.

What happens here is that we have a specific rule which enables us from the fact 3 to directly derive a conclusion concerning Tom's coming on the one hand, while on the other hand, we indirectly derive another conclusion from an intermediary conclusion obtained through a certain rule.

It is extremely important to notice that the two rules 1 and 2 cannot be combined in order to generate a new rule of the form "if Bob comes, then generally Tom comes". Indeed if we only know that all the times that Bob comes, Mary comes also and that most of the times when Mary comes, Tom comes, we have no idea of the number of times when Bob and Tom comes together as pointed out by Zadeh [13]. It may happen that in fact Tom comes when Mary comes but Bob does not, since Bob's coming is not a necessary condition for Mary's coming. Thus the rule 4 in our example is not inconsistent with the two other rules. Note that if all the rules were certain, the knowledge base would become inconsistent. However if we know that Mary comes it is legitimate to conclude from rule 2 that there is only a possibility equal to λ that Tom does not come. The fact that Mary comes can be definitely established from fact 3 and rule 1 in our example. In the absence of rule 4, the conclusion regarding Tom's coming is correct with respect to our state of knowledge although we have not the rule "if Bob comes, then generally Tom comes". When we have the rule 4, which is more specific, we must prefer the conclusion directly obtained from fact 3 and rule 4. See Prade [7], Ginsberg [3], Reiter and Criscuolo [8] for related discussions.

When the rule 1 is an equivalence or is at least almost reversible, a new (uncertain) rule can be legitimately built by chaining rules 1 and 2. Indeed it can be easily shown (see Zadeh [13]) that if we use a probabilistic approach we have

$$\mathrm{Prob}(T|B) \geq \max(0, \mathrm{Prob}(T|M) + \mathrm{Prob}(B|M) - 1) \qquad (20)$$

This formula can be easily extended to fuzzy-valued probabilities [13]. In case of the possibilistic approach it can be shown that we must have (see [1])

$$\max(\pi(\neg T|B),\ \pi(\neg B|M)) \geq \pi(\neg T|M) \qquad (21)$$

which can be translated in terms of necessity using (2)

$$\min(n(T|B),n(B|M)) \leq n(T|M) \qquad (22)$$

and which expresses that our certainty that Tom comes when Mary comes cannot be less than both the "degree of reversibility" of rule 1 and the certainty that Tom comes according to rule 4, elsewhere the knowledge base may be considered as inconsistent.

Note also that in order to be valuable a specific rule ought to lead to less uncertain conclusions than the ones obtainable by indirect derivations in case of consonance of the conclusions.

The above discussion remains preliminary but points out a central issue in inexact or default reasoning which must be dealt with by using sophisticated control strategies and/or by being cautious in establishing a knowledge base. It is thus important to be aware of this problem when building inference systems.

5 - Concluding remarks

The approach to uncertainty in reasoning we propose here remains quite simple in its principles. The knowledge representation, in terms of possibility distributions, which is used seems more suitable for representing uncertainty here than a probability-based method, since we are able to distinguish between a total lack of certainty that p is satisfied ($\pi(\neg p) = 1$) and the certainty that p is not satisfied ($\pi(p) = 0$), which is not possible in probility theory where $\text{prob}(\neg p) = 1 \Leftrightarrow \text{prob}(p) = 0$. Max and min are "qualitative" operations which are in agreement with the possible lack of precision of the different possibility degrees, what really matters is only that some alternatives are certainly more possible that others. They are only sensitive to the orderings of possibilities. Operations used in probability theory are more sensitive to changes (even limited) in probability values.

The basic ideas of this paper are currently experimented in an implemented system dealing with financial analysis. First results suggest that the approach works well in practice.

Besides the general problem of knowing when we have to block a combination operation, discussed in section 4.2., is a topic for further research.

REFERENCES

[1] - Dubois, D., Prade, H. (1985) The management of uncertainty in fuzzy expert systems and some applications. In : The Analysis of Fuzzy Information (J.C. Bezdek, ed.), CRC Press, Boca Raton, Fl., Vol.2, to appear.

[2] - Farreny, H. , Prade, H. (1985) A possibility theory-based approach to default and inexact reasoning. Computers and Artificial Intelligence, (Bratislava) , 4, to appear.

[3] - Ginsberg, M.L. (1984) Non-monotonic reasoning using Dempster's rule. Proc. Nat. Conf. on Artificial Intelligence (AAAI-84), Austin, Tx, Aug. 6-10, 1984, 126-129.

[4] - Pólya, G. (1954) Mathematics and Plausible Reasoning. Vol. II : Patterns of Plausible Inference. Princeton University Press, 2nd edition 1968.

[5] - Prade, H. (1985) A computational approach to approximate and plausible reasoning with applications to expert systems. IEEE Trans. Pattern Analysis & Machine Intelligence, 7, n° 3, 260-283.

[6] - Prade, H. (1985) A simple inference technique for dealing with uncertain facts in terms of possibility. Kybernetes, to appear.

[7] - Prade, H. (1985) Some issues in approximate and plausible reasoning in the framework of a possibility theory-based approach. In : Matters of Intelligence (L.M. Vaina, ed.), Academic Press, to appear.

[8] - Reiter, R., Criscuolo, G. (1983) Some representational issues in default reasoning. Computers & Mathematics with Applications, 9, n° 1, 15-27.

[9] - Rich, E. (1983) Default reasoning as likelihood reasoning. Proc. Nat. Conf. Artificial Intelligence (AAAI-83), Washington, DC, 348-351.

[10] - Shafer, G. (1976) A Mathematical Theory of Evidence. Princeton Univ. Press.

[11] - Shortliffe, E.H., Buchanan, B.G. (1975) A model of inexact reasoning in medicine. Mathematical Biosciences, 23, 351-379.

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

[13] - Zadeh, L.A. (1984) A theory of commonsense knowledge. In : Aspects of Vagueness (H.J. Skala, S. Termini, E. Trillas, Eds.), D. Reidel, 257-295.