Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge

Didier DUBOIS and Henri PRADE
Institut de Recherche en Informatique de Toulouse (I.R.I.T.)
Université Paul Sabatier – 118 route de Narbonne
31062 Toulouse Cedex 4 – France
Email: {dubois, prade}@irit.fr

Abstract: This paper advocates a basic distinction between two epistemic operations called focusing and revision, which can be defined in any, symbolic or numerical, representation framework which is rich enough for acknowledging the difference between factual evidence and generic knowledge. Revision amounts to modifying the generic knowledge when receiving new pieces of generic knowledge (or the factual evidence when obtaining more factual information), while focusing is just applying the generic knowledge to the reference class of situations which exactly corresponds to all the available evidence gathered on the case under consideration. Various settings are considered, upper and lower probabilities, belief functions, numerical possibility measures, ordinal possibility measures, conditional objects, nonmonotonic consequence relations.

1 - Introduction

Some basic modes of belief change have been laid bare by Levi (1980): an *expansion* corresponds to adding the new piece of information without rejecting previous beliefs, a *contraction* is the converse operation by which some piece of information is lost, and *revision* itself strictly speaking corresponds to accepting a piece of information that partially contradicts previous beliefs and modifying the latter accordingly. Revision and expansion are coherent in the sense that they coincide when the input information is consistent with previous beliefs. Another mode of belief change that has been more recently recognized is called updating (Katsuno and Mendelzon, 1991). The difference between updating and revision is that an updating operation takes into account the fact that the world referred to in the body of knowledge, has evolved, and so the set of beliefs must evolve accordingly; on the contrary, belief revision presupposes that some previous beliefs were wrong, and are corrected by the input information.

A last distinction, which is often absent from the symbolic approaches but which probabilists often endorse, even if not always explicitly, is between revision and what we have called 'focusing' (Dubois and Prade, 1992a). This is the topic of this paper. It relies on the idea that a body of knowledge often consists in two parts: background knowledge and particular pieces of evidence on a case at hand. The input can either alter the background knowledge (revision), or complete the available evidence, which then points to a new reference class (focusing).

To handle the distinction properly, we separate the generic knowledge K pertaining to classes of situations, from the factual evidence E pertaining to a particular case to which the generic knowledge is applied. In the following the term 'knowledge' refers to generic information, while information about a case is referred to as '(factual) evidence'. These two kinds of information may convey some form of partial ignorance. Factual evidence consists of information gathered on a case at hand, or the description of the actual world in a given situation. This information can be more or less precise and more or less reliable. Generic knowledge pertains to a class of situations considered as a whole, but does not refer to a particular case. Sometimes this class of situations is well defined (it is a population in the sense of statistics) and the generic knowledge is of a frequentist nature. Sometimes the relevant class of population is much more vaguely described (as in the famous "birds fly" example), and the generic knowledge describes rules of plausible inference of the form "if all I know is p then plausibly conclude q" (when knowledge is pervaded with exceptions). The levels of confidence expressed by the rules can be numerically modelled via a conditional probability for instance, or be handled in a purely ordinal setting (as in default reasoning). The factual evidence, which might be also somewhat uncertain, would be however assumed to be represented by propositional statements in the following (except if the contrary is explicitly acknowledged). The difference between generic knowledge and factual evidence can be illustrated by a diagnosis problem. The generic knowledge of a clinician consists in his/her knowledge about the links between the diseases and the symptoms and the distribution of the diseases in the population (in practice, the likelihoods and the prior probabilities). The factual evidence consists in the symptoms collected from the patient under consideration.

The next section provides a general discussion of focusing versus revision. See also Dubois and Prade (1994a), Dubois et al. (1996) for detailed discussions on this topic. Section 3 recalls focusing rules in numerical settings while Section 4 summarizes how focusing is handled in non-numerical settings. Section 5 briefly points out further developments and provides concluding remarks.

2 - Focusing vs. Revision — A General Discussion

The distinction between evidential and generic information is crucial for a proper understanding of belief revision processes and commonsense inference. By applying the generic knowledge K to the evidence E, a set C of consequences can be produced. Thus C is made of default conclusions, or of propositions having high probabilities, according to the representation framework used in K. Generally speaking, C is the set of the most plausible conclusions which can be inferred from E and K. Clearly, the consequences in C are open to revision and should be defeasible. Indeed when more evidence is available and E is changed into E', the consequences which can be obtained from E' and K may differ, and even contradict those derivable from E and K. In Gärdenfors(1988)'s view of revision, no distinction is made between evidence and knowledge. Beliefs are supposed to be maintained in a so-called, (propositional) belief set B closed under logical consequences. This belief set is to be revised when an input information is received which contradicts B. The lack of distinction between evidence and knowledge in Gärdenfors' approach puts all the information at the same level and limits the expressive power. However, considering revision and nonmonotonic reasoning as the two sides of the same coin (Makinson and Gärdenfors, 1991), reconciles the two views. Indeed if we consider that the belief set B corresponds exactly to the set of plausible consequences C, the epistemic entrenchment which underlies any revision process obeying the postulates proposed by Alchourrón, Gärdenfors and Makinson (see Gärdenfors, 1988) can be turned into a plausibility ordering between interpretations which implicitly gives birth to a conditional knowledge base K (Gärdenfors and Makinson, 1994). Then, revising B = C into C' when a new information is received (i.e., when E is expanded into E') amounts to computing C' as the set of propositions p such that the default rule E' ~> p is a piece of conditional knowledge deducible from K using the rational closure entailment proposed by Lehmann and Magidor (1992). But this corresponds to a particular modelling of the generic knowledge supposed to be stored in K. The distinction between E, K and C leaves open the way the generic knowledge is represented (K may also be modelled in terms of a family of probability measures, or by a Bayesian network, for instance).

An example of probabilistic knowledge base that can be represented by means of a family of probability measures consists in a set of statements of the form "most A_i's are B_i's", "few A_j's are B_j's", ... etc. More generally K = {"Q_i A_i's are B_i's", i = 1,n} where Q_i is an imprecise quantifier of the form $[\alpha,1]$ (for "most") or $[0,\alpha]$ (for "few"), or $[\alpha,\beta]$, expressing an ill-known proportion. Each such statement can be modelled by a constraint of the form $P(B_i \mid A_i) \in Q_i$. The evidence E is supposed to be incomplete but not uncertain here. It is represented by a proposition or an event. Assume a new piece of information arrives. The problem addressed here is what to do in order to account for the new information in the pair (K,E) so as to produce new plausible conclusions. When the input information is evidential it only affects the evidence E. Then, focusing consists in applying the generic knowledge to the context of interest as described by the evidential information (including the new one), in order to answer questions of interest pertaining to this particular situation. In such a process the knowledge should remain completely unchanged. Namely we know that the case at hand is in class A (E = A). Then, considering the above probabilistic example, suppose it is asked whether the case at hand has property B. What is to be computed is $P(B \mid A)$. However since there maybe more than one probability measure restricted by the statements in K, only upper and lower bounds of $P(B \mid A)$ can be computed.

This has to be contrasted with the situation where the input information is generic. It thus should be understood as a new constraint refining the available knowledge. In such a case, a genuine revision takes place since the knowledge has to be modified in order to incorporate the new piece of information as a new constraint (except if it is redundant with respect to what is already known). In the case of revision, the new piece of information is of the form P(A) = 1, which means that all cases (that the knowledge base refers to) satisfy property A. Then K becomes $K' \cup \{P(A) = 1\}$ and new values for the upper and lower bounds for P(B) where B is a property of interest can be then computed. These bounds do not coincide with the ones of $P(B \mid A)$ calculated above in the focusing case, and may even not exist if K implies that $P(A) \neq 1$. Revision of generic knowledge would correspond to the following types of example: a probability distribution over sizes of adults in a given country is available, and some input information comes in that nobody in the population is more than 6ft tall. These remarks make it clear that revising is not focusing. However, note that for both revision and focusing on the basis of uncertainty functions, such as probability measures, the basic tool is the conditioning operation.

A revision process can be iterated, and it progressively modifies the initial knowledge base. Iterated focusing only refines the body of evidence and the same generic knowledge is applied to a new reference class of situations pointed at by the evidence. In some problems, only evidential information is available, and new pieces of possibly uncertain evidence lead to a non-trivial revision process due to uncertainty, as in Shafer (1976). It is also worth pointing out that in case of multiple sources of information, the factual evidence is obtained through a symmetric combination process (for instance by applying Dempster's rule of combination in the belief function framework). A similar process should apply at the generic knowledge level if K is built from several sources.

The distinction between focusing and (generic) revision can only be made in uncertainty frameworks where the difference between generic knowledge and factual evidence can be captured and where generic knowledge is represented and processed as distinct from factual evidence. Such a distinction is not relevant in propositional logic for instance, since every piece of information takes the form of a propositional sentence. In the setting of Bayesian probability the distinction is conceptually meaningful but focusing and revising are both expressed by the same conditioning rule, that is Bayes rule. Hence some disputes among probabilists as to the meaning of conditioning. The most widely found view (e.g., De Finetti, 1974) is that Bayes rule operates a change of reference class, namely going from a prior to a posterior probability is not a revision process. Posterior probabilities are precalculated and the input information just prompts the selection of a particular posterior. With this view, the prior probability together with the likelihoods determine a unique joint probability over a space of interest, construed as a body of generic knowledge (the clinician's experience on a certain disease), and conditioning means integrating factual evidence so as to configurate the generic knowledge properly with respect to the reference class of the object on which this factual evidence bears (test results on the patient). This point of view, which is shared by the rule-based expert system literature, culminates with the advent of Bayesian networks.

The revision view of conditioning is typically advocated by people working in probability kinematics. Philosophers like Jeffrey (1983), Domotor (1985), Williams (1980), and Gärdenfors (1988), understand Bayes rule as a revision process, by which a prior probability P is changed into a new one P' due to the input information A. This latter view is supported by maximal cross-entropy arguments whereby it is established that Bayes rule obeys a minimal change requirement. This school of thought is called probability kinematics: the input information A is understood as the discovery that $P'(A) = 1$ while the prior is $P(A) < 1$, and must be modified. The possibility that the input information is itself probabilistic is commonly envisaged in probability kinematics: the input information enforces the probability of some event to take on a value which differs from the prior value. In this view the input information is at the same level as the prior information: both are (possibly uncertain) factual evidence or pieces of generic knowledge.

3 - Numerical Settings

3.1 - Belief Functions — Upper and Lower Probabilities

While revision and focusing do coincide in probability theory as already said, they no longer coincide in more general settings such as belief functions or upper and lower probabilities. Revision in belief function theory is defined by Dempster rule of conditioning that combines the conjunctive revision mode of the crude partial ignorance model, with Bayes rule normalization underlying a stability of the degrees of uncertainty in relative value. Let Bel be a belief function defined by $Bel(A) = \sum_{\emptyset \neq E_i \subseteq A} m(E_i)$ from a set of masses m_i on focal elements E_i's such that $\sum_i m(E_i) =$

1, $E_i \subseteq \Omega$, and Pl the associated plausibility measure $Pl(A) = 1 - Bel(\bar{A})$. Indeed Dempster rule of conditioning can be described as follows:

- Given the new piece of information A, turn each focal element E_i into $E_i \cap A$, and attach the mass m_i to it, adding the masses m_i and m_j if $E_i \cap A = E_j \cap A$.
- Renormalize the masses allocated to non-empty subsets $E_i \cap A \neq \emptyset$ as with Bayes rule, so as to reallocate the masses m_i such that $E_i \cap A = \emptyset$ proportionally.

It is well-known (Shafer, 1976) that it comes down to compute the expected degree of potential support:

$$Pl(B \mid A) = \frac{Pl(A \cap B)}{Pl(A)}. \tag{1}$$

Interestingly, Dempster's rule of conditioning can also be derived in a pure upper and lower probability context. If the belief function is viewed as characterizing a set of probabilities $\mathbb{P} = \{P \leq Pl\}$, then it can be proved that:

$$Pl(B \mid A) = \frac{Pl(A \cap B)}{Pl(A)} = \sup\left\{\frac{P(A \cap B)}{P(A)},\ P \leq Pl,\ P(A) = Pl(A)\right\}. \tag{2}$$

This result is due to the fact that the constraint $P(A) = Pl(A)$ never forbids $\sup\{P(A \cap B), P \leq Pl, P(A) = Pl(A)\}$ to be equal to $Pl(A \cap B)$, i.e., we can always have $P(A) = Pl(A)$ and $P(A \cap B) = Pl(A \cap B)$ for the same probability measure P, if Pl is a plausibility function. Equation (2) makes it clear the kind of revision at work with Dempster rule in the upper and lower probabilities context: the constraint $P(A) = Pl(A)$ corresponds to the selection of a maximum likelihood probability, which is very usual in statistics. As a generalization of this principle, Moral and De Campos (1991) have suggested that the magnitude of P(A) reflects the possibility of accepting the corresponding conditional probability $P(\cdot \mid A)$ in the updated set of probabilities.

Gilboa and Schmeidler (1992) have given a decision-theoretic interpretation of $Pl(B| A)$ in the setting of upper and lower probabilities, i.e., when the set $\mathbb{P}$ does not necessarily represent a belief function. Indeed equation (2) does hold when $Pl = P^*$ satisfies order-2 subadditivity only, that is $P^*(A \cup B) \leq P^*(A) + P^*(B) - P^*(A \cap B)$. The reason is that the property stating that we can always have $P(A) = P^*(A)$ and $P(B) = P^*(B)$ for a subset B of A for the same probability measure P is characteristic of order-2 subadditivity (see Huber, 1981). However equation (2) does not hold for upper probabilities deriving from any set of probabilities $\mathbb{P}$.

The alternative conditioning rule, already suggested by Dempster (1967), De Campos et al. (1990), Fagin and Halpern (1989) and Jaffray (1992) and here referred to as focusing, is defined in the setting of upper and lower probabilities as:

$$P^*_A(B) = \sup\left\{\frac{P(A \cap B)}{P(A)},\ P \leq P^*\right\} = \frac{P^*(A \cap B)}{P^*(A \cap B) + P_*(A \cap \bar{B})} \tag{3}$$

where P^* denotes an upper probability function and P_* the associated lower probability with $P^*(A) = 1 - P_*(\bar{A})$. As already said, the idea of focusing is to compute the probability of B in the state when A is supposed to be true without making any assumption about how the set of probabilities should be revised if A were actually true, especially without considering any probability measure in the set $\{P \leq P^*\}$ as impossible except those such that $P(A) = 0$. It leads to performing a sensitivity analysis on Bayes rule, when the probability function ranges over the set $\{P \le P^*\}$. Interestingly, if $P^* = Pl$ is a plausibility function then P^*_A is still a plausibility function. Note that while Dempster rule of conditioning (2) can easily be iterated, this is not so with the focusing rule, because the set $\{P(\cdot \mid A), P \le P^*\}$ is generally a proper subset of the set $\{P, P \le P^*_A\}$ induced by the upper probability bounds obtained via focusing (see Jaffray, 1992) even if P^* is a plausibility measure. But iterating the focusing operation by "conditioning again" P^*_A on a new event C makes no sense. What must be done is to start from the original function P^* and focus on the new reference class $A \cap C$.

P^*_A is generally much less informative than $P^*(\cdot \mid A)$ (applying (2) with P^* instead of Pl) and even sometimes less informative than P^* itself. This has been noticed by Kyburg (1987) for belief functions where $Pl_A \ge Pl(\cdot \mid A) \ge Bel(\cdot \mid A) \ge Bel_A$ holds. For instance if $A \cap E_i \ne \emptyset$, and $E_i \not\subseteq A\ \forall i$, then $P^*_A(B) = 1$ and $P^*_A(\bar{B}) = 1$, $\forall\ B \ne A$; $B \subseteq A$, i.e., we get a total ignorance function on the referential set A. This would be surprising in a learning process where A is a new piece of information and revision should improve our knowledge. The reason is that focusing is not made for revision and achieves no learning. To make it clear, suppose that the set of probabilities $\mathbb{P}$ represents knowledge stored in a database. Then $P^*_A(B)$ is part of the response of a query asking for the probability of being in B for an individual in A (focusing rule). The response should give $P_{*A}(B) = 1 - P^*_A(\bar{B})$ as well. On the other hand $P^*(B \mid A)$ is the result of modifying the database by enforcing $P(A) = 1$ (revision).

Focusing can be justified in terms of belief functions only, namely $Pl_A(B)$, as obtained in (3), can be viewed as the upper limit of a family of belief functions obtained by transferring for all focal elements such that $E_j \cap A \ne \emptyset$, $E_j \cap \bar{A} \ne \emptyset$ only one part of the mass m_j to the set $E_j \cap A$; see (De Campos et al., 1990). Each possible partial mass transfer from each focal element E_j to $E_j \cap A$ determines a possible way of changing the belief function into another one. The focusing process refrains from selecting among these potentially resulting belief functions and rather considers their lower envelope.

Examples: Consider the following small knowledge base

$$P(\text{young} \mid \text{student}) \ge 0.9$$
$$P(\text{single} \mid \text{young}) \ge 0.7$$

The focusing rule deals with the following problem: Tom is a student; what is the probability that he is single. It can be checked that P(single | student) is totally unknown: $P_*(\text{single} \mid \text{student}) = 0$, $P^*(\text{single} \mid \text{student}) = 1$. In particular nothing is known about P(student). The revision rule consists in entering the new piece of knowledge P(student) = 1, that expresses the fact that the knowledge base applies only to a population of students. Then it can be deduced that P(young) = P(young | student) ≥ 0.9 and P(single) = P(single | student) $\ge$ P(single | young) $\cdot$ P(young) = 0.63. However it is a mistake to model "Tom is a student" by P(student) = 1 since the former statement applies to Tom only while P(student) = 1 refers to the whole population under concern.

Another example illustrating the difference between focusing and revision can be given in the setting of belief functions: a die has been thrown by a player but he does not know the outcome yet. The player expects a rather high number (5 or 6), does not believe too much in a medium number (3, 4), and almost rules out the possibility of getting a small number (1, 2). He assigns a priori probability masses m({5,6}) = 0.7, m({3,4}) = 0.2 and m({1,2}) = 0.1 to focal elements. If he asks himself what if the outcome were not a six, then he has to focus on the situations where he would get no 6 and compute $Bel_A(B)$, $Pl_A(B)$ for A = {1, 2, 3, 4, 5}, and $B \subseteq \{1, 2, \ldots, 5\}$, where $Bel(B) = \sum_{E \subseteq B} m(E)$. For instance

$$Bel_A(5) = 0 = Bel(5);\ Pl_A(5) = \frac{0.7}{0.7 + 0.3} = 0.7 = Pl(5).$$

This is because he cannot rule out the situation where his belief in outcome 5 precisely would be zero. But if a friend tells him that the outcome is not a 6, then by Dempster's conditioning one gets Bel(5 | A) = 0.7. What happens is that, insofar as outcome 5 receives all the mass, 6 is now definitely ruled out. ■

In Dempster's approach to belief function theory, a finite probability space (Ω,P) is projected to a set U through a multiple-valued mapping Γ thus inducing belief and plausibility functions on U, such that

$$\forall\, B \subseteq U,\ Bel(B) = P(B_*) / P(U^*) \text{ and } Pl(B) = P(B^*) / P(U^*)$$

where $B_* = \{\omega \mid \Gamma(\omega) \subseteq B\}$, $B^* = \{\omega \mid \Gamma(\omega) \cap B \neq \emptyset\}$ are the lower and upper images of B via Γ. Revision by a constraint $A \subseteq U$ comes down to modifying the multiple-valued mapping $\Gamma: \Omega \to 2^U$ into $\Gamma_A: \Omega \to 2^A$ such that $\Gamma_A(\omega) = \Gamma(\omega) \cap A$, and normalizing the result. Denoting B_{*A} and B^{*A} the lower and upper images of B via Γ_A, yields

$$Pl(B \mid A) = P(B^{*A}) / P(A^{*A}) = P((B \cap A)^*) / P(A^*)$$

since $B^{*A} = (B \cap A)^* = \{\omega \mid \Gamma_A(\omega) \cap A \cap B \neq \emptyset\}$ as defined above. On the contrary focusing on $A \subseteq U$ comes down to envisaging all the possible inverse images of A through mappings $f: \Omega \to U$ compatible with Γ, letting

$$\Gamma^{-1}(A) = \{f^{-1}(A) \mid f \in \Gamma\},$$

where $f \in \Gamma$ means $\forall\omega$, $f(\omega) \in \Gamma(\omega)$, and to condition P on all possible such $f^{-1}(A)$ in Ω. Noticing that $\Gamma^{-1}(A) = \{C \mid A_* \subseteq C \subseteq A^*\}$, it can be proved that (Dubois and Prade, 1994a)

$$Pl_A(B) = \max_{A_* \subseteq C \subseteq A^*} P(B^* \mid C). \tag{4}$$

In other words, focusing on $B \subseteq U$ comes down to conditioning on an ill-known event in Ω, due to the imprecision expressed via the multiple-valued mapping. On the contrary, Dempster's conditioning rule assumes that C = B* in the above theorem. This result may help in the practical computing of the focusing rule since it comes down to maximizing over a finite set (Ω is assumed to be finite).

3.2 - Focusing in Numerical Possibility Theory

Since numerical possibility measures are particular cases of plausibility functions that are max-decomposable, it makes sense to envisage possibility degrees as upper bounds of probabilities and to apply conditioning (3) to possibility and necessity functions ∏ and N:

$$\forall B,\ \Pi_A(B) = \frac{\Pi(B \cap A)}{\Pi(B \cap A) + N(\bar{B} \cap A)} \tag{5}$$

and

$$N_A(B) = 1 - \Pi_A(\bar{B}) = \frac{N(B \cap A)}{N(B \cap A) + \Pi(\bar{B} \cap A)}. \tag{6}$$

When $N(A) > 0$, (5) and (6) do correspond to $\Pi_A(B) = \sup\{P(\cdot \mid A) \mid P \in \mathcal{P}(\Pi)\}$ and $N_B(A) = \inf\{P(\cdot \mid A) \mid P \in \mathcal{P}(\Pi)\}$, where $\mathcal{P}(\Pi) = \{P \mid P(B) \leq \Pi(B),\ \forall\ B \subseteq U\}$. Despite the strong probabilistic flavor of the above conditioning, it has been shown (Dubois and Prade, 1996) that Π_A and N_A are still possibility and necessity measures, i.e., Π_A is max-decomposable. This result was also recently pointed out by Walley (1996) for possibility measures. Moreover, given a possibility measure Π on a set U, the conditional set-function Π_A is associated with the possibility distribution

$$\begin{cases} \pi_A(u) = \max\left(\pi(u),\ \dfrac{\pi(u)}{\pi(u) + N(A)}\right) & \text{if } u \in A \\ \pi_A(u) = 0 & \text{if } u \notin A \end{cases} \tag{7}$$

with $\Pi_A(B) = \sup_{u \in B} \pi_A(u)$. The following properties are easy to check and expected as characterizing a focusing operation:

- if $N(A) = 0$, $\pi_A(u) = 1$, $\forall\ u \in A$ (total ignorance inside A when A is not an accepted belief, i.e., $N(A) > 0$ does not hold). This means that if nothing is known about A, focusing on A does not bring any information;
- $N_A(B) = N_A(A \cap B) = N_A(\bar{A} \cup B)$ (same property for Π_A). It means that when we focus on A, inquiring on B, $A \cap B$ or on $B \cup \bar{A}$ are the same. This is due to the fact that like in probability theory, $N_A(B)$ is the certainty degree of the conditional event $B \mid A$ (Dubois and Prade, 1994b), viewed as the family $\{X \mid B \cap A \subseteq X \subseteq \bar{A} \cup B\}$. Indeed, $N_A(B)$ only depends on $N(A \cap B)$ and $N(\bar{A} \cup B)$.

Other, more usual forms of conditioning exist in possibility theory in the numerical or in the ordinal setting. In case of a numerical possibility scale, we can also use Dempster rule of conditioning, specialized to possibility measures, i.e., consonant plausibility measures of Shafer (1976). It leads to the definition:

$$\forall B,\ B \cap A \neq \emptyset,\ \Pi(B \mid A) = \frac{\Pi(A \cap B)}{\Pi(A)} \tag{8}$$

provided that $\Pi(A) \neq \emptyset$. The conditional necessity function is defined by $N(B|A) = 1 - \Pi(\bar{B} \mid A)$, by duality.

As already said, focusing leads to a rather uninformative conditioning process. Indeed, it leads to belief-plausibility intervals that are wider and contain the interval obtained by Dempster rule. Dempster rule leads to a shrinkening the belief-plausibility intervals, thus acknowledging that the available knowledge has been enriched. On the contrary the focusing rule loses all information (except the assumption that the class of envisaged situations lies in A), because, in the absence of revision, no mass allocated to focal sets is allowed to flow to one of its proper subsets. The particular case of possibilistic focusing (7) indicates that this loss of information is systematic since π_A is always less specific that π on the set A. In other words, if some uncertain and incomplete information is available about a class of situations described on Ω, the induced information on a subclass of it characterized by a subset of Ω can be noticeably more imprecise.

4 - Non-Numerical Settings

The distinction made in the ordinal (and logical) setting between factual evidence and generic knowledge is in good agreement with the one arising in the previous section. A set of default rules $K = \{p_i \sim> q_i, i = 1,n\}$ can be considered as being analogous to a set $\mathbb{P}(K)$ of probabilities induced by conditional probability bounds $K = \{P(B_i \mid A_i) \geq a_i, i = 1,n\}$, where B_i (resp.: A_i) is the set of models of q_i (resp.: p_i). Focusing on a reference class A gathering the models of the proposition p leads to compute bounds of $P(B \mid A)$ induced by $\mathbb{P}(K)$ for some B of interest as previously explained. Let B be the set of models of q. If the coefficients a_i are infinitesimally close to 1, that is, of the form $1 - \varepsilon$, finding that $P(B \mid A) \geq 1 - \varepsilon$ in $\mathbb{P}(K)$ is strictly equivalent to inferring the rule $p \sim> q$ from K in the preferential system of Kraus et al. (1990)). This is not surprizing because the rule $p \sim> q$ can be viewed as a conditional objects $B \mid A$, in the sense of De Finetti (1974, pp. 307-312), such that $P(B \mid A)$ is the probability of $p \sim> q$, and the logic of conditional objects (Dubois and Prade, 1994b) is a model of the preferential system of Kraus et al. (1990). This is also closely related to the approach to default reasoning in the ordinal possibility setting.

The revision problem in the ordinal approach to possibility theory is treated in details in (Dubois and Prade, 1992b). Let us briefly recall it. When revising by A, each subset B of Ω is changed into $B \cap A$, and the level of possibility $\Pi(B)$ is carried over to $B \cap A$ if $B \cap A \neq \emptyset$. Finally, all subsets $C \subseteq \Omega$ such that $\Pi(C) = \max\{\Pi(B \cap A), B \subseteq \Omega\}$ are assigned a maximal level of possibility 1. It corresponds to a notion of conditional possibility $\Pi(B \mid A)$ defined as:

$$\Pi(B \mid A) = \sup \{a \in [0,1], \Pi(A \cap B) = \min(a, \Pi(A))\}.$$

This type of ordinal conditioning has been also recently studied by Williams (1994) under the name "adjustment". Recalling that in the ordinal case, the partial ignorance function pair (N, Π) is equivalent to a subset E of possible worlds equipped with a complete ordering relation expressing plausibility, revision by subset A comes down to restricting (without changing) this ordering to $A \cap E$. Just as in the case of Bayes conditioning the ordering of elements in $A \cap E$ is left unchanged. This revision satisfies all the Alchourrón-Gärdenfors-Makinson revision postulates, and the N function coincides with what Gärdenfors (1988) calls an "*epistemic entrenchment*".

The question of representing generic knowledge in ordinal possibility theory can be addressed if possibility theory is related to conditional theories of default reasoning (Kraus et al., 1990; Lehmann and Magidor, 1992). A body of generic knowledge K is encoded as a set of default rules $p \sim> q$, each being interpreted as a nonmonotonic inference rule "if p is true then it is normal that q be true". K is also called a conditional knowledge base. As proved in (Benferhat et al., 1992), it is natural to interpret a default rule $p \sim> q$, as a constraint on possibility measures of the form $\Pi(p \wedge q) > \Pi(p \wedge \neg q)$, i.e., such a default rule can be viewed as expressing that $p \wedge q$ is more normal than $p \wedge \neg q$. A set of default rules $K = \{p_i \rightarrow q_i, i = 1,n\}$ with noncontradictory condition parts can be viewed as a family of constraints

$$\mathcal{C}(K) = \{\Pi(p_i \wedge q_i) > \Pi(p_i \wedge \neg q_i), i = 1,n\}$$

restricting a family $\mathbf{\Pi}(K) = \{\Pi, \Pi \in \mathcal{C}(K), \Pi(p_i) > 0,\ i = 1,n\}$ of possibility distributions over the interpretations of a language. Given a piece of evidence encoded as a propositional sentence p and the conditional knowledge base K, the sentence q is a plausible consequence of (K,p) iff the rule $p \sim> q$ can be deduced from K. This inference has a precise meaning in the framework of possibility theory namely the (classical) inference of the strict inequality $\Pi(p \wedge q) > \Pi(p \wedge \neg q)$ for all Π in $\mathbf{\Pi}(K)$ derived from the set of constraints $\mathcal{C}(K)$. This inference also perfectly fits with the so-called "preferential inference system" of Kraus et al. (1990) and with the logic of conditional objects (Dubois and Prade, 1995). Another more productive type of inference is the so-called "rational closure" (Lehmann and Magidor, 1992). It can be captured in the possibilistic setting by selecting a particular possibility measure Π^* in $\mathbf{\Pi}(K)$ (the least informed one, see (Benferhat et al., 1992)) and checking the condition $\Pi^*(p \wedge q) > \Pi^*(p \wedge \neg q)$. These methods of inference of a default rule $p \sim> q$ from a body of generic knowledge K can be viewed as a focusing on the reference class pointed at by p, for the purpose of plausibly inferring q. On the contrary, a revision process consists in modifying the set of rules K by adding or deleting rules, of by restricting the set of possible worlds.

The logic of conditional objects (or any equivalent approach) has also the merit of displaying the difference between two modes of belief revision: evidence focusing and knowledge expansion that can be defined (conditional objects encode a default rule $p \sim> q$ as an entity $q \mid p$ which can be symbolically processed in agreement with a 3-valued semantics; see (Dubois and Prade, 1994b)), as follows:

– *Evidence focusing*: a new piece of evidence p arrives and makes the available information on the case at hand more complete. Then E is changed into $E \cup \{p\}$ (supposedly consistent). K remains untouched. But the plausible conclusions from K and $E \cup \{p\}$, i.e., c such that $K \vDash c \mid E \wedge p$ may radically differ from those derived from K and E.
– *Knowledge expansion*: it corresponds to adding new generic rules tainted with possible exceptions. Insofar as the new knowledge base is consistent it is clear that due to the monotonicity of inference $\vDash$, all plausible conclusions derived from K can still be derived from K' since if K is a subset of K' and $K \vDash c \mid E$ then $K' \vDash c \mid E$. But more conclusions may perhaps be obtained by K'.
– *Knowledge revision*: it encompasses the situation when the result of adding new generic rules to K leads to an inconsistency. In that case some mending of the knowledge base must be carried out in order to recover consistency. Preliminary results along this line are in Boutilier and Goldszmidt (1993, 1995).

5 - Concluding remarks

We have seen that the distinction between factual evidence and generic knowledge has important consequences for a proper handling of information in flux. It may also contribute to put the Ramsey test, which is an important issue in connection with belief change, in the correct perspective.

The idea of the Ramsey test is to model belief change operations by means of conditionals expressing propositions of the form "in the context of a belief set accepting an input A leads to accept proposition B" and to assume that these conditionals belong to the belief set. A well-known result of Gärdenfors (1986) shows that putting such conditionals in the belief set is inconsistent with basic postulates of belief change, if the minimal change principle is understood as using a mere expansion when the input does not contradict the state of belief. In particular these principles imply that revision is monotonic with respect to the belief set. But is is easy to convince oneself that belief revision, because of the possible inconsistency between the input information and the belief set cannot behave in a monotonic way.

The distinction between evidence E, knowledge K and plausible consequences C, makes also clear the need for other knowledge operations. In particular, if we want to modify the set C into a set C' by either adding desirable, or deleting undesirable conclusions, for a given evidence E, we would have to modify the generic knowledge K. This operation has been called "repairing K", in (Benferhat et al., 1996) where a method is proposed for tuning the set of conclusions derivable from K by applying rational monotony postulate (Gärdenfors and Makinson, 1994), without however questioning the conclusions obtained by preferential entailment (in the sense of Kraus, Lehmann and Magidor, 1990).

Besides, the introduction of new types of conditioning corresponding to the focusing operation, in the belief functions, and numerical possibility theory settings, suggests their possible use in the modelling of default rules. Preliminary investigations can be found in Dubois and Prade (1994a, 1997).

References

Benferhat S., Dubois D., Prade H. (1992) Representing default rules in possibilistic logic", Proc. of the 3rd Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'92), Cambridge, Mass., Oct. 26-29, 673-684.

Benferhat S., Dubois D., Prade H. (1996) Beyond counter-examples to nonmonotonic formalisms: A possibility-theoretic analysis. Proc. of the 12th Europ. Conf. on Artificial Intelligence (ECAI'96) (W. Wahlster, ed.), Budapest, Hungary, Aug. 11-16, 1996, John Wiley & Sons, New York, 652-656.

Boutilier C., Goldszmidt M. (1993) Revision by conditionals beliefs. Proc. of the 11th National Conf. on Artificial Intelligence (AAAI'93), Washington, DC, July 11-15, 649-654.

Boutilier C., Goldszmidt M. (1995) On the revision of conditional belief sets. In: Conditionals: From Philosophy to Computer Science (G. Crocco, L. Fariñas del Cerro, A. Herzig, eds.), Oxford University Press, Oxford, UK, 267-300.

De Campos L.M., Lamata M.T., Moral S. (1990) The concept of conditional fuzzy measure. Int. J. of Intelligent Systems, 5, 237-246.

De Finetti B. (1974) Theory of Probability. Wiley, New York.

Dempster A.P. (1967) Upper and lower probabilities induced by a multiple-valued mapping. Annals of Mathematical Statistics, 38, 325-339.

Domotor Z. (1985) Probability kinematics — Conditionals and entropy principles. Synthese, 63, 74-115.

Dubois D., Prade H. (1992a) Evidence, knowledge, and belief functions. Int. J. of Approximate Reasoning, 6, 295-319.

Dubois D., Prade H. (1992b) Belief change and possibility theory. In: Belief Revision (P. Gärdenfors, ed.), Cambridge Univ. Press, Cambridge, UK, 142-182.

Dubois D., Prade H. (1994a) Focusing versus updating in belief function theory. In: Advances in the Dempster-Shafer Theory of Evidence (R.R. Yager, M. Fedrizzi, J. Kacprzyk, eds.), Wiley, New York, 71-95.

Dubois D., Prade H. (1994b) Conditional objects as nonmonotonic consequence relationships. IEEE Trans. on Systems, Man and Cybernetics, 24(12), 1724-1740.

Dubois D., Prade H. (1995) Conditional objects, possibility theory, and default rules. In: Conditional Logics: From Phisolophy to Computer Science (G. Crocco, L. Fariñas del

Cerro, A. Herzig, eds.), Oxford University Press, Oxford, UK, 301-336.
Dubois D., Prade H. (1996) Focusing vs. revision in possibility theory. Proc. of the 5th IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE'96), New Orleans, LO, Sept. 8-11, 1700-1705.
Dubois D., Prade H. (1997) Bayesian conditioning in possibility theory. Fuzzy Sets and Systems, to appear.
Dubois D., Prade H., Smets P. (1996) Representing partial ignorance. IEEE Trans. on Systems, Man and Cybernetics, 26(3), 361-377.
Fagin R., Halpern J. (1989) A new approach to updating belief. Research Report RJ 7222, IBM Research Division, San Jose, CA.
Gärdenfors P. (1986) Belief revision and the Ramsay test for conditionals. Philosophical Review, 91, 81-83.
Gärdenfors P. (1988) Knowledge in Flux — Modeling the Dynamics of Epistemic States. The MIT Press, Cambridge, MA.
Gärdenfors P., Makinson D. (1994) Non-monotonic inference based on expectations. Artificial Intelligence, 65, 197-245.
Gilboa I., Schmeidler D. (1992) Updating ambiguous beliefs. Proc. of the 4th Conf. on Theoretical Aspects of Reasoning about Knowledge (TARK'92) (Y. Moses, ed.), Monterey, CA, March 22-25, 143-162.
Huber P.J. (1981) Robust Statistics. Wiley, New York.
Jaffray J.Y. (1992) Bayesian updating and belief functions. IEEE Trans. on Systems, Man and Cybernetics, 22, 1144-1152.
Jeffrey R. (1965) The Logic of Decision. McGraw-Hill, New York. 2nd edition, University of Chicago Press, 1983.
Katsuno H., Mendelzon A.O. (1991) On the difference between updating a knowledge base and revising it. Proc. of the 2nd Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'91) Cambridge, MA, April 22-25, 387-394.
Kraus S., Lehmann D., Magidor M. (1990) Nonmonotonic reasoning, preferential models and cumulative logics. Artificial Intelligence, 44, 167-207.
Kyburg H.E., Jr. (1987) Bayesian and non-Bayesian evidential updating. Artificial Intelligence, 31, 271-293.
Lehmann D., Magidor M. (1992) What does a conditional knowledge base entail? Artificial Intelligence, 55(1), 1-60.
Levi (1980) The Enterprize of Knowledge. The MIT Press, Cambridge, MA.
Makinson D., Gärdenfors P. (1991) Relations between the logic of theory change and nonmonotonic logic. The Logic of Theory Change (Proc. of the Workshop, Konstanz, Germany, Oct. 1989) (A. Fuhrmann, M. Morreau, eds.), Lecture Notes in Artificial Intelligence, Vol. 465, Springer Verlag, Berlin, 185-205.
Moral S., De Campos L. (1991) Updating uncertain information. In: Uncertainty in Knowledge Bases (B. Bouchon-Meunier, R.R. Yager, L.A. Zadeh, eds.), Springer Verlag, Berlin, 58-67.
Shafer G. (1976) A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ.
Walley P. (1996) Measures of uncertainty in expert systems. Artificial Intelligence, 83, 1-58.
Williams M.A. (1994) Transmutations of knowledge systems. Proc. of the 4th Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'94), Bonn, Germany, 619-632.
Williams P.M. (1980) Bayesian conditionalization and the principle of minimum information. British J. for the Philosophy of Sciences, 31, 131-144.