Belief revision and updates in numerical formalisms —An overview, with new results for the possibilistic framework—

Didier DUBOIS - Henri PRADE
Institut de Recherche en Informatique de Toulouse (I.R.I.T.), Universite Paul Sabatier - CN.R.S.
118 route de Narbonne, 31062 TOULOUSE Cedex (FRANCE)

Abstract

The difference between Bayesian conditioning and Lewis' imaging is somewhat similar to the one between Gardenfors' belief revision and Katsuno and Mendelzon' updating in the logical framework. Counterparts in possibility theory of these two operations are presented, including the case of conditioning upon an uncertain observation. Possibilistic conditioning satisfies all the postulates for belief revision, and possibilistic imaging all the updating postulates. Lastly, a third operation called "focusing" is naturally introduced in the setting of belief and plausibility functions.

1 Introduction

Numerical formalisms for the representation of uncertainty usually describe stales of knowledge in terms of possible states of the world $\omega \in \Omega$. These states w are supposed to be mutually exclusive and usually Ω is assumed to gather all the possible states of the world. Both in probability theory and in possibility theory, to each state w is attached a degree $d(\omega) \in [0,1]$ which estimates the extent to which co may represent the real state of the world. These states can be put in correspondence with the models used in logical formalisms. By convention, $d(\omega)=0$ means that we are completely certain that co cannot be the real slate of the world. But the meaning of $d(\omega)=1$ is completely different in probability theory where it means that co is *the* real state (complete knowledge), and in possibility theory where it only expresses that nothing prevents co from being the real state of the world.

In these two formalisms, the change of the current state of knowledge (called 'epistemic state' in the following), upon the arrival of a new information stating that the real world is in $A \subseteq \Omega$ corresponds to a modification of the assignment function d into a new assignment d\ This change should obey general principles which guarantee that i) d' is of the same nature as d (preservation of the representation principles); ii) $\bar{A}$, which denotes 'not A', is excluded by d', i.e., $\forall \omega \notin A, d'(\omega)=0$ is observed is held as certain after the revision or the updating); iii) some informational distance between d' and d is minimized (principle of minimal change). Counterparts to these principles are also at the basis of revision and updating in logical formalisms [12].

The probabilistic framework offers at least two ways of modifying a probability distribution upon the arrival of a new and certain information: the Bayesian conditioning, but also D. Lewis [21]'s 'imaging' which consists in translating the weights originally on models outside A to models which are their closest neighbours in A. This paper shows that the existence of these two modes, which can be also defined in the possibilistic framework, is analogous to the distinction between belief revision based on Alchourr6n, Gardenfors and Makinson (AGM) postulates [12] and updating based on Katsuno and Mendelzon [18]' postulates.

The paper is organized in three main parts. The next section surveys basic results on probabilistic conditioning and imaging. Section 3 introduces these two operations in the possibilistic framework and provides new results and justifications for them. Section 4 briefly considers belief and plausibility functions and then emphasizes the existence of a third operation, called 'focusing', in this setting.

2 The Probabilistic Framework

In this setting, an epistemic state is represented by a probability measure P on the set Ω of possible worlds, (assumed finite for simplicity), such that

$$\forall A,B \subseteq \Omega,\ A \cap B \neq \emptyset \Rightarrow P(A \cup B)=P(A)+P(B).$$

Complete knowledge is represented by $P(\{\omega_0\})=1$ for some ω_0. An epistemic state is characterized by the distribution $\{p(\omega), \omega \in \Omega\}$ from which $P(A)=\sum_{\omega \in A} p(\omega)$ can be computed.

2.1 Probabilistic Conditioning

Upon learning that event A has occurred, i.e. we are certain that A is true, the a priori epistemic state P is revised by Bayes theorem into the conditional probability

$$P(B|A)=P(B \cap A)/P(A)=P(A|B)\cdot P(B)/P(A) \quad (1)$$

where P(B) is changed into P(B|A). It is not defined if A is a priori judged to be impossible. Minimal requirements are satisfied by (1); namely i) $P(A|A)=1$, when $P(A)>0$ (priority to the input information); ii) $P(A)=1 \Rightarrow P(\cdot|A)=P$ (an already known input information does not modify the epistemic state). This form of revision via a scaling factor is not so natural (except if one considers that P(B|A) should be a relative frequency) and needs some justification. Several exist.

– Numerical justifications: (1) is the only possibility if for any $B,C \subseteq A$, $P(B|A)/P(C|A)$ should be equal to $P(B)/P(C)$, i.e. there is no *relative* change. Another justification is given in [12] where conditioning is proved to be the only rule such that:

$$A \cap A' = \emptyset \Rightarrow P(B|A \cup A')=\lambda P(B|A)+(1-\lambda)P(B|A') \quad (2)$$

where $\lambda=P(A)/P(A \cup A')=P(A)/(P(A)+P(A'))$. This result assumes that $P(A|A)=1$, and $P(\cdot|\Omega)=P$.

– Algebraic justification: Cox [1] and his followers (e.g. [14]) have proved that Bayes rule, as well as probability theory itself can be justified through consistency between the Boolean structure of the subsets of Ω and three simple axioms for the measure of uncertainty g, and the conditioned one $g(\cdot|A)$, i.e. ($\bar{A}$ is the complement of A)

a) $g(B \cap A) = f(g(B|A), g(A))$; b) $g(\bar{A}) = s(g(A))$, $\forall A \subseteq \Omega$;

c) f is a continuous *strictly* monotonic function in both places; s is a continuous, strictly decreasing function.

Then g should be a probability measure and f the product. This is more a justification of conditional probability than of revision itself. See [14] for a Coxian-like justification of conditional probability as actually performing a revision.

– Information-theoretic justification: the information content of a probabilistic epistemic state is Shannon entropy $S(P) = -\sum_{\omega \in \Omega} p(\omega) \mathrm{Log}\, p(\omega)$. The less ambiguous P, the smaller S(P); particularly S(P)=0 if P is a complete epistemic state and S(P) is maximal iff $p(\omega)=p(\omega')$, $\forall \omega, \omega' \in \Omega$. Kulback & Leibler' informational distance is then defined by

$$I(P,P') = +\sum_{\omega \in \Omega} p'(\omega) \mathrm{Log}(p'(\omega)/p(\omega)) \quad (3)$$

The conditional probability $P'=P(\cdot|A)$ minimizes $I(P,P')$ under the constraint $P'(A)=1$, see [29].

Note that the understanding of Bayesian conditioning as a revision process (as above) is challenged by many Bayesians who view conditioning as a mere change of reference class reflecting the available evidence, i.e. a "focusing process". However because of the existence of a single conditioning rule, there is no way to distinguish between both views.

2.2 Probabilistic Imaging

Another path in the problem of updating probabilities is the one followed by Lewis [21]. Assume that Ω is such that for any world $\omega \in \Omega$, and for any set $A \subseteq \Omega$, $\exists \omega_A$ such that ω_A is the closest world from ω, that belongs to A. Then the principle of minimal change can be expressed as an advice to move probability weights as little as possible away from the worlds that become impossible upon learning that some event $A \subseteq \Omega$ has occurred. This is formally expressed as

$$\forall \omega' \in A,\ p_A(\omega') = \sum_{\omega'=\omega_A} p(\omega). \quad (4)$$

This rule is called 'imaging' because p_A is the image of p on A obtained by moving the masses $p(\omega)$ for $\omega \notin A$ to $\omega_A \in A$, with the natural convention that $\omega_A = \omega$ if $\omega \in A$. This rule comes from the study of conditional logics. It has been generalized by Gärdenfors [12] to the case when the set of worlds in A closest to a given world ω contains more than one element. If $A(\omega) \subseteq A$ is the subset of closest worlds from ω, $p(\omega)$ can be shared among the various worlds $\omega' \in A(\omega)$ instead of being allocated to a unique world. Gärdenfors has proved that general imaging is the only updating rule that is homomorphic, i.e. (P_A is the measure based on p_A)

$$(\lambda P + (1-\lambda)P')_A = \lambda P_A + (1-\lambda)P'_A. \quad (5)$$

It expresses invariance under convex combination. Imaging can turn impossible worlds into possible ones, e.g. one may have $p_A(\omega_A)>0$ while $p(\omega_A)=0$. As a consequence a sure fact B a priori, i.e. such that $P(B)=1$ may become uncertain, i.e. $P_A(B)<1$. This is not the case with Bayesian conditioning. However as with the Bayesian rule, $P(A)=1 \Rightarrow P_A=P$.

2.3 Illustrative Example (inspired by Morreau)

A box contains either an apple (a) or a banana (b). Let ω_1, $\omega_2, \omega_3, \omega_4$ denote the states where $a \wedge b$ is true, $a \wedge \neg b$ is true, $\neg a \wedge b$ is true, $\neg a \wedge \neg b$ is true respectively. Our epistemic state is represented by $p(\omega_1)=p(\omega_4)=0$, $p(\omega_2)>0$, $p(\omega_3)>0$. For instance $p(\omega_2)=0.7$, $p(\omega_3)=0.3$, i.e. an apple is more probably present than a banana in the box. Upon the occurrence of $A=\{\omega_3,\omega_4\}$ (no apple) Bayes rule yields $p(\omega_3|A)=1$; i.e. there is a banana in the box (Gärdenfors' revision leads to the same result in the logical setting). Let us now apply Lewis' imaging. The closest "neighbour" of ω_2 ($\notin A$) in A is ω_4 (both agree that b is false). Then moving $p(\omega_2)$ to $\omega_{2A}=\omega_4$ (and $\omega_{3A}=\omega_3$ since ω_3 is in A) gives the update $p_A(\omega_4)=0.7$; $p_A(\omega_3)=0.3$; i.e. the most probable situation is that the box is empty. This is in agreement with a reasoning by case: either the box was containing an apple or a banana; if the apple (if any) has been taken out of the box, either the box is now empty or there is still the banana. This agrees with Katsuno and Mendelzon [18]' approach to updating. In Bayes rule, A is understood as "there is no apple" (static world), while with imaging, A rather means "there is no longer any apple" (world change).

2.4 Uncertain Inputs

The Bayesian setting has been extended to the case of uncertain inputs. An uncertain piece of evidence corresponds to an event $A \subseteq \Omega$ along with a probability α that this event did happen. The updated probability measure $P(B|(A,\alpha))$ can be computed using Jeffrey [17]'s rule as

$$P(B|(A,\alpha)) = \alpha P(B|A) + (1-\alpha) P(B|\bar{A}) \quad (6)$$

where $P(B|A)$ and $P(B|\bar{A})$ are obtained by regular conditioning. In the case of a set of possible observations, one of which is the true one, which forms a partition $\{A_1,\ldots,A_n\}$ of Ω, (6) is extended into

$$P(B|\{(A_i,\alpha_i)\}_{i=1,n}) = \sum_{i=1,n} \alpha_i P(B|A_i) \quad (7)$$

where the probability that A_i is the actual observation is α_i, with $\sum \alpha_i = 1$. In a strict Bayesian view of (6) and (7), α_i is interpreted as a conditional probability $P(A_i|E)$ where E denotes the (sure) event underlying the uncertain information. Then (7) assumes that $P(B|A_i)=P(B|A_i \cap E)$, i.e. that for all A_i, E is independent of B in the context A_i (e.g. [23]). (6) and (7) have been justified by Williams [29] on the basis of the distance $I(P,P')$ under the constraints $P'(A_i)=\alpha_i$.

(7) can be also justified at the formal level by the fact that the only way of combining the conditional probabilities $P(B|A_i)$ in an eventwise manner (i.e. using the same combination law for all events B) is to use a linear weighted combination such as (7) [19]. Lastly, pushing (7) to the limit by assuming $\Omega=\{\omega_1,\ldots,\omega_n\}$ and choosing the finest partition $A_i=\{\omega_i\}$, and then letting $\alpha_i=P_2(\{\omega_i\})$, Jeffrey's rule (7) comes down to a simple substitution of P_1 by P_2. The new piece of evidence totally destroys the epistemic state. Priority is given to the new information.

3 The Possibilistic Framework

A possibility distribution π is a mapping from Ω to a totally ordered set V containing a greatest element (denoted 1) and a least element (denoted 0), e.g. V=[0,1]. However any finite, or infinite and bounded, chain will do as well. A consistent epistemic state π is such that $\pi(\omega)=1$ for some ω, i.e. at least one of the worlds is considered as completely possible in Ω. Let π and π' be two possibility distributions on Ω describing epistemic states. If $\pi \leq \pi'$, π is said to be more specific than π' [30], i.e. the epistemic state described by π is more complete, contains more information than the one described by π'. If $\exists \omega_0 \in \Omega$, $\pi(\omega_0)=1$, and $\pi(\omega)=0$ for

$\omega \neq \omega_0$, π corresponds to a *complete* epistemic state.

Interpreted in this framework, the three basic forms of belief dynamics described in [12], namely expansion, contraction and revision can easily be depicted. The result of an expansion, that stems from receiving new information consistent with a previously available epistemic state described by π, is another possibility distribution π' that is more specific than π. Note that π', by definition, is also such that $\pi'(\omega)=1$. Hence if we let $C(\pi)=\{\omega | \pi(\omega)=1\}$ be the core of π (i.e. the set of preferred worlds in a given epistemic state), we have $C(\pi')\neq\emptyset$ and $C(\pi')\subseteq C(\pi)$. A contraction, i.e. the result of forgetting some piece of information among those that form an epistemic state, will be expressed by going from π to a less specific possibility distribution $\pi'\geq\pi$. The term revision will be interpreted as any other belief change which is neither a contraction nor an expansion. Namely, it is when from π, we reach π' where neither $\pi\geq\pi'$ nor $\pi'\leq\pi$ hold. More specifically we may speak of "strict" revision when $C(\pi)\cap C(\pi')=\emptyset$.

Similarly to the probabilistic case, a possibility distribution generates a set function Π called a possibility measure [31] defined by (for simplicity V=[0,1])

$$\Pi(A)=\max_{\omega\in A} \pi(\omega) \quad (8)$$

and satisfying $\Pi(A\cup B) = \max(\Pi(A),\Pi(B))$ as a basic axiom. $\Pi(A)$ evaluates to what extent the subset A of possible worlds is consistent with the epistemic state π. $\Pi(A)=0$ indicates that A is impossible. $\Pi(A)=1$ only means that A is consistent with π, and it may happen that $\Pi(A)=\Pi(\bar{A})=1$ ($\bar{A}$ is the complement of A), in which case, it expresses ignorance about A. The degree of certainty of A is measured by means of the necessity function $N(A)=1-\Pi(\bar{A})$, whose characteristic axiom is $N(A\cap B)=\min(N(A),N(B))$. A is considered as a sure fact in an epistemic state π whenever $N(A)=1$. If $S(\pi)$ is the support $\{\omega | \pi(\omega)>0\}$ of π, then $N(A)=1 \Leftrightarrow S(\pi)\subseteq A$, while $N(A)>0 \Leftrightarrow C(\pi)\subseteq A$ (Ω finite) means that A is credible.

3.1 Possibilistic Conditioning

If the new information A is such that $\Pi(A)=1$, i.e. A is consistent with the epistemic state π, π is expanded into π'

$$\forall\omega,\ \pi'(\omega)=\pi^+_A(\omega)=\min(\mu_A(\omega),\pi(\omega)) \quad (9)$$

where μ_A is the characteristic function of the subset A. When $\Pi(A)\neq 1$, π is revised into π' defined through conditionalization. Revision in possibility theory is performed by means of a conditioning device similar to the probabilistic one, obeying an equation of the form [15]

$$\forall B,\ \Pi(A\cap B)=\Pi(B|A)*\Pi(A),\ \text{if } \Pi(A)>0. \quad (10)$$

Possible choices for * are min and the product [4]. In case of *=min, choosing the least specific solution (i.e. the solution with the greatest possibility degrees in agreement with the constraint (10)) yields,

$$\Pi(B|A) = 1 \text{ if } \Pi(A\cap B)=\Pi(A)>0$$
$$= \Pi(A\cap B) \text{ otherwise.}$$

In particular $\Pi(B|A)=0$ if $A\cap B=\emptyset$. The conditional necessity function is defined by $N(B|A)=1-\Pi(\bar{B}|A)$, by duality. The possibility distribution underlying the conditional possibility measure $\Pi(\cdot|A)$ is defined by

$$\pi(\omega|A) = 0 \text{ if } \omega\notin A\ ;\ \pi(\omega|A) = 1 \text{ if } \pi(\omega)=\Pi(A),\ \omega\in A\ ;$$
$$\pi(\omega|A) = \pi(\omega) \text{ if } \pi(\omega)<\Pi(A),\ \omega\in A \quad (11)$$

When *=product, the corresponding expression is

$$\forall B,\ \Pi(B|A) = \Pi(A\cap B)/\Pi(A) \quad (12)$$

provided that $\Pi(A)\neq\emptyset$. This is Dempster rule of conditioning, specialized to possibility measures, i.e. consonant plausibility measures of Shafer [25]. The corresponding revised possibility distribution is

$$\pi(\omega|A)=\pi(\omega)/\Pi(A),\ \forall\omega\in A\ ;\ \pi(\omega|A)=0 \text{ otherwise.} \quad (13)$$

This rule is much closer to Bayesian conditioning than the ordinal rule (11) which is purely based on comparing numbers; (12) requires more of the structure of the unit interval (a product operation). In both cases the set function itself remains ordering-based. (12) makes sense if Ω is a continuous universe, if π is continuous and revision is required to preserve continuity. Both (11) and (13) satisfy

- $N(A)=1 \Rightarrow \pi(\cdot|A)=\pi$ (no revision if A was already certain)
- $\Pi(A)=1 \Rightarrow \pi(\cdot|A)=\min(\mu_A,\pi)=\pi^+_A$ (revision=expansion in case of consistency)
- $N(A|A)=1$ (priority to the new information).

Counterparts of the other AGM [12] postulates for revision hold as well with the two definitions; see [9] for a detailed study of the possibilistic counterparts of expansion, revision, and contraction. Especially (11) embodies a principle of minimal change. If π and π' define two real-valued possibility distributions on a finite Ω then the Hamming distance between π and π' is defined by $H(\pi,\pi')= \sum_{\omega\in\Omega} |\pi(\omega)-\pi'(\omega)|$. Then we have the following result [9]: $\pi(\cdot|A)$ is the possibility distribution the closest to π that complies with counterparts of Gärdenfors' postulates, as long as there is a single world ω_A where $\pi(\omega_A)=\Pi(A)$. $H(\pi;\pi(\cdot|A))$ is thus minimal under the constraint $N(A|A)=1$.

It is worth noticing that the revision rule (11) satisfies the counterpart of (2) for probabilistic conditioning with respect to the disjunction. It can be checked that

$$\Pi(B|A\cup A')=\max(\min(\alpha,\Pi(B|A)),\min(\alpha',\Pi(B|A'))) \quad (14)$$

with $\alpha=\Pi(A')\rightarrow\Pi(A)$ and $\alpha'=\Pi(A)\rightarrow\Pi(A')$ where $\rightarrow$ is the multiple-valued implication $a\rightarrow b=1$ if $a\leq b$ and b otherwise. The function from $[0,1]^2$ to $[0,1]$ defined by $M_{\alpha,\alpha'}(x,y)= \max(\min(\alpha,x),\min(\alpha',y))$, where $\max(\alpha,\alpha')=1$ is the possibilistic counterpart of the weighted arithmetic mean (or convex mixture) in probability theory. Condition $\max(\alpha,\alpha')=1$ is indeed verified in (14). The behavior of the product-based definitions of $\Pi(B|A)$ with respect to a disjunction of input is similar to the min-based definition: (14) remains true when $A\cap A'=\emptyset$ provided that we define $\alpha= \Pi(A')\rightarrow\Pi(A)$ and $\alpha'=\Pi(A)\rightarrow\Pi(A')$ with the implication $a\rightarrow b=\min(1,b/a)$ and min is changed into product.

While (13) looks as a more natural counterpart of probabilistic conditioning, N(B|A) stemming from (11) is more closely akin to the concept of "would counterfactual" following Lewis [20], and denoted A□→B, which is intended to mean "if it were the case that A, then it would be the case that B". Lewis proposes to consider A□→B as true in world ω if and only if some accessible world in $A\cap B$ is closer to ω than any world in $A\cap\bar{B}$, if there are worlds in A. Let us interpret "closer to world ω" as "preferred" in the sense of possibility degrees (ω thus denote the "ideal world"). Let us notice that we do have when $\Pi(A)>0$,

$$\forall B, B\cap A\neq\emptyset, N(B|A)>0 \Leftrightarrow \Pi(B\cap A)>\Pi(\bar{B}\cap A) \quad (15)$$

where $N(\cdot|A)$ is the necessity measure based on $\pi(\cdot|A)$. The latter inequality means that there is a world in $B\cap A$ which is more possible than any world in $\bar{B}\cap A$. Hence $N(B|A)>0$ agrees with the truth of A□→B. The counterpart of Lewis' "might conditional" $A \Diamond\!\!\rightarrow B$ is of course $\Pi(B|A)$ in the sense of (10) with $*=\min$.

3.2 Possibilistic Imaging

It is easy to envisage the possibilistic counterpart to Lewis' imaging since this type of belief change is based on mapping each possible world to the closest one that accomodates the input information. As in Sec. 2, define for any $\omega \in \Omega$, and non-empty set $A \subseteq \Omega$ the closest world $\omega_A \in A$ to ω. Then the image of an epistemic state π in A is

$$\pi^{\circ}{}_A(\omega') = \max_{\omega' = \omega_A} \pi(\omega) \text{ if } \omega' \in A;\ \pi^{\circ}{}_A(\omega') = 0 \text{ if } \omega' \notin A. \quad (16)$$

If there is more than one world ω_A closest to ω, then the weight $\pi(\omega)$ can be allocated to each of the closest worlds, forming the set $A(\omega)$, and the above updating rule becomes

$$\pi^{\circ}{}_A(\omega') = \max_{\substack{\omega \\ \omega' \in A(\omega)}} \pi(\omega) \text{ if } \omega' \in A;\quad \pi^{\circ}{}_A(\omega') = 0 \text{ if } \omega' \notin A. \quad (17)$$

If we define R_A as the relation that to each ω assigns its closest neighbours in A, the above update formula is nothing but Zadeh [31]'s extension principle that characterizes the fuzzy image of the fuzzy set whose membership function is π, via relation R_A.

It is easy to check that the updating rule (16) satisfies all postulates of Katsuno and Mendelson [18]'s updates namely:

U1) $\pi^{\circ}{}_A \le \mu_A$; U2) $\pi \le \mu_A \Rightarrow \pi^{\circ}{}_A = \pi$;
U3) if $A \ne \emptyset$ and π is normalized then $\pi^{\circ}{}_A$ is normalized ;
U4) $A=B \Rightarrow \pi^{\circ}{}_A = \pi^{\circ}{}_B$; U5) $\min(\pi^{\circ}{}_A, \mu_B) \le \pi^{\circ}{}_{A \cap B}$;
U6) $\pi^{\circ}{}_A \le \mu_B, \pi^{\circ}{}_B \le \mu_A \Rightarrow \pi^{\circ}{}_A = \pi^{\circ}{}_B$;
U7) π maximally specific $\Rightarrow \min(\pi^{\circ}{}_A, \pi^{\circ}{}_B) \le \pi^{\circ}{}_{A \cup B}$;
U8) $[\max(\pi, \pi')]^{\circ}{}_A = \max(\pi^{\circ}{}_A, \pi'^{\circ}{}_A)$.

Defining $A(\omega)$ precisely as $\{\omega' | \pi(\omega') = \Pi(A) = \max\{\pi(\omega), \omega \in A\}\}$ then $\pi^{\circ}{}_A = \pi(\cdot|A)$, i.e. we recover the revision based on conditioning. Clearly in this setting, we see that Katsuno and Mendelzon's approach subsumes the AGM framework. In the setting of possibility theory, the difference between Katsuno and Mendelson framework and the AGM approach are patent. While the AGM approach tries to use the input information so as to reduce incompleteness, the update view tries to carry the incompleteness of the epistemic state to a new epistemic state that agrees with the input information, assuming that the shift of the "real world" is minimal.

When $A(\omega)$ is a fuzzy subset of A, (17) reads

$$\pi^{\circ}{}_A(\omega') = \max_{\omega} \pi(\omega) * \mu_{A(\omega)}(\omega') \text{ if } \omega' \in A;\ \pi^{\circ}{}_A(\omega') = 0 \text{ if } \omega' \notin A$$

where $*=\min$ or product; we have to assume $\mu_{R_A}(\omega, \omega') = 1$ if $\omega = \omega' \in \mathrm{Supp}(A) = \{\omega \in A, \mu_A(\omega) > 0\}$, and 0 if $\omega' \notin \mathrm{Supp}(A)$. It can be easily checked that this updating is invariant under weighted max-combination, i.e., the counterpart of (5) reads:

$$[\max_i (\lambda_i * \pi_i)]^{\circ}{}_A = \max_i \lambda_i * (\pi_i)^{\circ}{}_A \text{ with } \max_i \lambda_i = 1.$$

Clearly U8 is a particular case of this invariance under max-weighted combination.

The possibilistic framework would enable us to deal with the apple and banana example in a way similar to the probabilistic solution, although in a more qualitative way. The example is really too elementary to exhibit significant differences between the two approaches.

3.3 Uncertain Inputs

Belief change can be extended to uncertain inputs of the form $N(A)=\alpha$. The main question here is how to interpret such an uncertain input. Two interpretations make sense [9]:

i) $N(A)=\alpha$ is taken as a constraint that the new epistemic state must satisfy; it means that if π' is obtained by revising π with information (A, α), the resulting necessity measure N' must be such that $N'(A)=\alpha$;
ii) $N(A)=\alpha$ is interpreted as an extra piece of information that may be useful or not useful to *refine* the current epistemic state; in that case α is viewed as a degree of reliability or priority of information A.

Interpretation i) is in the spirit of Jeffrey's rule. Clearly $N(A)=1$ will lead to an expansion of π into $\pi^{+}{}_A$ or a revision $\pi(\cdot|A)$, while $N(A)=0$ will force a contraction. In contrast, ii) corresponds to either a revision or an expansion but is never a contraction, since if α is too low, the input information will be discarded.

The input information is not modelled in the same way whether it is a constraint or an additional information. In the first case (i), $N(A)=\alpha$ is interpreted as $\Pi(A)=1$ and $\Pi(\bar{A}) = 1-\alpha$, and the belief change rule is of the form [9]

$$\begin{aligned} \pi(\omega|(A,\alpha)) &= \pi(\omega|A) \text{ if } \omega \in A \\ &= (1-\alpha) * \pi(\omega|\bar{A}) \text{ if } \omega \in \bar{A} \end{aligned} \quad (18)$$

where $*=\min$ or product according to whether $\pi(\omega|A)$ is the ordinal or Bayesian-like revised possibility distribution. Note that when $\alpha=1$, $\pi(\omega|(A,\alpha)) = \pi(\omega|A)$, but when $\alpha=0$, we obtain a possibility distribution less specific than π such that the associated necessity of A is zero.

In the second case (ii), the additional information $N(A)=\alpha$ is represented by a fuzzy set F with membership function μ_F

$$\mu_F(\omega) = 1 \text{ if } \omega \in A\ ;\ \mu_F(\omega) = 1-\alpha \text{ otherwise.}$$

Letting $F_\lambda = \{\omega | \mu_F(\omega) \ge \lambda\}$, each F_λ is viewed as the (non-fuzzy) regular input information underlying F, with guaranteed possibility λ and the revised epistemic state $\pi(\cdot|F)$ is defined by analogy with Jeffrey's rule as [6]

$$\pi(\omega|F) = \max_{\lambda \in (0,1]} \lambda * \pi(\omega|F_\lambda)$$

where the convex mixing is changed into the weighted maximum and $*$ is min or product again. In our particular case, it gives, for $\alpha > 0$

$$\pi(\omega|F) = \pi(\omega|A) \text{ if } \omega \in A;\ \pi(\omega|F) = \pi(\omega) * (1-\alpha) \text{ if } \omega \in \bar{A}. \quad (19)$$

Note that $\pi(\omega|F) \le \mu_F = \max(\mu_A, 1-\alpha)$; moreover $\pi(\omega|F) = \pi(\omega)$ if $\alpha=0$ since then $F=\Omega$, i.e. the operation is never a contraction. This behavior is very different from the case when $N(A) = \alpha$ is taken as a constraint.

The first revision rule (18) under uncertain inputs can be extended to a set of constraints $\Pi(A_i)=\lambda_i$, $i=1,n$, where $\{A_i, i=1,n\}$ forms a partition of Ω, and it gives

$$\pi(\omega|\{(A_i, \lambda_i)\}) = \lambda_i * \pi(\omega|A_i),\ \forall \omega \in A_i \quad (20)$$

where $*=$minimum or product whether $\pi(\omega|A_i)$ is ordinal or Bayesian-like. In the limit case when $A_i=\{\omega_i\}$, $\forall i$, the input is equivalent to a fuzzy input F with $\mu_F(\omega_i)=\lambda_i$. The above belief change rule reduces to a simple substitution of π by μ_F, just as for Jeffrey's rule for probabilities.

To conclude, while the belief change rule (19) is formally analogous to Jeffrey's rule, its behavior is very much akin to a revision à la Gärdenfors. See Section 3.4. On the contrary the other rule (18)-(20) is very close to the spirit of Jeffrey's rule, and has been proposed in another setting by Spohn [27] who uses the integers as a scale rather than [0,1] with the convention that 0 corresponds to the minimum impossibility (i.e. the maximal possibility), see [6].

3.4 Epistemic States as Weighted Propositions

The set of possible worlds in which a possibilistic logic formula $(\varphi\ \alpha)$ is true is the fuzzy set $[\varphi\ \alpha]$ on Ω given by [7]

$\mu_{[\varphi\ \alpha]}(\omega) = 1$ if $\omega \in [\varphi]$; $\mu_{[\varphi\ \alpha]}(\omega) = 1-\alpha$ otherwise ;

($[\varphi]$ is the set of possible worlds where φ is true); $\mu_{[\varphi\ \alpha]}$ is the least specific possibility distribution π such that $N([\varphi]) = \inf_{\omega \notin [\varphi]} 1-\pi(\omega) \geq \alpha$. The fuzzy set of worlds which satisfy a possibilistic knowledge base $K=\{(\varphi_i,\alpha_i), i=1,m\}$ is defined by the possibility distribution π_K

$$\forall \omega \in \Omega,\ \pi_K(\omega) = \min_{i=1,m} \max(\mu_{[\varphi_i]}(\omega), 1-\alpha_i) \quad (21)$$

which extends $[K]=[\varphi_1] \cap [\varphi_2] \cap \ldots \cap [\varphi_n]$ from a set of sentences to a set of weighted sentences. Semantic entailment is defined in terms of specificity ordering ($\pi \leq \pi'$). Namely $K \models (\varphi\ \alpha)$ if and only if $\pi_K \leq \max(\mu_{[\varphi]}, 1-\alpha)$. This notion of semantic entailment is exactly the one of Zadeh [31]. Note that here each possibilistic formula is viewed as an additional piece of information. Indeed $\min(\pi, \max(\mu_{[\varphi]}, 1-\alpha)) < \pi$ only if (φ,α) brings information not deducible from π. These definitions make sense for consistent possibilistic knowledge bases, i.e. such that $\pi_K(\omega)=1$ for some $\omega \in \Omega$. Consistency of K is equivalent to the consistency of the classical knowledge base K^* obtained by removing the weights. When $\max_{\omega \in \Omega} \pi(\omega)=\gamma<1$, K is said to be partially inconsistent, γ being the degree of consistency of K, i.e. $\text{inc}(K)=1-\gamma$. It $\gamma=0$, K is completely inconsistent.

Let us consider the case when K is consistent but $K'=K \cup \{(\varphi\ 1)\}$ is not, and let $\alpha=\text{inc}(K \cup \{(\varphi\ 1)\})>0$. The following identity is proved in [7]:

$K \cup \{(\varphi\ 1)\} \vdash (\psi\ \beta)$ with $\beta>\alpha$ if and only if $N(\psi|\varphi)>0$

where $N(\psi|\varphi)$ is the necessity measure induced from $\pi(\cdot|[\varphi])$ the possibility distribution expressing the content of K, revised with respect to the set of models of φ. Indeed let π' be the possibility distribution on Ω induced by K' then

$\pi' = \min(\pi_K, \mu_{[\varphi]})$ and $0 < \max_{\omega \in \Omega} \pi'(\omega) = 1-\alpha < 1$.

The possibility distribution $\tilde{\pi}'$ induced from the consistent part of K' made of sentences whose weights is higher than α, is defined as

$$\tilde{\pi}'(\omega) = \begin{cases} \pi(\omega) & \text{if } \omega \in [\varphi] \text{ and } \pi'(\omega)<1-\alpha \\ 1 & \text{if } \omega \in [\varphi] \text{ and } \pi(\omega)=1-\alpha \\ \pi'(\omega)=0 & \text{otherwise.} \end{cases}$$

Hence $\tilde{\pi}'=\pi(\cdot|[\varphi])$, the result of revising π by $[\varphi]$. The corresponding revision is rather drastic since all sentences $(\varphi_i\ \alpha_i)$ with weights $\alpha_i \leq \alpha$ are thrown away, and replaced by $(\varphi\ 1)$. Note that when $\Pi(\varphi)>0$, $N(\psi|\varphi)>0$ is equivalent to $N(\neg\varphi \vee \psi) > N(\neg\varphi \vee \neg\psi)$, or equivalently in terms of the associate epistemic entrenchment $>_c$: $\neg\varphi \vee \psi >_c \neg\varphi \vee \neg\psi$ [5], and corresponds to a characteristic condition for having ψ in the (ordered) belief set obtained by revising the deductive closure of K by φ, in the sense of Gärdenfors [12].

4 Revision, Updating and Focusing

Let us now consider Shafer's evidence theory [25][26]. The set of possible worlds is called frame of discernment. In this framework the available knowledge is represented in terms of a basic probability assignment m, which is a set function from the set of subsets 2^Ω to [0,1] with the constraints $m(\emptyset)=0$ and $\sum_A m(A)=1$. The subsets $A \subseteq \Omega$ such that $m(A)>0$ are called focal elements. Note that there is no constraint on the structure of the set $\mathcal{F}$ of focal elements (here supposed to be finite and which does not make a partition in general). Each focal element A_i represents the most accurate description, with certainty $m(A_i)$, of the available evidence pertaining to the location of the actual world in Ω. The subsets A_i are the possible realizations of an imprecise observation pervaded with uncertainty. Due to the incompleteness of the available information, A_i is not necessarily a singleton. A plausibility function Pl and a belief function Bel are bijectively associated with m [25] and are defined by

$$Pl(B) = \sum_{A: A\cap B \neq \emptyset} m(A) \quad (22)$$

$$Bel(B) = 1-Pl(\bar{B}) = \sum_{\emptyset \neq A \subseteq B} m(A) \quad (23)$$

Dempster rule of conditioning is expressed by

$$Pl(B|A) = Pl(A\cap B)/Pl(A)\ ;\ Bel(B|A) = 1-Pl(\bar{B}|A) \quad (24)$$

This rule of conditioning can be justified on the basis of Cox's axiom that defines a conditional function associated to any uncertainty measure g defined on Ω as in Sec. 2 [4]. Cox's axiom justifies Dempster's conditioning rule as well as the geometric rule of conditioning [28]

$$Bel_g(B|A)=Bel(A \cap B)/Bel(A)\ ;\ Pl_g(B|A)=1-Bel_g(\bar{B}|A) \quad (25)$$

In terms of basic probability assignments, $Pl(\cdot|A)$ defined by (25) is obtained by transferring all masses m(B) over to $A \cap B$, followed by a normalization step, while $Bel_g(\cdot|A)$ is obtained by letting $m_g(B|A)=m(B)$ if $B \subseteq A$ and 0 otherwise, followed by normalization, i.e. a more drastic way of conditioning. Dempster's rule of conditioning looks more attractive from the point of view of updating since $Pl(B|A)$ is undefined only if $Pl(A)=0$ (i.e. A is impossible) while $Bel_g(B|A)$ is undefined as soon as $Bel(A)=0$ (i.e. A is unknown). This unability to revise with a vacuous prior is counterintuitive, with the geometric rule.

Dempster rule of conditioning is a mixture of AGM-type expansion (when m(B) carries over to $A \cap B$ if A becomes true) and Bayesian updating. On the contrary, the geometric rule has nothing to do with an expansion on Ω and is more in the spirit of imaging since all masses outside A are moved inside. Dempster rule of conditioning subsumes conditional possibility based on product, i.e. (12).

Another approach to conditioning has been proposed by De Campos et al. [3] and Fagin and Halpern [11] under the form

$$P^*(B|A) = Pl(A\cap B)/Pl(A\cap B)+Bel(A\cap \bar{B}) \quad (26)$$

$$P_*(B|A) = Bel(A\cap B)/Bel(A\cap B)+Pl(A\cap \bar{B}) \quad (27)$$

These definitions can be justified by interpreting belief and plausibility functions as lower and upper probabilities, since it has been proved that

$$P^*(B|A)=\sup\{P(B|A) | P \in \mathcal{P}(Bel)\} \quad (28)$$

$$P_*(B|A)=\inf\{P(B|A) | P \in \mathcal{P}(Bel)\} \quad (29)$$

where $\mathcal{P}(Bel)=\{P | Bel(B) \leq P(B) \leq Pl(B), \forall B\}$. These conditional functions are actually upper and lower conditional probabilities and have been considered by Dempster [2] himself and Ruspini [24]. $P_*(\cdot|A)$ has been proved to be still a belief function [11][16]. Although very satisfying from a probabilistic point of view, these definitions lead to a rather uninformative conditioning process since $P^*(\cdot|A) \geq Pl(\cdot|A) \geq Bel(\cdot|A) \geq P_*(\cdot|A)$ as noticed by Dempster [2]. Especially, complete ignorance is obtained ($P^*(B|A)=1, P_*(B|A)=0$) as soon as $Bel(A\cap \bar{B})=0$ and $Bel(A\cap B)=0$, i.e. as soon as the conditioning set A intersects each focal element without including any one of them, thus making all focal elements smaller. In that case the revising process would correspond to oblivion rather than learning. It has been shown elsewere that (28)-(29) is not a rule for revision but a "focusing rule", by which one only changes the reference class, *without forcing* $P(\bar{A})=0$ [8]. Especially (28)-(29) does not modify the constraints specified by the belief function. Dempster rule of conditioning comes down to add the constraint P(A)=1 to the set $\mathcal{P}$(Bel), in the case when Pl(A)=1, i.e. A is viewed as a new piece of information to be integrated in the current knowledge, and leads to a revision, and not only a change of reference class. When Pl(A)≠1, the constraint P(A)=Pl(A) can be added to the set of constraints in (28)-(29), and the upper conditional probability thus obtained corresponds to Dempster rule in the general case, as proved by Gilboa and Schmeidler [13], who call it maximum likelihood revision. A more refined proposal can be found in [22] where the distributions which do not maximize P(A) are also somewhat taken into account.

5 Conclusion

This paper has emphasized that belief revision in the sense of Gardenfors, as well as updating in the sense of Katsuno and Mendelzon can be defined through conditioning and imaging respectively both in the probabilistic and in the possibilistic settings. The possibilistic framework leads to a more complete agreement with the two sets of postulates (first stated for propositional logic) than the probabilistic setting. The paper also has tried to relate the revision of a possibility distribution on a set of possible worlds to the revision of a knowledge base made of uncertain logical formulas. More work is needed to relate probabilistic rules to the axiomatic approaches to belief change in the logical framework, despite the existing bridges between probability and possibility theories. Namely we might consider devising revision and updating rules in logics of uncertainty different from possibilistic logic, and especially probabilistic logic. Indeed while the problem of change has been thoroughly studied for probabilistic representations of epistemic states on a set of possible worlds, nothing has been done at the syntactic level. Besides the justification of the different rules in evidence theory is in its infancy. The idea of focusing, i.e. changing the reference class as opposed to revising a body of knowledge might be worth introducing in the logical setting also.

The coherence between numerical versus symbolic approaches to knowledge representation is still present in the revising and updating problems. Pushing further the consequences of such a coherence looks like a challenging task.

The reader is referred to a more complete version of this paper for further discussions and proofs [10].

Acknowledgements: This work has been partially supported by the European ESPRIT Basic Research Action n° 6156 entitled "Defeasible Reasoning and Uncertainty Management Systems (DRUMS-2)".

References

1 Cox R.T. Probability, frequency and reasonable expectation. American J. of Physics 14:1-13, 1946.
2 Dempster A.P. Upper and lower probabilities induced by a multivalued mapping. Ann.Math.Stat. 38:325-339, 1967.
3 De Campos L.M.. Lamata M.T., Moral S. The concept of conditional fuzzy measure. Int. J. Intelligent Systems 237-246, 1990.
4 Dubois D., Prade H. The logical view of conditioning and its application to possibility and evidence theories. Int. J. Approximate Reasoning 4:23-46, 1990.
5 Dubois D., Prade H. Epistemic entrenchment and possibilistic logic. Artif. Intelligence 50:223-239, 1991.
6 Dubois D., Prade H. Updating with belief functions, ordinal conditional functions and possibility measures. In: Uncertainty in Artificial Intelligence 6 (P.P. Bonissone et al., eds.), North-Holland, 311-329, 1991.
7 Dubois D., Prade H. Possibilistic logic, preferential models, non-monotonicity and related issues. IJCAI'91, 419-424.
8 Dubois D., Prade H. Evidence, knowledge and belief functions. Int. J. Approx. Reasoning 6:295-319, 1992.
9 Dubois D., Prade H. Belief change and possibility theory. In: Belief Revision (P. Gardenfors, ed.), Cambridge University Press, 142-182. 1992.
10 Dubois D., Prade H. A survey of belief revision and updating rules in numerical uncertainty models. Forthcoming,
11 Fagin R., Halpern J.Y. A new approach to updating beliefs. In: Research Report n° RJ 7222 (67989), IBM Research Division, San Jose, CA, 1989.
12 Gardenfors P. Knowledge in Flux: Modeling the Dynamics of Epistemic States. MIT Press, 1988.
13 Gilboa I., Schmeidler D. Updating ambiguous beliefs. J. of Economic Theory, TARK'92.
14 Heckerman D.E. An axiomatic framework for belief updates. In: Uncertainty in Artificial Intelligence 2 (J.F. Lemmer, L.N. Kanal, eds.). North-Holland, 11-22, 1988.
15 Hisdal E. Conditional possibilities - Independence and non-interactivity. Fuzzy Sets & Syst. 1:283-297, 1978.
16 Jaffray J.Y. Bayesian updating and belief functions. IEEE Trans, on Systems, Man and Cybern. 22:1144-1152, 1992.
17 Jeffrey R. The Logic of Decision. McGraw-Hill, 1965.
18 Katsuno H., Mendelzon A.O. On the difference between updating a knowledge base and revising it. KR'91, 387-394.
19 Lehrer K., Wagner C.G. Rational Consensus in Science and Society. D. Reidel Publ. Co., Boston, 1981.
20 Lewis D.K. Counterfactuals and comparative possibility. J. Philos. Logic 2:418-446, 1973.
21 Lewis D.K. Probabilities of conditionals and conditional probabilities. The Philosophical Rev. 85:297-315, 1976.
22 Moral S., De Campos L.M. Updating uncertain information. In: Uncertainty in Knowledge Bases (B. Bouchon-Meunier et al., eds.), Springer Verlag, 58-67, 1991.
23 Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan-Kaufmann,1988.
24 Ruspini E.H. Approximate deduction in single evidential bodies. Proc. 2nd Workshop on Uncertainty in Artificial Intelligence, Pennsylvania, 1986, 215-222.
25 Shafer G. A Mathematical Theory of Evidence. Princeton University Press, Princeton, 1976.
26 Smets P. Belief functions. In: Non-Standard Logics for Automated Reasoning (P. Smets et al., eds.), Academic Press, 253-286, 1988.
27 Spohn W. Ordinal conditional functions: a dynamic theory of epistemic states. In: Causation in Decision, Belief Change, and Statistics, Vol. 2 (W.L. Harper, B. Skyrms, eds.), D. Reidel, Dordrecht, 105-134, 1988.
28 Suppes P., Zanotti M. On using random relations to generate upper and lower probabilities. Synthese 36:427-440. 1977.
29 Williams P. Bayesian conditionalization and the principle of minimum information. British J. of the Philosophy of Science 31:131-144, 1980.
30 Yager R.R. Measuring tranquility and anxiety in decision-making: an application of fuzzy sets. Int. J. General Systems 8:139-146, 1982.
31 Zadeh L.A. Possibility theory and soft data analysis. In: Mathematical Frontiers of the Social and Policy Sciences (L. Cobb, R.M. Thrall, eds.). West View Press Boulder Co., 69-127, 1981.