Possibilistic and standard probabilistic semantics of conditional knowledge

Salem Benferhat, Didier Dubois and Henri Prade

Institut de Recherche en Informatique de Toulouse (IRIT) – CNRS
Université Paul Sabatier, 118 route de Narbonne
31062 Toulouse Cedex 4, France
Email: {benferha, dubois, prade@}irit.fr

Abstract

The authors have proposed in their previous works to view a set of default pieces of information of the form, "generally, from α_i deduce β_i", as the family of possibility distributions satisfying constraints expressing that the situations where $\alpha_i \wedge \beta_i$ is true are more possible than the situations where $\alpha_i \wedge \neg\beta_i$ is true. A representation theorem in terms of this semantics, for default reasoning obeying the System P of postulates proposed by Kraus, Lehmann and Magidor, has been obtained. This paper offers a detailed analysis of the structure of this family of possibility distributions by exploiting two different orderings between them: Yager's specificity ordering and a new refinement ordering. It is shown that from a representation point of view, it is sufficient to consider the subset of linear possibility distributions which corresponds to all the possible completions of the default knowledge in agreement with the constraints. There also exists a semantics for system P in terms of infinitesimal probabilities. Surprisingly, it is also shown that *a standard* probabilistic semantics can be equivalently given to System P, without referring to *infinitesimals,* by using a special family of probability measures, that two of the authors have called acceptance functions, and that has been also recently considered by Snow in that perspective.

1. Introduction

An important feature of human reasoning is its ability to draw plausible conclusions from available information (although this information is often incomplete). The conclusions drawn in this way may have then to be revised in the light of new information. Default information is made of rules of the form "generally, if α then β", where α and β are propositional formulas; these rules are then subject to exceptions. A typical example of a default information is "generally, birds fly". Kraus et al. (1990) have proposed a set of postulates, known as System P, for nonmonotonic consequence relations, that are commonly regarded as the minimal core of any "reasonable" default reasoning system. These postulates are recalled Section 2. Section 3 recalls the possibilistic semantics of default rules, which provides a representation of System P. The basic idea of possibility theory (Zadeh, 1978) is to represent incomplete information by means of a set of mutually exclusive situations equipped with an ordering relation expressing that some situations are more plausible than others. In general, this relation is encoded by associating to each interpretation ω a possibility level in a totally ordered scale, for instance a positive real number between 0 and 1, denoted by $\pi(\omega)$. Each default "generally, α's are β" is then viewed as a constraint expressing that the most plausible situation where $\alpha \wedge \beta$ is true has a greater plausibility than the most plausible one where $\alpha \wedge \neg\beta$ is true (Benferhat et al., 1992). This writes $\Pi(\alpha\wedge\beta) > \Pi(\alpha\wedge\neg\beta)$ in the possibility theory framework, where Π is a possibility measure. Given a set of defaults Δ, we thus construct a family $\mathbb{\Pi}(\Delta)$ of comparative possibility distributions that satisfy the constraints induced by Δ.

Sections 4 to 6 further investigate the possibilistic semantics and then relate it to a new, non-infinitesimal, probabilistic semantics of system P. This is because constraints of the form $\Pi(\alpha\wedge\beta)>\Pi(\alpha\wedge\neg\beta)$ can be equivalently written $P(\alpha\wedge\beta)>P(\alpha\wedge\neg\beta)$ for a special family of probability measures which is such that the inequality between the probabilities of two disjoint events is determined by the probability ordering between the most probable situations compatible with these events.

2. Postulates for Nonmonotonicity

By a conditional or default information we mean a rule of the form "generally, if α then β" having possibly some exceptions. This rule is denoted by "$\alpha \rightarrow \beta$" where $\rightarrow$ is a *non-classical* arrow relating two classical formulas. In the whole paper the arrow $\rightarrow$ has this non-classical meaning. A *default base* is a set $\Delta = \{\alpha_i \rightarrow \beta_i, i=1,\ldots,n\}$ of default rules.

Kraus et al. (1990) have characterized a nonmonotonic consequence relation by a set of rationality postulates, known as System P (P for "preferential"), which are composed of an axiom schema ($\alpha \mid\!\sim \alpha$) and 5 inference rules:

Left logical equivalence: from $\alpha \Leftrightarrow \alpha' = T$ and $\alpha \mid\!\sim \beta$ deduce $\alpha' \mid\!\sim \beta$
Right weakening: from $\beta \models \beta'$ and $\alpha \mid\!\sim \beta$ deduce $\alpha \mid\!\sim \beta'$
OR: from $\alpha \mid\!\sim \delta$ and $\beta \mid\!\sim \delta$ deduce $\alpha\vee\beta \mid\!\sim \delta$
Cautious monotony: from $\alpha \mid\!\sim \beta$ and $\alpha \mid\!\sim \delta$ deduce $\alpha\wedge\beta \mid\!\sim \delta$
Cut: from $\alpha\wedge\beta \mid\!\sim \delta$ and $\alpha \mid\!\sim \beta$ deduce $\alpha \mid\!\sim \delta$.

A syntactic conditional entailment, denoted by $\mid\!\sim_P$, from a set of conditional assertions Δ, can then be defined as done by Kraus, Lehmann and Magidor (1990). Namely, $\Delta \mid\!\sim_P \phi \rightarrow \psi$ if and only if $\phi \mid\!\sim \psi$ can be derived from Δ using $\phi \mid\!\sim \phi$ as an axiom schema and the inference rules of system P. It is generally considered in the nonmonotonic community that none of the conclusions obtained in the set $\Delta^P=\{\phi \rightarrow \psi / \Delta \mid\!\sim_P \phi \rightarrow \psi\}$ can be challenged in practice, while all the proposed super-sets of Δ^P contain conclusions which may be debatable in particular cases. Moreover the set $\{\psi / \Delta \mid\!\sim_P \phi \rightarrow \psi\}$ for some ϕ is deductively closed. A default base Δ is said to be *consistent* if and only if there is no formula ϕ such that $\{\psi / \Delta \mid\!\sim_P \phi \rightarrow \psi\}$ is classically inconsistent. Following an idea due to Makinson, Lehmann and Magidor (1992) have proposed to augment System P with the so-called *rational monotony* inference rule defined by: from $\alpha \mid\!\sim \delta$ and $\alpha \not\mid\!\sim \neg\beta$ deduce $\alpha \wedge \beta \mid\!\sim \delta$, where $\alpha \not\mid\!\sim \neg\beta$ means that $\alpha \mid\!\sim \neg\beta$ does not hold. This property states that in the absence of relevant information in the conditional knowledge base Δ expressing that the "α's are (not β)'s", one can deduce the same thing from α, or from $\alpha \wedge \beta$.

Definition 1. A consistent set of conditional assertions is said to be a rational extension of Δ, denoted by Δ^R, if it contains Δ and it is closed under System P and the rational monotony property.

In general, and contrary to Δ^P, Δ^R is not unique. The intersection of all the rational extensions Δ^R simply gives back Δ^P (Lehmann and Magidor, 1992).

3. Possibility Theory and Default Rules

3.1. Basic Definitions

The basic object in possibility theory is the notion of *possibility distribution*, which is here a mapping, denoted by π, from the set of classical interpretations Ω to the interval [0,1], taken as a prototype of infinite bounded totally ordered scale. A possibility distribution corresponds to a ranking on Ω such that the most plausible worlds get the highest value. The possibility distribution π represents the available knowledge about where the real world is. By convention, $\pi(\omega)=1$ means that it is totally possible that ω be the real world, $\pi(\omega)>0$ means that ω is only somewhat possible, while $\pi(\omega)=0$ means that ω is certainly not the real world. The inequality $\pi(\omega)>\pi(\omega')$ means that the situation ω is a priori more plausible than ω'. A possibility distribution π induces two mappings which respectively grade the possibility $\Pi(\phi)$ and the certainty $N(\phi)$ of a formula ϕ:

- $\Pi(\phi)=\sup\{\pi(\omega) \mid \omega \models \phi\}$ which evaluates to what extent ϕ is consistent with the available knowledge expressed by π. We have: $\forall \phi,\psi\ \Pi(\phi \vee \psi) = \max(\Pi(\phi), \Pi(\psi))$;
- $N(\phi) = \inf\{1-\pi(\omega) \mid \omega \models \neg\phi\}$ which evaluates to what extent ϕ is entailed by the available knowledge. We have: $\forall \phi,\psi\ N(\phi \wedge \psi)=\min(N(\phi), N(\psi))$.

Certainty and possibility are dual via the relation $N(\phi)=1-\Pi(\neg\phi)$.

3.2. Qualitative Semantics of Possibilistic Logic

Possibility theory only needs the ordinal (and not the numerical) properties of [0,1]. As such, possibility theory provides a qualitative model of uncertainty. To each possibility distribution π, we associate its comparative counterpart, an ordering relation on Ω denoted by $>_\pi$, defined by $\omega >_\pi \omega'$ iff $\pi(\omega) > \pi(\omega')$. It can be viewed as a well-ordered partition $(E_1,\ldots,E_n)$ of Ω such that:

$$\omega >_\pi \omega' \text{ iff } \omega \in E_i, \omega' \in E_j \text{ and } i<j \quad (\text{for } 1 \le i,j \le n).$$

In a similar way, a complete pre-order $\geq_\pi$ is defined as:

$$\omega \geq_\pi \omega' \text{ iff } \omega \in E_i, \omega' \in E_j \text{ and } i \le j \quad (\text{for } 1 \le i,j \le n).$$

By convention E_1 gathers the worlds which are totally possible (i.e., $\forall \omega \in E_1$, $\pi(\omega)=1$). $>_\pi$ (or $\geq_\pi$) is called a comparative representation. Comparative possibility distributions allow us to represent a possibility distribution in terms of classes of equally possible worlds. Comparative possibility distributions are equivalent to what Lehmann and Magidor (1992) called ranked models. The following two definitions introduce the notion of π-entailment which can be defined in the spirit of Shoham's proposal:

Definition 2: An interpretation ω is a π-preferential model of a consistent formula ϕ w.r.t. the comparative possibility distribution $>_\pi$ iff:

(i) $\omega \models \phi$ and ii) $\nexists \omega'$, $\omega' \models \phi$ and $\omega' >_\pi \omega$.

We denote by $[\phi]_\pi$ the set of π-preferential models of the formula ϕ. This set is not defined when the formula is inconsistent. A comparative possibility relation on the set of formulas can also be defined for any pair ϕ and ψ as $\phi \geq_\Pi \psi \Leftrightarrow \Pi(\phi) \geq \Pi(\psi)$. Or equivalently, given $>_\pi=(E_1,\ldots, E_n)$, $\phi >_\Pi \psi$ iff there exists $\omega \in [\phi]_\pi$ such that for each $\omega' \in [\psi]_\pi$, we have $\omega >_\pi \omega'$.

Definition 3: ψ is a possibilistic consequence of $\phi \neq \bot$ w.r.t. $>_\pi$, denoted by $\phi \models_\pi \psi$, iff each π-preferential model of ϕ satisfies ψ, i.e, iff $\forall \omega \in [\phi]_\pi$, $\omega \models \psi$

We restrict this definition to formulas ϕ such that $\Pi(\phi)>0$, $\forall \phi \neq \bot$. It can easily checked that:

$$\phi \models_\pi \psi \text{ iff } \phi \wedge \psi >_\Pi \phi \wedge \neg\psi.$$

The possibilistic entailment is a nonmonotonic inference relation. It has been shown in (Dubois and Prade, 1991; Benferhat et al., 1992) that the inference $\models_\pi$ satisfies all the properties proposed in (Kraus et al., 1990) and (Gärdenfors and Makinson, 1994), including rational monotony, and can represent any of these relations, up to some minor technical details which are not compulsory.

3.3. Default Rules as a Family of Possibility Distributions

Benferhat et al. (1992) have proposed to model default rules of the form "normally if α then β" by "$\alpha \wedge \beta$ is more possible than $\alpha\wedge\neg\beta$". In other words, it is expressed by the strict inequality $\alpha\wedge\beta >_{\Pi} \alpha\wedge\neg\beta$. All possibility measures satisfying this inequality do express that if α is true then β is normally true. They correspond to all epistemic states where the rule is accepted. This minimal requirement is very natural since it guarantees that all rules in the default base are preserved. A set of defaults $\Delta=\{\alpha_i\rightarrow\beta_i, i=1,n\}$ with consistent conditions (i.e., $\forall i, \alpha_i\neq\bot$), can thus be viewed as a family of constraints $\mathcal{C}(\Delta)$ restricting a family $\Pi(\Delta)$ of comparative possibility distributions, which are said to be *compatible* with Δ:

Definition 4: A comparative possibility distribution $>_\pi$ is said to be compatible with Δ iff it holds that for each default rule $\alpha_i\rightarrow\beta_i$ of Δ, $\alpha_i\wedge\beta_i >_{\Pi} \alpha_i\wedge\neg\beta_i$.

$\Pi(\Delta)$ is empty iff Δ is inconsistent. In the following we assume that $\Pi(\Delta)$ is not empty.

4. Characterizing the Set $\Pi(\Delta)$

4.1. Two Orderings of the Elements in $\Pi(\Delta)$

This section analyses the structure of $\Pi(\Delta)$ on the basis of two natural orderings between comparative possibility distributions. The first one is Yager's *specificity ordering* that compares possibility distributions via their amount of incompleteness. The second ordering between possibility distributions, called *refinement ordering*, agrees with the idea of informativeness understood in the following sense: A preferential consequence relation $\models_\pi$ induced by a possibility distribution $>\pi$ is said to be more *informative* (or more *productive*) than $\models_{\pi'}$ iff $\forall\phi,\psi$, $\phi\models_{\pi'}\psi$ implies $\phi\models_\pi\psi$. For classical inference these two notions coincide in the sense that the more complete a set of formulas, the more informed it is, and the more conclusions can be obtained from ϕ and this set of formulas. However this equivalence no longer holds with a non-monotonic inference. These orderings are defined as:

Definition 5. Let $>_\pi=(E_1,\ldots,E_n)$ and $>_{\pi'}=(E'_1,\ldots,E'_{n'})$ be two comparative possibility distributions. Then:

1. $>_\pi$ refines $>_{\pi'}$ iff $\omega_1 >_{\pi'} \omega_2$ implies $\omega_1 >_\pi \omega_2$.
2. $>_\pi$ is less specific than $>_{\pi'}$ in the wide sense iff $\forall\omega$, if $\omega\in E'_i$ then $\omega\in E_j$ with $j\leq i$.

Equivalently, $>_\pi$ is less specific than $>_{\pi'}$ in the wide sense iff $\forall$ j=1, max(n,n'), $\bigcup_{i=1,j} E'_i\subseteq\bigcup_{i=1,j} E_i$, where for j>min(n,n') we use $E_j=\emptyset$ if n<n'. It is clear that the less specific $>_\pi$ is, the more numerous are the elements in the classes E_j of low rank j. Hence minimizing specificity comes down to minimizing the number k of equivalence classes, so as to assign as many worlds as possible to classes of lower ranks. Refining a possibility distribution comes down to splitting the elements of the associated partition:

Proposition 1. $>_\pi$ refines $>_{\pi'}$ iff $\forall i\ \exists$ j,k, such that $j\leq k$, and $E'_i=\bigcup_{p=j,k} E_p$

Proofs of propositions are available in a technical report. This proposition is important since it shows at the semantical level how, in a unique way, to augment a set of plausible conclusions produced by a given possibility distribution. The following proposition exhibits the connection of the two orderings with informativeness.

Proposition 2. The following statements are valid:

1. $>_\pi$ refines $>_{\pi'}$ iff $\models_\pi$ is more productive than $\models_{\pi'}$.
2. if $>_\pi$ refines $>_{\pi'}$ then $>_{\pi'}$ is less specific than $>_\pi$. The converse is false.
3. For $>_\pi$ less specific than $>_{\pi'}$, if $\top\models_\pi\phi$ then $\top\models_{\pi'}\phi$.

For the converse of 2 consider $\Delta=\{\beta\rightarrow\alpha\}$. Then the two following comparative possibility distributions: $>_\pi = (E_1=\{\alpha\beta, \alpha\neg\beta\}, E_2=\{\neg\alpha\beta, \neg\alpha\neg\beta\})$, and $>_{\pi'}= (E'_1=\{\alpha\beta\}, E'_2=\{\neg\alpha\beta, \neg\alpha\neg\beta, \alpha\neg\beta\})$ belong to $\Pi(\Delta)$. It is clear that $>_\pi$ is less specific than $>_{\pi'}$ but $>_{\pi'}$ does not refine $>_\pi$.

4.2. Maximal specificity

$\Pi_{max}(\Delta)=\{>_\pi / \nexists >_{\pi'}\in \Pi(\Delta)$ s.t. $>_\pi$ less specific than $>_{\pi'}\}$ denotes the set of the most specific $>_\pi$'s in $\Pi(\Delta)$. This subsection studies the structure of the possibility distributions in $\Pi_{max}(\Delta)$. First, as a corollary of Proposition 1, splitting any element of the well-ordered partition associated with $>_\pi$ in $\Pi(\Delta)$ again leads to a compatible comparative possibility distribution, i.e.,

Corollary. Let $>_\pi=(E_1,\ldots E_i,\ldots,E_n)\in \Pi(\Delta)$. Let $>_{\pi'}= (E_1,\ldots,E'_i, E''_i,\ldots E_n)$ obtained from $>_\pi$ by splitting E_i into $E'_i\cup E''_i$. Then $>_{\pi'}$ belongs to $\Pi(\Delta)$.

We now introduce *linear* possibility distributions:

Definition 6. A comparative possibility distribution $>_\pi=(E_1,\ldots,E_n\}$ of $\Pi(\Delta)$ is said to be linear iff each E_i is a singleton (i.e., contains exactly one interpretation).

In other words Ω is totally ordered by a linear possibility distribution. If π is a linear possibility distribution then for any α,β it holds that either $\alpha\models_\pi\beta$ or $\alpha\models_\pi\neg\beta$. This property is called the conditional excluded middle by Stalnaker. A second corollary of Prop. 1 is that there exists at least one linear possibility distribution exists in $\Pi(\Delta)$.

The most specific possibility distributions in $\Pi(\Delta)$ are exactly those which are linear (and thus the most refined and the most specific possibility distributions coincide):

Proposition 3: A possibility distribution $>_\pi$ of $\Pi(\Delta)$ is linear iff it belongs to $\Pi_{max}(\Delta)$.

4.3. Minimal specificity

The least specific comparative possibility distribution in $\Pi(\Delta)$ is unique. To show this, we first define the maximum of two possibility distributions, and show that this maximum operation is closed in $\Pi(\Delta)$, and that it computes the least upperbound of the two possibility distributions in the lattice-theoretic sense, for the specificity ordering.

Definition 7. Let $>_\pi=(E_1, \ldots, E_n)$ and $>_{\pi'}=(E'_1, \ldots, E'_m)$. Then the operator $\mathcal{M}ax$ is defined as:

$$\mathcal{M}ax(>_\pi, >_{\pi'}) = (E''_1, \ldots, E''_{\min(n,m)})$$

such that $E''_1=E_1 \cup E'_1$ and

$$E''_k=(E_k \cup E'_k) - (\bigcup_{i=1,k-1} E''_i) \text{ for } k=2,\min(n,m).$$

To make this definition more concrete let us consider an example: $\Omega=\{a,b,c,d,e\}$, $>_\pi=(E_1=\{a,b\}, E_2=\{c\}, E_3=\{d,e\})$ and $>_{\pi'}=(E'_1=\{a,c\}, E'_2=\{b,d\}, E'_3=\{e\})$. Then:

$$E''_1 = E_1 \cup E'_1 = \{a,b,c\};$$
$$E''_2 = E_2 \cup E'_2 - E''_1 = \{d\};$$
$$E''_3 = E_3 \cup E'_3 - E''_1 \cup E''_2 = \{e\}.$$

So $\text{Max}(>_\pi, >_{\pi'}) = (E''_1= \{a,b,c\}, E''_2= \{d\}, E''_3= \{e\})$.

Proposition 4. Let $>_\pi$ and $>_{\pi'}$ be two elements of $\Pi(\Delta)$. Then:

1. $\mathcal{M}ax\{>_\pi, >_{\pi'}\} \in \Pi(\Delta)$,
2. $\mathcal{M}ax\{>_\pi, >_{\pi'}\}$ is less specific than $>_\pi$ and $>_{\pi'}$.
3. If $>_{\pi^*}$ is less specific than both $>_\pi$ and $>_{\pi'}$ then it is less specific (in the wide sense) than $\mathcal{M}ax\{>_\pi, >_{\pi'}\}$
4. It is generally not true that $>_\pi$ and $>_{\pi'}$ refine $\mathcal{M}ax\{>_\pi, >_{\pi'}\}$ (in a non-trivial way).

For the counter-example of the point 4.4. Let $\Delta=\{\alpha \rightarrow \beta\}$. Let

$>_\pi=(E_1=\{\neg\alpha\beta\}, E_2=\{\alpha\beta\}, E_3=\{\alpha\neg\beta\}, E_4=\{\neg\alpha\neg\beta\})$,
$>_{\pi'}=(E'_1=\{\neg\alpha\neg\beta\}, E'_2=\{\alpha\beta\}, E'_3=\{\alpha\neg\beta\}, E'_4=\{\neg\alpha\beta\})$.

Then:

$$\mathcal{M}ax\{>_\pi, >_{\pi'}\}=(E''_1=\{\neg\alpha\beta, \neg\alpha\neg\beta\}, E''_2=\{\alpha\beta\}, E''_3=\{\alpha\neg\beta\}).$$

All the previous possibility distributions are compatible with Δ, but neither $>_\pi$ nor $>_{\pi'}$ refines $\mathcal{M}ax\{>_\pi, >_{\pi'}\}$.

Corollary. There exists exactly one possibility distribution in $\Pi(\Delta)$ which is the least specific one, denoted by $>_{\pi spe}$, and defined in the following way:

$$>_{\pi spe} = \mathcal{M}ax\{>_\pi \,/\, >_\pi \in \Pi(\Delta)\}.$$

Although, the least specific possibility distribution is unique, there is generally more than one least refined possibility distribution. Indeed, $(E_1=\{\neg\alpha\beta, \alpha\beta\}, E_2=\{\alpha\neg\beta, \neg\alpha\neg\beta\})$ and $(E'_1=\{\neg\alpha\neg\beta, \alpha\beta\}, E'_2=\{\alpha\neg\beta, \neg\alpha\beta\})$ are two least refined possibility distributions in $\Pi(\Delta)$ with $\Delta= \{\alpha \rightarrow \beta\}$) that are not comparable w.r.t. the refinement relation.

It has been shown (Benferhat et al., 1992) that the deductive closure of Δ based on the least specific possibility distribution $>_{\pi spe}$ corresponds to a particular rational extension of Δ^P called by Lehmann and Magidor (1992) "the rational closure". Equivalently $>_{\pi spe}$ induces an ordering of the rules in Δ whereby $\alpha \rightarrow \beta$ has priority over $\alpha' \rightarrow \beta'$ if and only if $\neg\alpha \vee \beta >_{Nspe} \neg\alpha' \vee \beta'$ using the necessity ordering induced by $>_{\pi spe}$ on the formulas; we recover Pearl's (1990) system Z.

5. Entailment Based on the Most Specific Possibility Distributions

This section considers the inference relation based on possibility distributions in $\Pi_{Max}(\Delta)$. It is shown that taking only one possibility distribution in $\Pi_{Max}(\Delta)$ corresponds to considering a closure of Δ^P under the so-called completion postulate ($\alpha \mid\!\sim \delta$ implies either $\alpha \wedge \beta \mid\!\sim \delta$ or $\alpha \wedge \beta \mid\!\sim \neg\delta$). Moreover, if the universal entailment based on all the linear possibility distributions (namely the entailment from α to β w.r.t. Δ defined by for all $>_\pi \in \Pi_{Max}(\Delta)$, $\alpha \models_\pi \beta$) yields the preferential closure Δ^P exactly. Since it is proved elsewhere (Dubois and Prade, 1995a) that the entailment based on all the possibility distributions in $\Pi(\Delta)$ yields Δ^P exactly, it means that $\Pi_{Max}(\Delta)$ is enough to characterize $\Pi(\Delta)$. Results of this section will be also very helpful to find a non-infinitesimal probabilistic semantics for Δ^P.

Definition 8: A consistent set of defaults Δ^C is said to be a complete extension of Δ if it contains Δ and it is closed under System P and the completion rule.

Proposition 5: Each complete extension of a conditional knowledge base is rational.

Let $\Delta^\pi = \{\alpha \rightarrow \beta$ such that $\alpha \models_\pi \beta\}$ be the set of default rules which are inferred from $>_\pi$. Then:

Proposition 6: Let $>_\pi$ belongs to $\Pi_{max}(\Delta)$. Then Δ^π satisfies the completion rule.

The converse is false. Indeed, let us consider a language with two propositional symbols b,f. Assume that Δ contains one default $b \rightarrow f$. Let $>_\pi= (E_1=\{bf, \neg bf, \neg b\neg f\}, E_2=\{b\neg f\}) \in \Pi(\Delta)$. We can check that Δ^π, which contains only $\{b \rightarrow f, b \wedge f \rightarrow f, b \wedge \neg f \rightarrow \neg f\}$ and the rules obtained from them via Right Weakening, is complete but π is not linear. To obtain the converse of the previous proposition, we must consider only *maximal* and complete extensions defined as:

Definition 9. A complete extension Δ^C is said to be maximal if it does not exist a complete extension $\Delta^{C'}$ such that $\Delta^C \subset \Delta^{C'}$.

Hence the following result:

Proposition 7. Δ^π is a maximal complete extension iff $>_\pi$ belongs to $\Pi_{Max}(\Delta)$.

The next result shows that taking the intersection of all the complete extensions of Δ leads to Δ^P. The two following lemmas help us to prove this result, especially the second lemma which shows that each rational extension of Δ derives from some linear possibility distributions.

Lemma 1. Let $>_\pi=(E_1,\ldots E_i,\ldots,E_n) \in \Pi(\Delta)$. Let $>_{\pi_1}=(E_1,\ldots,E'_i,E''_i,\ldots E_n)$ and $>_{\pi_2}=(E_1,\ldots,E''_i,E'_i,\ldots E_n)$ obtained from $>_\pi=\mathcal{M}ax\{>_{\pi_1}, >_{\pi_2}\}$ by splitting E_i in $E'_i \cup E''_i$. Then:

$$\alpha \models_\pi \beta \text{ iff } \alpha \models_{\pi_1} \beta \text{ and } \alpha \models_{\pi_2} \beta$$

Lemma 2. Let $>_\pi=(E_1,\ldots,E_n) \in \Pi(\Delta)$. Then there exists a subset Π of $\Pi_{max}(\Delta)$ such that:

1. $\Delta^{\pi} = \bigcap_{\pi' \in \Pi} \Delta^{\pi'}$;

2. $>_{\pi} = \mathcal{Max}\{>_{\pi'} / >_{\pi'} \in \Pi\}$.

The main result follows, namely :

Proposition 8: $\Delta^{P} = \bigcap_{\pi \in \Pi \max(\Delta)} \Delta^{\pi}$

Proposition 8 means that we do not need all the elements of $\Pi(\Delta)$ to recover Δ^{P}, but that a subset (here $\Pi_{Max}(\Delta)$) is enough. But $\Delta^{P} = \bigcap_{\pi \in \Pi(\Delta)} \Delta^{\pi}$ results as a corollary.

N.B.: The converse of Lemma 2 is not true. Namely, we cannot use any proper subset Π of $\Pi_{max}(\Delta)$ and build a rational extension of Δ from it. In other words, there may be no π is in $\Pi(\Delta)$ such that $\Delta^{\pi} = \bigcap_{\pi' \in \Pi} \Delta^{\pi'}$. Indeed, let us consider Δ containing only one rule "generally, we have p", this rule is denoted by "$\top \to p$". We assume that our language only contains two propositional symbols p and q. We can easily check that : $>_{\pi 1} = (E_1, E_2, E_3, E_4)$, $>_{\pi 2} = (E_3, E_2, E_1, E_4)$ with: $E_1 = \{pq\}$, $E_2 = \{\neg p \neg q\}$, $E_3 = \{p \neg q\}$, $E_4 = \{\neg pq\}$ are compatible with Δ. We have: $\top \to p \in (\Delta^{\pi 1} \cap \Delta^{\pi 2})$, $\top \to q \notin (\Delta^{\pi 1} \cap \Delta^{\pi 2})$ and $\neg q \to p \notin (\Delta^{\pi 1} \cap \Delta^{\pi 2})$. We can show that there is no π in $\Pi(\Delta)$ such that $\Delta^{\pi} = (\Delta^{\pi 1} \cap \Delta^{\pi 2})$. Indeed, if such π exists then from $\top \to p \in \Delta^{\pi}$ and $\top \to q \notin \Delta^{\pi}$ we necessary infer that $\neg q \to p \in \Delta^{\pi}$, since the entailment based on one possibility distribution π satisfies rational monotony.

6. A New Probabilistic Semantics for System P

Pearl (1988), after Adams, has investigated a way of assigning probabilities to default information and has studied the properties of the associated nonmonotonic consequence relation. Both use infinitesimal probability assignments, whereby probabilities committed to default information are all very close to 1. One may wonder whether there exists another noticeable interpretation of default information in terms of probability. By "noticeable" we mean a probabilistic interpretation which allows us to get at least the set Δ^{P}. A natural interpretation of "generally, if α then β" is $P(\beta|\alpha) > 1/2$ (namely "the majority of α's are β"). This is simpler than resorting to infinitesimals, and intuitively more convincing. However, for instance, the property OR defined in System P will not be satisfied (see (Pearl, 1988, page 494) for a counter-example). So, if we want to recover Δ^{P} and keep $P(\beta|\alpha) > 1/2$ as the interpretation of defaults $\alpha \to \beta$, we must restrict ourselves to a particular class of probability distributions.

Snow (1994) has introduced a class of probability distributions called "atomic bound system":

Definition 10: An atomic bound system (ABS) on Ω is a set of probability distributions P each inducing a total strict ordering > between elements of Ω, such that for each interpretation ω , $P(\omega) > \sum_{\omega':\omega>\omega'} P(\omega')$, and $P(\omega)>0$ for the interpretation of the lowest rank.

Probability measures à la Snow are closely related to probability measures which are *acceptance functions*, in the sense defined in (Dubois & Prade, 1995b). Acceptance functions g are mappings from Ω to [0,1] such that:

i) if $g(\phi)>g(\neg\phi)$ and $\phi \vdash \psi$ then $g(\psi)>g(\neg\psi)$ and
ii) if $g(\phi)>g(\neg\phi)$ and $g(\psi)>g(\neg\psi)$ then $g(\phi\wedge\psi)>g(\neg\phi\vee\neg\psi)$.

When $g(\phi) > g(\neg\phi)$ we say that ϕ is accepted. The requirements (i) and (ii) ensure that the set $\{\phi, g(\phi) > g(\neg\phi)\}$ is closed under deduction. Examples of acceptance functions are necessity measures (then $g(\phi) > g(\neg\phi)$ writes $N(\phi)>0$). Dubois and Prade (1995b) found the only probability measures P that are acceptance functions:

Proposition 9: The only probability measures P that are acceptance functions are such that $P(\omega_0)>0.5$ for some interpretation ω_0 and the probability measures such that $\exists\omega, \omega', P(\omega)=P(\omega')=0.5$.

Probability measures which are acceptance functions such that $P(\{\omega_0\}) > 1/2$ possess a "usual value", that is, an element which, in frequentistic terms, occurs more often than all the other together. Hence probability measures à la Snow are acceptance functions since choosing ω_0 as the maximal element in the ordering >, $P(\omega_0)>\sum_{\omega\neq\omega_0} P(\omega)$ implies $P(\omega_0)>0.5$. The converse does not hold due to the pathological case in the above proposition. Acceptance functions can be requested to remain so under conditioning. In the case of probabilities, ϕ is accepted when ψ is true if and only if $P(\phi|\psi)>1/2$, i.e., $P(\phi\wedge\psi)>P(\neg\phi\wedge\psi)$. Then it is natural to require that the set $\mathcal{A}(\psi)=\{\phi, P(\phi\wedge\psi)>P(\neg\phi\wedge\psi)\}$ be also deductively closed. This requirement presupposes that the following reinforcement of ii) holds

$$\text{if } P(\phi_1\wedge\psi)>P(\neg\phi_1\wedge\psi) \text{ and } P(\phi_2\wedge\psi)>P(\neg\phi_2\wedge\psi)$$
$$\text{then } P(\phi_1\wedge\phi_2\wedge\psi)>P((\neg\phi_1\vee\neg\phi_2)\wedge\psi).$$

(CA: Conditional And).

The meaning of this property is that if $\phi_1 \in \mathcal{A}(\psi)$ and $\phi_2 \in \mathcal{A}(\psi)$ then $\phi_1\wedge\phi_2 \in \mathcal{A}(\psi)$ as well. When $[\psi]=\Omega$, the CA rule gives back condition ii). An acceptance function g that satisfies the CA rule is a useful revision tool in the sense that upon learning that ψ is true, the belief set induced by g is revised into another belief set where ψ is true, (which, in the case of probabilities, is obtained by conditioning since $\mathcal{A}(\psi)=\{\phi, P(\phi|\psi)>1/2\}$). In the following they will be called revision-proof acceptance functions. Then, the following results hold:

Proposition 10. Any probability measure in an ABS is a revision-proof acceptance function.

Proposition 11. Any positive probability measure that induces a linear ordering on Ω and that satisfies the CA rule belongs to an ABS.

A probability measure in an atomic bound system can be constructed as follows: select a total order > on Ω, such that $\omega_1>\omega_2>\ldots>\omega_n$. Let $f(\omega_n)=\varepsilon_n$. Then $f(\omega_{n-1}) = \varepsilon_n+\varepsilon_{n-1}\cdot f(\omega_{n-2}) = 2\varepsilon_n+\varepsilon_{n-1}+\varepsilon_{n-2},\ldots, f(\omega_i)=\varepsilon_i+\sum_{j=i+1,n} 2^{j-(i+1)}\varepsilon_j$, for i=1,n. The numbers $\varepsilon_1\ldots\varepsilon_n$ are any positive real numbers. Note that $\sum_{i=1,n} f(i)=\sum_{j=1,n}$

$2^{j-1}\varepsilon_j$ since in the summation the term ε_j appears with coefficient $1+\Sigma_{k=0,j-2}\, 2^k=2^{j-1}$. Then let $P(\omega_i)= f(\omega_i)/(\Sigma_{i=1,n} f(\omega_i))$ for i=1,n. Clearly, any probability measure in an atomic bound system can be generated this way, choosing $\varepsilon_n=P(\omega_n)$, $\varepsilon_{n-1}=P(\omega_{n-1})-P(\omega_n)$, $\varepsilon_i= P(\omega_i)-\Sigma_{j>i} P(\omega_j),\ldots, \varepsilon_1=P(\omega_1)-\Sigma_{j>1} P(\omega_j)$. In particular, $f(\omega_i)>2^{n-i+1}f(\omega_n)$ so that $\forall i$, $P(\omega_i)> 2^{n-i+1}P(\omega_n)$. However this condition is not sufficient to ensure that P belongs to an atomic bound system.

An ABS on Ω can be partitioned into as many subsets of probability measures as total orderings on Ω. Let $\mathcal{P}_{ABS}$ be the atomic bound system on Ω, and $\mathcal{P}^{<}_{ABS}$ be the subset of $\mathcal{P}_{ABS}$ such that $P(\omega_{i_1})>P(\omega_{i_2})>\ldots >P(\omega_{i_n})$ if $|\Omega|=n$. Clearly > can also be viewed as a linear possibility ordering, i.e., $>=>_\pi$. Note that if $P(\omega_i)= \max\{P(\omega), \omega\in[\phi]\}$, $P(\omega_j)=\max\{P(\omega), \omega\in[\psi]\}$, then $P(\phi)>P(\psi)$ if and only if $P(\omega_i)>P(\omega_j)$ if and only if $\omega_i>_\pi\omega_j$, if and only if $\Pi(\phi)>\Pi(\psi)$. Hence the ordering of formulas induced by any probability distribution in $\mathcal{P}^{<}_{ABS}$ is the same as the qualitative possibility ordering induced by >. This remains true under conditioning:

Proposition 12. Given $P\in\mathcal{P}^{<}_{ABS}$, and $<_\pi=<$; then $\forall\alpha,\beta$, $P(\beta|\alpha)>1/2 \Leftrightarrow \alpha\wedge\beta >_\Pi \alpha\wedge\neg\beta$.

From Proposition 12 it is clear that the inference $\models_\pi$ induced by a linear possibility ordering $>_\pi$ can be encoded by means of any probability distribution $P\in\mathcal{P}^{<}_{ABS}$ where $>=>_\pi$, and $\alpha\models_\pi\beta \Leftrightarrow P(\beta|\alpha)>1/2$. This result appears in Snow (1996).

In the previous section it was proved that system P can be recovered by considering only the linear orderings on the set of interpretations (Proposition 8), so that $\Delta^P= \bigcap_{\pi\in\mathbb{\Pi}_{max}(\Delta)} \Delta^\pi$. Let $\mathcal{P}(\Delta) = \{P \,/\, P(\beta_i|\alpha_i)>1/2$ for each $\alpha_i\to\beta_i$ in Δ and $P\in\mathcal{P}^{ABS}\}$. This set is not empty as long as Δ is consistent since there are linear comparative possibility distributions in $\mathbb{\Pi}(\Delta)$. then:

Proposition 13: $\alpha\to\beta\in\Delta^P$ iff $\forall P\in\mathcal{P}(\Delta)$, $P(\beta|\alpha)>1/2$.

The proof is obvious using Proposition 12 and Proposition 8 since all probability measures in $\mathcal{P}(\Delta)$, corresponding to the same total ordering of interpretations yield the same set of rules $\alpha\to\beta$ such that $P(\beta|\alpha)>1/2$. This result exhibits a non-infinitesimal probabilistic semantics for System P.

7. Conclusion

This paper first provides a standard probability semantics for System P, which has already been equipped with many different semantics: the original two-layered preferential one provided by Kraus et al. (1990), the infinitesimal probability semantics (Pearl, 1988), the conditional object semantics (Dubois and Prade, 1995a), and the possibility theory-based semantics. Second, the possibility theory-based semantics has been improved by showing that it is enough to consider linear possibility distributions. These results suggest that system P has semantics in terms of uncertainty measures that are neither possibility nor probability measures, namely any class of set-functions that are monotonic under inclusion and satisfy the conditional AND rule (the latter clearly inducing a very strong constraint). Laying bare these set functions would bridge the gap between conditioning and belief revision in a very general sense. Related results along this line appear in Friedman and Halpern (1996).

In the long range, the close relationship between linear possibility distributions and probability distributions à la Snow, should contribute to the progress of several issues, such as: the study of acceptance functions, the comparison between qualitative independence in possibility theory and probabilistic independence, the study of relations and differences between probabilistic decision theory and possibility theory-based decision.

References

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

Dubois D., and Prade H. 1991. Possibilistic logic, preferential models, non-monotonicity and related issues. In Proc. of the Inter. Joint Conf. on Artificial Intelligence (IJCAI'91), Sydney, 419-424.

Dubois D., and Prade H. 1995a. Conditional objects, possibility theory and default rules. In Conditionals: From Philosophy to Computer Sciences (G. Crocco et al., eds.), 311-346. Oxford, UK: Oxford Univ. Press.

Dubois D., and Prade H. 1995b. Numerical representation of acceptance. In Proc. of the 11th Conf. on Uncertainty in Artificial Intelligence (UAI'95), Montreal, 149-156.

Friedmann, and Halpern. 1996. Plausibility measures and default reasoning. In Proc. of the National Conf. on Artificial Intelligence (AAAI'96), 1297-1304.

Gärdenfors P., and Makinson D. 1994. Nonmonotonic inference based on expectations. *Artificial Intelligence* 65:197-245.

Kraus S., Lehmann D., and Magidor M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44:167-207.

Pearl J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Pearl J. 1990. System Z: A natural ordering of defaults with tractable applications to default reasoning. In Proc. of the TARK'90, 121-135.

Snow P. 1994. The emergence of ordered belief from initial ignorance. In Proc. of the 12th National Conf. on Artificial Intelligence (AAAI'94), 281-286.

Snow P. 1996. Standard Probability Distributions Described by Rational Default Entailment, Technical Report.

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