Precise Propagation of Upper and Lower Probability Bounds in System P

Angelo Gilio
Dipartimento di Matematica e Informatica, Città Universitaria
Viale A. Doria, 6 - 95125 Catania (Italy).
e-mail: Gilio@dmi.unict.it

Abstract

In this paper we consider the inference rules of System P in the framework of coherent imprecise probabilistic assessments. Exploiting our algorithms, we propagate the lower and upper probability bounds associated with the conditional assertions of a given knowledge base, automatically obtaining the precise probability bounds for the derived conclusions of the inference rules. This allows a more flexible and realistic use of System P in default reasoning and provides an exact illustration of the degradation of the inference rules when interpreted in probabilistic terms. We also examine the *disjunctive Weak Rational Monotony* of System P^+ proposed by Adams in his extended probability logic.

Keywords: Nonmonotonic reasoning, System P, Conditional probability bounds, Precise propagation, Coherence.

Introduction

In the applications of intelligent systems to automated uncertain reasoning the explicit knowledge of the agent is represented by a knowledge base $\mathcal{K}$, constituted by a set of *conditional assertions* (i.e. *defaults*). The nonmonotonic inferential process is developed using a suitable set of rules. Among the many formalisms which have been proposed for default reasoning, the so-called System P developed in (Kraus, Lehmann, and Magidor 1990) is widely accepted and has a probabilistic semantics based on infinitesimal probabilities, see (Adams 1975, Pearl 1988). An extended probability logic has been proposed in (Adams 1986) by allowing *disjunction of conditionals.* The corresponding system P^+ has been studied in (Schurz 1998) where the perspectives of the infinitesimal probability semantics and that of a noninfinitesimal one, based on probabilistic inequality relations, have been unified. In (Hawthorne 1996) many logics of nonmonotonic conditionals, that behave like conditional probabilities at various levels of probabilistic support, have been examined. In the quoted paper the author shows that, for each given conditional $\rightarrow$, there is a probabilistic support level r and a conditional probability P such that, for all sentences B, A, it is $B \rightarrow A$ only if $P(A|B) \geq r$. We recall that an early examination of Adams rules by means of imprecise probabilities has been given in (Dubois, and Prade 1991). In the quoted paper a semantic interpretation in terms of intervals has been given for the relations of negligibility, closeness and comparability examined in the system of qualitative reasoning proposed in (Raiman 1986). Moreover, an application to the inference rules of Adams has been given interpreting "$P(A|B)$ is large" by means of the relation of closeness to 1 (making the infinitesimal parameter ϵ explicit). In this way a quantification of the degradation of the validity of Adams' rules when reasoning with noninfinitesimal probabilities has been obtained. Based on a particular class of probability distributions, called *big-stepped probabilities* (introduced with the name of *atomic bound system* in some papers of Snow, see, e.g., Snow 1996), in (Benferhat, Dubois, and Prade 1999) a standard probabilistic semantics has been provided to System P, without referring to *infinitesimals.* While in practical applications the semantics of infinitesimal probabilities may involve some difficulties, the approach based on lower (and possibly upper) probability bounds, proposed also in (Bourne, and Parsons 1998), is clearly more flexible and realistic. In this paper the propagation of probability bounds to the conditional assertions associated with the rules of System P is examined in the framework of the de Finetti's probabilistic methodology, based on the coherence principle. Notice that a coherent set of conditional probability assessments satisfies all the usual properties of conditional probabilities. A short examination of the logic of conditionals of Adams from the point of view of coherence has been given in (Gilio 1997), where the propagation of probabilistic bounds has not been considered. We point out that the peculiarity of our approach is given by the possibility of looking at the conditional probability $P(A|B)$ as a primitive concept,with no need of assuming that the probability of the *conditioning* event B be positive. On the contrary, in the probabilistic approaches usually adopted in the literature, see, e.g., (Adams 1975, Hawthorne

1996, Schurz 1998), by definition the quantity $P(A|B)$ is the *ratio* of $P(AB)$ and $P(B)$ if $P(B) \neq 0$, with $P(A|B) = 1$ if $P(B) = 0$. Notice that a clear *rationale* of this latter assumption does not seem to exist. Indeed, in the framework of coherence, this assumption is not made and the case of *conditioning* events of zero probability is managed without any problem: as an example, the condition $P(A|B)+P(A^c|B) = 1$, where A^c denotes the negation of A, is satisfied also when $P(B) = 0$. We think that the opportunity, offered by the probabilistic approach based on coherence, of developing a completely general treatment of probabilistic default reasoning is important specially in the field of nonmonotonic reasoning where infinitesimal probabilities play a significant role. Moreover, exploiting the algorithms developed in our framework, the lower and upper probability bounds associated with the conditional assertions of a given knowledge base can be propagated to further conditional assertions, obtaining in all cases the tightest probability bounds. Beside allowing a more flexible and realistic approach to probabilistic default reasoning, this provides an exact illustration of the degradation of System P rules when interpreted in probabilistic terms. The probabilistic approach based on coherence has been adopted in many recent papers, see, e. g., (Biazzo, and Gilio 2000), (Capotorti, and Vantaggi 1999), (Coletti 1994), (Coletti, and Scozzafava 1996), (Coletti, and Scozzafava 1999), (Gilio 1995a), (Gilio 1995b), (Gilio 1999), (Gilio, and Ingrassia 1998), (Gilio, and Scozzafava 1994), (Lad 1999), (Lad, Dickey, and Rahman 1990), (Scozzafava 1994). With the aim of reducing the computational difficulties, the algorithms described in (Biazzo, and Gilio 2000), implemented with Maple V, have been modified in (Biazzo, Gilio, and Sanfilippo 2000). A relevant sector (rapidly growing in the recent years) for applications of such algorithms may be that of probabilistic databases, (see, e.g., Laks, Lakshmanan, Leone, Ross, and Subrahmanian 1997; Biazzo, Ferro, Gilio, and Giugno 2000). The content of the paper is organized as follows:

- We first give some preliminary concepts and results on coherence and probability logic.
- We consider the definitions of probabilistic consistency and entailment given by Adams; then we consider a concept of strict probabilistic consistency and we prove its equivalence with that of probabilistic consistency.
- We consider the propagation of lower and upper probability bounds in System P. We also examine the disjunctive Weak Rational Monotony of System P$^+$.
- We examine the propagation of lower bounds with real $\epsilon-$values.
- We apply the results to (a slightly modified version of) an example given in (Bourne, and Parsons 1998).
- Finally, we give some conclusions.

Some preliminaries

In this section we recall some preliminary concepts and results on coherence checking and propagation of precise and imprecise probability assessments. We also show that our approach, being directly based on conditional assessments, can be framed in a setting of probability logic more general than the usual one. Given an arbitrary family $\mathcal{K}$ of conditional events, let P be a real function defined on $\mathcal{K}$. Given an integer n we denote by $\mathcal{F}_n = \{E_1|H_1, \ldots, E_n|H_n\}$ a finite subfamily of $\mathcal{K}$ and by $\mathcal{P}_n$ the vector $(p_1, \ldots, p_n)$, where $p_i = P(E_i|H_i), i = 1, 2, \ldots, n$. Denoting by the same symbols the events and their indicators, let G be the random quantity defined as

$$G = \sum_{i=1}^{n} s_i H_i (E_i - p_i),$$

where $s_1, s_2, \ldots, s_n$ are n real numbers. As well known, G can be interpreted as the random gain corresponding to a combination of n bets on $E_1|H_1, \ldots, E_n|H_n$, with stakes $s_1, s_2, \ldots, s_n$. Moreover, denoting by H_0 the union $H_1 \vee \cdots \vee H_n$, consider the restriction $G|H_0$ of G to H_0. Then, based on the *betting scheme*, we have

Definition 1 The function P is said coherent if, for every integer $n = 1, 2, \ldots$, for every subfamily $\mathcal{F}_n \subseteq \mathcal{K}$ and for every real numbers $s_1, s_2, \ldots, s_n$, the following condition is satisfied :

$$MinG|H_0 \leq 0 \leq MaxG|H_0. \qquad (1)$$

If the function P is coherent, then it is said a conditional probability on $\mathcal{K}$. In equivalent terms, we also say that P is a coherent probability assessment on $\mathcal{K}$. As well known, a coherent function P satisfies all the axioms of a conditional probability, while the converse is not true. We recall the following well known result (Lehman 1955, Williams 1975, Holzer 1985, Regazzini 1985).

Theorem 1 Let $\mathcal{K}$ be a family of conditional events and P be a coherent conditional probability assessment on $\mathcal{K}$. Given a family $\mathcal{K}^*$ such that $\mathcal{K} \subset \mathcal{K}^*$, there exists an extension P^* of P to $\mathcal{K}^*$, that is a coherent probability assessment P^* on $\mathcal{K}^*$ such that, for every $E|H \in \mathcal{K}$, it is $P^*(E|H) = P(E|H)$.

In particular we have

Theorem 2 Given a (precise) coherent probability assessment $\mathcal{P}_n = (p_1, \ldots, p_n)$ on a family $\mathcal{F}_n = \{E_1|H_1, \ldots, E_n|H_n\}$, with $p_i = P(E_i|H_i)$, and a further conditional event $E_{n+1}|H_{n+1}$, let $\mathcal{F}_{n+1}$ be the family $\mathcal{F}_n \cup \{E_{n+1}|H_{n+1}\}$. Then, there exists a suitable closed interval $[p', p''] \subseteq [0, 1]$ such that the extension $\mathcal{P}_{n+1} = (p_1, \ldots, p_n, p_{n+1})$ of $\mathcal{P}_n$ to $\mathcal{F}_{n+1}$ is coherent if and only if $p_{n+1} \in [p', p'']$.

In the case of unconditional events, the above result is known as the *fundamental theorem* of de Finetti.
Concerning imprecise probability assessments, let $\mathcal{A}_n = (\alpha_1, \ldots, \alpha_n)$ be the a vector of probability lower bounds on the family $\mathcal{F}_n = \{E_1|H_1, \ldots, E_n|H_n\}$ defined as

$$P(E_i|H_i) \geq \alpha_i \ , \quad i \in J_n = \{1, \ldots, n\} \ .$$

We consider the following definition of generalized coherence (g-coherence), given in (Biazzo, and Gilio 2000), which essentially coincides with a previous one introduced in (Gilio 1995a).

Definition 2 The vector of lower bounds $\mathcal{A}_n$ on $\mathcal{F}_n$ is said g-coherent if and only if there exists a precise coherent assessment $\mathcal{P}_n = (p_1, \ldots, p_n)$ on $\mathcal{F}_n$, with $p_i = P(E_i|H_i)$, which is consistent with $\mathcal{A}_n$, that is such that $p_i \geq \alpha_i$ for each $i \in J_n$.

The Definition 2 can be also applied to imprecise assessments like

$$\alpha_i \leq P(E_i|H_i) \leq \beta_i \;, \;\; i \in J_n,$$

since each inequality $P(E_i|H_i) \leq \beta_i$ amounts to the inequality $P(E_i^c|H_i) \geq 1 - \beta_i$.

Then, given an imprecise assessment $\mathcal{A}_n = (\alpha_1, \ldots, \alpha_n)$ on $\mathcal{F}_n$, a suitable procedure, given in (Gilio 1995b), can be used to check the g-coherence of $\mathcal{A}_n$. The g-coherent extension of $\mathcal{A}_n$ to a further conditional event $E_{n+1}|H_{n+1}$ has been studied in (Biazzo, and Gilio 2000) where, defining a suitable interval $[p_\circ, p^\circ] \subseteq [0, 1]$, the following result has been obtained.

Theorem 3 Given a g-coherent imprecise assessment $\mathcal{A}_n = ([\alpha_i, \beta_i], i \in J_n)$ on the family $\mathcal{F}_n = \{E_i|H_i, \; i \in J_n\}$, the extension $[\alpha_{n+1}, \beta_{n+1}]$ of $\mathcal{A}_n$ to a further conditional event $E_{n+1}|H_{n+1}$ is g-coherent if and only if the following condition is satisfied

$$[\alpha_{n+1}, \beta_{n+1}] \cap [p_\circ, p^\circ] \neq \emptyset \;. \tag{2}$$

Remark 1 In the quoted paper an algorithm has been proposed to determine $[p_\circ, p^\circ]$. Notice that, starting with a g-coherent assessment $\mathcal{A}_n$, by the same algorithm it is possible to determine the corresponding assessment $\mathcal{A}_n^*$ coherent w.r.t. the definition given in (Walley 1991). As an example, given the imprecise assessment $\mathcal{A}_2 = ([\frac{1}{5}, \frac{3}{5}], [\frac{1}{4}, \frac{4}{7}])$ on $\mathcal{F}_2 = \{E|H, E^c|H\}$, we have that $\mathcal{A}_2$ is g-coherent because, for example, the precise assessment $\mathcal{P}_2 = (\frac{1}{2}, \frac{1}{2})$ on $\mathcal{F}_2$, that is $P(E|H) = P(E^c|H) = \frac{1}{2}$, is coherent and it is $\frac{1}{2} \in [\frac{1}{5}, \frac{3}{5}], \frac{1}{2} \in [\frac{1}{4}, \frac{4}{7}]$. But, we observe that in $[\frac{1}{5}, \frac{3}{5}] \times [\frac{1}{4}, \frac{4}{7}]$ there doesn't exist a precise assessment like $\mathcal{P}_2 = (\frac{1}{5}, y)$, or $\mathcal{P}_2 = (x, \frac{1}{4})$, on $\mathcal{F}_2$ such that $\mathcal{P}_2$ is coherent. In this sense, $\mathcal{A}_2$ is not coherent. The algorithm given in (Biazzo, and Gilio 2000) modifies the initial (g-coherent) assessment $\mathcal{A}_2$, by determining the coherent assessment $\mathcal{A}_2^* = ([\frac{3}{7}, \frac{3}{5}], [\frac{2}{5}, \frac{4}{7}])$.

Our approach to the propagation of conditional probability bounds can be seen as a problem of *probabilistic deduction* in *probability logic*, see, e. g., (Frisch, and Haddawy 1994), (Lukasiewicz 1998), (Nilsson 1986).

We associate to each conditional assertion $H \mid\!\sim E$ in the knowledge base $\mathcal{K}$ a probability interval $[\alpha, \beta]$. In particular, given a family $\mathcal{F}_n$ of n conditional assertions, we consider an interval-valued probability assessment $\mathcal{A}_n = ([\alpha_i, \beta_i], i = 1, \ldots, n)$. Then, we can look at the pair $(\mathcal{F}_n, \mathcal{A}_n)$ as a probabilistic knowledge base, where each imprecise assessment $\alpha_i \leq P(E_i|H_i) \leq \beta_i$ is a probabilistic formula denoted by $(E_i|H_i)[\alpha_i, \beta_i]$. In our approach a probabilistic interpretation is just a coherent precise conditional probability assessment $\mathcal{P}_n$ on $\mathcal{F}_n$. A probabilistic interpretation $\mathcal{P}_n = (p_1, \ldots, p_n)$ is a *model* of a probabilistic formula $(E_i|H_i)[\alpha_i, \beta_i]$ iff $\mathcal{P}_n \models (E_i|H_i)[\alpha_i, \beta_i]$, that is $\alpha_i \leq p_i \leq \beta_i$. $\mathcal{P}_n$ is a model of the probabilistic knowledge base $\mathcal{K} = (\mathcal{F}_n, \mathcal{A}_n)$, denoted $\mathcal{P}_n \models \mathcal{K}$, iff $\mathcal{P}_n \models (E|H)[\alpha, \beta]$ for every $(E|H)[\alpha, \beta] \in \mathcal{K}$. A set of probabilistic formulas $\mathcal{K}$ is *satisfiable* iff a model of $\mathcal{K}$ exists, therefore the concept of satisfiability of $\mathcal{K} = (\mathcal{F}_n, \mathcal{A}_n)$ coincides with that of g-coherence of $\mathcal{A}_n$ on $\mathcal{F}_n$. A probabilistic formula $(E_{n+1}|H_{n+1})[\alpha_{n+1}, \beta_{n+1}]$ is a *logical consequence* of $\mathcal{K} = (\mathcal{F}_n, \mathcal{A}_n)$, denoted $\mathcal{K} \models (E_{n+1}|H_{n+1})[\alpha_{n+1}, \beta_{n+1}]$, iff

$$\alpha_{n+1} \leq inf\, \mathcal{I} \;, \quad \beta_{n+1} \geq sup\, \mathcal{I} \;,$$

where $\mathcal{I}$ is the set of the real values p such that there exists a model of $\mathcal{K} \cup \{(E_{n+1}|H_{n+1})[p, p]\}$. As shown by the condition (2), in our approach this amounts to

$$[p_\circ, p^\circ] \subseteq [\alpha_{n+1}, \beta_{n+1}].$$

A probabilistic formula $(E_{n+1}|H_{n+1})[\alpha_{n+1}, \beta_{n+1}]$ is a *tight logical consequence* of $\mathcal{K} = (\mathcal{F}_n, \mathcal{A}_n)$, denoted $\mathcal{K} \models_{tight} (E_{n+1}|H_{n+1})[\alpha_{n+1}, \beta_{n+1}]$, iff

$$\alpha_{n+1} = inf\, \mathcal{I} \;, \quad \beta_{n+1} = sup\, \mathcal{I} \;,$$

that is

$$\alpha_{n+1} = p_\circ \;, \quad \beta_{n+1} = p^\circ \;.$$

Considering a *probabilistic query* $(E_{n+1}|H_{n+1})[\alpha, \beta]$, where α and β are two different variables, to a probabilistic knowledge base $\mathcal{K} = (\mathcal{F}_n, \mathcal{A}_n)$ a *correct answer* is any $[\alpha, \beta] = [\alpha_{n+1}, \beta_{n+1}] \supseteq [p_\circ, p^\circ]$, that is such that $\mathcal{K} \models (E_{n+1}|H_{n+1})[\alpha_{n+1}, \beta_{n+1}]$. The *tight answer* is $[\alpha, \beta] = [p_\circ, p^\circ]$.

Remark 2 We point out that our probability logic setting is more general that the usual one. In fact, we don't rely on unconditional probability assessments. Hence, denoting by Ω the certain event (i.e. a tautology), if $\mathcal{K} = (\mathcal{F}_n, \mathcal{A}_n)$ is satisfiable, then it may happen that for some (or possibly all the) antecedents H_i's it results

$$\mathcal{K} \models_{tight} (H_i|\Omega)[0, 0] \;.$$

In fact, in our approach based on coherence, a conditional probability $P(E|H)$ represents the degree of belief (of someone) on "*E given H*" and has a natural meaning in all cases, included that one in which $P(H) = 0$. The only relevant thing is the assumption "*H true*".

Probabilistic consistency and entailment

We recall that in (Adams 1975) the conditional assertion "*if A then B* " is looked at as $P(B|A) \geq 1 - \epsilon \quad (\forall \epsilon > 0)$. Adopting a more realistic point of view we may look at the same conditional assertion as the probabilistic formula $(B|A)[\alpha, \beta]$, with $0 \leq \alpha \leq \beta \leq 1$, where usually $\beta = 1$. Then, a (probabilistic) knowledge base might be defined as a family of probabilistic formulas $\mathcal{K} = \{(E|H)[\alpha, \beta]\}$. In (Adams 1975) the following definition has been given.

Definition 3 The knowledge base $\mathcal{K}$ is *probabilistically consistent* (*p-consistent*) if, for every $\epsilon > 0$, there exists a probability P on A, proper for $\mathcal{K}$, such that $P(E|H) \geq 1-\epsilon$ for every $E|H \in \mathcal{K}$.

In our framework the concept of probabilistic consistency can be defined in the following way.

Definition 4 The knowledge base $\mathcal{K}$ is *probabilistically consistent* (*p-consistent*) if, for every set of lower bounds $\mathcal{A} = \{\alpha_{E|H}, E|H \in \mathcal{K}\}$ on $\mathcal{K}$, there exists a precise coherent probability assessment $P = \{p_{E|H}, E|H \in \mathcal{K}\}$ on $\mathcal{K}$, with $p_{E|H} = P(E|H)$, such that, for each $E|H \in \mathcal{K}$, $p_{E|H} \geq \alpha_{E|H}$.

We recall the concept of probabilistic entailment as defined in (Adams 1975).

Definition 5 A p-consistent knowledge base $\mathcal{K}$ *probabilistically entails* (*p-entails*) the conditional $B|A$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that for all probabilities P, proper for $\mathcal{K} \cup \{B|A\}$, if $P(E|H) \geq 1-\delta$ for each $E|H \in \mathcal{K}$, then $P(B|A) \geq 1-\epsilon$.

In our framework the concept of probabilistic entailment can be defined in the following way.

Definition 6 A p-consistent knowledge base $\mathcal{K}$ *probabilistically entails* the conditional $B|A$ if there exists a subfamily $\mathcal{F} \subseteq \mathcal{K}$ such that, for every $\alpha_{B|A}$, there exists a set of lower bounds $\mathcal{A} = \{\alpha_{E|H}, E|H \in \mathcal{F}\}$ on $\mathcal{F}$ such that for all precise coherent probability assessment $P = \{p_{B|A}, p_{E|H}, E|H \in \mathcal{F}\}$ on $\mathcal{F} \cup \{B|A\}$, with $p_{B|A} = P(B|A), p_{E|H} = P(E|H)$, if $p_{E|H} \geq \alpha_{E|H}$ for each $E|H \in \mathcal{F}$, then $p_{B|A} \geq \alpha_{B|A}$.

The probabilistic entailment of $B|A$ by $\mathcal{K}$ is denoted by the symbol $\mathcal{K} \Rightarrow B|A$. In (Adams 1975) a suitable set $\mathcal{R}$ of seven inference rules has been introduced, by means of which an inferential system can be developed to deduce in an automatic way all the plausible conclusions of a knowledge base $\mathcal{K}$.See also (Pearl 1988). The fundamental rules of the set $\mathcal{R}$ are the following ones:

$R1.$ $A \mid\!\sim C \;,\; A \mid\!\sim B \;\Rightarrow\; AB \mid\!\sim C$
(*Triangularity*)

$R2.$ $AB \mid\!\sim C \;,\; A \mid\!\sim B \;\Rightarrow\; A \mid\!\sim C$
(*Bayes*)

$R3$ $A \mid\!\sim C \;,\; B \mid\!\sim C \;\Rightarrow\; A \vee B \mid\!\sim C$
(*Disjunction*)

The previous rules, among others, have been used (with different names and in the framework of symbolic approaches too) by many authors, with the aim of developing some nonmonotonic logics to formalize the plausible reasoning (see e.g. (Kraus, Lehmann, and Magidor 1990), (Lehmann, and Magidor 1992), (Dubois, and Prade 1994)). See also the survey given in (Benferhat, Dubois, and Prade 1997).
We recall that in (Gilio 1997) the following concept of *strict probabilistic consistency* has been introduced.

Definition 7 The knowledge base $\mathcal{K}$ is strictly p-consistent if the probability assessment P on $\mathcal{K}$, such that $P(E|H) = 1$ for each $E|H \in \mathcal{K}$, is coherent.

Then the following result (given, without proof, in Gilio 1997) is obtained.

Theorem 4 $\mathcal{K}$ is p-consistent if and only if $\mathcal{K}$ is strictly p-consistent.

Proof. The proof of Theorem 4 is based on the following three assertions:

1. If $\mathcal{K}$ is strictly p-consistent, then $\mathcal{K}$ is p-consistent;
2. If $\mathcal{K}$ is p-consistent, then $\mathcal{K}$ is consistent;
3. If $\mathcal{K}$ is consistent, then $\mathcal{K}$ is strictly p-consistent.

The assertion 1 trivially holds because the strict p-consistency implies the p-consistency. The assertions 2 and 3 will be proved by the betting criterion.
2. We recall (Adams 1975) that, given $\mathcal{F} = \{E_i|H_i, i = 1, \ldots, n\}$, the *quasi-conjunction* $\mathcal{C}(\mathcal{F})$ of $E_1|H_1, \ldots, E_n|H_n$ is defined as

$$(E_1H_1 \vee H_1^c) \wedge \cdots \wedge (E_nH_n \vee H_n^c)|(H_1 \vee \cdots \vee H_n) \,.$$

Moreover, $\mathcal{K}$ is said consistent if and only if, for every $\mathcal{F} \subseteq \mathcal{K}$, the *quasi-conjunction* $\mathcal{C}(\mathcal{F})$ is verifiable. Given $\mathcal{F} = \{E_i|H_i, i = 1, \ldots, n\}$, let $C_1, \ldots, C_m$ be the constituents contained in H_0. Considering the random gain

$$G = \sum_{i=1}^{n} s_i H_i[E_i - (1-\varepsilon)],$$

with $\varepsilon > 0$ and $s_1, \ldots, s_n$ real numbers, with each C_h we associate a point $V_h = (v_{h1}, v_{h2}, \ldots, v_{hn}) \in [0,1]^n$, where

$$v_{hi} = \begin{cases} 1\,, & \text{if } C_h \subseteq E_iH_i\,, \\ 0\,, & \text{if } C_h \subseteq E_i^cH_i\,, \\ 1-\varepsilon\,, & \text{if } C_h \subseteq H_i^c\,. \end{cases}$$

Then, the value g_h of G associated with C_h is

$$g_h = \sum_{i:C_h \subseteq E_iH_i} s_i - (1-\varepsilon) \sum_{i:C_h \subseteq H_i} s_i = \sum_{i=1}^{n} s_i[q_{hi} - (1-\varepsilon)] \,.$$

Let us consider the following systems of linear inequalities:

$$(\mathcal{S}^*) \quad \begin{array}{l} \varepsilon \sum_{i:C_h \subseteq H_i} s_i - \sum_{i:C_h \subseteq E_i^cH_i} s_i \;<\; 0\,, \\[2ex] s_i \geq 0\,,\; i, \ldots, n;\;\; h = 1, \ldots, m. \end{array}$$

$$(\mathcal{S}) \quad \begin{array}{l} \sum_{h=1}^{m} \lambda_h[v_{hi} - (1-\varepsilon)] \;\geq\; 0\,,\;\; i = 1, \ldots, n; \\[2ex] \sum_{h=1}^{m} \lambda_h \;>\; 0\,,\;\; \lambda_h \geq 0\,,\;\; h = 1, \ldots, m. \end{array}$$

Based on a suitable alternative theorem, we have that exactly one of the following alternative holds:

- $(\mathcal{S}^*)$ has a nonnegative solution $(s_1, \ldots, s_n)$;
- (S) has a solution $(\lambda_1, \ldots, \lambda_m)$ semipositive, that is such that $\sum_{h=1}^{m} \lambda_h > 0,\ \lambda_h \geq 0,\ h = 1, \ldots, m$.

If $\mathcal{K}$ is p-consistent, then for every subfamily $\mathcal{F} = \{E_i|H_i, i = 1,\ldots,n\} \subseteq \mathcal{K}$ and for every $\varepsilon > 0$ there exists a coherent probability assessment $\mathcal{P} = (p_1,\ldots,p_n)$ on $\mathcal{F}$ such that $p_i \geq 1-\varepsilon$ for every i. This implies the compatibility of the system

$$\sum_{h=1}^{m} \lambda_h(q_{hi} - p_i) \;=\; 0\;,\;\; i = 1,\ldots,n;$$

$$\sum_{h=1}^{m} \lambda_h \;>\; 0\;,\;\; \lambda_h \geq 0\;,\;\; h = 1,\ldots,m,$$

where the quantities q_{hi} are obtained by the quantities v_{hi} replacing $1-\varepsilon$ with p_i. It follows

$$\sum_{h=1}^{m} \lambda_h[v_{hi} - (1-\varepsilon)] = \sum_{h=1}^{m} \lambda_h(q_{hi} - p_i) \;=\; 0\;,$$

so that $(\mathcal{S})$ is compatible and $(\mathcal{S}^*)$ is not. The compatibility of $(\mathcal{S})$ implies that $\mathcal{C}(\mathcal{F})$ is verifiable. In fact, if for each constituent C_h there existed a subscript i such that $C_h \subseteq E_i^c H_i$, then choosing $s_i = \cdots = s_n = s > 0$, for the value g_h we would obtain

$$g_h = \varepsilon \sum_{i:C_h \subseteq H_i} s - \sum_{i:C_h \subseteq E_i^c H_i} s \;<\; ns\varepsilon - s\;.$$

Then, for every $\varepsilon < \frac{1}{n}$, it would be $ns\varepsilon - s < 0$. Then, the system $(\mathcal{S}^*)$ would be compatible, which is absurd.
3. If $\mathcal{K}$ is consistent, then for every finite family $\mathcal{F} \subseteq \mathcal{K}$, the quasi-conjunction $\mathcal{C}(\mathcal{F})$ is verifiable, that is there exist a constituent C_h and a subscript j such that

$$C_h \subseteq E_i H_i \vee H_i,\; i = 1,\ldots,n;\;\; C_h \subseteq E_j H_j.$$

Then, let P be a probability assessment on $\mathcal{K}$, such that $P(E|H) = 1$ for every $E|H \in \mathcal{K}$, and let $\mathcal{P} = (1,1,\ldots,1)$ its restriction on $\mathcal{F}$. The random gain associated with $\mathcal{P}$ is given by

$$G = \sum_{i=1}^{n} s_i H_i(E_i - p_i) = \sum_{i=1}^{n} s_i H_i(E_i - 1) = -\sum_{i=1}^{n} s_i E_i^c H_i\;.$$

Then, the value g_h associated with C_h is

$$g_h \;=\; -\sum_{i:C_h \subseteq E_i^c H_i} s_i \;=\; 0\;,$$

for every $s_1,\ldots,s_n$. Therefore the coherence condition

$$MinG|H_0 \leq 0 \leq MaxG|H_0$$

is satisfied, for every finite family $\mathcal{F} \subseteq \mathcal{K}$ and the assessment P on $\mathcal{K}$, with $P(E|H) = 1$ for every $E|H \in \mathcal{K}$, is coherent. Hence $\mathcal{K}$ is strictly p-consistent. $\square$
Based on the Theorem 4, we can redefine the concept of probabilistic consistency in the following way.

Definition 8 The conditional knowledge base $\mathcal{K}$ is p-consistent if and only if the probability assessment P on $\mathcal{K}$, such that $P(E|H) = 1$ for every $E|H \in \mathcal{K}$, is coherent.

Moreover, the following well known result

Theorem 5 If $\mathcal{K}$ is consistent, then $\mathcal{K} \Rightarrow B|A$ if and only if $\mathcal{K} \cup \{B^c|A\}$ is inconsistent.

can be formulated in the following way:

Theorem 6 Given a consistent knowledge base $\mathcal{K}$ and a conditional $B|A$, $\mathcal{K}$ p-entails $B|A$ if and only if the probability assessment P on $\mathcal{K} \cup \{B^c|A\}$, with $P(B^c|A) = P(E|H) = 1$ for each $E|H \in \mathcal{K}$, is not coherent.

An example(*the penguin case*).
Given the knowledge base $\mathcal{K} = \{f|b, b|p, \neg f|p\}$, where

$$b = bird\;,\;\; f = fly\;,\;\; p = penguin\;,$$

we want to check the condition: $\mathcal{K} \models \neg f|p \wedge b$. To prove this conditional entailment we must verify that the assessment $\mathcal{P} = (1,1,1,1)$, defined on $\mathcal{F} = \{f|b, b|p, \neg f|p, \neg f|p \wedge b\}$, is not coherent. If we assume $p \wedge b = p$, then

$$P(\neg f|p) + P(f|p \wedge b) = P(\neg f|p) + P(f|p) = 1,$$

and from $P(\neg f|p) = 1$ it follows $P(f|p \wedge b) = 0$, so that $\mathcal{P}$ is not coherent and then $\mathcal{K} \models \neg f|p \wedge b$. If, for the sake of generality, we assume $p \wedge b \neq p$, the constituents generated by $\mathcal{F}$ and contained in $H_0 = b \vee p$ are

$$bfp\;,\;\; bf\neg p\;,\;\; b\neg fp\;,\;\; b\neg f\neg p\;,\;\; \neg bfp\;,\;\; \neg b\neg fp\;.$$

Based on an algorithm given in (Gilio 1995b), to check the coherence of $\mathcal{P}$ we start by checking the compatibility of the following system

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_5 + \lambda_6 = 1 \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1 \\ \lambda_2 + \lambda_3 + \lambda_4 + \lambda_6 = 1 \\ \lambda_1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_6 = 1 \\ \lambda_1 + \cdots + \lambda_6 = 1 \\ \lambda_h \geq 0,\;\; h = 1,\ldots,6. \end{cases}$$

As it can be verified, the system above is compatible and has the unique solution

$$\lambda_1 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 0\;,\; \lambda_2 = 1.$$

This means that the unique propagation of the assessment $\mathcal{P}$, defined on $\mathcal{F} = \{f|b, b|p, \neg f|p, \neg f|p \wedge b\}$, to $p|(b \vee p)$ is $P(p|b \vee p) = 0$ and that we need to complete the checking of coherence considering the assessment $\mathcal{P}_0 = (1,1,1)$ on $\mathcal{F}_0 = \{b|p, \neg f|p, \neg f|p \wedge b\}$. Then, we study the system

$$\begin{cases} x_1 + x_2 = 1 \\ x_2 + x_4 = 1 \\ x_1 + x_3 + x_4 = 1 \\ x_1 + \cdots + x_4 = 1 \\ x_h \geq 0,\;\; h = 1,\ldots,4, \end{cases}$$

which is *not compatible*. Therefore, $\mathcal{P}$ is *not coherent* and we conclude that $\{f|b, b|p, \neg f|p\} \models \neg f|p \wedge b$.

Exact propagation of probability bounds in System P

We recall that the inference rules in System P are the following ones:

1. $A \mid\!\sim A$
   $(Reflexivity)$

2. $\models A \leftrightarrow B,\ A \mid\!\sim C \implies B \mid\!\sim C$
   $(Left\ Logical\ Equivalence)$

3. $\models B \to C,\ A \mid\!\sim B \implies A \mid\!\sim C$
   $(Right\ Weakening)$

4. $A \mid\!\sim B,\ A \mid\!\sim C \implies A \mid\!\sim BC$
   (And)

5. $A \mid\!\sim C,\ A \mid\!\sim B \implies AB \mid\!\sim C$
   $(Cautious\ Monotonicity)$

6. $A \mid\!\sim C,\ B \mid\!\sim C \implies A \vee B \mid\!\sim C$
   (Or)

Two derived rules in System P are

7. $AB \mid\!\sim C,\ A \mid\!\sim B \implies A \mid\!\sim C \qquad (Cut)$

8. $AB \mid\!\sim C \implies A \mid\!\sim B \to C \qquad (S)$

As we can see, the rules $R1, R2, R3$ coincide respectively with the rules *Cautious Monotonicity, Cut, Or* in System P. In (Adam86) an extended probability logic (System P^{+}) was developed by allowing the disjunction of conditionals in the conclusion of inferences and by adding the following dWRM rule.

9. $A \mid\!\sim C \implies A \mid\!\sim B^c \ \vee\ AB \mid\!\sim C$
   $(disjunctive\ Weak\ Rational\ Monotony)$

Now we will show how the probability intervals associated with the antecedents in each rule of System P propagate in a precise way to the consequent of the given rule. We will also examine the rules Cut, S and dWRM. These bounds will provide an exact illustration of the degradation of System P rules when interpreted in probabilistic terms. Perhaps, as a by-product, also a better appreciation of the results given in (Hawthorne 1996) on the interpretation of nonmonotonic conditionals in terms of probabilistic support levels could be obtained. We assume that the basic events A, B, C are logically independent.

1. *Reflexivity rule.* As for every assertion A the assessment $P(A|A) = p$ is coherent if and only if $p = 1$, to the conditional assertion $A \mid\!\sim A$ we associate the interval $[\alpha, \beta] = [1, 1]$.

2. *Left Logical Equivalence rule.* If two assertions A, B are equivalent, then for every assertion C the assessment (x, y) on the pair of conditional events $C|A, C|B$ is coherent if and only $x = y$. Therefore we associate the same probability interval to the conditional assertions $C|A, C|B$. In other words, the assessment $[\alpha, \beta]$ on $C|A$ propagates to the same interval on $C|B$.

3. *Right Weakening rule.* If $B \subseteq C$ then, defining the inclusion among conditional events as in (Goodman, and Nguyen 1988), it is $B|A \subseteq C|A$ and then the assessment (x, y) on the pair of conditional events $B|A, C|A$ is coherent if and only $x \leq y$, see (Gilio 1993). Therefore, the assessment $[\alpha, \beta]$ on $B|A$ propagates to the interval $[\alpha, 1]$ on $C|A$.

4. *And rule.* Given the assessment (x, y) on the pair of conditional events $B|A, C|A$, as well known the extension $P(BC|A) = z$ is coherent if and only if

$$Max\ \{0, x + y - 1\} = z' \leq z \leq z'' = Min\ \{x, y\}.$$

Therefore, the probability intervals $[\alpha_1, \beta_1], [\alpha_2, \beta_2]$ on the antecedents $B|A, C|A$ of the rule propagate to the exact interval $[\alpha_3, \beta_3]$, with

$$\begin{aligned} \alpha_3 &= Min\ z' = Max\ \{0, \alpha_1 + \alpha_2 - 1\}\ , \\ \beta_3 &= Max\ z'' = Min\ \{\beta_1, \beta_2\}\ , \end{aligned} \tag{3}$$

on the consequent $BC|A$.

5. *Cautious Monotonicity rule.* Given the assessment (x, y) on the pair of conditional events $C|A, B|A$, as proved in (Gilio 1995b), the extension $P(C|AB) = z$ is coherent if and only if $z \in [z', z'']$, with

$$z' = \begin{cases} \frac{x+y-1}{y}\ , & \text{if} \quad x + y > 1 \\ 0\ , & \text{if} \quad x + y \leq 1 \end{cases}\ ,$$

$$z'' = \begin{cases} \frac{x}{y}\ , & \text{if} \quad x < y \\ 1\ , & \text{if} \quad x \geq y \end{cases}\ .$$

We observe that, for $(x, y) \in [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]$ the function $f(x, y)$ attains its minimum value at the point (α_1, α_2). Then, it follows

$$\alpha_3 = \begin{cases} \frac{\alpha_1+\alpha_2-1}{\alpha_2}, & \text{if} \quad \alpha_1 + \alpha_2 > 1 \\ 0, & \text{if} \quad \alpha_1 + \alpha_2 \leq 1 \end{cases} \tag{4}$$

Moreover, the function $g(x, y)$ attains its maximum value at the point (β_1, α_2). Then it follows

$$\beta_3 = \begin{cases} \frac{\beta_1}{\alpha_2}, & \text{if} \quad \beta_1 < \alpha_2 \\ 1, & \text{if} \quad \beta_1 \geq \alpha_2 \end{cases} \tag{5}$$

Then, in order to determine the interval $[\alpha_3, \beta_3]$ we have to consider the position of the vertices $(\alpha_1, \alpha_2), (\beta_1, \alpha_2)$ wrt diagonals of the unitary square $[0, 1]^2$.

6. *Or rule.* Given the assessment (x, y) on the pair of conditional events $C|A, C|B$, it can be proved that the extension $P(C|(A \vee B)) = z$ is coherent if and only if $z \in [z', z'']$, with

$$z' = \begin{cases} \frac{xy}{x+y-xy}, & \text{if} \quad (x, y) \neq (0, 0) \\ 0, & \text{if} \quad (x, y) = (0, 0) \end{cases}\ ,$$

$$z'' = \begin{cases} \frac{x+y-2xy}{1-xy}, & \text{if} \quad (x, y) \neq (1, 1) \\ 1, & \text{if} \quad (x, y) = (1, 1) \end{cases}\ .$$

Moreover, we observe that both z' and z'' increase as either x or y increase. Therefore, the probability intervals $[\alpha_1,\beta_1],[\alpha_2,\beta_2]$ on the antecedents $C|A, C|B$ of the rule propagate, under the condition $(\alpha_1,\alpha_2) \neq (0,0)$, $(\beta_1,\beta_2) \neq (1,1)$, to $[\alpha_3,\beta_3]$, with

$$\alpha_3 = \frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2-\alpha_1\alpha_2}\ , \qquad (6)$$

$$\beta_3 = \frac{\beta_1+\beta_2-2\beta_1\beta_2}{1-\beta_1\beta_2}\ , \qquad (7)$$

on the consequent $C|(A \vee B)$.

Concerning the rules Cut and S we have the following results.

(e) *Cut rule.* Given the assessment (x,y) on the pair of conditional events $C|AB, B|A$, it can be proved that the extension $P(C|A) = z$ is coherent if and only if

$$xy \leq z \leq xy + 1 - y\ .$$

Therefore, the probability intervals $[\alpha_1,\beta_1],[\alpha_2,\beta_2]$ on the antecedents $C|AB, B|A$ of the rule propagate to $[\alpha_3,\beta_3]$, with

$$\alpha_3 = \alpha_1\alpha_2\ , \quad \beta_3 = \beta_1\alpha_2 + 1 - \alpha_2\ , \qquad (8)$$

on the consequent $C|A$.

(f) *S rule.* As $C|AB \subseteq (B^c \vee C)|A$, then the assessment (x,y) on the conditional events $C|AB, (B^c \vee C)|A$ is coherent if and only if $x \leq y$. Therefore, the probability interval $[\alpha_1,\beta_1]$ on the antecedent $C|AB$ of the rule propagates to $[\alpha_2,\beta_2]$, with

$$\alpha_2 = \alpha_1, \quad \beta_2 = 1\ ,$$

on the consequent $(B^c \vee C)|A$.

(g) *dWRM rule.* Let $\mathcal{P} = (x,y,z)$ a probability assessment on the family $\{C|A, B^c|A, C|AB\}$. For this family the constituents (possible worlds) are

$$C_0 = A^c\ ,\ \ C_1 = ABC\ ,\ \ C_2 = ABC^c\ ,$$

$$C_3 = AB^cC\ ,\ \ C_4 = AB^cC^c\ .$$

To the constituents $C_1,\ldots,C_4$ we associate the points

$$Q_1 = (1,0,1)\ ,\ \ Q_2 = (0,0,0)\ ,$$

$$Q_3 = (1,1,z)\ ,\ \ Q_4 = (0,1,z)\ .$$

Then, based on the method given in (Gilio 1995b) and denoting by $\mathcal{I}$ the convex hull of the points $Q_1,\ldots,Q_4$, it can be proved that the coherence of $\mathcal{P}$ amounts to the condition $\mathcal{P} \in \mathcal{I}$. Notice that in general this condition is necessary but not sufficient for the coherence of an assessment $\mathcal{P}_n = (p_1,\ldots,p_n)$ on a family $\mathcal{F}_n = \{E_1|H_1,\ldots,E_n|H_n\}$. The study of the condition $\mathcal{P} \in \mathcal{I}$ requires considering the equations of the four planes determined respectively by the terns of points

$$\{Q_1,Q_2,Q_3\}\ ,\quad \{Q_2,Q_3,Q_4\}\ ,$$

$$\{Q_1,Q_2,Q_4\}\ ,\quad \{Q_1,Q_3,Q_4\}\ .$$

Denoting by X, Y, Z the axes' coordinates, the equations are given respectively by

$$Z = X + (z-1)Y\ ,\quad Z = zY\ ,$$

$$Z = X + zY\ ,\quad Z = (z-1)Y + 1\ .$$

Then, given the values x, y, it is

$$\mathcal{P} \in \mathcal{I} \iff z' \leq z \leq z''\ ,$$

where

$$z' = \begin{cases} \frac{x-y}{1-y}\ , & \text{if} \quad x > y \\ 0\ , & \text{if} \quad x \leq y \end{cases}\ ,$$

$$z'' = \begin{cases} \frac{x}{1-y}\ , & \text{if} \quad x+y < 1 \\ 1\ , & \text{if} \quad x+y \geq 1 \end{cases}\ .$$

In order to examine the probabilistic interpretation of the rule we introduce a partition $\{\mathcal{R}_1,\mathcal{R}_2,\mathcal{R}_3,\mathcal{R}_4\}$ of the unitary square $[0,1]^2$, with

$$\mathcal{R}_1 = \{(x,y) : x+y < 1, x \geq y\}\ ,$$

$$\mathcal{R}_2 = \{(x,y) : x+y < 1, x < y\}\ ,$$

$$\mathcal{R}_3 = \{(x,y) : x+y \geq 1, x < y\}\ ,$$

$$\mathcal{R}_4 = \{(x,y) : x+y \geq 1, x \geq y\}\ .$$

We have to examine the case in which x is "high", therefore $\mathcal{R}_2$ is not of interest. In $\mathcal{R}_3$, since $x < y$, if x is "high" then y is "high" too. In $\mathcal{R}_1$ and $\mathcal{R}_4$ it is $z \geq z' = \frac{x-y}{1-y}$ so that, if x is "high" and y is "not high", then z is "high".

Concerning propagation of probability intervals, if we consider the assessments $[\alpha_1,\beta_1],[\alpha_2,\beta_2]$ on the conditional events $C|A, B^c|A$, then for the interval $[\alpha_3,\beta_3]$ associated with $C|AB$ we first observe that the quantity $\frac{x}{1-y}$ attains its maximum value at the point (β_1,α_2), while the quantity $\frac{x-y}{1-y}$ attains its minimum value at the point (α_1,β_2). Then, we have:

$$\alpha_3 = \begin{cases} \frac{\alpha_1-\beta_2}{1-\beta_2}, & \text{if} \quad \alpha_1 \geq \beta_2 \\ 0, & \text{if} \quad \alpha_1 < \beta_2 \end{cases} \qquad (9)$$

$$\beta_3 = \begin{cases} \frac{\beta_1}{1-\alpha_2}, & \text{if} \quad \beta_1+\alpha_2 < 1 \\ 1, & \text{if} \quad \beta_1+\alpha_2 \geq 1 \end{cases} \qquad (10)$$

Remark 3 Using the lower bounds computed previously, we can verify the probabilistic entailment in each inference rule on the basis of Definition 6. We have

- *And rule.* For each given value α_3, from (3) we have that, for every $(\alpha_1,\alpha_2) \in [\alpha_3,1] \times [\alpha_3,1]$ such that $\alpha_1+\alpha_2 = 1+\alpha_3$, if $P(B|A) \geq \alpha_1, P(C|A) \geq \alpha_2$ then $P(BC|A) \geq \alpha_3$.
- *Cautious Monotonicity rule.* For each given value α_3, from (4) we have that, for every $(\alpha_1,\alpha_2) \in [\alpha_3,1] \times (0,1]$ such that $\alpha_1 + (1-\alpha_3)\alpha_2 = 1$, if $P(C|A) \geq \alpha_1, P(B|A) \geq \alpha_2$ then $P(C|AB) \geq \alpha_3$.

- *Or rule.* For each given value α_3, from (6) we have that, for every $(\alpha_1, \alpha_2) \in [\alpha_3, 1] \times [\alpha_3, 1]$ such that $\alpha_2 = \frac{\alpha_1 \alpha_3}{\alpha_1(1+\alpha_3)-\alpha_3}$, if $P(C|A) \geq \alpha_1, P(C|B) \geq \alpha_2$ then $P(C|A \vee B) \geq \alpha_3$.
- *Cut rule.* For each given value α_3, from (8) we have that for every $(\alpha_1, \alpha_2) \in [\alpha_3, 1] \times [\alpha_3, 1]$ such that $\alpha_2 = \frac{\alpha_3}{\alpha_1}$, if $P(C|AB) \geq \alpha_1, P(B|A) \geq \alpha_2$ then $P(C|A) \geq \alpha_3$.

Propagation with $\epsilon-$values

The results of the previous section can be examined in the particular case in which for $i = 1, 2$ it is $[\alpha_i, \beta_i] = [1 - \epsilon_i, 1]$. As it can be verified, the $\epsilon-$values propagate in the following way.

- *And rule.* From (3), the probability bounds $[1-\epsilon_1, 1], [1-\epsilon_2, 1]$ on the antecedents $B|A, C|A$ of the rule propagate, on the consequent $BC|A$, to the exact bounds $[1-\epsilon_3, 1]$, with

$$\epsilon_3 = \epsilon_1 + \epsilon_2 \tag{11}$$

- *Cautious Monotonicity rule.* From (4), the probability intervals $[1-\epsilon_1, 1]$, $[1-\epsilon_2, 1]$ on the antecedents $C|A, B|A$ of the rule propagate, on the consequent $C|AB$, to $[1-\epsilon_3, 1]$, with

$$\epsilon_3 = \frac{\epsilon_1}{1-\epsilon_2} \tag{12}$$

- *Or rule.* From (6), the probability intervals $[1-\epsilon_1, 1], [1-\epsilon_2, 1]$ on the antecedents $C|A, C|B$ of the rule propagate, on the consequent $C|(A \vee B)$, to $[1-\epsilon_3, 1]$, with

$$\epsilon_3 = \frac{\epsilon_1 + \epsilon_2 - 2\epsilon_1\epsilon_2}{1 - \epsilon_1\epsilon_2} \tag{13}$$

- *Cut rule.* From (8), the probability intervals $[1-\epsilon_1, 1], [1-\epsilon_2, 1]$ on the antecedents $C|AB, B|A$ of the rule propagate, on the consequent $C|A$, to $[1-\epsilon_3, 1]$, with

$$\epsilon_3 = \epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2 \tag{14}$$

Remark 4 Our results concerning the value of ϵ_3 coincide with that ones obtained in (Bourne, and Parsons 1998) for the rules *And* and *Cautious Monotonicity* and are better for the rules *Or* and *Cut*, as from (13) and (14) one has respectively

$$\epsilon_3 \;=\; \tfrac{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2}{1-\epsilon_1\epsilon_2} \;<\; \epsilon_1+\epsilon_2 \qquad (\epsilon_1 < 1, \epsilon_2 < 1),$$

$$\epsilon_3 \;=\; \epsilon_1+\epsilon_2-\epsilon_1\epsilon_2 \;<\; \epsilon_1+\epsilon_2 \quad (\epsilon_1 > 0, \epsilon_2 > 0).$$

The use of the precise bounds may have some relevance when the inference rules are applied with real $\epsilon-$values.

An application

We will now examine an example to give an idea, on one hand, of how much the conclusions may be sensible to the use of methods of exact propagation of probability bounds and, on another hand, of the related phenomenon of degradation of inference rules when interpreted in probabilistic terms. The example is a modified version of an application considered in (Bourne, and Parsons 1998) which was inspired by examples given in (Kraus, Lehmann, and Magidor 1990). We consider a probabilistic knowledge base consisting of some conditional assertions, which concern the fact that a given party has various attributes (the party is *great, noisy*, *Linda* and *Steve* are *present*, and so on). By the symbol $A \mid\!\sim_\epsilon B$ we denote the assessment $P(B|A) \geq 1 - \epsilon$. We start with a knowledge base which has the following rules and $\epsilon-$values:

1. $Linda \mid\!\sim_{0.05} great$
2. $Linda \mid\!\sim_{0.2} Steve$
3. $Linda \wedge Steve \mid\!\sim_{0.1} \neg noisy$
4. $Steve \mid\!\sim_{0.05} Linda$
5. $\neg noisy \mid\!\sim_{0.2} \neg great$

Notice that the conditional "$Linda \mid\!\sim_{0.05} great$" means that the probability of the conditional event

"(The party will be great | Linda goes to the party)"

is greater than or equal to $1 - 0.05$, and so on. We are interested in propagating the previous bounds to find the $\epsilon-$values of the following conditionals:

(a) $Linda \mid\!\sim_\epsilon \neg noisy$

(b) $\top \mid\!\sim_\epsilon \neg Linda$

(c) $Linda \wedge Steve \mid\!\sim_\epsilon great \wedge \neg noisy$

(d) $Steve \mid\!\sim_\epsilon \neg noisy$

(e) $Linda \vee Steve \mid\!\sim_\epsilon \neg noisy$

By the symbol $\top$ we denote (any tautology representing) the certain event.
Applying the Cut rule to the conditionals

$$Linda \mid\!\sim_{0.2} Steve\ , \quad Linda \wedge Steve \mid\!\sim_{0.1} \neg noisy\ ,$$

we obtain

(a) $Linda \mid\!\sim_{0.28} \neg noisy$.

Then, applying the And rule to the conditionals

$$Linda \mid\!\sim_{0.28} \neg noisy\ , \quad Linda \mid\!\sim_{0.05} great\ ,$$

we obtain the conditional

$$Linda \mid\!\sim_{0.33} great \wedge \neg noisy\ .$$

Applying the S rule to

$$Linda \mid\!\sim_{0.33} great \wedge \neg noisy$$

we obtain

$$\top \mid\!\sim_{0.33} \neg Linda \vee great \wedge \neg noisy\ .$$

Applying the S rule to

$$\neg noisy \mid\!\sim_{0.2} \neg great$$

we obtain

$$\top \mid\!\sim_{0.2} noisy \vee \neg great\ .$$

Finally, applying the And rule to the conditionals

$$\top \mid\!\sim_{0.33} \neg Linda \vee great \wedge \neg noisy,\ \top \mid\!\sim_{0.2} noisy \vee \neg great,$$

we obtain the conditional

$$\top \mid\!\sim_{0.53} \neg Linda \wedge (noisy \vee \neg great).$$

Then, applying the Right Weakening rule to

$$\top \mid\!\sim_{0.53} \neg Linda \wedge (noisy \vee \neg great)$$

we obtain

$$(b) \qquad \top \mid\!\sim_{0.53} \neg Linda\ .$$

Concerning the conditional (c), applying the Cautious Monotonicity rule to

$$Linda \mid\!\sim_{0.05} great\ ,\quad Linda \mid\!\sim_{0.2} Steve\ ,$$

we obtain

$$Linda \wedge Steve \mid\!\sim_{0.0625} great\ .$$

Then, applying the And rule to the conditionals

$$Linda \wedge Steve \mid\!\sim_{0.0625} great,\ Linda \wedge Steve \mid\!\sim_{0.1} \neg noisy,$$

we obtain

$$(c) \qquad Linda \wedge Steve \mid\!\sim_{0.0725} great \wedge \neg noisy\ .$$

Applying the Cut rule to the conditionals

$$Linda \wedge Steve \mid\!\sim_{0.1} \neg noisy\ ,\quad Steve \mid\!\sim_{0.05} Linda\ ,$$

we obtain

$$(d) \qquad Steve \mid\!\sim_{0.145} \neg noisy\ . \qquad (15)$$

Then, applying the Or rule to

$$Steve \mid\!\sim_{0.145} \neg noisy\ ,\quad Linda \mid\!\sim_{0.28} \neg noisy\ ,$$

we obtain

$$(e) \qquad Linda \vee Steve \mid\!\sim_{0.358} \neg noisy\ . \qquad (16)$$

We observe that, propagating the bounds with $\epsilon_3 = \epsilon_1 + \epsilon_2$, instead of the conditionals (15) and (16) we would obtain respectively

$$Steve \mid\!\sim_{0.15} \neg noisy\ ,$$

and

$$Linda \vee Steve \mid\!\sim_{0.425} \neg noisy\ .$$

Conclusions

In this paper the inference rules of System P have been considered in the framework of coherence. We have also examined the *disjunctive Weak Rational Monotony* proposed by Adams in his extended probability logic, corresponding to System P^+. Differently from the probabilistic approaches generally given in the literature, see, in particular, (Hawthorne 1996) and (Schurz 1998), within our framework we can directly manage conditional probability assessments, even if some (or possibly all the) conditioning events have zero probability. We think that this opportunity is important specially in the field of nonmonotonic reasoning where infinitesimal probabilities play a significant role. Exploiting our algorithms, the lower and upper probability bounds associated with the conditional assertions of a given knowledge base can be propagated to further conditional assertions, obtaining in all cases the precise probability bounds. In particular, this provides an exact illustration of the degradation of System P rules when interpreted in probabilistic terms. Moreover, our approach allows a more flexible and realistic treatment of uncertainty in probabilistic default reasoning and related applications such as probabilistic databases.

Acknowledgments

The author is grateful to the referees for the very helpful criticisms and suggestions.

References

Adams, E. W. 1975. *The Logic of Conditionals.* Dordrecht, Netherlands: Reidel.

Adams, E. W. 1986. On the logic of high probability, *Journal of Philosophical Logic* 15: 255-279.

Benferhat, S.; Dubois, D.; and Prade, H. 1997. Nonmonotonic reasoning, conditional objects and possibility theory. *Artificial Intelligence* 92(1-2): 259-276.

Benferhat, S.; Dubois, D.; and Prade, H. 1999. Possibilistic and Standard Probabilistic Semantics of Conditional Knowledge Bases. *J. Logic Computat.* 9(6): 873-895.

Biazzo, V.; Ferro, A.; Gilio, A.; and Giugno, R. 2000. A general probabilistic database model. In *Proceedings of IPMU'2000*, Forthcoming.

Biazzo, V.; and Gilio A. 2000. A generalization of the fundamental theorem of de Finetti for imprecise conditional probability assessments. *International Journal of Approximate Reasoning.* Forthcoming.

Biazzo, V.; Gilio, A.; and Sanfilippo, G. 2000. Computational Aspects in Checking of Coherence and Propagation of Conditional Probability Bounds. In *Proceedings of WUPES'2000*, Forthcoming.

Bourne, R.; and Parsons, S. 1998. Propagating probabilities in System P. In *Proceedings of the 11th International FLAIRS Conference*, 440-445.

Capotorti, A.; and Vantaggi, B. 1999. A general interpretation of conditioning and its implication on coherence. *Soft Computing* 3/3: 148-153.

Coletti, G. 1994. Coherent numerical and ordinal probabilistic assessments. *IEEE Trans. on Systems, Man, and Cybernetics* 24(12): 1747-1754.

Coletti, G.; and Scozzafava, R. 1996. Characterization of coherent conditional probabilities as a tool for their assessment and extension. *Journal of Uncertainty, Fuzziness and Knowledge-based Systems* 4(2): 103-127.

Coletti, G.; and Scozzafava, R. 1999. Conditioning and inference in intelligent systems. *Soft Computing* 3/3: 118-130.

Dubois, D.; and Prade, H. 1991. Semantic consideration on order of magnitude reasoning. In *Decision support systems and qualitative reasoning*, Singh, M. G.; and Trave-Massuyes L. eds., 223-228. Elsevier Science Publishers B. V., North Holland.

Dubois, D.; and Prade, H. 1994. Conditional Objects as Nonmonotonic Consequence Relationships. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(12): 1724-1740.

Frisch, A. M.; and Haddawy, P. 1994. Anytime Deduction for Probabilistic Logic, *Artificial Intelligence* 69: 93-122.

Gilio, A. 1993. Conditional events and subjective probability in management of uncertainty. In *Uncertainty in Intelligent Systems - IPMU'92*, Bouchon-Meunier, B.; Valverde, L.; and Yager, R. R. eds., 109-120. Elsevier Science Publ. B. V., North-Holland.

Gilio, A. 1995a. Probabilistic consistency of conditional probability bounds. In *Advances in Intelligent Computing - IPMU '94, Lecture Notes in Computer Science* 945, Bouchon-Meunier, B.; Yager, R. R.; and Zadeh, L. A. eds., 200-209. Berlin Heidelberg: Springer-Verlag.

Gilio, A. 1995b. Algorithms for precise and imprecise conditional probability assessments. In *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence*, Coletti, G.; Dubois, D.; and Scozzafava, R. eds., 231-254. New York: Plenum Press.

Gilio, A. 1997. Probabilistic modelling of uncertain conditionals. In *Procedings of European Symposium on Intelligent Techniques*, 54-57. Bari, Italy.

Gilio, A. 1999. Probabilistic relations among logically dependent conditional events, *Soft Computing* 3/3: 154-161.

Gilio, A.; and Ingrassia, S. 1998. Totally coherent set-valued probability assessments, *Kybernetika* 34(1): 3-15.

Gilio, A.; and Scozzafava, R. 1994. Conditional events in probability assessments and revision, *IEEE Trans. on Systems, Man and Cybernetics* 24(12): 1741-1746.

Goodman, I. R.; Nguyen H. T. 1988. Conditional objects and the modeling of uncertainties. In *Fuzzy Computing Theory, Hardware and Applications*, Gupta, M. M.; and Yamakawa, T. eds., 119-138. New York: North-Holland.

Hawthorne, J. 1996. On the logic of nonmonotonic conditionals and conditional probabilities, *Journal of Philosophical Logic* 25: 185-218.

Holzer, S. 1985. On coherence and conditional prevision, *Boll. Un. Mat. Ital.* 4(6) : 441-460.

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

Lad, F. 1999. Assessing the foundations for Bayesian networks: a challenge to the principles and the practice, *Soft Computing* 3/3: 174-180.

Lad, F. R.; Dickey, J. M.; and Rahman, M. A. 1990. The fundamental theorem of prevision, *Statistica*, anno L(1): 19-38.

Laks, V.; Lakshmanan, S.; Leone, N.; Ross, R.; and Subrahmanian, V. S. 1997. ProbView: A Flexible Probabilistic Database System, *ACM Transactions on Database Systems* 22(3): 419-469.

Lehmann, D.; and Magidor, M. 1992. What does a conditional knowledge base entail?, *Artificial Intelligence* 55: 1-60.

Lehman, R. S. 1955. On confirmation and rational betting, *The Journal of Symbolic Logic*: 251-262.

Lukasiewicz, T. 1998. Magic inference rules for probabilistic deduction under taxonomic knowledge. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 354-361. Madison, Wisconsin.

Nilsson, N. J. 1986. Probabilistic logic, *Artificial Intelligence* 28: 71-87.

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

Raiman, O. (1989). Le raisonnement sur les ordres de grandeur, *Revue d'Intelligence Artificielle* 3(4): 55-67.

Schurz, G. 1998. Probabilistic semantics for Delgrande's conditional logic and a counterexample to his default logic, *Artificial Intelligence* 102(1): 81-95.

Regazzini, E. 1985. Finitely additive conditional probabilities, *Rend. Sem. Mat. Fis.* 55: 69-89.

Scozzafava, R. 1994. Subjective probability versus belief functions in artificial intelligence, *Int. J. General Systems* 22: 197-206.

Walley, P. 1991. *Statistical reasoning with imprecise probabilities*. London: Chapman and Hall.

Snow, P. 1996. Warranting non-demonstrative reasoning upon the common ground of contending uncertainty approaches. In *Proceedings of the 6th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (IPMU-96), Vol. 2, 1381-1386. Granada, Spain.

Williams, P. M. 1975. Notes on conditional previsions, Technical Report, School of Mathematical and Physical Sciences, University of Sussex.