A Consistency Problem for Imprecise Conditional Probability Assessments

Renato Pelessoni
Dip. di Matematica Applicata 'B. de Finetti'
Università di Trieste
Piazzale Europa, 1 I – 34127 Trieste, Italy
E-mail: renatop@econ.univ.trieste.it

Paolo Vicig
Dip. di Matematica Applicata 'B. de Finetti'
Università di Trieste
Piazzale Europa, 1 I – 34127 Trieste, Italy
E-mail: paolov@econ.univ.trieste.it

Abstract

In this paper we introduce an operational procedure which, given an avoiding sure loss (ASL) imprecise probability assessment $\mathcal{P}$ on an arbitrary finite set of conditional events, determines its 'least-committal' coherent correction, i.e. the coherent imprecise probability assessment which reduces the imprecision of $\mathcal{P}$ as little as possible, without ever increasing it. Besides, a new proof of the consistency of a known procedure, which checks the ASL condition, is supplied by introducing a technique employed also in the proof of the previous procedure. It is then shown that the two procedures can be 'merged' to obtain an algorithm to be used when it is not known a priori whether $\mathcal{P}$ is ASL.

1 Introduction

A basic problem in handling uncertainty in Expert Systems is that of verifying whether a probabilistic evaluation $\mathcal{P}$ which forms (part of) the knowledge base is consistent.

In several common situations $\mathcal{P}$ is a precise or imprecise conditional probability assessment on a set $\mathfrak{I}$ of conditional events, which is finite but usually arbitrary ($\mathfrak{I}$ is not necessarily, for instance, an algebra); this makes it hard to apply 'traditional' definitions of (precise or imprecise) conditional probability to $\mathcal{P}$, while concepts of consistency based on de Finetti's coherence principle are well-suited for such instances, since they do not impose any constraint on the domain of $\mathcal{P}$. However, while a unique notion of (precise) coherent conditional probability is commonly used (although in various forms: see, for instance, [3], [8], [10]), different notions of coherence for imprecise probabilities have been proposed in literature ([2], [9], [11], [14], [15], [16]); they weaken in various ways the precise probability coherence condition. So, if $\underline{P}$ is a coherent lower probability (2.1.4), it does, for instance, not necessarily satisfy additivity and generally obeys only weak forms of the product rule like $\underline{P}(H)\cdot\underline{P}(E|H) \leq \underline{P}(E \wedge H)$. Besides, the avoiding sure loss (ASL) condition recalled in 2.1.3 is even weaker than 2.1.4 and does not necessarily preserve several important properties of coherent imprecise probabilities (among these, non-negativity, monotonicity, weak product rules). On the other hand, checking the ASL condition (an algorithm is given in [2], [6]; see also [7]) requires operationally fewer computations than checking coherence (an algorithm is given in [11]). Further, it is also often simpler for an Expert System user to elicit his/her opinions through an ASL rather than a coherent assessment.

A practical problem which may arise is then that of correcting an ASL assessment on $\mathfrak{I}$ into a coherent one, without modifying 'too much' the initial assessment. The correction should not increase the imprecision of the probability assessment on any event of $\mathfrak{I}$ and at the same time should also reduce this imprecision as little as possible, just what suffices to achieve a coherent evaluation.

This kind of correction is the *least-committal* imprecise probability defined by (1) in 2.1.7, where it is also shown that it always exists.

The main aim of this paper is to introduce and demonstrate an algorithm which finds the least-committal imprecise probability, starting from an ASL assessment on $\mathfrak{I}$.

The least-committal probability is also, from the viewpoint of the theory of imprecise probabilities developed by P. Walley in [14], [15], a special case of natural extension, so that the algorithm is a computational procedure to find the natural extension of $\mathcal{P}$ on $\mathfrak{I}$.

Section 2 recalls the known definitions and results from the theory of precise as well as imprecise probabilities which are needed in the sequel.

A known algorithm for checking the ASL condition is concisely discussed in section 3. In particular, a proof of its consistency, different from the existing ones, is given in 3.4: this introduces some aspects of the technique which is then employed also for proving the consistency (4.3) of the algorithm 4.2 for finding the least-committal probability (some details in the proofs are omitted: refer to [12] for an exhaustive exposition of similar techniques, applied to different problems).

Operationally, all the algorithms quoted consist of a sequence of linear programming (LP) problems. As shown in 4.4 and in example 4.5, it is then possible to extend the algorithm 4.2 to handle the case where it is not known whether the assessment $\mathcal{P}$ is ASL, with the purpose of both checking this condition and finding the least-committal imprecise probability (this includes checking coherence as well) at the same time.

2 Preliminaries

2.1 Coherence for imprecise probabilities

We shall consider in the sequel lower imprecise probabilities $\underline{P}(\cdot|\cdot)$. The algorithms in sections 3, 4 can be easily modified to deal with upper probabilities $\overline{P}(\cdot|\cdot)$. A theoretical reason to refer to lower (or upper) probabilities only is the coniugacy relation $\overline{P}(E\,|\,H) = 1 - \underline{P}(\overline{E}\,|\,H)$, which is often assumed to hold for various reasons ([11], [14]).

After introducing some notations, we recall firstly a definition of coherence for precise conditional probabilities [10], outlining concise comments and comparisons among it and the notions of coherence for imprecise probabilities in 2.1.3 and 2.1.4.

2.1.1 Notation

Logical notation will be used for operations and relations involving events. Particularly, '$\wedge$', '$\vee$' and '$\Rightarrow$' are used for logical product, sum and implication of events respectively. $\overline{E}$ (or $\neg E$), $\varnothing$, Ω indicate the negation of event E, the impossible event and the sure one respectively.

$|E|$ denotes the random number equal to 1 if the event E is true, to zero otherwise.

In the sequel, we shall denote with $\mathcal{E}$ an arbitrary (finite or infinite) not empty set of conditional events.

2.1.2 Definition

$P(\cdot|\cdot)$ is a *coherent* conditional probability on $\mathcal{E}$ iff, $\forall m$, $\forall E_i|H_i \in \mathcal{E}$, $\forall s_i \in \mathfrak{R}$, $i = 1,\ldots,m$, defining

$$G = \sum_{i=1}^{m} s_i\, |H_i|(|E_i| - P(E_i|H_i)) \text{ and } H = \vee_{i=1}^{m} H_i,$$

it is $\max G|H \geq 0$.

Defs. 2.1.2, 2.1.3 and 2.1.4 all require a conditional random number $G|H$ not to be strictly negative. $G|H$ can be interpreted as the *gain* a subject can obtain from betting on an arbitrary finite number of events in $\mathcal{E}$; the elementary bet on $E_i|H_i$ whose gain is $g_i = |H_i|(|E_i| - P(E_i|H_i))$ is called off iff $|H_i| = 0$. In this case, g_i gives a null contribution to G, since the subject bets on $E_i|H_i$ with the proviso that H_i is true. In 2.1.2 it is possible to bet both *in favour* (if $s_i > 0$) and *against* (if $s_i < 0$) any $E_i|H_i$. Defs. 2.1.3, 2.1.4 modify this betting scheme allowing, respectively, no bet (2.1.3) or at most one elementary bet (2.1.4) against an event of $\mathcal{E}$. Corresponding definitions for upper probabilities may be referred to modified betting schemes allowing no or at most one elementary bet in favour of an event of $\mathcal{E}$. See also [1], [2], [3], [5], [10], [14] for further information on the notions of coherence for precise and imprecise probabilities, as well as for other concepts of imprecise probabilities.

2.1.3 Definition

$\underline{P}(\cdot|\cdot)$ is an *avoiding sure loss* (ASL) lower probability on $\mathcal{E}$ iff, $\forall m$, $\forall E_i|H_i \in \mathcal{E}$, $\forall s_i \geq 0$, $i = 1,\ldots,m$, defining

$$G = \sum_{i=1}^{m} s_i\, |H_i|(|E_i| - \underline{P}(E_i|H_i)) \text{ and } H = \vee_{i=1}^{m} H_i,$$

it is $\max G|H \geq 0$.

2.1.4 Definition

$\underline{P}(\cdot|\cdot)$ is a *coherent* lower probability on $\mathcal{E}$ iff, $\forall$m, $\forall E_i|H_i \in \mathcal{E}$, $\forall s_i \geq 0$, i = 0,…,m, defining

$$G = \sum_{i=1}^{m} s_i \; |H_i|(|E_i| - \underline{P}(E_i|H_i)) - s_0|H_0|(|E_0| - \underline{P}(E_0|H_0))$$

and $H = \vee_{i=0}^{m} H_i$,

it is $\max G|H \geq 0$.

Def. 2.1.3 is that of 'avoiding uniform loss' given in [14], [15], referring to lower conditional previsions. Def. 2.1.4 is given in [16] for upper conditional previsions.

To adhere to the problems discussed in the paper, from now onwards (except for 2.2.5) we shall consider assessments given on a *finite* set of events $\mathfrak{I} = \{E_1|H_1,\ldots,E_n|H_n\}$.

In this framework, 2.1.5 and 2.1.6, whose proofs can be found in [2] and in [11] or [16] respectively, state conditions equivalent to 2.1.3 and 2.1.4. A definition equivalent to 2.1.3, based on the dominance condition in 2.1.5, is given in [6].

2.1.5 Theorem

$\underline{P}(\cdot|\cdot)$ is an ASL lower probability on $\mathfrak{I}$ iff there exists a coherent conditional probability $P(\cdot|\cdot)$ *dominating* $\underline{P}(\cdot|\cdot)$ on $\mathfrak{I}$, i.e.

$\underline{P}(E_i|H_i) \leq P(E_i|H_i)\ \forall E_i|H_i \in \mathfrak{I}$.

2.1.6 Lower envelope theorem

$\underline{P}(\cdot|\cdot)$ is a coherent lower probability on $\mathfrak{I}$ iff there exists a set $\mathcal{M}$ of coherent conditional probabilities on $\mathfrak{I}$ such that

$\underline{P}(E_i|H_i) = \min_{P \in \mathcal{M}} \{P(E_i|H_i)\}\ \forall E_i|H_i \in \mathfrak{I}$.

2.1.7 Proposition

Let $\underline{P}(\cdot|\cdot)$ be an ASL lower probability on $\mathfrak{I}$. Let $\mathcal{P}$ be the set of the coherent probabilities defined on $\mathfrak{I}$ and dominating $\underline{P}$. Then,

(1) $\qquad \underline{P}^*(E_i|H_i) = \min_{P \in \mathcal{P}} \{P(E_i|H_i)\}\ \forall E_i|H_i \in \mathfrak{I}$

is a coherent lower probability dominating $\underline{P}$. Moreover, every coherent lower probability dominating $\underline{P}$ on $\mathfrak{I}$ dominates $\underline{P}^*$.

Proof $\underline{P}^*$ is coherent by 2.1.6 and obviously dominates $\underline{P}$. Let $\underline{P}'$ be a coherent lower probability dominating $\underline{P}$ on $\mathfrak{I}$. By 2.1.6 there exists a set of coherent conditional probabilities $\mathcal{M}$ such that $\underline{P}'(E_i|H_i) = \min_{P \in \mathcal{M}} \{P(E_i|H_i)\}\ \forall E_i|H_i \in \mathfrak{I}$. Clearly, $\mathcal{M} \subseteq \mathcal{P}$ and so $\underline{P}^*(E_i|H_i) \leq \underline{P}'(E_i|H_i)\ \forall E_i|H_i \in \mathfrak{I}$. ■

Given an ASL lower probability $\underline{P}$ on $\mathfrak{I}$, (1) defines the *least-committal* lower probability $\underline{P}^*$ (see also [13] for other applications of the concept), which can be interpreted as the minimal coherent correction of $\underline{P}$ that dominates $\underline{P}$ on $\mathfrak{I}$. Since the dominance condition is necessary to avoid increasing the degree of imprecision when modifying $\underline{P}$, determining $\underline{P}^*$ on $\mathfrak{I}$ is a natural way of correcting $\underline{P}$.

It can be seen that the least-committal lower probability $\underline{P}^*$ is the natural extension on $\mathfrak{I}$ (as defined in [15]) of the ASL lower probability $\underline{P}$.

2.2 Some subjective probability results

As appears from 2.1.5, 2.1.6, results on precise coherent probabilities may be relevant in problems concerning imprecise probabilities. In particular, a characterisation theorem for coherent conditional probabilities (introduced in [4]) is recalled in 2.2.4 in a simplified version (proved in [5]).

2.2.1 Definitions

Symbol $\mathcal{P}$ is used to denote a partition of Ω. The not impossible events of $\mathcal{P}$ are called *atoms*. Given $\mathcal{P}$ and an event $K \neq \emptyset$, the *conditional partition* $\mathcal{P}|K$ is formed by the events of $\mathcal{P}$ conditioned on K, i.e. the atoms of $\mathcal{P}|K$ are obtained by the atoms of $\mathcal{P}$ still possible after conditioning on K.

We say that an event E (E|K) is *logically dependent* from $\mathcal{P}$ ($\mathcal{P}|K$) if every atom of $\mathcal{P}$ ($\mathcal{P}|K$) implies either E (E|K) or E (E|K). This happens iff E (E|K) is a logical sum of atoms of $\mathcal{P}$ ($\mathcal{P}|K$).

Given $\mathcal{P}$, a partition $\mathcal{P}'$ is said *coarser* than $\mathcal{P}$ if every atom of $\mathcal{P}'$ is logically dependent from $\mathcal{P}$.

$\mathcal{A}(\mathcal{P})$ ($\mathcal{A}(\mathcal{P}|K)$) is the set of all events logically dependent from $\mathcal{P}$ ($\mathcal{P}|K$). We set also $\mathcal{A}^{\emptyset}(\mathcal{P}) = \mathcal{A}(\mathcal{P}) - \{\emptyset\}$, $\mathcal{A}^{\emptyset}(\mathcal{P}|K) = \mathcal{A}(\mathcal{P}|K) - \{\emptyset|K\}$ and $\mathcal{A}(\mathcal{P})|\mathcal{A}^{\emptyset}(\mathcal{P}) = \{E|H : E \in \mathcal{A}(\mathcal{P}),\ H \in \mathcal{A}^{\emptyset}(\mathcal{P})\}$.

2.2.2 Definition

Given a finite partition $\mathcal{P}$, a triple $(\mathcal{P}', >, \{\pi_c(\cdot)\}_{c \in \mathcal{P}'})$ is a *weight system* on $\mathcal{P}$ if

– $\mathcal{P}'$ is a partition coarser than $\mathcal{P}$;
– $>$ is a total order on $\mathcal{P}'$;
– every $\pi_c(\cdot)$ (*weight function*) is a positive real-valued function defined, up to a constant factor, on the set of the atoms of $\mathcal{P}$ which imply C; $\pi_c(e_i)$ is called *weight* of e_i.

2.2.3 Definitions

Let $(\mathcal{P}',>,\{\pi_c(\cdot)\}_{c\in\mathcal{P}'})$ be a weight system on $\mathcal{P}$ and let e_i, e_j be atoms of $\mathcal{P}$. Then, e_i, e_j *have weights of the same order* if they imply the same $C\in\mathcal{P}'$. Instead, e_i *has weight of higher order* than that of e_j if $e_i \Rightarrow C$, $e_j \Rightarrow D$ $(C, D\in\mathcal{P}')$ and $C > D$. If $E\in\mathcal{A}^{\varnothing}(\mathcal{P})$, E^* denotes the logical sum of the atoms of $\mathcal{P}$, among those implying E, with weight of highest order.

The *sum function* $\sigma(E)$ is the sum of weights of the atoms implying E^*. Obviously, it is $\sigma(E) = \sigma(E^*)$. Conventionally, $\sigma(\varnothing) = 0$.

2.2.4 Characterisation theorem

Let $\mathcal{P}$ be a finite partition. $P(\cdot|\cdot)$ is a coherent conditional probability on $\mathcal{A}(\mathcal{P})|\mathcal{A}^{\varnothing}(\mathcal{P})$ iff there exists a (unique) weight system on $\mathcal{P}$ such that

$$P(E\,|\,H) = \frac{\sigma(E\wedge H^*)}{\sigma(H)} \quad \forall E\in\mathcal{A}(\mathcal{P}),\ \forall H\in\mathcal{A}^{\varnothing}(\mathcal{P}).$$

Finally, the following Thm. 2.2.5 [10] allows to extend any coherent conditional probability.

2.2.5 Extension theorem

Any coherent conditional probability on a not empty set $\mathcal{E}$ of conditional events can be extended to a coherent conditional probability on any superset of $\mathcal{E}$.

3 An algorithm for checking the ASL condition

It is useful for understanding the final procedure in 4.2 to discuss briefly a known algorithm for checking the ASL condition for an imprecise probability assessment.

3.1 Definitions

Let $\underline{P}$ be a lower probability assessment on $\mathfrak{I} = \{E_1|H_1,\dots,E_n|H_n\}$, $\underline{P}(E_i|H_i) = a_i$, $\mathfrak{I}_h \subset \mathfrak{I}$.

Define then $\mathcal{P}$ ($\mathcal{P}_h$) as the set of all coherent conditional (precise) probabilities dominating $\underline{P}$ on $\mathfrak{I}$ (on $\mathfrak{I}_h$), and, for $E_r|H_r\in\mathfrak{I}$ ($\in\mathfrak{I}_h$),

$m_r = \min_{P\in\mathcal{P}}\{P(E_r|H_r)\}$ ($m_r^{(h)} = \min_{P\in\mathcal{P}_h}\{P(E_r|H_r)\}$).

By 2.2.5, it is not restrictive to consider the probabilities in $\mathcal{P}$ ($\mathcal{P}_h$) on an arbitrary domain which includes $\mathfrak{I}$ ($\mathfrak{I}_h$).

Define $\mathcal{P}$ as the partition whose atoms are the not impossible logical products obtained developing the expression $\wedge_{i=1}^{n}[(E_i\wedge H_i)\vee(\overline{E}_i\wedge H_i)\vee\overline{H}_i]$. Note that $\mathfrak{I}\subset\mathcal{A}(\mathcal{P})|\mathcal{A}^{\varnothing}(\mathcal{P})$.

Let $I_1 = \{1,\dots,n\}$, $K_1 = \vee_{i\in I_1}H_i$. If $\overline{K}_1 \neq \varnothing$, then $\overline{K}_1$ is an atom of $\mathcal{P}$; we call $e_1,\dots,e_m$ the remaining atoms of $\mathcal{P}$. Define then $J_1 = \{j: e_j \Rightarrow K_1\} = \{1,\dots,m\}$, $\mathfrak{I}_0 = \mathfrak{I}$, $\mathcal{P}_0 = \mathcal{P}$.

The function $\delta(E\Rightarrow F)$ is equal to one if $E\Rightarrow F$ holds, to zero otherwise.

3.2 Checking the ASL condition

Step h (first step: h = 1)

Consider the linear system (Sh):

$$(S_h)\begin{cases}\sum_{j\in J_h} x_j\delta(e_j \Rightarrow E_i\wedge H_i) - a_i\sum_{j\in J_h} x_j\delta(e_j\Rightarrow H_i) \geq 0,\\ \text{for } i\in I_h;\\ \sum_{j\in J_h} x_j = 1,\ x_j \geq 0\ \ (j\in J_h)\end{cases}$$

If (Sh) has no solution, $\underline{P}$ is not ASL.

If $(x_j(h))_{j\in J_h}$ is a solution for (Sh), since $(x_j(h))_{j\in J_h}$ is a non-negative vector whose components sum up to one, it is also a coherent probability P_h on the conditional partition $\mathcal{P}|K_h$, putting $P_h(e_j|K_h) = x_j(h)$. As well known, P_h has a unique coherent extension on every $E|K_h\in\mathcal{A}(\mathcal{P}|K_h)$ by additivity. Then, also observing that for $e_j \Rightarrow K_h$ and $E \Rightarrow K_h$ it is

(2) $\quad \delta(e_j|K_h \Rightarrow E|K_h) = \delta(e_j \Rightarrow E)$,

we have $P_h(E|K_h) = \Sigma_{j\in J_h}x_j(h)\delta(e_j \Rightarrow E)$.

Referring to the solution found for (Sh), define then $I_{h+1} = \{i\in I_h: P_h(H_i|K_h) = 0\}$, $K_{h+1} = \vee_{i\in I_{h+1}}H_i$, $J_{h+1} = \{j\in J_h: e_j \Rightarrow K_{h+1}\}$, $\mathfrak{I}_h = \{E_i|H_i: H_i \Rightarrow K_{h+1}\}$.

We define also $K_h^+ = \vee\{e_j\in\mathcal{P}: P_h(e_j|K_h) > 0\}$, which will be used in 3.4 (c). Note that $K_h^+ \neq \varnothing$,

$K_h^+ \wedge K_k = \varnothing$ if $h < k$, $K_h^+ \wedge K_k^+ = \varnothing$ if $h \neq k$.

If $K_{h+1} = \varnothing$, the procedure terminates and $\underline{P}$ is ASL; otherwise continue with step h+1.

3.3 Remarks on system (Sh)

(a) If $P_h \in \mathcal{P}_{h-1}$, then $P_h(e_j|K_h)$, $j \in J_h$, is a solution for (Sh).

In fact, recalling (2), the i-th inequality in (Sh) may be written in terms of P_h as

$$(3) \qquad P_h(E_i \wedge H_i|K_h) - a_i P_h(H_i|K_h) \geq 0,$$

and (3) holds either trivially (if $P_h(H_i|K_h) = 0$) or else because of the following sequence (where the relation $H_i \Rightarrow K_h$ and the product rule are also exploited):

$$a_i = \underline{P}(E_i|H_i) \leq P_h(E_i|H_i) = P_h(E_i|H_i \wedge K_h) =$$

$$= \frac{P_h(E_i \wedge H_i \mid K_h)}{P_h(H_i \mid K_h)}.$$

The conclusion follows.

(b) Note also that (3) cannot hold trivially for all $i \in I_h$ ($P_h(\vee_{i\in I_h} H_i|K_h) = P_h(K_h|K_h) = 1$); I_{h+1} is precisely the set of all $i \in I_h$ such that the i-th inequality is trivially ($0 - a_i \cdot 0 \geq 0$) verified by the solution found for (Sh). It is easily seen that the inequalities in system (Sh+1) are then a proper subset of the inequalities of (Sh).

3.4 Algorithm consistency

This algorithm is discussed from various viewpoints in [2], [6], [7]. We give here a different proof of its consistency, which will be useful in the next section. Precisely:

(a) *the algorithm terminates in a finite number of steps*.

In fact, the number of inequalities in the first row of (Sh) is finite (n for h = 1) and strictly decreasing as h increases, by 3.3 (b).

(b) *If the last system* (St) *is incompatible*, $\underline{P}$ *is not* ASL.

In fact, by 3.3 (a) the set $\mathcal{P}_{t-1}$ is then empty, and so is also $\mathcal{P} \subset \mathcal{P}_{t-1}$ (3.1). Then $\underline{P}$ is not ASL by 2.1.5.

(c) *If the last system* (St) *is consistent*, $\underline{P}$ *is* ASL.

To prove this, it is sufficient to determine a coherent conditional probability P* which dominates $\underline{P}$ on $\mathfrak{I}$ (2.1.5) and is defined on $\mathcal{A}(\mathcal{P})|\mathcal{A}^{\varnothing}(\mathcal{P})$ ($\supset \mathfrak{I}$). By 2.2.4, we can assign P* by means of its weight system, defined as follows:

- $\mathcal{P}' = \{K_1^+,\ldots,K_t^+,D\}$, $D = \neg(\vee_{h=1}^{t} K_h^+)$. It is easy to verify that $\mathcal{P}'$ is a partition;
- $K_1^+ > \ldots > K_t^+ > D$;
- $\pi_h(e_j) = x_j(h)$, $\forall e_j$: $e_j \Rightarrow K_h^+$, $h = 1,\ldots,t$; π_D arbitrary and positive.

It is not difficult to see that, given P* just defined, choosing arbitrarily $E_r|H_r \in \mathfrak{I}$ there exists (a unique) h such that $H_r^* \Rightarrow K_h^+$; moreover, given $E \in \mathcal{A}(\mathcal{P})$, the following equality holds (σ is the sum function (2.2.3) of P*):

$$(4) \qquad \sigma(E \wedge H_r^*) = \Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow E \wedge H_r).$$

Then, applying 2.2.4, (4) twice (for $E = E_r$, $E = H_r$) together with $\sigma(H_r^*) = \sigma(H_r)$, and the r-th constraint in (Sh), we get:

$$P^*(E_r|H_r) = \frac{\sigma(E_r \wedge H_r^*)}{\sigma(H_r)} =$$

$$= [\Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow E_r \wedge H_r)] \,/\, [\Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow H_r)] \geq$$

$$\geq a_r = \underline{P}(E_r|H_r).$$

Hence $P^* \geq \underline{P}$. ∎

3.5 Remarks

(a) It has been proved in the final part of 3.4 (c) that the probability P* dominates $\underline{P}$ on $\mathfrak{I}$. Further, it is easily seen that the restriction of P* on $\mathcal{P}|K_1$ is the solution found for (S1): $P^*(e_j|K_1) = x_j(1) = P_1(e_j|K_1)$.

It follows from these two facts that, if $\underline{P}$ is ASL, any solution of (S1) is a probability on $\mathcal{P}|K_1$ which can be always coherently extended to a probability dominating $\underline{P}$ on $\mathfrak{I}$.

An analogue conclusion may be drawn for the solutions of system (Sh). In fact, by applying the same argument of 3.4 (c) to a sequence of systems starting with (Sh) instead of (S1), we determine a probability P_h^* on $\mathcal{A}(\mathcal{P})|\mathcal{A}^{\varnothing}(\mathcal{P})$ dominating $\underline{P}$ on $\mathfrak{I}_{h-1}$. Further, it is $P_h^*(e_j|K_h) = x_j(h) = P_h(e_j|K_h)$.

It follows that, if $\underline{P}$ is ASL on $\mathfrak{I}_{h-1}$, any solution of (Sh) can be always extended to a probability dominating $\underline{P}$ on $\mathfrak{I}_{h-1}$.

(b) If $\underline{P}$ is ASL, a solution for system (Sh) can be found operationally by optimising a linear function f, subject to (Sh). The choice of f has a crucial importance in the procedure 4.2 for finding the least-committal lower probability, because of the additional probabilistic information it gives us. Observe for this that if $f = \Sigma_{j\in J_h} x_j \delta(e_j \Rightarrow E)$, $E \in \mathcal{A}(\mathcal{P})$, $E \Rightarrow K_h$, by minimising f subject to (Sh) we obtain $\min f = \min P_h(E|K_h)$, where the minimum is over all coherent probabilities P_h dominating $\underline{P}$ on $\mathfrak{I}_{h-1}$ (3.3 (a), 3.5 (a)).

(c) Let $E_r|H_r$ be an event of $\mathfrak{I}_{h-1}$. We shall consider in section 4 the following system

$$(T_h)\begin{cases}\sum_{j\in J_h} x_j\delta(e_j \Rightarrow E_i \wedge H_i) - a_i \sum_{j\in J_h} x_j\delta(e_j \Rightarrow H_i) \geq 0, \\ \text{for } i \in I_h; \\ \sum_{j\in J_h} x_j\delta(e_j \Rightarrow H_r) = 1, x_j \geq 0 \ (j \in J_h)\end{cases}$$

Every solution $(x_j(h))_{j\in J_h}$ of (Th) corresponds to a solution $(x_j(h)/k)_{j\in J_h}$, $k = \Sigma_{j\in J_h} x_j(h)$, of (Sh) and it is therefore proportional to a probability P_h ($x_j(h) = kP_h(e_j|K_h)$, $\forall j \in J_h$). By the normalisation constraint in (Th), $P_h(H_r|K_h)$ is strictly positive: $P_h(H_r|K_h) = \Sigma_{j\in J_h} P_h(e_j|K_h)\delta(e_j \Rightarrow H_r) = 1/k > 0$. Clearly, P_h can be extended, by 3.5 (a), to a probability dominating $\underline{P}$ on $\mathfrak{I}_{h-1}$.

4 From ASL to least-committal probabilities

Suppose now that $\underline{P}$ is ASL. To find its least-committal coherent correction on $\mathfrak{I}$, by 2.1.7 we have to find $m_r = \min_{P\in\mathcal{P}} \{P(E_r|H_r)\}$, for $r = 1,\ldots,n$. The following proposition solves the problem in a special case.

4.1 Proposition

Let $\underline{P}$ be an ASL lower probability assessment on $\mathfrak{I}_{h-1}$ (the notation is as in sect. 3). Consider the linear programming (LP) problem (Ph):

(Ph) $\mu_r^{(h)} = \min \Sigma_{j\in J_h} x_j\delta(e_j \Rightarrow H_r)$, subject to (Sh).

If $\mu_r^{(h)} > 0$, then $m_r^{(h-1)}$ (3.1) is the solution of the following auxiliary LP problem (Ah):

(Ah) $m_r^{(h-1)} = \min \Sigma_{j\in J_h} x_j\delta(e_j \Rightarrow E_r \wedge H_r)$, subject to (Th).

Proof It ensues from 3.3 (a) and 3.5 (a) that the solutions of system (Sh) correspond to all probabilities of $\mathcal{P}_{h-1}$.

It ensues from an argument similar to that of 3.3 (a) that every probability belonging to the set $\mathcal{P}_{h-1}^+ = \{P_h \in \mathcal{P}_{h-1}: P_h(H_r|K_h) > 0\}$ is proportional to a solution of system (Th) and from 3.5 (c) that every solution of (Th) is proportional to a probability of $\mathcal{P}_{h-1}^+$. Given a solution $(x_j(h))_{j\in J_h}$ for (Th), it is then, $\forall j \in J_h$, $x_j(h) = kP_h(e_j|K_h)$, $k = \Sigma_{j\in J_h} x_j(h)$, $P_h \in \mathcal{P}_{h-1}^+$.

When the 'if' condition in the hypothesis of this proposition holds (that is, by 3.5 (b), when $P_h(H_r|K_h) > 0 \ \forall P_h \in \mathcal{P}_{h-1}$), it is $\mathcal{P}_{h-1}^+ = \mathcal{P}_{h-1}$, so that also the solutions of (Th) identify all $P_h \in \mathcal{P}_{h-1}$.

Therefore, for each $P_h \in \mathcal{P}_{h-1}$, the summations in (Th) may be written as:

$\Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow E_r \wedge H_r) = kP_h(E_r \wedge H_r|K_h)$,

$\Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow H_r) = kP_h(H_r|K_h)$.

From this, relation $H_r \Rightarrow K_h$, the product rule and the normalisation constraint in (Th) we get:

$$P_h(E_r|H_r) = P_h(E_r|H_r \wedge K_h) = \frac{P_h(E_r \wedge H_r | K_h)}{P_h(H_r | K_h)} =$$

$$= [\Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow E_r \wedge H_r)] / [\Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow H_r)] =$$

$$= \Sigma_{j\in J_h} x_j(h)\delta(e_j \Rightarrow E_r \wedge H_r).$$

It ensues that solving the LP problem (Ah) we determine $m_r^{(h-1)}$. ■

By applying 4.1 with $h = 1$, if the solution of the LP problem (P1) is such that $\min \Sigma_{j\in J_1} x_j\delta(e_j \Rightarrow H_r) > 0$, i.e. $P_1(H_r|K_1) > 0$, $\forall P_1 \in \mathcal{P}_0 = \mathcal{P}$, it is possible to find $m_r = m_r^{(0)}$ by solving the LP problem (A1). Otherwise, the procedure continues as shown in 4.2.

4.2 Finding the least-committal lower probability

We describe now a general procedure for finding the least-committal lower probability $\underline{P}^*$ for $\underline{P}$, supposing that $\underline{P}$ is ASL. It determines $m_r = \underline{P}^*(E_r|H_r)$ for each $E_r|H_r$. Clearly, $\underline{P}^* = \underline{P}$ iff $\underline{P}$ is coherent.

Step h (first step: h = 1)

Find a solution for the LP problem (Ph) (4.1):

(Ph) $\mu_r^{(h)} = \min \Sigma_{j\in J_h} x_j\delta(e_j \Rightarrow H_r)$, subject to (Sh).

If $\mu_r^{(h)} > 0$, then m_r is the solution of the LP problem (Ah) (4.1):

(Ah) $m_r = \min \Sigma_{j\in J_h} x_j\delta(e_j \Rightarrow E_r \wedge H_r)$, subject to (Th).

If $\mu_r^{(h)}=0$, and $x_j(h)=P_h(e_j|K_h)$, $j\in J_h$, is the solution found for (S_h) in problem (P_h), define, as in 3.2, $I_{h+1}=\{i\in I_h: P_h(H_i|K_h)=0\}$, $K_{h+1}=\vee_{i\in I_{h+1}}H_i$, $J_{h+1}=\{j\in J_h: e_j\Rightarrow K_{h+1}\}$, and continue with step h+1.

4.3 Algorithm consistency theorem

Let $\underline{P}$ be ASL on $\mathfrak{I}$. Algorithm 4.2 determines $m_r=\underline{P}^*(E_r|H_r)$ $\forall E_r|H_r\in\mathfrak{I}$.

Proof 4.2 terminates in a finite number of steps. In fact, since the sequence $\{K_h\}$ is strictly monotone and decreasing (because so is the sequence $\{I_h\}$ by 3.4 (a)) and $H_r\Rightarrow K_h$ holds for each problem (P_h), the stopping condition $\mu_h>0$ (i.e. $\min P_h(H_r|K_h)>0$) is sooner or later met.

At the last step s, by 4.1, the algorithm determines $m_r^{(s-1)}$. The case $s=1$ is trivial. Suppose then $s>1$.

Clearly, it is $m_r^{(s-1)}\le m_r$, since each probability dominating $\underline{P}$ on $\mathfrak{I}$ dominates $\underline{P}$ on $\mathfrak{I}_{s-1}$.

Hence the thesis is proved if we can find $P_r^*\in\mathcal{P}$ such that $P_r^*(E_r|H_r)=m_r^{(s-1)}$. To do this, by 2.2.4, we can assign the weight system of P_r^* on the partition $\mathcal{P}$, thereby defining P_r^* on $\mathcal{A}(\mathcal{P})|\mathcal{A}^{\varnothing}(\mathcal{P})$ ($\supset\mathfrak{I}$). Calling $(\bar{x}_j(s))_{j\in J_s}$ the solution found for system (T_s) in the LP problem (A_s), complete for this the sequence of solutions for systems (S_1),…,(S_{s-1}) by assigning to system (S_s) the solution $(y_j(s))_{j\in J_s}$ obtained by multiplying each component of $(\bar{x}_j(s))_{j\in J_s}$ by $1/\Sigma_{j\in J_s}\bar{x}_j(s)$. ($S_1$),…,($S_s$) is a particular subsequence in the procedure for checking the ASL condition for $\underline{P}$; let us complete it in the way described in sect. 3, if necessary (i.e., if $K_{s+1}\neq\varnothing$), getting in this way a sequence of solutions for (S_1),…,(S_t) ($t\ge s$). Referring to this sequence, define the weight system of P_r^* as in 3.4 (c) (so that, for instance, $\mathcal{P}'=\{K_1^+,\dots,K_t^+,D\}$; the only difference is that $\pi_s(e_j)$ is equal to $y_j(s)$ instead of $x_j(s)$).

In order to prove that it is

$$P_r^*(E_r|H_r)=m_r^{(s-1)}=\Sigma_{j\in J_s}\bar{x}_j(s)\delta(e_j\Rightarrow E_r\wedge H_r)$$

note firstly that by construction $H_r^*=H_r\wedge K_s^+$ and that, given $E\in\mathcal{A}(\mathcal{P})$, the following equality, analogue of (4) in 3.4 (c), holds (σ is the sum function of P_r^*):

$$\text{(4')}\qquad \sigma(E\wedge H_r^*)=\Sigma_{j\in J_s}y_j(s)\delta(e_j\Rightarrow E\wedge H_r).$$

We obtain now the following equalities (in the fourth, apply (4') twice, for $E=E_r$ and $E=H_r$):

$$m_r^{(s-1)}=\Sigma_{j\in J_s}\bar{x}_j(s)\delta(e_j\Rightarrow E_r\wedge H_r)=$$
$$=[\Sigma_{j\in J_s}\bar{x}_j(s)\delta(e_j\Rightarrow E_r\wedge H_r)]\,/\,[\Sigma_{j\in J_s}\bar{x}_j(s)\delta(e_j\Rightarrow H_r)]=$$
$$=[\Sigma_{j\in J_s}y_j(s)\delta(e_j\Rightarrow E_r\wedge H_r)]\,/\,[\Sigma_{j\in J_s}y_j(s)\delta(e_j\Rightarrow H_r)]=$$
$$=\frac{\sigma(E_r\wedge H_r^*)}{\sigma(H_r)}=P_r^*(E_r|H_r).\qquad\blacksquare$$

4.4 An operational generalisation

In order to apply 4.2, it is assumed that $\underline{P}$ is ASL. Nevertheless, when this is not known a priori it is possible to run one procedure (‘merging’ 3.2 and 4.2) which both checks the ASL condition for $\underline{P}$ and finds its least-committal lower probability, thereby verifying also whether $\underline{P}$ is coherent (therefore, it improves the algorithm for checking coherence introduced in [11]).

Refer for this to $E_r|H_r\in\mathfrak{I}$ and operate as follows at each step h ($h\ge 1$):

h.a) consider problem (P_h) in step h of 4.2. If (P_h) is infeasible, $\underline{P}$ is not ASL and the procedure stops; otherwise define I_{h+1}, K_{h+1}, J_{h+1} as in 4.2, both when $\mu_r^{(h)}=0$ and when $\mu_r^{(h)}>0$;

h.b.1) if $\mu_r^{(h)}=0$, go to step h+1.a);

h.b.2) if $\mu_r^{(h)}>0$ and $K_{h+1}=\varnothing$, $\underline{P}$ is ASL on $\mathfrak{I}$; solve then (A_h), whose solution is $\underline{P}^*(E_r|H_r)$;

h.b.3) if $\mu_r^{(h)}>0$ and $K_{h+1}\neq\varnothing$, execute from step h+1 onwards the procedure 3.2;

- if doing so a system is incompatible, $\underline{P}$ is not ASL on $\mathfrak{I}$;
- if all systems are consistent, $\underline{P}$ is ASL on $\mathfrak{I}$; in this case, solve (A_h) to find $\underline{P}^*(E_r|H_r)$.

If the above procedure states that $\underline{P}$ is ASL, execute 4.2 for all $E_i|H_i\in\mathfrak{I}$, $E_i|H_i\neq E_r|H_r$, finding $\underline{P}^*$, which is equal to $\underline{P}$ if and only if $\underline{P}$ is coherent.

4.5 An example

Given the partition $\mathcal{P}=\{e_1,e_2,e_3,e_4\}$ and the events $E=e_1\vee e_2\vee e_4$, $F=e_1\vee e_4$, $H=e_1\vee e_2\vee e_3$, consider the assignment $\underline{P}(F|E\wedge H)=1/3$, $\underline{P}(E|H)=3/4$, $\underline{P}(F|H)=1/2$, $\underline{P}(F|E)=1/2$. Since we do not know whether $\underline{P}$ is ASL on $\mathfrak{I}=\{F|E\wedge H, E|H, F|H, F|E\}$, let us apply 4.4, starting from $F|E\wedge H$:

1.a) Problem (P_1) is

$\mu^{(1)}=\min(x_1+x_2)$, subject to

$$\begin{cases} x_1 - 1/3(x_1 + x_2) \geq 0 \\ x_1 + x_2 - 3/4(x_1 + x_2 + x_3) \geq 0 \\ x_1 - 1/2(x_1 + x_2 + x_3) \geq 0 \\ x_1 + x_4 - 1/2(x_1 + x_2 + x_4) \geq 0 \\ x_1 + x_2 + x_3 + x_4 = 1, \quad x_j \geq 0 \ (j = 1,\dots,4) \end{cases}$$

1.b.1) Since $\mu^{(1)} = 0$ (obtained for $x_j = 0$ $(j \neq 4)$ and $x_4 = 1$), we pass to step 2.a), observing that $K_2 = e_1 \vee e_2 \vee e_3$.

2.a) Problem (P_2) is

$\mu^{(2)} = \min(x_1 + x_2)$, subject to

$$\begin{cases} x_1 - 1/3(x_1 + x_2) \geq 0 \\ x_1 + x_2 - 3/4(x_1 + x_2 + x_3) \geq 0 \\ x_1 - 1/2(x_1 + x_2 + x_3) \geq 0 \\ x_1 + x_2 + x_3 = 1, \quad x_j \geq 0 \ (j = 1,2,3) \end{cases}$$

2.b.2) Since $\mu^{(2)} = 3/4$ (obtained for $x_1 = 1/2$, $x_2 = x_3 = 1/4$) and $K_3 = \emptyset$, $\underline{P}$ is ASL. To find $\underline{P}^*(F|E \wedge H)$ we solve the following problem (A_2):

$m^{(1)} = \min x_1$, subject to

$$\begin{cases} x_1 - 1/3(x_1 + x_2) \geq 0 \\ x_1 + x_2 - 3/4(x_1 + x_2 + x_3) \geq 0 \\ x_1 - 1/2(x_1 + x_2 + x_3) \geq 0 \\ x_1 + x_2 = 1, \ x_j \geq 0 \ (j = 1,2,3) \end{cases}$$

We obtain (for $x_1 = x_2 = 1/2$ and $x_3 = 0$) $m^{(1)} = \underline{P}^*(F|E \wedge H) = 1/2 > \underline{P}(F|E \wedge H)$.

By applying 4.2 to the events of $\Im' = \Im - \{F|E \wedge H\}$, it can be verified that $\underline{P}^* = \underline{P}$ on $\Im'$. Therefore, we conclude that $\underline{P}$ is not coherent on $\Im$ and we must raise our evaluation on $F|E \wedge H$ from 1/3 to 1/2 in order to achieve coherence.

References

[1] S. Amarger, D. Dubois, H. Prade. Constraint Propagation with Imprecise Conditional Probabilities. In *Proc. of the 7th Conference on Uncertainty in Artificial Intelligence*, 26-34, Los Angeles, July 1991.

[2] G. Coletti. Coherent Numerical and Ordinal Probabilistic Assessments. *IEEE Trans. Systems Man Cybernet.*, 24 (12), 1747-1754, 1994.

[3] G. Coletti, R. Scozzafava. Characterization of Coherent Conditional Probabilities as a Tool for their Assessment and Extension. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, 4 (2), 103-127, 1996.

[4] L. Crisma. Prolungamenti di probabilità condizionate localmente coerenti. *Quad. n.5/93 del Dip. Mat. Appl. 'B. de Finetti'*, Univ. di Trieste, 1993.

[5] L. Crisma. Lezioni di calcolo delle probabilità. Edizioni Goliardiche, Trieste, 1997.

[6] A. Gilio. Probabilistic Consistency of Conditional Probability Bounds. In *Proc. of IPMU '94*, 455-460, Paris, July 1994.

[7] A. Gilio. Algorithms for Precise and Imprecise Conditional Probability Assessments. In *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence* (G. Coletti, D. Dubois, R. Scozzafava eds.), 231-254, Plenum Press, 1995.

[8] A. Gilio. Algorithms for Conditional Probability Assessments. In *Bayesian Statistics and Econometrics* (D. A. Berry, K. M. Chalower, J. K. Geweke eds.), 29-39, Wiley, 1996.

[9] A. Gilio, S. Ingrassia. Totally coherent interval-valued probability assessments. In *Proc. of the 4th WUPES*, 50-61, Prague, January 1997.

[10] S. Holzer. On Coherence and Conditional Prevision. *Boll. Un. Mat. Ital.*, Serie VI (IV-C), 441-460, 1985.

[11] P. Vicig. An Algorithm for Imprecise Conditional Probability Assessments in Expert Systems. In *Proc. of IPMU'96*, vol. 1, 61-66, Granada, July 1996.

[12] P. Vicig. Upper and Lower Bound for Coherent Extensions of Conditional Probabilities given on Finite Sets. *Quad. n.7/97 del Dip. Mat. Appl. 'B. de Finetti'*, Univ. di Trieste, 1997.

[13] P. Walley. Coherent Lower (and Upper) Probabilities. *Research Report of the Dep. of Statistics*, Univ. of Warwick, 1981

[14] P. Walley. Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, 1992.

[15] P. Walley. Coherent upper and lower previsions. *The Imprecise Probabilities Project* (http://eepkibm2.rug.ac.be/~ipp), 1997.

[16] P. M. Williams. Notes on Conditional Previsions. *Research Report of the School of Math. and Phys. Sci.*, Univ. of Sussex, 1975.