A simplified algorithm for inference by lower conditional probabilities

Andrea Capotorti
Dip. di Matematica e Informatica,
Perugia, Italy
capot@dipmat.unipg.it

Barbara Vantaggi
Dip. di Metodi e Modelli Matematici,
Univ. "La Sapienza", Roma, Italy
vantaggi@dmmm.uniroma1.it

Abstract

Thanks to the notion of locally strong coherence, the satisfiability of proper logical conditions on subfamilies of the initial domain helps to simplify inferential processes based on lower conditional assessments. Actually, these conditions avoid also round errors that, on the other hand, appear solving numerical systems. In this paper we introduce new conditions to be applied to sets of particular pairs of events. With respect to more general conditions already proposed, they avoid an exhaustive search, so that a sensible time-complexity reduction is possible. The usefulness of these rules in inferential processes is shown by a diagnostic medical problem with thyroid pathology.

Keywords. Partial lower conditional probability assessments, coherent inference, locally strong coherence.

1 Introduction

Decision models based on partial (precise or interval valued) conditional probabilities are nowadays well established. The adoption of partial assessments is due to their generality and wide applicability. Its roots are in the de Finetti's ideas (see for instance, [13, 14]) and nowadays has found a wide application (as an incomplete, but relevant, reference see [2, 6, 8, 10, 12, 17]). The peculiarity of incompleteness (i.e. the domain of the valuation is not necessarily a well structured set, like an algebra or a σ-algebra) distinguishes this theory from the usual ones (based on a measure-theoretic approach), where a complete initial evaluation is required. Sometimes the completeness hypothesis is not realistic in applications, hence partial assessments can better represent uncertainty.

On the other hand, such generality has the drawback that the assessment must be coherent with a specified numerical framework and this must be operationally tested.

In our paper we will adopt an interval valued conditional probability representation of uncertainty. *Imprecise* probabilities are nowadays taken in consideration by many authors with different motivations (for a wide exposition of the latest "trends" refer to [15]). Without facing the problem of the meaningfulness of "imprecise" probabilities, we intend extreme values (possibly coinciding) assigned either as a (initial) precise evaluation or as lower and upper probabilities (possibly obtained by previous inferences). For a deeper analysis refer to [11, 12].

Usually, precise or mixed models are treated as special cases of interval valued assessments. Starting with assessments $\big(\underline{P}(E_i|H_i), \overline{P}(E_i|H_i)\big)$ on generic conditional events $E_i|H_i$, $i = 1, \ldots, n$, they are all reduced to lower probability assessments by the duality property

$$\overline{P}(E_i|H_i) = 1 - \underline{P}(E_i^c|H_i). \tag{1}$$

In [10], for coherent inference with lower probabilities, a sequence of unconditional probabilities must be found for *each* element of the domain. This is operationally obtained by solving particular sequences of linear systems. Anyhow, because of the possibility of giving zero probabilities to conditioning events (see [8]) and of the derived notion of *locally strong coherence* [4], we avoid to build and solve some of such linear systems. In fact, the constraints represented by the linear systems can be equivalently fulfilled by solvability of particular logical configurations among the events.

A further sensible complexity reduction is also possible with a careful reduction from the lower-upper to the lower assessment. In fact, if (1) is applied to the entire evaluation domain, there is a duplication of data with respect to the events with a precise assessment associated. Operationally this duplication brings to build and solve not needed linear optimization problems. But we will avoid this unpleasant feature keeping trace of events, with precise probability

values, along all the procedure.

Moreover, by (1) pairs of elements with common conditioning event ($E_i|H_i$ and $E_i^c|H_i$) are introduced and they must be always considered together for locally strong coherence applicability (as it will clearly appear in the next sections). For this purpose, suitable logical conditions will be introduced to deal with such "twin" pairs.

We will also present a diagnostic problem about thyroid pathology to show applicability of the proposed machinery.

2 Preliminaries

We represent the starting domain of evaluation (knowledge base) as a family of conditional events $\mathcal{E} = \{E_1|H_1, \ldots, E_n|H_n\}$ with a set of logical constraints $\mathcal{C}$ among unconditional parts (representing incompatibility, inclusion or identity relations). Starting by $\mathcal{E}$ and $\mathcal{C}$, it is possible to build (just as an operational tool) the set of atoms $\mathcal{A}_\mathcal{E}$ contained in $H^0 = \bigvee_{i=1}^n H_i$. Atoms are generated by the unconditional parts $\mathcal{U}_\mathcal{E} = \{E_1, \ldots, E_n, H_1, \ldots, H_n\}$ and are built combining the possible occurrences of $E_i \wedge H_i$, $E_i^c \wedge H_i$ and H_i^c for each $E_i|H_i \in \mathcal{E}$, so that their cardinality will be of $O(3^n)$. In the sequel, when it will not cause misunderstanding, we will omit the conjunction connective $\wedge$.

An assessment $\mathbf{P}_\mathcal{E}$ on the domain $\mathcal{E}$ is given and, without loss of generality, we can suppose there are exactly k precise probability evaluations $P(E_i|H_i) = p_i$, $i = 1 \ldots k$, $k \in \{0, 1, \ldots, n\}$, while uncertainty on the other conditional events is expressed by distinct lower and upper probability bounds

$$\underline{P}(E_j|H_j) = \underline{p}_j \text{ and } \overline{P}(E_j|H_j) = \overline{p}_j \quad j = k+1, \ldots, n.$$

Starting from such partial assessment $\mathbf{P}_\mathcal{E}$, an inference goal is posed: to find *all coherent* conditional probabilities for a new event $E|H$, with E and H arbitrarily chosen, also among events logically not dependent on $\mathcal{U}_\mathcal{E}$. In this way the classical statistical inference based on Bayes Theorem is generalized.

In the following we sketch how a sound inference could be performed.

If $\mathbf{P}_\mathcal{E}$ is transformed by (1) in a lower probability assessment [1], we look for a class $\mathcal{P}$ of coherent "precise" conditional probabilities on $\mathcal{E}$ whose lower envelope coincides with $\mathbf{P}_\mathcal{E}$ (i.e. $\mathbf{P}_\mathcal{E}$ is coherent), so that we can compute

$$\underline{p} = \inf_{P \in \mathcal{P}} \underline{P}(E|H) \quad \overline{p} = \sup_{P \in \mathcal{P}} \overline{P}(E|H) \tag{2}$$

(where the single $\underline{P}(E|H)$ and $\overline{P}(E|H)$ are obtained applying de Finetti's fundamental theorem of prevision for each conditional probability distribution $P \in \mathcal{P}$).

Anyhow, thanks to Theorem 4 in [9] it is possible to skip to build the entire class $\mathcal{P}$. In fact, one of the three equivalence stated by the quoted theorem is:

Theorem 1 *Let $\mathcal{E} = \{E_1|H_1, \ldots, E_{2n}|H_{2n}\}$ be a finite set of conditional events and $\mathbf{P}_\mathcal{E}$ a lower probability assessment on it. Then the following statements are equivalent:*

i) there exists a class $\mathcal{P}$ of coherent conditional probabilities on $\mathcal{E}$ whose lower envelope coincide with $\mathbf{P}_\mathcal{E}$;

ii) for any $E_i|H_i \in \mathcal{E}$ there exists a class (usually not unique) of unconditional probabilities $\{P_0^i, P_1^i, \ldots, P_{\omega_i}^i\}$, each probability $P_{\alpha_i}^i$, $\alpha_i = 0, \ldots, \omega_i$ defined on a proper subset of atoms $\mathcal{A}_{\alpha_i}^i \subseteq \mathcal{A}_\mathcal{E}$, such that there is a unique $\beta_i \in \{0, \ldots, \omega_i\}$ with

$$\begin{aligned} \sum_{A_r \subseteq H_i} P_{\beta_i}^i(A_r) &> 0 \\ \underline{p}_i &= \frac{\sum_{A_r \subseteq E_i H_i} P_{\beta_i}^i(A_r)}{\sum_{A_r \subseteq H_i} P_{\beta_i}^i(A_r)} \end{aligned} \tag{3}$$

and, for each other $E_j|H_j \in \mathcal{E}$, there exists a unique $\delta_{ij} \in \{0, \ldots, \omega_i\}$ with

$$\begin{aligned} \sum_{A_r \subseteq H_j} P_{\delta_{ij}}^i(A_r) &> 0 \\ \underline{p}_j &\leq \frac{\sum_{A_r \subseteq E_j H_j} P_{\delta_{ij}}^i(A_r)}{\sum_{A_r \subseteq H_j} P_{\delta_{ij}}^i(A_r)} \end{aligned} \tag{4}$$

and moreover $\mathcal{A}_{\alpha_i'}^i \subset \mathcal{A}_{\alpha_i''}^i$ for $\alpha_i' > \alpha_i''$ and $P_{\alpha_i''}^i(A_r) = 0$ if $A_r \in \mathcal{A}_{\alpha_i'}^i$.

Practically speaking, the class $\mathcal{P}$ can be composed by conditional probabilities P represented by the classes $\{P_0^i, P_1^i, \ldots, P_{\omega_i}^i\}$, $i = 1, \ldots, 2n$ and, in the same paper, it is shown such classes are attainable as solution of linear programming sequences $\mathcal{L}_i = \{L_{\alpha_i}^i\}$,

[1] Note that applying (1) the domain $\mathcal{E}$ of the assessment $\mathbf{P}_\mathcal{E}$ has a double number of events $\mathcal{E} = \{E_1|H_1, \ldots, E_{2n}|H_{2n}\}$ with $E_{n+j}|H_{n+j} \equiv E_j^c|H_j$ and $\underline{p}_{n+j} = 1 - \overline{p}_{n+j}$, $j = 1, \ldots, n$. Anyhow the atoms $\mathcal{A}_\mathcal{E}$ remain the same after this transformation.

$i = 1, \ldots, 2n$, $\alpha_i = 0, \ldots, \omega_i$, of the form

$$L^i_{\alpha_i} = \begin{cases} \sum_{A_r \subseteq E_i H_i} P^i_{\alpha_i}(A_r) = \underline{p}_i \sum_{A_r \subseteq H_i} P^i_{\alpha_i}(A_r) \\ \qquad\qquad \text{if } P^i_{\alpha_i - 1}(H_i) = 0 \\ \sum_{A_r \subseteq E_j H_j} P^i_{\alpha_i}(A_r) \geq \underline{p}_j \sum_{A_r \subseteq H_j} P^i_{\alpha_i}(A_r) \\ \qquad\qquad \text{if } P^i_{\alpha_i - 1}(H_j) = 0 \\ \sum_{A_r \in \mathcal{A}_{\alpha_i}} P^i_{\alpha_i}(A_r) = 1 \end{cases} \tag{5}$$

The sequences $\mathcal{L}_i$ are $2n$ because the first "strict equality" constraint in (5) must rotate among the elements of $\mathcal{E} = \{E_1|H_1, \ldots, E_{2n}|H_{2n}\}$, with $E_{n+j}|H_{n+j} \equiv E^c_j|H_j$. So the index i in each sequence $\mathcal{L}_i$ (and hence also in each system $L^i_{\alpha_i}$) refers to a "leading" conditional event $E_i|H_i \in \mathcal{E}$.

As skillfully proved in [10], since elements $P \in \mathcal{P}$ are characterized by satisfying constraints $L^i_{\alpha_i}$, bounds (2) will be reached by those classes of solutions with the minimum number of relevant constraints. Such minimum number of constraints will be reached trying to force H to have probability zero, so that the optimization problem will be performed with the biggest α_i's indexes (note that the greater is α_i the lesser will be the number of constraints in $L^i_{\alpha_i}$).

Operationally speaking, actual extreme values $\underline{p}$ and $\overline{p}$ are obtained simply imposing the further constraint

$$\sum_{A_r \subseteq H} P^i_{\alpha_i}(A_r) = 0 \tag{6}$$

to the linear systems in $\mathcal{L}_i$ until it is possible (in the sequel we will denote by $\widetilde{\mathcal{L}}_i$ the sequences of systems $L^i_{\alpha_i}$ with such additional constraint). When, at level $\bar{\alpha}_i$, there will not be solution any more (or when all modified systems in $\widetilde{\mathcal{L}}_i$ will have solution) the extra constraint can be modified as

$$\sum_{A_r \subseteq H} P^i_{\bar{\alpha}_i}(A_r) = 1 \tag{7}$$

so that values

$$\begin{aligned} \underline{P}^i(EH) &= \min \sum_{A_r \subseteq EH} P^i_{\bar{\alpha}_i}(A_r) \\ \overline{P}^i(EH) &= \max \sum_{A_r \subseteq EH} P^i_{\bar{\alpha}_i}(A_r) \end{aligned}$$

can be computed under constraints $L^i_{\bar{\alpha}_i}$ and (7) (or simply set to 0 and 1, respectively). At the end it will result

$$\underline{p} = \inf_{i \in \{1, \ldots, 2n\}} \underline{P}^i(EH) \qquad \overline{p} = \sup_{i \in \{1, \ldots, 2n\}} \overline{P}^i(EH) \tag{8}$$

(note that the problem of obtaining $\underline{p}$ and $\overline{p}$ is reduced to linear programming problems, even if at the beginning it could appear a fractional problem).

In the context of coherence expressed trough a betting scheme, the new value $\underline{p}$ (called "natural extension") is defined as the supremum value of μ for which there are $\lambda_1 \geq 0, \ldots, \lambda_{2n} \geq 0$ such that

$$\sup \sum_{i=1}^{2n} \lambda_i H_i (E_i - \underline{p}_i) - H(E - \mu) < 0 \tag{9}$$

where the supremum is done over all atoms inside $H^0 \cup H$ and upper-case letters represent the indicator functions of corresponding events.

For the upper bound $\overline{p}$, we need to compute as before the lower bound $\underline{p}^c$ for $E^c|H$, so that $\overline{p} = 1 - \underline{p}^c$ is easily obtained.

Anyhow, in this approach a simultaneous consideration of all the atoms is needed, with all the complexity problems this implies. In [1], in the similar context of g-coherence for imprecise probabilities, it is shown how to simplify the procedure using linear properties of the gain. On the contrary we will proceed in a different direction.

Finally, since the solutions of the systems $\widetilde{\mathcal{L}}_i$ are also solutions of the systems in $\mathcal{L}_i$ (in fact they just fulfill the additional constraint (6) until the level $\bar{\alpha}_i$) they can be used to characterize the conditional probabilities $P \in \mathcal{P}$.

Anyway, note that whenever in $\mathbf{P}_{\mathcal{E}}$ there is some precise assessment (i.e. $k > 0$), all probability distributions $P^l_{\alpha_i}$ actually "strictly" fulfill the constraints associated to them, i.e. for $i = 1, \ldots, k$ and for $l = 1, \ldots, 2n$

$$P^l_{\alpha_i}(E_i H_i) = p_i P^l_{\alpha_i}(H_i) \tag{10}$$

(recall that for $i \in \{1, \ldots, k\}$ we have $\underline{p}_i = \overline{p}_i = p_i$).

Hence the linear systems $L^i_{\alpha_i}$ as shown in (5) can be modified imposing the strict equality constraint also for all conditional events $E_j|H_j$ with $j \in \{1, \ldots, k\}$ and not only for $E_i|H_i$. In this way, all the systems $L^i_{\alpha_i}$ would be exactly the same for $i = 1, \ldots, k$, so that the number of sequences strictly needed is just $1 + 2(n - k)$ instead of $2n$.

For this reason we return to the initial notation for the assessment domain $\mathcal{E} = \{E_1|H_1, \ldots, E_n|H_n\}$ and $\mathbf{P}_{\mathcal{E}}$ will denote a "mixed" assessment on it with $k \geq 0$ precise values. About the sequences of linear systems, $\widetilde{\mathcal{L}}_1$ will be the sequence associated to $E_1|H_1$ together with all conditional events with a precise assessment (if there are), while the other $\widetilde{\mathcal{L}}_l$'s will have an index that immediately refer to the "leading" conditional

event for that sequence (i.e. in $\widetilde{\mathcal{L}}_j$ a strict equality constraint will be associated to $E_j|H_j$ while in $\widetilde{\mathcal{L}}_{j^c}$ it will be done for $E_j^c|H_j$, with $j = k+1, \ldots, n$).

3 Locally strong coherence

As it has been described in Section 2, the steps to obtain the class $\mathcal{P}$ present a not negligible computational problem due to the exponential number of atoms $\mathcal{A}_{\mathcal{E}}$. Anyhow a simplification is possible.

Because of the possibility of giving zero probabilities to conditioning events (one of the main characteristics of de Finetti's approach) each linear system $L^i_{\alpha_i}$ can be simplified (following Coletti and Scozzafava's hints in [8]) looking for constraints trivially satisfied as $0 = 0$. In this way the number of actual inequalities (or equality) to fulfill in each single linear system $L^i_{\alpha_i}$ is reduced.

Note that conditioning events with zero probabilities are not "pathological" cases but can naturally arise in very simple situations (see the examples reported in [6]) also when the entire conditional assessment $\mathbf{P}_{\mathcal{E}}$ is strictly positive. Hence a sound methodology must include such a feature which, as we will see, can be suitably used (or "exploited" as said in [8]) to computationally simplify the entire procedure. So we will introduce "conditioning events with zero probabilities" even when they are not "strictly" needed, because they turn out to be helpful (all our results are based on this "positive drawback" of the zeroes).

To achieve a simplification it is necessary to "drive in the good direction" the zeroes. As already done for precise conditional probabilities in [3, 5] and for lower conditional probabilities in [4], this is possible thanks to the notion of *locally strong coherence* that will actually avoid to build (and solve) at least some of the first systems in the $\widetilde{\mathcal{L}}_i$'s, or possibly some entire sequence $\widetilde{\mathcal{L}}_i$.

To better understand our results we report the formal definitions of strong and locally strong coherence as already introduced in the quoted papers and adapted to our "mixed" assessment $\mathbf{P}_{\mathcal{E}}$ (anyhow the quoted papers remain the reference for a fully detailed exposition of this subject).

Strong coherence can be referred to any unconditional event:

Definition 1 *Let B be an unconditional event, $\mathcal{A}_{\mathcal{E}}^{|B}$ the set of atoms $A_r \in \mathcal{A}_{\mathcal{E}}$ such that $A_r B \neq \emptyset$, and $\mathbf{P}_{\mathcal{E}}$ an assessment such that the first k evaluations are precise. Then, $\mathbf{P}_{\mathcal{E}}$ is* strong coherent *with respect to B if and only if $H_i B \neq \emptyset$ for $i = 1, \ldots, n$ and the probability assessment P' defined as*

$$P'(E_i|H_iB) = \begin{cases} p_j & \text{if } \exists j \leq k \text{ s.t. } E_i|H_i \equiv E_j|H_j \\ \underline{p}_j & \text{if } \exists j > k \text{ s.t. } E_i|H_i \equiv E_j|H_j \\ 1-\overline{p}_j & \text{if } \exists j > k \text{ s.t. } E_i|H_i \equiv E_j^c|H_j \end{cases}$$

and such that $P'(H_iB) > 0$, for $i = 1, \ldots, n$, is coherent.

Note that assumption $H_iB \neq \emptyset$ assures that the conditional events $E_i|H_iB$ are well defined.

Such notion can be suitably used with the purpose to force a sub-group of conditional events to stay in a common layer (so that the application is "localized" to such sub-group).

Let $\mathcal{G}$ be a subfamily of $\mathcal{E}$, $\mathcal{R} = \mathcal{E} \setminus \mathcal{G}$ the residual and let us make a particular choice for the event B appearing in Definition 1

$$B = \left(\bigvee_{i:\, E_i|H_i \in \mathcal{R}} H_i \right)^c = \bigwedge_{i:\, E_i|H_i \in \mathcal{R}} H_i^c. \qquad (11)$$

Let $\mathbf{P}_{\mathcal{G}}$ be the restriction of $\mathbf{P}_{\mathcal{E}}$ to $\mathcal{G}$. If $\mathbf{P}_{\mathcal{G}}$ is strong coherent with respect to such B, then there exists an unconditional probability P_0 defined on $\mathcal{A}_{\mathcal{G}}^{|B}$ such that $P_0(H_iB) > 0$ for all $E_i|H_i \in \mathcal{G}$ and

$$\begin{aligned} p_i &= \frac{P_0(E_iH_iB)}{P_0(H_iB)} \quad \text{if } i \leq k \\ \underline{p}_i &\leq \frac{P_0(E_iH_iB)}{P_0(H_iB)} \quad \text{if } i > k \\ 1-\overline{p}_i &\leq \frac{P_0(E_i^cH_iB)}{P_0(H_iB)} \quad \text{if } i > k \end{aligned}$$

It is immediate now to extend P_0 on $\mathcal{A}_{\mathcal{E}}$ by imposing $P_0(A_r) = 0$ for all $A_r \in \mathcal{A}_{\mathcal{E}} \setminus \mathcal{A}_{\mathcal{G}}^{|B}$, and so, by additivity, $P_0(H_j) = 0$ for all $E_j|H_j \in \mathcal{R}$.

From this we can the state the following.

Definition 2 *Let $\mathbf{P}_{\mathcal{E}}$ be a "mixed" conditional probability and $\mathcal{G} \subseteq \mathcal{E}$. If $\mathbf{P}_{\mathcal{G}}$ is* strong coherent *with respect to the event $B = \bigwedge_{i:\, E_i|H_i \in \mathcal{R}} H_i^c$, then $\mathbf{P}_{\mathcal{E}}$ is* locally strong coherent *on $\mathcal{G}$.*

Proposition 1 *Let $\mathbf{P}_{\mathcal{G}}$ be the restriction of $\mathbf{P}_{\mathcal{E}}$ to $\mathcal{G} \subseteq \mathcal{E}$ and $\mathbf{P}_{\mathcal{R}}$ the restriction of $\mathbf{P}_{\mathcal{E}}$ to $\mathcal{R} = \mathcal{E} \setminus \mathcal{G}$. If $\mathbf{P}_{\mathcal{E}}$ is locally strong coherent on $\mathcal{G}$ then*

$$\mathbf{P}_{\mathcal{E}} \text{ is coherent} \Leftrightarrow \mathbf{P}_{\mathcal{R}} \text{ is coherent.}$$

Note that for extension purposes on $E|H$ (i.e. when also (6) must holds) it is enough to choose an event B (and so $\mathcal{G}$) such that it does not influence (i.e. shares atoms with) H. By Proposition 1, whenever locally strong coherence is detected on such $\mathcal{G}$, neither the coherence of $\mathbf{P}_{\mathcal{E}}$ nor the extension of the assessment to $E|H$ are actually influenced by the subfamily, so that $\mathcal{G}$ can be "neglected" focusing attention only on the rest $\mathcal{R} = \mathcal{E} \backslash \mathcal{G}$. Obviously the property can be iterated reducing, as much as possible, the initial domain and so the number of atoms to compute (that can be 0).

Once again, by Theorem 1 it would be possible to operationally check the locally strong coherence by solving a linear system. Anyhow this will not help the complexity reduction because we will continue to have an exponential number of unknowns. On the other hand, we will see that whenever the cardinality of $\mathcal{G}$ is small and some configuration among its components is allowed, the use of this linear system can be skipped, with all the benefits this brings.

The computational simplification depends on the capability to detect under which conditions the constraints associated to $\mathcal{G}$ in the $L^i_{\alpha_i}$'s are surely solved "properly" (i.e. not as $0 = 0$). This is possible singling out configurations that guarantee (i.e. imply) the existence of a non trivial solution only for events in $\mathcal{G}$ and null for all the others. Therefore, referring to the linear systems , it is like to solve the systems $L^i_{\alpha_i}$'s in (6) with less unknowns. Without entering into details, we can say that the cited particular configurations among conditional events in $\mathcal{G}$ are expressed by the compatibility of some specific combination among the events (logical sum of atoms in $\mathcal{A}_{\mathcal{E}}$). In such cases it is useless to actually build and solve the systems because a simple solution is always obtainable, and so it is enough to test the satisfiability of that configurations.

Satisfiability will depend on the particular logical relations $\mathcal{C}$ among the unconditional parts $\mathcal{U}_{\mathcal{E}}$, so that $\mathcal{C}$ has also in this case a relevant role on the assessment (in the previous approach it determines the set of atoms $\mathcal{A}_{\mathcal{E}}$). Note that when $\mathcal{C}$ is empty (i.e. when all the events are logically independent), any configuration is trivially satisfied. In this situation it is possible to choose as $\mathcal{G}$ one conditional event $E_i|H_i$ (or a pair $\{E_i|H_i, E_i^c|H_i\}$) and iterate this procedure. At the end all the elements of $\mathcal{E}$ can be eliminated, obtaining that $\mathbf{P}_{\mathcal{E}}$ is surely coherent (if, obviously, all its values are in $[0, 1]$).

For precise assessments the reductions are "permanent" (since there is actually only one sequence $\widetilde{\mathcal{L}}$), while for conditional lower probabilities we can adjoin configurations that "influence" only some branches of the procedure (i.e. only some sequence $\widetilde{\mathcal{L}}_j$ or parts of them). In the last situation, applicability of locally strong coherence of $\mathbf{P}_{\mathcal{E}}$ on $\mathcal{G}$ must also mention on which sequences it will be.

From the previous consideration, we can state that it is absolutely equivalent to solve linear systems like (5) (used just to check coherence), or (5) + (6) (used to check and extend), and to verify the satisfiability of the logical conditions.

Hence, for each conditional event $E_j|H_j$ a sequence $\widetilde{\mathcal{L}}_j$ of *sets of constraints* can be considered and, for the equivalence stated before, they can actually be fulfilled either by a set of logical configurations or by solution of linear systems. To each set of constraints, identified by an index α_j, a subset of conditional events $\mathcal{R}^j_{\alpha_j} \subseteq \mathcal{E}$ is associated.

Fixing $E_j|H_j$, each set of constraints is used to search for a precise probability P such that

$$\begin{array}{rcll} P(E_jH_j) & = & \underline{p}_j\ P(H_j) & \text{if } E_j|H_j \in \mathcal{R}^j_{\alpha_j} \\ P(E_iH_i) & = & p_i\ P(H_i) & \forall E_i|H_i \in \mathcal{R}^j_{\alpha_j} \\ & & & \text{with } i \in \{1, \ldots, k\} \quad (12) \\ P(E_lH_l) & \geq & \underline{p}_l\ P(H_l) & \forall E_l|H_l \in \mathcal{R}^j_{\alpha_j} \\ & & & \text{with } l \in \{k+1, \ldots, n\} \\ P(E_l^cH_l) & \geq & (1-\overline{p}_l)\ P(H_l) & \text{and } l \neq j \end{array}$$

Obviously, changing the reference event to $E_{j'}|H_{j'}$, the first strict equality constraint in (12) will be shifted to the index j'.

In particular, we denote by $\widetilde{L}^i_{\alpha_i}$ those configurations or systems to reach the "deepest" layer for H (i.e. when the explicit or implicit constraint (6) is added) and by $L^i_{\alpha_i}$ those to check the coherence on the events $\mathcal{R}^i_{\alpha_i}$, with $\alpha_i = 0, \ldots, \omega_i$.

As already noted, for each sequence, there will be a layer $\overline{\alpha}_i$ where the extreme values for $P(E|H)$ are computed. Such computation is trivial when the layer $\overline{\alpha}_i$ is reached just by logical conditions (in this situation $\mathcal{R}^i_{\overline{\alpha}_i} = \emptyset$ and hence extreme values are 0 and 1), otherwise it requires a linear programming optimization. So in general, for each sequence $\widetilde{\mathcal{L}}_i$ there are $\tilde{\gamma}_i$ sets of constraints $\widetilde{L}^i_0, \ldots, \widetilde{L}^i_{\tilde{\gamma}_i}$ fulfilled by logical configurations, followed by sets $\widetilde{L}^i_{\tilde{\gamma}_i+1}, \ldots, \widetilde{L}^i_{\overline{\alpha}_i-1}, L^i_{\overline{\alpha}_i}$ of linear systems. If, eventually, $\mathcal{R}^i_{\overline{\alpha}_i+1} \neq \emptyset$ other logical configurations $L^i_{\overline{\alpha}_i+1}, \ldots, L^i_{\gamma_i}$ and other linear systems $L^i_{\gamma_i+1}, \ldots, L^i_{\omega_i}$ are needed just to check coherence. Note that $0 \leq \tilde{\gamma}_i \leq \overline{\alpha}_i$, $\overline{\alpha}_i + 1 \leq \gamma_i \leq \omega_i$ where lower bounds represent computational worst cases, while upper bounds are the best situations (no linear system is needed).

Our procedure build and solve the first sequence $\widetilde{\mathcal{L}}_1$.

At the same time, when each single constraint is not trivially satisfied (i.e. it does not reduce to $0 = 0$), the procedure records if it is fulfilled as an equality or as a strict inequality. This is operationally done introducing variables ν_j^1 and $\nu_{j^c}^1$ for $E_j|H_j$ and $E_j^c|H_j$, respectively. In the sequel we will denote by j^ϵ indexes that can be both j and j^c.

To each $\nu_{j^\epsilon}^1$ will be assigned the string "=" if the particular solution in $\widetilde{\mathcal{L}}_1$ fulfills the constraint in (12) associated to $E_l^\epsilon|H_l$ as an equality, the string ">" otherwise. This will be useful in the other sequences $\widetilde{\mathcal{L}}_{j^\epsilon}$, $j = k+1, \ldots, n$, because the $\nu_{j^\epsilon}^1$'s show if solutions of the first sequence can be compatible with solutions of other sequences. In particular, if the constraints in $\widetilde{L}_{\alpha_{j^\epsilon}}^{j^\epsilon}$ or in $L_{\alpha_{j^\epsilon}}^{j^\epsilon}$ agree with $\nu_{l^\epsilon}^1$ for all $E_l^\epsilon|H_l \in \mathcal{R}_{\alpha_{j^\epsilon}}^{j^\epsilon}$, the constraints can be not tested and $\mathcal{R}_{\alpha_{j^\epsilon}}^{j}$ can be neglected.

The sequence $\widetilde{\mathcal{L}}_1$ is taken as a reference point because, in it, solutions for all the layers are actually detected and they could be taken as potential solutions of some other $\widetilde{\mathcal{L}}_{l^\epsilon}$, $l \in \{k+1, \ldots, n\}$.

To better describe how the procedure works, we introduce the further notation: let I_{α_1} be the set of indexes of the events in $\mathcal{R}_{\alpha_1}^1$. In this way we can associate to each conditional event $E_j^\epsilon|H_j$ the set of indexes $I_{\beta_{j^\epsilon}}$ with

$$\beta_{j^\epsilon} = \max_{\alpha_1 \in \{0,1,\ldots,\omega_1\}} \{\alpha_1 : j^\epsilon \in I_{\alpha_1}\}.$$

So β_{j^ϵ} represents the deepest layer reached by $E_j^\epsilon|H_j$ in the first sequence.

Relating to the previous notation, when $\beta_{j^\epsilon} < \overline{\alpha}_1$ the conditional event $E_j^\epsilon|H_j$ reaches a layer less deep than that of $E|H$, so the probability of $E_j^\epsilon|H_j$ is not a constraint for the inference step. This implies that the constraint associated to $E_j^\epsilon|H_j$ appears only in $\widetilde{L}_0^1, \ldots, \widetilde{L}_{\beta_{j^\epsilon}}^1$. On the other hand, if $\beta_{j^\epsilon} \equiv \overline{\alpha}_1$ then $E_j^\epsilon|H_j$ contributes to the computation of the extreme values for $P(E|H)$, but it does not influence the coherence on the residual events $\mathcal{R}_{\overline{\alpha}_1+1}^1$. Eventually, when $\beta_{j^\epsilon} > \overline{\alpha}_1$, the event $E_j^\epsilon|H_j$ is relevant for both inference on $E|H$ and check of coherence on $\mathcal{R}_{\overline{\alpha}_1+1}^1$.

We can now simplify the other sequences $\widetilde{\mathcal{L}}_{j^\epsilon}$, $j = k+1, \ldots, n$. In fact, when $\nu_{j^\epsilon}^1$ turns out to be "=", for any value of β_{j^ϵ} the sequence $\widetilde{\mathcal{L}}_{j^\epsilon}$ can be reduced just to compute the extreme values for $P(E|H)$ under the constraints $L_{\overline{\alpha}_{j^\epsilon}}^{j^\epsilon}$ (note that $I_{\overline{\alpha}_{j^\epsilon}} = I_{\overline{\alpha}_1}$ but $L_{\overline{\alpha}_{j^\epsilon}}^{j^\epsilon} \neq L_{\overline{\alpha}_1}^1$ because constraints must respect the rules in (12)). On the other hand, when $\nu_{j^\epsilon}^1$ is equal to ">", if $\beta_{j^\epsilon} < \overline{\alpha}_1$ then $\widetilde{\mathcal{L}}_{j^\epsilon} \equiv \{\widetilde{L}_{\beta_{j^\epsilon}}^{j^\epsilon}, \ldots, L_{\omega_{j^\epsilon}}^{j^\epsilon}\}$, while if $\beta_{j^\epsilon} \geq \overline{\alpha}_1$ then $\widetilde{\mathcal{L}}_{j^\epsilon} \equiv \{L_{\overline{\alpha}_{j^\epsilon}}^{j^\epsilon}, \ldots, L_{\omega_{j^\epsilon}}^{j^\epsilon}\}$.

Of course, whenever $\beta_{j^\epsilon} = 0$ and $\nu_{j^\epsilon}^1 =$ " $>$ " we will have a fictitious reduction, while when $\beta_{j^\epsilon} = \omega_{j^\epsilon}$ (and $\nu_{j^\epsilon}^1 =$ " $>$ ") only the last set of constraints must be built and solve.

4 Further elimination conditions

As already mentioned, it is possible to detect particular configurations among events of $\mathcal{G} \subseteq \mathcal{E}$ that guarantee the solvability of specific linear systems.

In [3, 4, 5], it has been done proposing explicit configurations for precise and lower assessments. The conditions are classified according to the cardinality of $\mathcal{G}$ and they are detected testing the satisfiability in $\mathcal{G}$ (endowed with $\mathbf{P}_\mathcal{G}$ and $\mathcal{C}$) of particular logical (and numerical) properties. We do not report them again because we prefer to focus attention on further conditions suitable for "twin" pairs of events $\big(E_j|H_j, E_j^c|H_j\big)$, $j \in \{k+1, \ldots, n\}$. In fact, having the same conditioning event H_j they must be eliminated simultaneously (i.e. in the same layer). Such peculiarity leads us to refine the conditions in [4] that works for more general subfamilies. This specification is not a merely formal refinement, but it is useful from a computational point of view, because it shrinks the search-space to set of twin pairs avoiding the procedure to make an exhaustive search. In this way there is a sensible reduction in time-complexity.

We introduce now the explicit conditions for single or double pairs, denoted with C_i and C_{ij} respectively. For the stated relevance of the first sequence $\widetilde{\mathcal{L}}_1$, we also give the corresponding elimination signs expressed by the vectors $\boldsymbol{\nu}_{C_i}$ and $\boldsymbol{\nu}_{C_{ij}}$.

To eliminate a pair $C_i = (E_i|H_i, E_i^c|H_i)$ the following rules can be applied:

i) we can differentiate two cases by the numerical value of $\underline{p}_i$

   i1) if $\underline{p}_i = 0$ and $E_i^c H_i \bigwedge_{l \neq i} H_l^c \neq \emptyset$ then $\mathbf{P}_\mathcal{E}$ is locally strong coherent on C_i for all potential sequences $\widetilde{\mathcal{L}}_k$ and, in particular, if $\overline{p}_i < 1$ then $\boldsymbol{\nu}_{C_i}^1 \equiv (=, >)$, otherwise $\boldsymbol{\nu}_{C_i}^1 \equiv (=, =)$;

   i2) whenever $\underline{p}_i > 0$ and both $E_i H_i \bigwedge_{l \neq i} H_l^c$ and $E_i^c H_i \bigwedge_{l \neq i} H_l^c$ are possible then $\mathbf{P}_\mathcal{E}$ is locally strong coherent on C_i for all potential sequences $\widetilde{\mathcal{L}}_k$ and, in particular, $\boldsymbol{\nu}_{C_i}^1 \equiv (=, =)$.

For two pairs $\mathcal{C}_{ij} = \big(E_i|H_i, E_i^c|H_i, E_j|H_j, E_j^c|H_j\big)$ we can detect more "sophisticated" configurations (in

these cases the sign string $\boldsymbol{\nu}^1_{C_{ij}}$ will have components $\left(\nu^1_i, \nu^1_{i^c}, \nu^1_j, \nu^1_{j^c}\right)$):

ii) when $\underline{p}_i = \underline{p}_j = 0$ and $E_i^c H_i E_j^c H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset$ then $\mathbf{P}_{\mathcal{E}}$ is locally strong coherent on C_{ij} for the potential sequences $\widetilde{\mathcal{L}}_k$ with $k \neq i^c, j^c$, so that we will surely have $\boldsymbol{\nu}^1_{C_{ij}} \equiv (=,>,=,>)$.
Anyhow some further specification can be done to obtain better signs (remember that whenever $\nu^1_k \equiv$ " $=$ " we have a simplification on the computation):

ii1) if also $E_i H_i E_j^c H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset$ and $E_i^c H_i E_j H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset$ hold, then $\boldsymbol{\nu}^1_{C_{ij}} \equiv (=,=,=,=)$;

ii2) if, until valid the logical condition in ii), only $E_i H_i E_j^c H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset$ holds, then $\boldsymbol{\nu}^1_{C_{ij}} \equiv (=,=,=,>)$;

ii3) otherwise, if $\overline{p}_i < \overline{p}_j$ and
$E_i H_i \wedge (E_j H_j \vee E_j^c H_j \vee H_j^c) \bigwedge_{l \neq i,j} H_l^c =$
$= E_i H_i \bigwedge_{l \neq i,j} H_l^c \neq \emptyset$,
then $\boldsymbol{\nu}^1_{C_{ij}} \equiv (=,=,=,>)$;

iii) if $\underline{p}_i$ or $\underline{p}_j$ (or both) is (are) different from 0, then we can distinguish four situations:

iii1) whenever $\overline{p}_i = \overline{p}_j = 1$ and
$E_i H_i E_j H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset$,
then $\mathbf{P}_{\mathcal{E}}$ is locally strong coherent on C_{ij} for the potential sequences $\widetilde{\mathcal{L}}_k$ with $k \neq i, j$, so that $\boldsymbol{\nu}^1_{C_{ij}} \equiv (>,=,>,=)$ if $i, j \neq 1$;

iii2) whenever $\max\{\underline{p}_i, \underline{p}_j\} \leq \min\{\overline{p}_i, \overline{p}_j\}$ and

$$\begin{cases} E_i H_i E_j H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \\ E_i^c H_i E_j^c H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \end{cases},$$

denoting with m and n the indexes s.t. $\underline{p}_m = \max\{\underline{p}_i, \underline{p}_j\}$ and $\overline{p}_n = \min\{\overline{p}_i, \overline{p}_j\}$, respectively, then $\mathbf{P}_{\mathcal{E}}$ is locally strong coherent on C_{ij} for the potential sequences $\widetilde{\mathcal{L}}_k$ with $k \neq m^c, n$ and so

$$\boldsymbol{\nu}^1_{C_{ij}} \equiv \begin{cases} (=,=,>,>) & \text{if } m = n = i \text{ and } j \neq 1 \\ (>,>,=,=) & \text{if } m = n = j \text{ and } i \neq 1 \\ (=,>,>,=) & \text{if } m = i, n = j \text{ and } j \neq 1 \\ (>,=,=,>) & \text{if } m = j, n = i \text{ and } i \neq 1 \end{cases}$$

iii3) if

$$\begin{cases} E_i H_i E_j H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \\ E_i^c H_i \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \\ E_j^c H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \end{cases}$$

then $\mathbf{P}_{\mathcal{E}}$ is locally strong coherent on C_{ij} for all the potential sequences $\widetilde{\mathcal{L}}_k$ and so $\boldsymbol{\nu}^1_{C_{ij}} \equiv (=,=,=,=)$;

iii4) if

$$\begin{cases} E_i^c H_i E_j^c H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \\ E_i H_i \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \\ E_j H_j \bigwedge_{l \neq i,j} H_l^c \neq \emptyset \end{cases}$$

then $\mathbf{P}_{\mathcal{E}}$ is locally strong coherent on C_{ij} for all the potential sequences $\widetilde{\mathcal{L}}_k$ and so $\boldsymbol{\nu}^1_{C_{ij}} \equiv (=,=,=,=)$.

The previous conditions can be combined with those for precise assessments, obtaining conditions to eliminate "mixed" subfamilies (i.e. composed by events with precise probabilities and twin pairs). In particular, when precise assessments associated to some $E_{i_0}|H_{i_0}$ are 0 and those to $E_{j_1}|H_{j_1}$ are 1, it suffices to modify all the logical expressions in i)–iii) replacing $H_{i_0}^c$ by $E_{i_0}^c H_{i_0}$ and $H_{j_1}^c$ by $E_{j_1} H_{j_1}$, respectively.

The entire "machinery" allows to perform inference obtaining "exact" results (i.e. the interval values $[\underline{p}, \overline{p}]$ are the same we would obtain by directly solving all the linear systems $\widetilde{L}^i_{\alpha_i}$) because the logical conditions do not require any heuristic. Actually, avoiding to build and solve huge linear systems, heavy round numerical errors are also skipped.

On the other hand, our procedure could be improved combining it with other techniques based on the linear structure of the random gain in the betting criterion, similarly to those presented for imprecise probabilities in [1]. Other benefits would be brought applying numerical techniques (like the column generation [16]) when we cannot avoid to use linear programming.

5 A practical application to a diagnostic problem

We present how our procedure works in a medical diagnosis. We consider a problem about thyroid pathology where the different hypotheses suggested by the physician are:

- morphological pathology D_M;
- nodular pathology (comprehending both benign (adenomas) and malignant (carcinomas)) D_N;
- functional pathology (hyper and hypothyroidism) D_F;
- compression or deviation of surrounding structures (specially of the trachea) D_H;
- absence of thyroid pathology D_A.

These hypotheses are endowed with the following logical constraints

$$\mathcal{C} = \{D_H \subset D_M \vee D_N;\ D_A \subset D_M^c \wedge D_N^c \wedge D_F^c\}.$$

Note that the five hypotheses do not cover all the possible thyroid pathology, but just those relevant to a physician when he has to perform a diagnosis on a patient with a set of symptoms like "arrhythmia, palpitations, nervousness, anxiety, weight loss and dysphony". First of all, the physician looks at the possible ultrasound results: *normal glandular volume* E_n*, high glandular volume* E_a *or tracheal compression and/or deviation* E_c. Obviously these results form a partition, while the logical relations among pathology and ultrasound results are

$$E_a \subset D_M \quad E_c \subset D_H \quad D_A \subset E_n.$$

From the physician's data-set[2] the following values are assessed

$$\begin{array}{ll}
P(E_n|D_M) = 0 & \\
P(E_c|D_M D_N D_H D_F) = 1 & \\
\underline{P}(E_a|D_M) = \frac{13}{14} & \overline{P}(E_a|D_M) = 1 \\
\underline{P}(E_n|D_N) = \frac{2}{11} & \overline{P}(E_n|D_N) = 1 \\
\underline{P}(E_n|D_M D_N) = \frac{1}{27} & \overline{P}(E_n|D_M D_N) = \frac{218}{243} \\
\underline{P}(E_a|D_M D_N D_H) = 0 & \overline{P}(E_a|D_M D_N D_H) = \frac{2}{5} \\
\underline{P}(E_c|D_M D_N D_H) = \frac{3}{5} & \overline{P}(E_c|D_M D_N D_H) = 1 \\
\underline{P}(E_a|D_M D_N D_F) = \frac{5}{38} & \overline{P}(E_a|D_M D_N D_F) = 1.
\end{array}$$

[2]The values derive from 377 clinical cases observed by Italian Hospital "Unità Operativa di Endocrinologia del Policlinico dell'Università di Chieti" during years '97-'98.

It can be proved (applying the procedure) that "prior" evaluations lying in $[0, 1]$ on all pathology are coherent with this assessment. The physician is looking for the probability that a patient is suffering of both morphological and nodule pathology supposing he has an high glandular volume (i.e. to perform inference on $D_M D_N|E_a$).

Applying the classical procedure, based on solving sequences of linear optimization programs, we would need a lot of computations: 16 sequences, 4 from the two precise assessments transformed in lower and 12 from the other interval evaluations. Each sequence would start with a system of 16 unknowns and 16 inequalities.

On the other hand, our procedure eliminates in the first sequence the twin evaluation on $E_n|D_N$ by condition $i1$) applied to the inverted pair $(E_n^c|D_N, E_n|D_N)$. At the second step, it is not anymore possible to impose $P(E_a) = 0$, so the procedure finds the minimum and the maximum values for $D_M D_N|E_a$, both being 0 in this sequence. Moreover, the upper bound cannot be increased in any other sequence. Computing these bounds the procedure find a solution that gives positive probability to the conditioning event D_M, so that $\{E_n|D_M,\ E_a|D_M,\ E_a^c|D_M\}$ do not pass to the next layer. Then, only for the purpose of checking coherence, the following elimination sequence is detected:

- $\{E_n|D_M D_N,\ E_n^c|D_M D_N, \}$ by condition i2);
- $\{E_a|D_M D_N D_H,\ E_a^c|D_M D_N D_H,\ E_c|D_M D_N D_H,\ E_c^c|D_M D_N D_H\}$ by condition ii4);
- $\{E_a^c|D_M D_N D_F,\ \ E_a|D_M D_N D_F\}$ by condition i1);
- $\{E_c|D_M D_N D_H D_F\}$ by elimination condition for precise values.

The same sequence is useful for all the other sequences and hence we use linear programming techniques only for the computations of the bounds.

Since the coherent value for $P(D_M D_N|E_a)$ is 0, the physician will disbelieve on the combined disease $D_M D_N$ for a patient with, apart from the initial symptoms, an augmented thyroid volume.

Anyhow, the physician has information about further cases not well reported (i.e. with some missing data). He could enlarge his data-set with this cases obtaining a change of the following likelihood values:

$$P(E_n|D_M) = 0.02 \quad \text{and} \quad \overline{P}(E_a|D_M) = 0.98$$

(while all the other remain as before). Performing the same inference on $D_M D_N|E_a$ the "elimination se-

quence" remains exactly the same, while the new coherent interval will be $[0, 0.56]$ (attained again on the first sequence). This time a mild confidence on the combined disease will support future decisions.

References

[1] V. Biazzo, A. Gilio and G. Sanfilippo. Computational Aspects in Checking of Coherence and Propagation of Conditional Probability Bounds. *Proc. of The* 5th *Workshoop on Uncertainty Processing WUPES'2000*, Czech Republic, 1-13, 2000.

[2] V. Biazzo, A. Ferro, A. Gilio and R. Giugno. A General Probabilistic Database Model. *Proc. of IPMU2000, Madrid, Spain*, 2:1107–1114, 2000.

[3] A. Capotorti and B. Vantaggi. An Algorithm for Coherent Conditional Probability Assessment. In *SIMAI'98*, Giardini Naxos, Italy, 2:144-148, 1998.

[4] A. Capotorti, L. Galli and B. Vantaggi. How to Use Locally Strong Coherence in an Inferential Process Based on Upper-Lower Probabilities. *Proc.* 5th *Workshoop on Uncertainty Processing WUPES'2000*, Czech Republic, 14–28, 2000 (Extended version submitted to *Soft Computing*).

[5] A. Capotorti and B. Vantaggi. Locally Strong Coherence in an Inference Process. To appear in *Annals of Mathematics and Artificial Intelligence*.

[6] G. Coletti. Coherence Principles for Handling Qualitative and Quantitative Partial Probabilistic Assessments. *Mathware & Soft Computing*, 3, 159–172, 1996.

[7] G. Coletti and R. Scozzafava. Characterization of Coherent Conditional Probabilities as a Tool for their Assessment and Extension. *Int. Journ. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 4(2): 103-127, 1996.

[8] G. Coletti and R. Scozzafava. Exploiting Zero Probabilities. *Proc. of EUFIT'97*, Aachen, Germany, ELITE Foundation, 1499–1503, 1997.

[9] G. Coletti and R. Scozzafava. Conditional Measures: Old and New. In *New trends in Fuzzy Systems*, World Scientific: 107-120, 1998.

[10] G. Coletti and R. Scozzafava. Conditioning and Inference in Intelligent Systems. *Soft Computing*, 3: 118-130, 1999.

[11] G. Coletti and R. Scozzafava. Coherent Upper and Lower Bayesian Updating. *Proocedings of the First International Symposium on Imprecise Probabilities and Their Applications*, de Cooman, G., Cozman, F.G., Moral, S., Walley, P. (eds.), Imprecise Probabilities Project, Universiteit Gent, Belgium, 101-110, 1999.

[12] G. Coletti and R. Scozzafava. The Role of Coherence in Eliciting and Handling "Imprecise" Probabilities and its Application to Medical Diagnosis. *Information Science*, 130: 41-65, 2000.

[13] B. de Finetti. Sull'Impostazione Assiomatica del Calcolo delle Probabilitá. *Annali Univ. di Trieste* 19: 29-81, 1949. (English translation in: (1972) Probability, Induction, Statistics. Ch. 5. London Wiley) .

[14] B. de Finetti. *Theory of Probability*, vol. 1-2, Wiley, New York, 1974.

[15] ISIPTA'99. *Proocedings of the First International Symposium on Imprecise Probabilities and Their Applications*, de Cooman, G., Cozman, F.G., Moral, S., Walley, P. (eds.), Imprecise Probabilities Project, Universiteit Gent, Belgium, 1999.

[16] B. Jaumard, P. Hansen and M. Poggi de Aragão Column Generation Methods for Probabilistic Logic. *ORSA Journal on Computing*, 3: 135-148, 1991.

[17] F. Lad. *Operational Subjective Statistical Methods: a Mathematical, Philosophical, and Historical Introduction*. Wiley, New York, 1996.