Coherence Principles for Handling Qualitative and Quantitative Partial Probabilistic Assessments

Giulianella Coletti
Dip. Matematica, Universit di Perugia
Via Vanvitelli 1, 06100 Perugia, Italy.
e-mail: giulicol@ipguniv.unipg.it

Abstract

In this paper we present an overview of mathematical models for handling partial entailments and their extensions in a probabilistic frame.

1 Introduction

Many criticisms concerning the adoption of the probability paradigm for representing the partial and uncertain knowledge in Artificial Intelligence are related to the following "putative obligations": giving, a priori, both complete description of the domain, containing all possible envisaged situations and moreover provided of a precise algebraic structure, and the overall assessment of probability.

However the probability model that gave rise to the above criticism has the merit of being the most known. Actually, it deals with a very particular case, hence it has the advantage of giving the possibility to elaborate a very elegant mathematical theory, but, when the probability must model actual uncertain situations, there is the need for a more general theory, which permits a gradual and coherent probability assignment, starting from a set of events containing only those apt to represent the initial cognitive domain. The goal of this general approach to probability theory is based on some very simple ideas: regard the events as propositions or pairs of propositions; consider probability as a linear operator rather than as a measure; regard the boolean algebra (or product of boolean algebras) spanned by the given events only as a provisional tool (possibly changing, when new events and new information are considered) apt to test the consistence of the partial assessment; forgo the myth of the unicity and accept that several probability distributions (infinitely many, in general) on the atoms are compatible with the given assessment.

The aim of this paper is to present an overview of mathematical models, based on the above principles, in order to rule partial (quantitative and qualitative) probabilistic entailments and their extensions. Particular attention is paid to focus on the strict connections between these pragmatic models and de Finetti's theory, more known for his semantic valence (discussed in [28]). In fact, what is put here in evidence is usually taken in second place (often relegated to the proofs of theorems), also because such conditions appear not as elegant and concise as the other ones. Nevertheless this approach supplies us with actual tools which are suitable to manage partial probabilistic information and to be used in automatic inferential processes. The problems related to the elaboration of optimal algorithms for a profitable use of some of these models in knowledge based systems have been recently studied by many authors (see, for instance, [6], [7], [11], [16], [17] and [19]).

2 Coherent probability assessments on unconditional events

In the following, for simplicity of exposition, we consider only the case of finite sets of events. We first illustrate in detail the concept of coherence for a numerical probability assessment.

As it is well known, given a non-empty set Ω and a Boolean algebra E of subsets of Ω, a (finitely additive) probability on (Ω, E) is a real-valued set function P satisfying

(A1) $P(\Omega) = 1$;

(A2) $P(A \cup B) = P(A) + P(B)$ for any disjoint $A, B \in \mathcal{E}$;

(A3) $P(E)$ is non-negative for any $E \in \mathcal{E}$.

But, if $\mathcal{E}$ is just any collection of subsets of Ω, representing events (that can be interpreted, for example, as propositions), then (A1-A3) are insufficient to characterize P as a probability on (Ω, E): for example, if $\mathcal{E}$ contains no union of disjoint sets, (A2) is vacuously satisfied and hence a "natural" property as the monotonicity of P is not ensured. More generally it may not exist an extension of P to any algebra containing $\mathcal{E}$ and satisfying (A1-A3), and then P can not be considered a partial definition of a function characterizable as a probability. Consider for instance the following example:

Example 1. Let $\{E_1, E_2, E_3\}$ be a partition of Ω: consider the family $\mathcal{E} = \{E_1 \cup E_2, E_1 \cup E_3, E_2 \cup E_3, \Omega\}$ and the assignment

$$P(E_1 \cup E_2) = 4/9,\ P(E_1 \cup E_3) = 5/9,\ P(E_2 \cup E_3) = 2/3,\ P(\Omega) = 1.$$

This assessment satisfies (A1-A3) on $\mathcal{E}$, but, since $P(E_1 \cup E_2) + P(E_1 \cup E_3) + P(E_2 \cup E_3) = 2[P(E_1) + P(E_2) + P(E_3)] = 15/9$, (A2) is not satisfied on the set $\mathcal{E} \cup \{E_1, E_2, E_3\}$, which is the minimal algebra containing $\mathcal{E}$.

Therefore it is necessary to find a rule able to define when an assessment is a probability, without using the facilities supplied by the algebraic structure of the set of events in a determining way; in other words, to focus on the specific properties of the probability function. Such a rule, which we will call "coherence" is based on the concepts of decomposability and linearity.

More precisely, if we consider an assessment $\mathcal{P} = (p_1, ..., p_n)$ on an arbitrary finite family $\mathcal{F} = \{E_1, ..., E_n\}$ (i.e. $P(E_i) = p_i$, $i = 1, 2, ...n$), and denote by $A_1, A_2, ..., A_n$ the atoms generated by these events, then, referring to the unknowns $x_r = P(A_r)$, coherence of $\mathcal{P}$ amounts to the compatibility of the following system

$$\begin{cases} \sum_{\substack{r \\ A_r \subseteq E_i}} x_r = p_i, & i = 1, \ldots, n, \\ \sum_{r=1}^{m} x_r = 1, x_r \geq 0, & r = 1, \ldots, m. \end{cases} \qquad (\mathcal{S}_1)$$

So coherence is equivalent to the existence of an extension of P from the given events E_i $(i = 1, 2, ..., n)$ to the atoms generated by them.

The connection between the above definition of coherence and that - well know - given by de Finetti in terms of coherent bets or equivalently of minimum penalty is based on classic alternative theorem. In fact, by denoting with e_i the indicator function of E_i, a real-valued function P defined on $\mathcal{E}$ is coherent (following the betting scheme of de Finetti) if, for any $E_1, ..., E_n \in \mathcal{F}$ and for any real numbers $\lambda_1, ..., \lambda_n$, the possible values of the random gain

$$G = \sum_{i=1}^{n} \lambda_i [e_i - P(E_i)],$$

are neither all positive nor all negative (in other words, if the values of $P(E)$ are interpreted as betting rates in the usual way, it must be impossible to bet at these rates in such a way as to be a sure loser or winner) and since de Finetti's coherent condition is equivalent to require that

$$\sup_{A_r} G \geq 0$$

for every choice of $\lambda_1, ..., \lambda_n$.

But, if I_{E_i} indicates the m-vector $(c_1, ..., c_n)$ with $c_k = 1$, if the atom $A_k \subset E_i$ and $c_k = 0$ otherwise, the above condition can be expressed by requiring that the following system, with unknowns λ_i

$$\begin{cases} \sum_{i=1}^{n} (\lambda_i I_{E_i} - p_i I_\Omega) > 0, \\ \lambda_i \in R, \; i = 1, \ldots, n. \end{cases} \qquad (\mathcal{S}_1')$$

has no solution. For a classical alternative theorem such a condition is necessary and sufficient for the solvability of (S_1).

Moreover, given a further event E_{n+1} and the corresponding extended family $\mathcal{K} = \mathcal{F} \cup \{E_{n+1}\}$, consider the case in which E_{n+1} is union of some atoms, i.e. E_{n+1} is *logically dependent* on the events of $\mathcal{F}$: then, putting $P(E_{n+1}) = p_{n+1}$, one has

$$p_{n+1} = \sum_{A_r \subseteq E_{n+1}}{}_r \; x_r.$$

Letting the vector $(x_1, x_2, ..., x_m)$ assume each value in the set X of solutions of system (S_1), i.e. of all possible extensions of $\mathcal{P}$ to the atoms, the probability p_{n+1} describes an interval $[p', p"] \subseteq [0,1]$, with $p' = \inf_X p_{n+1}, \quad p" = \sup_X p_{n+1}$.

If E_{n+1} is not logically dependent on the events of $\mathcal{F}$, there will exist two events E_* and E^* (possibly $E_* = \emptyset$ and $E^* = \Omega$) that are, respectively, the "maximum" and the "minimum" union of atoms such that

$$E_* \subseteq E_{n+1} \subseteq E^*.$$

So, given the probabilities x_r of the atoms, coherent assessments of $P(E_{n+1})$ are all those of the closed interval $[P(E_*), P(E^*)]$, i.e.

$$\sum_{A_r \subseteq E_*}{}_r \; x_r \leq p_{n+1} \leq \sum_{A_r \subseteq E^*}{}_r \; x_r.$$

Letting again the vector $(x_1, x_2, ..., x_m)$ assume each value in the set X, the probability p_{n+1} describes an interval $[p', p"] \subseteq [0,1]$, with

$$p' = \inf_X P(E_*) \quad , \quad p" = \sup_X P(E^*).$$

This result is dubbed as the fundamental theorem of probabilities of de Finetti ([14], p.78; [15], p.112: both pages refer to the English translations). The values $p', p"$ can be determined, for instance, by the simplex method of linear programming, as done in [3] ; see also [26]. Similar ideas appeared already in the classical work of Boole [2], that attracted little attention until it was revived in [19]. The same technique has been used, with a different aim, in [1]. An extensive discussion of the fundamental theorem, with several computational and geometrical examples, is in [23], [24]. More recently many authors are interested of computability problems related to the extension of probability (see for instance [16] and [21]).

It seems interesting to note that the extension theorem can be profitable used to put right a not coherent probability assessment. In fact, if for instance n-1 (but not n) evaluations are coherent, it is possible to choose which of evaluations changing, and in which way, to make coherent the total assessment. For example one can choose that implying the minimum alteration.

The above theory is extendible to different, less strict tools for the measurement of the degree of belief, however related with the probability frame. In fact it is possible that, at some stage of knowledge process, the probability assessment does not consist on a set of numbers, but on a set of intervals (imprecise probability), or on a set of ordinal relations (comparative or qualitative probability), or, finally on a mixage of numbers, intervals and relations. This is, for instance, the case that the field expert is not capable to summarize with numbers his degree of belief about the considered events or when different information must be assembled to act an inferential process.

We define coherent an imprecise probability assessment $Q = \{q(E_i) = [p'_i, p"_i], E_i \in \mathcal{F}\}$ defined on an arbitrary set of events $\mathcal{F}$, if there exists a coherent numerical probability assessment $\mathcal{P} = (p_1, ..., p_n)$ on $\mathcal{F}$, agreeing with $\mathcal{Q}$, in the sense that $p_i \in [p'_i p"_i]$; on the other words if there exists a solution of the following system, with unknowns $x_r = P(A_r)$ (where $A_1, ..., A_m$ are the atoms generated by events of $\mathcal{F}$)

$$\begin{cases} p'_i \leq \sum_{\substack{r \\ A_r \subseteq E_i}} x_r \leq p''_i, \quad i = 1, \dots, n, \\ \\ \sum_{r=1}^{m} x_r = 1, \, x_r \geq 0, \quad i = 1, \dots, m. \end{cases} \qquad (\mathcal{S}_2)$$

By using the same considerations made for numerical (precise) probabilities, we can give also in this case a condition equivalent to the solvability of the system (S_2), which may be read in terms of coherent bets as follows: function Q defined on $\mathcal{E}$, with values in the set of the closed intervals contained on $[0, 1]$ is coherent if, by indicate $Q(E_i) = [p'_i, p"_i]$, for any $E_1, ..., E_n \in \mathcal{F}$ and for any $\lambda_1, ..., \lambda_n \geq 0$ and $\gamma_1, \dots, \gamma_n \leq 0$, the possible values of the random gain

$$G = \sum_{i=1}^{n} \lambda_i [e_i - p'_i] + \sum_{i=1}^{n} \gamma_i [e_i - p''_i]$$

are all negative. In other words, if the values of p' and $p"$ are interpreted as betting rates of the gambler and of the bookie, respectively, to play on E_i, it must be impossible to bet in such a way as to be a sure loser.

For imprecise probability we can generalize in a natural way the results relating to the extendibility of a coherent assessment to a new event, and can essentially use the same techniques to determine the interval of coherent values with the previous ones.

Let us suppose now that we have an arbitrary set of events $\mathcal{F}$, and on $\mathcal{F}$ a comparative probability assessment, that is a (not necessarily complete) binary relation $\leq^*$ expressing the idea of "no more probable than". Also in this case we define the (ordinal) assessment $\leq^*$ to be coherent if there exists a coherent numerical probability assessment $\mathcal{P} = (p_1, ..., p_n)$ on $\mathcal{F}$, agreeing with $\leq^*$, in the sense that, if $<^*$ indicates the strict relation obtained when $E \leq^* F$ but not $F \leq^* E$, then

$$p_i \leq p_j \text{ if } E_i \leq^* E_j \quad \text{and} \quad p_i < p_j \text{ if } E_i <^* E_j.$$

That is $\leq^*$ is coherent if there exists a solution of the following system, with unknowns $x_r = P(A_r)$ (where $A_1, ..., A_m$ are the atoms generated by events of $\mathcal{F}$)

$$\begin{cases} \sum_{A_r \subseteq E_i} {}_r \; x_r \leq^* \sum_{A_r \subseteq E_i} {}_r \; x_r & E_i \leq^* E'_i \\ \\ \sum_{A_r \subseteq E_j} {}_r \; x_r <^* \sum_{A_r \subseteq E_j} {}_r \; x_r & E_j <^* E'_j \\ \\ \sum_{r=1}^{m} x_r = 1, x_r \geq 0, & r = 1, \ldots, m. \end{cases} \qquad (\mathcal{S}_3)$$

Also in this case it is possible (see [5]) to give an equivalent condition which can be significantly interpreted on the betting scheme. A comparative probability $\leq^*$ defined on $\mathcal{E}$, is coherent if, by for any $n \in N$, $E_1, ..., E_n$, $E'_1, ..., E'_n \in \mathcal{F}$, with $E_s \leq^* E'_s$, and for any $\lambda_1, \ldots, \lambda_n > 0$, the possible values of the random gain

$$G = \sum_{i=1}^{n} \lambda_i [e_i - e'_j]$$

can be all negative or null only in the case that $E_s =^* E'_s$, for every $s = 1, .., n$. The interpretation in terms of betting scheme is immediate, if one consider bets where even money is betted on E_i versus E'_i. About the extension of coherent comparative probabilities there are results analogous those related to the numerical ones (see [5]). In fact, given a further event E_{n+1} and the corresponding extended family $\mathcal{K} = \mathcal{F} \cup \{E_{n+1}\}$, coherence condition permits to detect the elements of $\mathcal{F}$ which must be less probable, those equivalent and finally those more probable than E_{n+1}, to have a coherent extension. For the other events, any choice of relation with E_{n+1} give a coherent proper extension of $\leq^*$ to $\mathcal{K}$.

The techniques to identify the above classes of events of $\mathcal{F}$ are the same used in the numerical case.

In the case that the probabilistic assessment contains both numerical and comparative evaluations, to test coherence consists on the proof of solvability of a system with equations and inequalities of the same kind of systems $(\mathcal{S}_1), (\mathcal{S}_2), (\mathcal{S}_3)$.

Remark 1. We note that when for representing the uncertainty we use a less rigid tool, such as a comparative probability, the incoherent assessments are obviously less frequent; on the contrary, the proof of coherence is computationally more complicate. This fact is true also when we consider not probabilistic tool, such as belief function or lower probability. Also for this functions there is the problem to test coherence if the assessment is given on a arbitrary set of events and also in this case the conditions of coherence are essentially related to solvability of some linear systems (for coherence condition see [27], for examples of incoherent assessments see [4]). But these system are certainly more complicate than those related to probability.

3 Coherent probability assessments on conditional events

Many interesting and unexpected features come to the fore when one tries to extend the above theory to conditional events and to the ensuing relevant concept of conditional probability, once a suitable extension of the concepts of coherence are introduced.

In this approach conditional probability is directly introduced as a function whose domain is a set of conditional events (regarded as three-valued logical entities). For more details see the paper [28], in this same issue.

In the usual approach instead the conditional probability $P(E|H)$ is introduced *by definition* as the ratio between the probabilities $P(E \cap H)$ and $P(H)$, assuming positive probability for the conditioning event, or, if this strong restriction is not required, conditional probability is defined in an axiomatic way, by considering the conditional events $E_i|H_i$ as pairs of unconditional events E_i, H_i, belonging to a cartesian product $\mathcal{E} \times \mathcal{H}$, $\mathcal{E}$ is Boolean algebra, $\mathcal{H}$ an additive class, with $\emptyset \notin \mathcal{H}$, and P a function from $\mathcal{E} \times \mathcal{H}$ to R, satisfying the following properties:

(B1) given any $H \in \mathcal{H}$ and $A_1, ..., A_n \in \mathcal{E}$ such that $A_i A_j \subseteq H^c$ $(i \neq j)$, the function $P(\cdot|H)$ defined on $\mathcal{E}$ satisfies

$$P\left(\left(\bigcup_{k=1}^{n} A_k\right)|H\right) = \sum_{k=1}^{n} P(A_k|H), \quad P(\Omega|H) = 1;$$

(B2) $P(H|H) = 1$ for any $H \in \mathcal{E} \cap \mathcal{H}$;

(B3) given E, H, A such that E, $H \in \mathcal{E}$, with A, $EA \in \mathcal{H}$, then

$$P(EH|A) = P(E|A)P(H|EA).$$

Notice that (B3) reduces, when $A = \Omega$, to the classical product rule for probability.

Obviously, as discussed for unconditional events, if $\mathcal{C}$ is just any collection of conditional events, then (B1-B3) are insufficient to characterize P as a conditional probability. Therefore also in this case is interesting to find a rule (which will be called "coherence") to define when an assessment on an arbitrary set of events is a probability. Then coherence of P on $\mathcal{C}$ entails that P is the restriction on $\mathcal{C}$ of a *conditional probability* defined on any $\mathcal{E} \times \mathcal{H} \supseteq \mathcal{C}$, where $\mathcal{E}$ is Boolean algebra and $\mathcal{H}$ an additive class, with $\emptyset \notin \mathcal{H}$.

For a set $\mathcal{C} = \{E_1|H_1, ..., E_n|H_n\}$, let us consider the set $\mathcal{A}$ of atoms generated by $\{E_1, ..., E_n, H_1, ..., H_n\}$. It should be clear that, while testing coherence of a given assessment $\{P(E_1H_1), ..., P(E_nH_n),\ P(H_1), ..., P(H_n)\}$ for the (unconditional) events $E_1H_1, ..., E_nH_n$, $H_1, ..., H_n$ is equivalent (as discussed in Sect.2) to test the solvability of the linear system (S_1). If we consider instead an assessment

$p_i = P(E_i|H_i)$ (for $i = 1, ..., n$) on the set $\mathcal{F}$ of conditional events $E_i|H_i$, then the existence of an extension of the probability distribution P on the set $\mathcal{A}$, satisfying the system with unknowns $P(A_r) \geq 0$, $r = 1, ..., m$,

$$\begin{cases} \sum\limits_{A_r \subseteq E_i H_i} P(A_r) = P(E_i|H_i) \sum_{r A_r \subseteq H_I} P(A_r), \quad i = 1, \dots, n \\ \\ \sum\limits_{A_r \subseteq H_0}^{r} P(A_r) = 1, \end{cases} \qquad (S_4)$$

with $H_0 = H_1 \cup ... \cup H_n$, is necessary, but not sufficient, to ensure coherence of P. In fact it does not even assure axioms B1-B3 for the considered events, as the following example shows.

Example 2. Consider the events $E|H$, $E^c|H$, $E|H^c$ and the assessment $P(E|H) = P(E^c|H) = 1/3$, $P(E|H^c) = 1$, which does not satisfy B1. Nevertheless the probability distribution $P(A_1) = P(A_2) = P(A_3) = 0$, $P(A_4) = 1$ on the atoms $A_1 = EH$, $A_2 = E^cH$, $A_3 = E^cH^c$, $A_4 = EH^c$, satisfies (S_4), which therefore is not sufficient to ensure coherence.

Notice that the first equation of (S_4) corresponds to the product rule

$$P(EH) = P(E|H)P(H)$$

for conditional probability, which is trivially satisfied by any value of $P(E|H)$ when $P(H) = 0$.

Condition (S_4) becomes sufficient, if also the following (not necessary) condition holds:

$$\sum_{A_r \subseteq H_i}^{r} P(A_r) > 0 \text{ for every conditioning event } H_i. \qquad (*)$$

So testing coherence of $P(E_i|H_i)$ would be equivalent to test the solvability of the linear system (S_4), with unknowns $P(A_r)$, under the condition (*). But, even if the latter may seem an "almost natural" condition on a finite set of events, it introduces some computational complications, due to the further strict inequalities involving the unknowns. But above all, ignoring the possible existence of null events drastically restricts the class of admissible probability assessments (see Example 3) and moreover, the possibility of extending in any case a coherent conditional probability is lost (see Example 4).

Example 3. Given three conditional events $E_1|H_1$, $E_2|H_2$, $E_3|H_3$ such that $\mathcal{A} = \{A_1, ..., A_5\}$, with $H_1 = A_1 \cup A_2 \cup A_3 \cup A_4$, $H_2 = A_1 \cup A_2$, $H_3 = A_3 \cup A_4$, $E_1H_1 = A_1$, $E_2H_2 = A_2$, $E_3H_3 = A_3$, consider the assessment

$$p_1 = P(E_1|H_1) = 3/4, \quad p_2 = P(E_2|H_2) = 1/4, \quad p_3 = P(E_3|H_3) = 1/2.$$

If we require positivity of the probability of conditioning events, we must adjoin to the system (S_4) also the conditions $P(A_1) + P(A_2) > 0$, $P(A_3) + P(A_4) > 0$, and this enlarged system has no solutions.

On the contrary the assessment is coherent. In fact, if we consider $\mathcal{E} \times \mathcal{E}^*$, where $\mathcal{E}$ is the algebra spanned by $\mathcal{A}$ and $\mathcal{E}^* = (\mathcal{E}\backslash\{\emptyset\})$, then the above assessment results a restriction of conditional probability P defined on $\mathcal{E} \times \mathcal{E}^*$, by extending for additivity the following values:

$$\begin{array}{l} P(A_1|A_1 \cup A_2) = P(A_1|A_1 \cup A_2 \cup A_3) = P(A_1|A_1 \cup A_2 \cup A_3 \cup A_4) = 3/4 \\ P(A_2|A_1 \cup A_2) = P(A_2|A_1 \cup A_2 \cup A_3) = P(A_2|A_1 \cup A_2 \cup A_3 \cup A_4) = 1/4 \\ P(A_j|A_i \cup A_j) = P(A_j|A_i \cup A_3 \cup A_4) = 0,\ P(A_j|A_3 \cup A_4) = 1/2 \\ i = 1, 2, j = 3, 4. \end{array}$$

The following theorem, which may be proved by using a relevant result given in [22] and the concept of coherence for probability assessments on unconditional events, focuses the meaning of coherence for a probabilistic assessment on conditional events. In fact coherence does not coincide with the existence of a probability distribution but with the existence of a class of probability distributions defined on suitable subsets of the set $\mathcal{A}$ of atoms generated by the given conditional events, which permits to write a probability of a conditional event as a ratio of two (unconditional) probability.

Theorem 1. *Let $\mathcal{C}$ be an arbitrary finite family of conditional events and let $\mathcal{A}_0$ be the relevant set of atoms. For a real function P on $\mathcal{C}$ the following two statements are equivalent:*

(i) P is a restriction of a conditional probability on $\mathcal{E} \times (\mathcal{E}\backslash\{\emptyset\})$, where $\mathcal{E}$ is the algebra spanned by $\mathcal{A}_0$;

(ii) there exists (at least) a class of probabilities $\{P_0, P_1, ...\}$, each probability P_k being defined on a suitable subset $\mathcal{A}_k \subseteq \mathcal{A}_0$, such that for any $E_i|H_i \in \mathcal{C}$ there is a unique P_k with

$$\sum_{A_r \subseteq H_i}^{r} P_k(A_r) > 0 \text{ and } P(E_i|H_i) = \frac{\sum_{A_r \subseteq E_i H_i}^{r} P_k(A_r)}{\sum_{A_r \subseteq H_i}^{r} P_k(A_r)};$$

moreover $\mathcal{A}_{k'} \subseteq \mathcal{A}_{k''}$ for $k' > k$" and $P_{k''}(A_r) = 0$ if $A_r \in \mathcal{A}_{k'} \subseteq \mathcal{A}_{k''}$.

In the literature (see, [21], [25]) the following condition of coherence, related to de Finetti's betting scheme, is present.

A real function $P : \mathcal{C} \to R$ is de Finetti-coherent if, for every finite subfamily $\mathcal{F} \subseteq \mathcal{C}$, and for every choice of $\lambda_1, \ldots, \lambda_n \in R$, the possible values of the restriction of the corresponding random gain $G = \sum_{i=1}^{n} \lambda_i h_i(e_i - p_i)$, to $H_0 = H_1, ..., H_n$, are neither all positive nor all negative, that is if

$$\sup_{A_r \subseteq H_0} G \geq 0.$$

The meaning in terms of coherent bets is similar to that of coherence for unconditional events, with the additional condition that the bet is annulled if no hypothesis H_i is true.

In [6] and in [12] equivalence between de Finetti coherence and condition (ii) of Theorem 1 is proved in a direct way. This proof sketches an algorithm to test coherence of P, based on the equivalence between condition (ii) and the compatibility of all the systems (S_k), with unknowns $P_k(A_r) \geq 0$, $A_r \in \mathcal{A}_{k'}$

$$\begin{cases} \sum\limits_{A_r \subseteq E_i H_i} P_k(A_r) = P(E_i|H_i) \sum_{A_r \subseteq H_I} P_k(A_r), \quad [\text{ if } P_{k-1}(H_i) = 0] \\ \\ \sum\limits_{A_r \subseteq H_0^k}{}_r \; P_k(A_r) = 1, \end{cases} \qquad (S_k)$$

where $P_{-1}(H_i) = 0$ for all H_i's, and H_0^k denotes, for $k \geq 0$, the union of the H_i's such that $P_{\alpha-1}(H_i) = 0$; so, in particular, $H_0^0 = H_0 = H_1 \cup ... \cup H_n$.

The above result permits (see [12]) to show (similarly to the case of unconditional events) that, given a "new" conditional event $E|H$, the two bounds p' and p" of the de Finetti's fundamental theorem are related, respectively, to the probabilities of the "greatest" union of atoms contained in $E|H$ and of the "smallest" union containing $E|H$, where the inclusion between conditional events is defined as in [18], that is

$$A|H \subseteq \circ B|K \Leftrightarrow AH \subseteq BK \text{ and } B^c K \subseteq A^x H. \qquad (*)$$

For connections between this inclusion and the numerical inequality related to the 3-valued entities $A|H$ and $B|K$ restricted to $H \cup K$ and moreover between the inclusion and coherent probability, see [10].

More precisely, consider a finite family $\mathcal{F}$ of conditional events and a coherent probability P on $\mathcal{F}$. Given a further event $E|H$, and denote by $E_*|H_*[E^*|H^*]$ the "maximum" ["minimum"] event logically dependent on $\mathcal{F}$ contained in $E|H$ [containing $E|H$] in the sense of (*) and by $\mathcal{P}(H)$ the set of all classes $\{P_0, P_1, ...\}$ (defined in Theorem 1). Then a coherent assessment of the conditional probability $P(E|H)$ is any value in the closed interval $p_* \leq P(E|H) \leq p^*$, where $p_* = 0$ $[p^* = 1]$ if on $\mathcal{P}(H_*)$ [on $\mathcal{P}(H^*)$] there is some class whose probabilities P_k satisfy $P_k(H_*) = 0$ $[P_k(H^*) = 0]$ for all k (such that $H_* \subseteq \mathcal{A}_k$ $[H^* \subseteq \mathcal{A}_k]$). On the other cases

$$p_* = \inf_{\mathcal{P}(H_*)} P(E_*|H_*), \quad p^* = \sup_{\mathcal{P}(H^*)} P(E^*|H^*)$$

In [12] is sketched an algorithm to find, for any new event, the interval $[p_*, p^*]$. It is based on solving some problems of linear programming.

The following two examples show that ignoring the possible existence of null events reduces the possibility of extending in any case a coherent conditional probability and that even when this extension exists, the corresponding interval may be smaller than whose values correspond to coherent assignments.

Example 4. Consider the first two conditional events and the corresponding assignments P of Example 3. Introduce as new conditional event $E|H$ the third one, i.e. $E|H$: since $P_k(H) = 0$ (this class has only one element, corresponding to $k = 0$), any value between 0 and 1 is a coherent assessment for $P(E|H)$. But if null events are not allowed as conditioning ones, we need the assumption $x_3 + x_4 > 0$, while coherence of the two given assignments implies $x_3 + x_4 = 0$. *So P cannot be extended to $E|H$.*

Example 5. Given two conditional events $E_1|H_1$, $E_2|H_2$, such that $\mathcal{A}_0 = \{A_1, ..., A_4, A_5\}$, with $H_1 = A_1 \cup A_2 \cup A_3$, $H_2 = A_1 \cup A_3 \cup A_4$, $E_1H_1 = A_1$, $E_2H_2 = A_1$, consider the assessment $P(E_1|H_1) = P(E_2|H_2) = 1/9$ and introduce the new conditional event $E|H = A_4|(A_2 \cup A_4 \cup A_5)$. Now, putting $x_r = P_0(A_r) \geq 0$, the system

$$\begin{cases} x_1 = \frac{1}{9}(x_1 + x_2 + x_3) \\ x_1 = \frac{1}{9}(x_1 + x_3 + x_4 + x_5) \\ x_1 + x_2 + x_3 + x_4 + x_5 = 1 \end{cases}$$

has among its solution also $x_2 = x_4 = x_5 = 0$, $x_1 = \frac{8}{9}x_3$, so that any value between 0 and 1 is a coherent extension of P to $E|H$. Instead, if we require $x_2 + x_4 + x_5 > 0$, since the above system implies anyway that $x_2 = x_4 + x_5$, the only admissible value for $P(E|H)$ is 1/2.

The above results can be extended in a natural way to imprecise probabilities, such as in the case of unconditional events. For instance the condition to check coherence may be obtained by the above one changing the first equation on two inequalities related to the extremes of the interval $P'(E_i|H_i)$, $P"(E_i|H_i)$. The same results instead can not translated in a natural way from the numerical framework in the comparative one, also if they can be used to introduce and to menage a concept of coherence for comparative probability defined on conditional events. Like for the case of unconditional events, we define a comparative probability, defined on an arbitrary set of conditional events $\mathcal{C}$, to be coherent if there exists a coherent numerical probability assessment $\mathcal{P} = (p_1, ..., p_n)$ on $\mathcal{C}$, agreeing with $\leq^*$, in the sense that

$$p_i \leq p_j \text{ if } E_i|H_i \leq^* E_j|H_j \text{ and } p_i < p_j \text{ if } E_i|H_j <^* E_j|H_j.$$

In this case we can not repeat the process made for unconditional events by passing from system (S_1) to system (S_3). For conditional events the process is in fact theoretically more complicate, nevertheless it is possible (as proved in [13]) to express the coherence condition essentially in terms of solvability of parametric linear systems. For the particular case of a comparative probability containing all the relations $\emptyset \leq^* H_i|\Omega$, coherence conditions are studied in [8], [9] and [10].

Remark 2. All the above results are extendible to the infinite case (see the same bibliography quoted above). In the particular case of comparative probability the condition of coherence must be reenforced on the infinite case with an archimedean condition.

Remark 3. We note that when coherence is checked or when, for a new event, the coherence interval needs to be found, any numerical or qualitative conditions (such as independence) can be taken into account. Such conditions will be translated on equations which will be added to the other ones on the relevant systems.

Remark 4. Notice that in this general context nothing can be used in a "superficial" way: for instance if we want to update a (coherent) probability asssessment on some hypotheses using Bayes theorem, we first need to prove that the probabilities of the likelihoods are coherent and moreover that probability assessments of hypoteses and of likelihoods are coherent together. In the usual Bayesian model in fact the particular choice of logical structure of the hypotheses (exhaustive and incompatible) grants the coherence. In the case that we have coherence, then Bayes theorem can be applied, but the result is in general an interval, not a number (see [11]).

References

[1] Adams, E.W. & Levine H.P. (1975). On the uncertainties transmitted from premises to conclusions in deductive inference, *Synthese,* **30**, 429-460

[2] Boole, G. (1854) *An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities.* Walton and Maberley, London. (reprint N.Y. Dover 1958)

[3] Bruno, G. & Gilio, A. (1980). Applicazione del metodo del simplesso al teorema fondamentale per le probabilita' nella concezione soggettivistica. *Statistica*, **40**, 337-344.

[4] Capotorti, A & Vantaggi, B. (1996) *Consistency and Logic Frames in Belief and Probability Assessments.* Preprint.

[5] Coletti, G. (1990). Coherent qualitative probability. *Journal of Mathem. Psych.,* **34**, 297-310.

[6] Coletti, G. (1993). Numerical and qualitative judgments in probabilistic expert systems. In *International Workshop on Probabilistic Methods in Expert Systems,* Ed. R.Scozzafava, 37-55. Roma: SIS.

[7] Coletti, G. (1994). Coherent numerical and ordinal probabilistic assessments. *IEEE Transactions on Systems, Man, and Cybernetics,* **24**, n.12, 1747-1754.

[8] Coletti, G., Gilio, A. & Scozzafava, R. (1990). Coherent qualitative probability and uncertainty in Artificial Intelligence.Proceed. *8th Intern. Congress of Cybernetic and Systems,* vol. I (Ed. C.N. Manikopoulos) NJIT Press, N.Y., 132-138.

[9] Coletti, G., Gilio, A. & Scozzafava, R. (1991). Assessment of qualitative judgements for conditional events in expert systems. *Lecture Notes in Computer Sciences* n.548 (Eds. R. Kruse and P. Siegel) 135-140. Berlin: Springer-Verlag.

[10] Coletti, G., Gilio, A. & Scozzafava, R. (1993). Comparative probability for conditional events: a new look through coherence. *Theory and Decision,* 35, 237-258.

[11] Coletti, G. & Scozzafava, R. (1996). - Learning from data by coherent probabilistic reasoning, *Proc. of ISUMA-NAFIPS'95,* University of Maryland September 17-20, 1993.

[12] Coletti, G. & Scozzafava, R. (1995). Characterization of coherent conditional probabilities as a tool for their assessment and extension. *Jour. of Uncertainty, Fuzziness and Knowledge-based Systems,* 4, n. 3.

[13] Coletti, G. & Scozzafava, R. (1996). Characterization of coherent comparative probabilities on conditional events. (In preparation)

[14] de Finetti, B. (1949). Sull'impostazione assiomatica del calcolo delle probabilita'. *Annali Univ. Trieste,* **19**, 3-55 (Engl. transl. (1972): Ch.5 in Probability, Induction, Statistics. London: Wiley).

[15] de Finetti, B. (1970). *Teoria della probabilita'.* Torino: Einaudi (Engl. transl. (1974): Theory of Probability, Chichester: Wiley).

[16] Druzdzel, M.J. & van der Gaag L.C. (1995) Elicitation of Probabilities for Belief Networks: Combining Qualitative and Quantitative Information. *Proceed. of UAI-95,* Montreal, 141-148. San Mateo: Morgan Kaufman

[17] Gilio, A. (1995) Algorithms for Precise and Imprecise Conditional Probability Assessments. In *Mathematical Models for handling Partial knowledge in Artificial Intelligence* (Eds. G.Coletti, D.Dubois, R. Scozzafava), N.Y.: Plenum.

[18] Goodman, I.R. & Nguyen, H.T. (1988). Conditional objects and the modeling of uncertainties. In *Fuzzy Computing*, Eds. Gupta, M. and Yamakawa, T., 119-138. Amsterdam: North Holland.

[19] Hailperin, T. (1965). Best Possible Inequalities for the Probability of a Logical Function of Events. *Am. Math. Monthly*, **72**, 343-359.

[20] Hansen, P., Douanya Nguets, G.B., Joumard, B. & Poggi de Aragao, M. (1995). Models and Algorithms for Probabilistic and Bayesian Logic, *Proceed. of IJCAI- 95*, vol. 2, 1862-1868.

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

[22] Krauss, P.H. (1968). Representation of conditional probability measures on Boolean algebras. *Acta Mat. Acad.Scient. Hungar.,* **19**, 229-241.

[23] Lad, F., Dickey, J.M. & Rahman, M.A. (1990). The fundamental theorem of prevision. *Statistica*, **50**, 19-38.

[24] Lad, F., Dickey, J.M. & Rahman, M.A. (1992). Numerical application of the fundamental theorem of prevision. *J. Statist. Comput. Simul.,* **40**, 135-151.

[25] Lehman, R.S. (1955). On confirmation and rational betting. *The J. of Symbolic Logic,* **20**, 251-262.

[26] Nilsson, N.J. (1986). Probabilistic logic. *Artif. Intel.,* **28**, 71-87.

[27] Regoli, G. (1994). Rational Comparisons and Numerical Representation. In *Decision Theory and Decision Analysis: Trends and Challenges,* Acad. Publ. Dordrecht, 113-126.

[28] Scozzafava, R. (1996). Subjective conditional probability and coherence principles for handling partial information. *In this issue.*