Locally additive comparative probabilities

Giulianella Coletti*
University of Perugia (Italy)
coletti@dipmat.unipg.it

Romano Scozzafava
University of Roma "La Sapienza" (Italy)
romscozz@dmmm.uniroma1.it

Abstract

We characterize binary relations (defined on an *arbitrary* family $\mathcal{E}$ of unconditional events) that are representable by a coherent conditional probability defined on $\mathcal{E} \times (\mathcal{E} \setminus \emptyset)$, and those that are representable by a weakly decomposable conditional measure. Both these relations are locally "additive".

Keywords. Comparative probability, conditional probability, coherence conditions.

1 Introduction

In decision problems, uncertain knowledge may be represented by a probability measure. However, when information is partial and not easily summarizable by a reliable numerical evaluation, then the natural tool for dealing with uncertain knowledge is *comparative* (or qualitative) probability (for the possible use of comparative probability in expert system see, for instance [7], [8], [3], [4]). In this approach, one (the decision maker, the field expert, ...) merely states his preferences (or his degrees of belief) on a set of *propositions* (events) without any quantification, but only through an *ordinal relation*.

The main problem, for an ordinal relation expressing a comparative degree of belief, is the setting up of a system of rules assuring coherence of the relation with respect to the idea that it intends to convey (such as "no less probable than", "no less believable than" and so on). Usually such a problem amounts to the consistency of the ordinal relation with some (numerical) theoretical model.

More precisely, given a numerical framework (probability, belief functions, lower probability, etc.) one finds the properties which are necessary and those which are sufficient for the existence of a numerical assessment (probability, or belief, etc.) on the events, agreeing – in some way – with the ordinal relation.

Let $\mathcal{E}$ be any set of events: denote by $\preceq$ a binary relation in $\mathcal{E}$ and with $\prec$ and $\sim$ the strict relation and the equivalence relation, respectively. If we give the sentence "agreeing with $\preceq$" the meaning of "representing $\preceq$", that is "being strictly monotone with $\preceq$", then for any choice of a *capacity* function as numerical framework of reference, it is necessary that an extension of $\preceq$ to the algebra $\mathcal{A}$ spanned by $\mathcal{E}$ exists, satisfying the following conditions:

($c1$) $\emptyset \preceq E$ for every $E \in \mathcal{A}$, and $\emptyset \prec \Omega$;
($c2$) $\preceq$ is a total preorder;
($c3$) for every $E, F \in \mathcal{A}$, $E \subset F \Rightarrow E \preceq F$,

where $\emptyset$ and Ω are, respectively, the *impossible* and the *certain* event.

When we specialize the capacity function (probability, belief, plausibility, and so on) representing $\preceq$, then we need adding to the above axioms a specific relevant condition, which essentially expresses a (more or less strong) sort of "qualitative additivity". The first (and the most known) additivity axiom (de Finetti [13], Koopman [19]) is the following

(p) for every $E, F, H \in \mathcal{A}$, with $E \wedge H = F \wedge H = \emptyset$, both the following implications hold:

$$E \preceq F \Rightarrow E \vee H \preceq F \vee H$$
$$E \prec F \Rightarrow E \vee H \prec F \vee H\,.$$

In fact the above axiom is necessary for the representability of $\preceq$ with any *additive function* with values in a totally ordered set (also, for instance, the set $\mathbb{R}^*$ of nonstandard real numbers).

If we refer instead to more general measures of uncertainty, such as belief functions, plausibilities and so on, then it is easy to see that (p) can be violated.
Nevertheless, also in this case a weaker additivity axiom is necessary; see, for this aspect, the following condition (b) introduced in [24], characterizing relations representable by a belief function, and conditions (pl), (l), (u) introduced in [1], characterizing

*Dip. Matematica - Via Vanvitelli - 06123 Perugia (Italy) relations representable by a plausibility, a lower probability, an upper probability, respectively:

(b) $\forall E, F, H \in \mathcal{A}$, with $E \subseteq F$ and $F \wedge H = \emptyset$,
$$E \prec F \Rightarrow E \vee H \prec F \vee H$$

(pl) $\forall E, F, H \in \mathcal{A}$, with $E \subseteq F$ and $F \wedge H = \emptyset$,
$$E \sim F \Rightarrow E \vee H \sim F \vee H$$

(l) $\forall E, F \in \mathcal{A}$, with $E \wedge F = \emptyset$,
$$\emptyset \prec E \Rightarrow F \prec F \vee E$$

(u) $\forall E, F \in \mathcal{A}$, with $E \wedge F = \emptyset$,
$$\emptyset \sim E \Rightarrow F \sim E \vee F\,.$$

As proved in [24] and [1], for a binary relation defined on a finite set of events satisfying $(c1)-(c3)$, conditions (b), (pl), (l) and (u) are also sufficient for the representability of $\preceq$ by a belief, a plausibility, a lower or an upper probability, respectively.
We note that none of the above conditions requires that either of the two implications in (p) be satisfied. In fact all of them involve only events related by inclusion relations ($E \subseteq F$ or $\emptyset \subseteq E$).

On the contrary, notice that the binary relation induced by a conditional probability by putting:

$$\begin{array}{lll} E \sim F & \text{if} & P(E|E \vee F) = P(F|E \vee F) \\ E \prec F & \text{if} & P(E|E \vee F) < P(F|E \vee F) \end{array}$$

satisfies the first implication of (p), while the second one can be violated when $P(E \vee F | E \vee F \vee H) = 0$.

The main aim of this paper is to characterize binary relations "locally representable" (see below) by a conditional probability P. We will study this problem in a completely general context, *i.e.* for a binary relation defined on an *arbitrary* set of events and *not necessarily complete*. Our numerical framework of reference will be the theory of coherent conditional probabilities and their characterizations in terms of families of probabilities (see for instance [5, 10]).

More precisely, if $\mathcal{E}$ is an arbitrary set of events E_i, and $\preceq$ a (possibly partial) binary relation, we find necessary and sufficient conditions for the existence of a coherent conditional probability, defined on the following subset of $\mathcal{E} \times \mathcal{E}$

$$\mathcal{E}^* = \{E|E \vee F \,:\, E \preceq F \text{ or } F \preceq E\}\,,$$

representing $\preceq$, that is such that:

$$(*) \qquad \begin{array}{l} E \preceq F \Rightarrow P(E|E \vee F) \leq P(F|E \vee F) \\ E \prec F \Rightarrow P(E|E \vee F) < P(F|E \vee F) \end{array}\,.$$

We characterize also comparative relations representable by more general conditional measures: we find necessary and sufficient conditions for the existence of a weakly $(\oplus, \odot)$-decomposable measure φ satisfying condition $(*)$ with φ in place of P, where $\oplus$ and $\odot$ are arbitrary operations on $[0,1]^2$ satisfying suitable properties on particular sets.

2 The numerical model of reference

What is usually emphasized in the literature – when a conditional probability $P(E|H)$ is taken into account – is only the fact that $P(\cdot|H)$ *is a probability for any given* H: this is a very restrictive (and misleading) view of conditional probability, corresponding trivially to just a modification of the so-called "sample space" Ω.

It is instead essential – for a correct handling of the subtle and delicate problems concerning the use of conditional probability – to regard the conditioning event H as a "variable", *i.e.* the "status" of H in $E|H$ is not just that of something representing a given *fact*, but that of an (uncertain) *event* (like E) for which the knowledge of its truth value is not required (this means, using a terminology due to Koopman [19], that H must be looked on – even if *asserted* – as being *contemplated*: similar terms are, respectively, *acquired* versus *assumed*).

We generalize (or better, in a sense, we give up) the idea of de Finetti of looking at a conditional event $E|H$, with $H \neq \emptyset$, as a 3–valued logical entity (*true* when both E and H are true, *false* when H is true and E is false, "undetermined" when H is false) by letting the third value *suitably depend on the given ordered pair* (E, H) and not being just an undetermined *common value* for all pairs: it turns out (as explained in detail in [11], [12]) that this function is a measure of the degree of belief in the conditional event $E|H$, which under suitable (and natural) conditions is the conditional probability $P(E|H)$ (in its most general sense related to the concept of *coherence*, and satisfying the classic axioms as given by de Finetti [14], Rényi [22], Dubins [15]), or, more generally, a decomposable conditional measure (see below).

A peculiarity (which entails a large flexibility in the management of any kind of uncertainty) of this concept of *coherent* conditional probability is that, due to its *direct* assignment as a whole, the knowledge (or the assessment) of the "joint" and "marginal" unconditional probabilities $P(E \wedge H)$ and $P(H)$ is not required; moreover, the *conditioning* event H (which *must* be a *possible* event) may have *zero probability.*

The classic *axioms for conditional probability* (given a set $\mathcal{C} = \mathcal{G} \times \mathcal{B}^o$ of conditional events $E|H$ such that $\mathcal{G}$ is a Boolean algebra and $\mathcal{B} \subseteq \mathcal{G}$ is closed with respect to (finite) logical sums, and putting $\mathcal{B}^o = \mathcal{B} \setminus \{\emptyset\}$) are:

(i) $P(H|H) = 1$, for every $H \in \mathcal{B}^o$,

(ii) $P(\cdot|H)$ is a (finitely additive) probability on $\mathcal{G}$ for any given $H \in \mathcal{B}^o$,

(iii) $P\big((E \wedge A)|H\big) = P(E|H) \cdot P\big(A|(E \wedge H)\big)$, for

every E, $A \in \mathcal{G}$ and E, $E \wedge H \in \mathcal{B}^o$.

A conditional probability P is defined on $\mathcal{G} \times \mathcal{B}^o$: however it is possible, through the concept of *coherence*, to handle also those situations where we need to assess P on an *arbitrary* set of conditional events $\mathcal{C} = \{E_1|H_1, \ldots, E_n|H_n\}$.

Definition 1 - *The assessment $P(\cdot|\cdot)$ on $\mathcal{C}$ is coherent if, given $\mathcal{C}' \supset \mathcal{C}$, with $\mathcal{C}' = \mathcal{G} \times \mathcal{B}^o$, it can be extended from $\mathcal{C}$ to $\mathcal{C}'$ as a conditional probability.*

A characterization of coherence is given by the following theorem (see, e.g., [5], [10], [11]).

Theorem 1 - *Let $\mathcal{C}$ be an arbitrary finite family of conditional events and $\mathcal{A}_o$ denote the set of atoms A_r generated by the events $E_1, H_1, \ldots, E_n, H_n$. For a real function P on $\mathcal{C}$ the following two statements are equivalent:*

(i) P is a coherent conditional probability on $\mathcal{C}$;

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

$$\sum_{A_r \subseteq H_i}^{r} P_\alpha(A_r) > 0\,, \quad P(E_i|H_i) = \frac{\sum_{A_r \subseteq E_i \wedge H_i}^{r} P_\alpha(A_r)}{\sum_{A_r \subseteq H_i}^{r} P_\alpha(A_r)}\;;$$

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

Any class $\{P_\alpha\}$ singled-out by the condition (ii) is said *to agree* with the conditional probability P.

The proof of the equivalence between conditions (i) and (ii) gives rise to an algorithm to test the coherence of the assessment P, based on the equivalence between condition (ii) and the compatibility of a sequence of systems $(\mathcal{S}_\alpha)$ with unknowns $P_\alpha(A_r) \geq 0$, $A_r \in \mathcal{A}_\alpha$,

$$(\mathcal{S}_\alpha) \quad \begin{cases} \sum_{A_r \subseteq E_i \wedge H_i}^{r} P_\alpha(A_r) = P(E_i|H_i) \sum_{A_r \subseteq H_i}^{r} P_\alpha(A_r) \\ \qquad\qquad\qquad\qquad [\text{if } P_{\alpha-1}(H_i) = 0]\,, \\ \sum_{A_r \subseteq H_0^\alpha}^{r} P_\alpha(A_r) = 1 \end{cases}$$

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

As proved in the aforementioned papers, conditions (i) and (ii) are equivalent also to the following *de Finetti's coherence* (as expressed, for example, in [21]), where $p_i = P(E_i|H_i)$:

(iii) for any choice of the real numbers $\lambda_1, ..., \lambda_n$

$$\sup_{A_r \wedge H_o} \sum_{i=1}^{n} \lambda_i H_i (E_i - p_i) \geq 0\,,$$

where $H_o = \bigvee_{i=1}^{n} H_i$.

The random quantity

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

can be interpreted as the gain corresponding to a combination of n bets of amounts $\lambda_1 p_1, \ldots, \lambda_n p_n$ on $E_1|H_1, \ldots, E_n|H_n$, with arbitrary stakes $\lambda_1, \ldots, \lambda_n$.

The previous theory has been extended in [12] to general (decomposable) conditional measures; we recall here some definitions and results:

Definition 2 - *Given a boolean algebra $\mathcal{E}$, a* weakly *$\oplus$–decomposable measure $\varphi : \mathcal{E} \to [0, 1]$ is a* capacity *such that there exists an operation $\oplus$ from $\varphi(\mathcal{E}) \times \varphi(\mathcal{E})$ to $I\!R^+$ satisfying the following condition: for every $E_i, E_j \in \mathcal{E}$, with $E_i \wedge E_j = \emptyset$,*

$$\varphi(E_i \vee E_j) = \varphi(E_i) \oplus \varphi(E_j)\,.$$

It is easily seen that, with respect to the elements of the following subset of $\varphi(\mathcal{E}) \times \varphi(\mathcal{E})$

$$\mathcal{K} = \{(\varphi(E_i), \varphi(E_j)) : E_i, E_j \in \mathcal{E}, E_i \wedge E_j = \emptyset\}\,,$$

the operation $\oplus$ is commutative, associative, increasing and admits 0 as neutral element. Nevertheless, as proved by Example 2 of [12], it need not be extensible to a function defined on the whole $\varphi(\mathcal{E}) \times \varphi(\mathcal{E})$ (and so neither on $[0, 1]^2$) and *satisfying the same properties.*

Definition 3 - *Given a family $\mathcal{C} = \mathcal{E} \times \mathcal{H}^0$ of conditional events, where $\mathcal{E}$ is a boolean algebra, $\mathcal{H}$ an additive set, with $\mathcal{H} \subseteq \mathcal{E}$ and $\mathcal{H}^0 = \mathcal{H} \backslash \{\emptyset\}$, a real function φ defined on $\mathcal{C}$ is a weakly $(\oplus, \odot)$–decomposable conditional measure if*

(γ_1) $\varphi(E|H) = \varphi(E \wedge H|H)$, for every $E \in \mathcal{E}$ and $H \in \mathcal{H}^o$,

(γ_2) there exists an operation $\oplus : \varphi(\mathcal{C}) \times \varphi(\mathcal{C}) \to \varphi(\mathcal{C})$ whose restriction to the set

$$\Delta = \{(\varphi(E_i|H), \varphi(E_j|H)) : E_i, E_j \in \mathcal{E},\, H \in \mathcal{H}^0\},$$

with $E_i \wedge E_j \wedge H = \emptyset$, is (commutative, associative and) increasing, admits 0 as neutral element, and is such that, for any given $H \in \mathcal{H}^o$, $\varphi(\cdot|H)$ is a weakly $\oplus$–decomposable measure,

(γ_3) there exists an operation $\odot : \varphi(\mathcal{C}) \times \varphi(\mathcal{C}) \to \varphi(\mathcal{C})$ whose restriction to the set

$\Gamma = \{(\varphi(E|H), \varphi(A|E \wedge H)) : A \in \mathcal{E},\ E, H, E \wedge H \in \mathcal{H}^o\}$ *is (commutative, associative and) increasing, admits* 1 *as neutral element and is such that, for any* $A,\ E \in \mathcal{E}$ *and* $E,\ E \wedge H \in \mathcal{H}^o$,

$$\varphi\big((E \wedge A)|H\big) = \varphi(E|H) \odot \varphi\big(A|(E \wedge H)\big)\,,$$

(γ_4) The operation $\odot$ is distributive over $\oplus$ for relations of the kind

$$\varphi(H|K) \odot \big(\varphi(E|H \wedge K) \oplus \varphi(F|H \wedge K)\big)\,,$$

with $K, H \wedge K \in \mathcal{H}^0\,,\ E \wedge F \wedge H \wedge K = \emptyset\,.$

Definition 4 - *$\mathcal{E}$ is a finite Boolean algebra, $\mathcal{H}$ an additive set, with $\mathcal{H} \subseteq \mathcal{E}$ and $\mathcal{H}^0 = \mathcal{H} \setminus \{\emptyset\}$, and $\mathcal{A} = \{A_r\}_{r=1,2,\ldots,m}$ is the set of atoms of $\mathcal{E}$. Let $\{\mathcal{A}_\alpha\}$ be a class of subsets of atoms, with $\mathcal{A}_{\alpha"} \subset \mathcal{A}_{\alpha'}$ for $\alpha" > \alpha'$, $\mathcal{A}_0 = \mathcal{A}$, and, given two operations $\oplus$ and $\odot$ from $I\!R^+ \times I\!R^+$ to $I\!R^+$, let $\{\varphi_0, \varphi_1, \ldots\}$ be a relevant class of $\oplus$–decomposable measures defined on $\mathcal{E}$ such that, for any α, the equation*

$$(1) \qquad \varphi_\alpha(E_i H_i) = x \odot \varphi_\alpha(H_i)\,.$$

has a solution $x \in [0,1]$. Moreover $\varphi_{\alpha"}(A_r) = 0$ for every $A_r \in \mathcal{A} \setminus \mathcal{A}_{\alpha"}$, and an atom A_r belongs to $\mathcal{A}_{\alpha"}$, with $\alpha" \geq 1$, if and only if there exists $H_i \in \mathcal{H}^0$, with $A_r \subseteq H_i$, such that, for every $\alpha < \alpha"$, there exists $E_i \in \mathcal{E}$ for which there is not a unique *solution of equation (1).*

The elements of the class $\{\varphi_0, \varphi_1, \ldots, \varphi_k\}$, with $k \leq m$, will be called *almost generating measures.* If $\odot$ is distributive over $\oplus$, they will be called *generating measures.*

In [12] a general result related to characterization of weakly $(\oplus, \odot)$–decomposable conditional measures is proved: we state here a theorem which is a corollary of that one.

Theorem 2 - *Let $\mathcal{C} = \mathcal{E} \times \mathcal{H}^0$, with $\mathcal{E}$ a boolean algebra, $\mathcal{H}$ an additive set, $\mathcal{H} \subseteq \mathcal{E}$ and $\mathcal{H}^0 = \mathcal{H} \setminus \{\emptyset\}$, a finite family of conditional events, and let $\mathcal{A} = \{A_r\}$ denote the set of atoms of $\mathcal{E}$. Let φ be a real function defined on $\mathcal{C}$, and $\oplus, \odot$ two operations from $\varphi(\mathcal{C}) \times \varphi(\mathcal{C})$ to $I\!R^+$. Then the following two statements are equivalent:*

(a) φ is a weakly $(\oplus\,,\ \odot)$–decomposable conditional measure on $\mathcal{C}$, with $\oplus$ and $\odot$ strictly increasing on Δ and Γ respectively and $\odot$ distributive over $\oplus$;

(b) there exists a (unique) class of generating $\oplus$–decomposable measures such that, for any $E_i|H_i \in \mathcal{C}$, there is a unique α such that $x = \varphi(E_i|H_i)$ is the unique solution of the equation

$$(2) \qquad \bigoplus_{A_r \subseteq E_i H_i} \varphi_\alpha(A_r) = x \odot \bigoplus_{A_r \subseteq H_i} \varphi_\alpha(A_r)\,.$$

3 Some Examples

We discuss now some example to introduce the relevant topics.

Example 1 - Given an experiment consisting of two tosses of a coin, consider the following events:

A_1 = "In the first toss the coin stands up (or is lost) and in the second toss it shows heads",

A_2 = "In the first toss the coin stands up (or is lost) and in the second toss it shows tails",

A_3 = "In both tosses the coin shows heads".

Certainly, if we have a very low degree of belief in the coin standing up (or being lost), the most reasonable ordinal relation $\preceq$ expressing the comparative degree of belief on the occurrence of the above events, is the following:

$$\emptyset \prec A_1 \sim A_2 \prec A_1 \vee A_2 \prec A_3 \sim A_1 \vee A_3 \sim$$
$$\sim A_2 \vee A_3 \sim A_1 \vee A_2 \vee A_3.$$

The next example takes in consideration the case that the expert, or the decision maker, orders with respect to his degree of belief some (possibly, a finite number) events picked out from a necessarily infinite class constituting the model of the problem.

Example 2 - Consider the process of recording the rain quantity fallen on New York during June. The data base consists of ten numbers $x_1, \ldots, x_n$, $x_i \neq x_j$ for $i \neq j$, representing the rain quantities of the last ten years. Let now x_0 be a quantity different from all the previous ones and, putting X= *"rain quantity in New York in the next month of June"*, consider the following events:

$$B_1 = \{X = x_0\}, \qquad B_2 = \bigvee_{i=1}^{10} \{X = x_i\},$$

$$B_3 = \{X < \tfrac{1}{2} min\{x_i\}\},$$

$$B_4 = \{min\{x_i\} < X < max\{x_i\},\ X \neq x_i\}.$$

One possible "natural" relation expressing the degrees of belief on the occurrence of the given events is the following:

$$\emptyset \prec B_1 \prec B_2 \prec B_1 \vee B_2 \prec B_3 \sim B_1 \vee B_3 \sim B_2 \vee B_3 \sim$$
$$\sim B_1 \vee B_2 \vee B_3 \prec B_4 \sim B_1 \vee B_2 \vee B_4 \prec$$
$$\prec B_4 \vee B_3 \sim B_4 \vee B_1 \vee B_2 \vee B_3\,.$$

Going back to Example 1 and taking into account that $A_1 \wedge A_3 = (A_1 \vee A_2) \wedge A_3 = \emptyset$ and $A_1 \subset A_1 \vee A_2$, then the relations $A_1 \prec A_1 \vee A_2$ and $A_1 \vee A_3 \sim A_1 \vee A_2 \vee A_3$ imply that there exists neither an additive nor a belief function representing $\preceq$.

With similar considerations, we can conclude that also in Example 2 there is neither an additive nor a belief function representing $\preceq$. Yet, there exists a plausibility (and so an upper probability) representing the comparative structures of Examples 1 and 2.

Nevertheless, we notice that both comparative assessments satisfy a condition stronger than (*pl*). In fact they satisfy the first implication of condition (*p*), and moreover they are locally representable by a coherent conditional probability. In particular, the comparative structure of Example 1 can be locally represented by the following conditional probability ($i = 1, 2$):

$$P(A_i|A_1 \vee A_2) = 1/2\,,$$
$$P(A_i|A_i \vee A_3) = P(A_i|A_1 \vee A_2 \vee A_3) =$$
$$P(A_1 \vee A_2|A_1 \vee A_2 \vee A_3) = 0\,,$$
$$P(A_3|A_i \vee A_3) = P(A_3|A_1 \vee A_2 \vee A_3) =$$
$$= P(A_1 \vee A_2|A_1 \vee A_2) = P(A_3 \vee A_i|A_1 \vee A_2 \vee A_3) = 1\,.$$

Analogously, the comparative structure of Example 2 can be locally represented by a conditional probability, which is the additive extension of the following assessment:

$$P(B_1|B_1 \vee B_2) = \tfrac{1}{3}\,,\ P(B_2|B_1 \vee B_2) = \tfrac{2}{3}\,,$$
$$P(B_i|H) = P(B_i|K) = 0\,,\ P(B_3|H) = P(B_4|K) = 1\,,$$
$$\text{with}\, H \supseteq B_i \vee B_3,\ K \supseteq B_i \vee B_4,\ i = 1, 2\,;$$
$$P(B_j|B_1 \vee B_2 \vee B_j) = 1,\ j = 3, 4\,,$$
$$P(B_3|W) = \tfrac{1}{4}\,,\ P(B_4|W) = \tfrac{3}{4}\,,\ \text{for}\ W \supseteq B_3 \vee B_4\,.$$

4 Local representation as a tool to manage partial knowledge

A comparative structure "local representable" by a conditional probability can be also a good model for comparative degrees of belief between default rules.

Example 3 - Consider the rule: A= "Typically, birds can fly", and the following comparative structure

$$\emptyset \prec \neg A \prec A \sim A \vee \neg A\,.$$

The conditional probability P such that

$$P(\emptyset|\neg A \vee \emptyset) = P(\neg A|\neg A \vee A) = 0\,;$$
$$P(\neg A|\neg A \vee \emptyset) = P(A|A \vee \neg A) =$$
$$= P(A \vee \neg A|A \vee \neg A) = 1.$$

locally represents this binary relation.

We consider now a situation arising in inferential (Bayesian) statistics, concerning the so–called "improper" distributions: we recall the notion of *pseudodensity* introduced in [23].

Definition 5 - *Given a comparative probability $\preceq$ on a set $\mathcal{C}$ of atoms, let X be a random variable (a map from $\mathcal{C}$ to $\Theta \subseteq I\!R$). A pseudodensity α of X is a function defined on $I\!R$, positive on Θ, representing $\preceq$, i.e., given $x, y \in \Theta$ and putting $X^{-1}(x) = C_x$,*

$$\alpha(x) \leq \alpha(y) \Longleftrightarrow C_x \preceq C_y.$$

Trivial examples are the following: (*i*) Let X be a *discrete* random variable with values in $\Theta \subseteq I\!R$ and with a discrete (everywhere positive) probability distribution $P(X = x) > 0$ for every $x \in \Theta$: clearly, the function $\alpha(x) = P(X = x)$ is a pseudodensity of X. (*ii*) If X is a continuous random variable with probability *density* $f(x)$ (> 0 for $x \in \Theta$), then f is a pseudodensity of X.

Notice that every point x of the support Θ of α corresponds to the atom $C_x = \{X = x\}$. The function defined on $\mathcal{C} = \{C_x|C_x \vee C_y\}$ by putting, for $x \neq y$,

$$P(C_x|C_x \vee C_y) = \frac{\alpha(x)}{\alpha(x) + \alpha(y)}$$

is a *coherent* conditional probability, as can be easily proved using Theorem 5 of [10] (for a direct proof, see [18]). It locally represents $\preceq$, that in general may not be representable by a (non–conditional) probability.

Consider the comparative probability on the set of atoms $\mathcal{C} = \{C_x : x \in [0, 1]\}$, defined as follows: $\emptyset \prec C_x$ and $C_x \sim C_y$ for every $x, y \in [0, 1]$. This ordinal relation is represented by the class of constant pseudodensities $\alpha(x) = k$ for every $x \in [0, 1]$, with $k > 0$. Hence, $\preceq$ is locally represented by the conditional probability

$$P(C_x|C_x \vee C_y) = \frac{1}{2}$$

for every pair of atoms $C_x, C_y \in \mathcal{C}$. Notice that for $k = 1$ the pseudodensity $\alpha(x) = k$ can be seen also as a uniform *density* on the bounded interval $[0, 1]$. But when x belongs to an unbounded interval, or, more generally, to an *arbitrary* subset of $I\!R$, α is not a density: in the statistical literature it is dubbed as an "improper" distributions (because its integral is not finite). Nevertheless, in our framework α is a proper tool, since it is just a point function, with no underlying measure. The pseudodensity $\alpha(x) = k$ for every $x \in \Theta$ (*arbitrary* subset of $I\!R$, bounded or not, measurable or not) is called *uniform pseudodensity*.

5 Weak local coherence

We consider now a comparative probability (possibly *partial*, and translating the idea of *not more probable than*) on a set of (unconditional) events $\mathcal{E}$. Let $\prec$ denote the strict relation (*i.e., less probable than*) and let $\sim$ be the equivalence relation (*i.e., equally probable as*).

Let $\mathbf{S} = \{(E, F) : E \prec F\}$, $\mathbf{E} = \{(E, F) : E \sim F\}$, $\mathbf{T} = \{(E, F) : E \preceq F\}$. We have $\mathbf{S} \cap \mathbf{E} = \emptyset$ and $\mathbf{S} \cup \mathbf{E} \subset \mathbf{T}$, where the inclusion can be strict, if there is some pair (E, F) such that E is judged not more probable than F, but there is no information (at present) that allows to be more specific.

Definition 6 - *A* weakly locally coherent comparative probability $\preceq$ *is a comparative probability satisfying* $(c3)$ *and the following axioms:*

$(c1')$ $\emptyset \prec E$ *for every* $E \in \mathcal{E}$, $E \neq \emptyset$

$(c2')$ $\preceq$ *has no intransitive cycles.*

(cp) *for every* $E, F, H, E\vee H, F\vee H \in \mathcal{E}$ *with* $E\wedge H = F \wedge H = \emptyset$

$$E \preceq F \Rightarrow \neg(F \vee H \prec E \vee H)$$

moreover, if $F \prec F \vee H$ *or* $F \sim H$, *then*

$$E \prec F \Rightarrow \neg(F \vee H \preceq E \vee H).$$

The above system of axioms, introduced in [6], is the natural generalization of that proposed in [9], which referes to a complete binary relation defined on an algebra of events. These axioms are $(c1'), (c2), (c3)$ and the following:

$(C4)$ for every $E, F, H \in \mathcal{E}$, with $H \wedge (E \vee F) = \emptyset$, if $F \prec F \vee H$ or $F \sim H$, then

$$E \prec F \Leftrightarrow (E \vee H \prec F \vee H).$$

The axioms of Definition 6 in fact are necessary to extend $\preceq$ to a complete relation on a Boolean algebra satisfying the axioms given in [9], and they are sufficient if (cp) is required on the transitive closure (*i.e.*, the smallest, with respect to $\subseteq$, transitive relation extending $\preceq$). In this case, we call *almost complete* the comparative probability $\preceq$.

We just note that, by (c1'), for a weakly locally coherent comparative probability, any *possible* event is strictly "more probable" than the *impossible* one. This intuitive axiom (that was already in de Finetti [13]) has been later weakened, essentially in order to represent $\preceq$ by a (non–conditional) probability. Moreover, axiom (cp) is an actual weakening of axiom (p); in fact it requires the additivity only for events E, F of the "same order of probability", in the following sense.

For a comparative structure $(\mathcal{E}, \preceq)$, with $\preceq$ a weak locally coherent comparative probability, we can associate to every event E the family $\mathcal{A}(E)$ of the events "*infinitely less probable*" than E, and then the family $\mathcal{B}(E)$ of the events which are of the "*same order of probability*" as E, with respect to the comparative probability $\preceq$. So we define:

$$\mathcal{A}(E) = \{F \in \mathcal{E} : \exists E_i \sim F_i \preceq E,\ F_i \subset E_i\},$$

with $i = 1, \ldots, n$, and $F \subseteq \bigvee_{i=1}^{n} (E_i \wedge F_i^c)$.

We note that if $\preceq$ is *complete*, satisfies $(B1) - (B4)$ and $\mathcal{E}$ is an *algebra*, then for every $E \in \mathcal{E}$ the set $\mathcal{A}(E)$ coincides with the set $\mathcal{A}_E$, whose definition clearly intends to express the meaning as a class of events "*infinitely less probable than E*":

$$\mathcal{A}_E = \{F \in \mathcal{E} : E \sim E \vee F \sim E \wedge F^c\}.$$

The proof that $\mathcal{A}(E) \supseteq \mathcal{A}_E$ is immediate, considering $E_1 = E \vee F$ and $F_1 = E \wedge F^c$. We prove now that $\mathcal{A}(E) \subseteq \mathcal{A}_E$. First notice that if $G \subseteq F$ and $F \in \mathcal{A}_E$, then, by definition of $\mathcal{A}_E$, using (cp), we have that $G \in \mathcal{A}_E$. Therefore it is sufficient to prove that any $F \in \mathcal{A}(E)$, with $F = \bigvee_{i=1}^{n}(E_i \wedge F_i^c)$, is an element of $\mathcal{A}_E$. Putting $K_i = (E_i \wedge F_i^c)$, taking into account the definition of $\mathcal{A}(E)$ and axiom (cp), for every $i = 1, \ldots n$ we have

$$E \sim E_i = F_i \vee K_i \sim E \vee K_i\,,$$

$$E \sim F_i = F_i \wedge K_i^c \sim E \wedge K_i^c.$$

By (cp) and $(c1)$ we get $E \sim E \vee F \sim E \wedge F^c$.

We can now define the set of events with *the same order of probability* of E

$$\mathcal{B}(E) = \{F \in \mathcal{E} : F \notin \mathcal{A}(E) \,\text{and}\, E \notin \mathcal{A}(F)\}.$$

If $\preceq$ is complete, satisfies $(c1'), (c2), (c3)$ and (cp), and $\mathcal{E}$ is an algebra, the sets $\mathcal{A}_E$ and $\mathcal{B}(E)$ satisfy many structural properties, as proved in the quoted paper [9].

We only recall here that $\{\mathcal{B}(E) : E \in \mathcal{E}\}$ is a partition of $\mathcal{E}$ (independently of the logical structure of $\mathcal{E}$). Moreover, we note that for every $E, F \in \mathcal{E}$, putting $G = \max_{\preceq}\{E, F\}$, we have $E \vee F \in \mathcal{B}(G)$.

Finally, we notice that if, in particular, (p) holds, then for every $E \in \mathcal{E}$ the set $\mathcal{A}(E)$ is empty: in fact in this case all the events are element of $\mathcal{B}(\Omega)$.

Consider now the problem of the local representability of a comparative probability $\preceq$. We first note that if a comparative probability $\preceq$, satisfying axiom $(c1')$, is representable by a (strictly positive) coherent probability, then $\preceq$ is obviously locally representable by a (coherent) conditional probability. Using the examples of Section 3, it is immediate to see that the converse is not true. The following Proposition, whose proof is straightforward, gives a necessary condition for the local representability.

Proposition - *Let* $\preceq$ *be a comparative probability defined on an* arbitrary *family of events* $\mathcal{E}$, *containing the impossible event* $\emptyset$. *If there exists a coherent conditional probability* P, *defined on* $\mathcal{F} \subseteq \mathcal{E} \times \mathcal{E}_0$, *representing* $\preceq$, *then* $\preceq$ *is a weakly locally coherent comparative probability.*

The converse is not true, that is axioms $(c1'), (c2'), (c3), (cp)$ are not sufficient to guarantee the existence of a conditional probability locally representing a comparative probability $\preceq$, even if the latter is a *complete* relation and $\mathcal{E}$ is an *algebra*. Consider in fact the well known example, given in [20], consisting of an algebra spanned by five atoms, and a comparative probability satisfying axioms $(c1)', (c2), (c3), (p)$, but not representable by any additive function.

Clearly, in this case, since (p) holds, the existence of a conditional probability locally representing $\preceq$ coincides with the existence of a (strictly positive) probability representing $\preceq$. Nevertheless, we can give a characterization of weakly locally coherent comparative probabilities in terms of weak decomposable conditional measures.

Theorem 3 - *Let $\mathcal{E}$ be a finite family of events containing $\emptyset$, and $\mathcal{A} = \{A_r\}$ the set of atoms generated by $\mathcal{E}$. If $\mathcal{A}^*$ is the algebra spanned by $\mathcal{A}$, for a comparative probability $\preceq$ in $\mathcal{E}$ the following statements are equivalent:*

(i) $\preceq$ is an almost complete weakly locally coherent comparative probability;

(ii) there exists a class of subsets of atoms $\{\mathcal{A}_\alpha\}$ (with $\mathcal{A}_{\alpha''} \subset \mathcal{A}_{\alpha'}$ for $\alpha'' > \alpha'$ and $\mathcal{A}_0 = \mathcal{A}$) and a relevant class of weakly $\oplus_\alpha$–decomposable measures φ_α, with $\oplus_\alpha$ strictly increasing, defined on $\mathcal{A}^*$ and such that for every $E \preceq F \in \mathcal{E}$ $[E \prec F \in \mathcal{E}]$ there exists a unique α with $\varphi_\alpha(F) > 0$ and $\varphi_\alpha(E) \leq \varphi_\alpha(F)$ $[\varphi_\alpha(E) < \varphi_\alpha(F)]$. Moreover $\varphi_{\alpha"}(A_r) = 0$ for every $A_r \in \mathcal{A} \setminus \mathcal{A}_{\alpha"}$, and an atom A_r belongs to $\mathcal{A}_{\alpha"}$, with $\alpha" \geq 1$, iff $A_r \in \mathcal{A}_{\alpha"-1}$ and $\varphi_{\alpha"-1}(A_r) = 0$.

(iii) there exists a weakly $(\oplus, \odot)$–decomposable conditional measure φ, with $\oplus$ and $\odot$ strictly increasing on Δ and Γ respectively and $\odot$ distributive over $\oplus$, locally representing $\preceq$.

Proof - We prove implication $(i) \Rightarrow (ii)$. We denote by the same symbol $\preceq$ any complete relation extending $\preceq$ in $\mathcal{A}$ and satisfying $(c1'), (c2'), (c3), (cp)$. Let E_0 be a maximal element of $\mathcal{E}$ with respect to $\preceq$ (*i.e.*, E_0 is such that there exists no $E \in \mathcal{E}$ with $E_0 \prec E$). We consider on $\mathcal{E}$ the relation $\preceq_0$, defined by putting: $E_i \preceq_0 F_i$ if $E_i \preceq F_i$ and $F_i \in \mathcal{B}(E_0)$, and $E_k \sim_0 \emptyset$ if $E_k \in \mathcal{A}(E_0)$. Since $\preceq_0$ is a total preorder and $\mathcal{A}^*$ is finite, then there exists a function φ_0 representing $\preceq_0$. By axiom (c1') it follows that φ_0 is positive for all the events in $\mathcal{B}(E_0)$, and by axiom $(c2)$ we have that φ_0 is monotone with respect to $\subseteq$. Define now $\oplus_0$ by putting, for every E, F such that $E \wedge F = \emptyset$, $\varphi_0(E \vee F) = \varphi_0(E) \oplus_0 \varphi_0(F)$. It is easy to prove that φ_0 is strictly monotone, symmetric, associative and admits 0 as neutral element in $\mathcal{K}$. Let $\mathcal{E}_1 = \mathcal{A}(E_0)$ and $\mathcal{A}_1$ the relevant set of atoms (and let $\mathcal{A}_1^*$ be the algebra spanned by it). Denote by E_1 a maximal element of $\mathcal{E}_1$, and define in $\mathcal{E}_1$ the relation $\preceq_1$ by putting:

$E_i \preceq_1 F_i$ if $E_i \preceq F_i$ and $F_i \in \mathcal{B}(E_1)$
$E_k \sim_1 \emptyset$ if $E_k \in \mathcal{A}(E_1)$.

By the same considerations made for $\preceq_0$, we obtain a capacity $\varphi_1 : \mathcal{A}_1^* \to [0,1]$ representing $\preceq_1$ and $\oplus_1$. In a finite number n of steps we get that $\mathcal{E}_n$ contains only the impossible event $\emptyset$.

We prove now the implication $(ii) \Rightarrow (iii)$. Define $\oplus$ by putting, for every $E_i, F_i \in \mathcal{K}$, with $E_i \preceq F_i$,

$$E_i \oplus F_i = E_i \oplus_\alpha F_i\,,$$

where α is the index such that $\varphi_\alpha(F_i) > 0$. Since $\{\mathcal{B}(E) : E \in \mathcal{E}\}$ is a partition and $\oplus_\alpha$ has 0 as neutral element for every α, the operation $\oplus$ is well defined. Consider now an arbitrary operation $\odot$ defined on $\{\varphi(E|H) : E, H \in \mathcal{A}^*, H \neq \emptyset\}$, commutative, distributive over $\oplus$ and whose restriction to Δ is associative, strictly monotone and admits 1 as neutral element, and put, for every $E, F \in \mathcal{A}^*$,

$$\varphi(E|F) = \varphi_\alpha(E \wedge F) \odot \varphi_\alpha(F)\,,$$

where φ_α is the relevant $\oplus$–decomposable measure, with $\varphi_\alpha(F) > 0$. By Theorem 2, φ is a weakly $(\oplus, \odot)$–decomposable conditional measure, and, since $\odot$ is strictly monotone, we have, for every $E_i \preceq F_i$,

$$\varphi(E_i|E_i \vee F_i) \leq \varphi(F_i|E_i \vee F_i)$$

and similarly for the strict inequalities.

The proof of implication $(iii) \Rightarrow (i)$ is straightforward.

Remark - We notice that Example 2 of [12] shows that it is impossible to prove that $\oplus$ is extensible to a symmetric, strictly increasing and associative operation on the whole $\varphi(\mathcal{E}) \times \varphi(\mathcal{E})$ (and so neither on $[0,1]^2$). This leads to some comments concerning a well-known result by Fine ([17], Chapter II, Theorem 4): in fact, if, in particular, (p) holds, then (in a finite setting) the assumptions of the latter theorem and those of our Theorem 3 coincide, but Fine's theorem asserts that $\oplus$ is commutative, associative and strictly increasing on $[0,1]^2$, even if his proof (given in the Appendix of [17]) actually shows that $\oplus$ is commutative, associative and strictly increasing *only for pairs belonging to $\mathcal{K}$.*

In [12] we mentioned that also J. Halpern has noticed a flaw in Fine's theorem.

6 Local coherence

We give now a condition of *local coherence* (already introduced in [6] and generalizing the coherence condition given in [2]), which is necessary and sufficient for the existence of a coherent probability representing a not necessarily complete comparative probability. Condition (lc) is stated in terms of sums of indicator functions, then it is essentially an algebraic condition.

Definition 7 - *We say that a binary relation $\preceq$, defined on a set of events $\mathcal{E}$ is a locally coherent comparative probability if it satisfies the following condition*

(lc) *for every* $n \in \mathbf{N}$, $E_i, F_i \in \mathcal{E}$, $c_i > 0$,

if $E_i \preceq E_i$ *and* $\sup \sum_i c_i(I_{F_i} - I_{E_i}) \leq 0$,

then either of the following conditions hold

(a) $E_i \sim F_i$, *for every* i,
(b) *if* $E_i \prec F_i$ *for some* i, *then there exists* $j \neq i$, *with* $j \in \{1, \dots, n\}$, *such that* $F_i \in \mathcal{A}(F_j)$.

It is possible to give an interpretation of (lc) in terms of coherent bets. In fact we may regard $c_i(I_{F_i} - I_{E_i})$ as an exchange between a bookie and a gambler, which yields an amount c_i to the bookie if F_i happens, and the same amount c_i to the gambler if E_i happens. This is betting even money on F_i versus E_i. Suppose to have this rule: if $E_i \preceq F_i$ for $i = 1, \dots n$, the bookie should accept any combination of bets, with $c_i > 0$, on F_i versus E_i. The relation $\preceq$ is incoherent if there exists one of these combinations, with a surely not positive gain and at least a pair of events $E_i \prec F_i$ not infinitely less probable than some other.

The following Proposition, whose proof is direct, studies connections between "local coherence" and "weak local coherence".

Proposition - *Let* $\preceq$ *be a comparative probability defined in an arbitrary family of events* $\mathcal{E}$. *If* $\preceq$ *is not trivial and* (lc) *holds, then* $\preceq$ *satisfies* $(c1), (c2), (c3)$ *and* (cp).

Notice that condition (lc) does not imply axiom $(c1')$, but only a weak form of it.

7 Extending locally coherent comparative probability

The following theorem deals with extensibility of a locally coherent assessment. The proof follows the line of similar theorems in [2] and is inspired by classic de Finetti–extension theorem for coherent probabilities.

Theorem 4 - *Let* $\mathcal{E}$ *be a family of events and* $\preceq$ *a corresponding comparative assessment; then there exists a (possibly not unique) total locally coherent extension* $\preceq^*$ *of* $\preceq$ *to an arbitrary family* $\mathcal{G}$ *of events, with* $\mathcal{G} \supseteq \mathcal{E}$, *if and only if* $\preceq$ *is locally coherent on* $\mathcal{E}$. *In particular, if* $\mathcal{G} = \mathcal{E} \vee \{G\}$, *there exists a unique suitable partition of* $\mathcal{E}$ *in families* $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4$ *such that locally coherent extensions of* $\preceq$ *are all relations with* $G \prec E$ *for every* $E \in \mathcal{E}_1$, $G \sim F$ *for every* $F \in \mathcal{E}_2$, $H \prec G$ *for every* $H \in \mathcal{E}_3$.

Proof - For the sake of brevity, we give only the proof of the second part (from which the first one follows in the usual way). Let us consider the following sets: $\mathcal{E}_1$ is the set of events E of $\mathcal{E}$ such that there exist $\alpha_i > 0$, $E_i \preceq F_i$ $(i = 1, \dots m)$, $E_1 \prec F_1$, and $F_i \in \mathcal{B}(\bigvee_2^m F_i)$ such that

$$\sup \sum_i [\alpha_i(I_{F_i} - I_{E_i}) + (I_E - I_G)] \leq 0\,;$$

$\mathcal{E}_2$ is the set of events $F \in \mathcal{E}$ such that there exist $\beta_j > 0$, $E_j \preceq F_j$ $(j = 1, \dots r)$, $E_1 \prec F_1$, and $F_1, F \in \mathcal{B}(\bigvee_2^r F_j)$ such that

$$\sup \sum_j [\beta_j(I_{F_j} - I_{E_j}) + (I_G - I_F)] \leq 0\,;$$

$\mathcal{E}_3$ is the set of events $H \in \mathcal{E}$ such that there exist $\delta_k > 0$, $E_k \sim F_k$ $(k = 1, \dots s)$ and $H \in \mathcal{B}(\bigvee_1^s F_k)$ such that

$$\sup \sum_k [\delta_k(I_{F_k} - I_{E_k}) + (I_H - I_G)] \leq 0\,.$$

Notice that if we have the latter inequality, we have also a similar inequality with $(I_G - I_H)$ in place of $(I_H - I_G)$. We prove now that for every $E \in \mathcal{E}_1$, $F \in \mathcal{E}_2$, $H \in \mathcal{E}_3$, we have $F \prec H \prec E$. Suppose $H \prec F$. Then by definition of $\mathcal{E}_2$ and $\mathcal{E}_3$ we have

$$\sup[\sum_k \delta_k(I_{F_k} - I_{E_k}) + \sum_j \beta_j(I_{F_j} - I_{E_j}) + (I_H - I_F)] \leq 0$$

contrary to the hypothesis of local coherence of $\preceq$. The proof of the second inequality is similar. Now we can proceed to assign the following relations: for every $E \in \mathcal{E}_1$, put $G \prec E$; for every $F \in \mathcal{E}_2$, put $F \sim G$; for every $H \in \mathcal{E}_3$, put $H \prec G$. Moreover, for every $K \in \mathcal{E} \setminus (\mathcal{E}_1 \vee \mathcal{E}_2 \vee \mathcal{E}_3)$ any relation is locally coherent.

We give now a theorem characterizing local coherence.

Theorem 5 - *Let* $\mathcal{E}$ *be a finite family of events containing* $\emptyset$, $\mathcal{E}^* = \{E|E \vee F \ : \ E \preceq F$ or $F \preceq E\}$, *and* $\preceq$ *a comparative probability in* $\mathcal{E}$. *The following statements are equivalent:*
(i) $\preceq$ *satisfies* $(c1')$ *and* (lc);
(ii) there exists a coherent (conditional) probability $P : \mathcal{E}^* \to [0, 1]$ *locally representing* $\preceq$.

Proof - In the proof of the implication $(i) \Rightarrow (ii)$ we actually build a coherent conditional P , locally representing $\preceq$, by a suitable set of probability distributions on relevant families of atoms satisfying condition (ii) of Theorem 1. Let $\preceq_0$ be as in Theorem 3. We first note that the unique locally coherent extension of $\preceq$ (and so of $\preceq_0$) to the pairs $E_1 \wedge E_k^c$, $E_1 \vee E_k$ is $E_1 \wedge E_k^c \sim E_1 \vee E_k$ (and so $E_1 \wedge E_k^c \sim_0 E_1 \vee E_k$). Let $\mathcal{A}$ the set of atoms generated by $\mathcal{E}$ and consider the following linear system S_0, where the unknown is the m-vector $W_0 = (w_1^0, \dots, w_m^0)$ (m is the cardinality of the set $\mathcal{A}_0 = \mathcal{A}$) and I_G denotes the indicator vector

$$(S_0) \begin{cases} (I_{F_i} - I_{E_i})W_0 > 0 & \text{if } E_i \prec_0 F_i \\ (I_{F_j} - I_{E_j})W_0 \geq 0 & \text{if } E_j \preceq_0 F_j \\ I_{E_k} W_0 = 0 & \text{if } E_k \sim_0 \emptyset \\ W_0 \geq 0\,. \end{cases}$$

Such a system is equivalent to the following

$$(S'_0)\begin{cases}(I_{F_i}-I_{E_i})W_0>0 & \text{if } E_i\prec_0 F_i\\ I_{(F_j}-I_{E_j})W_0\geq 0 & \text{if } E_j\preceq_0 F_j\\ (I_{E_1\wedge E_k^c}-I_{E_1\vee E_k})W_0=0 & \text{if } E_k\sim_0\emptyset\\ W_0\geq 0\,. \end{cases}$$

By using a well known theorem of alternative (see, for instance [16]), it is easy to prove that S'_0 has a solution if (and only if) the following system T'_0 (where $E'_k=E_1\wedge E^c_k$ and $E_k"=E_1\vee E_k$) has no solution (we put $G_i=I_{F_i}-I_{E_i}$)

$$(T'_0)\begin{cases}\sum\lambda_iG_i+\sum\mu_jG_j+\sum\xi_k(E'_k-E_k")\leq 0\\ \lambda_i,\mu_j,\xi_k\geq 0,\ \sum\lambda_i>0\,.\end{cases}$$

It is easy to see that (T'_0) has a solution if (and only if) $\preceq$ does not satisfy condition (lc). The function $P_0:\mathcal{A}_0\to[0,1]$ defined by putting, for any $A_k\in\mathcal{A}_0$,

$$P_0(A_k)=\frac{w^0_k}{\sum_1^n w^0_i}$$

is a probability distribution, and its (coherent) extension on the events E_i represents $\preceq_0$, since it satisfies system S_0. Let $\preceq_1$ be as in Theorem 3. Going on by the same procedure as that followed for $\preceq_0$, we obtain a probability distribution $P_1:\mathcal{A}_1\to[0,1]$, whose extension represents $\preceq_1$; and so on. After a finite number n of steps, we get that $\mathcal{E}_n$ contains only the impossible event $\emptyset$. Now, for every $E|E\vee F\in\mathcal{E}^*$, we define

$$P(E|E\vee F)=\frac{P_k(E)}{P_k(E\vee F)}\,,$$

where k is such that $E\vee F\in\mathcal{B}(E_k)$. By Theorem 2 it follows that P is a coherent conditional probability. It is immediate to prove that P locally represents $\preceq$. We prove now the implication $(ii)\Rightarrow(i)$. By Theorem 2 we have that there exist $P_0,\ldots,P_k$ such that $P_0=P(\cdot|\bigvee_{E\in\mathcal{E}}E)$ represents the restriction of $\preceq$ to the set $\mathcal{C}_0$ of the pairs $E_i\preceq F_i$ with $P_0(F_i)>0$, and give the same probability $(P_0(F_j)=P_0(E_j)=0)$ to the pairs $E_j\preceq F_j$ with $P_0(F_j)=0$. So P_0 is solution of system S_0. On the other hand, $P_1(\cdot)=P(\cdot|\bigvee_{E:P_0(E)=0}E)$ represents the restriction of $\preceq$ to the relevant set $\mathcal{C}_1$ and is zero elsewhere, therefore is the solution of system (S_1), and so on. The conclusion follows, since by the theorem of alternative all the systems $S_0,\ldots,S_k$ can have solution if (and only if) $\preceq$ satisfy (lc).
To prove that $\preceq$ satisfy $(c1')$ it is sufficient to recall that $P(\emptyset|E)=0$ and $P(E|E)=1$ for every E.

8 Strong local coherence

For infinite set of events, condition (lc) is not sufficient to assure that there exists a coherent conditional probability P representing $\preceq$. We introduce now a condition of *strong local coherence* (slc)

Definition 8 - *A binary relation $\preceq$, defined on a set of events $\mathcal{E}$, is a* strongly locally coherent *comparative probability if it satisfies the following condition*

(slc) *for every $E_i\preceq F_i$ there exists $\delta_i\geq 0$, with $\delta_i>0$ for $E_i\prec F_i$, such that for every $n\in\mathbf{N}$, $E_i,F_i\in\mathcal{E}$, $c_i>0$, one has that*

$E_i\preceq F_i$ *and* $\sup\sum_i c_i(I_{F_i}-I_{E_i}-\delta_iI_{E_i\vee F_i})\leq 0$

imply either of the following conditions

(a) $E_i\sim F_i$, for every i,
(b) if $E_i\prec F_i$ for some i, then there exists $j\neq i$, with $j\in\{1,\ldots,n\}$, such that $F_i\in\mathcal{A}(F_j)$.

It is possible to give an interpretation of (slc) (in terms of coherent bets) similar to that for condition (lc), by regarding δ_i as a penalty that one must pay to bet on a more probable event.

It is immediate to prove the following result.

Proposition - Let $\preceq$ be a comparative probability defined in an arbitrary family of events $\mathcal{E}$. If $\preceq$ satisfies (slc), then $\preceq$ satisfies (lc)

A result very similar to that of Theorem 4 can be proved in the case of strong local coherence.

Local strong coherence characterizes comparative probabilities locally representable by coherent conditional probability on an arbitrary set of events:

Theorem 6 - *Let $\mathcal{E}$ be an arbitrary family of events containing $\emptyset$, $\mathcal{A}$ the set of relevant atoms, and $\preceq$ a comparative probability in $\mathcal{E}$. The following statements are equivalent:*
(i) $\preceq$ satisfies $(c1')$ and (slc);
(ii) there exists a coherent conditional probability $P:\mathcal{E}^*\to[0,1]$ locally representing $\preceq$.

The theorem can be proved by using a compactification theorem and the following

Lemma - *Let $\mathcal{E}$ be an arbitrary family of events containing $\emptyset$ and $\preceq$ a comparative probability in $\mathcal{E}$. The following statements are equivalent:*
(i) $\preceq$ satisfies $(c1')$ and (slc);
(ii) for every finite set $\mathcal{F}$ contained in $\mathcal{E}$, there exists a coherent conditional probability $P_{\mathcal{F}}$ defined on $\mathcal{F}^*=\{E|E\vee F\,:\,E,F\in\mathcal{F},\,E\preceq F \text{ or } F\preceq E\}$, locally representing the restriction of $\preceq$ to $\mathcal{F}$ and such that, for every $E_i,F_i\in\mathcal{F}$ with $E_i\prec F_i$, we have $P_{\mathcal{F}}(F_i)-P_{\mathcal{F}}(E_i)\geq\delta_i$.

References

[1] A. Capotorti, G. Coletti, and B. Vantaggi. Non Additive Ordinal Relations Representable by Lower or Upper Probabilities. *Kybernetika*, 34 (1):79–90, 1998.

[2] G. Coletti. Coherent Qualitative Probability. *Jour. Math. Psych.*, 34:297-310, 1990.

[3] G. Coletti. A weak coherence condition for conditional comparative probabilities. In *Uncertainty in Intelligent Systems* (Bouchon-Meunier et al. eds.), Elsevier Sci. Publ.B.V.:195-201, 1993.

[4] G. Coletti. Comparative probabilities ruled by coherence conditions and its use in expert systems. *Intern. Jour. of General Systems*, 22:93-101, 1993.

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

[6] G. Coletti. Non additive ordinal relations representable by conditional probabilities and its use in expert systems, *Proc. of IPMU'96*, Granada:43-48, 1996.

[7] G. Coletti, A. Gilio, and R. Scozzafava. Coherent qualitative probability and uncertainty in Artificial Intelligence. *Proceed. 8th Int. Conf. on Cybernetics and Systems*, vol.1 (C.N.Manikopoulos ed.), NJIT Press, New York:132-138,1990.

[8] G. Coletti, A. Gilio, and R. Scozzafava. Assessment of qualitative judgements for conditional events in expert systems. In "Symbolic and qualitative approaches to uncertainty" (R.Kruse and P.Siegel Eds.). *Lecture Notes in Comp. Science*, 548:135–140, 1991.

[9] G. Coletti and G. Regoli. Probabilitá qualitative non archimedee e realizzabilitá. *Riv. Mat. Sci. Econom. Soc.*, 6:79-99, 1983.

[10] G. Coletti and R. Scozzafava. Characterization of coherent conditional probabilities as a tool for their assessment and extension. *Jour. of Uncertainty, Fuzziness and Knowledge-based Systems*, 4(3):103–127, 1996.

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

[12] G. Coletti and R. Scozzafava. From conditional events to conditional measures: a new axiomatic approach. *Annals of Mathematics and Artificial Intelligence*, Special Issue on "Representations of Uncertainty", to appear.

[13] B. de Finetti. Sul significato soggettivo della probabilità. *Fundamenta Matematicae*, 17:293–329, 1931.

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

[15] L.E. Dubins. Finitely additive conditional probability, conglomerability and disintegrations. *Ann. Probab.*, 3:89-99, 1975.

[16] W. Fenchel. Convex cones, sets and functions. *Lectures at Princeton University*, Spring term, 1951.

[17] T.L. Fine. *Theory of Probability.* Academic Press, London, 1973.

[18] A. Gilio and R. Scozzafava. Coerenza di probabilità condizionate realizzate mediante pseudodensità. *Metron*, 47:65-76, 1989.

[19] B.O. Koopman. The axioms and algebra of intuitive probability. *Ann. Math.*, 41:269-292, 1940.

[20] C.H. Kraft, J. Pratt, and A. Seidenberg. Intuitive probability in finite sets. *Ann. Mat. Stat.*, 30:408-419, 1959.

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

[22] A. Rényi. On conditional probability spaces generated by a dimensionally ordered set of measures. *Theor. Probab. Appl.*, 1:61–71, 1956.

[23] R. Scozzafava. A survey of some common misunderstandings concerning the role and meaning of finitely additive probabilities in statistical inference. *Statistica*, 44:21-45, 1984.

[24] S.K.M. Wong, Y.Y. Yao, P. Bollman, and H.C. Bürger. Axiomatization of qualitative belief structures. *IEEE Transactions on Systems, Man and Cybernetics*, 21 (4):726-734, 1991.