Conditional Random Quantities and Iterated Conditioning in the Setting of Coherence

Angelo Gilio[1] and Giuseppe Sanfilippo[2]

[1] Dipartimento di Scienze di Base e Applicate per l'Ingegneria,
University of Rome "La Sapienza", Italy
angelo.gilio@sbai.uniroma1.it

[2] Dipartimento di Matematica e Informatica, University of Palermo, Italy
giuseppe.sanfilippo@unipa.it

Abstract. We consider conditional random quantities (c.r.q.'s) in the setting of coherence. Given a numerical r.q. X and a non impossible event H, based on betting scheme we represent the c.r.q. $X|H$ as the unconditional r.q. $XH + \mu H^c$, where μ is the prevision assessed for $X|H$. We develop some elements for an algebra of c.r.q.'s, by giving a condition under which two c.r.q.'s $X|H$ and $Y|K$ coincide. We show that $X|HK$ coincides with a suitable c.r.q. $Y|K$ and we apply this representation to Bayesian updating of probabilities, by also deepening some aspects of Bayes' formula. Then, we introduce a notion of iterated c.r.q. $(X|H)|K$, by analyzing its relationship with $X|HK$. Our notion of iterated conditional cannot formalize Bayesian updating but has an economic rationale. Finally, we define the coherence for prevision assessments on iterated c.r.q.'s and we give an illustrative example.

Keywords: Coherence, betting scheme, conditional random quantities, conditional previsions, Bayesian updating, iterated conditioning.

1 Introduction

Probabilistic reasoning under coherence allows a consistent treatment of uncertainty in many applications of statistical analysis, economy, decision theory, fuzzy set theory, psychology and artificial intelligence. This probabilistic approach allows to manage incomplete probabilistic assignments in a situation of vague or partial knowledge, see e.g. [9, 11, 13–15, 32]; see also [18, 21, 22, 24–28, 36] where a flexible probabilistic approach to inference rules in nonmonotonic reasoning and to the psychology of uncertain reasoning is developed. Based on

[1] The final publication is available at
http://link.springer.com/chapter/10.1007/978-3-642-39091-3_19
Angelo Gilio, Giuseppe Sanfilippo, "Conditional Random Quantities and Iterated Conditioning in the Setting of Coherence", Chapter in Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Lecture Notes in Computer Science, Springer Berlin Heidelberg, vol. 7958, pp. 218-229, 2013.

coherence, we can develop a numerical approach to conditional events consistent with the three-valued logic proposed in the pioneering paper [16] by de Finetti; in this work we extend the approach to conditional random quantities (c.r.q.'s). Based on the betting scheme ([17], see also [31]), if for a numerical r.q. X we evaluate μ its prevision $\mathbb{P}(X)$, then we agree to pay (resp., to receive) an amount μ and to receive (resp., to pay) the random amount X. Analogously, given any non impossible event H, if we assess $\mathbb{P}(X|H) = \mu$ for the prevision of X *conditional on* H, then we agree to pay (resp., to receive) μ and to receive (resp., to pay) an amount, denoted $X|H$, which coincides with X, or μ, according to whether H is true, or false, i.e. $H = 1$, or $H = 0$ (in terms of indicators); then, *operatively*, $X|H = XH + \mu(1 - H)$. Thus, one of the values of $X|H$ is the prevision $\mathbb{P}(X|H) = \mu$, which is *subjectively evaluated*. In particular, if for a conditional event $A|H$ we assess $P(A|H) = p$, then (the indicator of) $A|H$ is the r.q. $AH + p(1 - H)$, with set of possible values $\{1, 0, p\}$. The problem of suitably defining the third value for (the indicators of) conditional events has been studied in some works by Coletti and Scozzafava (see, e.g., [13]).
We point out that, differently from other authors (see, e.g., ([37]; see also [34]), in our approach a c.r.q. $X|H$ is explicitly managed as an 'unconditional object' which among its possible values admits (the conditional prevision) μ. We also observe that the generalization of our results to imprecise conditional prevision assessments is out of the scope of the paper.
By exploiting this representation of c.r.q.'s, we obtain some basic results which concern an algebra of c.r.q.'s. Among other things, given any events H, K and any r.q.'s X, Y, we examine the condition under which $X|H$ and $Y|K$ coincide; in particular, we show that $X|HK$ can be represented as a suitable c.r.q. $Y|K$. Then, we use this representation in the context of Bayesian updating of probabilities and we deepen some aspects of Bayes' formula in the setting of coherence. As a natural consequence, we introduce the iterated c.r.q. $(X|H)|K$, which is defined as a suitable c.r.q. $Y|K$; then, we analyze its relationship with $X|HK$. However, the Bayesian updating for the probability of any hypothesis H cannot be formalized by our notion of iterated conditioning. Finally, we define the coherence for prevision assessments on iterated c.r.q.'s and we illustrate this notion by an example.

2 Preliminary Notions and Results

In our approach an event A represents an uncertain fact described by a (non ambiguous) logical proposition; hence we look at A as a two-valued logical entity which can be true (T), or false (F). The indicator of A, denoted by the same symbol, is a two-valued numerical quantity which is 1, or 0, according to whether A is true, or false. The sure event is denoted by Ω and the impossible event is denoted by $\emptyset$. Moreover, we denote by $A \wedge B$ (resp., $A \vee B$) the logical conjunction (resp., logical disjunction). In many cases we simply denote the conjunction between A and B as the product AB. By the symbol A^c we denote the negation of A. Given any events A and B, we simply write $A \subseteq B$ to denote

that A logically implies B, i.e. $AB^c = \emptyset$. We recall that n events are logically independent when the number of atoms, or constituents, generated by them is 2^n. In case of some logical dependencies among the events, the number of atoms is less than 2^n. Given any events A and B, with $A \neq \emptyset$, the conditional event $B|A$ is looked at as a three-valued logical entity which is true (T), or false (F), or void (V), according to whether AB is true, or AB^c is true, or A^c is true.
Given an event $H \neq \emptyset$ and a r.q. X, we denote by V_H, the set of possible values of X restricted to H and, if X is finite, we set $V_H = \{x_1, x_2, \dots, x_r\}$. In the setting of coherence, agreeing to the betting metaphor the prevision of $''X$ *conditional on* H'' (also named $''X$ *given* H''), $\mathbb{P}(X|H)$, is defined as the amount μ you agree to pay (resp., to receive), by knowing that you will receive (resp., to pay) the amount X if H is true, or you will receive back (resp., to pay back) the amount μ if H is false (bet called off). Agreeing with the operational subjective approach given in [31], we denote by $X|H$ the amount that you receive when a conditional bet is stipulated on $''X$ *given* H''. Then, it holds that $X|H = XH + \mu H^c$, where $\mu = \mathbb{P}(X|H)$, so that *operatively* we can look at the c.r.q. $X|H$ as the *unconditional* r.q. $XH + \mu H^c$. If X is finite and $\mu \notin V_H$, then $X|H \in \{x_1, x_2, \dots, x_r, \mu\}$. Moreover, denoting by A_i the event $(X = x_i)$, $i \in J_r$, the family $\{A_1H, \dots, A_rH, H^c\}$ is a partition of Ω and we have $X|H = XH + \mu H^c = x_1A_1H + \cdots + x_rA_rH + \mu H^c$. In particular, when X is an event A, the prevision of $X|H$ is the probability of $A|H$ and, if you assess $P(A|H) = p$, then for the indicator of $A|H$, denoted by the same symbol, we have $A|H = AH + pH^c \in \{1, 0, p\}$. The choice of p as the third value of $A|H$ has been proposed in some previous works, see e.g. [13, 19, 31].

Coherence for Conditional Prevision Assessments.
Given a prevision function $\mathbb{P}$ defined on an arbitrary family $\mathcal{K}$ of c.r.q.'s, let $\mathcal{F}_n = \{X_i|H_i,\, i \in J_n\}$ be any finite subfamily of $\mathcal{K}$; we set $\mathcal{M}_n = (\mu_i,\, i \in J_n)$, where $\mu_i = \mathbb{P}(X_i|H_i)$. With the pair $(\mathcal{F}_n, \mathcal{M}_n)$ we associate the random gain $\mathcal{G} = \sum_{i \in J_n} s_i H_i (X_i - \mu_i)$; moreover, we set $\mathcal{H}_n = H_1 \vee \cdots \vee H_n$ and we denote by $G_{\mathcal{H}_n}$ the set of values of $\mathcal{G}$ restricted to the disjunction $\mathcal{H}_n$ of the conditioning events $H_1, \dots, H_n$. Then, by de Finetti's *betting scheme*, we have

Definition 1. The function $\mathbb{P}$ defined on a finite family $\mathcal{K}$ is coherent if and only if, $\forall n \geq 1$, $\forall \mathcal{F}_n \subseteq \mathcal{K}$, $\forall s_1, \dots, s_n \in \mathbb{R}$, it holds that: $\inf G_{\mathcal{H}_n} \leq 0 \leq \sup G_{\mathcal{H}_n}$. When $\mathcal{K}$ is infinite, we say that $\mathbb{P}$ is coherent if its restriction $\mathcal{M}_n$ on $\mathcal{F}_n$ is coherent, for every $\mathcal{F}_n \subset \mathcal{K}$.

Remark 1. Given a finite c.r.q. $X|H$, with $\mathbb{P}(X|H) = \mu$ and $V_H = \{x_1, \dots, x_r\}$, we have that μ is coherent if and only if $min\, V_H \leq \mu \leq max\, V_H$. In particular, if $V_H = \{c\}$, then $X|H = cH + \mu H^c$; in this case μ is coherent if and only if $\mu = c$. Of course, for $X = H$ (resp. $X = H^c$) it holds that $\mu = 1$ (resp. $\mu = 0$) and hence $H|H = 1$, $H^c|H = 0$.

Checking of Coherence for Conditional Prevision Assessments.
Given a family of n finite c.r.q.'s $\mathcal{F}_n = \{X_1|H_1, \dots, X_n|H_n\}$, for each $i \in J_n$ we denote by $\{x_{i1}, \dots, x_{ir_i}\}$ the set of possible values for the restriction of X_i to

H_i; then, for each $i \in J_n$ and $j = 1, \ldots, r_i$, we set $A_{ij} = (X_i = x_{ij})$. Of course, for each $i \in J_n$, the family $\{H_i^c, A_{ij}H_i\,,\ j = 1, \ldots, r_i\}$ is a partition of the sure event Ω. Then, the constituents generated by the family $\mathcal{F}_n$ are (the elements of the partition of Ω) obtained by expanding the expression $\bigwedge_{i\in J_n}(A_{i1}H_i \vee \cdots \vee A_{ir_i}H_i \vee H_i^c)$. We set $C_0 = H_1^c \cdots H_n^c$ (it may be $C_0 = \emptyset$); moreover, we denote by $C_1, \ldots, C_m$ the constituents contained in $\mathcal{H}_n = H_1 \vee \cdots \vee H_n$. Hence $\bigwedge_{i\in J_n}(A_{i1}H_i \vee \cdots \vee A_{ir_i}H_i \vee H_i^c) = \bigvee_{h=0}^m C_h$. With each C_h, $h \in J_m$, we associate a vector $Q_h = (q_{h1}, \ldots, q_{hn})$, where

$$q_{hi} = x_{ij},\ \text{if}\ C_h \subseteq A_{ij}H_i\,,\ j = 1, \ldots, r_i;\ q_{hi} = \mu_i,\ \text{if}\ C_h \subseteq H_i^c\,.$$

In more explicit terms, for each $j \in \{1, \ldots, r_i\}$ the condition $C_h \subseteq A_{ij}H_i$ amounts to $C_h \subseteq A_{i1}^c \cdots A_{i,j-1}^c A_{ij} A_{i,j+1}^c \cdots A_{ir}^c A_{ir_i}^c H_i$. We observe that the vector $Q_h = (q_{h1}, \ldots, q_{hn})$ is the value of the random vector $(X_1|H_1, \ldots, X_n|H_n)$ when C_h is true; moreover, if C_0 is true, then the value of such a random vector is $\mathcal{M}_n = (\mu_1, \ldots, \mu_n)$. Denoting by $\mathcal{I}_n$ the convex hull of $Q_1, \ldots, Q_m$, the condition $\mathcal{M}_n \in \mathcal{I}_n$ amounts to the existence of a vector $(\lambda_1, \ldots, \lambda_m)$ such that: $\sum_{h\in J_m} \lambda_h Q_h = \mathcal{M}_n\,,\ \sum_{h\in J_m} \lambda_h = 1\,,\ \lambda_h \geq 0\,,\ \forall h$; in other words, $\mathcal{M}_n \in \mathcal{I}_n$ is equivalent to solvability of the following system Σ associated with the pair $(\mathcal{F}_n, \mathcal{M}_n)$, in the nonnegative unknowns $\lambda_1, \ldots, \lambda_m$,

$$\Sigma: \quad \sum_{h\in J_m} \lambda_h q_{hi} = \mu_i\,,\ i \in J_n\,;\ \sum_{h\in J_m} \lambda_h = 1\,;\ \lambda_h \geq 0\,,\ h \in J_m\,. \quad (1)$$

Given a subset $J \subseteq J_n$, we set $\mathcal{F}_J = \{X_i|H_i\,,\ i \in J\}\,,\ \ \mathcal{M}_J = (\mu_i\,,\ i \in J)\,$; then, we denote by Σ_J, where $\Sigma_{J_n} = \Sigma$, the system like (1) associated with the pair $(\mathcal{F}_J, \mathcal{M}_J)$. Then, it can be proved the following ([7])

Theorem 1. [*Characterization of coherence*]. Given a family of n finite c.r.q.'s $\mathcal{F}_n = \{X_1|H_1, \ldots, X_n|H_n\}$ and a vector $\mathcal{M}_n = (\mu_1, \ldots, \mu_n)$, the conditional prevision assessment $\mathbb{P}(X_1|H_1) = \mu_1, \ldots, \mathbb{P}(X_n|H_n) = \mu_n$ is coherent if and only if, for every subset $J \subseteq J_n$, defining $\mathcal{F}_J = \{X_i|H_i\,,\ i \in J\}$, $\mathcal{M}_J = (\mu_i\,,\ i \in J)$, the system Σ_J associated with the pair $(\mathcal{F}_J, \mathcal{M}_J)$ is solvable.

A characterization of coherence of conditional prevision assessments by non dominance with respect to proper scoring rules has been given in [8].

3 Deepenings on Conditional Random Quantities and Bayes Theorem

In this section, by exploiting the representation $X|H = XH + \mu H^c$, where $\mu = \mathbb{P}(X|H)$, we develop some elements of *an algebra of c.r.q.'s.* In particular, we recall a result which also concerns the general compound prevision theorem; then, we give some comments on the Bayesian updating of probabilities. We have

Theorem 2. Given any real quantities $a_1, \ldots, a_n$, any event $H \neq \emptyset$, any random quantities $X_1, \ldots, X_n$ and any coherent assessment $(\mu_1, \ldots, \mu_n, \nu)$ on $\{X_1|H, \ldots, X_n|H, (\sum_{i=1}^n a_i X_i)|H\}$, we have: $\sum_{i=1}^n a_i(X_i|H) = (\sum_{i=1}^n a_i X_i)|H$.

Proof. We have $(\sum_{i=1}^n a_i X_i)|H = (\sum_{i=1}^n a_i X_i)H + \nu H^c$; moreover, it holds that $\mathbb{P}[(\sum_{i=1}^n a_i X_i)|H] = \sum_{i=1}^n a_i \mathbb{P}(X_i|H)$; that is $\nu = \sum_{i=1}^n a_i \mu_i$. Then

$$\sum_{i=1}^n a_i(X_i|H) = \sum_{i=1}^n a_i(X_iH + \mu_i H^c) = \big(\sum_{i=1}^n a_i X_i\big)H + \nu H^c = \big(\sum_{i=1}^n a_i X_i\big)|H\,.$$

In particular: $a(X|H) = (aX)|H = aX|H$.

Theorem 3. Given any c.r.q.'s $X_1|H_1, \ldots, X_n|H_n$, with $\mathbb{P}(X_i|H_i) = \mu_i, \forall\, i$, and with $(\mu_1, \ldots, \mu_n)$ coherent, we have: $\mathbb{P}(\sum_{i=1}^n X_i|H_i) = \sum_{i=1}^n \mathbb{P}(X_i|H_i)$.

Proof. By linearity of prevision, we have

$$\mathbb{P}\Big(\sum_{i=1}^n X_i|H_i\Big) = \mathbb{P}\Big[\sum_{i=1}^n (X_iH_i + \mu_i H_i^c)\Big] = \sum_{i=1}^n \mathbb{P}\Big(X_iH_i + \mu_i H_i^c\Big) = \sum_{i=1}^n \mathbb{P}(X_i|H_i)\,.$$

We now consider the following questions:
(a) given two r.q. $X|H, Y|K$*, with* $H \neq K$*, may it happen that* $X|H = Y|K$ *?*
(b) given any events H, K*, with* $HK \neq \emptyset$*, and any r.q.* X*, with* $\mathbb{P}(X|HK) = \mu$*, does there exist a r.q.* Y *such that* $X|HK = Y|K$ *?*
We recall two results ([23, Theorems 7 and 9]) which show that the answers to both questions are positive. Concerning question (a) we have

Theorem 4. Given two c.r.q.'s $X|H, Y|K$, let (μ, ν) be a coherent prevision assessment on $\{X|H, Y|K\}$, with $\mathbb{P}(X|H) = \mu$, $\mathbb{P}(Y|K) = \nu$. Moreover, assume that $X|H = Y|K$ when the disjunction $H \vee K$ is true. Then $X|H = Y|K$.

The answer to question (b) is given in condition (i) of the result below, where (by Theorem 4) it is shown that $X|HK = Y|K$, where $Y = XH + yH^c$ and $y = \mathbb{P}(X|HK)$. The condition (ii), i.e. the general compound prevision theorem, is directly obtained by condition (i), by exploiting the linearity of prevision.

Theorem 5. Given two events $H \neq \emptyset, K \neq \emptyset$ and a r.q. X, let (x, y, z) be a coherent prevision assessment on $\{H|K, X|HK, XH|K\}$. Then:
(i) $X|HK = (XH + yH^c)|K$;
(ii) $z = xy$; that is: $\mathbb{P}(XH|K) = P(H|K)\mathbb{P}(X|HK)$.

In the next subsection, condition (i) of Theorem 5 will be applied to Bayesian updating of conditional probabilities.

3.1 An Application to Bayesian Inference

Given a hypothesis H, with $P(H) = p_0$, and a sequence of evidences $E_1, \ldots, E_n$, we set: $E_1 \cdots E_k = A_k\,,\;\; P(H|A_k) = p_k\,,\;\; Y_k = HA_{k-1} + p_k A_{k-1}^c\,,\;\; k = 1, \ldots, n$. By applying condition (i) of Theorem 5, with X, H, K replaced respectively by H, A_{k-1}, and E_k, we obtain

$$H|E_1 \cdots E_k = H|A_{k-1}E_k = Y_k|E_k = (HA_{k-1} + p_k A_{k-1}^c)|E_k\,,\; k = 1, \ldots, n\,.$$

We can verify that, in the previous equality, the prevision on the right-hand side coincides with that one on the left-hand side, which is the probability $P(H|E_1 \cdots E_k) = p_k$. Indeed, we have

$$\mathbb{P}(Y_k|E_k) = \mathbb{P}[(HA_{k-1} + p_k A^c_{k-1})|E_k] = P(HA_{k-1}|E_k) + p_k P(A^c_{k-1}|E_k) = \\ = p_k P(A_{k-1}|E_k) + p_k P(A^c_{k-1}|E_k) = p_k\,,\ k = 1, \ldots, n\,.$$

As we can see, the updating of the probability of H, on the basis of evidences $E_1, \ldots, E_n$, consists at each step in replacing a probability by the next one in the following sequence: $P(H)\,,\ P(H|E_1)\,,\ P(H|E_1E_2)\,,\ \cdots\,,\ P(H|E_1 \cdots E_k)\,,\ \cdots$; that is, using the Bayesian mechanism, at each step we replace $P(H|A_{k-1}) = p_{k-1}$ by $P(H|A_k) = p_k$ when the new evidence E_k is obtained. Of course, in order to compute p_k by Bayes' formula

$$p_k = P(H|A_k) = P(H|A_{k-1}E_k) = \frac{P(E_k|A_{k-1}H)P(H|A_{k-1})}{P(E_k|A_{k-1})}\,,$$

all the needed probabilities must be assigned and $P(E_k|A_{k-1})$ must be positive. If $P(E_k|A_{k-1}) = 0$, by the methods of coherence, e.g. by using the Algorithm 1 in [2], or the zero-layers procedure in [13], it easily follows $p_k \in [0, 1]$. More in general, if some of the values in Bayes' formula are not specified, then p_k is not uniquely determined and for its lower and upper bounds, l, u, there are different cases considered in the next subsection.

3.2 Lower/Upper Bounds on the Probability of $H|A_k$

We assume H, A_{k-1}, E_k logically independent and we set $A_{k-1} = A, E_k = E$; then $\{H|A_{k-1}, E_k|A_{k-1}, E_k|A_{k-1}H, H|A_{k-1}E_k\} = \{H|A, E|A, E|AH, H|AE\}$.
As a preliminary remark we note that, given any sub-family $\Gamma = \{E_1|H_1, E_2|H_2\}$ of the family $\{H|A, E|A, E|AH, H|AE\}$, it can be verified that the set of coherent assessments (x, y) on Γ coincides with the unit square $[0, 1]^2$. Then, if we assign only one of the quantities $P(E|A)$, or $P(E|AH)$, or $P(H|A)$, and we want to propagate it to $H|AE$, it holds that each value $z = P(H|AE) \in [0, 1]$ is a coherent extension of the given assignment; that is $l = 0, u = 1$.
We now consider the cases where we assign only two of the quantities $P(E|A)$, $P(E|AH)$, $P(H|A)$, by giving the lower/upper bounds on $P(H|AE)$. We have three cases (which, due to the lack of space, are discussed without proof):
(i) only $x = P(E|A)$ and $y = P(E|AH)$ are assigned; then, the assessment $\mathcal{P} = (x, y, z)$ on $\mathcal{F} = \{E|A, E|AH, H|AE\}$ is coherent if and only if $z \in [0, u]$, with $u = \frac{y(1-x)}{x(1-y)}$, or $u = 1$, according to whether $y < x$, or $y \geq x$.
(ii) only $x = P(H|A)$ and $y = P(E|AH)$ are assigned; then, the assessment $\mathcal{P} = (x, y, z)$ on $\mathcal{F} = \{H|A, E|AH, H|AE\}$ is coherent if and only if $z \in [l, 1]$, with $l = 0$, or $l = \frac{xy}{1-x+xy}$, according to whether $(x, y) = (1, 0)$, or $(x, y) \neq (1, 0)$.
(iii) only $x = P(H|A)$ and $y = P(E|A)$ are assigned; then, based on the probabilistic analysis of the *CM rule* given in [20], the assessment $\mathcal{P} = (x, y, z)$ on $\mathcal{F} = \{H|A, E|A, H|AE\}$ is coherent if and only if $l \leq z \leq u$, with

$$l = \begin{cases} \frac{x+y-1}{y}\,, & \text{if } x + y > 1\,, \\ 0\,, & \text{if } x + y \leq 1\,, \end{cases} \qquad u = \begin{cases} \frac{x}{y}\,, & \text{if } x < y\,, \\ 1\,, & \text{if } x \geq y\,. \end{cases}$$

Remark 2. Given n logically independent events $E_1, \ldots, E_{n-1}, H$, and any assessment $\mathcal{P} = (x_1, \ldots, x_{n-1}, p_0)$ on the family $\mathcal{F} = \{E_1, \ldots, E_{n-1}, H\}$, the extension $p_{n-1} = P(H|E_1 \cdots E_{n-1})$ is coherent if and only if: $l \leq p_{n-1} \leq u$, where

$$l = \begin{cases} max \left\{0, \frac{x_1+\cdots+x_{n-1}+p_0-(n-1)}{x_1+\cdots+x_{n-1}-(n-2)}\right\}, & \text{if } x_1 + \cdots + x_{n-1} > n-2\,, \\ 0\,, & \text{if } x_1 + \cdots + x_{n-1} \leq n-2\,; \end{cases}$$

$$u = \begin{cases} min \left\{1, \frac{p_0}{x_1+\cdots+x_{n-1}-(n-2)}\right\}, & \text{if } x_1 + \cdots + x_{n-1} > n-2\,, \\ 1\,, & \text{if } x_1 + \cdots + x_{n-1} \leq n-2\,. \end{cases}$$

The previous formulas are obtained from (and better represent) the lower and upper bounds, l and u, given for the *generalized Cautious Monotonicity rule* in [21]; indeed, the representation of the probability bounds given in [21, Theorem 11] only concerns the case where the condition $x_1+\cdots+x_{n-1} > n-2$ is satisfied. A similar comment applies to [21, Theorem 10].

Other aspects of Bayes' theorem have been analyzed in [13] and [39]. Theoretical aspects and algorithms concerning the set of probability assessments which are compatible with given initial ones have been studied in several fields, such as probabilistic reasoning under coherence, model-theoretic probabilistic logic, probabilistic satisfiability, credal networks, and others; see, e.g., [2–6, 9, 10, 13, 29, 35, 38]. In the next section we give some results on iterated conditioning and we make a critical comparison with Bayesian updating.

4 Iterated Conditioning

The basic intuition for our notion of iterated c.r.q. follows by the representation $X|H = XH + \mu H^c$, where $\mu = \mathbb{P}(X|H)$. After the definition we briefly discuss the meaning of the 'new object' $(X|H)|K$; then we give some results.

Definition 2. Given any events H, K, with $H \neq \emptyset, K \neq \emptyset$, and a finite r.q. X, with $\mathbb{P}(X|H) = \mu$, we define $(X|H)|K = (XH + \mu H^c)|K$ [3].

From the previous definition, as $H^c|H = 0$, it follows:
$(X|H)|H = (XH + \mu H^c)|H = XH|H + \mu H^c|H = XH|H = X|H$; then, if we set $Y = X|H = XH + \mu H^c$, from $(X|H)|H = X|H$ it follows $Y|H = Y$.

Remark 3. Does there exist a reasonable justification for Definition 2 ?
We can provide a rationale for Definition 2, by imagining a decision problem involving two prevision assessments:
1) an agent evaluates $\mathbb{P}(X|H) = \mu$, by accepting then any transaction where, by paying an amount μ, one receives the uncertain amount $Y = X|H = XH+\mu H^c$;
2) the same agent evaluates $\mathbb{P}(Y|K) = \nu$, with $Y = X|H$, by accepting then a transaction where, by paying ν, one receives the uncertain amount $Y|K$.

[3] This notion of iterated conditioning for c.r.q.'s is consistent with that one given for conditional events in [26].

Then, operatively: $\nu = \mathbb{P}(Y|K) = \mathbb{P}[(XH + \mu H^c)|K] = \mathbb{P}[(X|H)|K]$; that is, to evaluate the prevision of $Y|K$ amounts to evaluate the prevision of the iterated c.r.q. $(X|H)|K$. We point out that our notion of iterated conditioning does not concern those situations, typical of Bayesian updating, where a collection of pieces of evidence is synthesized by their conjunction and managed in a coherent way. Clearly, coherence plays a basic role also in our approach; for instance, concerning the discussion above, the agent must check coherence of the assessment (μ, ν) on $\{X|H, (X|H)|K\}$. This aspect will be considered in Section 5.
In the next result we show that $(X|H)|K$ may coincide with $X|H$, or $X|K$.

Proposition 1. Given any r.q. X and any nonimpossible events H, K, we have: (i) $(X|H)|K \neq (X|K)|H$; (ii) $(X|H)|K \neq X|HK$; (iii) if $H \subseteq K$, or $K \subseteq H$, then $(X|K)|H = (X|H)|K = X|HK$.

Proof. (i) The assertion follows by Definition 2.
(ii) Defining $\mathbb{P}(X|H) = \mu, \mathbb{P}(X|HK) = \eta$, in general it holds that $\mu \neq \eta$; thus, by condition (i) of Theorem 5, we have

$$X|HK = (XH + \eta H^c)|K \;\neq\; (XH + \mu H^c)|K = (X|H)|K\,.$$

(iii.a) If $H \subseteq K$, defining $\mathbb{P}(X|H) = \mu, \mathbb{P}(X|K) = z, \mathbb{P}(X|HK) = \eta$, we have $X|HK = X|H$ and $\eta = \mu$; then (by condition (i) of Theorem 5) we obtain $(X|H)|K = (XH + \mu H^c)|K = (XH + \eta H^c)|K = X|HK = X|H$. Moreover, $H \subseteq K$ implies $XK^c|H = zK^c|H = 0$; hence $XK|H = X(K + K^c)|H = X|H$. Then $(X|K)|H = (XK + zK^c)|H = XK|H = X|H = X|HK$.
(iii.b) If $K \subseteq H$, the assertion follows by a symmetric reasoning .

Remark 4. Note that, by condition (iii) in Proposition 1, $X|H = (X|H)|(H \vee K)$; then $\mathbb{P}(X|H) = \mathbb{P}[(X|H)|(H \vee K)]$. Indeed, defining $\mathbb{P}(X|H) = \mu$, we have $\mathbb{P}[(X|H)|(H \vee K)] = \mathbb{P}[(XH + \mu H^c)|(H \vee K)] = \mathbb{P}(XH|H \vee K) + \mathbb{P}(\mu H^c|H \vee K) =$ $= \mathbb{P}(X|H)P(H|H \vee K) + \mu P(H^c|H \vee K) = \mu$.

The next result shows that the sum $X|H + Y|K$ of two c.r.q.'s, with *different conditioning events* H, K, can be represented as a suitable c.r.q. $Z|(H \vee K)$.

Proposition 2. Given a coherent prevision assessment (μ, η) on $\{X|H, Y|K\}$, it holds that: $X|H + Y|K = Z|(H \vee K)$, where $Z = XH + \mu H^c + YK + \eta K^c$ and $\mathbb{P}[Z|(H \vee K)] = \mu + \eta$.

Proof. We observe that $H \subseteq (H \vee K), K \subseteq (H \vee K)$; then, from condition (iii) in Proposition 1, we have $X|H = (X|H)|(H \vee K) = (XH + \mu H^c)|(H \vee K)$, $Y|K = (Y|K)|(H \vee K) = (YK + \eta K^c)|(H \vee K)$. Then, by Theorem 2, we obtain $X|H + Y|K = (XH + \mu H^c + YK + \eta K^c)|(H \vee K) = Z|(H \vee K)$. Moreover, by Theorem 3 (see also Remark 4), $\mathbb{P}[Z|(H \vee K)] = \mu + \eta$.

We observe that, given any events A, H, K, if $H \subseteq K$, or $K \subseteq H$, then $(A|K)|H = (A|H)|K = A|HK$, and the Import-Export Principle ([33]) would be valid. But, in general we have $(A|H)|K \neq (A|K)|H$, $(A|H)|K \neq A|HK$, $(A|K)|H \neq A|HK$; that is, in agreement with other authors (see, e.g., [1, 30]), the Import-Export Principle does not hold, as illustrated by the example below.

Example 1. Given any events A, H, K, with $HK = \emptyset$, we denote by p the probability of $A|H$, $P(A|H)$, and by α the prevision of $(A|H)|K$, $\mathbb{P}[(A|H)|K]$. By Definition 2, $(A|H)|K = (AH + pH^c)|K = AHK + pH^cK + \alpha K^c = pK + \alpha K^c$; moreover, *conditionally on* K *being true* the r.q. $AH + pH^c$ is constant and equal to p; then, by Remark 1, $\alpha = \mathbb{P}[(AH + pH^c)|K] = p$. Therefore, from $HK = \emptyset$ it follows: $(A|H)|K = p$ (more in general, given any r.q. X, with $\mathbb{P}(X|H) = \mu$, if $HK = \emptyset$, then $(X|H)|K = \mu$). If the Import-Export Principle were valid, we would have $(A|H)|K = A|HK = A|\emptyset$, which makes no sense; indeed, in Bayesian updating it is absurd to consider two logically incompatible evidences H, K.

In the framework of Bayesian inference, given any uncertain hypothesis H and any evidences $E_1, E_2, \ldots, E_n$, we iteratively compute $P(H|E_1)$, $P(H|E_1E_2)$, $\cdots$, $P(H|E_1 \cdots E_n)$; this amounts to synthesizing the sequence $E_1, \ldots, E_n$ by the conjunction $E_1 \cdots E_n$. If iterated conditioning were defined in agreement with the Import-Export Principle, it would be $H|E_1E_2 = (H|E_1)|E_2$, and so on; then

$$P(H|E_1E_2) = P[(H|E_1)|E_2]\,,\; P(H|E_1E_2E_3) = P[((H|E_1)|E_2)|E_3]\,,\; \ldots\,.$$

But, in our approach we have $(H|E_1)|E_2 \neq H|E_1E_2$, and so on; thus, Bayesian updating cannot be formalized by our iterated conditioning. For instance, we cannot look at the prevision $\mathbb{P}[(H|E_1)|E_2]$ as the probability $P(H|E_1E_2)$. Indeed, defining $P(H|E_1) = p_1, P(H|E_1E_2) = p_2$, by condition (i) of Theorem 5 we have $H|E_1E_2 = (HE_1 + p_2E_1^c)|E_2 \neq (HE_1 + p_1E_1^c)|E_2 = (H|E_1)|E_2$.
As discussed in Remark 3, our notion of iterated conditioning is useful for applications different from Bayesian updating.

5 Coherent Prevision Assessments for Iterated Conditional Random Quantities

In this section we introduce the notion of coherent prevision assessments on iterated c.r.q.'s, like

$$\mathbb{P}[(X_1|H_1)|K_1] = \nu_1\,,\; \mathbb{P}[(X_2|H_2)|K_2] = \nu_2\,,\; \cdots\,,\; \mathbb{P}[(X_n|H_n)|K_n] = \nu_n\,;$$

then, we will discuss a simple example. We observe that the iterated conditional random quantities $(X_1|H_1)|K_1$, $\cdots$, $(X_n|H_n)|K_n$ involve the assessment $(\mu_1, \ldots, \mu_n)$ on $\{X_1|H_1, \ldots, X_n|H_n\}$; then, in the definition of coherence we must consider the global assessment $(\mu_1, \ldots, \mu_n, \nu_1, \ldots, \nu_n)$. We have

Definition 3. Given any random quantities $X_1, \ldots, X_n$ and any events $H_1, \ldots,$ $H_n, K_1, \ldots, K_n$, with $H_i \neq \emptyset, K_i \neq \emptyset$, $i = 1, \ldots, n$, the prevision assessment $(\mu_1, \ldots, \mu_n, \nu_1, \ldots, \nu_n)$ on $\mathcal{F} = \{X_1|H_1, \ldots, X_n|H_n, Y_1|K_1, \cdots, Y_n|K_n\}$, where $Y_1 = X_1|H_1, \ldots, Y_n = X_n|H_n$, is coherent if and only if, for every subfamily $\mathcal{S} \subseteq \mathcal{F}$, defining $\mathcal{H} = \bigvee_{i:X_i|H_i \in \mathcal{S}} H_i$, $\mathcal{K} = \bigvee_{i:Y_i|K_i \in \mathcal{S}} K_i$, and denoting by $G_{\mathcal{H} \vee \mathcal{K}}$ the set of possible values of the random gain

$$\mathcal{G} = \textstyle\sum_{i:X_i|H_i \in \mathcal{S}} s_i H_i (X_i - \mu_i) + \sum_{i:Y_i|K_i \in \mathcal{S}} \tau_i K_i (X_i H_i + \mu_i H_i^c - \nu_i)$$

restricted to $\mathcal{H} \vee \mathcal{K}$, with s_i, τ_i arbitrary real numbers for every i, it holds that $\inf G_{\mathcal{H} \vee \mathcal{K}} \leq 0 \leq \sup G_{\mathcal{H} \vee \mathcal{K}}$.

We observe that Definition 3 is nothing but Definition 1 applied to the family $\{X_i|H_i, Y_i|K_i, i = 1, \dots, n\}$, where $Y_i = X_i|H_i = X_iH_i + \mu_i H_i^c$, $\forall\, i$; hence the value $g_0 = 0$ of the random gain $\mathcal{G}$, associated with the atom $H_1^c \cdots H_n^c K_1^c \cdots K_n^c$ (*all the bets on* $X_i|H_i, (X_i|H_i)|K_i$, $i = 1, \dots, n$, *called off*), is discarded when defining coherence of the prevision assessment $(\mu_1, \dots, \mu_n, \nu_1, \dots, \nu_n)$ on the family $\{X_1|H_1, \dots, X_n|H_n, (X_1|H_1)|K_1, \cdots, (X_n|H_n)|K_n\}$. Moreover, the checking for coherence can be made by the usual methods already existing in literature (see, e.g., [7, 12, 13]). Based on the geometrical approach related to Theorem 1, we illustrate Definition 3 by the example below.

Example 2. Given a r.q. $X \in \{1, 2, \dots, 10\}$, we set $K = (X \in \{2, 4, \dots, 10\})$, $H = (X \leq 6)$, $\mathbb{P}(X|H) = \mu$, $\mathbb{P}[(X|H)|K] = \nu$, $\mathcal{M}_1 = (\mu)$, $\mathcal{S}_1 = \{X|H\}$, $\mathcal{M}_2 = (\nu)$, $\mathcal{S}_2 = \{Y|K\}$, $\mathcal{M}_3 = (\mu, \nu)$, $\mathcal{S}_3 = \{X|H, Y|K\}$, where $Y = X|H = XH + \mu H^c$.
As shown below, the set Π of coherent assessments (μ, ν) on $\{X|H, (X|H)|K\}$ is the (non convex) polygon whose boundary is the closed polygonal with vertices the points $(1, 1), (2, 2), (5, 2), (5, 5), (6, 6), (1, 6)$. We observe that Π is the union of the triangle T_1, with vertices the points $(1, 1), (6, 6), (1, 6)$, and the triangle T_2, with vertices the points $(2, 2), (5, 2), (5, 5)$.
We denote by $\mathcal{I}_j$ the convex hull associated with the pair $(\mathcal{S}_j, \mathcal{M}_j)$, $j = 1, 2, 3$. From a geometrical point of view, the coherence of (μ, ν) amounts to conditions $\mathcal{M}_j \in \mathcal{I}_j$, $j = 1, 2, 3$. Of course, $\mathcal{M}_1 \in \mathcal{I}_1$ if and only if $1 \leq \mu \leq 6$. If $2 \leq \mu \leq 6$, then $\mathcal{M}_2 \in \mathcal{I}_2$ is satisfied for every $\nu \in [2, 6]$; if $\mu < 2$, then $\mathcal{M}_2 \in \mathcal{I}_2$ for every $\nu \in [\mu, 6]$. To check if $\mathcal{M}_3 \in \mathcal{I}_3$, we determine the set of constituents contained in $H \vee K$, i.e. different from H^cK^c, which are obtained by expanding the expression $(HK \vee HK^c \vee H^cK) \wedge (A_1 \vee \cdots \vee A_{10})$, where $A_i = (X = i), i = 1, \dots, 10$. These constituents are A_2, A_4, A_6, A_1, A_3, A_5, A_8, A_{10}; the associated points Q_h's, for the pair $(\mathcal{S}_3, \mathcal{M}_3)$, are $(2, 2)$, $(4, 4)$, $(6, 6)$, $(1, \nu)$, $(3, \nu)$, $(5, \nu)$, (μ, μ), where with A_8 and A_{10} it is associated the same point (μ, μ).
We distinguish two cases: (i) $\mu \geq 2$; (ii) $\mu < 2$.
Case (i). For the convex hull it is enough to consider the points $(2, 2), (6, 6), (1, \nu)$, $(5, \nu)$. If $\mu \leq 5$ then (μ, ν) belongs to the segment with vertices $(1, \nu)$, $(5, \nu)$, so that the condition $\mathcal{M}_3 \in \mathcal{I}_3$ is satisfied and we have to continue by considering the condition $\mathcal{M}_2 \in \mathcal{I}_2$. If K is true, then $Y|K \in \{2, 4, 6, \mu\}$; hence, it must be $\nu \in [2, 6]$. If $\mu > 5$, then (μ, ν) belongs to the convex hull if and only if $\mu \leq \nu \leq 6$. In fact in this case (μ, ν) belongs to the triangle with vertices the points $(2, 2)$, $(6, 6)$, $(1, \nu)$. Of course, the condition $\nu \in [2, 6]$ is satisfied too.
Case (ii). For the convex hull it is enough to consider the points (μ, μ), $(6, 6), (1, \nu)$, $(5, \nu)$. Condition $\mathcal{M}_3 \in \mathcal{I}_3$ is satisfied because (μ, ν) belongs to the segment with vertices $(1, \nu), (5, \nu)$. Condition $\mathcal{M}_1 \in \mathcal{I}_1$ is satisfied because $1 \leq \mu < 2$; finally, the condition $\mathcal{M}_2 \in \mathcal{I}_2$ is satisfied for every $\nu \in [\mu, 6]$.

6 Conclusions

Based on betting scheme of de Finetti, we represented a c.r.q. as a suitable unconditional r.q., for which the assessed conditional prevision is one of the possible values. We obtained some results on basic operations among c.r.q.'s, by examining in particular a condition for the equality of two c.r.q.'s $X|H$ and $Y|K$. Then, we represented a c.r.q. $X|HK$ as a suitable c.r.q. $Y|K$ and we considered an application to Bayesian updating, by also deepening some aspects of Bayes' formula. We introduced a notion of iterated c.r.q. $(X|H)|K$, defined as a suitable c.r.q. $Y|K$, and we analyzed the relationship between $(X|H)|K$ and $X|HK$. Even if Bayesian updating cannot be formalized in our approach, we showed that our notion of iterated conditioning has an economic rationale. We discussed Bayesian updating in terms of iterated conditioning under the Import-Export Principle. But, such a principle is not valid in general and does not work in applications where our iterated conditioning does. Finally, we defined the notion of coherence for prevision assessments on iterated c.r.q.'s, by also giving an example. Future work should concern the extension of our results to the case of imprecise conditional prevision assessments.

Acknowledgments. The authors are grateful to the anonymous referees for they valuable criticisms and suggestions.

References

1. Adams, E.W.: The Logic of Conditionals. Reidel, Dordrecht (1975)
2. Biazzo, V., Gilio, A.: A generalization of the fundamental theorem of de Finetti for imprecise conditional probability assessments. Int. J. Approx. Reason. 24(2-3), 251–272 (2000)
3. Biazzo, V., Gilio, A.: On the linear structure of betting criterion and the checking of coherence. Ann. Math. Artif. Intell. 35(1-4), 83–106 (2002)
4. Biazzo, V., Gilio, A., Lukasiewicz, T., Sanfilippo, G.: Probabilistic logic under coherence: complexity and algorithms. Ann. Math. Artif. Intell. 45(1-2), 35–81 (2005)
5. Biazzo, V., Gilio, A., Sanfilippo, G.: Coherence checking and propagation of lower probability bounds. Soft Computing 7(5), 310–320 (2003)
6. Biazzo, V., Gilio, A., Sanfilippo, G.: On the Checking of G-Coherence of Conditional Probability Bounds. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 11, Suppl.2, 75–104 (2003)
7. Biazzo, V., Gilio, A., Sanfilippo, G.: Generalized coherence and connection property of imprecise conditional previsions. In: Proc. IPMU'08, Malaga, Spain, June22-27. pp. 907–914 (2008)
8. Biazzo, V., Gilio, A., Sanfilippo, G.: Coherent Conditional Previsions and Proper Scoring Rules. In: Advances in Computational Intelligence, CCIS, vol. 300, pp. 146–156. Springer (2012)
9. Brozzi, A., Capotorti, A., Vantaggi, B.: Incoherence correction strategies in statistical matching. Int. J. Approx. Reason. 53(8), 1124–1136 (2012)
10. de Campos, C.P., Cozman, F.G.: The inferential complexity of bayesian and credal networks. In: Kaelbling, L.P., Saffiotti, A. (eds.) IJCAI. pp. 1313–1318. Professional Book Center (2005)
11. Capotorti, A., Lad, F., Sanfilippo, G.: Reassessing Accuracy Rates of Median Decisions. The American Statistician 61(2), 132–138 (2007)

12. Capotorti, A., Vantaggi, B.: Locally Strong Coherence in Inference Processes. Ann. Math. Artif. Intell. 35(1-4), 125–149 (2002)
13. Coletti, G., Scozzafava, R.: Probabilistic logic in a coherent setting, Trends in logics, vol. 15. Kluwer, Dordrecht (2002)
14. Coletti, G., Scozzafava, R., Vantaggi, B.: Inferential processes leading to possibility and necessity. Information Sciences (2012), doi 10.1016/j.ins.2012.10.034
15. Dubois, D., Gilio, A., Kern-Isberner, G.: Probabilistic abduction without priors. Int. J. Approx. Reason. 47(3), 333–351 (2008)
16. de Finetti, B.: La Logique de la Probabilité. In: Actes du Congrès International de Philosophie Scientifique, Paris, 1935, pp. IV 1– IV 9. Hermann et C.ie (1936)
17. de Finetti, B.: Teoria delle probabilità. Ed. Einaudi, 2 voll., Torino (1970)
18. Gilio, A.: Probabilistic logic under coherence, conditional interpretations, and default reasoning. Synthese 146(1-2), 139–152 (2005)
19. Gilio, A.: Criterio di penalizzazione e condizioni di coerenza nella valutazione soggettiva della probabilità. Boll. Un. Mat. Ital. 4-B(3, Serie 7), 645–660 (1990)
20. Gilio, A.: Probabilistic Reasoning Under Coherence in System P. Ann. Math. Artif. Intell. 34(1-3), 5–34 (2002)
21. Gilio, A.: Generalizing inference rules in a coherence-based probabilistic default reasoning. Int. J. Approx. Reasoning 53(3), 413–434 (2012)
22. Gilio, A., Over, D.: The psychology of inferring conditionals from disjunctions: A probabilistic study. Journal of Mathematical Psychology 56(2), 118–131 (2012)
23. Gilio, A., Sanfilippo, G.: Conditional Random Quantities and Compounds of Conditionals. Studia Logica (submitted). http://arxiv.org/abs/1304.4990
24. Gilio, A., Sanfilippo, G.: Quasi Conjunction and p-entailment in Nonmonotonic Reasoning. In: Borgelt, C., et al. (eds.) Combining Soft Computing and Statistical Methods in Data Analysis, AISC, vol. 77, pp. 321–328. Springer (2010)
25. Gilio, A., Sanfilippo, G.: Quasi conjunction and inclusion relation in probabilistic default reasoning. In: Liu, W. (ed.) ECSQARU'11, LNCS, vol. 6717, pp. 497–508. Springer (2011)
26. Gilio, A., Sanfilippo, G.: Conjunction, Disjunction and Iterated Conditioning of Conditional Events. In: Synergies of Soft Computing and Statistics for Intelligent Data Analysis, AISC, vol. 190, pp. 399–407. Springer (2013)
27. Gilio, A., Sanfilippo, G.: Probabilistic entailment in the setting of coherence: The role of quasi conjunction and inclusion relation. Int. J. Approx. Reason. 54(4), 513–525 (2013)
28. Gilio, A., Sanfilippo, G.: Quasi conjunction, quasi disjunction, t-norms and t-conorms: Probabilistic aspects. Information Sciences (2013), doi 10.1016/j.ins.2013.03.019
29. Jaumard, B., Hansen, P., Poggi de Aragão, M.: Column generation methods for probabilistic logic. ORSA Journal on Computing 3(2), 135–148 (1991)
30. Kaufmann, S.: Conditionals Right and Left: Probabilities for the Whole Family. Journal of Philosophical Logic 38(1), 1–53 (2009)
31. Lad, F.: Operational Subjective Statistical Methods: a mathematical, philosophical, and historical introduction. John Wiley, New York (1996)
32. Lad, F., Sanfilippo, G., Agró, G.: Completing the logarithmic scoring rule for assessing probability distributions. AIP Conf. Proceedings 1490(1), 13–30 (2012)
33. McGee, V.: Conditional Probabilities and Compounds of Conditionals. Philosophical Review 98(4), 485–541 (1989)
34. Miranda, E., Zaffalon, M., Cooman de, G.: Conglomerable natural extension. Int. J. Approx. Reason. 53(8), 1200–1227 (2012)

35. Nilsson, N.J.: Probabilistic logic. Artificial Intelligence 28(1), 71–87 (1986)
36. Pfeifer, N., Kleiter, G.D.: Inference in conditional probability logic. Kybernetika 42, 391–404 (2006)
37. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London (1991)
38. Walley, P., Pelessoni, R., Vicig, P.: Direct algorithms for checking consistency and making inferences from conditional probability assessments. Journal of Statistical Planning and Inference 126(1), 119–151 (2004)
39. Wallmann, C., Kleiter, G.D.: Beware of too much information. In: Kroupa, T., Vejnarova, J. (eds.) Proceedings of WUPES 2012, Prague, pp. 214–225 (2012)