Coherent lower and upper conditional previsions defined by Hausdorff inner and outer measures to represent the role of conscious and unconscious thought in human decision making

Serena Doria[1]

Abstract

The model of coherent lower and upper conditional previsions, based on Hausdorff inner and outer measures, is proposed to represent the preference orderings and the equivalences, respectively assigned by the conscious and unconscious thought in human decision making under uncertainty. Complexity of partial information is represented by the Hausdorff dimension of the conditioning event. When the events, that describe the decision problem, are measurable is represented to the s-dimensional Hausdorff outer measure, where s is the Hausdorff dimension of the conditioning event, an optimal decision can be reached. The model is applied and discussed in Linda's Problem and the conjunction fallacy is resolved.

Keywords Coherent upper and lower conditional previsions · Hausdorff measures · Conscious and Unconscious thought · Linda's Problem · Conjunction fallacy

1 Introduction

Modeling human decisions under uncertainty is one of the issues that became crucial in AI in the recent years. Mathematical models for decision making under risk, such as those developed in the Expected Utility Theory, provide optimal decision. Nevertheless experiments and studies show that these rational decision models are not always able to describe the typical human approach to making decisions.

In the Prospect Theory [19] human reasoning and human decision making is investigated by different experiments which evidence biases of human intuition. Experimental methods lead to describe the dual process of the brain activity as regulated by two different ways of

---

✉ Serena Doria
serena.doria@unich.it

[1] University G. d'Annunzio, Chieti, Italy

thinking namely fast and slow thinking denoted also as System 1 and System 2, respectively. System 1 regulates intuitive, involuntary, unconscious and effortless activities while System 2 is the conscious part of the brain in charge of logical reasoning [20]. Also according to Matte Blanco [21] conscious and unconscious activities are two different modes of being. Specifically, he draws a distinction between logical conscious/asymmetrical thought - structured on the categories of time and space and ruled by Aristotle's principle of non-contradiction and unconscious/symmetrical thought - based upon the principles of symmetry and generalization. Both types of thoughts are supposed to combine in different human thinking experiences since they yield a bi-logic asset. Emotions are ways to reach and decode the unconscious since they function at the same way. Both scenarios stress the importance of emotion-driven, unconscious thinking in human decision-making so it is natural to wonder if mathematical models are able to describe and to re-produce these different aspects of human mind.

In this paper a new model of coherent upper and lower conditional previsions based on Hausdorff outer and inner measures [6–10, 11] is proposed to describe this dual aspect of human brain's activity.

Coherent upper and lower conditional previsions are non-linear functions satisfying the axioms of coherence and coherent lower and upper conditional probabilities are obtained when only indicator functions are considered [28]. In recent papers [14, 15] new mathematical approaches to construct coherent lower and upper previsions have been proposed.

Many properties of coherent lower previsions can be obtained by the conjugate coherent upper conditional previsions but the two non-linear functionals represent different binary relations between random variables; in fact preference orderings represented by the coherent lower previsions satisfy the antisymmetric property which is not satisfied by the binary relation represented by their conjugate coherent upper conditional previsions [8, 11, 12].

Coherent lower and upper conditional previsions defined by Hausdorff inner and outer measures are proposed to represent respectively a partial strict order and a complete equivalence relation between random variables. The two binary relations can describe the activity of the conscious human thought ruled by the antisymmetric property and of the unconscious human thought which is governed by the symmetric principle and the generalization principle according to the theory developed by Matte-Blanco. The model highlights also the role of unexpected events (event with zero probability) in the updating of knowledge and awareness respectively of the conscious and unconscious thought.

Let Ω be a non-empty set which represent the set of all possible results of a random phenomenon and let $\boldsymbol{B}$ be a partition of Ω. The partition represents the information we have about the phenomenon in the sense that we do not know which is the result ω of the experiment but we know which is the set B of the partition which contains the result ω.

If $\boldsymbol{B}$ is the partition of singletons of Ω we have precise information that is we know the exact result of the random experiment. In this case coherence assures that the conditional probability of an event A given $\boldsymbol{B}$ is equal to the indicator function of the event A.

This necessary condition of coherence is contradicted when conditional expectation is defined by the Radon-Nikodym derivative in the axiomatic way [1]; it occurs because one of the defining properties of the Radon-Nikodym derivative,that is to be measurable with respect to the σ-field of the conditioning events, contradicts the previous necessary condition [7].

For every $B \in \boldsymbol{B}$ the new model of coherent upper conditional probability based on Hausdorff outer measures is introduced. Each event is represented by a subset of Ω and his complexity is expressed by a number s, called Hausdorff dimension of the set. In particular for every conditioning event B let s be the Hausdorff dimension of the set B; the Hausdorff

measure of order s is called the *Hausdorff measure associated* with the coherent upper conditional probability $\overline{P}(E|B)$ for every event E. The coherent upper conditional probability of an event E given the information B, is defined by the Hausdorff measure of order s if s is the Hausdorff dimension of the set B. A priori measure of uncertainty is the level of knowledge each subject has before having a new piece of information, denoted by the set B; the measure of uncertainty that quantifies the level of knowledge each subject has on a posteriori situations is a coherent upper probability conditioned to the state B. The model of coherent upper conditional probabilities defined by Hausdorff outer measures has been proposed to affirm that there is updating of knowledge if the a priori and a posteriori measures of uncertainty are different, that is if they are defined by different Hausdorff outer measures. It is possible when the new piece of information represented by the set B, has a different complexity, measured in terms of Hausdorff dimension of the set, with respect to the previous information. In particular it occurs when the new information is represented by a set B with Hausdorff dimension less than the Hausdorff dimension t where t is the Hausdorff dimension of previous information. The t-dimensional Hausdorff measure defines the a priori coherent upper conditional probability which represent the knowledge before having information represented by B; the set B has zero a priori coherent upper conditional probability since its Hausdorff dimension is less then t. To exemplify how the model of coherent upper and lower conditional previsions, based on Hausdorff outer and inner measures, can be used to describe the two different activities of the human brain it is applied to solve the conjunction fallacy in Linda′s Problem; it is one of the most discussed heuristics [27] that exemplifies how real human decision-making under risk or uncertainty constantly violates the logic and most fundamental rules of probability calculus. Conjunction fallacy has been studied in literature [18, 22, 25, 26]. In [26] the conjunction fallacy is explained in the theoretic approach of inductive confirmation where the conjunction fallacy depends on the added conjunct being perceived as inductively confirmed. The incoherence of the conjunction fallacy in terms of Dutch book is also illustrated in [17]. In Doria and Cenci [13] coherent upper and lower conditional probabilities are proposed to solve the conjunction fallacy in Linda's Problem. In the quoted paper a monotone set function is assessed on the class of conditional events involved in Linda's Problem, it is extended to the field generated by the class by outer and inner measures, which are coherent upper and lower conditional probabilities, so that the conjunction fallacy does not occur and the preference ordering express by the participants to the Linda's experiment is mathematically represented. This approach, based on non-linear uncertainty measures, may preclude the possibility to have an optimal choice, that is to find a Bayes random variable under a coherent lower conditional prevision, which is a random variable that is maximal under a linear prevision on a given class of random variables. In this paper we show that the conjunction fallacy of Linda's Problem can be also solved if the preference ordering among events is represented by the linear model of coherent conditional probability defined by Hausdorff measures. It occurs when the considered events, involved in Linda's Problem are represented by sets which are measurable with respect to the Hausdorff measure of order s, where s is the Hausdorff dimension of the set B that represents the given information.

2 Coherent upper and lower conditional previsions defined by Hausdorff outer and inner measures

Let (Ω, d) be a metric space and let $\boldsymbol{B}$ be a partition of Ω. A bounded random variable is a function $X : \Omega \Re$ and $L(\Omega)$ is the class of all bounded random variables defined on Ω; for every

$B \in \boldsymbol{B}$ denote by $X \mid B$ the restriction of X to B and by $\sup(X \mid B)$ the supremum value that X assumes on B. Let $L(B)$ be the class of all bounded random variables $X \mid B$. Denote by I_A the indicator function of any event $A \in \wp(B)$, i.e. $I_A(\omega) = 1$ if $\omega \in A$ and $I_A(\omega) = 0$ if $\omega \in A^c$.

For every $B \in \boldsymbol{B}$ coherent upper conditional previsions $\overline{P}(\cdot|B)$ are functionals defined on $L(B)$ [28].

Definition 1 Coherent upper conditional previsions are functionals $\overline{P}(\cdot|B)$ defined on $L(B)$, such that the following axioms of coherence hold for every X and Y in $L(B)$ and every strictly positive constant λ:

•

$$\overline{P}(X|B) \leq \sup(X|B);$$

• $\overline{P}(\lambda X|B) = \lambda \overline{P}(X|B)$ (positive homogeneity);
• $\overline{P}(X + Y|B) \leq \overline{P}(X|B) + \overline{P}(Y|B)$ (subadditivity).

Suppose that $\overline{P}(X|B)$ is a coherent upper conditional prevision on a linear space $\mathcal{K} \subseteq L(B)$. Then its conjugate coherent lower conditional prevision is defined by the conjugacy property $\underline{P}(X|B) = -\overline{P}(-X|B)$. If for every X belonging to $\mathcal{K}$ we have $P(X|B) = \underline{P}(X|B) = \overline{P}(X|B)$ then $P(X|B)$ is called a coherent *linear* conditional prevision and if $\mathcal{K} = L(B)$ it is a linear, positive and positively homogenous functional in the sense of de Finetti [3, 4], Regazzini [23] and Walley [28].

From axioms 1)-3) and by the conjugacy property we have that

$$1 \leq \underline{P}(I_B|B) \leq \overline{P}(I_B|B) \leq 1$$

so that

$$\underline{P}(I_B|B) = \overline{P}(I_B|B) = 1.$$

In Walley [28] the functionals $\overline{P}(X|B)$ defined for $B \in \boldsymbol{B}$ and $X \in L(B)$ satisfying axioms 1)-3) and such that $\overline{P}(I_B|B) = 1$ are called *separately coherent*.

The unconditional coherent upper prevision $\overline{P} = \overline{P}(\cdot|\Omega)$ is obtained as a particular case when the conditioning event is Ω. Coherent upper conditional probabilities are obtained when only 0–1 valued random variables are considered.

An upper prevision is a real-valued function defined on some class of bounded random variables K. A necessary and sufficient condition for an upper prevision $\overline{P}$ to be coherent is to be the *upper envelope* of linear previsions, i.e. there is a class M of linear previsions such that $\overline{P} = \sup\{P : P \in M\}$.

A new model of coherent upper conditional probability based on Hausdorff outer measures is introduced in [6–8]. For the definition of Hausdorff outer measure and its basic properties see [16, 24].

The innovative aspect consists in the fact that the measure that is used to define the conditional upper probability depends on the complexity of the conditioning event, given in terms of Hausdorff dimension of the set B.

Therefore the events with a zero-value a priori probability determine the change of the measure of uncertainty that represents the level of knowledge of the subject.

Let $\delta > 0$ and let s be a non-negative number. The *diameter* of a non empty set U of Ω is defined as $|U| = \sup\{d(x, y) : x, y \in U\}$ and if a subset A of Ω is such that $A \subseteq \cup_i U_i$ and $0 < |U_i| \leq \delta$ for each i, the class $\{U_i\}$ is called a δ-cover of A.

The *Hausdorff s-dimensional outer measure* of A, denoted by $h^s(A)$, is defined on $\wp(\Omega)$, the class of all subsets of Ω, as

$$h^s(A) = \lim_{\delta \to 0} \inf \sum_{i=1}^{+\infty} |U_i|^s.$$

where the infimum is over all δ-covers $\{U_i\}$.

A subset A of Ω is called *measurable* with respect to the outer measure h^s if it decomposes every subset of Ω additively, that is if

$$h^s(E) = h^s(A \cap E) + h^s(E - A)$$

for all sets $E \subseteq \Omega$.

Hausdorff s-dimensional outer measures are submodular, continuous from below and their restriction on the Borel σ-field is countably additive.

The *Hausdorff dimension* of a set A, $dim_H(A)$, is defined as the unique value, such that

$$h^s(A) = +\infty \text{ if } 0 \leq s < dim_H(A),$$

$$h^s(A) = 0 \text{ if } dim_H(A) < s < +\infty.$$

For every $B \in \boldsymbol{B}$ denote by s the Hausdorff dimension of B and let h^s be the Hausdorff s-dimensional Hausdorff outer measure associated to the coherent upper conditional prevision. For every bounded random variable X a coherent upper conditional prevision $\overline{P}(X|B)$ is defined by the Choquet integral with respect to its associated Hausdorff outer measure if the conditioning event has positive and finite Hausdorff outer measure in its Hausdorff dimension. Otherwise if the conditioning event has Hausdorff outer measure in its Hausdorff dimension equal to zero or infinity it is defined by a 0–1 valued finitely, but not countably, additive probability. The following theorem has been proven in [7].

Theorem 1 Let (Ω, d) be a metric space and let $\boldsymbol{B}$ be a partition of Ω. For every $B \in \boldsymbol{B}$ denote by s the Hausdorff dimension of the conditioning event B and by h^s the Hausdorff s-dimensional outer measure. Let m_B be a 0–1 valued finitely additive, but not countably additive, probability on $\wp(B)$. Thus, for each $B \in \boldsymbol{B}$, the function defined on $\wp(B)$ by

$$\overline{P}(A|B) = \begin{cases} \dfrac{h^s(A \cap B)}{h^s(B)} & if \quad 0 < h^s(B) < +\infty \\ m_B & if \quad h^s(B) \in \{0, +\infty\} \end{cases}$$

is a coherent upper conditional probability.

If $B \in \boldsymbol{B}$ is a set with positive and finite Hausdorff outer measure in its Hausdorff dimension s the monotone set function μ_B^* defined for every $A \in \wp(B)$ by $\mu_B^*(A) = \frac{h^s(AB)}{h^s(B)}$ is a coherent upper conditional probability, which is submodular, continuous from below and such that its restriction to the σ-field of all μ_B^* measurable sets is a Borel regular countably additive probability.

The coherent upper unconditional probability $\overline{P} = \mu_\Omega^*$ defined on $\wp(\Omega)$ is obtained for B equal to Ω.

Denoted by h_s the Hausdorff inner measure of order s, which is the dual of the Hausdorff outer measures of order s h^s, we have that the conjugate lower conditional probability $\overline{\mu}_B^*$ of μ_B^* is

$$\overline{\mu}_B^*(B) = \mu_B^*(\Omega) - \mu_B^*(B^c) = 1 - 0 = 1 = \mu_B^*(B)$$

Moreover

$$\overline{\mu}_B^*(A) = \mu_B^*(\Omega) - \mu_B^*(A^c) = \frac{h^s(\Omega \cap B)}{h^s(B)} - \frac{h^s(A^c \cap B)}{h^s(B)}.$$

In the following theorem, proven in [7], the coherent upper conditional probability defined in Theorem 1 is extended to the class of all bounded random variables and, when the conditioning event B has positive and finite Hausdorff outer measure in its Hausdorff dimension, the coherent upper prevision id defined by the Choquet integral [2, 5].

Theorem 2 Let (Ω, d) be a metric space and let **B** be a partition of Ω. For every $B \in \mathbf{B}$ denote by s the Hausdorff dimension of the conditioning event B and by h^s the Hausdorff s-dimensional outer measure. Let m_B be a 0-1valued finitely additive, but not countably additive, probability on $\wp(B)$. Then for each $B \in \mathbf{B}$ the functional $\overline{P}(X|B)$ defined on $L(B)$ by

$$\overline{P}(X|B) = \begin{cases} \frac{1}{h^s(B)} \int_B X dh^s & \text{if} \quad 0 < h^s(B) < +\infty \\ \int_B X dm_B & \text{if} \quad h^s(B) \in \{0, +\infty\} \end{cases}$$

is a coherent upper conditional prevision.

When the conditioning event B has Hausdorff outer measure in its Hausdorff dimension equal to zero or infinity, an additive conditional probability is coherent if and only if it takes only $0 - 1$ values. Because linear previsions on $L(B)$ are uniquely determined by their restrictions to events, the class of linear previsions on $L(B)$ whose restrictions to events take only the values 0 and 1 can be identified with the class of $0 - 1$ valued additive probability defined on all subsets of B (Walley 1991). In Theorem 1 and Theorem 2 a different m_B is chosen for each B.
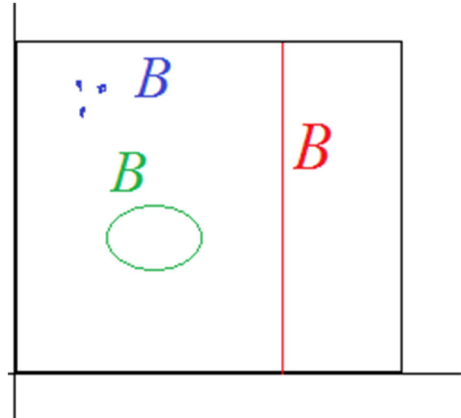
If the conditioning event B has positive and finite Hausdorff outer measure in its Hausdorff dimension the functional $\overline{P}(X|B)$ is proven to be monotone, comonotonically additive, submodular and continuous from below.

In Fig. 1 some examples of conditioning events B are given with different Hausdorff dimension: if B is a finite set of points its Hausdorff dimension is 0, if B is a segment its Hausdorff dimension is 1, if B is a set whose border is an ellipse then its Hausdorff dimension is 2; the conditional prevision is defined respectively by the Hausdorff measure of order 0, 1 and 2.

3 Preference ordering and indifference between random variables represented by coherent lower and upper conditional previsions

Linear functional are used to represent preference orderings in the following way.

Fig. 1 Examples of conditioning events B with different Hausdor dimension

Definition 2 A preference ordering $\succ$ on the class $L(B)$ of random variables defined on B is represented by a linear functional Γ if and and only if

$$X_i|B \succ X_j|B \Leftrightarrow \Gamma(X_i|B) > \Gamma(X_j|B)$$

and

$$X_i|B \approx X_j|B \Leftrightarrow \Gamma(X_i|B) = \Gamma(X_j|B)$$

Nevertheless there are preference orderings that cannot be represented by a linear functional as occurs in Example 1 of [11].

A partial strict order can be represented by the lower conditional prevision $\underline{P}(X|B)$.

In Walley [28] different preference orderings are defined with respect to lower and upper coherent previsions.

For each $B \in \mathbf{B}$ let $\underline{P}(\cdot|B)$ and $\overline{P}(\cdot|B)$ be respectively a lower coherent conditional prevision and its conjugate upper coherent conditional prevision defined on the class $L(B)$ of all random variables on B. Denote by K a sub-class of $L(B)$.

For each $B \in \mathbf{B}$ a partial strict order $\succ_*$, (i.e. a partial antisymmetric and transitive binary relation) is defined with respect to a coherent lower conditional prevision and a weak order $\succ^*$, (i.e. a complete reflexive and transitive binary relation) is defined with respect to a coherent upper conditional prevision.

Definition 3 We say that X is *preferable* to Y given B with respect to $\underline{P}$ i.e. $X \succ_* Y$ in B if and only if

$$\underline{P}((X-Y)|B) > 0$$

In particular we show that.

Proposition 1 the binary relation $\succ_*$ satisfies the antisymmetric property, i.e.

$$X \succ_* Y \Longleftrightarrow \underline{P}((X-Y)|B) > 0 \Longrightarrow$$

$$\underline{P}\big((Y-X|B) \leq 0 \Longleftrightarrow Y \, not \succ_* X.$$

Proof. We have

$$0 < \underline{P}\,((X-Y)|B) < \overline{P}((X-Y)|B) \Longrightarrow$$

$$\overline{P}\,((X-Y)|B) = -\,\underline{P}\,((Y-X)|B) > 0$$

so that $\overline{P}((Y-X)\,|B) < 0$ that is $Y not \succ_* X$. ∘.

Two random variables which have previsions equal to zero cannot be compared by the ordering $\succ_*$.

A weak order $\succ^*$ can be defined on $L(B)$ with respect to $\overline{P}$ but it cannot represent a strict preference ordering because it does not satisfied the antisymmetric property.

Definition 4 We say that $X \succ^* Y$ given B if and only if $\overline{P}((X-Y)\,|B) > 0$.

Example 1 Let $X, Y \in L(B)$ such that $\overline{P}((X-Y)\,|B) > 0$ and $\underline{P}\,((X-Y)\,|B) < 0$; then $\overline{P}((X-Y)\,|B) > 0$ does not imply $\overline{P}((Y-X)\,|B) < 0$ since

$$\overline{P}((Y-X)|B) < 0 \Longleftrightarrow -\,\underline{P}\,((X-Y)|B) < 0 \Longleftrightarrow \underline{P}\,((X-Y)|B) > 0$$

Two complete equivalence relations, which are complete reflexive, symmetric and transitive binary relations on $L(B)$ can be represented by the coherent upper conditional prevision $\overline{P}$ $(X|B)$.

Definition 5 Two random variables X and $Y \in L(B)$ are equivalent given B with respect to $\overline{P}$ if and only if $\overline{P}(X|B) = \overline{P}(Y|B)$.

Definition 6 We say that X and Y are *indifferent* given B with respect to $\overline{P}$, i.e. $X \approx Y$ in B if and only if

$$\overline{P}((X-Y)|B) = \overline{P}((Y-X)|B) = 0.$$

Remark 1 If the coherent conditional prevision $P(\cdot|\,B)$ is linear then

$$P((X-Y)|B) = P((Y-X)|B) = 0 \Longleftrightarrow P(X|B) = P(Y|B)$$

and two random variables X and Y are indifferent given B if and only if they are equivalent given B.

In Theorem 14 of [8] it has been proven that if two random variables are indifferent with respect to the coherent upper conditional prevision defined in Theorem 2 then they are equivalent. The result holds because Hausdorff outer measures are submodular.

Theorem 3 Let $X, Y \in L(B)$ be two random variables, which are indifferent given B with respect to $\overline{P}$ then they are indifferent with respect to the conjugate lower conditional prevision $\underline{P}$, that is $\underline{P}\,((X-Y)\,|B) = P((Y-X)\,|B) = 0.$

Proof. Since $X, Y \in L(B)$ are indifferent given B we have

$$-\underline{P}((Y-X)|B) = \overline{P}((X-Y)|B) = 0$$

$$-\underline{P}((X-Y)|B) = \overline{P}((Y-X)|B) = 0,$$

so that $\underline{P}((X-Y)|B) = \underline{P}((Y-X)|B) = 0.$ ∘

Theorem 4 Let $X, Y \in L(B)$ be two random variables, such that $X \succ_* Y$ given B with respect to $\underline{P}(\cdot|B)$ then X and Y are not indifferent given B with respect to $\overline{P}(\cdot|B)$.

Proof. If $X \succ_* Y$ given B with respect to $\underline{P}(\cdot|B)$ then

$$0 < \underline{P}((X-Y)|B) \leq \overline{P}((X-Y)|B)$$

so $\overline{P}((X-Y)|B) \neq 0$ and X and to Y are not indifferent given B with respect to $\overline{P}(\cdot|B)$ according to Definition 4. ∘

In the following example, discussed in Example 8 and Example 9 of [11], three random variables X_1, X_2, X_3 are given and it is shown that the lower vacuous conditional prevision does not represent the preference ordering $X_1 \succ_* X_2$ and the upper vacuous conditional prevision does not represent the indifference between X_2 and X_3; In Example 3 it is shown that these binary relations are represented by the model proposed in Theorem 2.

Example 2 Let $\boldsymbol{B} = \{B_1, B_2\}$ and $K = \{X_1, X_2, X_3\}$ with

random variables	B_1	B_2
X_1	0.3	0.3
X_2	0.7	0.0
X_3	0.0	0.7

The preference ordering $X_1 \succ X_2$ and $X_2 \approx X_3$ cannot be represented by the lower vacuous conditional prevision defined by $\underline{P}(X|\Omega) = inf\{X(\omega) : \omega \in \Omega\}$ since

$$\underline{P}((X_1-X_2)|\Omega) = -0.4 \ \text{(and)}$$

$$\underline{P}((X_2-X_3)|\Omega) = \underline{P}(X_3-X_2|\Omega) = 0$$

and it is not represented by the upper vacuous conditional prevision $\overline{P}(X|\Omega) = sup\{X(\omega) : \omega \in \Omega\}$ because

$$\overline{P}((X_1-X_2)|\Omega) = 0 \ \text{(and)}$$

$$\overline{P}((X_2-X_3)|\Omega) = \overline{P}((X_3-X_2)|\Omega) = 0.7.$$

In the next example it is shown that the preference ordering between the random variables X_i for $i = 1, \ldots, 3$, defined as in Example 2, can be represented by the coherent upper conditional prevision defined in Theorem 2 and by its conjugate lower conditional prevision.

Example 3 Let (Ω, d) be a metric space with $\Omega = N$ so that $dim_H(\Omega) = 0$ and $h^0 = (\Omega) = +\infty$. Let $\boldsymbol{B} = \{B_1, B_2\}$ be the partition of Ω where $B_1 = \{p \in N : p = 2n; n \in N\}$ and $B_2 = \{d \in N : d = 2n - 1; n \in N\}$ so that $dim_H(B_1) = dim_H(B_2) = 0$ and $h^0(B_1) = h^0(B_2) = +\infty$ and let X_i for $i = 1$, ..., 3 be the random variables defined in Example 2. By Theorem 2 we have the coherent lower and upper conditional previsions are equal to a 0–1 valued finitely additive, but not countably, probability

$$P((X_1 - X_2)|\Omega) = 1$$

$$P((X_2 - X_3|\Omega)) = P(X_3 - X_2|\Omega) = 0.$$

and the ordering $X_1 \succ X_1$ and $X_2 \approx X_2$ can be represented by the given coherent conditional prevision.

When only indicator functions are considered the preference ordering in Definition 3 and the equivalence and indifference relations given respectively in Definition 5 and Definition 6 permit to make a decision between events. It may occur that it is precluded to have a unique optimal choice if coherent upper and lower conditional previsions do not coincide.

Definition 7 A random variable X_i is admissible in K under $\underline{P}(\cdot|B)$ if no random variable $X_j \in K$ with $i \neq j$ is preferable to X_i with respect to $\succ_*$.

Definition 8 An admissible random variable X_i in K is maximal under $\underline{P}(\cdot|B)$ if it is preferable to X_j according to $\succ^*$ for all $X_j \in K$, so also the coherent upper conditional prevision is involved to determine a maximal random variable in a class K.

Definition 9 A Bayes random variable under a coherent lower conditional prevision is a random variable which is maximal under a linear prevision on the class of all random variables defined on B.

If information is represented by a set B with Hausdorff dimension s and if we consider a class of events represented by sets, which are measurable with respect to the s-Hausdorff outer measure, then coherent upper and lower probabilities coincide and an optimal choice can be obtained. It is represented by a Bayes random variable. In the next section coherent conditional probability defined by Hausdorff measures is considered in Linda's Problem to represent partial knowledge that the participants to the experiment have, given the information. The possibility to represent partial knowledge by a unique linear conditional probability permits to obtain a unique optimal choice.

4 Optimal choice in Linda's problem

A natural interpretation of the conditional probability is to represent the partial knowledge we have about a random phenomenon when information, represented by an event B, is provided. One of the main advantages of the concept of conditional probability is that it is functional to fostering the complexity of decision making but without precluding the possibility of an optimal choice. Namely, conditional probability can be used to represent the preference orderings between events with the aim to determine an optimal choice that can actually come

to. When coherent upper and lower conditional probabilities coincide, the preference ordering defining in Definition 3 and the binary relations given in Definition 5 and in Definition 6 are represented by a conditional probability in the following way:

Definition 10 E is preferable to F given B if the conditional probability probability $P(E|B)$ is greater than the conditional probability $P(F|B)$ and the two events are equivalent if the conditional probability probability $P(E|B)$ and $P(F|B)$ are equal, that is

$$E \succ F \text{ given } B \Leftrightarrow P(E|B) > P(F|B)$$

$$E \approx F \text{ given } B \Leftrightarrow P(E|B) = P(F|B).$$

When no information is given the conditioning event is assumed to be Ω, which is the set of all possible results of the random phenomenon, and the unconditional prevision is obtained. The unconditional probability of an event A is denoted by $P(A) = P(A|\Omega)$.

In Tverskyâ€™s and Kahnemanâ€™s heuristics-based explanation of rational behavior, as in the so-called Lindaâ€™s problem, preference orderings are represented by an unconditional probability (i.e., the standard notion of probability such as in EUT). It iswhat produces the well-known "conjunction fallacy" because the information about Linda given to the participants to the experiment cannot be considered in the illustration of the preference orderings.

But letâ€™s look at the famous experimental test more closely. Participants are introduced to the following dilemma: Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which current event about Linda is more likely?

- E: "Linda is a bank teller".
- $E \cap F$: "Linda is a bank teller and is active in the feminist movement".

The majority of the participants (85% of those asked) indicated the conjunction of sentences as more probable than one of its conjuncts.The majority of the participants choose/prefer the event $E \cap F$ to the event E. That is, the probability assigned to a conjunction of two events is larger than the probability assigned to each of these two events (conjunction fallacy). It is in contrast with the axioms of the standard probability calculus which states $P(F \cap E) \leq P(E)$ while, in the case represented in the experiment, respondents consider that $P(E) < P(E \cap F)$. This fallacious result became, among other things, the basis for the creation of the concept of representativeness heuristics which are fully developed in Kahneman [27] when the 48 heuristics coming from years of experimental work - Linda is the heuristics n. 15 - are used to explain the two modes functioning of the brain; namely,either fast or slow thinking as well as their implications for representing and explaining rational decisions under risk and uncertainty.

As stated above, it is believed that the main problem of the original experiment, and what causes the fallacy, rely on how the preference ordering is represented. Once following the axioms of EUT, the preference orderings cannot be represented - accordingto Definition 10 - by an unconditional probability P, in which the full information provided about Linda, can actually be used to offer both a better picture of the decision making and a consistent result. In

fact, if the preference ordering $E \cap F$ is preferable to E is represented by the unconditional probability, the conjunction fallacy occurs since $E \cap F$ is preferable to E i.e. $E \cap F \succ E$ if and only if $P(E \cap F) > P(E)$.

Conversely, the conjunction fallacy can be dissolved once the preference ordering is represented by conditional probability which is defined, precisely, on the basis of the complexity of the given information. In the original experiment, information givenabout Linda, produces, in the majority of the participants, the idea that it is more likely that Linda is a feminist than she is not. Instead, when applying the conditional probability concept, which depends on the complexity of the information, the level of knowledge about Linda is updated. Thus,the resulting preference ordering does not incur in the fallacy.

This can be demonstrated by reconsidering the following events in a new light:

- E: „Linda is a bank teller.“
- F: „Linda is active in the feminist movement.“
- $E \cap F$: „Linda is a bank teller and is active in the feminist movement.“
- B: „Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.“

In the case of unconditional probability, no information about Linda can be considered.

No mathematical model, where uncertainty due to partial knowledge, is represented by a monotone set function (as conditional probability or upper and lower conditional probabilities) can describe the belief of the majority of the participants in Linda's problem by the condition $P(E \cap F | B) > P(E | B)$; in this case the monotony is not assured and so the conjugacy fallacy occurs.

The model proposed in the paper wants to stress the role of information, represented by the set B, in the updating of knowledge not only about the events $E \cap F$ and E but also about the events F and F^c. The fact that the majority prefers $E \cap F \mid B$ to $E \mid B$ is expressed by the result that information B produces. In other words the information B produces an updating with respect to every event involved in the problem.

To know the event B produces a change in the degree of belief on the fact that Linda is feminist, so that we should have:

$$P(F|B) > P(F^c|B)$$

In this case the preference ordering has to be represented in terms of conditional probability, which embodies the new partial knowledge about Linda when information, represented by the set B, is provided. Here, an event is defined by a proposition which is true or false and can be represented by a set which is uniquely determined by a property i.e., a true or false statement. Accordingly, Linda's Problem can be reformulated in terms of sets in which the complexity of an event can be represented by the geometric complexity of the set. This kind of complexity is expressed by a number called Hausdorff dimension of the set. In this paper, the updating process of available knowledge about Linda is then, represented by a conditional probability that is defined by the s-dimensional Hausdorff measures where s is the Hausdorff dimension of the conditioning event represented by the set B. The Hausdorff dimension of a set is a number that measures the geometric complexity of the set B.

As a result, the model of conditional probability proposed that describes the partial knowledge in Linda's Problem as said above assures that:

- the events E and $E \cap F$ are equivalent given the information B. Thus, the conjunction fallacy is dissolved since $P(E|B) = P(E \cap F|B)$
- the fact that a majority of the participants prefer the event $E \cap F$ with respect to the event E given the information B is due to the fact that the updated information about Linda corroborates the idea that Linda is a feminist, i.e. $P(F|B) = 1 > 0 = P(F^c|B)$.

5 Preference orderings represented by coherent conditional probability defined by Hausdorff measures in Linda's problem

In this section the model of conditional probability proposed in Theorem 1 is applied to Linda's Problem and it solves the conjunction fallacy since the events E and $F \cap E$ are equivalent given B, that is $P(E|B) = P(E \cap F)$ and $P(F|B) = 1$ and $P(F^c|B) = 0$ and this is the explanation that the majority of the participants to Linda's experiment between two equivalent events choose the event $E \cap F$.

The previous events can be represented by the following sets; let $\Omega = [0, 1]$ and

$$E = \left[0, \frac{1}{4}\right], F = [0, 1] - \left\{\frac{1}{4}, \frac{1}{2}\right\} B = \left[0, \frac{3}{4}\right] - \left\{\frac{1}{8}, \frac{3}{8}\right\}.$$

So that Ω, B, E, F, $E \cap F$ have Hausdorff dimension equal to 1 while $F^c = \left\{\frac{1}{4}, \frac{1}{2}\right\}$ has Hausdorff dimension equal to 0.

Then by Theorem 1 the conditional probability of the following events given B is defined by the 1-dimensional Hausdorff measure and we obtain

$$P(E|B) = \frac{h^1(E \cap B)}{h^1(B)} = \frac{1}{3}, P(E \cap F|B) = \frac{h^1(E \cap F \cap B)}{h^1(B)} = \frac{1}{3}, P(F|B) = \frac{h^1(F \cap B)}{h^1(B)} = 1$$

while

$$P(F^c|B) = \frac{h^1(F^c \cap B)}{h^1(B)} = 0$$

the probability of the event F^c, the complement of F, i.e. the event that is true if F is false, given B is equal to zero because F^c is a finite set and so its Hausdorff dimension is less than the Hausdorff dimension of the conditioning event B, which represent information given about Linda.

Linda's Problem shows that the conditional probability defined by Hausdorff measures represents the preference ordering of the participants to the experiment, in fact the events E, $E \cap F$ given B have the same conditional probability and so they are equivalent and the event $F \mid B$, i.e. Linda is a feminist given information represented by the event B, has conditional probability 1 and its complement $F^c \mid B$, i.e. Linda is not a feminist given B, has conditional probability equal to zero. It occurs because the set F^c has Hausdorff dimension equal to 0 because it is a finite set, while the conditioning has Hausdorff dimension equal to 1.

We note that the unconditional probability of the events which are involved in Linda's Problem is

$$P(E|\Omega) = \frac{h^1(E\cap\Omega)}{h^1(\Omega)} = \frac{1}{4}, P(E\cap F|\Omega) = \frac{h^1(E\cap F\cap\Omega)}{h^1(\Omega)} = \frac{1}{4}$$

$$P(F|\Omega) = \frac{h^1(F\cap\Omega)}{h^1(\Omega)} = \frac{3}{4} \text{ and } P(F^c|\Omega) = \frac{h^1(F^c\cap\Omega)}{h^1(\Omega)} = 0$$

We can observe that different conditional probabilities arise from different conditioning events B and Ω.

6 Preference orderings represented by coherent upper and lower conditional probabilities in Linda's problem

In this section Linda's Problem is analyzed by using upper and lower conditional probabilities. The involved events are supposed to be represented by non-measurable sets. According to the results proven in Section 3, coherent upper conditional previsions are considered to assess equivalences, since they cannot represent a strict partial order because they do not satisfy the antisymmetric property (see Example 1); coherent lower condition previsions are considered to represent prefer-ence orderings since they satisfy the antisymmetric property as proven in Proposition 1.

Let $S = \{\emptyset, \Omega, E\cap F|B, E|B, F^c|B\}$; let μ be the monotone set function on S such that μ $(E\cap F|B) = \frac{1}{3}, \mu(E|B) = \frac{1}{3}, \mu(F^c|B) = \frac{3}{4}, \mu(\emptyset) = 0, \mu(\Omega) = 1$. Then, we extend the mono-tone set function μ to any subset A of Ω by its outer and inner measures, respectively defined by

$$\overline{P}(A) = \inf\{\mu(H) : H\supseteq A; H\in S\}\forall A\in\wp(\Omega)$$

and

$$\underline{P}(A) = \sup\{\mu(M) : M\subset A; M\in S\}\forall A\in\wp(\Omega).$$

They are related by the following equality:

$$\underline{P}(A) = \overline{P}(\Omega) - \overline{P}(A^c).$$

The monotone set function μ on S can be extended to any subset A of Ω by its outer measure and the conjugate lower conditional probability can be defined by the inner measure; for example the event $E^c\cap F \mid B$ does not belongto the class S so its upper conditional probability is equal to its outer measure; since the conditional event $E^c\cap F \mid B$ is the restriction of the event $E^c\cap F$ on B then its outer measure is equal to 1 since the only set in S containing $E^c\cap F\cap B$ is Ω.

So the following assessment can be considered:

	$E\cap F\|B$	$BF^C\|B$		$(E\cap E^c\|B$	$F\|B$	
	$F\|B$			$F)^c\|B$		
$\overline{P}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{1}$
$\underline{P}$	0	0	0	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{4}$

The conjunction fallacy does not occur because

$\underline{P}(E \cap F|B) \leq \underline{P}(E|B)$ and $\overline{P}(E \cap F|B) \leq \overline{P}(E|B)$.

Moreover, we obtain

$$\overline{P}(E \cap F|B) = \frac{1}{3} = \overline{P}(E|B)$$

and according to Definition 3 the indicator function I_F is preferable to the indicator function I_{F^c} since

$$\overline{P}(I_F - I_{F^c}|B) = 2\overline{P}(F|B) - 1\overline{P}(\Omega|B) > 0$$

and

$$\underline{P}(F^c|B) < \underline{P}(F|B)$$

The previous equalities and disequalities represent the result in Linda experiment, that is the events $E \cap F$ and E are equivalent but the majority of the participants assess a greater probability to the event F (Linda is feminist) than to the event F^c given the information B they receive about Linda.

This last result can be interpreted as describing a situation in which first System 1 assesses equivalence between the events $E \cap F$, E and then System 2, whose activity is represented by the lower conditional prevision, assesses the preferences expressed by the majority of the participants to the experiment.

Since

$$\underline{P}(E|B) = \underline{P}(F|B) = \underline{P}(E \cap F|B) = 0$$

the events $E \cap F$, E and F cannot be compared with respect to the lower conditional probability $\underline{P}(\cdot|B)$ according Definition 3 because they have probability zero given B and so we are not able to decide if they are admissible in the class S according to Definition 7 and which of them is maximal under the lower conditional probability according to Definition 8 .

It is interesting to note that if we firstly extend the assessment of μ on S by the inner measure and then we calculate the coherent upper conditional prevision by the conjugate property we cannot describe the preference of the majority in Linda's experiment. In fact in this case we obtain the following assessment:

$E \cap E\|BF^c\|B$	$(E \cap E^c\|BF\|B$		
$F\|B$	$F)^c\|B$		
$\overline{P}$ 1 1 1	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{4}$

which implies

$$\underline{P}(I_F - I_{F^c}|B) = 2\underline{P}(F|B) - \underline{P}(\Omega|B) < 0$$

and

$$\underline{P}(F|B) < \underline{P}(F^c|B)$$

so that, according to Definition 3, the indicator function I_F is not preferable to the indicator function I_{F^c}

This result could confirm that the preference of the majority of the participants in Linda's experiment are caused by System 1 faster than System 2.

7 Conclusions

In this paper coherent lower and upper conditional previsions are proposed as a mathematical tool to describe the two aspects of human brain's activity: the conscious/asymmetrical and unconscious/symmetrical thought. In fact the two non-linear functionals represent different binary relations between random variables; one of them, represented by the coherent lower previsions, satisfies the antisymmetric property which is not satisfied by the binary relation represented by their conjugate coherent upper conditional previsions. So coherent lower and upper conditional previsions can be used to represent respectively the partial strict preference order, which is the result of the conscious thought, and the equivalences assigned by the unconscious thought. When preference orderings and equivalence are represented by non linear functions may occur that it is not possible to obtain an optimal choice, that is to determine a Bayes random variable, which is a maximal random variable with respect to a linear prevision. According to the approach proposed in this paper to obtain an optimal choice means that the conscious and the unconscious thought lead to the same decision. The model of coherent upper and lower conditional previsions based on Hausdorff outer and inner measures assures that the optimal decision can be made if measurable sets are considered as it is shown for Linda's Problem. Hausdorff outer measures are defined in a metric space and examples of measurable sets with respect to Hausdorff outer measures are Borelian sets, which are sets generated by the open sets. If $|\Omega| \geq 2$ then any metric does not induce the topology $\mathcal{T} = \{\emptyset, \Omega\}$ since in a metric space points are closed sets while in $\mathcal{T}$ the only closed sets are $\emptyset$ and Ω. So that the class of the Borelian sets contains at least a pair of sets B and B^c different from the banal subsets of Ω. This result in general is not true for coherent upper an lower conditional probabilities which coincide only on the banal sets. According to this interpretation, by Theorem 3, we could conclude that if two random variables are indifferent with respect to the unconscious thought then one of them cannot be preferable to the other with respect to the conscious thought; by Theorem 4 we could obtain that if a random variable is preferable to another one with respect to the conscious mind then they cannot be indifferent with respect to the unconscious thought. Pathological situations can be obtained when two random variables are indifferent with respect to the unconscious thought but one of them if preferable to the other one with respect to the conscious thought. Or alternatively a random variable is preferable to another one with respect to the conscious thought but the two random variable are indifferent with respect to the unconscious thought.

These situations can be captured by the model because in this case the lower conditional prevision should be greater than the upper conditional prevision. The updating model based on Hausdorff outer and inner measures can represent respectively the awareness process of the unconscious and conscious thought which depend on unexpected events in all cases.

References

1. Billingsley, P.: Probability and Measure. Wiley (1986)
2. Choquet, G.: Theory of capacities. Annales de l'Institut Fourier. **5**, 131–295 (1953)
3. de Finetti, B.: Probability Induction and Statistics. Wiley, New York (1970)
4. de Finetti, B.: Theory of Probability. Wiley (1974)
5. Denneberg, D.: Non-additive measure and integral. Kluwer Academic Publishers. (1994)
6. Doria, S.: Probabilistic independence with respect to upper and lower conditional probabilities assigned by Hausdorff outer and inner measures. Int. J. Approx. Reason. **46**, 617–635 (2007)
7. Doria, S.: Characterization of a coherent upper conditional prevision as the Choquet integral with respect to its associated Hausdorff outer measure. Ann. Oper. Res. 33–48 (2012)
8. Doria, S.: Symmetric coherent upper conditional prevision by the Choquet integral with respect to Hausdorff outer measure. Ann. Oper. Res. 377–396 (2014)
9. Doria, S.: On the disintegration property of a coherent upper conditional prevision by the Choquet integral with respect to its associated Hausdorff outer measure. Ann. Oper. Res. **216**(2), 253–269 (2017)
10. Doria, S., Dutta, B., Mesiar, R.: Integral representation of coherent upper conditional prevision with respect to its associated Hausdorff outer measure: a comparison among the Choquet integral, the pan-integral and the concave integral. Int. J. Gen. Syst. **216**(2), 569–592 (2018)
11. Doria, S.: Preference orderings represented by coherent upper and lower previsions. Theor. Decis. **87**, 233–259 (2019)
12. Doria, S.: Representation of indifference and preference ordering between random variables by coherent upper and lower conditional previsions defined by Hausdorff outer and inner measures, Proceedings of the Thirty-Third International Florida Artificial Intelligence Research Society Conference, AAAI Press, 613-616, (2020)
13. Doria, S., Cenci, A.: Modeling decision in AI: Re-thinking Linda in terms of coherent lower and upper conditional previsions. Modeling decision in Artificial Intelligence, Lectures Notes in Artificial Intelligence, Torra, V., Narukawa, Y., Nin, J., Agell,N. (Eds.), Springer, (2020)
14. Doria, S., Mesiar, R., Šeliga, A.: Construction Method of Coherent Lower and Upper Previsions Based on Collection Integrals. Accepted in *Bollettino dell'Unione Matematica Italiana*, (2020),
15. Doria, S., Mesiar, R., Šeliga, A.: Integral representation of coherent lower previsions by super-additive integrals. Axioms **9**(2), art.no. 43, (2020)
16. Falconer, K.J.: The Geometry of Fractal Sets. Cambridge University Press (1986)
17. Gilio, A., Over, D.: The psychology of inferring conditionals from disjunctions: a probabilistic study. Journal of Mathematical Psychology. **56**(2), 118–131 (2012)
18. Hartmann, S., Meijs, W.: Walter the banker: the conjunction fallacy reconsidered. Synthese. **184**(1), 73–87 (2012)
19. Kahneman, D., Tversky, A.: Prospect Theory: An Analysis of Decision under Risk. Econometrica. **47**(2), 263–291 (1979)
20. Kahneman, D.: Thinking. Fast and Slow, Farrar, Straus and Giroux (2011)
21. Matte Blanco, I.: The Unconscious as Infinite Sets: an Essay on Bi-Logic. Gerald Duckworth, London (1975)
22. Pothos, E.M., Busemeyer, J.R.: Can quantum probability provide a new direction for cognitive modeling? Behav. Brain Sci. **36**, 255–327 (2013)
23. Regazzini, E.: De Finetti's coherence and statistical inference. Ann. Stat. **15**(2), 845–864 (1987)
24. Rogers, C.A.: Hausdorff Measures. Cambridge University Press (1970)
25. Tentori, K., Bonini, N., Osherson, D.: The conjunction fallacy:a misunderstanding about conjunction? Cognit.Sci. **28**, 467–477 (2004)
26. Tentori, K., Crupi, V., Russo, S.: On the determinants of the conjunction fallacy:probability versus inductive confirmation. J. Exp. Psychol. Gen. **142**, 235–255 (2013)

27. Tversky, A., Kahnemann, D.: Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. Psycol Rev. **90**, 4 (1983)
28. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London, (1991)