The Generalization of the Conjunctive Rule for Aggregating Contradictory Sources of Information Based on Generalized Credal Sets

Andrey G. Bronevich
National Research University
Higher School of Economics
Moscow, Russia
brone@mail.ru

Igor N. Rozenberg
JSC "Research, Development and Planning Institute for Railway Information Technology, Automation and Telecommunication"
Moscow, Russia
I.Rozenberg@gismps.ru

Abstract

In the paper we consider the generalization of the conjunctive rule in the theory of imprecise probabilities. Let us remind that the conjunction rule, produced on credal sets, gives their intersection and it is not defined if this intersection is empty. In the last case the sources of information are called contradictory[1]. Meanwhile, in the Dempster-Shafer theory it is possible to use the conjunctive rule for contradictory sources of information having as a result a non-normalized belief function that can be greater than zero at empty set. In the paper we try to exploit this idea and introduce into consideration so called generalized credal sets allowing to model imprecision (non-specificity), conflict, and contradiction in information. Based on generalized credal sets the conjunctive rule is well defined for contradictory sources of information and it can be conceived as the generalization of the conjunctive rule for belief functions. We also show how generalized credal sets can be used for modeling information when the avoiding sure loss condition is not satisfied, and consider coherence conditions and natural extension based on generalized credal sets.

Keywords. Imprecise probabilities, conjunctive rule, generalized credal sets, contradictory sources of information.

1 Introduction

In the theory of imprecise probabilities [18, 7, 1] there are many models for describing uncertainty: credal sets, upper and lower probabilities, lower and upper coherent previsions, sets of desirable gambles, etc. But in any case, we can equivalently represent the information with the help of sets of probability measures. As one can check, up to now there are no many works concerning the case when the available information is contradictory, i.e. the avoiding sure loss condition is not satisfied.

[1] We will use next the term "contradictory" because the traditional term "conflict" is also used by identifying another type of uncertainty described by probability measures.

By the way, in evidence theory [15, 8, 16] there is a possible way to describe contradiction based on transferable belief model. In this model we can describe contradictory information by assigning non-zero values to the corresponding belief function at empty set[2]. In this paper we will try to exploit this idea that leads to some generalizations of the theory of imprecise probabilities, in particular based on this idea it is possible to extend the conjunctive rule (C-rule) for aggregating belief functions for more general theories of imprecise probabilities [3, 4].

Let us notice that in the literature one can find results concerning the aggregation rules for imprecise probabilities [17, 9, 14, 13]. The rule from [17] deals with lower previsions and generalizes the pooling method for aggregation of probability measures. In [9] the aggregation rule is based on an idea that non-conflicting information should be aggregated in conjunctive manner and conflicting information should be aggregated in disjunctive manner. In [14] the proposed aggregation rules are based on modeling the interaction among expert's opinions. Authors of [13] try to get the aggregation rule for credal sets with properties close to the C-rule but their rule is based on some heuristic algorithmic procedure.

The paper has the following structure. Sections 2 and 3 remind some definitions from the theory of monotone measures, belief functions and the theory of imprecise probabilities. Then in Sections 4 and 5 we describe the basic rules of aggregation in general theories of imprecise probabilities and investigate the connection of these rules to the combination rules in evidence theory. After that we try to generalize the C-rule firstly (Section 6) for probability measures, and secondly (Section 7) for general models of imprecise probabilities using so-called generalized credal sets. Based on generalized credal sets it is possible to model contradiction in information and introduce analogous notions and constructions as in traditional theory of imprecise probabilities like coherence and natural extension, as shown in Section 8.

[2] This statement will be clarified in the next sections.

2 Some Definitions and Notations from the Theory of Non-additive Measures

Let X be a non-empty finite set and let 2^X be the power set of X. We will consider set functions on the algebra 2^X of various types: monotone measures, probability measures, lower and upper probabilities. A set function $\mu : 2^X \to [0,1]$ is called

1) *normalized* if $\mu(\varnothing) = 0$ and $\mu(X) = 1$;
2) *monotone* if $A, B \in 2^X$ and $A \subseteq B$ implies $\mu(A) \le \mu(B)$;
3) *additive* if $\mu(A) + \mu(B) = \mu(A \cap B) + \mu(A \cup B)$ for all $A, B \in 2^X$;
4) *2-monotone* if $\mu(A) + \mu(B) \le \mu(A \cap B) + \mu(A \cup B)$ for all $A, B \in 2^X$;
5) *2-alternating* if $\mu(A) + \mu(B) \ge \mu(A \cap B) + \mu(A \cup B)$ for all $A, B \in 2^X$;
6) a *monotone measure* if it is monotone and normalized;
7) a *probability measure* if it is additive and normalized;
8) a *belief function* if there is non-additive set function $m : 2^X \to [0,1]$ called the *basic belief assignment* (bba) such that $\sum_{A \in 2^X} m(A) = 1$ and $\mu(B) = \sum_{A \subseteq B} m(A)$.

The following operations on set functions are defined:

a) convex sum: $\mu = a\mu_1 + (1-a)\mu_2$, where $a \in [0,1]$, and $\mu(A) = a\mu_1(A) + (1-a)\mu_2(A)$ for all $A \in 2^X$;
b) $\mu_1 \le \mu_2$ if $\mu_1(A) \le \mu_2(A)$ for all $A \in 2^X$;
c) μ^d is the dual of μ if $\mu^d(A) = 1 - \mu(\overline{A})$ for all $A \in 2^X$, and $\overline{A}$ denotes the compliment of A.

Let us remind that the theory of evidence models uncertainty with the help of belief functions. In this theory (e.g. transferable belief model) we describe contradiction using non-normalized belief functions, i.e. it is possible that $Bel(\varnothing) > 0$ for belief function Bel. Let Bel be a belief function with the bba m. Then

- the set $A \in 2^X$ is a *focal element* for Bel if $m(A) > 0$;
- the set of all focal elements is called the *body of evidence*;
- Bel is called *categorical* if its body of evidence contains only one focal element. Any categorical belief function $\eta_{\langle B \rangle}$ with focal element B can be computed as

$$\eta_{\langle B \rangle}(A) = \begin{cases} 1, & B \subseteq A, \\ 0, & \text{otherwise.} \end{cases}$$

- Bel is a probability measure iff $m(A) = 0$ for all $A \in 2^X$ with $|A| \ge 2$. In this paper we also consider non-normalized probability measures P for which $P(\varnothing) > 0$.
- any belief function μ has the following representation through categorical belief functions:

$$Bel = \sum_{B \in 2^X} m(B)\eta_{\langle B \rangle}.$$

In the sequel we will use the following notations:

- M_{pr} is the set of all probability measures on 2^X and $\overline{M}_{pr}$ be the set of all probability measures including also non-normalized probability measures.
- M_{bel} and $\overline{M}_{bel}$ are sets of all belief functions on 2^X and the bar indicates that belief functions from $\overline{M}_{bel}$ may be non-normalized.
- M_{mon} is the set of all monotone measures on 2^X.
- $M_{2\text{-}mon}$ is the set of all 2-monotone measures on 2^X.
- if M is a family of set functions, then we denote $M^d = \{\mu^d | \mu \in M\}$. For example, M^d_{bel} denotes the set of all plausibility functions, which are dual to belief functions, or $M^d_{2\text{-}mon}$ is the set of all 2-alternating measures on 2^X.

3 Models of Imprecise Probabilities: Lower and Upper Probabilities and Credal Sets

Assume that $\mu : 2^X \to [0,1]$ is a set function that gives us lower bounds of probabilities. Then this function *avoids sure loss* iff there is a probability measure $P \in M_{pr}$ such that $\mu \le P$. If the avoiding sure loss condition is not fulfilled, then the information described by μ is *contradictory*. Any non-contradictory lower probability function μ defines the non-empty set of probability measures

$$\mathbf{P}(\mu) = \{P \in M_{pr} | P \ge \mu\}$$

called the credal set. Generally, a set $\mathbf{P}$ of probability measures is called a *credal set* if it is convex and closed.

Analogously the model of upper probabilities is introduced. Let us suppose that $\nu : 2^X \to [0,1]$ gives us the upper bounds of probabilities. Then this function avoids sure loss iff there is a probability measure $P \in M_{pr}$ such that $\nu \ge P$. In this case we call an upper probability function non-contradictory and describe it by a credal set

$$\mathbf{P}(\nu) = \{P \in M_{pr} | P \le \nu\}.$$

We can equivalently replace the model based on lower probabilities by the model based on upper probabilities. For this purpose we transform any lower probability μ to the upper probability μ^d. It easy to show that

$$\left\{P \in M_{pr} | P \le \mu^d\right\} = \{P \in M_{pr} | P \ge \mu\},$$

i.e. the corresponding credal sets coincide.

Let us introduce also coherent lower and upper probabilities. A non-contradictory lower probability μ is called *coherent* if for any $A \in 2^X$ there exists $P \in M_{pr}$ such that $\mu(A) = P(A)$ and $\mu \le P$, in other words,

$$\mu(A) = \inf\{P(A) | P \in \mathbf{P}(\mu)\},$$

where $\mathbf{P}(\mu) = \{P \in M_{pr} | P \ge \mu\}$.

Analogously, a non-contradictory upper probability ν is called *coherent* if for any $A \in 2^X$ there exists $P \in M_{pr}$ such that $\nu(A) = P(A)$ and $\nu \geq P$, in other words,

$$\nu(A) = \inf\{P(A)|P \in \mathbf{P}(\nu)\},$$

where $\mathbf{P}(\nu) = \{P \in M_{pr}|P \geq \nu\}$.

Coherent lower probabilities and coherent upper probabilities are connected with the dual relation, i.e. if μ is a coherent lower probability then μ^d is the coherent upper probability. We can also generate a coherent lower probability μ and coherent upper probability ν using a credal set $\mathbf{P}$ by formulas:

$$\mu(A) = \inf\{P(A)|P \in \mathbf{P}\},$$
$$\nu(A) = \sup\{P(A)|P \in \mathbf{P}\},$$

where $A \in 2^X$, and obviously, $\nu = \mu^d$ in this case.

Let μ be a non-contradictory lower probability. Then we can improve lower bounds of probabilities using the natural extension. It is defined as

$$\mu_{coh}(A) = \inf\{P(A)|P \in \mathbf{P}(\mu)\},$$

where $A \in 2^X$. Clearly, μ_{coh} is a coherent lower probability.

Let us remind that any credal set can be equivalently defined with the help of lower previsions. Let K' be a subset of the set K of all real functions of the type $f : X \to \mathbb{R}$. In some cases we assume that $K' = K$. Then *lower previsions* on K' are defined by the functional $\underline{E} : K' \to \mathbb{R}$. This functional defines the credal set

$$\mathbf{P}(\underline{E}) = \left\{P \in M_{pr} \middle| \forall f \in K' : \sum_{x\in X} f(x)P(\{x\}) \geq \underline{E}[f]\right\}.$$

If the credal set $\mathbf{P}(\underline{E})$ is empty then lower previsions do not satisfy the avoiding sure loss condition and we say that lower previsions contain contradiction. In some sense lower previsions can be understood as lower bounds of expectations of random variables in K'. The model based on lower previsions is more general than the model based on lower probabilities because we obtain the last model if we assume that $K' = \{1_A\}_{A\in 2^X}$, where

$$1_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

is the characteristic function of the set A. We can improve the lower bounds of expectations using the procedure called the natural extension

$$\underline{E}'[f] = \inf\left\{\sum_{x\in X} f(x)P(\{x\}) \middle| P \in \mathbf{P}(\underline{E})\right\}.$$

Note that this procedure is not defined if $\mathbf{P}(\underline{E}) = \varnothing$. Let us remind that the functional $\underline{E}$ defines coherent lower previsions if $\underline{E}'[f] = \underline{E}[f]$ for all $f \in K'$.

Analogously, upper previsions are introduced. Any functional $\overline{E} : K' \to \mathbb{R}$ can be conceived as upper previsions. The upper previsions are not contradictory (or avoid sure loss) iff the credal set

$$\mathbf{P}(\overline{E}) = \left\{P \in M_{pr} \middle| \forall f \in K' : \sum_{x\in X} f(x)P(\{x\}) \leq \overline{E}[f]\right\}.$$

is not empty. We can improve the upper bounds of expectations using the natural extension:

$$\overline{E}'[f] = \sup\left\{\sum_{x\in X} f(x)P(\{x\}) \middle| P \in \mathbf{P}(\overline{E})\right\}.$$

If $\overline{E}'[f] = \overline{E}[f]$ for all $f \in K'$, then $\overline{E}$ is a coherent lower prevision. Let us notice that we can equivalently describe uncertain information by lower or upper previsions. If the functional $\underline{E} : K' \to \mathbb{R}$ describes the lower previsions then we can equivalently describe the same information by upper previsions defined by

$$\overline{E}[f] = -\underline{E}[-f]$$

for all $-f \in K'$.

4 The Conjunctive and Disjunctive Rules for Aggregating Sources of Information

Consider n sources of information described by credal sets $\mathbf{P}_1, \ldots, \mathbf{P}_n$. Then there are several possible ways for aggregating this information that depends on prior assumptions. If we suppose that each source of information is reliable then we can aggregate them using intersection of the corresponding sets:

$$\mathbf{P} = \mathbf{P}_1 \cap \cdots \cap \mathbf{P}_n.$$

This rule of aggregation is called the *conjunctive rule (C-rule)*. It is easy to see that if we describe credal sets with the help of lower probability functions $\mu_1, \ldots, \mu_n$, then C-rule can be represented as

$$\mu = \mu_1 \vee \cdots \vee \mu_n,$$

where $\vee$ is the maximum operation. The last formula is justified because in this case

$$\mathbf{P}(\mu) = \mathbf{P}(\mu_1) \cap \cdots \cap \mathbf{P}(\mu_n)$$

If we describe sources of information by upper probabilities $\mu_1, \ldots, \mu_n$, then the C-rule is clearly expressed with the minimum operation $\wedge$ as

$$\mu = \mu_1 \wedge \cdots \wedge \mu_n$$

Analogously, the conjunctive rule is expressed in models based on lower previsions $E_i : K' \to \mathbb{R}$, $i = 1, \ldots, n$, or upper previsions $\overline{E}_i : K' \to \mathbb{R}$, $i = 1, \ldots, n$, as

$$\underline{E} = \underline{E}_1 \vee \cdots \vee \underline{E}_n, \quad \overline{E} = \overline{E}_1 \wedge \cdots \wedge \overline{E}_n. \tag{1}$$

We would like to emphasize that there are other rules for aggregation of information sources. If we know that at least one source of information is reliable and all sources of information are represented by credal sets $\mathbf{P}_1,\dots,\mathbf{P}_n$, then we can use the *disjunctive rule*, in which the result is the minimal credal set $\mathbf{P}$ that contains the corresponding credal sets $\mathbf{P}_i$, $i=1,\dots,n$. This disjunctive rule is expressed through lower previsions $E_i : K' \to \mathbb{R}$, $i=1,\dots,n$, or upper previsions $\overline{E}_i : K' \to \mathbb{R}$, $i=1,\dots,n$, as

$$\underline{E} = \underline{E}_1 \wedge \cdots \wedge \underline{E}_n, \quad \overline{E} = \overline{E}_1 \vee \cdots \vee \overline{E}_n.$$

The mixture rule can be used if we can evaluate the reliability of information. Let us assume this reliability is given by non-negative numbers a_i, $i=1,\dots,n$, such that $\sum_{i=1}^n a_i = 1$. Then we can aggregate sources of information described by credal sets $\mathbf{P}_i$, $i=1,\dots,n$, as

$$\mathbf{P} = \left\{ \sum_{i=1}^n a_i P_i \middle| P_i \in \mathbf{P}_i, i=1,\dots,n \right\}.$$

The counterparts of this rule for lower previsions $E_i : K' \to \mathbb{R}$, $i=1,\dots,n$, or upper previsions $\overline{E}_i : K' \to \mathbb{R}$, $i=1,\dots,n$, are

$$\underline{E} = \sum_{i=1}^n a_i \underline{E}_i \quad \text{or} \quad \overline{E} = \sum_{i=1}^n a_i \overline{E}_i.$$

Let us notice that other possible rules of aggregation have properties that more or less similar to the considered rules.

Let us observe that applying the C-rule is possible if the resulting credal set is not empty. In the opposite case we say that there is contradiction among sources of information. Meanwhile, in evidence theory the C-rule is also applicable if the sources of information are contradictory. In the next section we will introduce such C-rules, considered in [4], and give some hints about how they can be generalized in the theory of imprecise probabilities.

5 Conjunctive Rules of Aggregation in Evidence Theory, the Order of Specialization

Let $Bel_1 = \sum_{A \in 2^X} m_1(A)\eta_{\langle A \rangle}$ and $Bel_2 = \sum_{B \in 2^X} m_2(B)\eta_{\langle B \rangle}$ be belief functions. Then the *conjunctive combination rule* (C-rule)[3] [10, 4] is defined by

$$Bel = \sum_{A,B \in 2^X} m(A,B)\eta_{\langle A \cap B \rangle},$$

where $m : 2^X \times 2^X \to [0,1]$ is such that

$$\begin{cases} \sum_{B \in 2^X} m(A,B) = m_1(A), & A \in 2^X, \\ \sum_{A \in 2^X} m(A,B) = m_2(B), & B \in 2^X. \end{cases} \tag{2}$$

[3] In [4] such combination rules are called generalized Dempster-Shafer rules.

Observe that we get the classical C-rule [8] if $m(A,B) = m_1(A)m_2(B)$ for any $A,B \in 2^X$. The use of such general rule can be explained using the interpretation of belief functions through random sets. A random set ξ is a random variable taking its values in 2^X. Any such random variable can be defined by probabilities $P(\xi = A)$ and these probabilities can be identified with values $m(A)$ in evidence theory. Given two random sets ξ_1 and ξ_2 with values in 2^X. If we assume that these random sets are independent, then

$$P(\xi_1 = A, \xi_2 = B) = P(\xi_1 = A)P(\xi_2 = B).$$

The using of the classical C-rule means that from two sources of information described by independent random sets ξ_1 and ξ_2 we obtain a new random set ξ defined by

$$P(\xi = C) = \sum_{A \cap B = C} P(\xi_1 = A)P(\xi_2 = B).$$

Thus, the generalization of the classical C-rule can be got if we suppose that random sets ξ_1 and ξ_2 can be dependent. In this case we can only guarantee that the non-negative set function $m(A,B) = P(\xi_1 = A, \xi_2 = B)$ obeys (2).

Let us notice that the C-rule is not uniquely defined and it can be also applied in a case, when the sources of information are contradictory. The ways of choosing optimal conjunctive combination rules according to several justified criteria can be found in [4]. The main conclusion from [4] is that an optimal C-rule should be chosen among Pareto optimal C-rules w.r.t. the partial order on belief functions called specialization.

Let Bel_1 and Bel_2 be belief functions with bbas m_1 and m_2. We write $Bel_1 \preceq Bel_2$ if Bel_2 can be obtained from Bel_1 using a linear contraction transform $\Phi : 2^X \times 2^X \to [0,1]$, i.e. $m_2(B) = \sum_{A \in 2^X} \Phi(A,B)m_1(A)$, and the set function $\Phi : 2^X \times 2^X \to [0,1]$ has the following properties: $\sum_{B \in 2^X} \Phi(A,B) = 1$ for any $B \in 2^X$ and $\Phi(A,B) = 0$ if $B \not\subseteq A$. The partial order $\preceq$ is called *specialization*. It is easy to show [11] that $Bel_1 \preceq Bel_2$ implies $Bel_1 \leq Bel_2$, but the opposite is not true in general. The main results [4] showing the connections of C-rules and the order $\preceq$ are given in the next propositions.

Proposition 1 *If Bel is the result of a C-rule applied to* $Bel_1, Bel_2 \in \overline{M}_{bel}$*, then* $Bel_1 \preceq Bel$ *and* $Bel_2 \preceq Bel$*. Furthermore, each minimal element of the set*

$$\mathbf{Bel}(Bel_1, Bel_2) = \left\{ Bel \in \overline{M}_{bel} | Bel_1 \preceq Bel, Bel_2 \preceq Bel \right\}$$

w.r.t. the order $\preceq$ *for arbitrary* $Bel_1, Bel_2 \in \overline{M}_{bel}$ *can be obtained by a C-rule.*

This result shows that the optimal choice of a C-rule should be made to get the best approximation of the set function $\max\{Bel_1, Bel_2\}$ and this choice should obviously be made in the set of minimal elements of $\mathbf{Bel}(Bel_1, Bel_2)$ w.r.t. $\preceq$ that can be obtained by so called Pareto optimal C-rules.

Proposition 2 *The order $\preceq$ is equivalent to the order $\leq$ on the set $\overline{M}_{pr}$. In addition if $Bel \leq P$ for $P \in \overline{M}_{pr}$ and $Bel \in \overline{M}_{Bel}$, then $Bel \preceq P$. Furthermore,*

$$Bel(A) = \inf\{P(A) | P \in \mathbf{P}(Bel)\},$$

where $\mathbf{P}(Bel) = \left\{P \in \overline{M}_{pr} \middle| Bel \preceq P\right\}$.

Remark 1 Proposition 2 shows that in evidence theory any belief function can be equivalently represented by $\mathbf{P}(Bel)$ that may be called a generalized credal set. Such a construction with a slightly different definition will be introduced in the next section. Clearly, the above proposition allows us to write

$$\mathbf{P}(Bel) = \left\{P \in \overline{M}_{pr} \middle| Bel \leq P\right\}.$$

Let $Bel_1, Bel_2 \in \overline{M}_{bel}$. Then we denote by $R(Bel_1, Bel_2)$ the set of all possible belief measures that can be obtained by C-rules applied to Bel_1 and Bel_2. Then the amount of contradiction between Bel_1 and Bel_2 by C-rules can be computed as

$$Con(Bel_1, Bel_2) = \inf\left\{Bel(\varnothing) \middle| Bel \in R(Bel_1, Bel_2)\right\}.$$

Let us observe that this measure of contradiction (or conflict) is considered in many papers [4, 5, 6, 10], where authors show that $Con(Bel_1, Bel_2)$ has better properties than a measure of contradiction based on the classical C-rule.

Proposition 3 *Let* $\mathbf{P}(Bel_i) = \left\{P \in \overline{M}_{pr} \middle| Bel_i \leq P\right\}$, *where* $Bel_i \in \overline{M}_{bel}$, $i = 1, 2$. *Then*

$$Con(Bel_1, Bel_2) = \inf\{P(\varnothing) | P \in \mathbf{P}(Bel_1) \cap \mathbf{P}(Bel_2)\}.$$

Thus, in this section we has shown that it is possible to extend the model of non-normalized belief functions on more general theories of imprecise probabilities using generalized credal sets, and this problem will be investigated in the next sections.

6 The Conjunctive Rule for Probability Measures Admitting Contradiction

Let us consider the case when we have two sources of information described by probability measures P_1 and P_2. These sources of information are absolutely contradictory if we can divide the space X on two disjoint subsets A and B such that $P_1(A) = 1$ and $P_2(B) = 1$. In other words, sources of information support that events A and B are certain, but it is not possible because these events are disjoint. In classical logic false implies anything, thus we can write

$$P_1 \wedge P_2 = \bigwedge_{P_i \in M_{pr}} P_i = \eta^d_{\langle X \rangle},$$

where $\eta^d_{\langle X \rangle}$ describes the result of conjunction of all possible probability measures on 2^X. Now we will try to generalize the above rule for two probability measures that are not absolutely contradict each other. In this case we can divide probability measures on 2 parts:

$$P_1 = (1-a)P_1^{(1)} + aP_1^{(2)}, \quad P_2 = (1-a)P_2^{(1)} + aP_2^{(2)},$$

where $a \in [0,1]$, $P_k^{(i)} \in M_{pr}$, $i = 1, 2$, $k = 1, 2$, and $P_1^{(1)}$, $P_2^{(1)}$ are parts of probability measures that don't contradict each other, i.e. $P_1^{(1)} = P_2^{(1)}$, and probability measures $P_1^{(2)}$, $P_2^{(2)}$ are absolutely contradict each other. The value

$$Con(P_1, P_2) = a = 1 - \sum_{x_i \in X} \min\{P_1(\{x_i\}), P_2(\{x_i\})\}$$

is called the *amount of contradiction* and the above measures are defined by the following formulas:

$$P_1^{(1)}(\{x_i\}) = P_2^{(1)}(\{x_i\}) = \frac{1}{1-a} \min\{P_1(\{x_i\}), P_2(\{x_i\})\},$$

where $x_i \in X$ and $a < 1$ (if $a = 1$, then a measure $P_1^{(1)} = P_2^{(1)}$ is defined arbitrary);

$$P_1^{(2)}(\{x_i\}) = \frac{1}{a}\left(P_1(\{x_i\}) - (1-a)P_1^{(1)}(\{x_i\})\right),$$

$$P_2^{(2)}(\{x_i\}) = \frac{1}{a}\left(P_2(\{x_i\}) - (1-a)P_2^{(1)}(\{x_i\})\right),$$

where $x_i \in X$ and $a > 0$ (if $a = 0$, then absolutely contradictory measures $P_1^{(2)}$, $P_2^{(2)}$ are defined arbitrary).

Example 1 Assume that $X = \{x_1, x_2, x_3\}$. In this example any probability measure P can be described by a vector $\left(P(\{x_1\}), P(\{x_2\}), P(\{x_3\})\right)$. Let the probability measures P_1 and P_2 be defined by the following vectors: $P_1 = (0.4, 0.2, 0.4)$ and $P_2 = (0.2, 0.4, 0.4)$. Then $a = 0.2$, $P_1^{(1)} = P_2^{(1)} = (0.25, 0, 25, 0.5)$, $P_1^{(2)} = (1, 0, 0)$, and finally $P_2^{(2)} = (0, 1, 0)$.

Let us observe that measures $P_1^{(2)}$, $P_2^{(2)}$ are absolutely contradictory, because $P_1^{(2)}(\{x_1\}) = 1$ and $P_2^{(2)}(\{x_2\}) = 1$ for disjoint sets $\{x_1\}$ and $\{x_2\}$.

Summarizing we introduce the following definition.

Definition 1 The C-rule for probability measures $P_1, P_2 \in M_{pr}$ is defined as

$$P_1 \wedge P_2 = \sum_{x_i \in X} \min\{P_1(\{x_i\}), P_2(\{x_i\})\} \eta_{\langle \{x_i\} \rangle} + a\eta^d_{\langle X \rangle},$$

where $a = 1 - \sum_{x_i \in X} \min\{P_1(\{x_i\}), P_2(\{x_i\})\}$.

Example 2 Consider probability measures P_1 and P_2 from Example 1. Then

$$\begin{aligned} P_1 \wedge P_2 &= 0.8P_1^{(1)} + 0.2\eta^d_{\langle X \rangle} \\ &= 0.2\eta_{\langle \{x_1\} \rangle} + 0.2\eta_{\langle \{x_2\} \rangle} + 0.4\eta_{\langle \{x_1\} \rangle} + 0.2\eta^d_{\langle X \rangle}. \end{aligned}$$

Let $X = \{x_1, \ldots, x_n\}$. In the next we will describe the contradiction in information using measures of the type

$$P = \sum_{i=1}^{n} a_i \eta_{\langle \{x_i\} \rangle} + a_0 \eta^d_{\langle X \rangle}, \tag{3}$$

where $a_i \geq 0$, $i = 0, \ldots, n$, and $\sum_{i=0}^{n} a_i = 1$. Observe that $P \in M_{pr}$ if $a_0 = 0$ [4], and P is understood as a contradictory lower probability. If $a_0 > 0$, then the value a_0 gives us the amount of contradiction. The set of all possible measures, represented by (3), is denoted by $\overline{M}_{cpr}$. Let us notice that $M_{pr} \subseteq \overline{M}_{cpr}$.

Remark 2 Note that the set functions in $\overline{M}_{cpr}$ are plausibility functions with bba m such that $m(A) = 0$ if $1 < |A| < |X|$.

It is possible to describe the C-rule with the order $\leq$ on $\overline{M}_{cpr}$ considered as a partially ordered set.

Lemma 1 *Let $P_1, P_2 \in \overline{M}_{cpr}$ and $P_1 = \sum_{i=1}^{n} a_i \eta_{\langle \{x_i\} \rangle} + a_0 \eta^d_{\langle X \rangle}$, $P_2 = \sum_{i=1}^{n} b_i \eta_{\langle \{x_i\} \rangle} + b_0 \eta^d_{\langle X \rangle}$. Then $P_1 \leq P_2$ iff $a_i \geq b_i$, $i = 1, \ldots, n$.*

Corollary 1 *Let $P_1, \ldots, P_m \in \overline{M}_{cpr}$ and defined by*

$$P_k = \sum_{i=1}^{n} a_i^{(k)} \eta_{\langle \{x_i\} \rangle} + a_0^{(k)} \eta^d_{\langle X \rangle}$$

for $k = 1, \ldots, m$, then the exact upper bound of $P_1, \ldots, P_m$ in $\overline{M}_{cpr}$ is

$$P = \sum_{i=1}^{n} c_i \eta_{\langle \{x_i\} \rangle} + c_0 \eta^d_{\langle X \rangle},$$

where $c_i = \min\{a_i^{(1)}, \ldots, a_i^{(m)}\}$ for $i = 1, \ldots, n$, and $c_0 = 1 - \sum_{i=1}^{n} c_i$.

Remark 3 Corollary 1 implies that the C-rule of probability measures $P_1, P_2 \in M_{pr}$ is the exact upper bound of the set $\{P_1, P_2\}$. Thus, we define next the C-rule for arbitrary measures $P_1, \ldots, P_m \in \overline{M}_{cpr}$ as the exact upper bound of the set $\{P_1, \ldots, P_m\}$ in $\overline{M}_{cpr}$. This bound is denoted as $P_1 \wedge \cdots \wedge P_m$.

Example 3 Let we take probability measures P_1 and P_2 from Example 1, and the probability measure $P_3 = (0.4, 0.4, 0.2)$, then

$$P_1 \wedge P_2 \wedge P_3 = 0.2\eta_{\langle \{x_1\} \rangle} + 0.2\eta_{\langle \{x_2\} \rangle} + 0.2\eta_{\langle \{x_3\} \rangle} + 0.4\eta^d_{\langle X \rangle}.$$

7 Generalized Upper and Lower Credal Sets

Observe that using measures from $\overline{M}_{cpr}$ we can describe contradictory and conflicting information. Let us remind (see e.g. [12, 2] for details) that pure conflict is described by probability measures, and the theory of imprecise probabilities allows us to model conflict and non-specificity (imprecision) in information, and non-specificity is caused by uncertainty in choosing a "true probability measure" among possible alternatives. If we try to describe imprecise information with some contradiction and conflict we should consider subsets of $\overline{M}_{cpr}$. Let us observe the following. Let $P_1 \in \overline{M}_{cpr}$, then $P_2 \in \overline{M}_{cpr}$ with $P_2 \geq P_1$ can be used for describing the same information but with a greater amount of contradiction. Thus, the subset $\mathbf{P}$ in $\overline{M}_{cpr}$ describing imprecise information has to satisfy the following property:

a) $P_1 \in \mathbf{P}$, $P_2 \in \overline{M}_{pr}$, $P_1 \leq P_2$ implies that $P_2 \in \mathbf{P}$.

The next two properties are essential for the most models of imprecise probabilities (cf. credal sets).

b) If $P_1, P_2 \in \mathbf{P}$ then $aP_1 + (1-a)P_2 \in \mathbf{P}$ for any $P_1, P_2 \in \mathbf{P}$ and $a \in [0, 1]$.
c) The set $\mathbf{P}$ is closed in a sense that it can be considered as a subset of Euclidean space (any $P = a_0 \eta^d_{\langle X \rangle} + \sum_{i=1}^{n} a_i \eta_{\langle \{x_i\} \rangle}$ is a vector $(a_0, a_1, \ldots, a_n)$ in $\mathbb{R}^{n+1}$).

Summarizing we can introduce the following definition.

Definition 2 A subset $\mathbf{P} \subseteq \overline{M}_{cpr}$ is called an *upper generalized credal set* if it satisfies conditions a), b), c).

The C-rule for generalized upper credal sets can be defined analogously as for usual credal sets.

Definition 3 Let $\mathbf{P}_1, \ldots, \mathbf{P}_m$ be non-empty credal sets in $\overline{M}_{cpr}$. Then the credal set $\mathbf{P}$ produced by the C-rule is defined as $\mathbf{P} = \mathbf{P}_1 \cap \cdots \cap \mathbf{P}_m$.

Let us introduce new concepts that help to understand this definition. Let $\mathbf{P}$ be a credal set in $\overline{M}_{cpr}$. A subset consisting of all minimal elements in $\mathbf{P}$ is called the *profile* of $\mathbf{P}$ and it is denoted by $\mathit{profile}(\mathbf{P})$. Evidently, any profile uniquely defines the corresponding credal set. If $\mathbf{P}$ describes information without contradiction, then $\mathit{profile}(\mathbf{P})$ is a credal set in usual sense, i.e. $\mathit{profile}(\mathbf{P}) \subseteq M_{pr}$. In particular, if we have two credal sets $\mathbf{P}_1, \mathbf{P}_2$ in $\overline{M}_{cpr}$ with $\mathit{profile}(\mathbf{P}_i) \in M_{pr}$, $i = 1, 2$, then applying the C-rule gives us the profile:

$$\mathit{profile}(\mathbf{P}_1 \cap \mathbf{P}_2) = \mathit{profile}(\mathbf{P}_1) \wedge \mathit{profile}(\mathbf{P}_2).$$

Observe that any upper generalized credal set give us many lower possible bounds of probabilities and each possible value is characterized by contradiction. Let us denote the amount of contradiction in $P \in \overline{M}_{cpr}$ by $Con(P)$. Then to characterize the possible lower bounds of probabilities computed by an upper generalized credal set $\mathbf{P}$ we introduce into consideration the set function

$$\mu^r(A) = \inf\{P(A) | P \in \mathbf{P}, Con(P) \leq r\},$$

where $A \in 2^X$ and $r \in [0, 1]$ is the level of contradiction. The set function μ^r can be interpreted as a lower probability for the credal set $\mathbf{P}$ with a level of contradiction r.

Lemma 2 *For any upper generalized credal set $\mathbf{P}$:*

$$\mu^r(A) = \inf\{P(A) | P \in \mathit{profile}(\mathbf{P}), Con(P) \leq r\}.$$

[4] Observe that if $P \in M_{pr}$, then $P = \sum_{i=1}^{n} P(\{x_i\}) \eta_{\langle \{x_i\} \rangle}$.

Remark 4 We can consider the generalized upper credal sets whose profiles are credal sets in usual sense. In a case, when profiles of upper generalized credal sets are credal sets in usual sense, μ^r does not depend on r, and the considered model coincides with the model of imprecise probabilities based on usual credal sets.

Example 4 Let $X = \{x_1, x_2, x_3\}$. Then any

$$P = a_1\eta_{\langle\{x_1\}\rangle} + a_2\eta_{\langle\{x_2\}\rangle} + a_3\eta_{\langle\{x_3\}\rangle} + a_0\eta^d_{\langle X\rangle}$$

in $\overline{M}_{cpr}$ can be defined by the vector $P = (a_1, a_2, a_3, a_0)$. Consider upper generalized credal sets $\mathbf{P}_i$, $i = 1, 2, 3$, whose profiles are credal sets in usual sense:

$$\begin{aligned} &\mathit{profile}(\mathbf{P}_1) = \{aP_1 + (1-a)P_2 | t \in [0,1]\},\\ &\mathit{profile}(\mathbf{P}_2) = \{P_3\}, \quad \mathit{profile}(\mathbf{P}_3) = \{P_4\}, \end{aligned}$$

where

$$\begin{aligned} P_1 &= (2/3, 0, 1/3, 0), & P_2 &= (0, 2/3, 1/3, 0),\\ P_3 &= (1/3, 1/3, 1/3, 0), & P_4 &= (1/3, 1/2, 1/6, 0). \end{aligned}$$

Let us find the profile of $\mathbf{P}_1 \cap \mathbf{P}_2$. It obviously consists of minimal elements in the set

$$\begin{aligned} &\left\{P' \wedge P'' | P' \in \mathit{profile}(\mathbf{P}_1), P'' \in \mathit{profile}(\mathbf{P}_2)\right\} = \\ &\quad \{P | P = (1/3, t, 1/3, 1/3 - t), t \in [0, 1/3]\} \cup \\ &\quad\quad \{P | P = (t, 1/3, 1/3, 1/3 - t), t \in [0, 1/3]\}. \end{aligned}$$

The above set has only one minimal element, namely, $P_5 = (1/3, 1/3, 1/3, 0)$. Therefore, $\mathit{profile}(\mathbf{P}_1 \cap \mathbf{P}_2) = \{P_5\}$.

Analogously, let us find the profile of $\mathbf{P}_1 \cap \mathbf{P}_3$. It consists of minimal elements in the set

$$\begin{aligned} &\left\{P' \wedge P'' | P' \in \mathit{profile}(\mathbf{P}_1), P'' \in \mathit{profile}(\mathbf{P}_3)\right\} = \\ &\{P | P = (2t/3, 1/2, 1/6, 1/3 - 2t/3), t \in [0, 1/4]\} \cup \\ &\{P | P = (2t/3, 2(1-t)/3, 1/6, 1/6), t \in [1/4, 1/2]\} \cup \\ &\{P | P = (1/3, 2(1-t)/3, 1/6, 2t/3 - 1/6), t \in (1/2, 1]\}. \end{aligned}$$

The minimal elements of this set are $tP_6 + (1-t)P_7$, where $t \in [0, 1]$, and

$$P_6 = (1/6, 1/2, 1/6, 1/6), \quad P_7 = (1/3, 1/3, 1/6, 1/6).$$

Thus, $\mathit{profile}(\mathbf{P}_1 \cap \mathbf{P}_3) = \{tP_6 + (1-t)P_7 | t \in [0, 1]\}$.

Let us show next how it is possible to define lower bounds of expectation. Consider first expectations w.r.t. measures in $\overline{M}_{cpr}$. If $P \in M_{pr}$, then for any function $f : X \to \mathbb{R}$ we define the expectation $E_P(f)$ as

$$E_P(f) = \sum_{x \in X} f(x) P(\{x\}).$$

We can extend the functional E_P to the set of all measures in $\overline{M}_{cpr}$, using the considered interpretation of a measure $P \in \overline{M}_{cpr}$ through the C-rule. Obviously, $P = \wedge_{P_i \in M_{pr} | P_i \leq P} P_i$. Then this C-rule is expressed through expectations E_{P_i}, $P_i \leq P$, as (cf. formula (1))

$$\underline{E}_P = \bigvee_{P_i \in M_{pr} | P_i \leq P} E_{P_i}.$$

Lemma 3 *For any* $P = a_0\eta^d_{\langle X\rangle} + \sum_{i=1}^n a_i\eta_{\langle\{x_i\}\rangle}$ *and* $f : X \to \mathbb{R}$ *the value* $\underline{E}_P(f)$ *can be computed as*

$$\underline{E}_P(f) = a_0 \max_{x \in X} f(x) + \sum_{i=1}^n a_i f(x_i).$$

Let $\mathbf{P}$ be a credal set in $\overline{M}_{cpr}$. We will define first the lower expectation $\underline{E}_{\mathbf{P}}(f)$ for non-negative functions $f : X \to \mathbb{R}$. Let the set of all such functions be denoted by K^+. Because $\underline{E}_{\mathbf{P}}(f)$ is the lower expectation, we can define this value for any $f \in K^+$ as

$$\underline{E}_{\mathbf{P}}(f) = \inf_{P \in \mathbf{P}} \underline{E}_P(f).$$

Example 5 Let $\mathbf{P} = \mathbf{P}_1 \cap \mathbf{P}_3$, where $\mathbf{P}_1 \cap \mathbf{P}_3$ is defined in Example 4, then

$$\underline{E}_{\mathbf{P}}(f) = \min\left\{\underline{E}_{P_6}(f), \underline{E}_{P_7}(f)\right\},$$

where

$$\begin{aligned} \underline{E}_{P_6}(f) &= \frac{1}{6} f(x_1) + \frac{1}{2} f(x_2) + \frac{1}{6} f(x_3) + \frac{1}{6} \max_{x_i \in X} f(x_i),\\ \underline{E}_{P_7}(f) &= \frac{1}{3} f(x_1) + \frac{1}{3} f(x_2) + \frac{1}{6} f(x_3) + \frac{1}{6} \max_{x_i \in X} f(x_i). \end{aligned}$$

Let us indicate some properties of $\underline{E}_{\mathbf{P}}$ on K^+. In the next we denote by $\mathbb{R}^+$ the set of all non-negative real numbers. The function in K^+ with values equal to $a \in \mathbb{R}^+$ is denoted also by a. We write $f_1 \leq f_2$ for $f_1, f_2 \in K^+$ if $f_1(x) \leq f_2(x)$ for all $x \in X$.

Lemma 4 *The functional* $\underline{E}_{\mathbf{P}}$ *on* K^+ *has the following properties:*

1) $\underline{E}_{\mathbf{P}}(0) = 0$*;* $\underline{E}_{\mathbf{P}}(1) = 1$*;*

2) $\underline{E}_{\mathbf{P}}(f + a) = \underline{E}_{\mathbf{P}}(f) + a$ *for any* $f \in K^+$ *and* $a \in \mathbb{R}^+$*;*

3) $\underline{E}_{\mathbf{P}}(af) = a\underline{E}_{\mathbf{P}}(f)$ *for any* $f \in K^+$ *and* $a \in \mathbb{R}^+$*;*

4) $\underline{E}_{\mathbf{P}}(f_1) \leq \underline{E}_{\mathbf{P}}(f_2)$ *for* $f_1, f_2 \in K^+$ *if* $f_1 \leq f_2$*.*

Let us consider also the dual concept of generalized upper credal sets. In this case we describe uncertainty by set functions from the set $\overline{M}^d_{cpr}$. Any measure P in $\overline{M}^d_{cpr}$ is represented as

$$P = a_0\eta_{\langle X\rangle} + \sum_{i=1}^n a_i\eta_{\langle\{x_i\}\rangle},$$

where $a_i \geq 0$, $i = 0, \ldots, n$, and $\sum_{i=0}^n a_i = 1$, and it is conceived as an upper probability. The value a_0 shows the amount of contradiction. If $a_0 = 0$, then P is a probability measure. Evidently, measures from $\overline{M}^d_{cpr}$ describe conflict and contradiction in information and we can define the upper expectation $\overline{E}_P(f)$ for any $f \in K$ w.r.t. arbitrary P in $\overline{M}^d_{cpr}$ through the Choquet integral:

$$\overline{E}_P(f) = \int_X f(x)dP = a_0 \min_{x\in X} f(x) + \sum_{i=1}^{n} a_i f(x_i).$$

For describing conflict, contradiction and non-specificity with the help of measures in $\overline{M}^d_{cpr}$, we introduce the notion of lower generalized credal set.

Definition 4 A *lower generalized credal set* $\mathbf{P}$ is a non-empty subset of $\overline{M}^d_{cpr}$ with the following properties:

a) $P_1 \in \mathbf{P}$, $P_2 \in \overline{M}^d_{cpr}$, $P_1 \geq P_2$ implies that $P_2 \in \mathbf{P}$.
b) if $P_1, P_2 \in \mathbf{P}$, then $aP_1 + (1-a)P_2 \in \mathbf{P}$ for any $P_1, P_2 \in \mathbf{P}$ and $a \in [0,1]$.
c) $\mathbf{P}$ is closed set if we consider it as a subset of Euclidean space (any $P = a_0\eta^d_{\langle X\rangle} + \sum_{i=1}^n a_i \eta_{\langle\{x_i\}\rangle}$ is a vector $(a_0, a_1, \ldots, a_n)$ in $\mathbb{R}^{n+1}$).

The set of all maximal elements in a generalized lower credal set $\mathbf{P}$ is called the *profile* of $\mathbf{P}$ and it is denoted by *profile*$(\mathbf{P})$. Emphasize that generalized lower and upper credal sets are dual concepts, for instance, if $\mathbf{P}$ is a credal set in $\overline{M}_{cpr}$, then $\mathbf{P}^d$ is a credal set in $\overline{M}^d_{cpr}$; profiles of $\mathbf{P}$ and $\mathbf{P}^d$ are also connected with the dual relation: *profile*$(\mathbf{P})^d$ = *profile*$(\mathbf{P}^d)$; if $\mathbf{P}_1, \ldots, \mathbf{P}_m$ are credal sets in $\overline{M}_{cpr}$, then the expression for the C-rule is defined by the same way for the credal sets in $\overline{M}_{cpr}$ and $\overline{M}^d_{cpr}$, and

$$(\mathbf{P}_1 \cap \cdots \cap \mathbf{P}_m)^d = \mathbf{P}_1^d \cap \cdots \cap \mathbf{P}_m^d.$$

The upper expectation $\overline{E}_{\mathbf{P}}(f)$ of $f \in K^+$ w.r.t. the credal set $\mathbf{P}$ in $\overline{M}^d_{cpr}$ is defined as follows:

$$\overline{E}_{\mathbf{P}}(f) = \sup_{P\in\mathbf{P}} \overline{E}_P(f).$$

It is easy to check that the functional $\overline{E}_{\mathbf{P}}$ obeys the same properties as $\underline{E}_{\mathbf{P}}$ described in Lemma 4. The duality property of functionals $\underline{E}_{\mathbf{P}}$ and $\overline{E}_{\mathbf{P}}$ on K^+ is described in the following lemma.

Lemma 5 $\overline{E}_{\mathbf{P}^d}(f) = a - \underline{E}_{\mathbf{P}}(a-f)$, *where* $\mathbf{P}$ *is a credal set in* $\overline{M}_{cpr}$, $f \in K^+$, *and* $a = \max_{x\in X} f(x)$.

Remark 5 In the next we will extend functionals $\underline{E}_{\mathbf{P}}$ and $\overline{E}_{\mathbf{P}}$ on the set K of all real valued functions, assuming that the property 2) from Lemma 4 is valid for functions in K. Then for any $f \in K$ the values $\underline{E}_{\mathbf{P}}(f)$ and $\overline{E}_{\mathbf{P}}(f)$ are computed by

$$\underline{E}_{\mathbf{P}}(f) = \underline{E}_{\mathbf{P}}(\underline{f}) + a, \quad \overline{E}_{\mathbf{P}}(f) = \overline{E}_{\mathbf{P}}(\underline{f}) + a,$$

where $a = \min_{x\in X} f(x)$, and $\underline{f} = f - a$. Clearly $\underline{f} \in K^+$ and there exists $x \in X$ such that $\underline{f}(x) = 0$. We will call such functions *normalized* and keep the notation $\underline{f}$ (using lower bar).

Example 6 Let us consider the lower generalized credal set $\mathbf{P}^d = (\mathbf{P}_1 \cap \mathbf{P}_3)^d$, where $\mathbf{P}_1 \cap \mathbf{P}_3$ is defined in Example 4. Then we can compute $\overline{E}_{\mathbf{P}^d}(f)$ for any $f \in K^+$ as

$$\overline{E}_{\mathbf{P}^d}(f) = \max\left\{\overline{E}_{P_6^d}(f), \overline{E}_{P_7^d}(f)\right\},$$

where

$$\overline{E}_{P_6^d}(f) = \frac{1}{6}f(x_1) + \frac{1}{2}f(x_2) + \frac{1}{6}f(x_3) + \frac{1}{6}\min_{x_i\in X} f(x_i),$$
$$\overline{E}_{P_7^d}(f) = \frac{1}{3}f(x_1) + \frac{1}{3}f(x_2) + \frac{1}{6}f(x_3) + \frac{1}{6}\min_{x_i\in X} f(x_i).$$

Observe that for normalized functions $\frac{1}{6}\min_{x_i\in X} \underline{f}(x_i) = 0$. Let us compute $\overline{E}_{\mathbf{P}^d}(f)$ if $f = (f(x_1), f(x_2), f(x_3)) = (1,1,-3)$. Then $\min_{x_i\in X} f(x_i) = -3$, $\underline{f} = (4,4,0)$,

$$\overline{E}_{P_6^d}(f) = \overline{E}_{P_6^d}(\underline{f}) - 3 = \frac{4}{6} + \frac{4}{2} + 0 - 3 = -\frac{1}{3}.$$

Let us notice that all properties formulated in Lemma 4 remain valid for functionals $\underline{E}_{\mathbf{P}}$ and $\overline{E}_{\mathbf{P}}$ on K. The dual relation between $\underline{E}_{\mathbf{P}}$ and $\overline{E}_{\mathbf{P}}$ can be reformulated as $\overline{E}_{\mathbf{P}^d}(f) = -\underline{E}_{\mathbf{P}}(-f)$ for any credal set in $\overline{M}_{cpr}$ and $f \in K$.

The next lemma gives us the additional characteristic property of $\overline{E}_{\mathbf{P}}$, which, we will see later, helps us to describe the whole set of functionals $\underline{E}_{\mathbf{P}}$ and $\overline{E}_{\mathbf{P}}$.

Lemma 6 *Let* $\underline{f}_1, \underline{f}_2, \underline{f}_3$ *be normalized functions in* K^+ *such that* $\underline{f}_1 + \underline{f}_2 = \underline{f}_3$. *Then for any credal set* $\mathbf{P}$ *in* $\overline{M}^d_{cpr}$ *it is valid* $\overline{E}_{\mathbf{P}}(\underline{f}_1) + \overline{E}_{\mathbf{P}}(\underline{f}_2) \geq \overline{E}_{\mathbf{P}}(\underline{f}_3)$.

Theorem 1 *A functional* $\Phi : K^+ \to \mathbb{R}$ *coincides with* $\overline{E}_{\mathbf{P}}$ *on* K^+ *for some credal set* $\mathbf{P}$ *in* $\overline{M}^d_{cpr}$ *iff it has the following properties:*

1) $\Phi(0) = 0$; $\Phi(1) = 1$;
2) $\Phi(f + a) = \Phi(f) + a$ *for any* $f \in K^+$ *and* $a \in \mathbb{R}^+$;
3) $\Phi(af) = a\Phi(f)$ *for any* $f \in K^+$ *and* $a \in \mathbb{R}^+$;
4) $\Phi(f_1) \leq \Phi(f_2)$ *for* $f_1, f_2 \in K^+$ *if* $f_1 \leq f_2$;
5) $\Phi(\underline{f}_1) + \Phi(\underline{f}_2) \geq \Phi(\underline{f}_3)$ *for any normalized functions* $\underline{f}_1, \underline{f}_2, \underline{f}_3$ *in* K^+ *such that* $\underline{f}_1 + \underline{f}_2 = \underline{f}_3$.

8 Generalized Coherent Upper Previsions

Let $K' \subseteq K$, where K is the set of all functions of the type $f : X \to \mathbb{R}$, and let $\overline{E} : K' \to \mathbb{R}$ be the functional that defines the upper previsions, that may not satisfy the avoiding sure loss condition. Then $\overline{E}$ defines the non-empty lower generalized credal set $\mathbf{P}$ in $\overline{M}^d_{cpr}$ as follows:

$$\mathbf{P} = \left\{P \in \overline{M}^d_{cpr} | \forall f \in K' : \overline{E}_P(f) \leq \overline{E}(f)\right\} \qquad (4)$$

iff $\inf_{x\in X} f(x) \leq \overline{E}(f)$ for all $f \in K'$. Based on generalized credal set $\mathbf{P}$, we can define the natural extension of $\overline{E}$ by

$$\overline{E}'(f) = \sup\left\{\overline{E}_P(f) | P \in \mathbf{P}\right\} = \overline{E}_{\mathbf{P}}(f)$$

for all $f \in K$.

Theorem 2 *Let $\overline{E}: K' \to \mathbb{R}$ be the functional that defines the upper previsions. Then its natural extension $\overline{E}': K \to \mathbb{R}$ based on generalized credal sets can be computed as*

$$\overline{E}'(\underline{f}) =$$
$$\inf\left\{\sum_k a_k \overline{E}(\underline{f_k}) + a \middle| \sum_k a_k \underline{f_k} + a\mathbf{1} \geq \underline{f}, f_k \in K', a_k, a \geq 0\right\},$$

where $\underline{f}$ and $\underline{f_k}$ are normalized functions, and $\overline{E}'(\underline{f}) = \overline{E}'(f) - b$, $\overline{E}(\underline{f_k}) = \overline{E}(f_k) - b_k$, $b = \min_{x\in X} f(x)$, and $b_k = \min_{x\in X} f_k(x)$.

9 Conclusion

In this paper we generalize the C-rule for general theories of imprecise probabilities using the way of modeling contradiction (conflict) in evidence theory. This allows us to introduce upper and lower generalized credal sets and represent the C-rule as the intersection of corresponding generalized credal sets. The paper contains also some insights of how this model can be used in the theory of imprecise probabilities admitting contradiction.

Appendix[5]

Proof (Lemma 1) *Necessity.* Let $P_1 \leq P_2$, then in particular, $P_1(X \setminus \{x_i\}) \leq P_2(X \setminus \{x_i\})$, $i = 1, \ldots, n$, or equivalently, $1 - a_i \leq 1 - b_i$, or $a_i \geq b_i$, $i = 1, \ldots, n$.

Sufficiency. Let $a_i \geq b_i$, $i = 1, \ldots, n$, then

$$P_1 = \sum_{i=1}^{n}\left(b_i \eta_{\langle\{x_i\}\rangle} + (a_i - b_i)\eta_{\langle\{x_i\}\rangle}\right) + a_0 \eta^d_{\langle X\rangle}$$
$$\leq \sum_{i=1}^{n}\left(b_i \eta_{\langle\{x_i\}\rangle} + (a_i - b_i)\eta^d_{\langle X\rangle}\right) + a_0 \eta^d_{\langle X\rangle} = P_2.$$

Proof (Lemma 2) Because the set $\mathbf{P}$ is closed, we have $\mathbf{P} = \{P \in \overline{M}_{cpr} | \exists P' \in profile(\mathbf{P}) : P \geq P'\}$. This implies the required result.

Proof (Lemma 3) Because P is a plausibility function (2-alternating measure), the value $\underline{E}_P(f)$ is expressed through the Choquet integral:

$$\underline{E}_P(f) = \int_X f(x)dP = a_0 \int_X f(x) d\eta^d_{\langle X\rangle} + \sum_{i=1}^{n} a_i \int_X f(x) d\eta_{\langle\{x_i\}\rangle}$$
$$= a_0 \max_{x\in X} f(x) + \sum_{i=1}^{n} a_i f(x_i).$$

In the last expression we use the additivity of the Choquet integral w.r.t. the sum of measures, and also that $\int_X f(x) d\eta_{\langle\{x_i\}\rangle} = f(x_i)$ and $\int_X f(x) d\eta^d_{\langle X\rangle} = \max_{x\in X} f(x)$.

Proof (Lemma 5) Notice that the validity of $\overline{E}_{P^d}(f) = a - \underline{E}_P(a - f)$ for $P \in \overline{M}_{cpr}$ follows from the properties of the Choquet integral. By definition

$$\overline{E}_{\mathbf{P}^d}(f) = \sup_{P^d \in \mathbf{P}^d} \overline{E}_{P^d}(f) = \sup_{P\in\mathbf{P}} (a - \underline{E}_P(a - f))$$
$$= a - \inf_{P\in\mathbf{P}} \underline{E}_P(a - f) = a - \underline{E}_{\mathbf{P}}(a - f).$$

[5]Straightforward proofs are omitted.

Proof (Lemma 6) Because by definition the credal set $\mathbf{P}$ is closed, there exists $P \in \mathbf{P}$ such that $\overline{E}_P\left(\underline{f}_3\right) = \overline{E}_{\mathbf{P}}(\underline{f}_3)$. Assume that $P = a_0 \eta_{\langle X\rangle} + \sum_{i=1}^{n} a_i \eta_{\langle\{x_i\}\rangle}$. Notice that in this case

$$\overline{E}_P(\underline{f}_k) = \sum_{i=1}^{n} a_i \underline{f}_k(x_i), \quad k = 1, 2, 3,$$

since $\min_{x\in X} \underline{f}_k(x) = 0$. Thus, $\overline{E}_P(\underline{f}_1) + \overline{E}_P(\underline{f}_2) = \overline{E}_P(\underline{f}_3)$. In addition, clearly $\overline{E}_{\mathbf{P}}(\underline{f}_k) \geq \overline{E}_P(\underline{f}_k)$, $k = 1, 2$. This implies the inequality from the lemma.

Proof (Theorem 1) Necessity follows from Lemma 4 (see Remark 5) and Lemma 6. Let us prove sufficiency. It is sufficient to show that for any normalized function $\underline{f}$ there is a $P \in \overline{M}^d_{cpr}$ such that $\Phi(\underline{f}) = \overline{E}_P(\underline{f})$ and $\Phi \geq \overline{E}_P$. Because $\underline{f}$ is normalized there is $x_k \in X$ such that $\underline{f}(x_k) = 0$. Let us consider the set K' of all functions f in K^+ with $f(x_k) = 0$. Let us notice that the monotone functional Φ on K' is sublinear, and by Hahn-Banach's Theorem there is a linear functional on K'

$$\alpha(f) = \sum_{i=1}^{n} a_i f(x_i)$$

such that $a_i \geq 0$, $i = 1, \ldots, n$, $\sum_{i=1}^{n} a_i \leq 1$, $\alpha \leq \Phi$ and $\alpha(\underline{f}) = \Phi(\underline{f})$. Obviously, we can assume that $a_k = 0$. Introduce into consideration

$$P = a_0 \eta_{\langle X\rangle} + \sum_{i=1}^{n} a_i \eta_{\langle\{x_i\}\rangle},$$

where $a_0 = 1 - \sum_{i=1}^{n} a_i$ and show that $\Phi(\underline{f}) = \overline{E}_P(\underline{f})$ and $\Phi \geq \overline{E}_P$. The equality $\Phi(\underline{f}) = \overline{E}_P(\underline{f})$ is obvious. Let us show that $\Phi(g) \geq \overline{E}_P(g)$ for any $g \in K^+$. Obviously, $\Phi(g) \geq \overline{E}_P(g)$ iff $\Phi(\underline{g}) \geq \overline{E}_P(\underline{g})$. Notice that $\overline{E}_P(\underline{g}) = \overline{E}_P(\underline{g}')$, where $\underline{g}'(x_i) = \underline{g}(x_i)$ for $i \neq k$ and $\underline{g}'(x_i) = 0$ otherwise. Since $\underline{g}' \leq \underline{g}$, we get $\overline{E}_P(\underline{g}) = \overline{E}_P(\underline{g}') \leq \Phi(\underline{g}') \leq \Phi(\underline{g})$. The theorem is fully proved.

Proof (Theorem 2) Let us show first that functionals $\overline{E}$ and $\overline{E}'$ define the same credal set, i.e. the credal set $\mathbf{P}$ defined by (4) is equal to

$$\mathbf{P}' = \left\{P \in \overline{M}^d_{cpr} \middle| \forall f \in K : \overline{E}_P(f) \leq \overline{E}'(f)\right\}.$$

The inclusion $\mathbf{P}' \subseteq \mathbf{P}$ is obvious. Let $P \in \mathbf{P}$, then by our assumption $\overline{E}_P(\underline{f_k}) \leq \overline{E}(\underline{f_k})$ for $f_k \in K'$ and

$$\overline{E}_P(\underline{f}) = \sum_{i=1}^{n} P(\{x_i\}) \underline{f}(x_i) \leq \sum_{i=1}^{n} P(\{x_i\}) \left(\sum_k a_k \underline{f_k}(x_i) + a\right)$$
$$\leq \sum_{i=1}^{n} P(\{x_i\}) \sum_k a_k \underline{f_k}(x_i) + a$$
$$= \sum_k a_k \overline{E}_P(\underline{f_k}) + a$$
$$\leq \sum_k a_k \overline{E}(\underline{f_k}) + a.$$

Thus, $\mathbf{P} \subseteq \mathbf{P}'$, i.e. $\mathbf{P}' = \mathbf{P}$. Let us show that the functional $\overline{E}'$ obeys all properties on K^+ for functional Φ given in Theorem 1. It is easy to check that properties 1), 2), 3), 5) are valid. Let us show that the monotonicity property 4) is also satisfied. For this purpose introduce into consideration the functional

$$\Phi(f) = \inf\left\{\sum_k a_k \overline{E}(\underline{f_k}) + a \middle| \sum_k a_k \underline{f_k} + a\mathbf{1} \geq f, f_k \in K', a_k, a \geq 0\right\}$$

on K^+. Evidently, $\overline{E}'(\underline{f}) = \Phi(\underline{f})$ for every $f \in K^+$. It is easy to check that this functional on K^+ has the following properties:

1) $\Phi(\mathbf{0}) = 0$, $\Phi(\mathbf{1}) \leq 1$;
2) $\Phi(af) = a\Phi(f)$ for any $f \in K^+$ and $a \in \mathbb{R}^+$;
3) $\Phi(f_1) \leq \Phi(f_2)$ for $f_1, f_2 \in K^+$ if $f_1 \leq f_2$;
4) $\Phi(f_1) + \Phi(f_2) \geq \Phi(f_3)$ for any functions f_1, f_2, f_3 in K^+ such that $f_1 + f_2 = f_3$.

By Hahn-Banach's Theorem for every $f \in K^+$ there is a linear functional on K^+, $\alpha(f) = \sum_{i=1}^n a_i f(x_i)$, such that $a_i \geq 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n a_i \leq 1$, $\alpha \leq \Phi$ and $\alpha(f) = \Phi(f)$. We will use next this functional for proving monotonicity of $\overline{E}'$. Consider an arbitrary $f, g \in K^+$ such that $f \leq g$. Let $f = \underline{f} + c$. Then inequality $\overline{E}'(f) \leq \overline{E}'(g)$ is equivalent to $\overline{E}'(\underline{f}) \leq \overline{E}'(g')$, where $g' = g - c$. Obviously, $\overline{E}'(\underline{f}) = \Phi(\underline{f}) \leq \Phi(g')$. By previous conclusions, there is a linear functional $\alpha(f) = \sum_{i=1}^n a_i f(x_i)$ on K^+ such that $a_i \geq 0$, $i = 1, \ldots, n$, $\sum_{i=1}^n a_i \leq 1$, $\alpha \leq \Phi$ and $\alpha(g') = \Phi(g')$. Let $P = a_0 \eta_{\langle X \rangle} + \sum_{i=1}^n a_i \eta_{\langle \{x_i\} \rangle}$, where $a_0 = 1 - \sum_{i=1}^n a_i$. It is easy to see that $P \in \mathbf{P}$ and $\Phi(g') \leq \overline{E}_P(g') \leq \overline{E}'(g')$, i.e. $\overline{E}'(\underline{f}) \leq \overline{E}'(g')$ and $\overline{E}'(f) \leq \overline{E}'(g)$.

Thus, we prove that the functional $\overline{E}'$ obeys all properties from Theorem 1. This means that it is the natural extension of $\overline{E}$.

Acknowledgements

This study (Research Grant No.14-01-0015) was supported by The National Research University Higher School of Economics Academic Fund Program in 2014/2015.

Authors express their sincerely thanks to the anonymous reviewers for detailed and helpful comments.

References

[1] T. Augustin, F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes, eds. *Introduction to Imprecise Probabilities*. New York: Wiley, 2014.

[2] A. G. Bronevich and G. J. Klir, Measures of uncertainty for imprecise probabilities: an axiomatic approach. *International Journal of Approximate Reasoning* 51: 365-390, 2010.

[3] A. G. Bronevich and I. N. Rozenberg. The choice of generalized Dempster-Shafer rules for aggregating belief functions based on imprecision indices. *Belief Functions: Theory and Applications. Lecture Notes in Computer Science*, vol. 8764, Springer Verlag, Berlin, 2014, pp 21-28.

[4] A. G. Bronevich and I. N. Rozenberg. The choice of generalized Dempster-Shafer rules for aggregating belief functions. *International Journal of Approximate Reasoning* 56: 122-136, 2015.

[5] M. E. G. V. Cattaneo. Combining belief functions issued from dependent sources, in: J.-M. Bernard, T. Seidenfeld, M. Zaffalon (Eds.), ISIPTA '03, Proceedings in Informatics, vol. 18, Carleton Scientific,Waterloo, 2003, pp. 133-147.

[6] M. E. G. V. Cattaneo. Belief functions combination without the assumption of independence of the information sources. *International Journal of Approximate Reasoning* 52: 299-315, 2011.

[7] G. de Cooman and M. C. M. Troffaes. *Lower Previsions*. New York: Wiley, 2014.

[8] T. Denoeux. Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence* 172: 234-264, 2008.

[9] S. Destercke and V. Antoine. Combining imprecise probability masses with maximal coherent subsets: Application to ensemble classification. *Synergies of Soft Computing and Statistics for Intelligent Data Analysis Advances in Intelligent Systems and Computing,* Springer, vol. 190, 2013, pp 27-35.

[10] S. Destercke and T. Burger. Toward an axiomatic definition of conflict between belief functions, IEEE Trans. Syst. Man Cybern. 43: 585-596, 2013.

[11] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12:193-226, 1986.

[12] G. J. Klir. *Uncertainty and Information: Foundations of Generalized Information Theory.* Hoboken, NJ: Wiley-Interscience, 2006.

[13] S. Moral and J. Sagrado. Aggregation of imprecise probabilities. In B. Bouchon Meunier, ed., *Aggregation and Fusion of Imperfect Information*, pages 162-188. Physica-Verlag, Heidelberg, 1997.

[14] R. Nau. The aggregation of imprecise probabilities. *Journal of Statistical Planning end Inference* 105: 265-282, 2002.

[15] G. Shafer. *A mathematical theory of evidence*. Princeton, N.J.: Princeton University Press, 1976.

[16] P. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion* 8: 387-412, 2007.

[17] M. C. M. Troffaes. Generalising the conjunction rule for aggregating conflicting expert opinions. *International Journal of Intelligent Systems* 21: 361-380, 2006.

[18] P. Walley. *Statistical Reasoning with Imprecise Probabilities.* London: Chapman and Hall, 1991.