Generalized Hartley Measures on Credal Sets

Andrey G. Bronevich
Igor N. Rozenberg

JSC "Research and Design Institute for Information Technology, Signalling and Telecommunications on Railway Transport", Orlikov per.5, building 1 107996 Moscow, Russia

Abstract
The paper considers various extensions of the Hartley measure on credal sets and their investigation based on a system of axioms.

Keywords: uncertainty measures, generalized Hartley measure, credal sets

1. Introduction
In the theory of imprecise probabilities [4, 23], there are many functionals for measuring uncertainty [5, 11, 19]. Among them we distinguish measures of conflict, non-specificity and total uncertainty. Conflict is related to modeling uncertainty by probability measures; non-specificity is connected to the choice of a probability model among possible ones. The total uncertainty is conceived as aggregated uncertainty of these two types. Historically, measures of conflict have been firstly introduced in probability theory, and they are known as entropies [22, 20]. The most popular of them are the Shannon entropy and Rényi entropies. The measure of non-specificity has been firstly introduced for analyzing information that can be described by non-graded possibility measures and called the Hartley measure [18]. There are several attempts to extend these measures to various models of imprecise probabilities [1, 5], or especially to belief functions [16]. As we can see from the literature [2, 3, 5, 11, 19] the most justified of them are the maximal (upper) entropy for evaluating total uncertainty, the minimal (lower) entropy for measuring conflict and the generalized Hartley measure as a measure of non-specificity.

Unfortunately, the generalized Hartley measure was fully accepted only for belief functions. Although there are several extensions of it to coherent lower probabilities and credal sets [1, 5], the thorough investigation of their properties was not produced yet. To close these gaps, in this paper we consider three possible extensions of the Hartley measure to credal sets. The third one is new and based on an interpretation of the Hartley measure in decision theory as follows. An imprecise probability model can be viewed as a system of precise probability models. We call two probability models $P_1$ and $P_2$ fully contradictory if there is a decision $f$ on the set of alternatives $X = \{x_1, \ldots, x_n\}$ whose utility is the highest and equal to $\max_{x \in X} f(x)$ for $P_1$ and the utility of $f$ is the smallest and equal to $\min_{x \in X} f(x)$ for $P_2$. We show that the Hartley measure can be viewed as the logarithm of the maximal number of pairwise fully contradictory probability measures in the corresponding credal set. We also introduce the notion of $\varepsilon$-contradictory probability measures, and make hints of how to generalize the Hartley measure in this way.

The paper has the following structure. In Section 2, we recall the basic notions from the theory of belief functions and imprecise probabilities. Section 3 gives the axioms of classical uncertainty measures: the Shannon entropy and the Hartley measure. In Sections 4 and 5, we recall [5] the axioms for uncertainty measures on belief functions, the possible disaggregations of a measure of total uncertainty onto measures of conflict and non-specificity and the extension of these results on credal sets. In Section 6, we describe some properties of the linear extension of the generalized Hartley measure on coherent lower probabilities and credal sets. In Section 7, we introduce the Hartley measure on credal sets and investigate its properties. The paper is ended with the discussion of obtained results.

2. Monotone Measures and Credal Sets
Let $X$ be a finite non-empty set and let $2^X$ denote the powerset of $X$. Then a set function $\mu : 2^X \to [0, 1]$ is called a monotone measure or capacity [13, 15] if

1) $\mu(\emptyset) = 0$, $\mu(X) = 1$ (normalizing);
2) $\mu(A) \leq \mu(B)$ for every $A, B \in 2^X$ with $A \subseteq B$ (monotonicity).

Note that a probability measure is a special case of monotone measures. In this case, $\mu(A) + \mu(B) = \mu(A \cup B)$ for any disjoint $A, B \in 2^X$. We denote by $M_{mon}(X)$ the set of all monotone measures on $2^X$, and by $M_{pr}(X)$ the set of all probability measures on $2^X$. We will further use the following operations and relations on monotone measures.

Let $\mu_1, \mu_2 \in M_{mon}(X)$, then

1) $\mu = a\mu_1 + (1-a)\mu_2$, where $a \in [0, 1]$, if $\mu(A) = a\mu_1(A) + (1-a)\mu_2(A)$ for all $A \in 2^X$ ($\mu$ is called the convex sum of $\mu_1$ and $\mu_2$);

1. Sometimes, we use the convex sum $\mu = a\mu_1 + (1-a)\mu_2$, where $\mu_i \in M_{mon}(X)$, $i = 1, 2$, and $X_1 \cap X_2 = \emptyset$. In this case, $\mu \in M_{mon}(X)$,
2) $\mu_1 \leq \mu_2$ if $\mu_1(A) \leq \mu_2(A)$ for all $A \in 2^X$;
3) $\mu_1 \leq \mu_2$ if $\mu_1(A) = 1 - \mu_2(A')$ for all $A \in 2^X$, where $A'$ denotes the complement of $A$ ($\mu_2$ is called the dual of $\mu_1$).

$\mu$ is called a lower probability [4, 23] if the set $\mathcal{P}(\mu) = \{ P \in \mathcal{M}_p(X) | P \geq \mu \}$ is not empty, if, in addition, $\mu(A) = \inf_{P \in \mathcal{P}(\mu)} P(A)$ for all $A \in 2^X$, then $\mu$ is called a coherent lower probability. The family of probability measures $\mathcal{P}(\mu)$ is called the credal set. More generally, a credal set $\mathcal{P}$ is a non-empty convex and closed subset of $\mathcal{M}_p(X)$. The convexity means that if $P_1, P_2 \in \mathcal{P}$ implies that $\alpha P_1 + (1-\alpha)P_2 \in \mathcal{P}$ for every $\alpha \in [0, 1]$. The closeness of $\mathcal{P}$ means that if we represent every $P \in \mathcal{M}_p(X)$, where $X = \{x_1, ..., x_n\}$, as a point $P = (P(x_1), ..., P(x_n))$ in $\mathbb{R}^n$, then $\mathcal{P}$ is a closed subset of $\mathbb{R}^n$. Instead of defining credal sets using lower probabilities, we can do it by upper probabilities. A $\mu \in \mathcal{M}_{\text{mon}}(X)$ is called an upper probability if $\mu^d$ is a lower probability. It is easy to show that $\mathcal{P}(\mu)$ is a convex and closed set. It is possible to show that every functional $\mu$ is a coherent upper probability. The set of all coherent upper probabilities on $2^X$ is denoted by $\mathcal{M}_{2-\text{mon}}(X)$.

An important example of coherent lower probabilities (2-monotone measures) are belief functions [21]. A $\mu \in \mathcal{M}_{\text{mon}}(X)$ is called a belief function if there is $m : 2^X \to [0, 1]$ called the basic belief assignment (bba) with $\sum_{A \in 2^X} m(A) = 1$ and $m(\emptyset) = 0$ such that $\mu(A) = \sum_{B \subseteq A} m(B)$. The set of all belief functions on $2^X$ is denoted by $\mathcal{M}_{\text{bel}}(X)$. Assume that $\mu \in \mathcal{M}_{\text{bel}}(X)$ with the bba $m$. A set $A \in 2^X$ is called a focal element for $\mu$ if $m(A) > 0$. The set of focal elements is called the body of evidence. A belief function with only one focal element $B$ is called categorical and denoted by $\eta_B$. Obviously,

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

Every $B \in \mathcal{M}_{\text{bel}}(X)$ with the bba $m$ is represented as a convex sum of categorical belief functions:

$$\mathcal{B} = \sum_{B \subseteq 2^X} m(B) \eta_B.$$ 

In the sequel, we will consider credal sets with a finite number of extreme points, and the set of all such credal sets on $2^X$ is denoted by $\mathcal{C}(X)$. If $P_1, ..., P_m \in \mathcal{P}$ are extreme points of $\mathcal{P} \subset \mathcal{C}(X)$, then every $P \in \mathcal{P}$ can be represented as a convex sum of $P_1, ..., P_m$, i.e. $P = \sum_{i=1}^m a_i P_i$ for some $a_i \geq 0, i = 1, ..., m$, with $\sum_{i=1}^m a_i = 1$.

In notations like $M_{\text{pr}}(X), \mathcal{C}(X)$, we can drop the reference set $X$, if its definition in the context is not necessary.

### 3. Hartley Measure and Shannon Entropy

The Hartley measure $H$ evaluates uncertainty if we only know that the unknown value of a random variable $\xi$ is in a finite set $A$. Let us formulate the desirable properties of such non-specificity measure. We suppose that $H : \mathcal{P} \to [0, +\infty)$, where $\mathcal{P}$ is the family of all finite non-empty sets.

- **H1. Boundary condition:** $H(A) = 0$ for $A \in \mathcal{P}$ iff $|A| = 1$.
- **H2. Monotonicity:** $H(A) \leq H(B)$, $A, B \in \mathcal{P}$, if $A \subseteq B$.
- **H3. Label independency:** let $\phi : A \to B$ be a bijection between finite sets $A, B \in \mathcal{P}$, then $H(A) = H(B)$.
- **H4. Additivity:** $H(A \times B) = H(A) + H(B)$ for $A, B \in \mathcal{P}$.

Let us discuss the above properties. Property H1 says that only in the case of exact information, when $|A| = 1$, the value of $H$ is equal to zero. According to the Property H2, knowing that $\xi(\omega) \in A$ is more exact than $\xi(\omega) \in B$, therefore, $H(A) \leq H(B)$. The bijection $\phi$ in Property H3 can be conceived as giving new names to the elements of $A$. Thus, this renaming does not affect the uncertainty. Property H4 means the following. Assume that we have two random variables $\xi$ and $\eta$, and we only know that $\xi(\omega) \in A$ and $\eta(\omega) \in B$, then we can conclude that $\{\xi(\omega), \eta(\omega)\} \in A \times B$. We assume that if we aggregate two independent (non-interacted) pieces of information, then uncertainty behaves additively. This implies Property 4.

Properties H2 and H4 imply the subadditivity property.

**H5. Subadditivity:** Let $A \in \mathcal{P}$ be such that $A \subseteq X \times Y$ and let $A_X$ and $A_Y$ be its projections on $X$ and $Y$, respectively, i.e. $A_X = \{x \in X | \exists y \in Y : (x, y) \in A\}$ and $A_Y = \{y \in Y | \exists x \in X : (x, y) \in A\}$. Then $H(A_X) + H(A_Y) \geq H(A)$.

It is possible to show that every functional $H$ with Properties H1-H4 can be represented in the form: $H(A) = c \ln |A|$, where $c > 0$. If we additionally require that $H(A) = 1$ if $|A| = 2$, then $H(A) = \log_2 |A|$.

The situation, when we only know that $\xi(\omega) \in A$, can be described by the set of all possible probability distributions of $\xi$ that coincides with $M_{\text{pr}}(A)$, or by the categorical belief function $\eta_A$, because, $M_{\text{pr}}(A) = \{P \in M_{\text{pr}}(X) \mid P(A) \geq 0\}$.

Another special case is when a random variable $\xi$ takes its values in a finite set $X$ and we know the probability distribution of $\xi$. Before describing an uncertainty measure $S$ for this case, called the Shannon entropy, we will introduce the following constructions:

a) let $\phi : X \to Y$ and $\mu \in \mathcal{M}_{\text{mon}}(X)$, then the image $\mu^\phi$ of $\mu$ is defined by $\mu^\phi(B) = \mu(\phi^{-1}(B))$, where $\phi^{-1}(B) = \{x \in X | \phi(x) \in B\}$;
b) let \( \mu \in M_{moo}(X \times Y) \), then projections \( \mu_X \) and \( \mu_Y \) on \( X \) and \( Y \), respectively, are defined by \( \mu_X(A) = \mu(Ax) \) and \( \mu_Y(A) = \mu(AY) \) for every \( A \in 2^X \times Y \).

The Shannon entropy is the functional \( S : M_{pr} \to [0, +\infty) \) with the following properties:

S1. **Boundary condition**: \( S(P) = 0 \) for \( P \in M_{pr}(X) \) iff \( P = \eta_{\{x\}} \) for some \( x \in X \).

S2. **Label independency**: let \( \varphi : X \to Y \) be a bijection between finite sets \( X \) and \( Y \), and \( P \in M_{pr}(X) \), then \( S(P^\varphi) = S(P) \).

S3. **Expansibility**: let \( \varphi : X \to Y \) be an injection such that \( X \subseteq Y \) and \( \varphi(x) = x \) for every \( x \in X \), then \( S(P^\varphi) = S(P) \) for every \( P \in M_{pr}(X) \).

S4. **Additivity**: let \( P \in M_{pr}(X \times Y) \), then \( S(P) = \sum_{y \in Y} P_y(\{y\}) S(P_y) + S(P_y) \), where \( P_{\{y\}} \in M_{pr}(X) \) is defined by \( P_{\{y\}}(A) = P(A \times \{y\}) / P_{\{y\}}(\{y\}) \) for every \( A \in 2^X \).

S5. **Subadditivity**: let \( P \in M_{pr}(X \times Y) \), then \( S(P_X) + S(P_Y) \geq S(P) \).

Note that Properties S1-S2 have the same interpretation as for the Hartley measure. Property S3 means that adding dummy elements to \( X \) does not affect the value \( S(P) \). Properties S2-S3 are equivalent to S2-S3. Let \( \varphi : X \to Y \) be an injection, then \( S(P^\varphi) = S(P) \) for every \( P \in M_{pr}(X) \).

Property S4 has the following interpretation. Let \( P \in M_{pr}(X \times Y) \) describe the joint probability distribution of random variables \( \xi_X \) and \( \xi_Y \) with values in \( X \) and \( Y \), respectively. Then Property S4 is equivalent to \( H(\xi_X, \xi_Y) = H(\xi_X|\xi_Y) + H(\xi_Y|\xi_X) \). Hence, \( H(\xi_X, \xi_Y) \) is the entropy of the joint probability distribution of \( \xi_X \) and \( \xi_Y \), and \( H(\xi_X|\xi_Y) \) and \( H(\xi_Y|\xi_X) \) are the conditional entropies of \( \xi_X \) given \( \xi_Y \) and \( \xi_Y \) given \( \xi_X \). The weak form of Property S4 is

\[ S(P) = \sum_{y \in Y} P_y(\{y\}) \ln P_y(\{y\}) \]

for every \( P \in M_{pr}(X) \).

4. **Uncertainty Measures on Belief Functions**

While modelling uncertainty by belief functions, we distinguish two types of uncertainty: *non-specificity* and *conflict*. Conflict is related to modelling uncertainty by probability measures; non-specificity comes from the possible choices of a “true” probability model among admissible ones. We also need to introduce a measure of total uncertainty that aggregates uncertainty of these two types. Therefore, we should define three functionals:

- a measure of conflict \( U_C : M_{bel} \to [0, +\infty) \);
- a measure of non-specificity \( U_N : M_{bel} \to [0, +\infty) \);
- a measure of total uncertainty \( U_T : M_{bel} \to [0, +\infty) \).

In [5], the following system of axioms is proposed (see also a slightly different system of axioms for \( U_T \) in [17]).

B1. **Boundary condition**: \( U_T(\mu) = 0 \) for \( \mu \in M_{pr}(X) \) and \( U_C(\eta_{\{\cdot\}}) = 0 \) for \( B \in 2^X \setminus \{\emptyset\} \).

B2. **Expansibility and label independency**: let \( \varphi : X \to Y \) be an injection and \( \mu \in M_{bel}(X) \), then \( U_T(\varphi) = U_T(\mu) \).

B3. **Monotonicity w.r.t. mapping**: let \( \varphi : X \to Y \) be a mapping and \( \mu \in M_{bel}(X) \), then \( U_T(\varphi) \leq U_T(\mu) \).

B4. **Monotonicity**: let \( \mu_1, \mu_2 \in M_{bel}(X) \) and \( \mu_1 \leq \mu_2 \), then \( U_N(\mu_1) \leq U_N(\mu_2) \) and \( U_T(\mu_1) \geq U_T(\mu_2) \).

B5. **The first additivity property**: let \( \mu \in M_{bel}(X \times Y) \) be such that \( \mu = \sum_{A \in 2^X} m_X(A) \eta_{\{A\times\{\cdot\}\}} \), where \( B \in 2^Y \setminus \{\emptyset\} \),

\[ U_T(\mu) = U_T(\mu_X) + U_T(\mu_Y) \]

where \( m_X(A) = \frac{U_T(\varphi) - U_T(\mu_\varphi)}{\mu(\varphi)} \), \( A \in 2^X \).

B6. **The second additivity property**: let \( \mu \in M_{bel}(X \times Y) \) and \( \mu_X \in M_{pr}(Y) \), then

\[ U_T(\mu) = \sum_{y \in Y} \mu_Y(\{y\}) U_T(\mu_{\{y\}}) + U_T(\mu_Y) \]

where \( m_X(A) = \frac{U_T(\varphi) - U_T(\mu_\varphi)}{\mu(\varphi)} \), \( A \in 2^X \).

B7. **Subadditivity**: let \( \mu \in M_{bel}(X \times Y) \), then \( U_T(\mu_X) + U_T(\mu_Y) \geq U_T(\mu) \).

B8. **Disaggregation**: \( U_C(\mu) + U_N(\mu) = U_T(\mu) \) for every \( \mu \in M_{bel}(X) \).

Let us observe that if we consider the restriction of the above axioms for probability measures, then we get Properties S1-S5 of the Shannon entropy, analogously, the restriction of these axioms for categorical belief functions are Properties H1-H5 of the Hartley measure. Thus, \( U_T(P) = U_C(P) = S(P) \) for every \( P \in M_{pr} \) and \( U_T(\mu) = U_N(\mu) = H(\mu) \) for every categorical belief function \( \mu \). Because the axioms B1-B8 have the same interpretation as the basic properties formulated for the Shannon entropy and the Hartley measure, we will explain only some of them.

Consider the following explanation of Axiom B3 given in [5]. Assume that \( \mu \in M_{bel}(X) \) and \( \varphi : X \to Y \) is such that \( \varphi(x) = y_i \) if \( x \in X_i \), where \( \{X_1, ..., X_k\} \) is the partition of \( X \). Assume also that \( y_i \in X_i \), \( i = 1, ..., k \). Thus, \( \varphi \) has

\[ \varphi(x) = \begin{cases} y_i & \text{if } x \in X_i \\ \emptyset & \text{if } x \notin X_i \end{cases} \]
the following interpretation: if the true alternative is in \(X_i\), then it is \(y_i\). Because any additional information reduces uncertainty, we should require that \(U_T(\mu^0) \leq U_T(\mu)\).

Since in Axiom B4 \(P(\mu_1) \supseteq P(\mu_2)\), \(\mu_1\) looks as a model of uncertainty at least as with the same or higher non-specificity than \(\mu_2\). Therefore, we require that \(U_N(\mu_1) \geq U_N(\mu_2)\) and \(U_T(\mu_1) \geq U_T(\mu_2)\).

Axiom B5 is transformed to Property H4 for the Hartley measure, when \(\mu_x\) is a categorical belief function. Formally, in Axiom B5 the Mobius product \(\times_M\) of \(\mu_x\) and \(\mu_T\) is used, because by definition,

\[
\mu = \mu_X \times_M \mu_T = \sum_{A \in 2^X} \sum_{B \in 2^Y} m_X(A)m_Y(B)\eta_{\langle A \times B \rangle},
\]

where \(m_X\) and \(m_Y\) are bbas of \(\mu_x\) and \(\mu_T\), respectively.

Axiom B6 is the generalization of Property H4. Actually, we get Property H4 taking \(\mu_x \in M_{pr}(X)\). Note that the conditional belief function \(\mu_x\) is well justified in the theory of imprecise probabilities. It is possible equivalently to exchange Axiom B6 to

\[B^6.\] Let \(\{X_1, \ldots, X_n\}\) be a partition of \(X \) and \(\mu_x \in M_{be}(X_i)\), \(i = 1, \ldots, m\). Consider \(\mu = \sum_{i=1}^n a_i \mu_i\), where \(\sum_{i=1}^m a_i = 1\) and \(a_i > 0\), \(i = 1, \ldots, m\). Then \(U_T(\mu) = \sum_{i=1}^m a_i U_T(\mu_i) + U_T(P)\), where \(P \in M_{pr}(Y)\) is such that \(Y = \{1, \ldots, m\}\) and \(P(\{i\}) = a_i, i = 1, \ldots, m\).

There are several results [5], that we will use later. Axioms B1-B8 imply that

1) \(U_T \left( \sum_{i=1}^m a_i \mu_i \right) \geq \sum_{i=1}^m a_i U_T(\mu_i)\) for every \(\sum_{i=1}^m a_i = 1\), \(a_i > 0\), \(\mu_i \in M_{be}(X_i)\), \(i = 1, \ldots, m\).

2) The set \(\mathcal{S}\) of all functionals, satisfying Axioms B1-B8, is a convex cone, i.e. if \(U_T^{(1)} \in \mathcal{S}\), \(i = 1, 2\), then \(a U_T^{(1)} + b U_T^{(2)} \in \mathcal{S}\) for any \(a, b > 0\).

We can impose normalizing conditions \(U_T(\eta_{\langle X \rangle}) = a\) and \(U_T(P^0) = b\), where \(P^0 \in M_{pr}(X)\) defines the uniform probability distribution on \(X\) and \(|X| = 2\). Axiom B4 implies that \(a \geq b\) and we need to take \(a > b\) to 0 providing \(U_T\) to be not identical zero. We denote the set of all possible total uncertainty measures with such normalizing conditions by \(\mathcal{S}_{a,a}\). There are two different total uncertainty measures satisfying Axioms B1-B8. It is possible to prove that \(\mathcal{S}_{a,0} = a > 0\), is a singleton, i.e. such normalizing conditions uniquely define the total uncertainty measure that coincides with the generalized Hartley measure [16]:

\[GH(\mu) = \sum_{B \in 2^X} m(B)H(B),\]

where \(\mu \in M_{be}(X)\) and \(m\) is the bba of \(\mu\). The set \(\mathcal{S}_{a,a}, a > 0\), is also non-empty, since \(S_{\text{max}} \in \mathcal{S}_{a,a}\), where \(S_{\text{max}}\) is the maximal (upper) entropy, i.e. \(S_{\text{max}}(\mu) = \sup_{P \in P(\mu)} S(P)\).

3) There are several admissible disaggregations of \(U_T \in \mathcal{S}\):

a) \(U_C(\mu) = \inf_{P \in P(\mu)} U_T(\mu)\), where \(U_T\) is the Shannon entropy on \(M_{pr}, U_N = U_T - U_C\). This \(U_C\) is denoted by \(S_{\text{min}}\) and called the minimal (lower) entropy.

b) \(U_N(\mu) = \sum_{B \in 2^X} m(B)U_T(\eta_{\langle B \rangle})\), where \(m\) is the bba of \(\mu \in M_{be}(X)\), and \(U_C = U_T - U_N\). In this case, \(U_N\) is obviously the generalized Hartley measure.

5. Uncertainty Measures on Credal Sets

In [5], a reader can find a system of axioms for uncertainty measures on credal sets, that generalizes Axioms B1-B8. For describing them, we define the following operations on credal sets:

1) let \(\varphi : X \to Y\) be a mapping and \(P \in \mathcal{C}r(X)\), then \(P^\varphi = \{P^\varphi|P \in \mathcal{C}r(X)\};

2) \(P = aP_1 + (1 - a)P_2\) for \(a \in [0, 1]\) and \(P_1, P_2 \in \mathcal{C}r(X)\); if \(P \in \mathcal{C}r(X)\), then \(P_X = \{P_X|P \in \mathcal{C}r(X)\};

3) let \(P \in \mathcal{C}r(X \times Y)\), then \(P_X = \{P_X|P \in \mathcal{C}r(X)\}\); and \(P_Y = \{P_Y|P \in \mathcal{C}r(Y)\};

4) let \(P_X \in \mathcal{C}r(X)\) and \(P_Y \in \mathcal{C}r(Y)\), then \(P_{X \times N}P_Y = \{P_{X \times N}P_Y|P_X, P_Y \in \mathcal{C}r(X)\};

Using the above operations, we define axioms for \(U_T, U_C\) and \(U_N\) on \(\mathcal{C}r(X)\) as follows [5]:

C1. Boundary condition: let \(P \in \mathcal{C}r(X)\), then \(U_N(P) = 0\) if \(P = \emptyset\), and \(U_C(P) = 0\) if \(P = \emptyset\) for some \(B \in 2^X \setminus \emptyset\).

C2. Expansibility and label independency: let \(\varphi : X \to Y\) be an injection, then \(U_T(P^\varphi) = U_T(P)\), \(U_N(P) = U_N(P)\), \(U_C(P) = U_C(P)\) for any \(P \in \mathcal{C}r(X)\).

C3. Monotonicity w.r.t. mapping: let \(\varphi : X \to Y\) be a mapping, and \(P \in \mathcal{C}r(X)\), then \(U_T(P) \geq U_T(P^\varphi)\).

C4. Monotonicity: let \(P_1, P_2 \in \mathcal{C}r(X)\) and \(P_1 \supseteq P_2\), then \(U_N(P_1) \geq U_N(P_2)\) and \(U_T(P_1) \geq U_T(P_2)\).

C5. The first additivity property: let \(X, Y\) be non-empty finite sets, \(P_X = P(\eta_{\langle A \rangle})\), \(A \in 2^X \setminus \emptyset\), and \(P_Y \in \mathcal{C}r(Y)\), then \(U_T(P_{X \times N}P_Y) = U_T(P_X) + U_T(P_Y)\).

C6. The second additivity property: let \(P \in \mathcal{C}r(X \times Y)\) and \(P_Y = \{P_Y|P \in P \times Y\}\), then \(U_T(P) = \sum_{y \in Y} U_T(P_{Y(y)})\).

C7. Subadditivity: let \(X, Y\) be non-empty finite sets and \(P \in \mathcal{C}r(X \times Y)\), then \(U_T(P_X + U_T(P_Y) \geq U_T(P)\).

C8. Disaggregation: \(U_C(P) + U_N(P) = U_T(P)\) for every \(P \in \mathcal{C}r(X)\).

Axiom C6 can be equivalently exchanged to

\[C^6.\] Let \(\{X_1, \ldots, X_n\}\) be a partition of the set \(X\) and \(P = \sum_{k=1}^n a_k P_k\), where \(P_k \in \mathcal{C}r(X_k)\), \(a_k \geq 0\), \(k = 1, \ldots, m\), \(\sum_{k=1}^m a_k = 1\). Then

\[U_T(P) = \sum_{k=1}^m a_k U_T(P_k) + U_T(P),\]

where \(P \in M_{pr}(Y)\) is such that \(Y = \{1, \ldots, m\}\) and \(P(\{i\}) = a_i, i = 1, \ldots, m\).

What do we know about the possible functionals \(U_T\) and its disaggregations? One known possible choice of \(U_T\) is the maximal entropy \(S_{\text{max}}(P) = \sup_S S(P)\), where \(P \in \mathcal{C}r(X)\); and
the possible disaggregation can be based on the minimal entropy $S_{\min}(P)$, where $P \in \mathcal{C}$, i.e. $U_T = S_{\max}$. $U_C = S_{\min}$, and $U_N = U_T - U_C$. There are several ways to extend the generalized Hartley measure to credal sets. It should be noted that any such extension does not satisfy Axioms C1-C8, as we can see by the next example.

**Example 1** Assume that $P \in \mathcal{C}(X \times Y)$, where $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$. Assume also that any $P$ in $P$ is described by a point $P = (p(x_1,y_1), p(x_1,y_2), p(x_2,y_1), p(x_2,y_2))$, where $p(x_i,y_j) = P\{\{(x_i,y_j)\}\}$, $i, j = 1, 2$, and $P$ has two extreme points: $P_1 = (0, 0.5, 0.5, 0)$ and $P_2 = (0.5, 0, 0, 0.5)$, i.e. $P = \{aP_1 + (1 - a)P_2|a \in [0, 1]\}$. We see that $P_X = \{P_X\}$, where $P_X = (0.5, 0.5, 0.5)$, and $P_Y = \{P_Y\}$, where $P_Y = (0.5, 0.5)$. Assume that $U_N$ satisfies the subadditivity property, i.e. $U_N(P_X) + U_N(P_Y) \geq U_N(P)$. Because, in our example, $U_N(P_X) = U_N(P_Y) = 0$, we can conclude that $U_N(P) = 0$.

Let us check whether the monotonicity w.r.t. mapping is fulfilled in this case. Consider the mapping $\varphi : X \times Y \to Z$ such that $\varphi(x_1,y_1) = \varphi(x_2,y_2) = z_1$, $\varphi(x_1,y_2) = \varphi(x_2,y_1) = z_2$. Then $P^\varphi = P(\varphi_Z)$ and $U_N(P^\varphi) > 0$ for an non-trivial $U_N$, but this contradicts to the monotonicity w.r.t. mapping, since we see that $U_N(P^\varphi) > U_N(P)$.

The first known extension of $GH$ is proposed in [1] and based on the following construction. Let $P \in \mathcal{C}(X)$, then the corresponding coherent lower probability is defined as

$$
\mu(A) = \inf_{P \in \mathcal{C}} P(A), A \in 2^X.
$$

After that we calculate the Möbius transform [12] of $\mu$

$$
m(A) = \sum_{B \subseteq A} (-1)^{|A| - |B|} \mu(B),
$$

and finally,

$$GH_1(\mu) = \sum_{B \in 2^X} m(B)H(B).
$$

The proof of $GH_1$ monotonicity can be found in [1, 7]. Formally, $GH_1$ can be seen as the linear extension of $H$ to coherent lower probabilities, and if we use $GH_1$ on credal sets, then we have the same result if different credal sets generate the same coherent lower probability. $GH_1$ is not subadditive and monotone w.r.t. mapping as follows from the next example.

**Example 2** Consider the credal set from Example 1. Let us compute values of the coherent lower probability $\mu$. For this purpose, let us denote $u_1 = (x_1, y_1)$, $u_2 = (x_1, y_2)$, $u_3 = (x_2, y_1)$, $u_4 = (x_2, y_2)$. Then $\mu(\{u_1, u_2\}) = \mu(\{u_1, u_3\}) = \mu(\{u_1, u_4\}) = 0.5, \mu(\{u_2, u_3\}) = 0.5, \mu(\{u_2, u_4\}) = 0.5, \mu(\{u_3, u_4\}) = 0.5$ is equal to zero on other sets in $2^X$. The computation of $m$ results in $m(\{u_1, u_2\}) = m(\{u_1, u_3\}) = m(\{u_2, u_3\}) = m(\{u_3, u_4\}) = 0.5, m(\{u_1, u_2, u_3\}) = m(\{u_1, u_2, u_4\}) = m(\{u_1, u_3, u_4\}) = 0.5, m(\{u_2, u_3, u_4\}) = 0.5, m(\{u_1, u_2, u_3, u_4\}) = 1$. $m$ is equal to zero on other sets in $2^X$. Assume that $H(B) = \log_2 |B|$ for every $B \neq \emptyset$. Then $GH_1(\mu) = 2 - 2\log_2 3 + 2 \approx 0.83$. We see that both properties C3 and C7 are not fulfilled.

The second extension, introduced in [5], is based on the inner approximation of $GH$:

$$
GH_2(\mu) = \sup(GH(\mu)|P(\mu) \subseteq P, \mu \in M_{\text{cred}}(X)).
$$

In [5], a reader can find the proof that $GH_2$ can be used for disaggregation of $U_T$, and it is subadditive. Formally, we can also check the first and the second additivity properties defined as

**The first additivity property:** let $X, Y$ be non-empty finite sets, $P_X = P(\eta_X), A \in 2^X \setminus \emptyset$, and $P_Y \in \mathcal{C}(Y)$, then $U_N(P_X \times_N P_Y) = U_N(P_X) + U_N(P_Y)$.

**The second additivity property:** let $\{X_1, ..., X_k\}$ be a partition of the set $X$ and $P = \sum_{i=1}^k a_i P_i$, where $P_i \in \mathcal{C}(X_i), a_i \geq 0, i = 1, ..., k, \sum_{i=1}^k a_i = 1$. Then

$$U_N(P) = \sum_{i=1}^k a_i U_N(P_i).
$$

Clearly, the last property is the counterpart of C6*, in which we drop the term $U_N(P) = 0, P \in M_{\text{cred}}(Y)$.

**Remark 1** In [5] a reader can find results that $GH_2$ obeys the first and second additivity properties. Obviously, it can be used for disaggregation of a measure of total uncertainty. In the next section, we will check these properties for $GH_1$.

6. Properties of $GH_1$

**Theorem 1** $GH_1$ obeys the first additivity property.

**Theorem 2** $GH_1$ obeys the second additivity property.

**Lemma 1** Let $P \in M_{\text{cred}}(X)$ and $P(\{x\}) > 0$ for all $x \in X$. Then the largest credal set $P \in \mathcal{C}(X)$ such that $P \in P$ and $S_{\max}(P) = S(P)$ is $P = \{Q \in M_{\text{cred}}(X)\mid E_Q(f_S) \geq E_P(f_S)\}$, where $f_S(x) = \log P(\{x\}), x \in X$.

**Remark 2** Consider how we can generalize Lemma 1 for the case, when $P$ takes values equal to zero on some singletons. Assume that $P(\{x\}) = 0$ for some $x \in X$ and $Q(\{x\}) > 0$, then $\lim_{a \to 0} S(aQ + (1 - a)P) = +\infty$, i.e. $P(\{x\}) = 0$ implies that $Q(\{x\}) = 0$ for every $Q \in P$, i.e. we can reduce this problem considering only those elements of $X$, where $P(\{x\}) > 0$ like in Lemma 1.
Remark 3 Let $X = \{x_1, \ldots, x_n\}$, then the special case of Lemma 1 is when $P(\{x_i\}) = 1/n$, $i = 1, \ldots, n$. In this case, $E Q(f) - E r(f) = \sum_{x \in X} (Q(\{x\}) - P(\{x\})) \ln(1/n) = 0$, for every $Q \in M_{pr}(X)$, i.e. $P = M_{pr}(X)$. Note that for other cases $P$ is a boundary point of $\mathcal{P}$.

The next proposition shows the construction of coherent lower probabilities $\mu$ on $2^X$ with $P \in \mathcal{P}(\mu)$ and $S_{\max}(\mathcal{P}(\mu)) = S(P)$ for a given $P \in M_{pr}(X)$.

**Proposition 1** Let $P \in M_{pr}(X)$ such that $P(\{x\}) > 0$ for all $x \in X$. Consider a subset $\mathcal{A} \subseteq 2^X$ and a credal set defined by

$$
\mathcal{P} = \{Q \in M_{pr}(X) \mid A \in \mathcal{A} : Q(A) \geq P(A)\},
$$

(1)

Introduce the corresponding coherent lower probability $\mu$ defined by $\mu(A) = \inf(\{Q(A) \mid Q \in \mathcal{P}\})$, $A \in 2^X$. Then $P(\mu) = P \in \mathcal{P}(\mu)$, and $S_{\max}(\mathcal{P}(\mu)) = S(P)$ if there are $a_k \geq 0, A \in \mathcal{A}$, and $b \in \mathcal{B}$, such that $\sum A \in \mathcal{A} A_1 b + b X = f_5$, where $f_5$ is defined like in Lemma 1.

**Corollary 1** Let $P \in M_{pr}(X)$, $P(\{x\}) > 0$ for all $x \in X$, and $\{P(\{x\}) \mid x \in X, i = 1, \ldots, k\}$, where values $a_i$ are indexed such that $a_1 > a_2 > \ldots > a_k > 0$. Define a credal set $\mathcal{P}$ by formula (1), where and $A_i = \{x \in X \mid P(\{x\}) \geq a_i\}$, $i = 1, \ldots, k$. Then $P \in \mathcal{P}$ and $S_{\max}(\mathcal{P}(\mu)) = S(P)$.

Remark 4 It is well known that the credal set from Corollary 1 can be generated by the $\mu \in M_{prop}(X)$ whose body of evidence is $\mathcal{A}$ and the corresponding bba $m$ is defined as $m(A_i) = a_i - a_{i+1}$, $i = 1, \ldots, k$, where $a_{k+1} = 0$ by convention. Such a $\mu$ is a necessity measure, since focal elements of $\mu$ are linearly ordered w.r.t. the inclusion relation.

**Example 3** Let $X = \{x_1, x_2, x_3, x_4\}$ and $P \in M_{pr}(X)$ is defined by $P(\{x_i\}) = q_i$, $i = 1, \ldots, 4$, where $q$ is the positive root of the equation $q_1^2 + q_1^2 + q_2 + q_1 = 0$ (as $0.518 \ldots$). Let us construct $\mu$ like in Corollary 1. Then $m \in M_{prop}(X)$ whose body of evidence is $\mathcal{A} = \{A_i\}_{i=1}^4$, where $A_i = \{x_1, \ldots, x_i\}$, and $m(A_i) = q_i - q_i^{i+1}$, $i = 1, \ldots, 3$, $m(A_4) = q_4$. Let us construct $\mu$ like in Proposition 1 by the system of sets $\mathcal{A} = \{\{x_1, x_2\}, \{x_1, x_2, x_3\}\}$. Let us show that there is a representation

$$
y_1 x_1 + y_2 x_2 + y_3 x_3 + y_4 x_1 x_2 x_3 x_4 = f_5,
$$

(2)

where $y_1, y_2 \geq 0$, and $y_3 \in \mathbb{R}$. We see that (2) is equivalent to the following system of linear equations:

$$
\begin{align*}
y_1 + y_2 + y_3 &= \ln q_1, \\
y_1 + y_3 &= 2 \ln q_1, \\
y_2 + y_3 &= 3 \ln q_1, \\
y_3 &= 4 \ln q.
\end{align*}
$$

The solution of this system is $y_1 = -2 \ln q, y_2 = -\ln q, y_3 = 4 \ln q$, and $y_i, i = 1, \ldots, 3$, satisfy the required conditions. The values of $\mu$ and $\mu_1$ are given in Table 1.

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$\mu$</th>
<th>$\mu_1$</th>
<th>$\mu_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$q$</td>
<td>$q - q^4$</td>
<td>$q - q^4$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$q + q^2$</td>
<td>$q + q^2$</td>
<td>$q + q^2$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$q + q^2$</td>
<td>$q + q^2$</td>
<td>$q + q^2$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>$q$</td>
<td>$q - q^4$</td>
<td>$q - q^4$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$q + q^2$</td>
<td>$q + q^2$</td>
<td>$q + q^2$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$q$</td>
<td>$q - q^3$</td>
<td>$q - q^3$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$q$</td>
<td>$q - q^3$</td>
<td>$q - q^3$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$q$</td>
<td>$q - q^3$</td>
<td>$q - q^3$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$q$</td>
<td>$q - q^3$</td>
<td>$q - q^3$</td>
</tr>
</tbody>
</table>

By Proposition 1, $P \in \mathcal{P}(\mu)$, $P \in \mathcal{P}(\mu_1)$, and $S_{\max}(\mathcal{P}(\mu_1)) = S_{\max}(\mathcal{P}(\mu)) = S(P)$. Note that the set $\mathcal{M} = \{\mu \in M_{\text{prop}}(X) \mid P \in \mathcal{P}(\mu), S_{\max}(\mathcal{P}(\mu)) = S(P)\}$ does not contain the smallest element in general. To show this consider a coherent lower probability $\mu_2$, defined by $\mu_2(A) = \min(\mu(A), \mu_1(A)), A \in 2^X$. The values of $\mu_2$ are also given in Table 1. Consider a probability measure $P_1$ with values $P_1(\{x_1\}) = q - q^2$, $P_1(\{x_2\}) = q^2 + q^4$, $P_1(\{x_3\}) = (q^2 + q^4)/2$. It is easy to check that $P_1 \in \mathcal{M}(\mu_2)$.

$$
S(P_1) = (q - q^2) \ln(q - q^2) - (q^2 + q^4) \ln(q^2 + q^4) - (q^2 + q^4) \ln \frac{q^2 + q^4}{2} = 1.202\ldots
$$

therefore, $\mu_2 \notin \mathcal{M}$.

**Proposition 2** Let $P \in M_{pr}(X)$ with $P(\{x\}) > 0$ for all $x \in X$ and

$$
\mathcal{M} = \{v \in M_{2 - \text{mon}}(X) \mid P \in \mathcal{P}(v), S_{\max}(\mathcal{P}(v)) = S(P)\}.
$$

Consider a $\mu \in \mathcal{M}$ constructed like in Corollary 1. Then $\mu \geq \mu$ for every $v \in \mathcal{M}$.

**Remark 5** Note the result formulated in Proposition 2 does not contradict Example 3, where we construct $\mu_1$ such that $\mu_1 \not\geq \mu$, since $\mu_1 \notin M_{2 - \text{mon}}(X)$. This can be seen from the inequality:

$$
\mu_1(\{x_1, x_2\}) + \mu_1(\{x_1, x_3\}) = 2q + q^2 + q^3 > \mu_1(\{x_1, x_2, x_3\}) = 2q + q^2 - q^4.
$$

The above results can be seen as the investigation of the inner approximation of the credal set defined in Lemma 1 by coherent lower probabilities. The next proposition describes the upper approximation of this credal set.

**Proposition 3** Let $X = \{x_1, \ldots, x_n\}$ and $P \in M_{pr}(X)$ such that $P(\{x_1\}) \geq P(\{x_2\}) \geq \ldots \geq P(\{x_n\}) > 0$. Consider the corresponding credal set $\mathcal{P} = \{Q \in M_{pr}(X) \mid E Q(f_5) \geq E_{P}(f_5)\}$ from Lemma 1 and a coherent lower probability $\mu$ on $2^X$ defined by $\mu(A) = \inf(\{P(A) \mid P \in \mathcal{P}\})$, where $A \in 2^X$. Then
In the sequel, we will represent $k$ and $\sum_{i=1}^{m} x_i$ where values $a_i$ are indexed such that $a_1 > a_2 > \ldots > a_k > 0$. Let $P$ be the credal set defined in Lemma 1 and a coherent lower probability $\mu$ on $2^X$ defined by $\mu(A) = \inf (Q(A) | Q \in P)$. Then

$$P = \{Q \in M_{pr}(X)| \forall A \in \mathcal{A} : Q(A) \geq \mu(A)\},$$

where $\mathcal{A} = \{A_i\}_{i=1}^{k}$ and $A_i = \{x \in X | P(x) \geq a_i\}$, $i = 1, \ldots, k$.

**Remark 6** Since the set $\mathcal{A}$ defined in Corollary 2 is linearly ordered by the inclusion relation, the measure $\mu$ is a necessity measure (a consonant belief function). Let us compute at first values of $\mu$ on $\mathcal{A}$:

1) $\mu(A_i) = 0$ if $E_P(f_s) - f_s(a_{i+1}) \leq 0$;
2) $\mu(A_i) = (E_P(f_s) - f_s(a_{i+1}) - (f_s(a_i) - f_s(a_{i+1}))$ if $E_P(f_s) - f_s(a_{i+1}) > 0$ and $i \neq k$;
3) $\mu(A_k) = 1$, since $A_k = X$.

We can compute the corresponding bba $m$ on by $m(A_i) = \mu(A_i) - \mu(A_{i-1})$, $i = 1, \ldots, k$, where $A_0 = \emptyset$ by convention. Using the above result, we can compute $GH(\mu)$ by the formula:

$$GH(\mu) = \sum_{i=1}^{k} (\mu(A_i) - \mu(A_{i-1})) |A_i| = \sum_{i=1}^{k-1} \mu(A_i) |A_i| - |A_{i+1}| + \mu(A_k) |A_k|.$$

Let $m(A_k) \neq 1$ and $j = \min\{i \in \{1, \ldots, k - 1\} | E_P(f_s) - f_s(a_{i+1}) > 0\}$, then

$$GH(\mu) = \sum_{i=j}^{k-1} (E_P(f_{s_i}) - f_{s_i}(a_{i+1}) - (f_{s_i}(a_i) - f_{s_i}(a_{i+1})) |A_i| - |A_{i+1}| + \mu(A_k) |A_k|.$$

Since $E_P(f_s) - f_s(a_{i+1}) = 1 - \ln a_i - \ln a_{i+1}$ and $\sum_{i=j}^{k-1} (|A_i| - |A_{i+1}|) = |A_j| - |A_k|$, we get

$$GH(\mu) = \ln |A_j| + \sum_{i=j}^{k-1} \frac{(\ln a_i - E_P(f_{s_i}) - (\ln a_{i+1}) - |A_i|)}{|A_{i+1}| - |A_i|}.$$

In the sequel, we will represent $GH(\mu)$ as $GH(\mu) = \ln |A_j| + (\ln a_i - E_P(f_{s_i}))F(a)$, where

$$F(a) = \sum_{x \in X} \ln |A_{i+1}| - |A_i|.$$

**Example 4** Let us construct an example of a $P \in Cr(X)$ for which $GH_1(P) > S_{\max}(P)$. Assume that $P$ is constructed like in Lemma 1, i.e. there is a $P \in M_{pr}(X)$ with $P(x) > 0$ for all $x \in X$ and such that $P = \{Q \in M_{pr}(X) | E_Q(f_s) \geq E_P(f_{s_i})\}$. In our example we will assume that $A_1 = \{x_1\}$, $A_2 = \{x_1, \ldots, x_n\}$, $A_3 = X = \{x_1, \ldots, x_n\}$, $a_1 = q^2(a_3$ and $a_2 = qa_3$. In addition, these parameters are chosen such that $E_P(f_{s_i}) - \ln(a_2) = 0$. Then

$$E_P(f_{s_i}) - \ln(a_2) = a_1 \ln (a_1/a_2) + a_3(n-m) \ln (a_3/a_2) = 0,$$

or $a_3^2 \ln q - (n-m)a_3 \ln q = 0$. Thus, we can choose $q = \sqrt{n-m}$ and the norming condition implies that $(n-m)a_3 + (m-1)\sqrt{n-m}a_3 + (n-m)a_3 = 1$, i.e.

$$a_3 = \frac{1}{2(n-m) + (m-1)\sqrt{n-m}},$$

$$F(a) = \frac{\ln |A_3| - \ln |A_2|}{\ln a_1 - \ln a_3} = \frac{\ln n - \ln m}{\ln n - \ln m},$$

$$GH(\mu) = \ln |A_2| + (\ln a_1 - E_P(f_{s_i}))F(a) = \ln m + (\ln a_1 - \ln a_2)F(a) =$$

$$\ln m + 0.5\ln(n-m)\ln(n-m) = 0.5(\ln n + \ln m),$$

$$a_2 = \frac{\sqrt{n-m}}{2(n-m) + (m-1)\sqrt{n-m}} = \frac{1}{2\sqrt{n-m} + (m-1)},$$

$$S(P) = - E_P(f_{s_i}) = - \ln a_2 = \ln(2\sqrt{n-m} + m-1),$$

We see that

$$S(P) - GH(\mu) = \ln \left(\frac{2\sqrt{n-m} + m-1}{\sqrt{mn}}\right).$$

Let us denote $s = m/n$, then

$$\lim_{n \to \infty} (S(P) - GH(\mu)) = \lim_{n \to \infty} \left(2\sqrt{n-1} + nx - 1\right) = 0.5\ln x < 0,$$

i.e. $S(P) - GH(\mu) < 0$, if $n$ is sufficiently large. In particular, if $n = 100$ and $m = 36$, then $GH_1(P) = GH(\mu) = \ln 60$, $S_{\max}(P) = S(P) = \ln 51$, and we see that $GH_1(P) > S_{\max}(P)$.

**7. The Hartley Measure on Credal Sets**

Assume that uncertainty is described by a probability measure $P \in M_{pr}(X)$ and the choice of the optimal decision is based on the expected utility, i.e. every decision is described by a function $f : X \to \mathbb{R}$ and the expected utility is defined as

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

If uncertainty is described by a credal set $P$, then we only know the lower and upper bounds of expected utility defined by
Then these probability measures are pairwise fully contradictory probability measures in a credal set. The next lemma shows that we can check the full contradiction (inconsistent) if every two vertices in it are not adjacent. Thus, the problem of finding the maximal system of pairwise fully contradictory probability measures is equivalent to the problem of finding the independent set in $G$ with the largest cardinality.

**Remark 8** Assume that $P_1, \ldots, P_k$ are the extreme points of a credal set $P$ with the corresponding sets $A_i, i = 1, \ldots, k$ defined like in Proposition 4. Then the problem of finding the maximal system of pairwise fully contradictory probability measures in $P$ can be reformulated in terms of graph theory. Consider the undirected graph $G$, whose vertices are sets $A_i, i = 1, \ldots, k$, and there is an edge between $A_i$ and $A_j, i \neq j$, iff $A_i \cap A_j \neq \emptyset$. A set of vertices is called independent if every two vertices in it are not adjacent. Thus, the problem of finding the maximal system of pairwise fully contradictory probability measures is equivalent to the problem of finding the independent set in $G$ with the largest cardinality.

**Remark 9** We can conclude from Remark 8 that the search of the maximal system of pairwise fully contradictory probability measures is NP-hard, however, for special cases we have the solutions. For example, if $P = P(\eta_{\{A\}})$, then the extreme points of this set are probability measures $\eta_{\{x\}}, x \in A$, which are pairwise fully contradictory, and the cardinality of this system is $|A|$. Based on Remark 9, we introduce the following definition.

**Definition 1** Let $P \in Cr(X)$, then the Hartley measure $H(P) = \ln(n(P))$, where $n(P)$ is the largest number of pairwise fully contradictory probability measures in $P$.

**Lemma 4** Let $P \in Cr(X)$, then there is a credal set $P_1 \subseteq P$ and a mapping $\varphi : X \to Y$ such that $P_1^\varphi = P(\eta(Y))$ and $H(P) = |Y|$.

The main properties of $H$ are given in the following proposition.

**Proposition 6** The Hartley measure $H$ on $Cr$ has the following properties:

1. $H(P(\eta_{\{A\}})) = \ln |A|$ for every $A \neq \emptyset$;
2. $H(P) = 0$ if $P = \{P\}$, where $P \in M_P$;
3. let $P \in Cr(X)$ and $\varphi : X \to Y$, then $H(P^{\varphi}) \leq H(P)$; in addition, $H(P^{\varphi}) = H(P)$ if $\varphi$ is an injection;
4. let $P_1, P_2 \in Cr(X)$ and $P_1 \subseteq P_2$, then $H(P_1) \leq H(P_2)$;
5. let $X, Y$ be non-empty finite sets, $P_X = P(\eta_A), A \in 2^X \setminus \emptyset$, and $P_Y \in Cr(Y)$, then $H(P_X \times_A P_Y) \geq H(P_X) + H(P_Y);

6. let $\{X_1, X_2\}$ be a partition of the set $X$ and $P = aP_1 + (1 - a)P_2$, where $P_i \in Cr(X_i), i = 1, 2$, and $a \in (0, 1)$, then $U\eta(P) = \min\{U\eta(P_1), U\eta(P_2)\}$.

7. $H(P) \leq S_{\text{max}}(P)$ for every $P \in Cr$.

**Remark 10** Proposition 6 implies that $H$ can be served as a non-specificity measure for $S_{\text{max}}$, disgregation, however, it does not obey the second additivity property according to the statement 6) of Proposition 6, and as follows from the next examples it is not subadditive, and it does not obey the first additivity property.

**Example 5** Consider the credal set $P \in Cr(X \times Y)$ from Example 1. We see that $P_X = \{P_x\}$ and $P_Y = \{P_y\}$, i.e. $H(P_X) = H(P_Y) = 0$, and $P$ has only two extreme points $P_1 = (0, 0, 0, 0, 0, 0)$ and $P_2 = (0, 0, 0, 0, 0, 0)$. Because $P_1$ and $P_2$ are fully contradictory, $H(P) = \ln 2$, and $H(P) > H(P_X) + H(P_Y)$.

**Example 6** Let $P = P_X \times YP_Y$, where $X = \{x_1, x_2\}$, $Y = \{y_1, y_2, y_3\}$, $P_X = P(\eta_{\{x\}})$ and $P_Y \in Cr(Y)$, defined by extreme points $P_Y^{(i)}, i = 1, 2, 3$, with probabilities: $P_Y^{(1)}(\{y_1\}) = P_Y^{(1)}(\{y_2\}) = 0.5, P_Y^{(2)}(\{y_2\}) = P_Y^{(2)}(\{y_3\}) = 0.5, P_Y^{(3)}(\{y_1\}) = P_Y^{(3)}(\{y_3\}) = 0.5$. Consider probability measures $P^{(i)} \in P, i = 1, 2, 3$, defined by
w.r.t. mapping is not fulfilled for wise fully contradictory, we conclude that $H$ 

Then $S \in \{x, y\}$. Clearly, $S_{\text{max}}(\mu) = -0.5 \ln 0.25 - 0.5 \ln 0.5 = 1.5 \ln 2$, $S_{\text{min}}(\mu) = \ln 2$. Therefore, $S_{\text{max}}(\mu) - S_{\text{min}}(\mu) = 0.5 \ln 2$. Consider a mapping $\varphi : X \to X$ defined by $\varphi(x_i) = x_i$, $i = 1, 2, 3$. 

Then $\mu^\varphi = 0.5 \eta((x_1, x_2)) + 0.5 \eta((x_3))$, $S_{\text{max}}(\mu^\varphi) = \ln 2$, $S_{\text{min}}(\mu^\varphi) = 0$, i.e. $U_N(\mu^\varphi) > U_N(\mu)$ if $U_N = S_{\text{max}} - S_{\text{min}}$.

Based on Example 1, we see that all desirable properties of non-specificity measures cannot be fulfilled simultaneously, for example, if $U_N$ is subadditive, then it is not monotone w.r.t. mapping and weakly sensitive. In opinion, the sensitivity and monotonicity w.r.t. mapping of a non-specificity measure has a higher importance, than its subadditivity. It is possible to increase sensitivity of the Hartley measure on credal sets introducing $\varepsilon$-Hartley measures. The idea consists in the following. Probability measures $P_1, P_2 \in M_{pr}(X)$ are called $\varepsilon$-contradictory for $\varepsilon \in (0, 1]$ if $d(P_1, P_2) > \varepsilon$. Then $H_\varepsilon(P)$, where $P \in Cr(X)$ is the logarithm of the largest number of pairwise $\varepsilon$-contradictory measures in $P$. However, the investigation of these measures is not included in this paper.

It is possible to use the above non-specificity measures on coherent lower probabilities or 2-monotone measures. At the first glance, $G_{H_1}$ on 2-monotone measures looks optimal. This may be the topic for further research.

8. Discussion and Conclusion

In previous sections, we have analyzed properties of non-specificity measures. Now we are ready to present the obtained results in Table 2. In Table 2, you can see also properties that characterize sensitivity of non-specificity measures formulated as follows:

**Weak sensitivity:** let $P \in Cr(X)$, then $U_N(P) = 0$ iff $P = \{P\}$, where $P \in M_{pr}(X)$.

**Strong sensitivity:** let $P_1, P_2 \in Cr(X)$ and $P_1 \subset P_2$, then $U_N(P_1) < U_N(P_2)$.

A reader can check that $U_N = S_{\text{max}} - S_{\text{min}}$ is not subadditive using the credal set example from Table 1. The monotonicity w.r.t. mapping is not fulfilled for $U_N = S_{\text{max}} - S_{\text{min}}$ as shown in the next example.

**Example 7** Assume that $X = \{x_1, x_2, x_3\}$ and $\mu = 0.5 \eta((x_1, x_2)) + 0.5 \eta((x_3))$. Clearly, $S_{\text{max}}(\mu) = -0.5 \ln 0.25 - 0.5 \ln 0.5 = 1.5 \ln 2$, $S_{\text{min}}(\mu) = \ln 2$. Therefore, $S_{\text{max}}(\mu) - S_{\text{min}}(\mu) = 0.5 \ln 2$. Consider a mapping $\varphi : X \to X$ defined by $\varphi(x_i) = x_i$, $i = 1, 2, 3$.

Then $\mu^\varphi = 0.5 \eta((x_1, x_2)) + 0.5 \eta((x_3))$, $S_{\text{max}}(\mu^\varphi) = \ln 2$, $S_{\text{min}}(\mu^\varphi) = 0$, i.e. $U_N(\mu^\varphi) > U_N(\mu)$ if $U_N = S_{\text{max}} - S_{\text{min}}$.

**Table 2: Properties of non-specificity measures.**

<table>
<thead>
<tr>
<th>Properties \ $U_N$</th>
<th>$S_{\text{max}} - S_{\text{min}}$</th>
<th>$GH_1$</th>
<th>$GH_2$</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boundary conditions</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Monotonicity</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Monotonicity w.r.t. mapping</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>The first additivity property</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>The second additivity property</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Subadditivity</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Weak sensitivity</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Strong sensitivity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Disaggregation</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Since probability measures $P^{(i)} \in P$, $i = 1, 2, 3$, are pairwise fully contradictory, we conclude that $H(P) \geq \ln 3$. We see that $H(P_X) = \ln 2$ and $H(P_Y) = \ln 1 = 0$. Therefore, $H(P) > H(P_X) + H(P_Y)$, i.e. the first additivity property is not fulfilled.

References


