

Measure of Information Content of Basic Belief Assignments

Jean Dezert¹[0000–0003–3474–9186], Albena Tchamova²[0000–0002–8250–0521],
and Deqiang Han³[0000–0001–5603–796X]

¹ The French Aerospace Lab
Chemin de la Hunière, F-91761 Palaiseau, France.

`jean.dezert@onera.fr`

² Institute for IC & Technologies. Bulgarian Acad. of Sci., Sofia, Bulgaria.
`tchamova@bas.bg`

³ Institute of Integrated Automation, Xi'an Jiaotong University,
Xi'an, China 710049.
`deqhan@xjtu.edu.cn`

Abstract. In this paper, we present a measure of Information Content (IC) of Basic Belief Assignments (BBAs), and we show how it can be easily calculated. This new IC measure is interpreted as the dual of the effective measure of uncertainty (i.e. generalized entropy) of BBAs developed recently.

Keywords: Belief functions · Information content · Generalized entropy.

1 Introduction

Information quality (IQ) evaluation is of major importance for information processing and for helping the decision-making under uncertainty. In [1], the authors introduced the Accessibility, Interpretability, Relevance, and Integrity concepts as main attributes to describe the information quality in the context of assurance and belief networks, but unfortunately they present only general concepts without explicit formulas to evaluate quantitatively these attributes. In several recent books devoted to IQ [2–5], the authors proposed different models and methods of IQ evaluations. Recently in [6], Bouhamed et al. proposed a quantitative IQ evaluation using the possibility theory framework, which could be extended to the belief functions theory framework with further investigations. In this latter work, the information quantity component being necessary for the IQ evaluation is based on Gini's entropy rather than classical Shannon entropy. From the examination of these aforementioned references (and some references therein), it is far from obvious to make a clear justified choice among all these methods, especially when we model the uncertain information by belief functions (BF). What is clear however is that several distinct factors (or components) must be taken into account in the IQ evaluation mechanism. In this paper we focus on one of these components which is the Information Content (IC) component that we consider as the very (if not the most) essential component for IQ evaluation

and indispensable for developing an effective IQ evaluation method in future research works.

It is worth noting that we do not address directly the whole IQ evaluation problem in this work but to provide a mathematical solution for measuring the IC of any Basic Belief Assignments (BBA) in the belief functions (BF) framework. Our new IC measure is interpreted as the dual of an effective Measure of Uncertainty (MoU) developed recently [7]. We show how to calculate the IC of a BBA, and we also discuss the notion of information gain and information loss in the BF context. In our opinion, we cannot define a measure of Information Content independently of a Measure of Uncertainty (MoU) because they must be strongly related to each other. Actually these measures are two different sides of a same *abstract coin* we would say. On one side (the uncertainty side), more uncertainty content we have harder is the decision or choice to make, and on the other side (the information side) more information content we have easier and stronger is the decision or choice to make. This very simple and natural basic principle will be clarified mathematically next. So, the measure of information content of a BBA must reflect somehow the easiness and strength in the choice of an element of the frame of discernment drawn from the BBA (i.e. in the decision-making). This paper is organized as follows. After a brief recall of basics of belief functions in section 2, we recall the effective MoU adopted in this work in section 3. Section 4 defines the measure of information content of a BBA and the information granules vector. Section 5 introduces the notions of information gain and information loss. Conclusions and perspectives appear in the last section.

2 Belief functions

The belief functions (BF) were introduced by Shafer [8] for modeling epistemic uncertainty, reasoning about uncertainty and combining distinct sources of evidence. The answer of the problem under concern is assumed to belong to a known finite discrete frame of discernment (FoD) $\Theta = \{\theta_1, \dots, \theta_N\}$ where all elements (i.e. members) of Θ are exhaustive and mutually exclusive. The set of all subsets of Θ (including empty set \emptyset , and Θ) is the power-set of Θ denoted by 2^Θ . The number of elements (i.e. the cardinality) of the power-set is $2^{|\Theta|}$. A (normalized) basic belief assignment (BBA) associated with a given source of evidence is a mapping $m^\Theta(\cdot) : 2^\Theta \rightarrow [0, 1]$ such that $m^\Theta(\emptyset) = 0$ and $\sum_{X \in 2^\Theta} m^\Theta(X) = 1$. A BBA $m^\Theta(\cdot)$ characterizes a source of evidence related with a FoD Θ . For notation shorthand, we can omit the superscript Θ in $m^\Theta(\cdot)$ notation if there is no ambiguity on the FoD we work with⁴. The quantity $m(X)$ is called the mass of belief for X . The element $X \in 2^\Theta$ is called a focal element (FE) of $m(\cdot)$ if $m(X) > 0$. The set of all focal elements of $m(\cdot)$ is denoted⁵ by $\mathcal{F}_\Theta(m) \triangleq \{X \in 2^\Theta | m(X) > 0\}$.

⁴ However, we will keep $m^\Theta(\cdot)$ notation when very necessary.

⁵ \triangleq means *equal by definition*.

The belief and the plausibility of X are defined for any $X \in 2^\Theta$ by [8]

$$Bel(X) = \sum_{Y \in 2^\Theta | Y \subseteq X} m(Y) \quad (1)$$

$$Pl(X) = \sum_{Y \in 2^\Theta | X \cap Y \neq \emptyset} m(Y) = 1 - Bel(\bar{X}). \quad (2)$$

where $\bar{X} \triangleq \Theta \setminus \{X\}$ is the complement of X in Θ .

One has always $0 \leq Bel(X) \leq Pl(X) \leq 1$, see [8]. For $X = \emptyset$, $Bel(\emptyset) = 0$ and $Pl(\emptyset) = 0$, and for $X = \Theta$ one has $Bel(\Theta) = 1$ and $Pl(\Theta) = 1$. $Bel(X)$ and $Pl(X)$ are often interpreted as the lower and upper bounds of unknown probability $P(X)$ of X , that is $Bel(X) \leq P(X) \leq Pl(X)$. To quantify the uncertainty (i.e. the imprecision) of $P(X) \in [Bel(X), Pl(X)]$, we use the notation $u(X) \in [0, 1]$ defined by

$$u(X) \triangleq Pl(X) - Bel(X) \quad (3)$$

The quantity $u(X) = 0$ if $Bel(X) = Pl(X)$ which means that $P(X)$ is known precisely, and one has $P(X) = Bel(X) = Pl(X)$. One has $u(\emptyset) = 0$ because $Bel(\emptyset) = Pl(\emptyset) = 0$, and one has $u(\Theta) = 0$ because $Bel(\Theta) = Pl(\Theta) = 1$. If all focal elements of $m(\cdot)$ are singletons of 2^Θ the BBA $m(\cdot)$ is a Bayesian BBA because $\forall X \in 2^\Theta$ one has $Bel(X) = Pl(X) = P(X)$ and $u(X) = 0$. Hence the belief and plausibility of X coincide with a probability measure $P(X)$ defined on the FoD Θ . The vacuous BBA characterizing a totally ignorant source of evidence is defined by $m_v(X) = 1$ for $X = \Theta$, and $m_v(X) = 0$ for all $X \in 2^\Theta$ different from Θ . This particular BBA has played a major role in the establishment of a new effective measure of uncertainty of BBA defined in [7].

3 Generalized entropy of a BBA

In [9] we did analyze in details forty-eight measures of uncertainty (MoU) of BBAs by covering 40 years of research works on this topic. Some of these MoUs capture only a particular aspect of the uncertainty inherent to a BBA (typically, the non-specificity and the conflict). Other MoUs propose a total uncertainty measure to capture jointly several aspects of the uncertainty. Unfortunately, most of these MoUs fail to satisfy four very simple reasonable and essential desiderata, and so they cannot be considered as really effective and useful. Actually only six MoUs can be considered as effective from the mathematical sense presented next, but unfortunately they appear as conceptually defective and disputable, see discussions in [9]. That is why, a better effective measure of uncertainty (MoU), i.e. generalized entropy of BBAs has been developed and presented in [7]. The mathematical definition of this new effective entropy is given by

$$U(m) = \sum_{X \in 2^\Theta} s(X) \quad (4)$$

with

$$s(X) \triangleq -m(X)(1 - u(X)) \log(m(X)) + u(X)(1 - m(X)) \quad (5)$$

The quantity $-(1 - u(X)) \log(m(X)) = (1 - u(X)) \log(1/m(X))$ entering in $s(X)$ in (5) is the surprisal⁶ $\log(1/m(X))$ of X discounted by the confidence $(1 - u(X))$ one has on the precision of $P(X)$. The term $-m(X)(1 - u(X)) \log(m(X))$ is the weighted discounted surprisal of X . The term $u(X)(1 - m(X))$ entering in (5) corresponds to the imprecision of $P(X)$ discounted by $(1 - m(X))$ because the greater $m(X)$ the less one should take into account the imprecision $u(X)$ in the MoU. The quantity $s(X)$ is the uncertainty contribution related to element X (named the *entropiece* of X) in the MoU $U(m)$. This entropiece $s(X)$ involves $m(X)$ and the imprecision $u(X) = Pl(X) - Bel(X)$ about the unknown probability of X in a subtle interwoven manner. The cardinality of X is indirectly taken into account in the derivation of $s(X)$ thanks to $u(X)$ which requires the derivation of $Pl(X)$ and $Bel(X)$ functions depending on the cardinality of X . Because $u(X) \in [0, 1]$ and $m(X) \in [0, 1]$ one has $s(X) \geq 0$, and $U(m) \geq 0$. The quantity $U(m)$ is expressed in *nats* because we use the natural logarithm. $U(m)$ can be expressed in *bits* by dividing the $U(m)$ value in *nats* by $\log(2) = 0.69314718\dots$. This measure of uncertainty $U(m)$ is a continuous function in its basic belief mass arguments because it is a summation of continuous functions. In formula (5), we always take $m(X) \log(m(X)) = 0$ when $m(X) = 0$ because $\lim_{m(X) \rightarrow 0^+} m(X) \log(m(X)) = 0$ which can be proved using L'Hôpital rule [11]. Note that for any BBA m , one has always $s(\emptyset) = 0$ because $m(\emptyset) = 0$ and $u(\emptyset) = Pl(\emptyset) - Bel(\emptyset) = 0 - 0 = 0$. For the vacuous BBA, one has $s(\Theta) = 0$ because $m_v(\Theta) = 1$ and $u(\Theta) = Pl(\Theta) - Bel(\Theta) = 1 - 1 = 0$.

The set $\{s(X), X \in 2^\Theta\}$ of the entropieces values $s(X)$ can be represented by an entropiece vector $\mathbf{s}(m^\Theta) = [s(X), X \in 2^\Theta]^T$, where any order of elements X of the power set 2^Θ can be chosen. For simplicity, we suggest to use the classical N -bits representation (if $|\Theta| = N$) with the increasing order - see the next example.

This measure of uncertainty $U(m)$ is effective because it can be proved (see proofs in [7]) that it satisfies the following four essential properties:

1. $U(m) = 0$ for any BBA $m(\cdot)$ focused on a singleton X of 2^Θ .
2. $U(m_v^\Theta) < U(m_v^{\Theta'})$ if $|\Theta| < |\Theta'|$.
3. $U(m) = -\sum_{X \in \Theta} m(X) \log(m(X))$ if the BBA $m(\cdot)$ is a Bayesian BBA. Hence, $U(m)$ reduces to Shannon entropy [12] in this case.
4. $U(m) < U(m_v)$ for any non-vacuous BBA $m(\cdot)$ and for the vacuous BBA $m_v(\cdot)$ defined with respect to the same FoD.

The proof of the three first properties is quite simple to make. The proof of the last property is much more difficult. As explained in [7], we do not consider that the sub-additivity property [13] of $U(m)$ is a fundamental desideratum that an effective MoU must satisfy in general. In fact the sub-additivity desideratum is

⁶ This terminology is not used by Shannon in his original paper but it has been introduced by Tribus in [10] in the probabilistic context, and by analogy we adopt Tribus' terminology also for BBAs.

incompatible with the fourth important property $U(m) < U(m_v)$ above which stipulates that none non-vacuous BBA can be more uncertain (i.e. more ignorant about the problem under consideration) than the vacuous BBA. Actually, it does not make sense to have the entropy $U(m_v^{\Theta \times \Theta'})$ of the vacuous joint BBA $m_v^{\Theta \times \Theta'}$ defined on the cartesian product space $\Theta \times \Theta'$ smaller than (or equal to) the sum $U(m_v^\Theta) + U(m_v^{\Theta'})$ of entropies of vacuous BBAs m_v^Θ and $m_v^{\Theta'}$ defined respectively on Θ and Θ' . There is no theoretical justification, nor intuitive reason for this sub-additivity desideratum in the context of non-bayesian BBAs. Of course for Bayesian BBAs, $U(m)$ is equivalent to Shannon entropy which is in this case sub-additive.

It can be also proved, see [7] for details, that the entropy of the vacuous BBA m_v related to a FoD Θ is equal to

$$U(m_v^\Theta) = 2^{|\Theta|} - 2 \quad (6)$$

This maximum entropy value $U(m_v)$ makes perfect sense because for this very particular BBA there is no information at all about the conflicts between the elements of the FoD. Actually for all $X \in 2^\Theta \setminus \{\emptyset, \Theta\}$ one has $u(X) = 1$ because $[Bel(X), Pl(X)] = [0, 1]$, and one has $u(\emptyset) = 0$ and $u(\Theta) = 0$. Hence, the sum of all imprecisions of $P(X)$ for all $X \in 2^\Theta$ is exactly equal to $2^{|\Theta|} - 2$ which corresponds to $U(m_v^\Theta)$ as expected. Moreover, one has always $U(m_v^\Theta) > \log(|\Theta|)$ which means that the vacuous BBA has always an entropy greater than the maximum of Shannon entropy $\log(|\Theta|)$ obtained with the uniform probability mass function distributed on Θ .

Example 1 of entropy calculation: consider $\Theta = \{\theta_1, \theta_2\}$ and the BBA $m^\Theta(\theta_1) = 0.5$, $m^\Theta(\theta_2) = 0.3$ and $m^\Theta(\theta_1 \cup \theta_2) = 0.2$, then one has $[Bel(\emptyset), Pl(\emptyset)] = [0, 1]$ and $u(\emptyset) = 0$, $[Bel(\theta_1), Pl(\theta_1)] = [0.5, 0.7]$, $[Bel(\theta_2), Pl(\theta_2)] = [0.3, 0.5]$, and $[Bel(\Theta), Pl(\Theta)] = [1, 1]$. Hence, $u(\theta_1) = 0.2$, $u(\theta_2) = 0.2$ and $u(\Theta) = 0$. Applying (5), one gets $s(\emptyset) = 0$, $s(\theta_1) \approx 0.377258$, $s(\theta_2) \approx 0.428953$ and $s(\Theta) \approx 0.321887$. Using the 2-bits representation with increasing ordering⁷, we encode the elements of the power set as $\emptyset = 00$, $\theta_1 = 01$, $\theta_2 = 10$ and $\theta_1 \cup \theta_2 = 11$. The entropiece vector for this simple example is

$$\mathbf{s}(m^\Theta) = \begin{bmatrix} s(\emptyset) \\ s(\theta_1) \\ s(\theta_2) \\ s(\theta_1 \cup \theta_2) \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0.377258 \\ 0.428953 \\ 0.321887 \end{bmatrix} \quad (7)$$

If we use the classical N-bits (here $N = 2$) representation with increasing ordering (as we recommend) the first component of entropiece vector $\mathbf{s}(m^\Theta)$ will be $s(\emptyset)$ which is always equal to zero for any BBA m , hence the first component of $\mathbf{s}(m^\Theta)$ is always zero. By summing all the components of the entropiece vector $\mathbf{s}(m^\Theta)$ we obtain the entropy $U(m^\Theta) \approx 1.128098$ nats of the BBA $m^\Theta(\cdot)$. Note that the components $s(X)$ (for $X \neq \emptyset$) of the entropieces vector $\mathbf{s}(m^\Theta)$ are not

⁷ Once the binary values are converted into their digit value with the most significant bit on the left (i.e the least significant bit on the right).

independent because they are linked to each other through the calculation of $Bel(X)$ and $Pl(X)$ values entering in $u(X)$.

Example 2 of entropy calculation: for the vacuous BBA m_v^Θ , and when using the binary increasing encoding of elements of 2^Θ , the first component $s(\emptyset)$ and the last component $s(\Theta)$ of entropiece vector $\mathbf{s}(m_v^\Theta)$ will always be equal to zero, and all other components of $\mathbf{s}(m_v^\Theta)$ will be equal to one. For instance, if we consider $\Theta = \{\theta_1, \theta_2\}$ and the vacuous BBA $m_v^\Theta(\theta_1) = 0$, $m_v^\Theta(\theta_2) = 0$ and $m_v^\Theta(\theta_1 \cup \theta_2) = 1$, the corresponding entropiece vector $\mathbf{s}(m_v^\Theta)$ is

$$\mathbf{s}(m_v^\Theta) = \begin{bmatrix} s(\emptyset) \\ s(\theta_1) \\ s(\theta_2) \\ s(\theta_1 \cup \theta_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad (8)$$

By summing all the components of the entropiece vector $\mathbf{s}(m_v^\Theta)$ we obtain the entropy value $U(m_v^\Theta) = 2$ nats for this vacuous BBA $m_v^\Theta(\cdot)$, which is of course in agreement with the formula (6).

4 Information content of a BBA

We consider a (non-empty) FoD of cardinality $|\Theta| = N$, and we model our state of knowledge about the problem under consideration by a BBA defined on 2^Θ . Without more knowledge than the FoD itself (and its cardinality N), we are totally ignorant about the solution of the problem we want to solve, and of course we have no clue for making a decision/choice among the elements of the FoD. The BBA reflecting this total ignorant situation is the vacuous BBA $m_v(\cdot)$, whose maximal entropy is $U(m_v) = 2^N - 2$. In such case, we naturally expect that the information content we have⁸ is zero when the uncertainty measure is maximal. In the very opposite case, it is very natural to consider that the information content of a BBA is maximal if the entropy value (the MoU value) of a BBA $m(\cdot)$ is zero, meaning that we make a choice of one element of the FoD without hesitation. Based on these very simple ideas, we propose to define the information content of any BBA $m(\cdot)$ as the dual of the effective measure of uncertainty, more precisely by

$$IC(m^\Theta) \triangleq U(m_v^\Theta) - U(m^\Theta) = (2^{|\Theta|} - 2) - \sum_{X \in 2^\Theta} s(X) \quad (9)$$

where $s(X)$ is the *entropiece* of the element $X \in 2^\Theta$ given by (5), that is

$$s(X) \triangleq -(1 - u(X))m^\Theta(X) \log(m^\Theta(X)) + u(X)(1 - m^\Theta(X))$$

and where $u(X)$ is the level of imprecision of the probability $P(X)$ given by

$$u(X) = Pl^\Theta(X) - Bel^\Theta(X) = \sum_{Y \in 2^\Theta | X \cap Y \neq \emptyset} m^\Theta(Y) - \sum_{Y \in 2^\Theta | Y \subseteq X} m^\Theta(Y) \quad (10)$$

⁸ aside of the value of N of course.

From the definition (9), one sees that for $m^\Theta \neq m_v^\Theta$ one has $IC(m^\Theta) > 0$ because $U(m^\Theta) < U(m_v^\Theta)$, and for $m^\Theta = m_v^\Theta$ one has $IC(m_v^\Theta) = 0$, which is what we naturally expect.

It is worth mentioning that the information content $IC(m^\Theta)$ of a BBA depends not only of the BBA $m(\cdot)$ itself but also on the cardinality of the frame of discernment⁹ Θ because $IC(m^\Theta)$ requires the knowledge of $|\Theta|$ to calculate the max entropy value $U(m_v^\Theta) = 2^{|\Theta|} - 2$ entering in (9). This remark is very important to understand that even if two BBAs (defined on different FoDs) focus entirely on a same focal element, their information contents are necessarily different. For instance, if we consider the Bayesian BBA with $m^\Theta(\theta_1) = 1$ defined on the FoD $\Theta = \{\theta_1, \theta_2\}$, then

$$IC(m^\Theta) = U(m_v^\Theta) - U(m^\Theta) = (2^{|\Theta|} - 2) - 0 = 2 \text{ (nats)}$$

whereas if we consider the Bayesian BBA with $m^{\Theta'}(\theta_1) = 1$ defined on the larger FoD $\Theta' = \{\theta_1, \theta_2, \theta_3\}$ (for instance), then

$$IC(m^{\Theta'}) = U(m_v^{\Theta'}) - U(m^{\Theta'}) = (2^{|\Theta'|} - 2) - 0 = 6 \text{ (nats)}$$

So even if the decision θ_1 that we would make based either on m^Θ or on $m^{\Theta'}$ is the same, these decisions must not be considered actually with the same strength, and this is what reflects our information content measure.

From this very simple definition of information content, we can also define the Normalized Information Content (NIC) (if needed later in some applications), denoted by $NIC(m^\Theta)$ by normalizing $IC(m^\Theta)$ with respect to the maximal value of entropy $U(m_v^\Theta)$ as

$$NIC(m^\Theta) \triangleq \frac{U(m_v^\Theta) - U(m^\Theta)}{U(m_v^\Theta)} = 1 - \frac{U(m^\Theta)}{U(m_v^\Theta)} \quad (11)$$

Hence we will have $NIC(m^\Theta) \in [0, 1]$ and $NIC(m^\Theta) = 0$ for $m = m_v$, and $NIC(m^\Theta) = 1$ for $U(m) = 0$ which is obtained when $m(\cdot)$ is entirely focused on a singleton $\theta_i \in \Theta$, that is $m^\Theta(\theta_i) = 1$ for some $i \in \{1, 2, \dots, |\Theta|\}$.

In fact, the (total) information content of a BBA $IC(m^\Theta)$ is the sum of all the *information granules* $IG(X|m^\Theta)$ of elements $X \in 2^\Theta$ carried by a BBA m^Θ , that is

$$IC(m^\Theta) = \sum_{X \in 2^\Theta} IG(X|m^\Theta) \quad (12)$$

where

$$IG(X|m^\Theta) \triangleq \begin{cases} 0, & \text{if } X = \emptyset \\ -s(X), & \text{if } X = \Theta \\ 1 - s(X) & \text{otherwise} \end{cases} \quad (13)$$

⁹ That is why it is better, we think, to use the notation $IC(m^\Theta)$ instead of $IC(m)$.

We can define the information granules vector¹⁰ $\mathbf{IG}(m) = [IG(X|m^\Theta), X \in 2^\Theta]^T$ by

$$\mathbf{IG}(m^\Theta) \triangleq \mathbf{s}(m_v^\Theta) - \mathbf{s}(m^\Theta) \quad (14)$$

One sees that the (total) information content $IC(m^\Theta)$ of a BBA m^Θ is just the sum of all components $IG(X|m^\Theta)$ of the information granules vector $\mathbf{IG}(m)$. The information granules vector $\mathbf{IG}(m)$ is interesting and useful because it helps to see the contribution of each element X in the whole measure of the information content $IC(m^\Theta)$ of a BBA m^Θ .

Example 1 (continued): consider $\Theta = \{\theta_1, \theta_2\}$ and the BBA $m^\Theta(\theta_1) = 0.5$, $m^\Theta(\theta_2) = 0.3$ and $m^\Theta(\theta_1 \cup \theta_2) = 0.2$. The information granules vector $\mathbf{IG}(m^\Theta)$ is given by

$$\mathbf{IG}(m^\Theta) = \mathbf{s}(m_v^\Theta) - \mathbf{s}(m^\Theta) = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0.377258 \\ 0.428953 \\ 0.321887 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.622742 \\ 0.571047 \\ -0.321887 \end{bmatrix} \quad (15)$$

By summing all the components of the information granules vector $\mathbf{IG}(m^\Theta)$ we obtain the (total) information content $IC(m^\Theta) = 0.871902$ nats of the BBA m^Θ , which can of course be calculated directly also as

$$IC(m^\Theta) = U(m_v^\Theta) - U(m^\Theta) = 2 - 1.128098 = 0.871902$$

However, the information granules vector $\mathbf{IG}(m^\Theta)$ is interesting to identify the contribution of each element X in the whole measure of the information content.

5 Information gain and information loss

Once the IC measure is defined for a BBA, it is rather simple to define the information gain and information loss of a BBA with respect to another one, both defined on a same FoD Θ . Suppose that we have a first BBA m_1^Θ and a second BBA m_2^Θ , then we can calculate by formula (9) their respective information contents $IC(m_1^\Theta)$ and $IC(m_2^\Theta)$. The difference of information content measure of m_2^Θ with respect to m_1^Θ is defined by¹¹

$$\Delta_{IC}(m_2|m_1) \triangleq IC(m_2^\Theta) - IC(m_1^\Theta) \quad (16)$$

If we replace $IC(m_2^\Theta)$ and $IC(m_1^\Theta)$ by their expressions according to (9), it comes

$$\Delta_{IC}(m_2|m_1) = [U(m_v^\Theta) - U(m_2^\Theta)] - [U(m_v^\Theta) - U(m_1^\Theta)] = U(m_1^\Theta) - U(m_2^\Theta) \quad (17)$$

¹⁰ We suppose for convenience that the elements $X \in 2^\Theta$ are listed in increasing order using the classical $|\Theta|$ -bits representation with the least significant bit on the right.

¹¹ Similarly, we can define $\Delta_{IC}(m_1|m_2) \triangleq IC(m_1^\Theta) - IC(m_2^\Theta) = -\Delta_{IC}(m_2|m_1)$.

If $\Delta_{IC}(m_2|m_1) = 0$, the BBAs m_1^Θ and m_2^Θ have same measure of information content. So, there is no gain and no loss in information content if one switches from m_1^Θ to m_2^Θ or vice versa. $\Delta_{IC}(m_2|m_1) = 0$ does not mean that the decisions based on m_1^Θ and on m_2^Θ are the same. It does only means that the decision based on m_1^Θ must be as easy as the decision made based on m_2^Θ . It means that they have the same informational strength. That's it. If $\Delta_{IC}(m_2|m_1) > 0$, one has $IC(m_2^\Theta) > IC(m_1^\Theta)$, i.e. the BBA m_2^Θ is more informative than m_1^Θ . In this case we get an information gain if one switches from m_1^Θ to m_2^Θ , and by duality we get an uncertainty reduction by switching from m_1^Θ to m_2^Θ . It means that it must be easier to make a decision based on m_2^Θ rather on m_1^Θ . If $\Delta_{IC}(m_2|m_1) < 0$, one has $IC(m_2^\Theta) < IC(m_1^\Theta)$, i.e. the BBA m_2^Θ is less informative than m_1^Θ . In this case we get an information loss if one switches from m_1^Θ to m_2^Θ , and by duality we get an uncertainty raise by switching from m_1^Θ to m_2^Θ . It means that it must be easier to make a decision based on m_1^Θ rather on m_2^Θ .

As simple example, consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$. For the vacuous BBA one has $U(m_v^\Theta) = 2^3 - 2 = 6$ nats. Suppose at time $k = 1$ one has the BBA $m_1^\Theta(\theta_1 \cup \theta_2) = 0.2$, $m_1^\Theta(\theta_1 \cup \theta_3) = 0.3$, $m_1^\Theta(\theta_1 \cup \theta_2 \cup \theta_3) = 0.5$, then $U(m_1^\Theta) \approx 5.1493$ nats, and $IC(m_1^\Theta) = U(m_v^\Theta) - U(m_1^\Theta) \approx 0.8507$ nats. Suppose that after some information processing (belief revision, or fusion, etc) we come up with the BBA m_2^Θ at time $k = 2$ defined by $m_2^\Theta(\theta_1) = 0.2$ and $m_2^\Theta(\theta_1 \cup \theta_3) = 0.8$, then $U(m_2^\Theta) \approx 0.5004$ nats and $IC(m_2^\Theta) = U(m_v^\Theta) - U(m_2^\Theta) \approx 5.4996$ nats. In this case, we get $\Delta_{IC}(m_2|m_1) = 5.4996 - 0.8507 = 4.6489$ which is positive. Hence we get an information gain by switching from m_1^Θ to m_2^Θ thanks to the information processing applied.

6 Conclusions

In this paper we have introduced a measure of information content (IC) for any basic belief assignment (BBA). This IC measure based on an effective measure of uncertainty of BBAs is quite simple to calculate, and it reflects somehow the informational strength and easiness ability to make a decision based on any belief mass function. We have also shown how it is possible to identify the contribution of each focal element of the BBA to this information content measure thanks to the information granule vector. This new IC measure is also interesting because it allows to well quantify the information loss or gain between two BBAs, and thus as perspectives we could use it to quantify precisely and compare the performances of information processing using belief functions (like fusion rules, belief conditioning, etc). We hope that this new theoretical IC measure will open interesting tracks for forthcoming research works on reasoning about uncertainty with belief functions.

References

1. Bovee, M., Srivastava, R.S.: A Conceptual Framework and Belief-Function Approach to Assessing Overall Information Quality, *Int. J. of Intell. Systems*, Vol. 18(1), pp. 51–74, January (2003).

2. Floridi, L., Illari, P. (Editors): The philosophy of information quality, Springer International Publishing Switzerland (2014).
3. Batini, C., Scannapieco, M.: Data and Information Quality: Dimensions, Principles and Techniques, Springer International Publishing Switzerland (2016).
4. Kenett, R.S., Shmueli, G.: Information quality, John Wiley & Sons (2017).
5. Bossé E., Rogova, G.L. (Editors): Information Quality in Information Fusion and Decision Making, Information Fusion and Data Science, Springer Nature, Switzerland (2019).
6. Bouhamed, S.A., Kalle, I.K., Yager, R.R., Bossé, E., Solaiman, B.: An intelligent quality-based approach to fusing multi-source possibilistic information, *Information Fusion*, Vol. 55, pp. 68–90 (2020).
7. Dezert, J.: An Effective Measure of Uncertainty of Basic Belief Assignments, *Fusion 2022 Int. Conf.*, Linköping, Sweden, pp. 1–10, July (2022).
8. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ, U.S.A. (1976)
9. Dezert, J., Tchamova A.: On Effectiveness of Measures of Uncertainty of Basic Belief Assignments, *Information & Security Journal: An International Journal (ISIJ)*, Vol. 52, Feb. (2022).
10. Tribus, M.: Rational Descriptions, Decisions and Designs, Pergamon Press, New York, pp. 26–28 (1969).
11. Bradley, R.E., Petrilli, S.J., Sandifer, C.E.: L'Hôpital's analyse des infiniments petits (An annotated translation with source material by Johann Bernoulli), Birkhäuser, 311 pages (2015).
12. Shannon, C.E.: A mathematical theory of communication, *The Bell System Technical Journal*, Vol. 27, pp. 379-423 & 623-656, July & October (1948).
13. Klir, G.J.: Principles of uncertainty: What are they? Why do we need them?, *Fuzzy Sets and Systems*, Vol. 74, pp. 15–31 (1995).