International Journal of Approximate Reasoning 52 (2011) 449–460

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Conflict management in Dempster–Shafer theory using the degree of falsity ☆

Johan Schubert *

Division of Information Systems, Swedish Defence Research Agency, SE-164 90 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 3 February 2010
Received in revised form 26 October 2010
Accepted 27 October 2010
Available online 4 November 2010

Keywords:
Dempster–Shafer theory
Belief function
Conflict
Conflict management
Discounting

ABSTRACT

In this article we develop a method for conflict management within Dempster–Shafer theory. The idea is that each piece of evidence is discounted in proportion to the degree that it contributes to the conflict. This way the contributors of conflict are managed on a case-by-case basis in relation to the problem they cause. Discounting is performed in a sequence of incremental steps, with conflict updated at each step, until the overall conflict is brought down exactly to a predefined acceptable level.

1. Introduction

In this article we develop a method for conflict management within Dempster–Shafer theory [2–8] where it is assumed that all belief functions are referring to the same problem or alternatively that they are false.

In general a high degree of conflict is seen as if there is a representation error in the frame of discernment, while a small conflict may be the result of measuring errors.

One type of representation error resulting in high conflict is when belief functions concerning different subproblems that should be handled independently are erroneously combined [9,10]. When this is the case the assumption that all belief functions combined must refer to the same problem (*not* different subproblems) is violated.

We may interpret the conflict as metalevel evidence stating that at least one piece of evidence in the combination should not be part of that combination. By temporarily removing (and replacing) each belief function from the combination, one at a time, we induce a drop in conflict. This is used to derive metalevel evidence regarding each individual belief function indicating that this particular belief function does not belong to the problem in question.

When assuming that there is only one problem at hand, such metalevel evidence must be interpreted as a proposition about the falsity of this belief function. A normalization of the drop in conflict will be shown to be the degree of falsity of that belief function.

However, instead of directly discounting each piece of evidence to its individual degree of falsity we take an incremental step in that direction for all belief functions. Based on these initial discounts we recalculate conflict and update all degrees of

☆ A short version of this study was presented at the Twelfth International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2008) in Málaga, Spain [1]. This work was supported by the FOI research project "Real-Time Simulation Supporting Effects-Based Planning", which is funded by the R&D programme of the Swedish Armed Forces.

* Tel.: +46 8 55503702; fax: +46 8 55503700.
E-mail address: schubert@foi.se
URL: http://www.foi.se/fusion

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2010.10.004

falsities. The process is performed sequentially until a predefined level of maximal acceptable conflict is reached. With this sequential approach we obtain a smooth discounting process (compared to if we would have fully discounted each belief function to its degree of falsity) and we are able to exactly match any level of acceptable conflict without risk of overshooting.

An alternative way to manage the conflict is to assume that there are different subproblems where the set of basic belief assignments (bbas) may be distributed to different clusters that should be handled separately [9–18].

Another approach also using meta-knowledge regarding the reliability of the source is *contextual discounting* [19]. It is also possible to develop alternative distance measures between bodies of evidence [21,22]. In [22] Jousselme and Maupin compare several different distance measures. It is important to observe that different measures may measure different types of distances. Some distance measures measure the degree to which two bodies of evidence are different, while others such as conflict, measure the degree to which they are incompatible. For example, two propositions (corresponding to two focal elements) "a red car" and "a fast car" are different, but may be fully compatible if there is a red fast car in the frame of discernment.

A recent paper [23] also uses the idea of sequential discount to manage the conflict when combining belief functions. However, they use a distance measure by Jousselme et al. [20] that measures dissimilarity.

A recent overview of different alternatives for conflict management when combining conflicting belief functions was given by Smets, see [24].

In Section 2 we investigate the degree of falsity of a piece of evidence. In Section 3 we develop a method of sequential incremental discounting using the degree of falsity. We perform an experiment to investigate the behavior of an algorithm for conflict management in Section 4. Finally, conclusions are drawn in Section 5.

2. Degree of falsity

Let us recapitulate the interpretation of conflict as if there is at least one piece of evidence that violates the representation given by the frame of discernment, and thus can be said *not* to belong to the set of bbas that refer to this problem χ [10].

A conflict in χ is thus interpreted as a piece of metalevel evidence that there is at least one piece of evidence that does not belong to the subset

$$\begin{aligned} &m_\chi(\exists j.\ e_j \notin \chi) = c^{(0,0)}, \\ &m_\chi(\Theta) = 1 - c^{(0,0)}, \end{aligned} \tag{1}$$

where $c^{(0,0)}$ is the initial conflict in χ.

Let us observe one piece of evidence e_q in χ. If e_q is taken out from χ the conflict $c^{(0,0)}$ in χ decreases to $c^{(0,q)}$. This decrease in conflict can be interpreted as follows: there exists some metalevel evidence indicating that e_q does not belong to χ

$$\begin{aligned} &m_{\Delta\chi}(e_q \notin \chi), \\ &m_{\Delta\chi}(\Theta), \end{aligned} \tag{2}$$

and the remainder of the conflict $c^{(0,q)}$ after e_q has been taken out from χ is metalevel evidence that there is at least one other piece of evidence e_j, $j \neq q$, that does not belong to $\chi - \{e_q\}$

$$\begin{aligned} &m_{\chi-\{e_q\}}\big(\exists j \neq q.\ e_j \notin (\chi - \{e_q\})\big) = c^{(0,q)}, \\ &m_\chi(\Theta) = 1 - c^{(0,q)}, \end{aligned} \tag{3}$$

We will derive the basic belief number (bbn) $m_{\Delta\chi}(e_q \notin \chi)$ by stating that the belief in the proposition that there is at least one piece of evidence that does not belong to χ, $\exists j.\ e_j \notin \chi$, should be equal no matter whether we base that belief on the original piece of metalevel evidence, before e_q is taken out from χ, or on a combination of the other two pieces of metalevel evidence $m_{\Delta\chi}(e_q \notin \chi)$ and $m_{\chi-\{e_q\}}\big(\exists j \neq q.\ e_j \notin (\chi - \{e_q\})\big)$, after e_q is taken out from χ, i.e.:

$$\mathrm{Bel}_\chi(\exists j.\ e_j \notin \chi) = Bel_{\Delta\chi\oplus(\chi-\{e_q\})}(\exists j.\ e_j \notin \chi). \tag{4}$$

We have, on the left hand side (LHS)

$$\mathrm{Bel}_\chi(\exists j.\ e_j \notin \chi) = m_\chi(\exists j.\ e_j \notin \chi) = c^{(0,0)}, \tag{5}$$

and, on the right hand side (RHS)

$$\begin{aligned} \mathrm{Bel}_{\Delta\chi\oplus(\chi-\{e_q\})}(\exists j.\ e_j \notin \chi) &= m_{\Delta\chi\oplus(\chi-\{e_q\})}\big((e_q \notin \chi) \wedge \big(\exists_j \neq q.\ e_j \notin (\chi - \{e_q\})\big)\big) \\ &\quad + m_{\Delta\chi\oplus(\chi-\{e_q\})}\big(\exists j \neq q.\ e_j \notin (\chi - \{e_q\})\big) + m_{\Delta\chi\oplus(\chi-\{e_q\})}(e_q \notin \chi) \\ &= m_{\Delta\chi}(e_q \notin \chi) m_{(\chi-\{e_q\})}\big(\exists j \neq q.\ e_j \notin (\chi - \{e_q\})\big) + m_{\Delta\chi}(\Theta) m_{(\chi-\{e_q\})}\big(\exists j \neq q.\ e_j \notin (\chi - \{e_q\})\big) \\ &\quad + m_{\Delta\chi}(e_q \notin \chi) m_{(\chi-\{e_q\})}(\Theta) \\ &= m_{\Delta\chi}(e_q \notin \chi) c^{(0,q)} + \big[1 - m_{\Delta\chi}(e_q \notin \chi)\big] c^{(0,q)} + m_{\Delta\chi}(e_q \notin \chi)\big(1 - c^{(0,q)}\big) \\ &= c^{(0,q)} + m_{\Delta\chi}(e_q \notin \chi)\big(1 - c^{(0,q)}\big). \end{aligned} \tag{6}$$

Setting LHS = RHS, we get

$$
\begin{aligned}
m_{\Delta\chi}(e_q \notin \chi) &= \frac{c^{(0,0)} - c^{(0,q)}}{1 - c^{(0,q)}}, \\
m_{\Delta\chi}(\Theta) &= \frac{1 - c^{(0,0)}}{1 - c^{(0,q)}}.
\end{aligned}
\tag{7}
$$

This is the degree of falsity of e_q under the assumption that we are dealing with one problem, *not* several different subproblems.

3. Sequential incremental discounting

In this section we investigate how to manage the conflict on an individual case-by-case basis using the degree of falsity.

If $m_{\Delta\chi}(e_q \notin \chi) = 1$ then e_q is certainly false and must not be used in the combination. This becomes the situation when $c^{(0,0)} = 1$, for any $c^{(0,q)} < 1$. For $c^{(0,q)} = 1$ we define $m_{\Delta\chi}(e_q \notin \chi) = 0$ as the proposition is not supported when conflict remains unchanged, equal to 1. When $m_{\Delta\chi}(e_q \notin \chi) = 0$ then we have no indication regarding the falsity of e_q and will take no additional action. This is the situation when we observe no change in conflict $c^{(0,0)} = c^{(0,q)}$. When $0 < m_{\Delta\chi}(e_q \notin \chi) < 1$, then e_q contributes to the overall conflict and its conflict contribution must be managed. We would then like to pay less regard to a piece of evidence the higher the degree is that it is false, pay no attention to it when it is certainly false, and leave it unchanged when there is no indication as to its falsity. This can be done by using the discounting operation.

The discounting operation was introduced to handle the case when the source of some piece of evidence is lacking in credibility [4]. The credibility of the source, α, also became the credibility of the piece of evidence. The situation was handled by discounting each supported proposition other than Θ with the credibility α and by adding the discounted mass to Θ;

$$
m^{\%}(A_j) = \begin{cases} \alpha m(A_j), & A_j \neq \Theta, \\ 1 - \alpha + \alpha m(\Theta), & A_j = \Theta. \end{cases}
\tag{8}
$$

We will use the same discounting operation in this case when there is a direct indication for each separate piece of evidence regardless of which source produced it.

As the degree of falsity of e_q is proportional to the conflict that e_q contributes to the overall conflict we discount it using its credibility. The conflict in Dempster's rule when combining all pieces of evidence regarding e_q, as identical to one minus the credibility of the evidence;

$$
\alpha_q = 1 - m_{\Delta\chi}(e_q \notin \chi).
\tag{9}
$$

At step d, $c^{(d,0)}$ represents the conflict in χ after d sequential discounts of all bbas, and $c^{(d,q)}$ is the remaining conflict we would have in χ after d sequential discounts of all bbas if e_q is taken out from χ at this stage before combining.

Using the credibility (degree of falsity) we may derive a set of incrementally discounted bbas $\left\{m_q^d\right\}_q$ as

$$
\begin{aligned}
m_q^{d+1}(A) &= \left[1 - \varepsilon\left(\frac{c^{(d,0)} - c^{(d,q)}}{1 - c^{(d,q)}}\right)\right] m_q^d(A), \quad \forall A \subset \Theta,\ A \neq \emptyset, \\
m_q^{d+1}(\Theta) &= 1 - \sum_{A \subset \Theta} m_q^{d+1}(A),
\end{aligned}
\tag{10}
$$

where $\varepsilon \ll 1$ and $\left\{m_q^0\right\}_q$ is the initial set of bbas.

Alternatively, we can also rewrite Eq. (10) as

$$
\begin{aligned}
m_q^{d+1}(A) &= \prod_{i=0}^{d}\left[1 - \varepsilon\left(\frac{c^{(d,0)} - c^{(d,q)}}{1 - c^{(d,q)}}\right)\right] m_q^0(A), \quad \forall A \subset \Theta,\ A \neq \emptyset, \\
m_q^{d+1}(\Theta) &= 1 - \sum_{A \subset \Theta} m_q^{d+1}(A),
\end{aligned}
\tag{11}
$$

where

$$
\begin{aligned}
c^{(d,0)} &= m_\chi^d(\emptyset) = \oplus\left\{m_j^d\right\}_j(\emptyset), \\
c^{(d,q)} &= m_{\chi-\{e_q\}}^d(\emptyset) = \oplus\left(\left\{m_j^d\right\}_j - \left\{m_q^d\right\}\right)(\emptyset).
\end{aligned}
\tag{12}
$$

The combinations of all bbas in Eq. (12) using Dempster's rule is carried out by first converting all bbas to commonality functions [4]

$$
Q_j^d(A) = \sum_{B \supseteq A} m_j^d(B), \quad \forall j,\ A \subseteq \Theta.
\tag{13}
$$

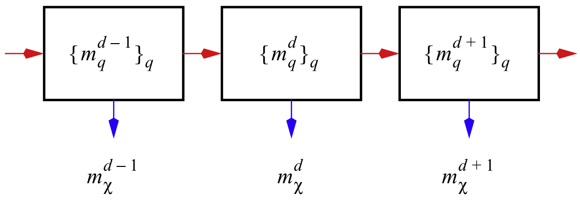

Fig. 1. The process of sequential discounting and combination. Red arrows are sequential discounting. Blue arrows are combination.

Secondly, we multiply all commonality functions

$$\begin{aligned} Q_\chi^d(A) &= \prod_{j=1}^{n} Q_j^d(A), \quad \forall A \subseteq \Theta, \\ Q_{\chi-\{e_q\}}^d(A) &= \prod_{\substack{j=1 \\ \neq q}}^{n} Q_j^d(A), \quad \forall A \subseteq \Theta, \end{aligned} \tag{14}$$

to obtain the unnormalized Dempster's rule.

Finally, we convert back to bbas in order to register the received conflict. We get

$$\begin{aligned} m_\chi^d(A) &= \sum_{B \supseteq A} (-1)^{|B-A|} Q_\chi^d(B), \\ m_{\chi-\{e_q\}}^d(A) &= \sum_{B \supseteq A} (-1)^{|B-A|} Q_{\chi-\{e_q\}}^d(B). \end{aligned} \tag{15}$$

When $A \equiv \emptyset$, this can be simplified to

$$\begin{aligned} m_\chi^d(\emptyset) &= \sum_{B} (-1)^{|B|} Q_\chi^d(B), \\ m_{\chi-\{e_q\}}^d(\emptyset) &= \sum_{B} (-1)^{|B|} Q_{\chi-\{e_q\}}^d(B), \end{aligned} \tag{16}$$

i.e., when we are only interested in the conflict. Here, $c^{(d,0)} = m_\chi^d(\emptyset)$ and $c^{(d,q)} = m_{\chi-\{e_q\}}^d(\emptyset)$. Using Eq. (10) we now obtain the sought after discounted bbas at the next step $d + 1$.

In each situation the bbas are sequentially discounted by repeated use of Eqs. (10) and (12), followed by combination using Eqs. (13)–(15), see Fig. 1.

In Algorithm 1 we describe an algorithm for performing sequential incremental discounting of all bbas.

The maximum conflict allowed is considered to be a domain dependent parameter.

4. An experiment

In this section we conduct an experiment with ten bbas over a frame of discernment with three elements and seven possible focal elements. We study the combination of the bbas and the use of conflict management through their sequential discounting using the degree of falsity and a gain factor of $\varepsilon = 0.1$. In an experiment with higher gain factors (not shown), e.g., $\varepsilon = 0.3$, the curves of m_χ^d evidently become step-wise linear.

Each bba has a random number of focal elements $n_q \in [1, |2^\Theta| - 1]$, where the number n_q is drawn with a uniform probability within the interval. The n_q focal elements are then drawn with an uniform probability $p = 1/(|2^\Theta| - 1)$ from the set $2^\Theta - \{\Theta\}$. With probability 1 we include Θ in the bba. Each focal element is given a random bbn drawn uniformly from $[0, \beta]$, $\beta \leqslant 1$, where β is chosen such that the bbns sum up to 1. As these bbas are constructed randomly, they are not constructed with any particular problem in sight; they are bound to be highly conflicting and a challenging test case.

Let us observe the process of sequential incremental discounting. At each step d in the sequential discounting we calculate the degree of falsity for all bbas. However, instead of discounting each bba to its full degree of falsity

$$\alpha^{(0,q)} = 1 - m_{\Delta\chi}(e_q \notin \chi), \tag{17}$$

as was done in [10], we take an incremental step in that direction by assigning

$$\alpha_\varepsilon^{(d,q)} = 1 - \varepsilon m_{\Delta\chi}(e_q \notin \chi), \tag{18}$$

where ε is a gain factor, $\varepsilon \ll 1$. In our experiments we use Algorithm 1 with $\varepsilon = 0.1$, Fig. 2. We have

$$\alpha_\varepsilon^{(d,q)} = 1 - \varepsilon \left(\frac{c^{(d,0)} - c^{(d,q)}}{1 - c^{(d,q)}} \right), \tag{19}$$

at step d, where $c^{(d,0)}$ and $c^{(d,q)}$ are calculated using Eq. (12).

INITIALIZE

$m_q^0(A), \;\; \forall q, A \subseteq \Theta$ (the bbas); k ($0 < k < 1$, the maximum conflict allowed); $\varepsilon = 0.1$

(a gain factor); $d = -1$;

REPEAT

$d = d + 1$;

Calculate $c^{(d,0)}$ using Eq. (12);

Calculate $c^{(d,q)}, \;\; \forall q$ using Eq. (12);

Calculate $m_q^{d+1}(A), \;\; \forall q, A \subseteq \Theta$ using Eq. (10);

UNTIL

$c^{(d,0)} \leq k$;

RETURN

$\{m_q^d\}_q$;

Fig. 2. Algorithm 1: Algorithm for sequential incremental discounting.

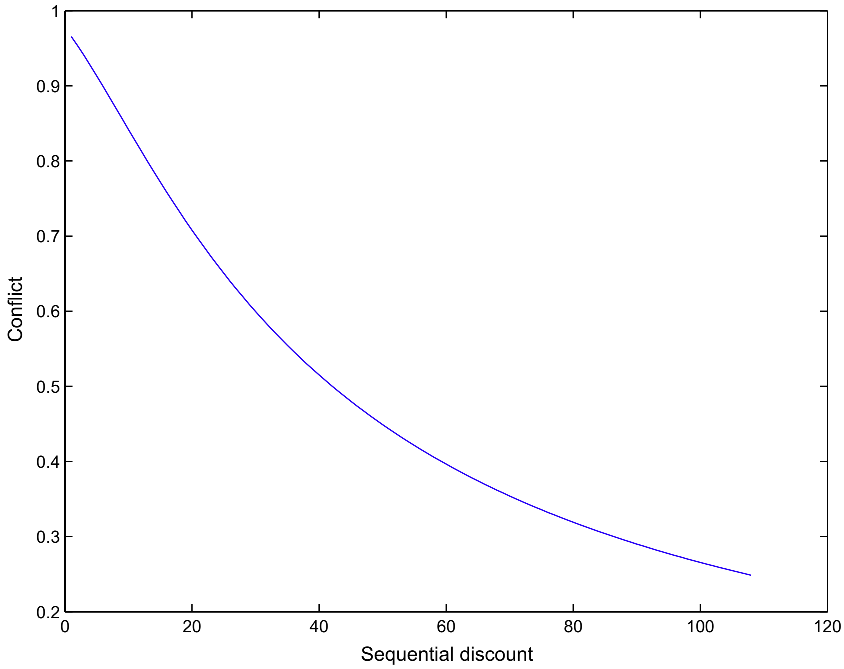

Fig. 3. Conflict decreasing with sequential incremental discounting.

In Fig. 3 we observe the conflict when we combine the ten bbas with Dempster's rule after different numbers of successively performed incremental discounts.

We notice an initial steady decline in conflict which is later somewhat moderated. As the conflict may be interpreted as a piece of metalevel evidence that there is something wrong with the representation of the problem we should at least request a conflict less or equal to 0.5. This level is reached after 42 incremental discounts.

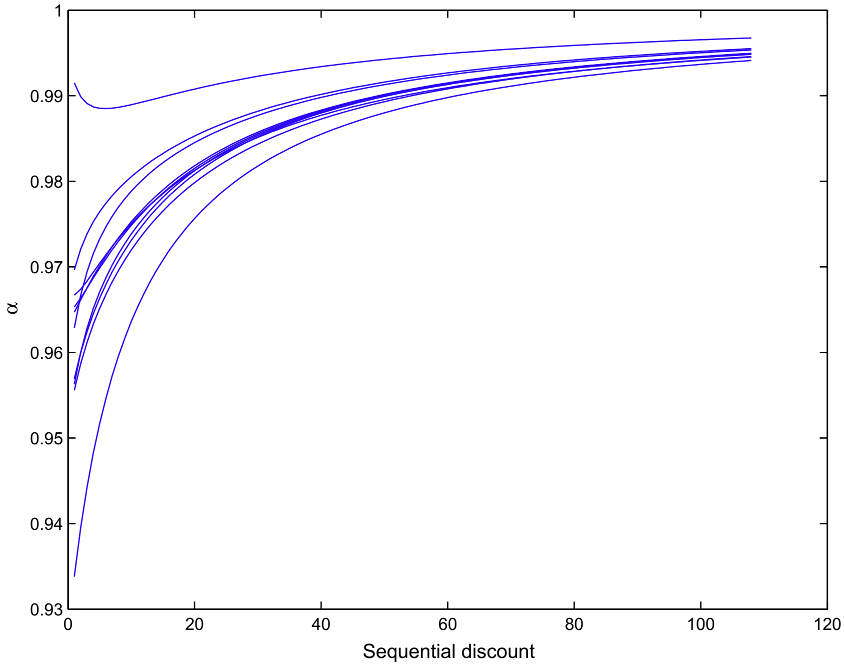

Fig. 4. Sequential $\alpha_\varepsilon^{(d,q)}$ for ten bbas, $\{e_q\}_{q=1}^{10}$ for successive steps d with $\varepsilon = 0.1$.

In Fig. 4 we observe the sequential incremental discounting factor $\alpha_\varepsilon^{(d,q)}$ for different bbas e_q. The initial discounting varies strongly between 1% and 7% where the discounting $\alpha_\varepsilon^{(d,q)}$ is proportional to the degree of credibility $\alpha^{(d,q)} = 1 - m(e_q \notin \chi)$.

As examples of how the bbas are changed by the sequential discounting, let us observe this for four different bbas in Fig. 5. In each case the blue line corresponds to $m_q^d(\Theta)$ and the other lines correspond to other focal elements $A \subset \Theta$. From these examples we notice especially the increase of nonspecificity in the bbas as support for $m_q^d(\Theta)$ increases with discounting.

The successive combinations of the ten bbas are shown in Figs. 6 and 7, without and with normalization, respectively. At each step d each bba is first discounted. After discounting, all bbas are combined. This process is illustrated in Fig. 1. In Fig. 6 we observe at each step the combined result, where the bbns are shown in blue (except for $m_\chi^d(\emptyset)$, red, and $m_q^d(\Theta)$, green). The bbns for identical focal elements at different steps are shown as curves.

In Fig. 7 we notice how the preferred hypothesis changes with sequential discounting as bbas which are highly conflicting have a high degree of falsity and are more strongly discounted than others. Here, the two hypotheses that initially were 1st and 3rd, become 2nd and 1st at the 42nd sequential discount (the 50% conflict level). We notice that this change in preference order takes place at the 20th sequential discount around the 70% conflict level. Thus, in practice k can be fairly high.

In Fig. 8 we observe for comparison direct discounting of each piece of evidence to its individual degree of falsity. In comparison with the sequential approach in Fig. 7 we notice that this corresponds roughly to the 20th sequential discount with a rather high 70% remaining conflict. In this example, this is also the point where the preferences order change. Thus, using sequential discounting that brings down the overall conflict somewhat further, e.g., towards the 50% level, obtains a stable preference order among the different alternatives.

Thus, sequential discounting is superior to direct discounting in that it can bring down the conflict to any predefined level, or be observed during the discounting process in order to find a stable preference order among alternatives.

In Table 1 we compare discounting with basic averaging of all mass functions. In column 2 we find $\oplus\left\{m_q^0(A)\right\}_q$ where all mass functions are combined without any discounting, in column 3 we perform a 1-step discount to each mass functions degree of falsity before combining $\oplus\left\{m_q^{\text{direct}}(A)\right\}_q$. In column 4–8 we tabulate the combination of all mass functions $\oplus\left\{m_q^{\alpha}(A)\right\}_q$ that are sequentially discounted 10–50 times before combination with $\varepsilon = 0.1$. In the last column we compare this to basic averaging of all mass functions, where

$$m_{average}^0 = \sum_q m_q^0(A). \tag{20}$$

Let us (in Table 2) observe the preference order of the top three preferred focal elements tabulated in Table 1.

We notice a small change in preference order between discounting and no discounting. However, the fundamental difference is noticed between that of all combination vs. averaging where the preference order is completely different. With combination of belief functions (with or without discounting) we prefer a singleton, with averaging Θ is the preferred focal element.

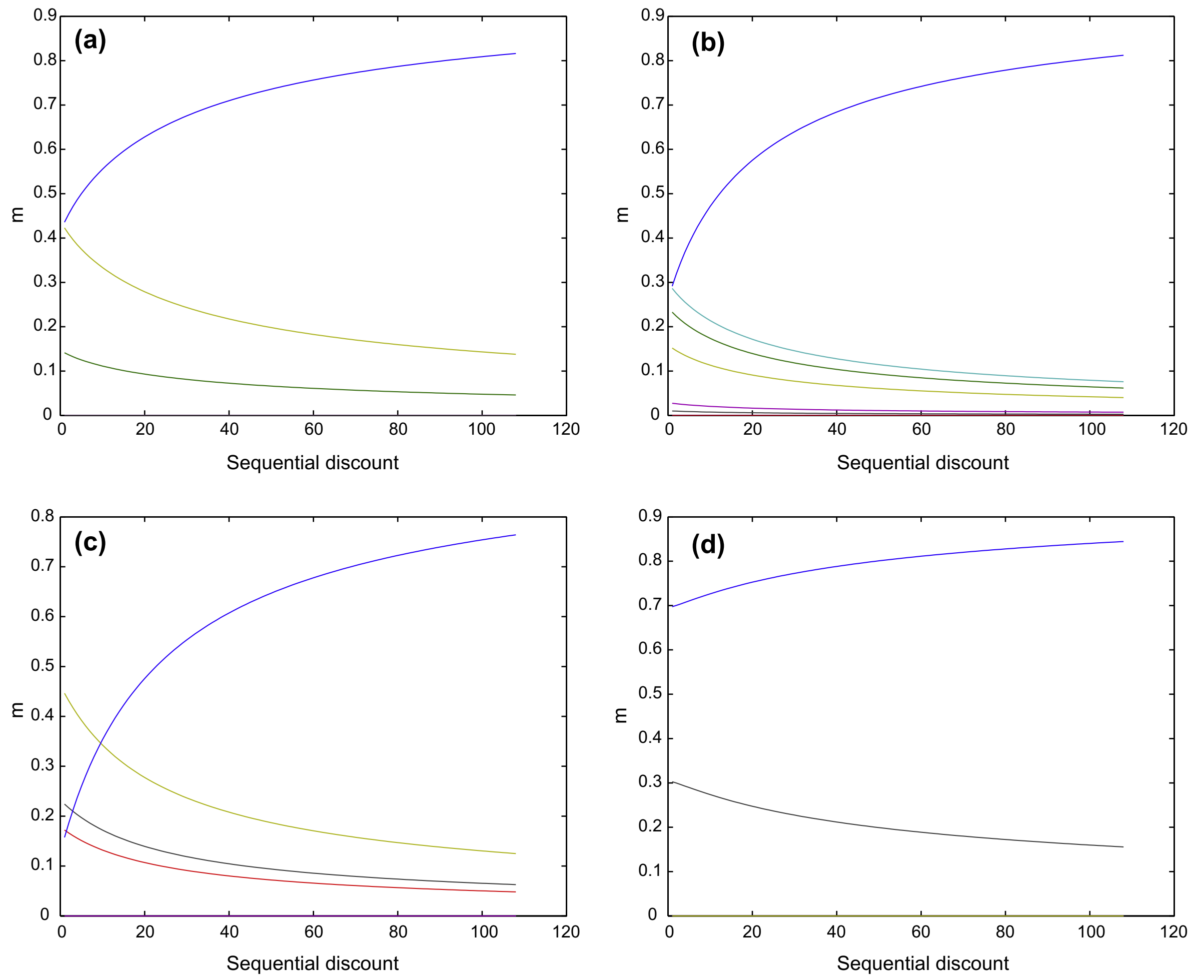

Fig. 5. Four different bbas sequentially discounted; $m_q^d(\Theta)$ blue line.

In an experiment with 2500 sequential discounts we notice that when $d \to \infty$ then $m_q^d(\Theta) \to 1$ slowly in a logarithm-like way, Fig. 9.

As information is lost by discounting it may be viewed as a necessary evil in order to manage the conflict (when this is high). Obviously, if a poor representation of the problem at hand (through the frame of discernment) is the cause of the conflict rather than poorly represented input data, we should change the frame of discernment. We measure the information lost by studying entropy measures as the sequential discounting progresses.

We prefer to see basic belief masses that are focused on as few and as small focal elements as possible. This can be measured by generalizing Shannon's entropy [25] and Hartley's information [26] measures, respectively. We will use a measure of aggregated uncertainty (*AU*) that takes both types of uncertainty into account.

The aggregated uncertainty functional *AU* is defined as

$$AU(\mathrm{Bel}) = \max_{\{p_x\}_{x\in\Theta}} \left\{ -\sum_{x\in\Theta} p(x)\log_2 p(x) \right\}, \tag{21}$$

where $\{p_x\}_{x\in\Theta}$ is the set of all probability distributions such that $p_x \in [0,1]$ for all $x \in \Theta$

$$\sum_{x\in\Theta} p(x) = 1, \tag{22}$$

and

$$\mathrm{Bel}(A) \leqslant \sum_{x\in A} p(x), \tag{23}$$

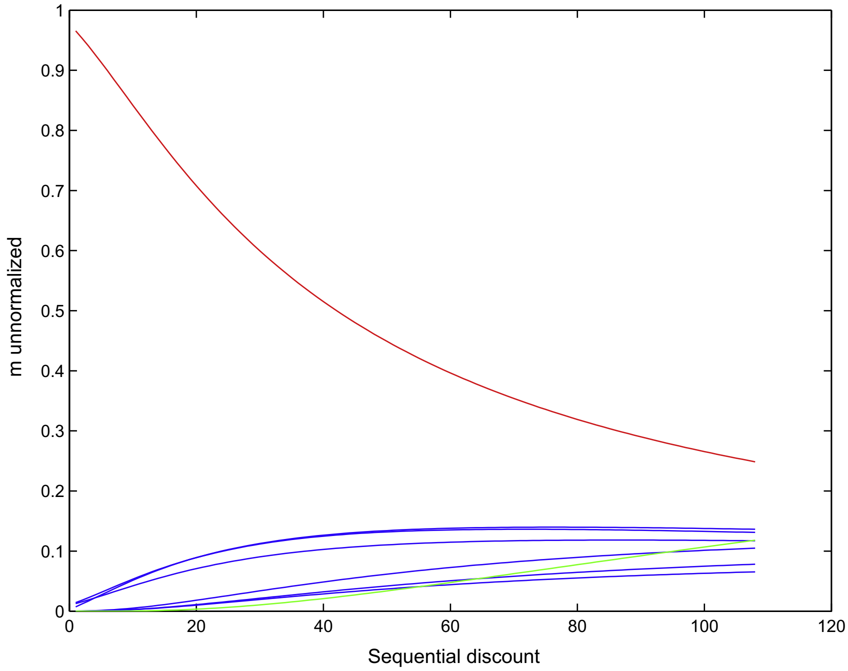

Fig. 6. Combination of sequentially discounted bbas without normalization. Red line is $m_q^d(\emptyset)$. Green line is $m_q^d(\Theta)$.

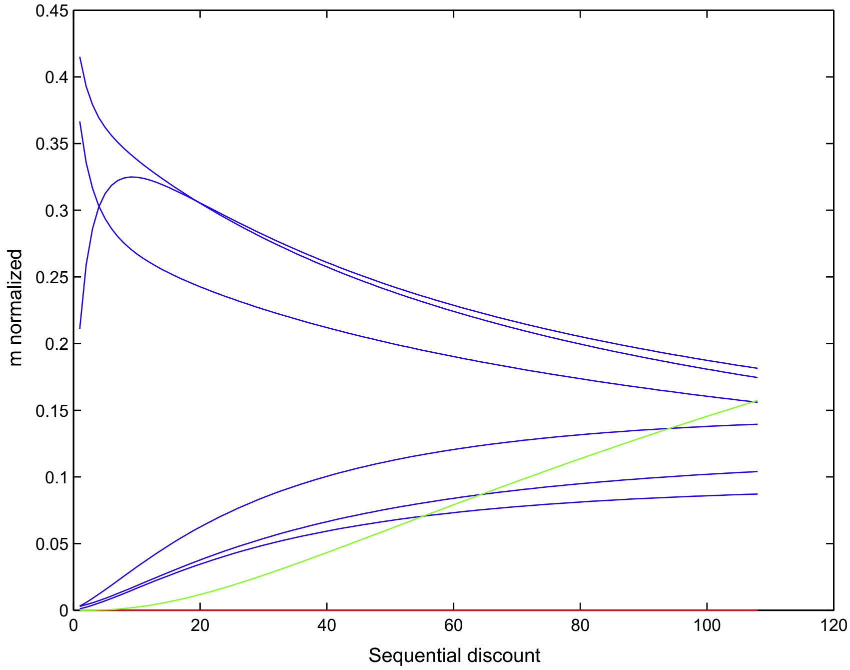

Fig. 7. Combination of sequentially discounted bbas with normalization. Red line is $m_q^d(\emptyset)$. Green line is $m_q^d(\Theta)$.

for all $A \subseteq \Theta$. *AU* was independently discovered by several authors about the same time [27–29].

Abellán et al. [30] suggested that *AU* could be disaggregated in separate measures of nonspecificity and scattering that generalize Hartley information [26] and Shannon entropy [25], respectively. Dubois and Prade [31] defined such a measure of nonspecificity as

$$I(m_x^d) = \sum_{A \in F} m_x^d \log_2 |A|, \tag{24}$$

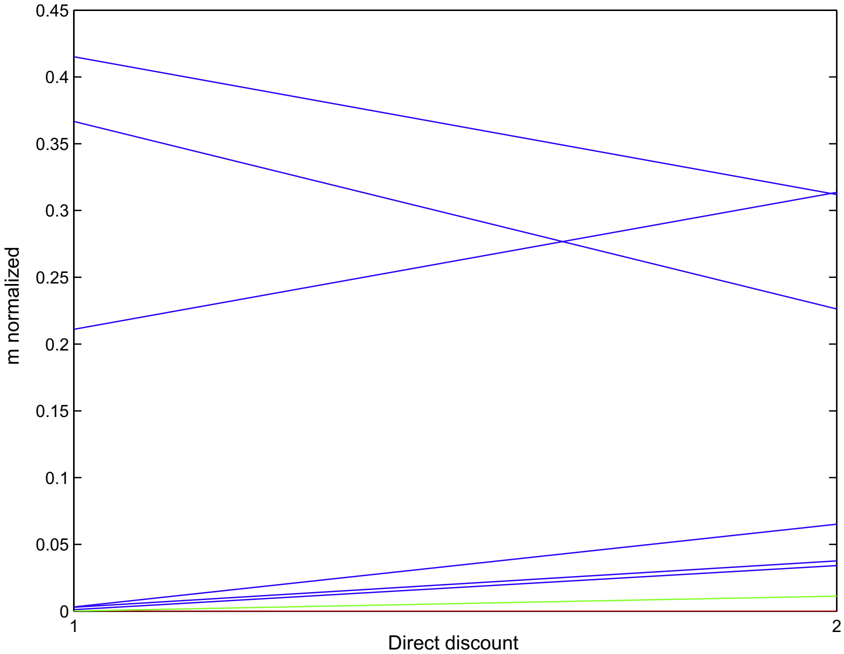

Fig. 8. Single step direct discounting to each piece of evidence individual degree of falsity.

Table 1
A comparison of initial combination direct and sequential discounting with averaging.

	No discount	Direct discount	Sequential discount (α)					Averaging
			10	20	30	40	50	
{e1}	0.211	0.314	0.324	0.303	0.279	0.259	0.242	0.079
{e2}	0.367	0.226	0.264	0.241	0.224	0.211	0.199	0.096
{e1,e2}	0.001	0.034	0.018	0.036	0.050	0.060	0.068	0.136
{e3}	0.415	0.312	0.334	0.302	0.277	0.256	0.238	0.070
{e1,e3}	0.003	0.065	0.036	0.065	0.086	0.102	0.113	0.186
{e2,e3}	0.003	0.038	0.020	0.040	0.055	0.068	0.077	0.144
Θ	0.000	0.011	0.003	0.013	0.028	0.045	0.063	0.289

Table 2
Preference order of discounting vs. averaging.

Preference order	No discount	Direct discount	Sequential discount at α = 50	Averaging
1st	{e3}	{e1}	{e1}	Θ
2nd	{e2}	{e3}	{e3}	{e1,e3}
3rd	{e1}	{e2}	{e2}	{e2,e3}

where $F \subseteq 2^{\Theta}$ is the set of focal elements. From Eqs. (21) and (24) we may define a generalized Shannon entropy [30] as

$$GS(m_x^d) = AU(m_x^d) - I(m_x^d). \tag{25}$$

An algorithm for computing AU was found by Meyerowitz et al. [32]. For the sake of completeness we cite the algorithm here, in the way it is described by Harmanec et al. [33], Fig. 10.

This measure reduces to Shannon's entropy [25] when m_χ^d represents a probability distribution (i.e., $\forall A. |A| = 1$) and to Hartley's information [26] when m_χ^d is certain (i.e., $\exists A. m_\chi^d(A) = 1$). Obviously, the aggregated uncertainty reaches its minimum $AU\left(m_\chi^d\right) = 0$ when both conditions apply, i.e., $\exists A. m_\chi^d(A) = 1 \; \& \; |A| = 1$.

In Fig. 11 we observe the entropy of m_χ^d at different stages d of the sequential discounting. We observe a rapid increase in aggregated uncertainty in the unnormalized case as mass is transferred towards Θ as discounting progresses, red line. For the normalized case we observe entropy quickly reaching close to its theoretical maximum of $\log_2(|\Theta|)$ which is 1.585 for $|\Theta| = 3$.

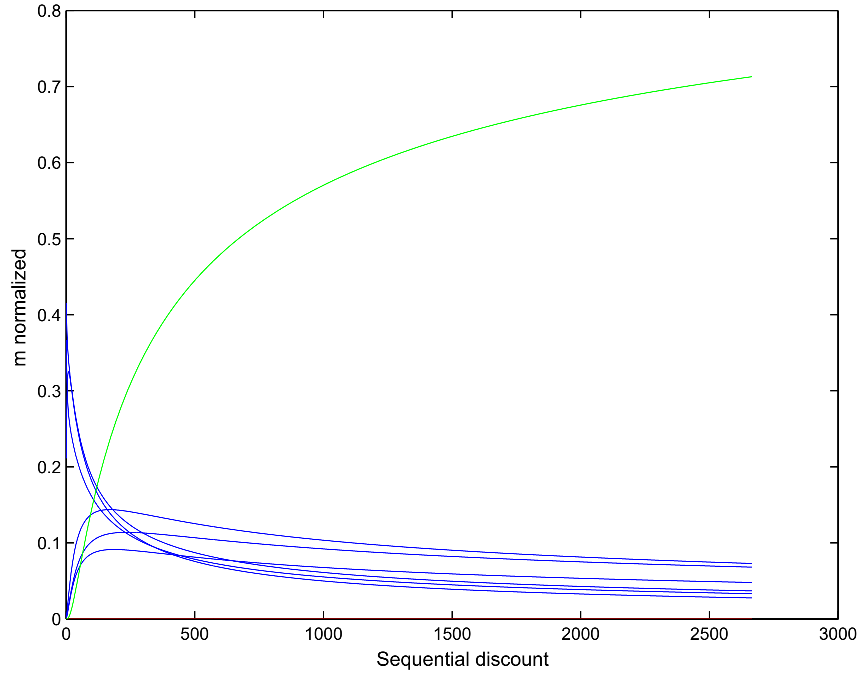

Fig. 9. $m_q^d(\Theta) \to 1$ as discounting progresses, green line.

Input: a frame of discernment X, a belief function Bel on X.

Output: $AU(Bel)$, $\{p_x\}_{x \in X}$ such that $AU(Bel) = -\sum_{x \in X} p_x \log_2 p_x$, $p_i \geq 0$, $\sum_{x \in X} p_x = 1$, and $Bel(A) \leq \sum_{x \in X} p_x$ for all $\emptyset \neq A \subseteq X$.

Step 1. Find a non-empty set $A \subseteq X$, such that $Bel(A) / |A|$ is maximal. If there are more than one such set A, take the one with maximal cardinality.

Step 2. For $x \in A$, put $p_x = Bel(A) / |A|$.

Step 3. For each $B \subseteq X - A$, put $Bel(B) = Bel(B \cup A) - Bel(A)$.

Step 4. Put $X = X - A$.

Step 5. If $X \neq \emptyset$ and $Bel(X) > 0$, then go to Step 1.

Step 6. If $Bel(X) = 0$ and $X \neq \emptyset$, then put $p_x = 0$ for all $x \in X$.

Step 7. Calculate $AU(Bel) = -\sum_{x \in X} p_x \log_2 p_x$.

Fig. 10. Algorithm 2: An algorithm for computing $AU(Bel)$.

We observe an increase in Hartley information $I\left(m_\chi^d\right)$ (blue line) as mass is transferred towards the frame as a whole $m_\chi^d(\Theta)$ as the sequential discounting progresses. As the aggregate uncertainty $AU\left(m_\chi^d\right)$ (black line) is close to constant throughout much of the discounting process we observe an early peak in generalized Shannon's entropy $GS\left(m_\chi^d\right)$, Eq. (25), followed by a counterintuitive decrease in generalized Shannon's entropy as Hartley's information increases with discounting. In other test cases with other random belief functions where the aggregate uncertainty has a more gradual increase

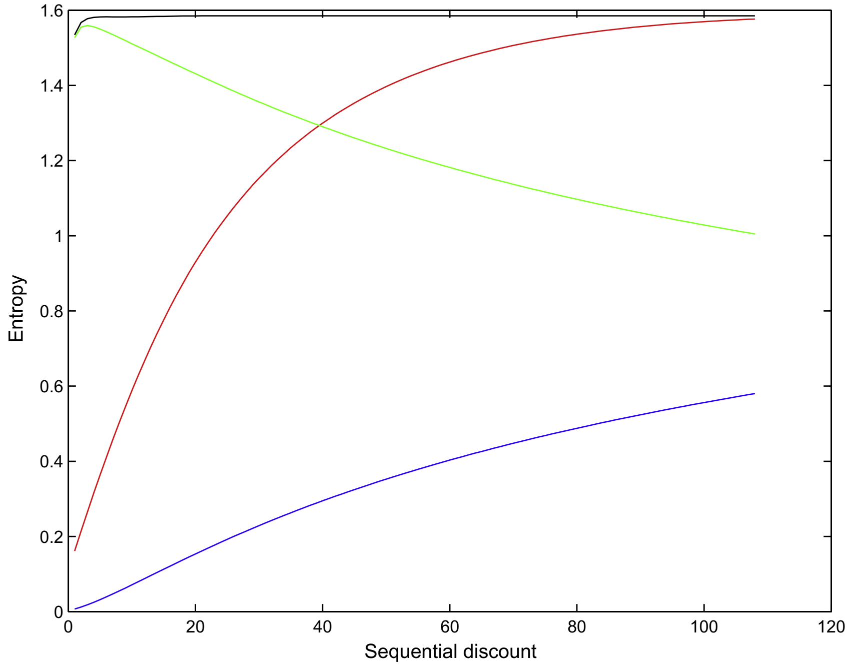

Fig. 11. Total entropy increases with discounting. Aggregated uncertainty, unnormalized (red); aggregated uncertainty, normalized (black); Hartley's information, normalized (blue); Shannon's entropy, normalized (green).

throughout the discounting process we observe the same effect but with a later peak for $GS\left(m_{\chi}^{d}\right)$ followed by the same decrease as Hartley's information increases more than the aggregated uncertainty $\left(AU\left(m_{\chi}^{d}\right)\right)$.

As is apparent from Fig. 7 the loss of information by discounting does not make the analysis difficult. Rather it makes the conclusions that may be drawn from the combination of discounted belief functions more reliable, as the conflict is reduced. For instance, after 20 sequential discounts (see Fig. 7) the preference order of supported focal elements becomes stable (in the region of reasonable discounting).

5. Conclusions

We have demonstrated that we can successfully manage the conflict of Dempster's rule by making well motivated and precise discounting of all belief functions. Such discounting is made individually for each belief function in proportion to its degree of falsity. We show that by performing the discounting process in a series of incremental steps we can reach any predefined acceptable level of conflict. In an experiment we find that this discounting does not normally make it more difficult to identify the most supported proposition. Rather it makes the selection process of the preferred proposition more robust when highly conflicting pieces of evidence are discounted down to a level they deserve.

References

[1] J. Schubert, Conflict management in Dempster–Shafer theory by sequential discounting using the degree of falsity, in: L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (Eds.), Proceedings of the Twelfth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, 2008, pp. 298–305.
[2] A.P. Dempster, Upper and lower probabilities induced by a multiple valued mapping, The Annals of Mathematical Statistics 38 (1967) 325–339.
[3] A.P. Dempster, A generalization of Bayesian inference, Journal of the Royal Statistical Society Series B 30 (1968) 205–247.
[4] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[5] G. Shafer, Perspectives on the theory and practice of belief functions, International Journal of Approximate Reasoning 4 (1990) 323–362.
[6] G. Shafer, Rejoinders to comments on perspectives on the theory and practice of belief functions, International Journal of Approximate Reasoning 6 (1992) 445–480.
[7] P. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (1994) 191–234.
[8] R.Y. Yager, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster–Shafer Theory of Evidence, Wiley, New York, 1994.
[9] J. Schubert, On nonspecific evidence, International Journal of Intelligent Systems 8 (1993) 711–725.
[10] J. Schubert, Specifying nonspecific evidence, International Journal of Intelligent Systems 11 (1996) 525–563.
[11] J. Schubert, Creating prototypes for fast classification in Dempster–Shafer clustering, in: D.M. Gabbay, R. Kruse, A. Nonnengart, H.J. Ohlbach (Eds.), Proceedings of the First International Joint Conference on Qualitative and Quantitative Practical Reasoning, LNAI, 1244, Springer, Berlin, 1997, pp. 525–535.
[12] J. Schubert, Managing inconsistent intelligence, in: Proceedings of the Third International Conference on Information Fusion, International Society of Information Fusion, Sunnyvale, CA, 2000. pp. TuB4/10–16.

[13] J. Schubert, Clustering belief functions based on attracting and conflicting metalevel evidence, in: B. Bouchon-Meunier, L. Foulloy, R.R. Yager (Eds.), Intelligent Systems for Information Processing: From Representation to Applications, Elsevier, Amsterdam, 2003, pp. 349–360.
[14] J. Schubert, Clustering belief functions based on attracting and conflicting metalevel evidence using Potts spin mean field theory, Information Fusion 5 (2004) 309–318.
[15] J. Schubert, Managing decomposed belief functions, in: B. Bouchon-Meunier, R.R. Yager, C. Marsala, M. Rifqi (Eds.), Uncertainty and Intelligent Information Systems, World Scientific Publishing, Singapore, 2008, pp. 91–103.
[16] J. Schubert, Clustering decomposed belief functions using generalized weights of conflicts, International Journal of Approximate Reasoning 48 (2008) 466–480.
[17] J. Schubert, H. Sidenbladh, Sequential clustering with particle filtering – estimating the number of clusters from data, in: Proceedings of the Eighth International Conference on Information Fusion, International Society of Information Fusion, Sunnyvale, CA, 2005, pp. 1–8 (paper A4-3).
[18] A. Ayoun, P. Smets, Data association in multi-target detection using the transferable belief model, International Journal of Intelligent Systems 16 (2001) 1167–1182.
[19] D. Mercier, B. Quost, T. Denœux, Refined modeling of sensor reliability in the belief function framework using contextual discounting, Information Fusion 9 (2006) 246–258.
[20] A.-L. Jousselme, D. Grenier, É. Bossé, A new distance between two bodies of evidence, Information Fusion 2 (2001) 91–101.
[21] A. Martin, A.-L. Jousselme, C. Osswald, Conflict measure for the discounting operation on belief functions, in: Proceedings of the Eleventh International Conference on Information Fusion, IEEE, Piscataway, NJ, 2008, pp. 1003–1010.
[22] A.-L. Jousselme, P. Maupin, On some properties of distances in evidence theory, in: Proceedings of the Workshop on the Theory of Belief Functions, Brest, France, 2010, pp. 1–6 (paper 137).
[23] J. Klein, O. Colot, Automatic discounting rate computation using a dissent criterion, in: Proceedings of the Workshop on the Theory of Belief Functions, Brest, France, 2010, pp. 1–6 (paper 124).
[24] P. Smets, Analyzing the combination of conflicting belief functions, Information Fusion 8 (2007) 387–412.
[25] C.E. Shannon, A mathematical theory of communication, The Bell Systems Technical Journal 27 (1948). pp. 379–423, 623–656.
[26] R.V.L. Hartley, Transmission of information, The Bell System Technical Journal 7 (1928) 535–563.
[27] C.W.R. Chau, P. Lingras, S.K.M. Wong, Upper and lower entropies of belief functions using compatible probability functions, in: J. Komorowski, Z.W. Ras (Eds.), Proceedings of the Seventh International Symposium on Methodologies for Intelligent Systems, LNAI, 689, Springer, Berlin, 1993, pp. 306–315.
[28] Y. Maeda, H. Ichihashi, An uncertainty measure with monotonicity under the random set inclusion, International Journal of General Systems 21 (1993) 379–392.
[29] D. Harmanec, G.J. Klir, Measuring total uncertainty in Dempster–Shafer theory: a novel approach, International Journal of General Systems 22 (1994) 405–419.
[30] J. Abellán, G.J. Klir, S. Moral, Disaggregated total uncertainty measure for credal sets, International Journal of General Systems 35 (2006) 29–44.
[31] D. Dubois, H. Prade, A note on measures of specificity for fuzzy sets, International Journal of General Systems 10 (1985) 279–283.
[32] A. Meyerowitz, F. Richman, E. Walker, Calculating maximum-entropy probability for belief functions, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2 (1994) 377–389.
[33] D. Harmanec, G. Resconi, G.J. Klir, Y. Pan, On the computation of uncertainty measure in Dempster–Shafer theory, International Journal of General Systems 25 (1996) 153–163.