Refined modeling of sensor reliability in the belief function framework using contextual discounting

David Mercier[1,2], Benjamin Quost[1] and Thierry Denœux[1]

(1) UMR CNRS 6599 Heudiasyc
Université de Technologie de Compiègne
BP 20529 - F-60205 Compiègne cedex - France

(2) SOLYSTIC, 14 avenue Raspail,
F-94257 Gentilly Cedex, France

August 30, 2006

Abstract

In belief functions theory, the discounting operation allows to combine information provided by a source in the form of a belief function with meta-knowledge regarding the reliability of that source, resulting in a "weakened", less informative belief function. In this article, an extension of the discounting operation is proposed, allowing to use more detailed information regarding the reliability of the source in different contexts, i.e., conditionally on different hypotheses regarding the variable on interest. This results in a *contextual discounting* operation parameterized with a discount rate vector. Some properties of this contextual discounting operation are studied, and its relationship with classical discounting is explained. A method for learning the discount rates is also presented.

Keywords: Evidence Theory, Dempster-Shafer Theory, Belief Functions, Transferable Belief Model, Uncertainty, Information Fusion.

1 Introduction

In information fusion, it is usually important to take into account the reliability of the different sources in the evidence aggregation process [7, 6, 10, 16, 5]. In this paper, this problem is addressed using the Dempster-Shafer theory of belief functions [11], a powerful and flexible framework for representing and reasoning with various forms of imperfect information and knowledge.

In the belief function framework, knowledge about the reliability of a source of information (or sensor) is achieved by the discounting operation, which transforms each belief function provided by a source into a weaker, less informative one [11]. The discounting operation is controlled by a discount rate α taking values between 0 and 1: if $\alpha = 0$, the belief function is unchanged; if $\alpha = 1$, the belief function is transformed into the vacuous belief function, meaning that the information provided by the sensor is completely discarded. As shown by Smets [13], the discounting operation is not *ad hoc*, but it can be derived from a simple model of sensor reliability. In this model, the sensor can be in two states: reliable or not. If we know that the sensor is reliable, the belief function it provides is accepted without any modification. If we know that it is not reliable, we consider the information coming from the source as irrelevant. In practice, we do not know for sure whether the source is reliable or not, but we have some degree of belief, equal to $1 - \alpha$, in this hypothesis. Each piece of knowledge in this model can be translated into a belief function, and the combination of these belief functions leads to the discounting operation.

In the above model, knowledge about sensor or expert reliability is described by a single number. In certain cases, however, more refined knowledge is available. In particular, the reliability of the source of information can be expected to depend on the true value of the variable of interest. In medical diagnosis, for instance, a physician may be, due to his/her past experience or training, particularly competent to diagnose some types of diseases, while being less competent for other types. In target recognition, the performances of a data acquisition system may depend not only on weather conditions, but also on background and target properties [1, 5], making the reliability of the decision system dependent on the target at hand.

To account for such refined knowledge, we proposed in [9] an extension of the above discounting model, in which the user is allowed to quantify his/her confidence in the reliability of the source, conditionally on values or sets of values taken by the variable of interest. Combination of this information yields a new *contextual discounting* operation, controlled no longer by a single discount rate, but by a vector of discount rates describing the expected reliability of the source in different contexts. Here, properties of this operation are thoroughly studied, and its relationship with classical discounting is explained. A method for learning discount rates from data, generalizing the expert tuning method presented in [4], is also introduced.

The rest of this paper is organized as follows. Background material on belief functions is first recalled in Section 2. Contextual discounting, and a more general notion, Θ-contextual discounting, are then introduced in Sections 3 and 4, respectively. The problem of learning discount rates is then addressed in Section 5, and Section 6 concludes the paper.

2 Background on Belief Functions

Belief functions were first introduced by Dempster as a tool for statistical inference [2], and were later proposed by Shafer [11] as a general formalism for representing partial information and reasoning under uncertainty. Since then, different models based on belief functions have been proposed, including the Hints model [8] and the Transferable Belief Model (TBM) [15, 18]. These models have in common the basic mathematical apparatus of belief functions, but they differ at the semantic level. In the TBM, belief functions are interpreted as expressing weighted opinions, irrespective of any underlying probability distributions, whereas an underlying probability space is postulated in the Hints model (which is actually quite close to the initial Dempster's model). Also, the concepts of unnormalized mass functions and pignistic transformation are specific to the TBM. A discussion of these two models (as well as other interpretations of belief functions, such as random sets) can be found in [14]. Such a discussion is clearly out of the scope of this paper, where the TBM interpretation will be adopted for clarity of exposition. However, our approach is fully compatible with other interpretations.

2.1 Basic Concepts

Let x be a variable taking values in a finite set $\Omega = \{\omega_1, \ldots, \omega_K\}$, called the frame of discernment. The knowledge held by a rational agent Y, regarding the actual value ω_0 taken by x, given an evidential corpus EC, can be quantified by a basic belief assignment (BBA) $m_Y^\Omega[EC]$, defined as a function from 2^Ω to $[0,1]$ verifying:

$$\sum_{A \subseteq \Omega} m_Y^\Omega[EC](A) = 1 \ .$$

When there is no ambiguity, the full notation $m_Y^\Omega[EC]$ will be simplified to m_Y^Ω or m^Ω. The vacuous mass function, defined by $m^\Omega(\Omega) = 1$, represents total ignorance. Note that, in the TBM, BBAs are not required to be normalized, i.e., we may have $m^\Omega(\emptyset) > 0$. The interpretation of $m^\Omega(\emptyset)$ is discussed in [12].

The belief and plausibility functions associated with a BBA are defined, respectively, as:

$$bel^\Omega(A) = \sum_{\emptyset \neq B \subseteq A} m^\Omega(B),$$

and

$$pl^\Omega(A) = \sum_{B \cap A \neq \emptyset} m^\Omega(B), \quad \forall A \subseteq \Omega.$$

These functions play a central role in the TBM as they have easy interpretation: $bel^\Omega(A)$ is interpreted as a degree of justified support given to proposition A by the available evidence, whereas $pl^\Omega(A)$ is a measure of the maximum potential support that could be given to A, if further evidence became available. Related to bel^Ω and pl^Ω is the implicability function [17] b^Ω defined by $b^\Omega(A) = bel^\Omega(A) + m^\Omega(\emptyset) = 1 - pl^\Omega(\overline{A})$, $\forall A \subseteq \Omega$, where $\overline{A}$ is the complement of A.

The basic operation for combining BBAs induced by distinct sources of information is the *conjunctive rule of combination* (CRC) defined as

$$m_1^\Omega \cap\!\!\!\!\bigcirc\, m_2^\Omega(A) = \sum_{B \cap C = A} m_1^\Omega(B) m_2^\Omega(C), \qquad \forall A \subseteq \Omega. \tag{1}$$

This rule is commutative and associative. A disjunctive counterpart of the CRC is the *disjunctive rule of combination* (DRC) [3, 13], defined as:

$$m_1^\Omega \cup\!\!\!\!\bigcirc\, m_2^\Omega(A) = \sum_{B \cup C = A} m_1^\Omega(B) m_2^\Omega(C), \quad \forall A \subseteq \Omega. \tag{2}$$

The CRC applies when both sources are known to be reliable, whereas the DRC corresponds the hypothesis that at least one of the two sources is reliable [13]. The DRC can be conveniently expressed using the implicability function: the implicability function $b^{\Omega}_{1\bigcirc\!\!\!\!\!\cap\,2}$ associated with $m^{\Omega}_1 \bigcirc\!\!\!\!\!\cup\, m^{\Omega}_2$ can be obtained from b^{Ω}_1 and b^{Ω}_2, the implicability functions associated with m^{Ω}_1 and m^{Ω}_2, by pointwise multiplication: $b^{\Omega}_{1\bigcirc\!\!\!\!\!\cup\,2}(A) = b^{\Omega}_1(A) b^{\Omega}_2(A)$, for all $A \subseteq \Omega$.

2.2 Marginalization and Vacuous Extension

A BBA defined on a product frame $\Omega \times \Theta$ may be marginalized on Ω, by transferring and summing each mass $m^{\Omega\times\Theta}(B)$ for $B \subseteq \Omega \times \Theta$ to its projection on Ω:

$$m^{\Omega\times\Theta\downarrow\Omega}(A) = \sum_{\{B\subseteq\Omega\times\Theta \mid B\downarrow\Omega=A\}} m^{\Omega\times\Theta}(B), \quad \forall A \subseteq \Omega, \tag{3}$$

where $B \downarrow \Omega$ denotes the projection of B onto Ω.

It is usually not possible to retrieve the original BBA $m^{\Omega\times\Theta}$ from its marginalization $m^{\Omega\times\Theta\downarrow\Omega}$ on Ω. However, the least committed BBA such that its projection on Ω is $m^{\Omega\times\Theta\downarrow\Omega}$ may be computed. This defines the *vacuous extension* of m^{Ω} in the product frame $\Omega \times \Theta$ [13], given by:

$$m^{\Omega\uparrow\Omega\times\Theta}(B) = \begin{cases} m^{\Omega}(A) & \text{if } B = A \times \Theta \text{ for some } A \subseteq \Omega, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

2.3 Conditioning and Ballooning Extension

Conditional beliefs represent knowledge which is valid provided that a hypothesis is satisfied. Let m^{Ω} be a BBA, $A \subseteq \Omega$ an hypothesis and m^{Ω}_A the categorical BBA such as $m^{\Omega}_A(A) = 1$; the conditional belief function $m^{\Omega}[A]$ is:

$$m^{\Omega}[A] = m^{\Omega} \bigcirc\!\!\!\!\!\cap\, m^{\Omega}_A. \tag{5}$$

If $m^{\Omega\times\Theta}$ is defined on the product frame $\Omega\times\Theta$, and C is a subset of Θ, the conditional BBA $m^{\Omega}[C]$ is defined by combining $m^{\Omega\times\Theta}$ with $m^{\Theta\uparrow\Omega\times\Theta}_C$ (where m^{Θ}_C is the categorical BBA verifying $m^{\Theta}_C(C) = 1$), and marginalizing the result on Ω:

$$m^{\Omega}[C] = \left(m^{\Omega\times\Theta} \bigcirc\!\!\!\!\!\cap\, m^{\Theta\uparrow\Omega\times\Theta}_C\right)^{\downarrow\Omega}. \tag{6}$$

Assume now that $m^\Omega[C]$ represents your beliefs on Ω conditionally on C, i.e., in a context where C holds. There are usually many BBAs on $\Omega \times \Theta$, whose conditioning on C yields $m^\Omega[C]$. Among these, the least committed one is the ballooning extension [13] defined by:

$$m^\Omega[C]^{\Uparrow \Omega\times\Theta}(A \times C \cup \Omega \times \overline{C}) = m^\Omega[C](A), \quad \forall A \subseteq \Omega, \tag{7}$$

where $\overline{C}$ is the complement of C.

2.4 Specialization and Generalization

Let us consider two BBAs m_1^Ω and m_2^Ω. BBA m_1^Ω is said to be a specialization of m_2^Ω, or to be strongly included in m_2^Ω [3], if it can be obtained from m_2^Ω by transferring each mass $m_2^\Omega(A)$ to subsets of A. We then say that m_2^Ω is a generalization of m_1^Ω. This property may be interpreted in terms of information content: if m_1^Ω is a specialization of m_2^Ω, it is considered to have greater information content than m_2^Ω.

Mathematically, strong inclusion may be conveniently expressed using matrix calculus [17]. Assume that the belief masses are arranged in vectors of dimension 2^Ω. Let $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ be the vectors corresponding to m_1^Ω and m_2^Ω, respectively. Then m_2^Ω is a generalization of m_1^Ω iff $\boldsymbol{m}_2 = \boldsymbol{G}\boldsymbol{m}_1$, where $\boldsymbol{G} = [G(A,B)]$, $A, B \subseteq \Omega$, is a stochastic matrix verifying $G(A,B) = 0$, for all A and B such that $B \not\subseteq A$. With usual notations,

$$m_2^\Omega(A) = \sum_{B\subseteq\Omega} G(A,B) m_1^\Omega(B). \tag{8}$$

The term $G(A,B)$ represents the fraction of the mass $m_1^\Omega(B)$ which "flows up" to A, with $A \supseteq B$. Matrix $\boldsymbol{G}$ is referred to as a generalization matrix.

2.5 Discounting

Let us assume that agent Y receives a BBA m_S^Ω from a source S, describing the source's beliefs regarding the actual value ω_0. Moreover, Y has some knowledge about the *reliability* of S, quantified by a BBA $m_Y^\mathcal{R}$ on the frame $\mathcal{R} = \{R, NR\}$, where R stands for "the source is reliable", and NR for "the source is not reliable" [13]. Let us assume that $m_Y^\mathcal{R}$ has the following form:

$$\begin{cases} m_Y^\mathcal{R}(\{R\}) & = & 1-\alpha \\ m_Y^\mathcal{R}(\mathcal{R}) & = & \alpha, \end{cases} \tag{9}$$

for some $\alpha \in [0,1]$. Thus, α represents the plausibility that the source is not reliable ($pl_Y^{\mathcal{R}}(\{NR\}) = \alpha$), whereas $1-\alpha$ represents the degree of belief that it is reliable ($bel_Y^{\mathcal{R}}(\{R\}) = 1-\alpha$).

If S is reliable, the information provided by S becomes Y's knowledge:

$$m_Y^\Omega[R] = m_S^\Omega, \tag{10}$$

where the notation $m_Y^\Omega[R]$ is used in place of $m_Y^\Omega[\{R\}]$ for simplicity. If S is not reliable, the information provided by S cannot be taken into account, and Y's knowledge is vacuous:

$$m_Y^\Omega[NR](\Omega) = 1. \tag{11}$$

Therefore, we have two non-vacuous pieces of evidence, $m_Y^{\mathcal{R}}$ and $m_Y^\Omega[R]$. Assuming that they are distinct, they can be combined by vacuously extending $m_Y^{\mathcal{R}}$ to $\Omega \times \mathcal{R}$, computing the ballooning extension of $m_Y^\Omega[R]$ in the same frame, applying the CRC, and marginalizing the result on Ω:

$$m_Y^\Omega[m_S^\Omega, m_Y^{\mathcal{R}}] = \left(m_Y^\Omega[R]^{\Uparrow \Omega\times\mathcal{R}} \bigcirc\!\!\!\!\cap\, m_Y^{\mathcal{R}\uparrow\Omega\times\mathcal{R}} \right)^{\downarrow\Omega}. \tag{12}$$

The resulting BBA $m_Y^\Omega[m_S^\Omega, m_Y^{\mathcal{R}}]$ (where the brackets [ ] indicate the evidential corpus, i.e., what is known to the belief holder) only depends on m_S^Ω and α. Let us denote it by ${}^\alpha m_Y^\Omega$. It is equal to

$$\begin{cases} {}^\alpha m_Y^\Omega(A) & = & (1-\alpha) m_S^\Omega(A), \quad \forall A \subset \Omega, \\ {}^\alpha m_Y^\Omega(\Omega) & = & (1-\alpha) m_S^\Omega(\Omega) + \alpha. \end{cases} \tag{13}$$

The discounting operation can also be simply expressed using the implicability functions:

$${}^\alpha b_Y^\Omega(A) = \begin{cases} (1-\alpha) b_S^\Omega(A) & \text{if } A \neq \Omega \\ 1 & \text{otherwise.} \end{cases} \tag{14}$$

This operation was called discounting by Shafer [11, page 251], who introduced it on intuitive grounds. The formal justification presented here was proposed by Smets [13].

Remark 1 Note that ${}^\alpha m_Y^\Omega$ is obtained from m_S^Ω by transfering to Ω a fraction α of each mass $m_S^\Omega(B)$, for $B \neq \Omega$, and leaving the rest on B. Consequently, ${}^\alpha m_Y^\Omega$ is a generalization of m_S^Ω: we have

$$ {}^{\alpha}m_Y^\Omega(A) = \sum_{B\subseteq\Omega} {}^{\alpha}G(A,B)m_S^\Omega(B), \tag{15} $$

where ${}^{\alpha}\boldsymbol{G}$ is the generalization matrix defined by

$$ {}^{\alpha}G(A,B) = \begin{cases} 1-\alpha & \text{if } A = B \neq \Omega, \\ \alpha & \text{if } A = \Omega \text{ and } B \subset A, \\ 1 & \text{if } A = B = \Omega \\ 0 & \text{otherwise.} \end{cases} \tag{16} $$

(We recall that ${}^{\alpha}G(A,B)$ is equal to the fraction of $m_S^\Omega(B)$ transferred to A.) This matrix is shown for the case $K = 3$ in Table 1.

Table 1: Generalization matrix associated with classical discounting in the case $K = 3$ (only non zero terms are represented).

$A \setminus B$	$\emptyset$	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_1,\omega_2\}$	$\{\omega_3\}$	$\{\omega_1,\omega_3\}$	$\{\omega_2,\omega_3\}$	Ω
$\emptyset$	$1-\alpha$							
$\{\omega_1\}$		$1-\alpha$						
$\{\omega_2\}$			$1-\alpha$					
$\{\omega_1,\omega_2\}$				$1-\alpha$				
$\{\omega_3\}$					$1-\alpha$			
$\{\omega_1,\omega_3\}$						$1-\alpha$		
$\{\omega_2,\omega_3\}$							$1-\alpha$	
Ω	α	α	α	α	α	α	α	1

Remark 2 We can see ${}^{\alpha}m_Y^\Omega$ as the disjunctive combination of m_S^Ω with m_0^Ω defined by $m_0^\Omega(\emptyset) = 1-\alpha$ and $m_0^\Omega(\Omega) = \alpha$, since we have:

$$ m_S^\Omega \ovee m_0^\Omega(A) = m_S^\Omega(A)m_0^\Omega(\emptyset) = (1-\alpha)m_S^\Omega(A) = {}^{\alpha}m_Y^\Omega(A) $$

for all $A \subset \Omega$ and

$$ m_S^\Omega \ovee m_0^\Omega(\Omega) = m_S^\Omega(\Omega)m_0^\Omega(\emptyset) + m_0^\Omega(\Omega)\sum_{A\subseteq\Omega} m_S^\Omega(A) = (1-\alpha)m_S^\Omega(\Omega) + \alpha = {}^{\alpha}m_Y^\Omega(\Omega). $$

Remark 3 If $m_Y^{\mathcal{R}}$ is a Bayesian BBA:

$$\begin{cases} m_Y^{\mathcal{R}}(\{R\}) & = & 1-\alpha, \\ m_Y^{\mathcal{R}}(\{NR\}) & = & \alpha, \end{cases} \tag{17}$$

the result of the discounting is the same [13]. Note that we might have considered this model in the first place, instead of (9), as it may seem more natural to readers familiar with probability theory. However, we find it remarkable that expression (13) can be obtained assuming only weaker information regarding the reliability of the source, as expressed by (9).

Example 1 As in [4], let us consider a simplified aerial target recognition problem, in which we have three classes: airplane ($\omega_1 \equiv a$), helicopter ($\omega_2 \equiv h$) and rocket ($\omega_3 \equiv r$). Let $\Omega = \{a, h, r\}$. Assume that a sensor S has provided the following BBA for a given target: $m_S^{\Omega}(\{a\}) = 0.5$, $m_S^{\Omega}(\{r\}) = 0.5$, meaning that the sensor hesitates between classifying the target as an airplane or a rocket. Assume that the agent has a degree of belief equal to $1-\alpha$ that the sensor is reliable. A discount rate α is then applied to the sensor's BBA. With $\alpha = 0.4$, the agent's BBA becomes: ${}^{\alpha}m_Y^{\Omega}(\{a\}) = 0.5(1-\alpha) = 0.3$, ${}^{\alpha}m_Y^{\Omega}(\{r\}) = 0.5(1-\alpha) = 0.3$, ${}^{\alpha}m_Y^{\Omega}(\Omega) = \alpha = 0.4$.

3 Contextual Discounting

In this section, the above discounting operation is extended to allow the representation of more refined meta-knowledge regarding the reliability of the source of information in different contexts.

3.1 Basic Assumptions

Let us now assume that we have evidence regarding the reliability of a source S, conditionally on each $\omega_k \in \Omega$, i.e., in a context where the quantity x of interest is known to be equal to ω_k. We thus have K conditional BBAs $m_Y^{\mathcal{R}}[\omega_k]$, $k = 1, \ldots, K$, instead of the single unconditional BBA in (9). (As before, the notation $m_Y^{\mathcal{R}}[\omega_k]$ is used in place of $m_Y^{\mathcal{R}}[\{\omega_k\}]$ for simplicity). Assume that these BBAs have the following form:

$$\begin{cases} m_Y^{\mathcal{R}}[\omega_k](\{R\}) & = & \beta_k, \\ m_Y^{\mathcal{R}}[\omega_k](\mathcal{R}) & = & \alpha_k, \end{cases} \tag{18}$$

where $\beta_k = 1 - \alpha_k$. Thus, β_k represents our degree of belief that the source is reliable, when it is known that the actual value of x is ω_k, whereas α_k is equal to the plausibility that the source is not reliable, in the same context.

As before, we adopt the source's beliefs on Ω when we know that the source is reliable, and we consider the information provided by the source as totally irrelevant when we know that it is not reliable, as expressed by Equations (10) and (11), respectively.

We thus have $K+1$ non vacuous BBAs: $m_Y^{\Omega}[R] = m_S^{\Omega}$, and $m_Y^{\mathcal{R}}[\omega_k]$, $k = 1, \ldots, K$. To combine these BBAs, we have to compute their ballooning extensions in $\Omega \times \mathcal{R}$ using (7), combine them using the CRC, and marginalize the result on Ω. This may be compactly expressed by the following equation, which generalizes (12):

$$m_Y^{\Omega}\left[m_S^{\Omega}, m_Y^{\mathcal{R}}[\omega_1], \ldots, m_Y^{\mathcal{R}}[\omega_K]\right] = \left(m_Y^{\Omega}[R]^{\Uparrow\Omega\times\mathcal{R}} \circledcap m_Y^{\mathcal{R}}[\omega_1]^{\Uparrow\Omega\times\mathcal{R}} \circledcap \ldots \circledcap m_Y^{\mathcal{R}}[\omega_K]^{\Uparrow\Omega\times\mathcal{R}}\right)^{\downarrow\Omega}. \tag{19}$$

The result of this combination only depends on m_S^{Ω} and on vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$ of discount rates. It will be noted ${}^{\boldsymbol{\alpha}}m_Y^{\Omega}$ (which is consistent with the notation used in Section 2.5 for classical discounting), and the transformation $m_Y^{\Omega} \to {}^{\boldsymbol{\alpha}}m_Y^{\Omega}$ will be referred to as the *contextual discounting operation*, with discount rate vector $\boldsymbol{\alpha}$.

3.2 Expression of ${}^{\boldsymbol{\alpha}}m_Y^{\Omega}$

Proposition 1 *Let m_S^{Ω} denote a BBA on Ω, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K) \in [0,1]^K$. The contextually discounted BBA ${}^{\boldsymbol{\alpha}}m_Y^{\Omega}$ with discount rate vector $\boldsymbol{\alpha}$ is given by*

$${}^{\boldsymbol{\alpha}}m_Y^{\Omega}(A) = \sum_{B \subseteq \Omega} {}^{\boldsymbol{\alpha}}G(A,B) m_S^{\Omega}(B), \quad \forall A \subseteq \Omega, \tag{20}$$

with:

$${}^{\boldsymbol{\alpha}}G(A,B) = \begin{cases} \displaystyle\prod_{\omega_k \in A \setminus B} \alpha_k \prod_{\omega_\ell \in \overline{A}} \beta_\ell & \text{if } B \subseteq A, \\ 0 & \text{otherwise,} \end{cases} \tag{21}$$

where, following the classical mathematical convention, a product of terms is equal to 1 if the index set is empty.

Proof: See Section A.1. □

Coefficients ${}^{\boldsymbol{\alpha}}G(A,B)$ for all $A, B \subseteq \Omega$ define a generalization matrix: ${}^{\boldsymbol{\alpha}}G(A,B)$ is equal to the fraction of $m_S^\Omega(B)$ transferred to A. As shown by (21), this fraction increases with:

- the plausibility α_k that the source is not reliable, given that the actual value of the variable of interest is equal to ω_k, for each $\omega_k \in A \setminus B$;
- the degree of belief β_ℓ that the source is reliable, given that the actual value of the variable of interest is equal to ω_ℓ, for each $\omega_\ell \notin A$.

The property of classical discounting mentioned in Remark 2 has a counterpart with contextual discounting, as shown by the following proposition.

Proposition 2 *${}^{\boldsymbol{\alpha}}m^\Omega$ is the disjunctive combination of m_S^Ω with a bbm m_0^Ω defined by*

$$m_0^\Omega(C) = \prod_{\omega_k \in C} \alpha_k \prod_{\omega_\ell \in \overline{C}} \beta_\ell, \quad \forall C \subseteq \Omega \ . \tag{22}$$

Proof: See Section A.1. □

The following two propositions show that both ${}^{\boldsymbol{\alpha}}\boldsymbol{G}$ of Proposition 1 and m_0^Ω of Proposition 2 can be decomposed into simpler mathematical entities, each one corresponding to an elementary contextual discounting operation.

Proposition 3 *The BBA m_0^Ω of Proposition 2, can be written as:*

$$m_0^\Omega = m_1^\Omega \circledcup m_2^\Omega \circledcup \ldots \circledcup m_K^\Omega,$$

where each BBA m_k^Ω is defined by:

$$\begin{aligned} m_k^\Omega(\emptyset) &= \beta_k, \\ m_k^\Omega(\{\omega_k\}) &= \alpha_k. \end{aligned}$$

Proof: Obvious from (22) and the definition of the DRC (2). □

We note that each BBA m_k^Ω has the same form as m_0^Ω of Proposition 2, corresponding to a contextual discounting with discount rate vector $\boldsymbol{\alpha}_k$ with all components equal to 0, except component k equal to α_k. This shows that contextual discounting with discount rate rate vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$ can be decomposed into a sequence of contextual discountings with discount rate vectors $\boldsymbol{\alpha}_1 = (\alpha_1, 0, \ldots, 0)$, $\boldsymbol{\alpha}_2 = (0, \alpha_2, 0, \ldots, 0)$, $\ldots$, $\boldsymbol{\alpha}_K = (0, \ldots, 0, \alpha_K)$; we can write

$$^{\boldsymbol{\alpha}}m_Y^\Omega = {}^{\boldsymbol{\alpha}_K}({}^{\boldsymbol{\alpha}_{K-1}}(\ldots({}^{\boldsymbol{\alpha}_1}m_Y^\Omega)\ldots)).$$

Proposition 4 *The generalization matrix $^{\boldsymbol{\alpha}}\boldsymbol{G}$, defined in Proposition 1, can be expressed as:*

$$^{\boldsymbol{\alpha}}\boldsymbol{G} = \prod_{k=1}^{K} {}^{\boldsymbol{\alpha}_k}\boldsymbol{G},$$

where $^{\boldsymbol{\alpha}_k}\boldsymbol{G}$ is the generalization matrix defined by:

$$^{\boldsymbol{\alpha}_k}G(A,B) = \begin{cases} 1 & \textit{if } A = B \textit{ and } \omega_k \in B, \\ \beta_k & \textit{if } A = B \textit{ and } \omega_k \notin B, \\ \alpha_k & \textit{if } A \neq B \textit{ and } A = \{\omega_k\} \cup B, \\ 0 & \textit{otherwise.} \end{cases} \tag{23}$$

Proof: Immediate from Proposition 3. Each generalization matrix $^{\boldsymbol{\alpha}_k}\boldsymbol{G}$ corresponds to a contextual discounting with discount rate vector $\boldsymbol{\alpha}_k$. □

As another consequence of Proposition 3, the contextual discounting operation has a simple expression in terms of implicability and plausibility functions, as shown in the following proposition.

Proposition 5 *Let $^{\boldsymbol{\alpha}}b_Y^\Omega$ and $^{\boldsymbol{\alpha}}pl_Y^\Omega$ be, respectively, the implicability and plausibility functions associated with $^{\boldsymbol{\alpha}}m_Y^\Omega$. They can be obtained from b_S^Ω and pl_S^Ω, the corresponding functions associated with m_S^Ω, as:*

$$^{\boldsymbol{\alpha}}b_Y^\Omega(A) = b_S^\Omega(A) \prod_{\omega_k \in \overline{A}} \beta_k, \quad \forall A \subseteq \Omega,$$

and

$$^{\boldsymbol{\alpha}}pl_Y^\Omega(A) = 1 - (1 - pl_S^\Omega(A)) \prod_{\omega_k \in A} \beta_k, \quad \forall A \subseteq \Omega\ .$$

Proof: Let b_k^Ω be the implicability function associated with m_k^Ω of Proposition 3, for $k = 1, \ldots, K$. We have

$$b_k^\Omega(A) = \begin{cases} \beta_k & \text{if } \omega_k \notin A \\ 1 & \text{otherwise.} \end{cases}$$

Now, from Propositions 2 and 3, we have ${}^{\boldsymbol{\alpha}}b_Y^{\Omega} = b_S^{\Omega} \prod_{k=1}^{K} b_k^{\Omega}$, which yields the expression of ${}^{\boldsymbol{\alpha}}b_Y^{\Omega}$. The expression of ${}^{\boldsymbol{\alpha}}pl_Y^{\Omega}$ is then easily obtained using the equality ${}^{\boldsymbol{\alpha}}pl_Y^{\Omega}(A) = 1 - {}^{\boldsymbol{\alpha}}b_Y^{\Omega}(\overline{A})$. □

3.3 Discussion

Proposition 4, and particularly Equation (23) shed some light on the nature of the contextual discounting operation. For each value of k, each mass $m_S^{\Omega}(B)$ for $B \subseteq \Omega$ is unchanged if $\omega_k \in B$, whereas a fraction α_k is transferred to $\{\omega_k\} \cup B$ if $\omega_k \notin B$. This may be interpreted as follows: if the true state is ω_k, the agent's degree of belief that the source is reliable is β_k; consequently, this fraction of the mass initially assigned by the source to B remains focused on B, whereas the remaining part is transferred to the union of B and $\{\omega_k\}$. This operation is repeated for each B and each k.

Note that contextual discounting as defined in this section does *not* generalize the classical discounting recalled in Section 2.5: as will be shown in the examples below, the solution obtained by discounting m_S^{Ω} with rates $\alpha_i = \alpha$, $i = 1, \ldots, K$ is different, in general, from the one obtained using the classical discounting operation with a single rate α. Both classical and contextual discounting appear in fact to be two instances of a more general concept, which will be introduced in Section 4.

One may also wonder what would happen if the assumption expressed by (18) was changed, without increasing the number of parameters. Two alternatives would be to assign the mass α_k to $\{NR\}$, and either leave the mass β_k on $\{R\}$ (yielding a Bayesian BBA), or transfer it to $\mathcal{R}$. As shown in Appendix A.2, the same solution as derived in Section 3.2 is recovered in the first case, whereas the vacuous BBA is obtained in the second case.

3.4 Special Cases and Example

The generalization matrix ${}^{\boldsymbol{\alpha}}G$ associated with contextual discounting in the cases where $K = 2$ and $K = 3$ are shown in Tables 2 and 3, respectively.

In the case $K = 2$, we have:

$$
\begin{aligned}
{}^{\boldsymbol{\alpha}}m_Y^\Omega(\emptyset) &= \beta_1\beta_2 m_S^\Omega(\emptyset) \\
{}^{\boldsymbol{\alpha}}m_Y^\Omega(\{\omega_1\}) &= \alpha_1\beta_2 m_S^\Omega(\emptyset) + \beta_2 m_S^\Omega(\{\omega_1\}) \\
{}^{\boldsymbol{\alpha}}m_Y^\Omega(\{\omega_2\}) &= \alpha_2\beta_1 m_S^\Omega(\emptyset) + \beta_1 m_S^\Omega(\{\omega_1\}) \\
{}^{\boldsymbol{\alpha}}m_Y^\Omega(\Omega) &= \alpha_1\alpha_2 m_S^\Omega(\emptyset) + \alpha_2 m_S^\Omega(\{\omega_1\}) + \alpha_1 m_S^\Omega(\{\omega_2\}) + m_S^\Omega(\Omega)\ .
\end{aligned}
$$

We can see that the fraction of the mass $m_S^\Omega(\{\omega_1\})$ which is transferred to Ω is equal to $\alpha_2 = 1 - \beta_2$, where β_2 is the degree of belief in R conditionally on ω_2. When $m_S^\Omega(\emptyset) = 0$ and $\alpha_1 = \alpha_2$, then contextual discounting is, in this case, equivalent to classical discounting. This is not true, however, for $K = 3$. In general, contextual discounting transfers each mass $m_S^\Omega(A)$ to all supersets of A, and not only to Ω as does classical discounting.

Table 2: Generalization matrix associated with the contextual discounting in the case $K = 2$.

$A \setminus B$	$\emptyset$	$\{\omega_1\}$	$\{\omega_2\}$	Ω
$\emptyset$	$\beta_1\beta_2$			
$\{\omega_1\}$	$\alpha_1\beta_2$	β_2		
$\{\omega_2\}$	$\beta_1\alpha_2$		β_1	
Ω	$\alpha_1\alpha_2$	α_2	α_1	1

Example 2 Continuing Example 1, let us assume that a sensor provides a BBA m_S^Ω such that $m_S^\Omega(\{a\}) = 0.5$ and $m_S^\Omega(\{r\}) = 0.5$, and let us now consider the following states of belief regarding the reliability of the sensor:

Case 1: The plausibility α_1 that the sensor is not reliable when the source is an airplane is equal to 0.4, whereas the sensor is known to be fully reliable ($\alpha_2 = \alpha_3 = 0$) when the target is a helicopter or a rocket. The corresponding discount rate vector is $\boldsymbol{\alpha}_1 = (0.4, 0, 0)$. The generalization matrix ${}^{\boldsymbol{\alpha}_1}G$ is (with the subsets of Ω ordered as

Table 3: Generalization matrix associated with the contextual discounting in the case $K = 3$.

$A \setminus B$	$\emptyset$	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_1, \omega_2\}$	$\{\omega_3\}$	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$	Ω
$\emptyset$	$\beta_1\beta_2\beta_3$							
$\{\omega_1\}$	$\alpha_1\beta_2\beta_3$	$\beta_2\beta_3$						
$\{\omega_2\}$	$\beta_1\alpha_2\beta_3$		$\beta_1\beta_3$					
$\{\omega_1, \omega_2\}$	$\alpha_1\alpha_2\beta_3$	$\alpha_2\beta_3$	$\alpha_1\beta_3$	β_3				
$\{\omega_3\}$	$\beta_1\beta_2\alpha_3$				$\beta_1\beta_2$			
$\{\omega_1, \omega_3\}$	$\alpha_1\beta_2\alpha_3$	$\beta_2\alpha_3$			$\alpha_1\beta_2$	β_2		
$\{\omega_2, \omega_3\}$	$\beta_1\alpha_2\alpha_3$		$\beta_1\alpha_3$		$\beta_1\alpha_2$		β_1	
Ω	$\alpha_1\alpha_2\alpha_3$	$\alpha_2\alpha_3$	$\alpha_1\alpha_3$	α_3	$\alpha_1\alpha_2$	α_2	α_1	1

in Tables 1 and 3):

$$
{}^{\boldsymbol{\alpha}_1}G = \begin{pmatrix}
0.6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.6 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.4 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.6 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.4 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.6 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.4 & 1
\end{pmatrix},
$$

and the dicounted BBA ${}^{\boldsymbol{\alpha}_1}m$ is

$$
{}^{\boldsymbol{\alpha}_1}m_Y^\Omega(\{a\}) = 0.5, \quad {}^{\boldsymbol{\alpha}_1}m_Y^\Omega(\{r\}) = 0.5 \times 0.6 = 0.3, \quad {}^{\boldsymbol{\alpha}_1}m_Y^\Omega(\{a, r\}) = 0.5 \times 0.4 = 0.2 .
$$

We can see that a fraction 0.4 of the mass initially assigned to $\{r\}$ has been transferred to $\{a, r\}$, which can be interpreted as follows: if the target is an airplane, then the source is not reliable, and it may erroneously declare it as a rocket; consequently, when the source reports a rocket, it may actually be a rocket or an airplane.

Case 2: The plausibility α_2 that the sensor is not reliable when the source is a helicopter is equal to 0.6, whereas the sensor is known to be fully reliable ($\alpha_1 = \alpha_3 = 0$)

when the target is an airplane or a rocket. The corresponding discount rate vector is $\boldsymbol{\alpha}_2 = (0, 0.6, 0)$. The generalization matrix ${}^{\boldsymbol{\alpha}_2}G$ is:

$$
{}^{\boldsymbol{\alpha}_2}G = \begin{pmatrix}
0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.6 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.6 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.4 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.6 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.6 & 0 & 1
\end{pmatrix},
$$

and the dicounted BBA ${}^{\boldsymbol{\alpha}_2}m$ is

$$
{}^{\boldsymbol{\alpha}_2}m_Y^{\Omega}(\{a\}) = 0.5 \times 0.4 = 0.2, \quad {}^{\boldsymbol{\alpha}_2}m_Y^{\Omega}(\{a, h\}) = 0.5 \times 0.6 = 0.3,
$$

$$
{}^{\boldsymbol{\alpha}_2}m_Y^{\Omega}(\{r\}) = 0.5 \times 0.4 = 0.2, \quad {}^{\boldsymbol{\alpha}_2}m_Y^{\Omega}(\{h, r\}) = 0.5 \times 0.6 = 0.3 \ .
$$

The interpretation is similar to that of Case 1: this time, the masses given to $\{a\}$ and $\{r\}$ are partially transferred, respectively, to $\{a, h\}$ and $\{h, r\}$, to account for the low reliability of the source when the target is actually helicopter.

Case 3: We both have $\alpha_1 = 0.4$ and $\alpha_2 = 0.6$, meaning that the sensor has plausiblity 0.4 of not being reliable when the target is an airplane, and 0.6 when it is a helicopter. As before, the sensor is assumed to be fully reliable when the target is a rocket. The discount rate vector is this $\alpha = (0.4, 0.6, 0)$. The generalization matrix ${}^{\boldsymbol{\alpha}}G$ is then:

$$
{}^{\boldsymbol{\alpha}}G = \begin{pmatrix}
0.24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.16 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.36 & 0 & 0.6 & 0 & 0 & 0 & 0 & 0 \\
0.24 & 0.6 & 0.4 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.24 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.16 & 0.4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.36 & 0 & 0.6 & 0 \\
0 & 0 & 0 & 0 & 0.24 & 0.6 & 0.4 & 1
\end{pmatrix},
$$

and the BBA ${}^{\boldsymbol{\alpha}}m_Y$ is

$$ {}^{\boldsymbol{\alpha}}m_Y^\Omega(\{a\}) = 0.2, \quad {}^{\boldsymbol{\alpha}}m_Y^\Omega(\{a,h\}) = 0.3, $$

$$ {}^{\boldsymbol{\alpha}}m_Y^\Omega(\{r\}) = 0.12, \quad {}^{\boldsymbol{\alpha}}m_Y^\Omega(\{a,r\}) = 0.08, $$

$$ {}^{\boldsymbol{\alpha}}m_Y^\Omega(\{h,r\}) = 0.18, \quad {}^{\boldsymbol{\alpha}}m_Y^\Omega(\Omega) = 0.12. $$

It can be checked that, as a consequence of Proposition 3, we have

$$ {}^{\boldsymbol{\alpha}}G = {}^{\boldsymbol{\alpha}_1}G\, {}^{\boldsymbol{\alpha}_2}G = {}^{\boldsymbol{\alpha}_2}G\, {}^{\boldsymbol{\alpha}_1}G, $$

and ${}^{\boldsymbol{\alpha}}m_Y^\Omega = {}^{\boldsymbol{\alpha}_1}({}^{\boldsymbol{\alpha}_2}m_Y^\Omega) = {}^{\boldsymbol{\alpha}_2}({}^{\boldsymbol{\alpha}_1}m_Y^\Omega)$. Contextual discounting with discount rate vector $\boldsymbol{\alpha} = (0.4, 0.6, 0)$ is thus equivalent to contextual discounting with rate vector $\boldsymbol{\alpha}_1 = (0.4, 0, 0)$ as in Case 1, followed by contextual discounting with rate vector $\boldsymbol{\alpha}_2 = (0, 0.6, 0)$ as in Case 2. For instance the mass ${}^{\boldsymbol{\alpha}}m_Y^\Omega(\Omega) = 0.12$ can be explained by the transfer of 40 % of the mass $m_S^\Omega(\{r\}) = 0.5$ to $\{a, r\}$, of which 60% are then transferred to Ω.

4 Contextual Discounting Based on a Coarsening

In the previous section, we have assumed that the reliability of a source of information can be assessed conditionally on each $\omega_k \in \Omega$. In some situations, however, such detailed information will not be available. Nevertheless, we can still assume that we can assess the reliability of the source, for each element of a coarsening of Ω. As will be shown in this section, this assumption results in a new family of operations, generalizing both classical and contextual discounting.

4.1 Basic Assumptions

Let $\Theta = \{\theta_1, \ldots, \theta_L\}$ be a coarsening of Ω, which means that $\theta_1, \ldots, \theta_L$ form a partition of Ω. Let $m_Y^{\mathcal{R}}[\theta_k]$ denote the BBA on $\mathcal{R}$ quantifying our belief in the reliability of the source, when we know that the actual value of x is in θ_k (θ_k then constitutes a more general context than considered in the previous section). We assume that each $m_Y^{\mathcal{R}}[\theta_k]$, $k = 1, \ldots, L$ is of the following form:

$$ \begin{cases} m_Y^{\mathcal{R}}[\theta_k](\{R\}) &= \beta_k, \\ m_Y^{\mathcal{R}}[\theta_k](\mathcal{R}) &= \alpha_k, \end{cases} \tag{24} $$

where, as before, $\beta_k = 1 - \alpha_k$.

As in Section 3.1, we want to combine these L conditional belief functions with $m_Y^\Omega[R]$, which can be done by computing all the ballooning extensions, combining them using the CRC, and marginalizing on Ω:

$$m_Y^\Omega\left[m_S^\Omega, m_Y^\mathcal{R}[\theta_1], \ldots, m_Y^\mathcal{R}[\theta_L]\right] =$$
$$\left(m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}} \textcircled{\scriptsize$\cap$} m_Y^\mathcal{R}[\theta_1]^{\Uparrow\Omega\times\mathcal{R}} \textcircled{\scriptsize$\cap$} \ldots \textcircled{\scriptsize$\cap$} m_Y^\mathcal{R}[\theta_L]^{\Uparrow\Omega\times\mathcal{R}}\right)^{\downarrow\Omega}. \quad (25)$$

The result of this combination only depends on m_S^Ω, Θ, and vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_L)$ of discount rates. It will be noted ${}^{\boldsymbol{\alpha}}_{\Theta}m^\Omega$, and the transformation $m_Y^\Omega \rightarrow {}^{\boldsymbol{\alpha}}_{\Theta}m_Y^\Omega$ will be referred to as the Θ-*contextual discounting operation*, with discount rate vector $\boldsymbol{\alpha}$. Note that the contextual discounting defined in Section 3.2 corresponds to the special case where $L = K$ and $\theta_k = \{\omega_k\}$, $k = 1, \ldots, K$; it will be called Ω-contextual discounting for short. As will be shown later, Θ-contextual discounting also generalizes classical discounting, which corresponds to the case where $\Theta = \{\Omega\}$.

4.2 Expression of ${}^{\boldsymbol{\alpha}}_{\Theta}m^\Omega$

For any $A \subseteq \Omega$, let

$$A_* = \bigcup_{\{\theta\in\Theta, \theta\subseteq A\}} \theta,$$
$$A^* = \bigcup_{\{\theta\in\Theta, \theta\cap A\neq\emptyset\}} \theta,$$

and

$$\mathcal{C} = \{A \subseteq \Omega \mid \exists I \subseteq \{1, \ldots, L\}, A = \bigcup_{i\in I} \theta_i\}.$$

A_* and A^* are thus, respectively, the largest element of $\mathcal{C}$ included in A, and the smallest element of $\mathcal{C}$ that contains A.

With these notations, we have the following propositions, which generalize the results obtained in Section 3.2:

Proposition 6 *Let m_S^Ω denote a BBA on Ω, $\Theta = \{\theta_1, \ldots, \theta_L\}$ a coarsening of Ω, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_L) \in [0, 1]^L$. The Θ-contextually discounted BBA ${}^{\boldsymbol{\alpha}}_{\Theta}m_Y^\Omega$ with discount rate vector $\boldsymbol{\alpha}$ is given by*

$${}^{\boldsymbol{\alpha}}_{\Theta}m^\Omega(A) = \sum_{B\subseteq A} {}^{\boldsymbol{\alpha}}_{\Theta}G(A, B) m_S^\Omega(B), \quad \forall A \subseteq \Omega, \quad (26)$$

with:

$$ {}^{\boldsymbol{\alpha}}_{\Theta}G(A,B) = \begin{cases} \prod_{\cup\theta_k=(A\setminus B)^*} \alpha_k \prod_{\cup\theta_\ell=\overline{A}^*} \beta_\ell & \text{if } \exists C \in \mathcal{C}, B \cup C = A, \\ 0 & \text{otherwise.} \end{cases} \tag{27} $$

Proof: See Section A.3. □

Proposition 7 *${}^{\boldsymbol{\alpha}}_{\Theta}m^{\Omega}_{Y}$ is the disjunctive combination of m^{Ω}_{S} with a bbm m^{Ω}_{0} defined by*

$$ m^{\Omega}_{0}(C) = \begin{cases} \prod_{\cup\theta_k=C} \alpha_k \prod_{\cup\theta_\ell=\overline{C}} \beta_\ell & \text{if } C \in \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases} \tag{28} $$

Proof: See Section A.3. □

As in the case of Ω-contextual discounting considered in Section 3.2, the BBA m^{Ω}_{0} of Proposition 7 and the generalization matrix ${}^{\boldsymbol{\alpha}}_{\Theta}\boldsymbol{G}$ of Proposition 6 both admit simple decompositions, as described in the following two propositions.

Proposition 8 *The BBA m^{Ω}_{0} defined in Proposition 7 can be rewritten as:*

$$ m^{\Omega}_{0} = m^{\Omega}_{1} \bigcirc\!\!\!\!\cup\, m^{\Omega}_{2} \bigcirc\!\!\!\!\cup \ldots \bigcirc\!\!\!\!\cup\, m^{\Omega}_{L}, $$

where the m^{Ω}_{ℓ} ($\ell = 1, \ldots, L$) are defined by:

$$ \begin{aligned} m^{\Omega}_{\ell}(\emptyset) &= \beta_\ell, \\ m^{\Omega}_{\ell}(\theta_\ell) &= \alpha_\ell. \end{aligned} $$

Proof: Obvious from (28) and the definition of the DRC (2). Each BBA m^{Ω}_{ℓ} corresponds to a Θ-contextual discounting with discount rate vector $\boldsymbol{\alpha}_\ell$ with all components equal to 0, except component ℓ equal to α_ℓ. This shows that Θ-contextual discounting with rate vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_L)$ can be decomposed into L Θ-contextual discountings with rate vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_L$. □

Proposition 9 *The generalization matrix ${}^{\boldsymbol{\alpha}}_{\Theta}\boldsymbol{G}$, defined in Proposition 6, can be expressed as:*

$$ {}^{\boldsymbol{\alpha}}_{\Theta}G = \prod_{\ell=1}^{L} {}^{\boldsymbol{\alpha}_\ell}_{\Theta}\boldsymbol{G}, $$

where generalization matrix ${}^{\boldsymbol{\alpha}_\ell}_{\Theta}\boldsymbol{G}$ *associated with BBA* m_ℓ^Ω *is defined by:*

$$
{}^{\boldsymbol{\alpha}_\ell}_{\Theta}G(A,B) = \begin{cases} 1 & \text{if } A = B \text{ and } \theta_\ell \subseteq B, \\ \beta_\ell & \text{if } A = B \text{ and } \theta_\ell \not\subseteq B, \\ \alpha_\ell & \text{if } A \neq B \text{ and } A = \theta_\ell \cup B, \\ 0 & \text{otherwise.} \end{cases} \tag{29}
$$

Proof: Immediate from Proposition 8. Each generalization matrix ${}^{\boldsymbol{\alpha}_\ell}_{\Theta}\boldsymbol{G}$ corresponds to disjunctive combination with m_ℓ^Ω, and to Θ-contextual discounting with rate vector $\boldsymbol{\alpha}_\ell$. □

Proposition 10 *Let* ${}^{\boldsymbol{\alpha}}_{\Theta}b_Y^\Omega$ *and* ${}^{\boldsymbol{\alpha}}_{\Theta}pl_Y^\Omega$ *be, respectively, the implicability and plausibility functions associated with* ${}^{\boldsymbol{\alpha}}_{\Theta}m_Y^\Omega$*. They can be obtained from* b_S^Ω *and* pl_S^Ω*, the corresponding functions associated with* m_S^Ω*, as:*

$$
{}^{\boldsymbol{\alpha}}_{\Theta}b_Y^\Omega(A) = b_S^\Omega(A) \prod_{\cup\theta_\ell = \overline{A}^*} \beta_\ell, \quad \forall A \subseteq \Omega
$$

and

$$
{}^{\boldsymbol{\alpha}}_{\Theta}pl_Y^\Omega(A) = 1 - (1 - pl_S^\Omega(A)) \prod_{\cup\theta_\ell = A^*} \beta_\ell, \quad \forall A \subseteq \Omega \ .
$$

Proof: Let b_ℓ^Ω be the implicability function associated with BBA m_ℓ^Ω of Proposition 8, for $\ell = 1, \ldots, L$. We have

$$
b_\ell^\Omega(A) = \begin{cases} \beta_\ell & \text{if } \theta_\ell \not\subseteq A \\ 1 & \text{otherwise.} \end{cases}
$$

Now, from Propositions 7 and 8, we have ${}^{\boldsymbol{\alpha}}_{\Theta}b_Y^\Omega = b_S^\Omega \prod_{\ell=1}^L b_\ell^\Omega$, which yields the expression of ${}^{\boldsymbol{\alpha}}_{\Theta}b_Y^\Omega$. The expression of ${}^{\boldsymbol{\alpha}}_{\Theta}pl_Y^\Omega$ is then easily obtained using the equality ${}^{\boldsymbol{\alpha}}_{\Theta}pl_Y^\Omega(A) = 1 - {}^{\boldsymbol{\alpha}}_{\Theta}b_Y^\Omega(\overline{A})$. □

As mentioned above, Θ-contextual discounting generalizes the contextual discounting introduced in Section 3. Interestingly, it also generalizes the classical discounting operation recalled in Section 2.5, as shown by the following proposition.

Proposition 11 *Let* $\Theta = \{\Omega\}$ *denote the trivial partition of* Ω *in one class. Then,* Θ*-contextual discounting is identical to classical discounting.*

Proof: From Proposition 7, ${}^{\boldsymbol{\alpha}}m^{\Omega}$ is the disjunctive combination of m_S^{Ω} with a BBA m_0^{Ω} defined by $m_0^{\Omega}(\emptyset) = 1 - \alpha$ and $m_0^{\Omega}(\Omega) = \alpha$. Hence, from Remark 2, ${}^{\boldsymbol{\alpha}}m^{\Omega}$ is equal to the classical discounting of m_S^{Ω}. □

Example 3 Returning to Example 1, let us now assume that the agent has degrees of belief equal to 0.4 and 0.9 that the source is reliable when the target is, respectively, an airplane, and either a helicopter, or a rocket. The relevant coarsening here is $\Theta = \{\theta_1, \theta_2\}$, with $\theta_1 = \{a\}$, $\theta_2 = \{h, r\}$. Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$. The generalization matrix associated with this Θ-discounting is shown in Table 4 (where the notation $\Omega = \{\omega_1, \omega_2, \omega_3\}$ is used to allow comparison with Table 3).

Here, we have $\alpha_1 = 0.6$ and $\alpha_2 = 0.1$. If, as before, the BBA provided by the sensor is $m_S^{\Omega}(\{a\}) = 0.5$, $m_S^{\Omega}(\{r\}) = 0.5$, the result of the Θ-contextual discounted is then:

$$^{\boldsymbol{\alpha}}_{\Theta}m_Y^{\Omega}(\{a\}) = 0.45, \quad ^{\boldsymbol{\alpha}}_{\Theta}m_Y^{\Omega}(\Omega) = 0.05 + 0.03 = 0.08,$$

$$^{\boldsymbol{\alpha}}_{\Theta}m_Y^{\Omega}(\{r\}) = 0.18, \quad ^{\boldsymbol{\alpha}}_{\Theta}m_Y^{\Omega}(\{a, r\}) = 0.27, \quad ^{\boldsymbol{\alpha}}_{\Theta}m_Y^{\Omega}(\{h, r\}) = 0.02.$$

Table 4: Generalization matrix associated with Θ-contextual discounting in the case $K = 3$ with $\Theta = \{\{\omega_1\}, \{\omega_2, \omega_3\}\}$.

	$\emptyset$	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_1, \omega_2\}$	$\{\omega_3\}$	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$	Ω
$\emptyset$	$\beta_1\beta_2$							
$\{\omega_1\}$	$\alpha_1\beta_2$	β_2						
$\{\omega_2\}$			$\beta_1\beta_2$					
$\{\omega_1, \omega_2\}$			$\alpha_1\beta_2$	β_2				
$\{\omega_3\}$					$\beta_1\beta_2$			
$\{\omega_1, \omega_3\}$					$\alpha_1\beta_2$	β_2		
$\{\omega_2, \omega_3\}$	$\beta_1\alpha_2$		$\beta_1\alpha_2$		$\beta_1\alpha_2$		β_1	
Ω	$\alpha_1\alpha_2$	α_2	$\alpha_1\alpha_2$	α_2	$\alpha_1\alpha_2$	α_2	α_1	1

5 Learning of Discount Rates

In practice, the discount rates in the above Ω-contextual or Θ-contextual discounting operations will have to be either elicited from experts or learnt from data. The latter approach is considered in this section, generalizing the expert tuning method introduced in [4] for learning the discount rates in classical discounting.

5.1 Single Sensor

Following [4], let us assume that we have a training set related to n objects $o_1, \ldots, o_n$. Each object o_i belongs to a class in $\Omega = \{\omega_1, \ldots, \omega_K\}$. The class of object o_i is encoded by K binary indicator variables $\delta_{i,k}$, with $\delta_{i,k} = 1$ if object o_i belongs to class ω_k, and $\delta_{i,k} = 0$ otherwise.

Assume that a sensor or expert S provides, for each object o_i, a BBA $m_S^\Omega\{o_i\}$, quantifying its belief concerning the class of object o_i. Following previous work by Zouhal and Denœux [19] in pattern classification, Elouedi et al. [4] proposed to find a scalar discount rate α minimizing the following measure of discrepancy between beliefs and observations:

$$E_{bet}(\alpha) = \sum_{i=1}^{n} \sum_{k=1}^{K} ({}^{\alpha}BetP^{\Omega}\{o_i\}(\omega_k) - \delta_{i,k})^2, \tag{30}$$

where ${}^{\alpha}BetP^{\Omega}\{o_i\}$ is the pignistic probability distribution [18] associated with ${}^{\alpha}m^{\Omega}$, defined as:

$${}^{\alpha}BetP^{\Omega}\{o_i\}(\omega_k) = \sum_{\{A \subseteq \Omega, \omega_k \in A\}} \frac{{}^{\alpha}m^{\Omega}\{o_i\}(A)}{(1 - m^{\Omega}\{o_i\}(\emptyset))|A|}. \tag{31}$$

The minimization of $E_{bet}(\alpha)$ is a scalar constrained nonlinear programming problem. When all BBAs are normalized, i.e. $m^{\Omega}\{o_i\}(\emptyset) = 0$ for all i, then $E_{bet}(\alpha)$ is a quadratic function of α, whose minimum can be found analytically, as shown in [4].

The above expert tuning approach can easily be generalized to contextual discounting (only Ω-contextual discounting will be considered in this section, although the same approach can be extended to Θ-contextual discounting, once a coarsening Θ has been chosen). Let $E_{bet}(\boldsymbol{\alpha})$ be the same error function as defined in (30), where contextual discounting with discount rate vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$ is used in place of classical discounting. This is now a nonlinear K-dimensional function, whose minimization with respect to $\boldsymbol{\alpha}$ under the constraints $0 \leq \alpha_k \leq 1$ can be performed using a constrained nonlinear minimization procedure. The problem happens to be simpler if we replace $E_{bet}(\boldsymbol{\alpha})$ by the following alternative error function:

$$E_{pl}(\boldsymbol{\alpha}) = \sum_{i=1}^{n} \sum_{k=1}^{K} ({}^{\alpha}pl^{\Omega}\{o_i\}(\{\omega_k\}) - \delta_{i,k})^2, \tag{32}$$

which seems just as reasonable as (30). From Proposition 5, we can see that ${}^{\alpha}pl^{\Omega}\{o_i\}(\{\omega_k\})$ is an affine function of α_k:

$$\begin{aligned} {}^{\alpha}pl^{\Omega}\{o_i\}(\{\omega_k\}) &= 1 - (1 - pl^{\Omega}\{o_i\}(\{\omega_k\}))\beta_k \\ &= \alpha_k(1 - pl^{\Omega}\{o_i\}(\{\omega_k\})) + pl^{\Omega}\{o_i\}(\{\omega_k\}). \end{aligned}$$

Equation (32) can be written in matrix form as follows. Let

$$\boldsymbol{pl}_i = (pl^{\Omega}\{o_i\}(\{\omega_1\}), \ldots, pl^{\Omega}\{o_i\}(\{\omega_K\}))^T$$

denote the K-dimensional column vector containing the plausibilites of the singletons for object o_i, $\mathrm{diag}(1 - \boldsymbol{pl}_i)$ the $K \times K$ diagonal matrix whose diagonal elements are the complements to one of the components of $\boldsymbol{pl}_i$, and $\boldsymbol{\delta}_i = (\delta_{i,1}, \ldots, \delta_{i,K})^T$ the K-dimensional column vector of 0-1 class indicator variables for object o_i. Let

$$\boldsymbol{Q} = \begin{bmatrix} \mathrm{diag}(1 - \boldsymbol{pl}_1) \\ \vdots \\ \mathrm{diag}(1 - \boldsymbol{pl}_n) \end{bmatrix}, \qquad \boldsymbol{d} = \begin{bmatrix} \boldsymbol{\delta}_1 - \boldsymbol{pl}_1 \\ \vdots \\ \boldsymbol{\delta}_n - \boldsymbol{pl}_n \end{bmatrix}.$$

Then (32) can be writen as

$$E_{pl}(\boldsymbol{\alpha}) = \|\boldsymbol{Q}\boldsymbol{\alpha} - \boldsymbol{d}\|^2. \tag{33}$$

It is clear from the above formulation that the minimization of $E_{pl}(\boldsymbol{\alpha})$ is a constrained least-squares problem, which can be solved efficiently using standard iterative algorithms.

Example 4 Let us consider again the target recognition problem used in Example 1, and the data of Table 5, taken from [4]. Four objects with known classification have been classified by two sensors S_1 and S_2. Each sensor has provided, for each

Table 5: Data of Example 4, taken from [4].

	a	h	r	$\{a,h\}$	$\{a,r\}$	$\{h,r\}$	Ω	ground truth
$m_{S_1}\{o_1\}$	0	0	0.5	0	0	0.3	0.2	a
$m_{S_1}\{o_2\}$	0	0.5	0.2	0	0	0	0.3	h
$m_{S_1}\{o_3\}$	0	0.4	0	0	0.6	0	0	a
$m_{S_1}\{o_4\}$	0	0	0	0	0.6	0.4	0	r
$m_{S_2}\{o_1\}$	0	0	0	0.7	0	0	0.3	a
$m_{S_2}\{o_2\}$	0.3	0	0	0.4	0	0	0.3	h
$m_{S_2}\{o_3\}$	0.2	0	0	0	0	0.6	0.2	a
$m_{S_2}\{o_4\}$	0	0	0	0	0	1	0	r

object, a BBA on Ω, as shown in Table 5. We first consider the use of this data to optimize the discount rates to apply to each sensor individually. As reported in [4], the expert tuning approach (with a single discount rate and minimization of (30)) yields $\alpha = 0.66$ for sensor S_1, and $\alpha = 0.52$ for sensor S_2. With our approach (three contextual discount rates and minimization of (33)), we obtain $\boldsymbol{\alpha} = (0.24, 0, 0)$ for sensor S_1, and $\boldsymbol{\alpha} = (0.26, 0, 0)$ for sensor S_2. With a larger training set, such a result would indicate that both sensors are less reliable in the case where the target is an airplane, and that S_1 is slightly more reliable that S_2 (smaller values of α_k correspond to higher reliability). However, it is clear that such an interpretation is not valid with such a small training set, which is only used here as an illustration.

Remark 4 Note that, in this section, only the case of Ω-contextual discounting has been addressed. The generalization to Θ-contextual discounting is straightforward as long as the coarsening Θ is fixed. This seems a reasonable assumption in many applications where the choice of Θ is likely to be guided by the available meta-knowledge regarding the reliability of the sensor. If such knowledge does not exist, then one could consider searching for the optimal Θ, which is a completely different problem. If the cardinality of Ω is small, then an exhaustive search for the optimal partition may be feasable. The development of more efficient solutions to this problem is left for further study.

5.2 Multiple Sensors

When we have several sensors providing independent measurements, a usual strategy is to discount each sensor, and then combine them conjunctively using the CRC. In such a case, it seems preferable to optimize the performance of the combination, instead of optimizing the performance of each sensor individually.

For simplicity, assume that we have two sensors S_1 and S_2. Let ${}^{\boldsymbol{\alpha}_1}pl^{\Omega}_{S_1}\{o_i\}$ be the plausibility provided by sensor S_1 discounted with rates $\boldsymbol{\alpha}_1$, and ${}^{\boldsymbol{\alpha}_2}pl^{\Omega}_{S_2}\{o_i\}$ the plausibility provided by sensor S_2 discounted with rates $\boldsymbol{\alpha}_2$. After conjunctive combination, the plausibility of each singleton is equal to the product of the plausibilities given by the two sources (this results from the fact that the plausibility of a singleton is equal to its commonality). The plausibility of $\{\omega_k\}$ is thus equal to

$$ {}^{\boldsymbol{\alpha}_1}pl^{\Omega}_{S_1}\{o_i\}(\{\omega_k\}) \times {}^{\boldsymbol{\alpha}_2}\, pl^{\Omega}_{S_2}\{o_i\}(\{\omega_k\}). $$

Discount rate vectors $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ can thus be determined so as to minimize the following error function:

$$ E_{pl}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = \sum_{i=1}^{n} \sum_{k=1}^{K} \left({}^{\boldsymbol{\alpha}_1}pl^{\Omega}_{S_1}\{o_i\}(\{\omega_k\}) \times {}^{\boldsymbol{\alpha}_2}\, pl^{\Omega}_{S_2}\{o_i\}(\{\omega_k\}) - \delta_{i,k} \right)^2 . \qquad (34) $$

Note that this criterion is no longer quadratic. It can be minimized using a standard constrained nonlinear optimization procedure.

Example 5 With the data of Table 5, we obtain $\boldsymbol{\alpha}_1 = (0.45, 0, 0)$ and $\boldsymbol{\alpha}_2 = (0.39, 1, 0)$. Note that, as remarked in [4], this result should not be compared to the previous one: in the first case, we optimized the performance of each sensor individually, whereas in this second case we optimize the performances of the two sensors used in combination.

6 Conclusion

A new unary operation on discrete belief functions has been introduced. This contextual discounting operation is parameterized by a vector of discount rates representing degrees of belief in sensor reliability, conditionally on the variable of interest taking certain values or sets of values. These discount rates thus allow to express meta-knowledge about sensor reliability, at a level of detail chosen by the user. The classical discounting operation is recovered as a special case and corresponds to coarse grained knowledge of sensor reliability.

Choosing which Θ-contextual discounting operation to apply in a given situation will involve both selecting a suitable coarsening Θ of the frame of discernment, and determining discount rates. Once Θ has been fixed, we have shown that discount rates can be learnt automatically from labeled data by minimizing an error function. The automatic determination of Θ from data, however, has been left for further study.

A Proofs

A.1 Proofs of Propositions 1 and 2

The ballooning extension of $m_Y^{\mathcal{R}}[\omega_k]$ is:

$$\begin{aligned} m_Y^{\mathcal{R}\Uparrow\Omega\times\mathcal{R}}(\{\omega_k\} \times \{R\} \cup \overline{\{\omega_k\}} \times \mathcal{R}) &= \beta_k, & (35)\\ m_Y^{\mathcal{R}\Uparrow\Omega\times\mathcal{R}}(\Omega \times \mathcal{R}) &= \alpha_k. & (36)\end{aligned}$$

Let $m_r^{\Omega\times\mathcal{R}}$ be the conjunctive combination of the $m_Y^{\mathcal{R}}[\omega_k]^{\Uparrow\Omega\times\mathcal{R}}$, $k = 1, \ldots, K$. Using the following equality, for any $k \neq \ell$:

$$(\{\omega_k\} \times \{R\} \cup \overline{\{\omega_k\}} \times \mathcal{R}) \cap (\{\omega_\ell\} \times \{R\} \cup \overline{\{\omega_\ell\}} \times \mathcal{R}) = \{\omega_k, \omega_\ell\} \times \{R\} \cup \overline{\{\omega_k, \omega_\ell\}} \times \mathcal{R},$$

we easily obtain the expression of $m_r^{\Omega\times\mathcal{R}}$ as:

$$m_r^{\Omega\times\mathcal{R}}(C \times \{R\} \cup \overline{C} \times \mathcal{R}) = \begin{cases} \prod_{\omega_k \in \overline{C}} \alpha_k \prod_{\omega_\ell \in C} \beta_\ell & \text{if } C \neq \emptyset \text{ and } C \neq \Omega, \\ \prod_{k=1}^{K} \alpha_k & \text{if } C = \emptyset, \\ \prod_{\ell=1}^{K} \beta_\ell & \text{if } C = \Omega. \end{cases} \quad (37)$$

By exchanging the roles of C and $\overline{C}$, we obtain:

$$m_r^{\Omega\times\mathcal{R}}(\overline{C}\times\{R\}\cup C\times\mathcal{R}) = \begin{cases} \prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell & \text{if } C\neq\emptyset \text{ and } C\neq\Omega, \\ \prod_{k=1}^{K}\alpha_k & \text{if } C=\Omega, \\ \prod_{\ell=1}^{K}\beta_\ell & \text{if } C=\emptyset, \end{cases} \tag{38}$$

which can be noted more simply:

$$m_r^{\Omega\times\mathcal{R}}(\overline{C}\times\{R\}\cup C\times\mathcal{R}) = \prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell, \quad \forall C\subseteq\Omega. \tag{39}$$

The above BBA must now be combined with $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}}$:

$${}^{\boldsymbol{\alpha}}m_Y^\Omega = \left(m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}} \textcircled{\cap}\, m_r^{\Omega\times\mathcal{R}}\right)^{\downarrow\Omega} \tag{40}$$

The BBAs $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}}$ and $m_r^{\Omega\times\mathcal{R}}$ have focal sets of the form $B\times\{R\}\cup\Omega\times\{NR\}$ and $\overline{C}\times\{R\}\cup C\times\mathcal{R}$, respectively, with $B,C\subseteq\Omega$. The intersection of two such focal sets is:

$$(\overline{C}\times\{R\}\cup C\times\mathcal{R})\cap(B\times\{R\}\cup\Omega\times\{NR\}) = B\times\{R\}\cup C\times\{NR\},$$

and it can be obtained only for a particular choice of B and C. Then:

$$m_Y[R]^{\Uparrow\Omega\times\mathcal{R}} \textcircled{\cap}\, m_r^{\Omega\times\mathcal{R}}(B\times\{R\}\cup C\times\{NR\}) = \left[\prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell\right] m_S^\Omega(B). \tag{41}$$

Marginalizing this BBA on Ω gives:

$${}^{\boldsymbol{\alpha}}m^\Omega(A) = \sum_{B\cup C=A}\left[\prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell\right] m_S^\Omega(B), \quad \forall A\subseteq\Omega, \tag{42}$$

which can also be written as:

$${}^{\boldsymbol{\alpha}}m^\Omega(A) = \sum_{B\subseteq A} {}^{\boldsymbol{\alpha}}G(A,B)m_S^\Omega(B), \quad \forall A\subseteq\Omega, \tag{43}$$

with:

$${}^{\boldsymbol{\alpha}}G(A,B) = \sum_{C:B\cup C=A}\prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell\ , \ \forall B\subseteq A\subseteq\Omega, \tag{44}$$

and ${}^{\boldsymbol{\alpha}}G(A,B) = 0$ if $B \not\subseteq A$. Now, we have

$$\begin{aligned} B \cup C = A &\Leftrightarrow \exists D \subseteq B,\ C = (A \setminus B) \cup D \\ &\Leftrightarrow \exists D \subseteq B,\ \overline{C} = \overline{A} \cup (B \setminus D), \end{aligned}$$

and, consequently, for all A and B such that $B \subseteq A$:

$$\begin{aligned} {}^{\boldsymbol{\alpha}}G(A,B) &= \sum_{D \subseteq B} \prod_{\omega_k \in (A \setminus B) \cup D} \alpha_k \prod_{\omega_\ell \in \overline{A} \cup (B \setminus D)} \beta_\ell \\ &= \prod_{\omega_k \in A \setminus B} \alpha_k \prod_{\omega_\ell \in \overline{A}} \beta_\ell \sum_{D \subseteq B} \prod_{\omega_k \in B \setminus D} \beta_k \prod_{\omega_\ell \in D} \alpha_\ell \\ &= \prod_{\omega_k \in A \setminus B} \alpha_k \prod_{\omega_\ell \in \overline{A}} \beta_\ell \prod_{\omega_k \in B} (\alpha_k + \beta_k) \\ &= \prod_{\omega_k \in A \setminus B} \alpha_k \prod_{\omega_\ell \in \overline{A}} \beta_\ell, \end{aligned}$$

which completes the proof of Proposition 1.

Proposition 2 is obvious from (42).

A.2 Solutions obtained by modifying Equation (18)

Case 1: Assume that (18) is replaced by

$$\left\{ \begin{aligned} m_Y^{\mathcal{R}}[\omega_k](\{R\}) &= \beta_k, \\ m_Y^{\mathcal{R}}[\omega_k](\{NR\}) &= \alpha_k. \end{aligned} \right. \tag{45}$$

In this case, the ballooning extension of $m_Y^{\mathcal{R}}[\omega_k]$ is:

$$m_Y^{\mathcal{R}\Uparrow\Omega\times\mathcal{R}}(\{\omega_k\} \times \{R\} \cup \overline{\{\omega_k\}} \times \mathcal{R}) = \beta_k, \tag{46}$$

$$m_Y^{\mathcal{R}\Uparrow\Omega\times\mathcal{R}}(\{\omega_k\} \times \{NR\} \cup \overline{\{\omega_k\}} \times \mathcal{R}) = \alpha_k. \tag{47}$$

Let $m_r^{\Omega\times\mathcal{R}}$ be the conjunctive combination of the $m_Y^{\mathcal{R}}[\omega_k]^{\Uparrow\Omega\times\mathcal{R}}$, $k = 1, \ldots, K$. Using the following equalities, for any $k \neq \ell$:

$$(\{\omega_k\} \times \{R\} \cup \overline{\{\omega_k\}} \times \mathcal{R}) \cap (\{\omega_\ell\} \times \{R\} \cup \overline{\{\omega_\ell\}} \times \mathcal{R}) = \\ \{\omega_k, \omega_\ell\} \times \{R\} \cup \overline{\{\omega_k, \omega_\ell\}} \times \mathcal{R}, \tag{48}$$

$$(\{\omega_k\} \times \{R\} \cup \overline{\{\omega_k\}} \times \mathcal{R}) \cap (\{\omega_\ell\} \times \{NR\} \cup \overline{\{\omega_\ell\}} \times \mathcal{R}) = \\ \{\omega_k\} \times \{R\} \cup \{\omega_\ell\} \times \{NR\} \cup \overline{\{\omega_k, \omega_\ell\}} \times \mathcal{R}. \tag{49}$$

we can obtain the expression of $m_r^{\Omega\times\mathcal{R}}$ as:

$$m_r^{\Omega\times\mathcal{R}}(\overline{C}\times\{R\}\cup C\times\{NR\}) = \prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell, \quad \forall C\subseteq\Omega. \tag{50}$$

Like before (40), the above BBA must now be combined with $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}}$, whose focal sets are of the form $B\times\{R\}\cup\Omega\times\{NR\}$, with $B\subseteq\Omega$. The intersection of focal sets of $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}}$ and $m_r^{\Omega\times\mathcal{R}}$ is:

$$(\overline{C}\times\{R\}\cup C\times\{NR\})\cap(B\times\{R\}\cup\Omega\times\{NR\}) = (B\cap\overline{C})\times\{R\}\cup C\times\{NR\}.$$

Then:

$$m_Y[R]^{\Uparrow\Omega\times\mathcal{R}} \textcircled{\cap}\, m_r^{\Omega\times\mathcal{R}}((B\cap\overline{C})\times\{R\}\cup C\times\{NR\}) = \left[\prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell\right] m_S^\Omega(B). \tag{51}$$

Marginalizing this BBA on Ω gives:

$$\begin{aligned} {}^{\boldsymbol{\alpha}}m^\Omega(A) &= \sum_{(B\cap\overline{C})\cup C=A}\left[\prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell\right] m_S^\Omega(B), \quad \forall A\subseteq\Omega, &(52)\\ &= \sum_{B\cup C=A}\left[\prod_{\omega_k\in C}\alpha_k \prod_{\omega_\ell\in\overline{C}}\beta_\ell\right] m_S^\Omega(B), \quad \forall A\subseteq\Omega, &(53)\end{aligned}$$

and, we recognize equation (42).

Case 2: Assume that (18) is replaced by

$$\begin{cases} m_Y^{\mathcal{R}}[\omega_k](\mathcal{R}) &= \beta_k,\\ m_Y^{\mathcal{R}}[\omega_k](\{NR\}) &= \alpha_k. \end{cases} \tag{54}$$

In this case, the ballooning extension of $m_Y^{\mathcal{R}}[\omega_k]$ is:

$$\begin{aligned} m_Y^{\mathcal{R}\Uparrow\Omega\times\mathcal{R}}(\Omega\times\mathcal{R}) &= \beta_k, &(55)\\ m_Y^{\mathcal{R}\Uparrow\Omega\times\mathcal{R}}(\{\omega_k\}\times\{NR\}\cup\overline{\{\omega_k\}}\times\mathcal{R}) &= \alpha_k. &(56)\end{aligned}$$

Let $m_r^{\Omega\times\mathcal{R}}$ be the conjunctive combination of the $m_Y^{\mathcal{R}}[\omega_k]^{\Uparrow\Omega\times\mathcal{R}}$, $k=1,\ldots,K$. Using the following equality, for any $k\neq\ell$:

$$(\{\omega_k\}\times\{NR\}\cup\overline{\{\omega_k\}}\times\mathcal{R})\cap(\{\omega_\ell\}\times\{NR\}\cup\overline{\{\omega_\ell\}}\times\mathcal{R}) = \{\omega_k,\omega_\ell\}\times\{NR\}\cup\overline{\{\omega_k,\omega_\ell\}}\times\mathcal{R}, \tag{57}$$

we obtain the expression of $m_r^{\Omega\times\mathcal{R}}$ as:

$$m_r^{\Omega\times\mathcal{R}}(C \times \{NR\} \cup \overline{C} \times \mathcal{R}) = \prod_{\omega_k \in C} \alpha_k \prod_{\omega_\ell \in \overline{C}} \beta_\ell. \tag{58}$$

The above BBA must now be combined with $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}}$, whose focal sets are of the form $B \times \{R\} \cup \Omega \times \{NR\}$, with $B \subseteq \Omega$. The intersection of focal sets of $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}}$ and $m_r^{\Omega\times\mathcal{R}}$ is:

$$(C \times \{NR\} \cup \overline{C} \times \mathcal{R}) \cap (B \times \{R\} \cup \Omega \times \{NR\}) = \Omega \times \{NR\} \cup (\overline{C} \cap B) \times \{R\}.$$

Thus, all focal sets of $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}} \bigcirc\!\!\!\!\cap\; m_r^{\Omega\times\mathcal{R}}$ contain $\Omega \times \{NR\}$. When marginalizing on Ω, the mass of each focal sets is then transferred to Ω, and

$$^{\boldsymbol{\alpha}}m_Y^\Omega(\Omega) = \left(m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}} \bigcirc\!\!\!\!\cap\; m_r^{\Omega\times\mathcal{R}}\right)^{\downarrow\Omega}(\Omega) = 1. \tag{59}$$

A.3 Proofs of Propositions 6 and 7

Let $m_r^{\Omega\times\mathcal{R}}$ denote the conjunctive combination of the $m_Y^{\mathcal{R}}[\theta_k]^{\Uparrow\Omega\times\mathcal{R}}$, $k = 1, \ldots, L$. A similar line of reasoning as performed in the proof of Proposition 1 shows that the focal elements of $m_r^{\Omega\times\mathcal{R}}$ are of the form $\overline{C} \times \{R\} \cup C \times \mathcal{R}$ with $C \in \mathcal{C}$, and

$$m_r^{\Omega\times\mathcal{R}}(\overline{C} \times \{R\} \cup C \times \mathcal{R}) = \prod_{\cup\theta_k = C} \alpha_k \prod_{\cup\theta_\ell = \overline{C}} \beta_\ell, \tag{60}$$

which is the equivalent of (39).

After combination with $m_Y^\Omega[R]^{\Uparrow\Omega\times\mathcal{R}}$ and marginalization on Ω, we obtain:

$$\begin{aligned}
{}^{\boldsymbol{\alpha}}_{\Theta}m^\Omega(A) &= \sum_{B \cup C = A} m_r^{\Omega\times\mathcal{R}}(\overline{C} \times \{R\} \cup C \times \mathcal{R}) m_S^\Omega(B), \quad \forall A \subseteq \Omega \\
&= \sum_{B \cup C = A} \left[\prod_{\cup\theta_k = C} \alpha_k \prod_{\cup\theta_\ell = \overline{C}} \beta_\ell\right] m_S^\Omega(B), \quad \forall A \subseteq \Omega \\
&= \sum_{B \subseteq A} {}^{\boldsymbol{\alpha}}_{\Theta}G(A,B) m_S^\Omega(B), \quad \forall A \subseteq \Omega,
\end{aligned} \tag{61}$$

where ${}^{\boldsymbol{\alpha}}_{\Theta}G(A,B)$ denotes a coefficient of the generalization matrix associated with the Θ-contextual discounting:

$${}^{\boldsymbol{\alpha}}_{\Theta}G(A,B) = \begin{cases} \sum_{B \cup C = A} \prod_{\cup\theta_k = C} \alpha_k \prod_{\cup\theta_\ell = \overline{C}} \beta_\ell & \text{if } \exists C \in \mathcal{C}, B \cup C = A, \\ 0 & \text{otherwise.} \end{cases} \tag{62}$$

Now, let $A, B \subseteq \Omega$, and $C \in \mathcal{C}$ such that $B \cup C = A$. For any $\theta \subseteq C$, we have either $\theta \subseteq B$ or $\theta \cap (A \setminus B) \neq \emptyset$. Hence, we have $C = (A \setminus B)^* \cup D$ for some $D \in \mathcal{C}$, $D \subseteq B_*$. Similarly, if $\theta \subseteq \overline{C}$, we have either $\theta \subseteq (B \setminus D)$, or $\theta \cap \overline{A} \neq \emptyset$. Hence, $\overline{C} = (B \setminus D)_* \cup \overline{A}^*$. We can thus write

$$
\begin{aligned}
{}^{\boldsymbol{\alpha}}_{\Theta}G(A,B) &= \sum_{D \subseteq B_*} \prod_{\cup\theta_k = (A\setminus B)^* \cup D} \alpha_k \prod_{\cup\theta_\ell = (B\setminus D)_* \cup \overline{A}^*} \beta_\ell \\
&= \prod_{\cup_i\theta_k = (A\setminus B)^*} \alpha_k \prod_{\cup\theta_\ell = \overline{A}^*} \beta_\ell \sum_{D \subseteq B_*} \prod_{\cup\theta_k = D} \alpha_k \prod_{\cup\theta_\ell = (B\setminus D)^*} \beta_\ell \\
&= \prod_{\cup\theta_k = (A\setminus B)^*} \alpha_k \prod_{\cup\theta_\ell = \overline{A}^*} \beta_\ell \prod_{\cup\theta_k = B_*} (\alpha_k + \beta_k) \\
&= \prod_{\cup\theta_k = (A\setminus B)^*} \alpha_k \prod_{\cup\theta_\ell = \overline{A}^*} \beta_\ell \ ,
\end{aligned}
$$

which completes the proof of Proposition 6.

Proposition 7 results directly from (61).

References

[1] A. Appriou. Uncertain data aggregation in classification and tracking processes. In B. Bouchon-Meunier, editor, *Aggregation and Fusion of imperfect information*, pages 231–260. Physica-Verlag, Heidelberg, 1998.

[2] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

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

[4] Z. Elouedi, K. Mellouli, and Ph. Smets. Assessing sensor reliability for multisensor data fusion within the Transferable Belief Model. *IEEE Transactions on Systems, Man and Cybernetics B*, 34(1):782–787, 2004.

[5] S. Fabre, A. Appriou, and X. Briottet. Presentation and description of two classification methods using data fusion based on sensor management. *Information Fusion*, 2:49–71, 2001.

[6] S. Le Hégarat-Mascle, I. Bloch, and D. Vidal-Madjar. Introduction of neighborhood information in evidence theory and application to data fusion of radar and optical images with partial cloud cover. *Pattern recognition*, 31(11):1811–1823, November 1998.

[7] H. Kim and P. H. Swain. Evidential reasoning approach to multisource-data classification in remote sensing. *IEEE Transactions on Systems, Man and Cybernetics*, 25(8):1257–1265, 1995.

[8] J. Kohlas and P.-A. Monney. *A Mathematical Theory of Hints. An Approach to the Dempster-Shafer Theory of Evidence.* Springer-Verlag, Berlin, 1995.

[9] D. Mercier, B. Quost, and T. Denœux. Contextual discounting of belief functions. In L. Godo, editor, *Proceedings of ECQSARU'2005*, pages 552–562, Barcelona, Spain, July 2005. Springer-Verlag.

[10] N. Milisavljević and I. Bloch. Sensor fusion in anti-personnel mine detection using a two-level belief function model. *IEEE Transaction on Systems, Man and Cybernetics C*, 33(2):269–283, 2003.

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

[12] Ph. Smets. The nature of the unnormalized beliefs encountered in the Transferable Belief Model. In D. Dubois, M. P. Wellman, and B. D'Ambrosio, editors, *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, pages 292–297, San Mateo, CA, 1992. Morgan Kaufmann.

[13] Ph. Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.

[14] Ph. Smets. What is Dempster-Shafer's model ? In R. R. Yager, M. Fedrizzi, and J. Kacprzyk, editors, *Advances in the Dempster-Shafer theory of evidence*, pages 5–34. Wiley, New-York, 1994.

[15] Ph. Smets. The Transferable Belief Model for quantified belief representation. In D. M. Gabbay and Ph. Smets, editors, *Handbook of Defeasible reasoning and uncertainty management systems*, volume 1, pages 267–301. Kluwer Academic Publishers, Dordrecht, 1998.

[16] Ph. Smets. Data fusion in the Transferable Belief Model. In *Proceedings of FUSION'2000*, Paris, France, 2000.

[17] Ph. Smets. The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, 31(1–2):1–30, 2002.

[18] Ph. Smets and R. Kennes. The Transferable Belief Model. *Artificial Intelligence*, 66:191–243, 1994.

[19] L. M. Zouhal and T. Denœux. An evidence-theoretic k-NN rule with parameter optimization. *IEEE Trans. on Systems, Man and Cybernetics C*, 28(2):263–271, 1998.