The Application of the Matrix Calculus to Belief Functions.

Philippe Smets
IRIDIA
Université Libre de Bruxelles
50 av. Roosevelt, CP 194-6, 1050 Bruxelles, Belgium
psmets@ulb.ac.be
http://iridia.ulb.ac.be/˜psmets

August 10, 2004

Abstract

The mathematics of belief functions can be handled with the use of the matrix notation. This representation helps greatly the user thanks to its notational simplicity and its efficiency for proving theorems. We show how to use them for several problems related to belief functions and the transferable belief model.

Keywords: belief functions, transferable belief model, matrix.

1 Introduction

The mathematics of belief functions is often cumbersome because of the many summations symbols and all its subscripts. The equations are often hard to read and might discourage potential readers for their complexity (just as combinatorial calculus can discourage newcomers in probability theory). Most of the operations encountered in belief function theory happen to be linear operations and can be represented using the matrix notation (Smets, 1998; Monney, 2001). As usual with matrices, it helps greatly for the readability of the equations and the easiness in their manipulations. We present how matrix calculus can be used to help those working with belief functions. This matrix representation seems to be poorly known, even among belief function specialists, so we feel useful to present it. No result is really new, but we feel the content of this paper can be very useful for future work in belief function theory. Elementary knowledge of matrix calculus is required. Proofs are trivial but tedious. Most are obtained recursively.

1.1 Ordering the Elements of the Vectors

The belief function bel^Ω defined on the finite frame of discernment Ω, as well as its related basic belief assignment (bba) m^Ω, plausibility function pl^Ω, commonality function q^Ω, implicability functions b^Ω, can be seen as vectors in $R^{|\Omega|}$.

position	$c\,b\,a$	Ω	**m**	**bel**	**pl**
1	000	$\emptyset$	$m(\emptyset)$	$bel(\emptyset)$	$pl(\emptyset)$
2	001	a	$m(a)$	$bel(a)$	$pl(a)$
3	010	b	$m(b)$	$bel(b)$	$pl(b)$
4	011	ab	$m(ab)$	$bel(ab)$	$pl(ab)$
5	100	c	$m(c)$	$bel(c)$	$pl(c)$
6	101	ac	$m(ac)$	$bel(ac)$	$pl(ac)$
7	110	bc	$m(bc)$	$bel(bc)$	$pl(bc)$
8	111	abc	$m(abc)$	$bel(abc)$	$pl(abc)$

Table 1: Order of the elements of the vectors **m**, **bel** and **pl** when $\Omega = \{a, b, c\}$. We write ab for $\{a, b\}$, etc...

The order of their elements can be arbitrary, but one particular order turns out to be extremely practical, and helps enormously in discovering the underlying patterns encountered in the many relations we present. Furthermore it leads to very efficient algorithms in MatLab or similar programming languages.

For pedagogical purpose, many examples are presented on the frames $\Omega = \{a, b, c\}$ or $\Omega = \{a, b\}$. Generalizations are immediate. In the matrices, dots replace zeros as it enhances the matrix structure.

Let m^Ω be a bba defined on the frame of discernment $\Omega = \{a, b, c\}$. The elements of m^Ω are put in binary order. It means that the first element of m^Ω is the empty set, the next is $\{a\}$, the next is $\{b\}$, the next is $\{a, b\}$, etc... Table 1 presents what are the vectors for $\Omega = \{a, b, c\}$. In general, the i-th element of the vector $\mathbf{v} = [v_i]$ corresponds to the set which elements are those indicated by a 1 in the binary representation of $i - 1$. Suppose $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Consider the element v_{14}. The binary representation of $14 - 1$ is 1101 and the set is thus $\{\omega_1, \omega_3, \omega_4\}$: so $v_{14} = v(\{\omega_1, \omega_3, \omega_4\})$.

We use the following notations and conventions.

1. Matrices and vectors are written in bold types, and their elements in normal types, like in $\mathbf{A} = [A_{i,j}]$. By default, the lengths of the vectors and matrices are $2^{|\Omega|}$. Vectors are column vectors.

2. 0 and 1 denote the two scalars.

3. $\mathbf{0}$ denotes the column vector of length $2^{|\Omega|}$ which components are 0.

4. $\mathbf{1}$ denotes the column vector of length $2^{|\Omega|}$ which components are 1.

5. $\mathbf{1_A}$ denotes the column vector of length $2^{|\Omega|}$ which components are 0 except the component corresponding to $A \subseteq \Omega$ which value is 1.

6. $\mathbf{m}'$ and $\mathbf{M}'$ denote the transpose of the vector $\mathbf{m}$ and the matrix $\mathbf{M}$, respectively.

7. $\mathbf{Diag(v)}$ is the diagonal matrix which diagonal elements are the elements of the $\mathbf{v}$ vector.

8. For two vectors $\mathbf{u}$ and $\mathbf{v}$ on the same domain, we write $\mathbf{u} \geq \mathbf{v}$ to mean that $u_i \geq v_i$, $\forall i$, and symmetrically for $\mathbf{u} \leq \mathbf{v}$.

$$\mathbf{J} = \begin{bmatrix} . & . & . & . & . & . & . & 1 \\ . & . & . & . & . & . & 1 & . \\ . & . & . & . & . & 1 & . & . \\ . & . & . & . & 1 & . & . & . \\ . & . & . & 1 & . & . & . & . \\ . & . & 1 & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . \\ 1 & . & . & . & . & . & . & . \end{bmatrix}$$

Table 2: The $\mathbf{J}$ matrix for $|\Omega| = 3$. Dots replace zeros.

9. For notational simplicity sake, we write a for $\{a\}$, ab for $\{a, b\}$, abc for $\{a, b, c\}$, etc. . . , thus a list made of some of the symbols of the elements of Ω denotes the set that contains exactly these elements.

10. $\mathbf{I}$ denotes the unitary matrix, i.e., its elements are zeros except those on the main diagonal that are ones.

11. $\mathbf{J}$ denotes the square matrix which elements are zeros except those on the secondary diagonal that are ones (see Table 2). $\mathbf{J}$ equals its own transpose and its own inverse: $\mathbf{J}\cdot\mathbf{J} = \mathbf{I}$, or equivalently $\mathbf{J} = \mathbf{J}^{-1}$. Its major properties are that it inverses the order of the rows of a matrix when placed before it, the first becoming the last, etc. . . , and it inverses the order of the columns of a matrix when placed behind it.

1.2 Belief Functions

This work in presented in the transferable belief model framework (Smets & Kennes, 1994; Smets & Kruse, 1997; Smets, 1998). It means in particular that belief functions are not necessarily normalized. That is, we accept that positive masses may be given to the empty set and that $bel(\Omega)$ and $pl(\Omega)$ may be smaller than 1, as encountered among others under the open world assumption (Smets, 1988).

Let bel be a belief function defined on a finite frame of discernment Ω (Shafer, 1976). Several functions can be defined from bel, which are all in one to one correspondence. They are the basic belief assignment m (denoted bba), the implicability function b, the commonality function q and the plausibility function pl. In order to get all these relations easily accessible, we present them here. Except when specifically mentioned by an extra line, all these relations hold for all $A \subseteq \Omega$.

$$
\begin{array}{llll}
m-b: & m(A) = \sum_{B\subseteq A}(-1)^{|A|-|B|}b(B) & & b(A) = \sum_{B\subseteq A} m(B)\\
m-q: & m(A) = \sum_{A\subseteq B}(-1)^{|B|-|A|}q(B) & & q(A) = \sum_{A\subseteq B} m(B)\\
m-bel: & m(A) = \sum_{B\subseteq A}(-1)^{|A|-|B|}bel(B) & & bel(A) = \sum_{\emptyset\neq B\subseteq A} m(B)\\
 & m(\emptyset) = 1-bel(\Omega) & & bel(\emptyset) = 0\\
m-pl: & m(A) = \sum_{B\subseteq A}(-1)^{|A|-|B|+1}pl(\overline{B}) & & pl(A) = \sum_{B\cap A\neq\emptyset} m(B)\\
 & m(\emptyset) = 1-pl(\Omega) & & pl(\emptyset) = 0\\
b-q: & b(A) = \sum_{B\subseteq \overline{A}}(-1)^{|B|}q(B) & & q(A) = \sum_{B\subseteq A}(-1)^{|B|}b(\overline{B})\\
b-bel: & b(A) = bel(A)+m(\emptyset) & & bel(A) = b(A)-m(\emptyset)\\
b-pl: & b(A) = 1-pl(\overline{A}) & & pl(A) = 1-b(\overline{A})\\
q-bel: & q(A) = \sum_{B\subseteq A}(-1)^{|B|}bel(\overline{B}) & & bel(A) = \sum_{B\subseteq \overline{A}}(-1)^{|B|}q(B)\\
 & q(\emptyset) = 1 & & bel(\emptyset) = 0\\
q-pl: & q(A) = \sum_{B\subseteq A}(-1)^{|B|+1}pl(B) & & pl(A) = \sum_{\emptyset\neq B\subseteq A}(-1)^{|B|+1}q(B)\\
 & q(\emptyset) = 1 & & pl(\emptyset) = 0\\
bel-pl: & bel(A) = pl(\Omega)-pl(\overline{A}) & & pl(A) = bel(\Omega)-bel(\overline{A})
\end{array}
$$

Often the domain of these functions must be made explicit, a notation maybe cumbersome but that avoids confusions. When useful, we put the domain in superscript as in m^{Ω}.

1.3 The Least Commitment Principle

If a belief function is not fully defined, and the actual one belongs to a family of belief functions, we choose the belief function that belongs to the family and that is as least committed as possible. This principle is called the least commitment principle (LCP). The issue is to decide what criteria must be used to order belief functions according to their degree of 'commitment'.

Dubois and Prade (1987) have made three proposals to order belief functions. Let m_1 and m_2 be two bba's on Ω. The statement that m_1 is at least as committed as m_2 is denoted $m_1 \sqsubseteq_x m_2$ corresponding to some x-ordering where the subscript x can take three values. One of them is based on the concept of specialization which is explained in section 7. Then m_2 is said to be x-less committed than m_1.

The proposed orderings are:

- *pl-ordering.* If $pl_1(A) \leq pl_2(A)$ for all $A \subseteq \Omega$, we write $m_1 \sqsubseteq_{pl} m_2$
- *q-ordering.* If $q_1(A) \leq q_2(A)$ for all $A \subseteq \Omega$, we write $m_1 \sqsubseteq_q m_2$

- *s-ordering.* If m_1 is a specialization of m_2, we write $m_1 \sqsubseteq_s m_2$

where pl denotes the plausibility function and q denotes the commonality function.

Among all belief functions on Ω, the least committed belief function is the vacuous belief function (i.e. its bba is $\mathbf{1}_\Omega$).

The concept of 'least commitment' permits the construction of a partial order on the set of belief functions (Yager, 1986; Dubois & Prade, 1986).

The Principle of Minimal Commitment consists in selecting the least committed belief function in a set of equally justified belief functions. The principle formalizes the idea that one should never give more support than justified to any subset of Ω. It satisfies a form of scepticism, of noncommitment, of conservatism in the allocation of our belief. In its spirit, it is not far from what the probabilists try to achieve with the maximum entropy principle (Dubois & Prade, 1987; Hsia, 1991).

1.4 The Negation of a Belief Function

Dubois and Prade (1986) defined the concept of the negation of a belief function defined on Ω and which bba is m. The bba of the negation of m, denoted $\overline{m}$, is defined by:

$$\overline{m}(A) = m(\overline{A}),\ \forall A \subseteq \Omega.$$

In that case, we have:

$$\overline{b}(A) = q(\overline{A}), \qquad \overline{q}(A) = b(\overline{A}), \qquad \forall A \subseteq \Omega$$

where $\overline{b}$ and $\overline{q}$ denote, respectively, the implicability and the commonality function corresponding to $\overline{m}$.

These relations become:

$$\overline{\mathbf{m}} = \mathbf{J} \cdot \mathbf{m} \qquad \overline{\mathbf{b}} = \mathbf{J} \cdot \mathbf{q} \qquad \overline{\mathbf{q}} = \mathbf{J} \cdot \mathbf{b}$$

2 The Möbius Transforms

The bba is in fact the so-called Möbius transform of the belief function. All other transformations between the m, b, q and pl functions can be put in that family.

Suppose a bba m on $\Omega = \{a, b, c\}$. The classical relation between m and b

$$b(A) = \sum_{B \subseteq A} m(B),\ \forall A \subseteq \Omega$$

can be represented by:

$$\mathbf{b} = \mathbf{BfrM} \cdot \mathbf{m}$$

where $\mathbf{m}$ is the bba (a $2^{|\Omega|}$ column vector), $\mathbf{b}$ is the implicability function (a $2^{|\Omega|}$ column vector), and $\mathbf{BfrM}$ is a $2^{|\Omega|} \times 2^{|\Omega|}$ matrix which values are $BfrM(A, B) = 1$ iff $B \subseteq A$ and 0 otherwise. The full matrix is presented in Table 3, where $\Omega = \{a, b, c\}$. Thanks to the order used for the vector elements, the pattern becomes clear. The matrix is build from the $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ building block.

	∅	a	b	a,b	c	a,c	b,c	a,b,c
∅	1	.	.	.	.	.	.	.
a	1	1	.	.	.	.	.	.
b	1	.	1	.	.	.	.	.
a,b	1	1	1	1	.	.	.	.
c	1	.	.	.	1	.	.	.
a,c	1	1	.	.	1	1	.	.
b,c	1	.	1	.	1	.	1	.
a,b,c	1	1	1	1	1	1	1	1

Table 3: Matrix **BfrM** when $\Omega = \{a, b, c\}$

This block would be what the **BfrM** would be if $|\Omega| = 1$. To get the matrix when $|\Omega| = 2$, we reproduce the same block at the upper left, lower left, lower right corner, and fill the last corner with zeros. To get the matrix when $|\Omega| = 3$, we use the matrix we have built at the previous step, and proceed identically. This construction pattern is clearly indicated by the borders in Table 3. In fact, going from a set with i elements to a set with $i + 1$ elements consists in multiplying the initial block $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ by the matrix obtained with i elements using Kronecker multiplication.

$$\mathbf{BfrM}_{i+1} = kron(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \mathbf{BfrM}_i) \qquad \mathbf{BfrM}_1 = 1$$

or equivalently

$$\mathbf{BfrM}_{i+1} = \begin{bmatrix} \mathbf{BfrM}_i & \mathbf{0} \\ \mathbf{BfrM}_i & \mathbf{BfrM}_i \end{bmatrix}, \qquad \mathbf{BfrM}_0 = [1].$$

The code of the MatLab program that builds **BfrM** is given by:

```
BfrM =[1];
for i=1:cardinalΩ
      BfrM = [BfrM zeros(2^(i−1));BfrM BfrM];
end
```

For practical purpose, the transformation matrix that transforms the **Y** vector into the **X** vector is denoted **XfrY**. This order simplifies the control of the order used for matrix multiplication, as the last letter of first term must be equal to the first of the second term like in $\mathbf{XfrY} \cdot \mathbf{YfrZ} = \mathbf{XfrZ}$, a valid relation for all the transformations considered in this paper. In particular $\mathbf{XfrY} \cdot \mathbf{YfrX} = \mathbf{I}$.

The transformations matrices between $\mathbf{m}, \mathbf{b}$ and $\mathbf{q}$ satisfy:

$$\mathbf{BfrM}: \quad BfrM(A, B) = 1 \quad \text{if } B \subseteq A \quad = 0 \text{ otherwise} \tag{1}$$

$$\mathbf{MfrB}: \quad MfrB(A, B) = (-1)^{|A|-|B|} \quad \text{if } B \subseteq A \quad = 0 \text{ otherwise} \tag{2}$$

$$\mathbf{QfrM}: \quad QfrM(A, B) = 1 \quad \text{if } A \subseteq B \quad = 0 \text{ otherwise} \tag{3}$$

$$\mathbf{MfrQ}: \quad MfrQ(A, B) = (-1)^{|B|-|A|} \quad \text{if } A \subseteq B \quad = 0 \text{ otherwise} \tag{4}$$

$$\mathbf{BfrQ}: \quad BfrQ(A, B) = (-1)^{|B|} \quad \text{if } B \subseteq \overline{A} \quad = 0 \text{ otherwise} \tag{5}$$

$$\mathbf{QfrB}: \quad QfrB(A, B) = (-1)^{|B|} \quad \text{if } \overline{A} \subseteq B \quad = 0 \text{ otherwise} \tag{6}$$

$$\mathbf{BfrM} = \begin{bmatrix} 1 & . & . & . & . & . & . & . \\ 1 & 1 & . & . & . & . & . & . \\ 1 & . & 1 & . & . & . & . & . \\ 1 & 1 & 1 & 1 & . & . & . & . \\ 1 & . & . & . & 1 & . & . & . \\ 1 & 1 & . & . & 1 & 1 & . & . \\ 1 & . & 1 & . & 1 & . & 1 & . \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \qquad \mathbf{MfrB} = \begin{bmatrix} 1 & . & . & . & . & . & . & . \\ -1 & 1 & . & . & . & . & . & . \\ -1 & . & 1 & . & . & . & . & . \\ 1 & -1 & -1 & 1 & . & . & . & . \\ -1 & . & . & . & 1 & . & . & . \\ 1 & -1 & . & . & -1 & 1 & . & . \\ 1 & . & -1 & . & -1 & . & 1 & . \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix}$$

Table 4: The matrices between **m** and **b** for $|\Omega| = \{a, b, c\}$.

$$\mathbf{QfrM} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ . & 1 & . & 1 & . & 1 & . & 1 \\ . & . & 1 & 1 & . & . & 1 & 1 \\ . & . & . & 1 & . & . & . & 1 \\ . & . & . & . & 1 & 1 & 1 & 1 \\ . & . & . & . & . & 1 & . & 1 \\ . & . & . & . & . & . & 1 & 1 \\ . & . & . & . & . & . & . & 1 \end{bmatrix} \qquad \mathbf{MfrQ} = \begin{bmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ . & 1 & . & -1 & . & -1 & . & 1 \\ . & . & 1 & -1 & . & . & -1 & 1 \\ . & . & . & 1 & . & . & . & -1 \\ . & . & . & . & 1 & -1 & -1 & 1 \\ . & . & . & . & . & 1 & . & -1 \\ . & . & . & . & . & . & 1 & -1 \\ . & . & . & . & . & . & . & 1 \end{bmatrix}$$

Table 5: The matrices between **m** and **q** for $|\Omega| = \{a, b, c\}$.

All transformation can be build using one matrix, like **BfrM**. Given **J** and **BfrM** that we denote hereafter **B** for simplicity sake, we get:

$$m - b: \qquad \mathbf{MfrB} = \mathbf{B}^{-1} \tag{7}$$

$$m - q: \quad \mathbf{QfrM} = \mathbf{J} \cdot \mathbf{B} \cdot \mathbf{J} \qquad \mathbf{MfrQ} = \mathbf{J} \cdot \mathbf{B}^{-1} \cdot \mathbf{J} \tag{8}$$

$$b - q: \quad \mathbf{BfrQ} = \mathbf{B} \cdot \mathbf{J} \cdot \mathbf{B}^{-1} \cdot \mathbf{J} \qquad \mathbf{QfrB} = \mathbf{J} \cdot \mathbf{B} \cdot \mathbf{J} \cdot \mathbf{B}^{-1} \tag{9}$$

In Tables 4, 5 and 6, we present these matrices when $|\Omega| = 3$.

Just to enhance to simplicity achieved by the use of the matrix notation, we can show, in one line, the next relation.

$$\mathbf{BfrQ} = \mathbf{BfrM} \cdot \mathbf{J} \cdot \mathbf{BfrM}^{-1} \cdot \mathbf{J} = \mathbf{BfrM} \cdot \mathbf{J} \cdot \mathbf{MfrB} \cdot \mathbf{J}$$

Just think about the difficulty one would face when trying to prove it using only detailed summations.

2.1 The Other Transformations

The bba is the pivotal function in belief function theory. The **b** and the **q** functions are also essential, they are highly symmetrical in that **b** is the sum of the masses 'below' and **q** of those 'above', what is a short cut to express that $b(A)$ is the sum of all the masses given to subsets of A and symmetrically $q(A)$ is the sum of all the masses given to supersets of A.

The other two functions **bel** and **pl** are mathematically less convenient as one must always handle the empty set case. They could have been neglected if

$$\mathbf{BfrQ} = \begin{bmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & . & -1 & . & -1 & . & 1 & . \\ 1 & -1 & . & . & -1 & 1 & . & . \\ 1 & . & . & . & -1 & . & . & . \\ 1 & -1 & -1 & 1 & . & . & . & . \\ 1 & . & -1 & . & . & . & . & . \\ 1 & -1 & . & . & . & . & . & . \\ 1 & . & . & . & . & . & . & . \end{bmatrix} \qquad \mathbf{QfrB} = \begin{bmatrix} . & . & . & . & . & . & . & 1 \\ . & . & . & . & . & . & -1 & 1 \\ . & . & . & . & . & -1 & . & 1 \\ . & . & . & . & 1 & -1 & -1 & 1 \\ . & . & . & -1 & . & . & . & 1 \\ . & . & 1 & -1 & . & . & -1 & 1 \\ . & 1 & . & -1 & . & -1 & . & 1 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix}$$

Table 6: The matrices between $\mathbf{b}$ and $\mathbf{q}$ for $|\Omega| = \{a, b, c\}$.

it where not for the fact that they represent what common sense considers as beliefs and plausibilities.

$$\mathbf{bel} = \mathbf{b} - b(\emptyset)\mathbf{1}$$

$$\mathbf{pl} = \mathbf{1} - \mathbf{J} \cdot \mathbf{b}$$

so $\mathbf{BELfrM} = \mathbf{BfrM} - b(\emptyset)\mathbf{1} \cdot \mathbf{1}'$

2.2 Matrix Patterns

The major transforms are all based on a Kronecker multiplication based on a 2×2 matrix with one zero, and the other elements being 1 or -1, with the further property that vectors with positive terms can be mapped into vector with positive terms. Let $\mathcal{P}M$ be that family of matrices.

So $\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$ is included in $\mathcal{P}M$ whereas $\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$ are not included in $\mathcal{P}M$ as they always map vectors with positive terms into vectors with at least one negative term.

The family of possible matrices in $\mathcal{P}M$ must satisfy:

- on the row with 0, the other term is 1
- on the row with no 0, at most one term is -1.

We list the 12 possible matrices of $\mathcal{P}M$ in Table 7. In fact all of them can be derived from the $\mathbf{BfrM} = \mathbf{B}$ matrix. The relation is listed atop of the matrices in Table 7, using relations 7, 8 and 9.

3 The Fast Möbius Transform

Just as the Fourier transform is computed today using the FFT (Fast Fourier Transform) algorithm of Cooley-Tuckey, the Möbius transform can similarly be executed by a Fast Möbius Transform (FMT). It has been presented in (Kennes & Smets, 1991, 1990; Kennes, 1992). It is based on the discovery that the implicability function b can be computed from its bba m by the algorithm presented graphically in Figure 1. Suppose a bba m on a three element frame Ω. The column m lists the indexes of the bba vector. In fact, the index denotes the value of the vector element corresponding to the index. So ab denotes $m(a, b)$.

BfrM · J		**MfrB · J**		**QfrB**	
0	1	0	1	0	1
1	1	1	−1	−1	1
BfrM		**QfrB · J**		**MfrB**	
1	0	1	0	1	0
1	1	1	−1	−1	1
QfrM		**MfrQ**		**BfrQ · J**	
1	1	1	−1	−1	1
0	1	0	1	0	1
QfrM · J		**BfrQ**		**MfrQ · J**	
1	1	1	−1	−1	1
1	0	1	0	1	0

Table 7: The possible patterns of matrices in $\mathcal{P}M$.

We iteratively build vectors v_i, $i = 1, \ldots, n$ of length $2^{|\Omega|}$. The 'trick' consists in drawing the lines in Figure 1 for $n = 3$. A link is drawn from every element of v_{i-1} to v_i (with $v_0 = m$).

Then another link is drawn from elements 1,3,5,7 of m to the elements 2,4,6,8 of v_1, from elements 1 2, 5,6 of v_1 to elements 3 4,7,8 of v_2, and from elements 1,2,3,4 of v_2 to elements 5,6,7,8 of v_3. The detailed algorithm is given in (Kennes, 1992) These links are then used as follow. The value of the element j of the next vector is the sum of the values of those elements of the previous one connected to j. It means in practice that there are 'sums' of one term and sums of two terms.

We illustrate the computation in Figure 1. We compute the vector v_1. The values of the components of v_1 are obtained by adding those values of the m vectors that are linked to v_1 by a line. So $v_1(\emptyset) = m(\emptyset)$, $v_1(a) = m(\emptyset) + m(a)$, $v_1(b) = m(b)$, $v_1(a, b) = m(a) + m(a, b)$,... The symbols listed in the v_1 vector indicate the subsets of m which masses are included in the v_1 value. Then we build the vector v_2 by a similar method, adding the values of the v_1 vectors that are linked by a line. So $v_2(a, b)$ is obtained by adding $v_1(a)$ and $v_1(a, b)$ hence the masses added in $v_2(a, b)$ are $m(\emptyset)+m(a)+m(b)+m(a, b)$ as indicated by the labels of v_2. The v_3 vector is built similarly. For instance, $v_3(a, c) = v_2(a) + v_2(a, c) = m(\emptyset) + m(a) + m(c) + m(a, c)$ which is $b(a, c)$.

The MatLab code of the transformation from m to b is given for illustrative purpose.

v = m';
for i = 1 : n
    k = $2^{(n-i)}$;
    v = reshape(v,$2^{(i-1)}$,$2^{(n+1-i)}$);
    v(:,(1:k)*2) = v(:,(1:k)*2) + v(:,(1:k)*2 - 1);
end
b = reshape(v,1,2^n)';

Table 8 shows the matrix decomposition that underlies the FMT, the right most matrix performs the task of the left transformation in Figure 1, and so on.

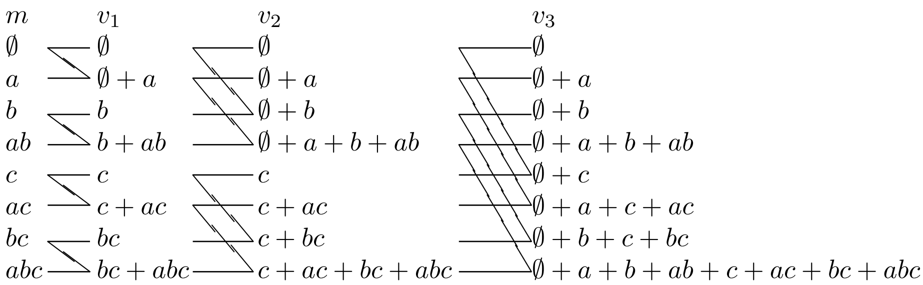

Figure 1: Detail of the FMT when $\Omega = \{a, b, c\}$. The symbol a denotes $m(a)$, ab denotes $m(a, b)$ etc...

$$\mathbf{BfrM} = \mathbf{M_3} \cdot \mathbf{M_2} \cdot \mathbf{M_1} \text{ where}$$

$$\mathbf{M_3} = \begin{bmatrix} 1 & . & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . \\ . & . & 1 & . & . & . & . & . \\ . & . & . & 1 & . & . & . & . \\ 1 & . & . & . & 1 & . & . & . \\ . & 1 & . & . & . & 1 & . & . \\ . & . & 1 & . & . & . & 1 & . \\ . & . & . & 1 & . & . & . & 1 \end{bmatrix} \qquad \mathbf{M_2} = \begin{bmatrix} 1 & . & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . \\ 1 & . & 1 & . & . & . & . & . \\ . & 1 & . & 1 & . & . & . & . \\ . & . & . & . & 1 & . & . & . \\ . & . & . & . & . & 1 & . & . \\ . & . & . & . & 1 & . & 1 & . \\ . & . & . & . & . & 1 & . & 1 \end{bmatrix}$$

$$\mathbf{M_1} = \begin{bmatrix} 1 & . & . & . & . & . & . & . \\ 1 & 1 & . & . & . & . & . & . \\ . & . & 1 & . & . & . & . & . \\ . & . & 1 & 1 & . & . & . & . \\ . & . & . & . & 1 & . & . & . \\ . & . & . & . & 1 & 1 & . & . \\ . & . & . & . & . & . & 1 & . \\ . & . & . & . & . & . & 1 & 1 \end{bmatrix}$$

Table 8: Decomposition of **BfrM** into the matrices that underlies the FMT when $\Omega = \{a, b, c\}$. $v_1 = M_1 \cdot m$, $v_2 = M_2 \cdot v_1$, $v_3 = M_3 \cdot v_2$ and $b = v_3$.

The other transformations from any of m, b, q, bel, pl into any of m, b, q, bel, pl are obtained similarly. All details are given in (Kennes, 1992). Their MatLab codes are available on http://iridia.ulb.ac.be/˜psmets

4 The Pignistic Transformation

When beliefs are represented by a belief function and a decision must be taken, we show that the transformation requires to build the needed probability function consists in building the so-called pignistic probability function. The justification of this transformation is presented in (Smets & Kennes, 1994; Smets, 2002). The equation of the pignistic transformation is given by:

$$BetP(A) = \sum_{B \subseteq \Omega} \frac{|A \cap B|}{|B|} \frac{m(B)}{1 - m(\emptyset)}$$

We consider the matrix notation for $\Omega = \{a, b, c\}$. The computation is done as if m was normalized. If it is not the case, proceed with the computation as such and normalize the results at the end.

We present both the relations to compute all the propositions at once, or only those on the singletons, the other probabilities being computed from these last. Obviously the second is computationaly more efficient. The first representation is given in order to be complete.

The pignistic transformation becomes:

$$\mathbf{BetP} = \mathbf{BetPfrM} \cdot \mathbf{m}$$

where **BetP** is the column vector which elements are $BetP(A), A \subseteq \Omega$, and

$$\mathbf{BetPfrM} = \begin{bmatrix} . & . & . & . & . & . & . & . \\ . & 1/1 & . & 1/2 & . & 1/2 & . & 1/3 \\ . & . & 1/1 & 1/2 & . & . & 1/2 & 1/3 \\ . & 1/1 & 1/1 & 2/2 & . & 1/2 & 1/2 & 2/3 \\ . & . & . & . & 1/1 & 1/2 & 1/2 & 1/3 \\ . & 1/1 & . & 1/2 & 1/1 & 2/2 & 1/2 & 2/3 \\ . & . & 1/1 & 1/2 & 1/1 & 1/2 & 2/2 & 2/3 \\ . & 1/1 & 1/1 & 2/2 & 1/1 & 2/2 & 2/2 & 3/3 \end{bmatrix}$$

BetPfrM can be also represented as: $\mathbf{BetPfrM} = \mathbf{CardAB} \cdot \mathbf{D}$ where

$$\mathbf{CardAB} = \begin{bmatrix} . & . & . & . & . & . & . & . \\ . & 1 & . & 1 & . & 1 & . & 1 \\ . & . & 1 & 1 & . & . & 1 & 1 \\ . & 1 & 1 & 2 & . & 1 & 1 & 2 \\ . & . & . & . & 1 & 1 & 1 & 1 \\ . & 1 & . & 1 & 1 & 2 & 1 & 2 \\ . & . & 1 & 1 & 1 & 1 & 2 & 2 \\ . & 1 & 1 & 2 & 1 & 2 & 2 & 3 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} . & . & . & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . \\ . & . & 1 & . & . & . & . & . \\ . & . & . & 1/2 & . & . & . & . \\ . & . & . & . & 1 & . & . & . \\ . & . & . & . & . & 1/2 & . & . \\ . & . & . & . & . & . & 1/2 & . \\ . & . & . & . & . & . & . & 1/3 \end{bmatrix}, \mathbf{CardA} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \\ 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}$$

$\mathbf{D} = \mathbf{Diag}(\mathbf{CardA}^{-})$ where $\mathbf{CardA}^{-}$ is the column vector which elements are the inverse of the elements of **CardA** and where by definition $1/0 = 0$. The matrix **CardAB** and the vector **CardA** can be build iteratively as follows. The i index denotes their value after having included i elements. The algorithm proceeds for $i = 1$ to $i = |\Omega| - 1$.

$$\mathbf{CardAB}_{i+1} = \begin{bmatrix} \mathbf{CardAB}_i & \mathbf{CardAB}_i \\ \mathbf{CardAB}_i & \mathbf{1} \cdot \mathbf{1}' + \mathbf{CardAB}_i \end{bmatrix} \qquad \mathbf{CardAB}_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{CardA}_{i+1} = \begin{bmatrix} \mathbf{CardA}_i \\ \mathbf{1} \cdot \mathbf{1}' + \mathbf{CardA}_i \end{bmatrix} \qquad \mathbf{CardA}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This approach is of course computationally inefficient. It is enough to compute **BetP** on the elements of Ω and to compute $BetP(A)$ for $A \subseteq \Omega$ by adding its values on the elements of A. Let **betP** be the column vector of length $|\Omega|$ and which elements $betP(x)$ are $BetP(\{x\})$ for $x \in \Omega$. Then $\mathbf{betP} = \mathbf{betPfrM} \cdot \mathbf{m}$ with:

$$\mathbf{betPfrM} = \begin{bmatrix} . & 1 & . & 1/2 & . & 1/2 & . & 1/3 \\ . & . & 1 & 1/2 & . & . & 1/2 & 1/3 \\ . & . & . & . & 1 & 1/2 & 1/2 & 1/3 \end{bmatrix}$$

The construction of **betPfrM** can be achieved using $\mathbf{betPfrM} = \mathbf{SupS} \cdot \mathbf{D}$ where:

$$\mathbf{SupS} = \begin{bmatrix} . & 1 & . & 1 & . & 1 & . & 1 \\ . & . & 1 & 1 & . & . & 1 & 1 \\ . & . & . & . & 1 & 1 & 1 & 1 \end{bmatrix}$$

and **SupS** is built iteratively as:

$$\mathbf{SupS}_{i+1} = \begin{bmatrix} \mathbf{SupS}_i & \mathbf{SupS}_i \\ \mathbf{0}' & \mathbf{1}' \end{bmatrix} \qquad \mathbf{SupS}_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

where $\mathbf{0}'$ and $\mathbf{1}'$ are the line vector of appropriate length which values are all 0 or 1, respectively.

It might be worth looking if one could describe operators to build $betP$ directly from b or q. The matrix for $\mathbf{betP} = \mathbf{betPfrB} \cdot \mathbf{b}$ given below does not seem to be easily synthesized. On the contrary, the matrix **betPfrQ** in $\mathbf{betP} = \mathbf{betPfrQ} \cdot \mathbf{q}$ is very simple. **betPfrQ** is obtained directly from **betPfrM** by multiplying by -1 the coefficients of the columns of **betPfrM** which index B had an even cardinality.

$$\mathbf{betPfrB} = 1/6 \cdot \begin{bmatrix} -2 & 2 & -1 & 1 & -1 & 1 & -2 & 2 \\ -2 & -1 & 2 & 1 & -1 & -2 & 1 & 2 \\ -2 & -1 & -1 & -2 & 2 & 1 & 1 & 2 \end{bmatrix}$$

$$\mathbf{betPfrQ} = \begin{bmatrix} . & 1 & . & -1/2 & . & -1/2 & . & 1/3 \\ . & . & 1 & -1/2 & . & . & -1/2 & 1/3 \\ . & . & . & . & 1 & -1/2 & -1/2 & 1/3 \end{bmatrix}$$

5 The Interaction Indices $\mathcal{I}_\omega$

Grabisch (1996) introduces the concept of interaction indices $\mathcal{I}_\omega$ for $\omega \subseteq \Omega$ in the belief function framework. Let the bba m^Ω, then the interaction indices are defined as:

$$\mathcal{I}_\omega = \sum_{A \subseteq \overline{\omega}} \frac{m^\Omega(\omega \cup A)}{|A|+1} \qquad \forall \omega \subseteq \Omega. \tag{10}$$

Note that when ω is a singleton of Ω and m^Ω is normalized, $\mathcal{I}_\omega = BetP^\Omega(\omega)$ (see section 4). The pignistic transformation produces the interaction index on the singletons, and these are just the so-called Shapley values described in cooperative games.

Let $\mathcal{I}$ be the vector of the interaction indices. The matrix representation of relation (10) is $\mathcal{I} = \mathbf{IfrM} \cdot \mathbf{m}$. When $|\Omega| = 3$, **IfrM** is:

$$
\mathbf{IfrM} = \begin{bmatrix}
1/1 & 1/2 & 1/2 & 1/3 & 1/2 & 1/3 & 1/3 & 1/4 \\
. & 1/1 & . & 1/2 & . & 1/2 & . & 1/3 \\
. & . & 1/1 & 1/2 & . & . & 1/2 & 1/3 \\
. & . & . & 1/1 & . & . & . & 1/2 \\
. & . & . & . & 1/1 & 1/2 & 1/2 & 1/3 \\
. & . & . & . & . & 1/1 & . & 1/2 \\
. & . & . & . & . & . & 1/1 & 1/2 \\
. & . & . & . & . & . & . & 1/1
\end{bmatrix}
$$

The coefficient $\mathbf{IfrM}(A, B)$ of $\mathbf{IfrM}$ is given by $1/(1 + |B \cap \overline{A}|)$ if $A \subseteq B$ and 0 otherwise. Inverse formulas are given in (Grabisch, 1996). The vectors m^Ω and $\mathcal{I}$ are in one to one correspondence as $\mathbf{IfrM}$ is not singular.

6 Transformations of bba into bba

Let $\mathcal{M}^\Omega$ be the set of bba defined on Ω. For any $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}^\Omega$, we can define a matrix $\mathbf{M}$ so that

$$\mathbf{m}_2 = \mathbf{M} \cdot \mathbf{m}_1.$$

One solution is $\mathbf{M} = \mathbf{m}_2 \cdot \mathbf{1}'$. Indeed $\mathbf{1}' \cdot \mathbf{m}_1 = 1$ and $\mathbf{m}_2 \cdot \mathbf{1}' \cdot \mathbf{m}_1 = \mathbf{m}_2$.

Definition 6.1 *A stochastic matrix $\mathbf{A} = [a_{ij}]$ is a square matrix with $a_{ij} \geq 0$ and*

$$\sum_i a_{ij} = 1, \ \forall j.$$

Let $\mathcal{SM}_\Omega$ denote the set of $2^{|\Omega|} \times 2^{|\Omega|}$ stochastic matrices.

Theorem 6.1 *The set of matrices that maps every element of $\mathcal{M}^\Omega$ into an element of $\mathcal{M}^\Omega$ is $\mathcal{SM}_\Omega$.*

Proof. 1. If $\mathbf{M} \in \mathcal{SM}_\Omega$, then for all $\mathbf{m} \in \mathcal{M}^\Omega$, one has $\mathbf{M} \cdot \mathbf{m} \in \mathcal{M}^\Omega$ as all its values are non negative and $\mathbf{1}' \cdot \mathbf{M} \cdot \mathbf{m} = \mathbf{1}' \cdot \mathbf{m} = 1$.

2. If $\mathbf{M} \notin \mathcal{SM}_\Omega$, then there is at least one column of $\mathbf{M}$, let it be the A column, where at least one element is negative, or where the sum of the elements is not 1. Then take the bba $\mathbf{1}_A$. It belongs to $\mathcal{M}^\Omega$, and $\mathbf{M} \cdot \mathbf{1}_A$ is the A column of $\mathbf{M}$. This column vector is not a bba as either an element is negative, or their sum is not 1. □

Two special families of stochastic matrices are the specialization and generalization matrices.

7 Specializations and Generalizations

Definition 7.1 *A specialization matrix $\mathbf{S} = [s(A, B)], A, B \subseteq \Omega$ is a stochastic matrix which coefficients $s(A, B) = 0, \ \forall A \not\subseteq B$.*

Definition 7.2 *A generalization matrix $\mathbf{G} = [g(A, B)], A, B \subseteq \Omega$ is a stochastic matrix which coefficients $g(A, B) = 0, \ \forall B \not\subseteq A$.*

$$
\mathbf{S} = \begin{bmatrix} 1 & .3 & .2 & .4 & .4 & .2 & .1 & .1 \\ . & .7 & . & .1 & . & .1 & . & .1 \\ . & . & .8 & .2 & . & . & .3 & .2 \\ . & . & . & .3 & . & . & . & .1 \\ . & . & . & . & .6 & .4 & .3 & .2 \\ . & . & . & . & . & .3 & . & .1 \\ . & . & . & . & . & . & .3 & .1 \\ . & . & . & . & . & . & . & .1 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} .1 & . & . & . & . & . & . & . \\ .1 & .3 & . & . & . & . & . & . \\ .1 & . & .3 & . & . & . & . & . \\ .2 & .3 & .4 & .6 & . & . & . & . \\ .1 & . & . & . & .3 & . & . & . \\ .2 & .3 & . & . & .2 & .8 & . & . \\ .1 & . & .1 & . & .1 & . & .7 & . \\ .1 & .1 & .2 & .4 & .4 & .2 & .3 & 1 \end{bmatrix}
$$

Table 9: Example of a specialization matrix $\mathbf{S}$ (left) and of a generalization matrix $\mathbf{G}$ (right) where $\mathbf{G} = \mathbf{J} \cdot \mathbf{S} \cdot \mathbf{J}$.

These definitions result from the property that if $\mathbf{m}_1 \in \mathcal{M}^\Omega$ and $\mathbf{m}_2 = S \cdot \mathbf{m}_1$, then the masses of $\mathbf{m}_1$ 'flow down' into $\mathbf{m}_2$, by what is meant that for each $A \subseteq \Omega$, the mass $\mathbf{m}_1(A)$ is distributed among the subsets of A when building $\mathbf{m}_2$. An important consequence of this property is that $\mathbf{b}_2 \geq \mathbf{b}_1$ and $\mathbf{q}_2 \leq \mathbf{q}_1$.

The generalization is doing just the reverse, masses 'flow up'. Table 9 presents a specialization matrix and a generalization matrix.

Theorem 7.1 *Let $\mathbf{m}_1 \in \mathcal{M}^\Omega$, $\mathbf{S}$ be a specialization matrix and $\mathbf{G}$ be a generalization matrix. Let $\mathbf{m}_2 = \mathbf{S} \cdot \mathbf{m}_1$ and $\mathbf{m}_3 = \mathbf{G} \cdot \mathbf{m}_1$. Then: $\mathbf{b}_2 \geq \mathbf{b}_1, \mathbf{q}_2 \leq \mathbf{q}_1$ and $\mathbf{b}_3 \leq \mathbf{b}_1, \mathbf{q}_3 \geq \mathbf{q}_1$.*

Proof. This theorem is proved in (Dubois & Prade, 1987). We only prove the first inequality using the matrix notation.

Let $\mathbf{m}_1 \in \mathcal{M}^\Omega$, $\mathbf{S}$ be a specialization matrix and $\mathbf{m}_2 = \mathbf{S} \cdot \mathbf{m}_1$. One has $\mathbf{b}_2 = \mathbf{BfrM} \cdot \mathbf{m}_2 = \mathbf{BfrM} \cdot \mathbf{S} \cdot \mathbf{m}_1$. We must prove $\mathbf{b}_2 \geq \mathbf{b}_1 = \mathbf{BfrM} \cdot \mathbf{m}_1$.

The colums of $\mathbf{S}$ are bba's which masses are given only to the subsets of the column index. Let $\mathbf{m}_B$ be the B column of $\mathbf{S}$. The B column of $\mathbf{BfrM} \cdot \mathbf{S}$ is the implicabilty function $\mathbf{b}_B$ built from the bba's $\mathbf{m}_B$ with the property that $b_B(B) = 1$.

The elements of $\mathbf{BfrM}$ satisfy $BfrM(A, B) = 1$ if $B \subseteq A$ and 0 otherwise (see relation 1). As $b_B(B) = 1$, we also have $b_B(A) = 1$ whenever $B \subseteq A$. Hence for every $X \subseteq \Omega$, we have $b_B(A) \geq BfrM(A, B)$.

In that case, $b_2(A) = \sum_{B \subseteq \Omega} b_B(A) m_1(B) \geq \sum_{B \subseteq \Omega} BfrM(A, B) m_1(B) = b_1(A)$, proving thus the first inequality. The others are proved similarly. □

Given any pair $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}^\Omega$, we can always find a specialization matrix $\mathbf{S}$ and a generalization matrix $\mathbf{G}$ such that $\mathbf{m}_2 = \mathbf{S} \cdot \mathbf{G} \cdot \mathbf{m}_1$. For instance let all columns of $\mathbf{G}$ be made of vector $\mathbf{1}_\Omega$, the Ω column of $\mathbf{S}$ be $\mathbf{m}_2$, and the other columns be the vector $\mathbf{1}_\emptyset$. With $\Omega = \{a, b\}$, it becomes:

$$
\mathbf{m}_2 = \begin{bmatrix} 1 & 1 & 1 & m_2(\emptyset) \\ . & . & . & m_2(a) \\ . & . & . & m_2(b) \\ . & . & . & m_2(\Omega) \end{bmatrix} \cdot \begin{bmatrix} . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \mathbf{m}_1
$$

Similarly for any pair $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}^\Omega$, we can always find a specialization matrix $\mathbf{S}$ and a generalization matrix $\mathbf{G}$ such that $\mathbf{m}_2 = \mathbf{G} \cdot \mathbf{S} \cdot \mathbf{m}_1$. For instance let all columns of $\mathbf{S}$ be made of vector $\mathbf{1}_\emptyset$, and the $\emptyset$ column of $\mathbf{G}$ be $\mathbf{m}_2$, and the other columns be the vector $\mathbf{1}_\Omega$. With $\Omega = \{a, b\}$, it becomes:

$$\mathbf{m}_2 = \begin{bmatrix} m_2(\emptyset) & . & . & . \\ m_2(a) & . & . & . \\ m_2(b) & . & . & . \\ m_2(\Omega) & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \end{bmatrix} \cdot \mathbf{m}_1$$

As we can see, the transformation of any bba can thus be achieved either by a stochastic matrix or by a pair of specialization and generalization matrices.

7.1 Iterations

It is easy to prove that the specialization of a specialization is a specialization, and the same with generalizations. This results from the fact that both matrices are triangular matrices.

Theorem 7.2 *Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be two specialization matrices on Ω, then $\mathbf{S}_1 \cdot \mathbf{S}_2$ is a specialization matrix on Ω.*

Theorem 7.3 *Let $\mathbf{G}_1$ and $\mathbf{G}_2$ be two generalization matrices on Ω, then $\mathbf{G}_1 \cdot \mathbf{G}_2$ is a generalization matrix on Ω.*

Other interesting theorems are listed below.

Theorem 7.4 *If $\mathbf{S}$ is a specialization matrix on Ω, then $\mathbf{J} \cdot \mathbf{S} \cdot \mathbf{J}$ is a generalization matrix on Ω. If $\mathbf{G}$ is a generalization matrix on Ω, then $\mathbf{J} \cdot \mathbf{G} \cdot \mathbf{J}$ is a specialization matrix on Ω.*

Theorem 7.5 *The set of specialization and generalization matrices on Ω are in one-to-one correspondence.*

Theorem 7.6 *The determinants of a specialization and of a generalization matrix are the product of their diagonal terms, respectively.*

Theorem 7.7 *If $\mathbf{S} = \mathbf{J} \cdot \mathbf{G} \cdot \mathbf{J}$ or $\mathbf{G} = \mathbf{J} \cdot \mathbf{S} \cdot \mathbf{J}$, then the determinants of $\mathbf{S}$ and $\mathbf{G}$ are equal.*

Sketches of the proofs are as follows. Theorem 7.4 results from the double mirror inversion that corresponds to the double $\mathbf{J}$ operation. Theorem 7.5 results from the previous one as $\mathbf{J}$ is not singular. Theorem 7.6 is a property of any triangular matrix. Theorem 7.7 results from the fact that the determinant of $\mathbf{J}$ is 1.

8 Revision

The revision of a bba $\mathbf{m}_1$ by a new piece of evidence Ev can always be represented by a stochastic matrix $\mathbf{M}(Ev, \mathbf{m}_1)$ that transforms $\mathbf{m}_1$ into $\mathbf{m}_1[Ev]$:

$$\mathbf{m}_1[Ev] = \mathbf{M}(Ev, \mathbf{m}_1) \cdot \mathbf{m}_1.$$

If the value of the matrix depend only on Ev and not on $\mathbf{m}_1$, we can write:

$$\mathbf{m}_1[Ev] = \mathbf{M}(Ev) \cdot \mathbf{m}_1.$$

In that case we say that the piece of evidence that induced $\mathbf{m}_1$ and Ev are 'distinct'pieces of evidence. We think this is indeed what was meant by the 'distinctness' requirement encountered in belief function theory.

In some cases $\mathbf{M}(Ev)$ is a specialization or a generalization matrix.

When the revision is a specialization (generalization) we call it a conjunctive (disjunctive) revision as beliefs become more concentrated, (more diffused). The terms come from the analogy with the AND that create 'smaller' sets and the OR that create 'larger' sets.

We say that $\mathbf{m}_2$ is a specialization of $\mathbf{m}_1$ if there exists a specialization matrix $\mathbf{S}$ so that $\mathbf{m}_2 = \mathbf{S} \cdot \mathbf{m}_1$.

We say that $\mathbf{m}_2$ is a generalization of $\mathbf{m}_1$ if there exists a generalization matrix $\mathbf{G}$ so that $\mathbf{m}_2 = \mathbf{G} \cdot \mathbf{m}_1$.

9 Dempster's Rule of Conditioning

Suppose a bba $\mathbf{m}_1$ and let $\mathcal{S}(\mathbf{m}_1)$ be the set of bba that are specializations of $\mathbf{m}_1$: $\mathcal{S}(\mathbf{m}_1) = \{\mathbf{m} : \exists \text{ a specialization } \mathbf{S}, \mathbf{m} = \mathbf{S} \cdot \mathbf{m}_1\}$. In that set consider the bba $\mathbf{m}_2$ that are specializations of $\mathbf{m}_1$ such that $pl_2(\overline{A}) = 0$, and among them select the least committed element $\mathbf{m}^*$. Thus $\mathbf{b}^* \leq \mathbf{b}, \ \forall \mathbf{m} \in \mathcal{S}(\mathbf{m}_1)$ with $pl(\overline{A}) = 0$.

The solution $\mathbf{m}^*$ is unique and well known (Klawonn & Smets, 1992): it is what is produced by Dempster's rule of conditioning (without the normalization) when conditioning on A. We have:

$$m^*(B) = m_1[A](B) = \sum_{C \subseteq \overline{A}} m_1(B \cup C), \ \forall B \subseteq A.$$

The other masses are 0. Note that we write the conditioning event between [and] (for any of $\mathbf{m}, \mathbf{b}, \mathbf{q}, \mathbf{bel}, \mathbf{pl}$). This helps as $\mathbf{m}[X]$ is thus a vector, whereas with classical notation we would have had to write $\mathbf{m}(.|X)$ or $\mathbf{m}(|X)$, what we feel inaesthetical and sometimes confusing.

The corresponding specialization matrix is given for $\Omega = \{a, b, c\}$ and $A = \{a, b\}$ by
$$\begin{bmatrix} 1 & . & . & . & 1 & . & . & . \\ . & 1 & . & . & . & 1 & . & . \\ . & . & 1 & . & . & . & 1 & . \\ . & . & . & 1 & . & . & . & 1 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \end{bmatrix}$$

This shows that the revision achieved by Dempster's rule of conditioning corresponds to generating the least committed belief function that is a specialization of the initial bba and so that the plausibility of $\overline{A}$ becomes 0. These are exactly what conditioning on A is about, $\overline{A}$ being impossible, its plausibility must become 0, and the conditioning event being all we know we select the least committed solution.

The dual of this 'conditioning' operator is obtained by the next generalization matrix $\begin{bmatrix} . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 1 & . & . & . & 1 & . & . & . \\ . & 1 & . & . & . & 1 & . & . \\ . & . & 1 & . & . & . & 1 & . \\ . & . & . & 1 & . & . & . & 1 \end{bmatrix}$

This corresponds to a form of deconditioning on $\{a, b\}$. The masses $m_1(X)$ for $X \subseteq \{a, b\}$ are transferred to $X \cup \{c\}$. This is what we get if we start by conditioning on $\bar{c}$, then decide to backtrack, to cancel the conditioning revision, but all we can do it to take the least committed bba among all the generalizations of $\mathbf{m}_1$ that satisfy $pl_2(c) = 1$. This deconditioning process is not often encountered in probability theory as in that context, it cannot be realized without introducing some further assumptions.

Some properties are easily expresses and proved. Let $\mathbf{C}_A$ ($\mathbf{D}_A$) be the matrix representing the conditioning (deconditioning) on A :

- $\mathbf{C}_A \cdot \mathbf{C}_B = \mathbf{C}_{A \cap B}$
- $\mathbf{C}_A \cdot \mathbf{C}_A = \mathbf{C}_A$
- $\mathbf{D}_A \cdot \mathbf{D}_B = \mathbf{D}_{A \cup B}$
- $\mathbf{D}_A \cdot \mathbf{D}_A = \mathbf{D}_A$
- $\mathbf{C}_A \cdot \mathbf{D}_A = \mathbf{C}_A$
- $\mathbf{D}_A \cdot \mathbf{C}_A = \mathbf{D}_A$
- $\mathbf{C}_A \cdot \mathbf{D}_A \cdot \mathbf{C}_A = \mathbf{C}_A$
- $\mathbf{D}_A \cdot \mathbf{C}_A \cdot \mathbf{D}_A = \mathbf{D}_A$

The last two equalities show that $\mathbf{C}_A$ and $\mathbf{D}_A$ are each other generalized inverses.

10 Conjunctive and Disjunctive Rules of Combination

Conjunctive revision of a bba by a distinct piece of evidence is achieved by a specialization matrix. Suppose we have several pieces of evidence and each is represented by a specialization matrix. A really conjunctive combination rule should satisfy commutativity and associativity requirements, the order under which the revisions are applied does not change the results.

The family of commutative and associative specialization matrices that contains the set of conditioning specialization is called the set of Dempsterian specialization matrix. In (Klawonn & Smets, 1992), we derive its structure.

Let $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}^\Omega$, and let the specialization matrix $\mathbf{S}_{\mathbf{m}_2}$ be such that $s_{\mathbf{m}_2}(A, B) = m_2[B](A)$. Then $\mathbf{S}_{\mathbf{m}_2} \cdot \mathbf{m}_1 = \mathbf{m}_2 \textcircled{\cap} \mathbf{m}_1$ hence what is obtained by conjunctive combination rule (which is equal to Dempster's rule of combination except for the normalization).

$$\mathbf{S} = \begin{bmatrix} 1.0 & .70 & .50 & .37 & .50 & .30 & .15 & .10 \\ . & .30 & . & .13 & . & .20 & . & .05 \\ . & . & .50 & .33 & . & . & .35 & .20 \\ . & . & . & .17 & . & . & . & .15 \\ . & . & . & . & .50 & .40 & .35 & .27 \\ . & . & . & . & . & .10 & . & .08 \\ . & . & . & . & . & . & .15 & .13 \\ . & . & . & . & . & . & . & .02 \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} .10 & . & . & . & . & . & . & . \\ .05 & .15 & . & . & . & . & . & . \\ .20 & . & .30 & . & . & . & . & . \\ .15 & .35 & .20 & .50 & . & . & . & . \\ .27 & . & . & . & .37 & . & . & . \\ .08 & .35 & . & . & .13 & .50 & . & . \\ .13 & . & .40 & . & .33 & . & .70 & . \\ .02 & .15 & .10 & .50 & .17 & .50 & .30 & 1.0 \end{bmatrix}$$

Table 10: Example of a dempsterian specialization matrix **S** and a dempsterian generalization matrix **G**.

$$\mathbf{S} = \mathbf{MfrQ} \cdot \begin{bmatrix} 1.0 & . & . & . & . & . & . & . \\ . & .30 & . & . & . & . & . & . \\ . & . & .50 & . & . & . & . & . \\ . & . & . & .17 & . & . & . & . \\ . & . & . & . & .50 & . & . & . \\ . & . & . & . & . & .10 & . & . \\ . & . & . & . & . & . & .15 & . \\ . & . & . & . & . & . & . & .02 \end{bmatrix} \cdot \mathbf{QfrM}$$

Table 11: The eigenvalues-eigenvectors decomposition of **S** of Table 10.

So we can justify the conjunctive combination rule from the commutativity and associativity requirements, assuming that conditioning is a special form of conjunctive combination rule.

Similar results are derived for the disjunctive combination rule: we just replace specialization by generalization, and conditioning by deconditioning.

Table 10 presents the dempsterian specialization and generalization matrices built from $\mathbf{m} \in \mathcal{M}_{\{a,b,c\}}$, where $\mathbf{m}$ is the Ω column of **S** and the $\emptyset$ column of **G**.

10.1 Eigenvalues and Eigenvectors

Let $\mathbf{S}_m$ be the dempsterian specialization matrix generated by $\mathbf{m} \in \mathcal{M}^{\Omega}$. The matrix with the eigenvectors of $\mathbf{S}_\mathbf{m}$ is the **QfrM** matrix (that does not depend on m), and the eigenvalues of $\mathbf{S}_\mathbf{m}$ are the elements of the commonality function q related to m (Klawonn & Smets, 1992). The q values happens also to be the elements of the diagonal of $\mathbf{S}_\mathbf{m}$. This results from the equality $q(A) = m[A](A)$. This of course explains the importance of the commonality functions when conjunctive combination rules are involved. Table 11 presents the decomposition for the **S** matrix of Table 10.

$$\mathbf{G} = \mathbf{MfrB} \cdot \begin{bmatrix} .10 & . & . & . & . & . & . & . \\ . & .15 & . & . & . & . & . & . \\ . & . & .30 & . & . & . & . & . \\ . & . & . & .50 & . & . & . & . \\ . & . & . & . & .37 & . & . & . \\ . & . & . & . & . & .50 & . & . \\ . & . & . & . & . & . & .70 & . \\ . & . & . & . & . & . & . & 1.0 \end{bmatrix} \cdot \mathbf{BfrM}$$

Table 12: The eigenvalues-eigenvectors decomposition of **G** in Table 10.

The matrix representation of the conjunctive combination rule becomes:

$$\begin{aligned} \mathbf{q}_{1\textcircled{\cap}2} &= \mathbf{QfrM} \cdot \mathbf{m}_{1\textcircled{\cap}2} \\ &= \mathbf{QfrM} \cdot \mathbf{S}_{\mathbf{m}_1} \cdot \mathbf{m}_2 \\ &= \mathbf{QfrM} \cdot \mathbf{MfrQ} \cdot \mathbf{Diag}(\mathbf{q}_1) \cdot \mathbf{QfrM} \cdot \mathbf{m}_2 \\ &= \mathbf{Diag}(\mathbf{q}_1) \cdot \mathbf{q}_2 \end{aligned}$$

The same holds with the dempsterian generalization matrix generated by $\mathbf{m} \in \mathcal{M}^\Omega$. The matrix with the eigenvectors of $\mathbf{G_m}$ is the **BfrM** matrix, and the eigenvalues of $\mathbf{G}_m$ are the elements of the implicability functions **b** related to **m** (Klawonn & Smets, 1992). The b values happens to be the elements of the diagonal of $\mathbf{G_m}$. This explains the importance of the implicability functions when disjunctive combination rules are involved. Table 12 presents the decomposition for the **G** matrix of Table 10.

The matrix representation of the disjunctive combination rule becomes:

$$\begin{aligned} \mathbf{b}_{1\textcircled{\cup}2} &= \mathbf{BfrM} \cdot \mathbf{m}_{1\textcircled{\cup}2} \\ &= \mathbf{BfrM} \cdot \mathbf{G}_{\mathbf{m}_1} \cdot \mathbf{m}_2 \\ &= \mathbf{BfrM} \cdot \mathbf{MfrB} \cdot \mathbf{Diag}(\mathbf{b}_1) \cdot \mathbf{BfrM} \cdot \mathbf{m}_2 \\ &= \mathbf{Diag}(\mathbf{b}_1) \cdot \mathbf{b}_2 \end{aligned}$$

10.2 De Morgan Algebra

Theorem 10.1 *Let* $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}^\Omega$. *Then:*

$$\overline{\mathbf{m}_1 \textcircled{\cap} \mathbf{m}_2} = \overline{\mathbf{m}_1} \textcircled{\cup} \overline{\mathbf{m}_2}$$

and

$$\overline{\mathbf{m}_1 \textcircled{\cup} \mathbf{m}_2} = \overline{\mathbf{m}_1} \textcircled{\cap} \overline{\mathbf{m}_2},$$

where the overline indicates the negation operator.

Proof.

$$
\begin{aligned}
\overline{\mathbf{m}_1 \circledcap \mathbf{m}_2} &= \mathbf{J} \cdot \mathbf{m}_1 \circledcap \mathbf{m}_2 \\
&= \mathbf{J} \cdot \mathbf{S}_{\mathbf{m}_1} \cdot \mathbf{m}_2 \\
&= \mathbf{J} \cdot \mathbf{MfrQ} \cdot \mathbf{Diag}(\mathbf{q}_1) \cdot \mathbf{QfrM} \cdot \mathbf{m}_2 \\
&= \mathbf{J} \cdot \mathbf{MfrQ} \cdot \mathbf{J} \cdot \mathbf{J} \cdot \mathbf{Diag}(\mathbf{q}_1) \cdot \mathbf{J} \cdot \mathbf{J} \cdot \mathbf{q}_2 \\
&= \mathbf{MfrB} \cdot \mathbf{Diag}(\overline{\mathbf{b}_1}) \cdot \overline{\mathbf{b}_2} \\
&= \mathbf{MfrB} \cdot \overline{\mathbf{b}_1} \circledcup \overline{\mathbf{b}_2} \\
&= \overline{\mathbf{m}_1} \circledcup \overline{\mathbf{m}_2}
\end{aligned}
$$

as $\mathbf{J} \cdot \mathbf{Diag}(\mathbf{q}_1) \cdot \mathbf{J} = \mathbf{Diag}(\overline{\mathbf{b}_1})$, $\mathbf{J} \cdot \mathbf{q}_2 = \overline{\mathbf{b}_2}$, and $\mathbf{J} \cdot \mathbf{MfrQ} \cdot \mathbf{J} = \mathbf{MfrB}$. The second part is proved similarly. □

The use of the matrix notation really simplifies the proof.

11 Canonical Representations

A simple support function is a belief function which masses are all 0 except for one subset of Ω and for Ω itself, the last two masses being positive and adding to 1. The non Ω set is called the focal set of the simple support function. A simple support function with focal set $X \subseteq \Omega$ and weight w_X which is the mass given to Ω is denoted as X^{w_X}.

A separable bba is defined as a bba that can be represented as the result of the conjunctive combination of simple support functions.

A non-dogmatic bba is a bba so that $m(\Omega) > 0$.

The inverse of the $\circledcap$ operator, denoted $\overline{\circledcap}$ is defined for non-dogmatic bba $\mathbf{m}_2$ as follows:

$$q_{1\overline{\circledcap}2}(X) = q_1(X)/q_2(X), \ \forall X \subseteq \Omega$$

In (Smets, 1995), we prove that any non-dogmatic bba can be represented as follows:

$$\mathbf{m} = \mathbf{m}_C \overline{\circledcap} \mathbf{m}_D$$

where $\mathbf{m}_C$ and $\mathbf{m}_D$ are separable bba and their underlying simple support functions are defined on different focal sets:

$$\mathbf{m}_C = \bigcirc\!\!\!\!\cap_{X \subseteq \Omega_+} X^{w_X}$$

$$\mathbf{m}_D = \bigcirc\!\!\!\!\cap_{X \subseteq \Omega_-} X^{w_X}$$

and $\Omega_+, \Omega_- \subseteq 2^\Omega$ with $\Omega_+ \cap \Omega_- = \emptyset$.

The algorithm to compute the w_X happens to be a Möbius transform applied to the logarithm of the commonality function. Let $lq(X) = \log(q(X)), \ \forall X \subseteq \Omega$. As the bba is non-dogmatic, $\mathbf{q} > 0$, and thus $\mathbf{lq}$ is well defined. Then the vector $\mathbf{lw}$ which elements are the logarithm of the w_X is given by:

$$\mathbf{lw} = -\mathbf{MfrQ} \cdot \mathbf{lq}.$$

They can thus be computed using the FMT.

The theory can be extended to dogmatic bba, but it requires subtleties out of focus here. The idea is to put an epsilon on Ω, to work with it and to take limits for epsilon going to 0.

Similar results are obtained with disjunctive decompositions, and even with α-junctive decompositions.

12 The α-junctions

In (Smets, 1997), we study the set of possible linear fusion - aggregation operators.

Let $\mathbf{m}_1$ and $\mathbf{m}_2$ be two bba on Ω. We want to build a bba $\mathbf{m}_{12}$ such that $\mathbf{m}_{12} = f(\mathbf{m}_1, \mathbf{m}_2)$, thus $\mathbf{m}_{12}$ depends only of $\mathbf{m}_1$ and $\mathbf{m}_2$. Thus we want to determine what are the operators that map $\mathcal{M}^\Omega \times \mathcal{M}^\Omega$ to $\mathcal{M}^\Omega$ and that satisfy the next requirements.

1. Linearity: $f(\mathbf{m},\ p\mathbf{m}_1 + q\mathbf{m}_2) = pf(\mathbf{m}, \mathbf{m}_1) + qf(\mathbf{m}, \mathbf{m}_2),\ p \in [0,1], q = 1 - p$
2. Commutativity: $f(\mathbf{m}_1, \mathbf{m}_2) = f(\mathbf{m}_2, \mathbf{m}_1)$.
3. Associativity: $f(f(\mathbf{m}_1, \mathbf{m}_2), \mathbf{m}_3) = f(\mathbf{m}_1, f(\mathbf{m}_2, \mathbf{m}_3))$.
4. Existence of a belief function $\mathbf{m}_{vac}$ such that $f(\mathbf{m}, \mathbf{m}_{vac}) = \mathbf{m}$ for any $\mathbf{m}$.
5. Anonymity: relabeling the elements of Ω does not affect the results.
6. Context preservation: if $pl_1(X) = pl_2(X) = 0$ for some $X \subseteq \Omega$, then $pl_{12}(X) = 0$.

The origin of these requirements is as follows.

1. Linearity: with probability p, You are in context C_1, in which case Your belief is represented by m_1, and with probability $q = 1 - p$, You are in context C_2, in which case Your belief is represented by m_2. Your belief before knowing the context You are in is $p\mathbf{m}_1 + q\mathbf{m}_2$ (proved in (Smets & Kennes, 1994)). You can then combine this belief with m. But You can also consider that with probability p (q) You are in context C_1 (C_2) in which case the result of the combination is $f(\mathbf{m}, \mathbf{m}_1)$, ($f(\mathbf{m}, \mathbf{m}_2)$), and You can then take their weighted average. The two results should be the same. The consequence is that we will end up with matrices.
2. Commutativity: the order of the combination is irrelevant.
3. Associativity: the order under which bba's are combined is irrelevant. Thus the family of acceptable matrices is strongly limited.
4. The belief function $\mathbf{m}_{vac}$ is a bba which combination with any other bba leaves Your beliefs unchanged. We thus have a zero element.
5. Anonymity: the results do not depend on the label given to the elements. Hence combination and permutation commute.
6. Context preservation: if X is not plausible for both bba's, X remains not plausible after their combination. This implies many zeros in the matrices.

The solutions are stochastic matrices and the operation is called an α-junction (as, among others, it will cover both conjunctions and disjunctions). We have (for proofs; see (Smets, 1997)):

$$\mathbf{m}_{12} = \mathbf{K}(\mathbf{m}_1) \cdot \mathbf{m}_2$$

where

$$\mathbf{K}(\mathbf{m}_1) = \sum_{X \subseteq \Omega} m_1(X) \cdot \mathbf{K}_X$$

.

The structure of the $2^{|\Omega|} \times 2^{|\Omega|}$ matrices $\mathbf{K}_X$ depends on $\mathbf{m}_{vac}$ and of one parameter $\alpha \in [0,1]$. We prove that there are only two solutions for $\mathbf{m}_{vac}$: either $\mathbf{m}_{vac} = \mathbf{1}_\Omega$ or $\mathbf{m}_{vac} = \mathbf{1}_\emptyset$. So there are only two sets of solutions, which will even satisfy De Morgan's laws.

The eigenvalues - eigenvectors decomposition of $\mathbf{K}_X$ is given by $\mathbf{K}_X = \mathbf{G}^{-1} \cdot \mathbf{V}_X \cdot \mathbf{G}$ where interestingly the $\mathbf{G}$ matrix does not depend on X. If we define $\mathbf{g} = \mathbf{G} \cdot \mathbf{m}$, then $g_{12}(X) = g_1(X) g_2(X)$ for all $X \subseteq \Omega$. This is the analogous of the pointwise product rule used with the commonality functions and the implicability functions to compute the conjunctive combination rule and the disjunctive combination rule, respectively.

We present the two sets of solutions.

12.1 The Conjunctive Case: $\mathbf{m}_{vac} = \mathbf{1}_\Omega$

The only matrices that satisfy all the above requirements with $\mathbf{m}_{vac} = \mathbf{1}_\Omega$ are given below, where $\alpha \in [0,1]$ and is constant for all $\mathbf{K}_X$ and $\mathbf{V}_X$. Their formal definitions are quite laborious, so we present them in the example 12.1. Their patterns are in fact quite simpler than what the equations might lead to think.

$$\begin{array}{lll} \mathbf{K}_\Omega = \mathbf{I} & \mathbf{V}_\Omega = \mathbf{I} & \\ \mathbf{K}_{\{\overline{x}\}} = [k_{\overline{x}}(A,B)] & \mathbf{V}_{\{\overline{x}\}} = [v_{\overline{x}}(A,B)] & \forall x \in \Omega \\ \mathbf{K}_X = \prod_{x \notin X} \mathbf{K}_{\{\overline{x}\}} & \mathbf{V}_X = \prod_{x \notin X} \mathbf{V}_{\{\overline{x}\}} & \forall X \subseteq \Omega \end{array}$$

where

$$\begin{array}{ll} k_{\overline{x}}(A,B) & v_{\overline{x}}(A,B) \\ = 1 \text{ if } x \notin A, B = A \cup \{x\} & = 1 \text{ if } x \notin A, A = B \\ = \alpha \text{ if } x \notin B, B = A & = \alpha - 1 \text{ if } x \in A, A = B \\ = 1 - \alpha \text{ if } x \notin B, A = B \cup \{x\} & = 0 \text{ if } A \neq B \\ = 0 \text{ otherwise} & \end{array}$$

The X column of the $2^{|\Omega|} \times 2^{|\Omega|}$ $\mathbf{G}$ matrix is $\mathbf{V}_X \cdot \mathbf{1}$.

We denote by $\mathbf{K}^{\cap,\alpha}(\mathbf{m})$ the matrix $\mathbf{K}(\mathbf{m})$ computed above. When $\alpha = 1$, $\mathbf{K}^{\cap,1}(\mathbf{m})$ become the dempsterian specialization matrix and $\mathbf{K}^{\cap,1}(\mathbf{m}_1) \cdot \mathbf{m}_2 = \mathbf{m}_1 \textcircled{\cap} \mathbf{m}_2$. This is why we use the $\cap$ index.

The case $\alpha = 0$ corresponds to the combination rule

$$m_{12}(Z) = \sum_{Z = (X \cap Y) \cup (\overline{X} \cap \overline{Y})} m_1(X) m_2(Y).$$

This would be the combination rule to be used when combining two distinct pieces of evidence such that all You know is that either both are reliable or none are reliable.

Example 12.1. **:** We present the various matrices when $\Omega = \{a, b\}$ where $\overline{\alpha} = 1 - \alpha$.

$$\mathbf{K}_{\{\overline{a}\}} = \begin{bmatrix} \alpha & 1 & . & . \\ \overline{\alpha} & . & . & . \\ . & . & \alpha & 1 \\ . & . & \overline{\alpha} & . \end{bmatrix}, \mathbf{K}_{\{\overline{b}\}} = \begin{bmatrix} \alpha & . & 1 & . \\ . & \alpha & . & 1 \\ \overline{\alpha} & . & . & . \\ . & \overline{\alpha} & . & . \end{bmatrix}, \mathbf{K}_{\emptyset} = \begin{bmatrix} \alpha^2 & \alpha & \alpha & 1 \\ \alpha\overline{\alpha} & . & \overline{\alpha} & . \\ \alpha\overline{\alpha} & \overline{\alpha} & . & . \\ \overline{\alpha}^2 & . & . & . \end{bmatrix}$$

$$\mathbf{V}_{\{\overline{a}\}} = \begin{bmatrix} 1 & . & . & . \\ . & -\overline{\alpha} & . & . \\ . & . & 1 & . \\ . & . & . & -\overline{\alpha} \end{bmatrix}, \mathbf{V}_{\{\overline{b}\}} = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & . & -\overline{\alpha} & . \\ . & . & . & -\overline{\alpha} \end{bmatrix}$$

$$\mathbf{V}_{\emptyset} = \begin{bmatrix} 1 & . & . & . \\ . & -\overline{\alpha} & . & . \\ . & . & -\overline{\alpha} & . \\ . & . & . & \overline{\alpha}^2 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -\overline{\alpha} & 1 & -\overline{\alpha} & 1 \\ -\overline{\alpha} & -\overline{\alpha} & 1 & 1 \\ \overline{\alpha}^2 & -\overline{\alpha} & -\overline{\alpha} & 1 \end{bmatrix}$$

In particular we have:

$$\mathbf{m}_{12} = m_1(\emptyset) \begin{bmatrix} \alpha^2 & \alpha & \alpha & 1 \\ \alpha\overline{\alpha} & . & \overline{\alpha} & . \\ \alpha\overline{\alpha} & \overline{\alpha} & . & . \\ \overline{\alpha}^2 & . & . & . \end{bmatrix} \mathbf{m}_2 + m_1(a) \begin{bmatrix} \alpha & . & 1 & . \\ . & \alpha & . & 1 \\ \overline{\alpha} & . & . & . \\ . & \overline{\alpha} & . & . \end{bmatrix} \mathbf{m}_2$$

$$+ m_1(b) \begin{bmatrix} \alpha & 1 & . & . \\ \overline{\alpha} & . & . & . \\ . & . & \alpha & 1 \\ . & . & \overline{\alpha} & . \end{bmatrix} \mathbf{m}_2 + m_1(\Omega)\mathbf{I} \cdot \mathbf{m}_2$$

12.2 The Disjunctive Case: $\mathbf{m}_{vac} = \mathbf{1}_{\emptyset}$

The only matrices that satisfy all the above requirements with $\mathbf{m}_{vac} = \mathbf{1}_{\emptyset}$ are given below, where $\alpha \in [0, 1]$ and is constant for all $\mathbf{K}_X$ and $\mathbf{V}_X$.

$$\begin{aligned} \mathbf{K}_{\emptyset} &= \mathbf{I} & \mathbf{V}_{\emptyset} &= \mathbf{I} & & \\ \mathbf{K}_{\{x\}} &= [k_x(A, B)] & \mathbf{V}_{\{x\}} &= [v_x(A, B)] & & \forall x \in \Omega \\ \mathbf{K}_X &= \prod_{x \in X} \mathbf{K}_{\{x\}} & \mathbf{V}_X &= \prod_{x \in X} \mathbf{V}_{\{x\}} & & \forall X \subseteq \Omega \end{aligned}$$

where

$$\begin{array}{ll} k_x(A,B) & v_x(A,B) \\ = 1 \text{ if } x \notin B, A = B \cup \{x\} & = 1 \text{ if } x \notin A, A = B \\ = \alpha \text{ if } x \in B, B = A & = \alpha - 1 \text{ if } x \in A, A = B \\ = 1-\alpha \text{ if } x \notin A, B = A \cup \{x\} & = 0 \text{ if } A \neq B \\ = 0 \text{ otherwise} & \end{array}$$

The X column of $\mathbf{G}$ is $\mathbf{V}_X \cdot \mathbf{1}$.

We denote by $\mathbf{K}^{\cup,\alpha}(\mathbf{m})$ the matrix $\mathbf{K}(m)$ computed above. When $\alpha = 1$, $\mathbf{K}^{\cup,1}(\mathbf{m})$ become the dempsterian generalization matrix and $\mathbf{K}^{\cup,1}(\mathbf{m}_1) \cdot \mathbf{m}_2 = \mathbf{m}_1 \circledcirc \mathbf{m}_2$. This is why we use the $\cup$ index.

The case $\alpha = 0$ corresponds to the combination rule

$$\mathbf{m}_{12}(Z) = \sum_{Z = X \underline{\cup} Y} \mathbf{m}_1(X)\mathbf{m}_2(Y)$$

where $\underline{\cup}$ is the exclusive OR. This would be the combination rule to be used when combining two distinct pieces of evidence such that all You know is that one is reliable and the other is not, but You don't know which one is the reliable one.

Example 12.2. **:** We present the various matrices when $\Omega = \{a, b\}$ where $\overline{\alpha} = 1 - \alpha$.

$$\mathbf{K}_{\{a\}} = \begin{bmatrix} . & \overline{\alpha} & . & . \\ 1 & \alpha & . & . \\ . & . & . & \overline{\alpha} \\ . & . & 1 & \alpha \end{bmatrix}, \mathbf{K}_{\{b\}} = \begin{bmatrix} . & . & \overline{\alpha} & . \\ . & . & . & \overline{\alpha} \\ 1 & . & \alpha & . \\ . & 1 & . & \alpha \end{bmatrix}, \mathbf{K}_{\Omega} = \begin{bmatrix} . & . & . & \overline{\alpha}^2 \\ . & . & \overline{\alpha} & \alpha\overline{\alpha} \\ . & \overline{\alpha} & . & \alpha\overline{\alpha} \\ 1 & \alpha & \alpha & \alpha^2 \end{bmatrix}$$

$$\mathbf{V}_{\{a\}} = \begin{bmatrix} 1 & . & . & . \\ . & -\overline{\alpha} & . & . \\ . & . & 1 & . \\ . & . & . & -\overline{\alpha} \end{bmatrix}, \mathbf{V}_{\{b\}} = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & . & -\overline{\alpha} & . \\ . & . & . & -\overline{\alpha} \end{bmatrix},$$

$$\mathbf{V}_{\Omega} = \begin{bmatrix} 1 & . & . & . \\ . & -\overline{\alpha} & . & . \\ . & . & -\overline{\alpha} & . \\ . & . & . & \overline{\alpha}^2 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\overline{\alpha} & 1 & -\overline{\alpha} \\ 1 & 1 & -\overline{\alpha} & -\overline{\alpha} \\ 1 & -\overline{\alpha} & -\overline{\alpha} & \overline{\alpha}^2 \end{bmatrix}$$

12.3 De Morgan Laws

The De Morgan laws can be applied to the α-junction operators, where $\mathbf{J}$ plays the role of the negation.

Theorem 12.1 *For* $\mathbf{m} \in \mathcal{M}^{\Omega}$, $\alpha \in [0, 1]$,

$$\mathbf{K}^{\cap,\alpha}(\mathbf{m}) = \mathbf{J} \cdot \mathbf{K}^{\cup,\alpha}(\mathbf{J} \cdot \mathbf{m}) \cdot \mathbf{J}.$$

Theorem 12.2 *Suppose* $\mathbf{m}_i = \mathbf{m}[Ev_i]$, *for* $i = 1, 2$, $\mathbf{m}[Ev_i \cap_\alpha Ev_2] = \mathbf{K}^{\cap,\alpha}(\mathbf{m}_1) \cdot \mathbf{m}_2$, *and* $\mathbf{m}[Ev_i \cup_\alpha Ev_2] = \mathbf{K}^{\cup,\alpha}(\mathbf{m}_1) \cdot \mathbf{m}_2$, *then*

$$\mathbf{m}[\neg(Ev_1 \cap_\alpha Ev_2)] = \mathbf{m}[\neg Ev_1 \cup_\alpha \neg Ev_2]$$

and

$$\mathbf{m}[\neg(Ev_1 \cup_\alpha Ev_2)] = \mathbf{m}[\neg Ev_1 \cap_\alpha \neg Ev_2]$$

Proof.

$$\begin{aligned}\mathbf{m}[\neg(Ev_1 \cap_\alpha Ev_2)] &= \mathbf{J} \cdot \mathbf{m}[Ev_1 \cap_\alpha Ev_2] \\ &= \mathbf{J} \cdot \mathbf{K}^{\cap,\alpha}(\mathbf{m}_1) \cdot \mathbf{m}_2 \\ &= \mathbf{J} \cdot \mathbf{J} \cdot \mathbf{K}^{\cup,\alpha}(\mathbf{J} \cdot \mathbf{m}_1) \cdot \mathbf{J} \cdot \mathbf{m}_2 \\ &= \mathbf{m}[\neg Ev_1 \cup_\alpha \neg Ev_2]\end{aligned}$$

The second property is proved similarly.

□

This theorem enhances the De Morgan duality between negated pieces of evidence and the conjunctive and disjunctive forms of the α-junctions.

12.4 Interpretation

The cases $\mathbf{K}^{\cup,\alpha}$ and $\mathbf{K}^{\cap,\alpha}$ for $\alpha = 0$ or 1 have a meaning as already explained. The practical meaning of these operators for the other α values in unclear.

13 Conclusions

We have presented how to handle the belief functions computations using matrix notations. As usual, this method greatly simplifies representation and computation. The order we use to represent basic belief assignment as vectors is in fact very important as it really simplifies matters.

Using this matrix notation, we show how classical relations described in belief function theory and in the transferable belief model are represented, that is to say, the Möbius transforms, the Fast Möbius transform, the specialization and generalization, the conditioning, the conjunctive combination rule and disjunctive combination rule, the pignistic transformation, the canonical decomposition, the α-junctions. We think researchers in belief functions will be helped by this notation and by the theorems presented here under this highly efficient form.

References

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

Dubois, D., & Prade, H. (1987). The principle of minimum specificity as a basis for evidential reasoning. In B. Bouchon & R. R. Yager (Eds.), *Uncertainty in knowledge-based systems* (pp. 75–84). Springer Verlag, Berlin.

Grabisch, M. (1996). k-order additive discrete fuzzy measures. In IPMU-96 (Ed.), *Information processing and management of uncertainty* (pp. 1345–1350).

Hsia, Y. T. (1991). Characterizing belief with minimum commitment. In IJCAI-91 (Ed.), *Int. joint conf. on artificial intelligence* (pp. 1184–1189). Morgan Kaufman, San Mateo, Ca.

Kennes, R. (1992). Computational aspects of the Möbius transform of a graph. *IEEE-SMC*, *22*, 201–223.

Kennes, R., & Smets, P. (1990). Fast algorithms for Dempster-Shafer theory. In B. Bouchon-Meunier, R. R. Yager, & L. A. Zadeh (Eds.), *Uncertainty in knowledge bases* (pp. 14–23). Springer Verlag, Berlin.

Kennes, R., & Smets, P. (1991). Computational aspects of the Möbius transform. In P. P. Bonissone, M. Henrion, L. N. Kanal, & J. F. Lemmer (Eds.), *Uncertainty in artificial intelligence 6* (pp. 401–416). North Holland, Amsterdam.

Klawonn, F., & Smets, P. (1992). The dynamic of belief in the transferable belief model and specialization- generalization matrices. In D. Dubois, M. P. Wellman, B. D'Ambrosio, & P. Smets (Eds.), *Uncertainty in artificial intelligence 92* (pp. 130–137). Morgan Kaufman, San Mateo, Ca.

Monney, P. (2001). Dempster specialization matrices and the combination of belief functions. In S. Benferhat & P. Besnard (Eds.), *Symbolic and quantitative approaches to reasoning with uncertainty* (pp. 316–327). Springer-Verlag, Berlin.

Shafer, G. (1976). *A mathematical theory of evidence.* Princeton Univ. Press. Princeton, NJ.

Smets, P. (1988). Belief functions. In P. Smets, E. H. Mamdani, D. Dubois, & H. Prade (Eds.), *Non standard logics for automated reasoning* (pp. 253–286). Academic Press, London.

Smets, P. (1995). The canonical decomposition of a weighted belief. In IJCAI-95 (Ed.), *Int. joint conf. on artificial intelligence* (pp. 1896–1901). Morgan Kaufman, San Mateo, Ca.

Smets, P. (1997). The alpha-junctions: combination operators applicable to belief functions. In D. Gabbay, R. Kruse, A. Nonnengart, & H. Ohlbach (Eds.), *Qualitative and quantitative practical reasoning* (p. 131-153). Springer.

Smets, P. (1998). The transferable belief model for quantified belief representation. In D. M. Gabbay & P. Smets (Eds.), *Handbook of defeasible reasoning and uncertainty management systems* (Vol. 1, pp. 267–301). Kluwer, Doordrecht, The Netherlands.

Smets, P. (2002). Decision making in a context where uncertainty is represented by belief functions. In R. P. Srivastava (Ed.), *Belief functions in business decisions.* Physica-Verlag, Forthcoming.

Smets, P., & Kennes, R. (1994). The transferable belief model. *Artificial Intelligence*, *66*, 191–234.

Smets, P., & Kruse, R. (1997). The transferable belief model for quantified belief representation. In A. Motro & P. Smets (Eds.), *Uncertainty in information systems: From needs to solutions* (pp. 343–368). Boston, MA.: Kluwer.

Yager, R. (1986). The entailment principle for Dempster-Shafer granules. *Int. J. Intell. Systems*, *1*, 247–262.