



**HAL**  
open science

## From set relations to belief function relations

Sébastien Destercke, Frédéric Pichon, John Klein

► **To cite this version:**

Sébastien Destercke, Frédéric Pichon, John Klein. From set relations to belief function relations. International Journal of Approximate Reasoning, 2019, 110, pp.46-63. 10.1016/j.ijar.2019.04.002 . hal-02108318

**HAL Id: hal-02108318**

**<https://hal.science/hal-02108318>**

Submitted on 24 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# From set relations to belief function relations

Sébastien Destercke, Frédéric Pichon, and John Klein

Sorbonne Universités, Université Technologique de Compiègne, CNRS, UMR 7253 -  
Heudiasyc, 57 Avenue de Landshut, Compiègne, France

Univ. Artois, EA 3926, Laboratoire de Génie Informatique et d'Automatique de  
l'Artois (LGI2A), Béthune, F-62400, France.

Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL - Centre de Recherche en  
Informatique Signal et Automatique de Lille, F-59000 Lille, France

`sebastien.destercke@hds.utc.fr`

`frederic.pichon@univ-artois.fr`

`john.klein@univ-lille.fr`

**Abstract.** In uncertainty theories, a common problem is to define how we can extend relations between sets (e.g., inclusion, ranking, consistency, ...) to corresponding notions between uncertainty representations. Such definitions can then be used to perform the same operations as those that are done for sets: comparing information content, ordering alternatives or checking consistency, to name a few. In this paper, we propose a general way to extend set relations to belief functions, using constrained stochastic matrices to identify those belief functions in relation. We then study some properties of our proposal, as well as its connections with existing works focusing on specific relations.

**Keywords:** set relations, belief functions, specificity, ranking, consistency.

## 1 Introduction

One can define many relations between two (or more) subsets  $A, B$  of some finite set  $X$ , i.e. between elements of some boolean algebra  $(2^X, \cap, \cup, \cdot^c)$ . Such relations can check whether the sets are consistent ( $A \cap B \neq \emptyset$ ); whether one set is more informative than another, or implies it ( $A \subseteq B$ ); when the space on which they are defined is ordered, whether one set is “higher” than another ( $A \prec B$ ); etc. These relations can then be related to practical problems such as restoring consistency or ranking alternatives.

To address the same questions in those uncertainty theories that formally generalise set theory (based, e.g., on possibility distributions, belief functions or sets of probabilities [12]), it is desirable to carry over relations between sets to uncertainty representations. Given the higher expressiveness of such theories, the problem is ill-posed in the sense that there is not a unique way to do so. We can cite as a typical example the notion of inclusion between belief functions, that has many definitions [15]. Yet, the works that deal with such issues usually focus

39 on extending one particular relation (e.g., inclusion, non-empty intersection) in  
40 meaningful ways.

41 In this paper, we propose a simple way to extend any set relation to an equiv-  
42 alent relation between belief functions, in the sense that the relation is exactly  
43 recovered when considering categorical belief functions (i.e., belief functions hav-  
44 ing a single focal element), that are equivalent to sets. Basically, for a pair of  
45 belief functions to be in relation, we require that there must exist at least one  
46 (left) stochastic matrix such that one of these belief functions is obtained as  
47 the dot product of the matrix with the other belief function. Additionally, the  
48 matrix is constrained to have null entries on pairs of focal sets not satisfying the  
49 relation to extend.

50 To our knowledge, no systematic ways of extending set relations has been  
51 proposed in the literature before, and while there may be other ways to perform  
52 such an extension, the presented solution has the advantage to be a formal ex-  
53 tension (as the relation is exactly recovered for the case of sets), and to connect  
54 with other more specific proposals of the literature. The proposal is presented in  
55 Section 2, along with the necessary reminders. To which extent it can preserve  
56 properties of the initial relation, including its compatibility with (multivariate)  
57 functions, is studied in Sections 3 (properties on initial spaces) and 5 (compat-  
58 ibility property). To make the approach more concrete, Section 4 relates it to  
59 existing works on specific relations, while Section 6 illustrates the results by  
60 applying them to simple examples, sometimes inspired from applications (sys-  
61 tem reliability and multi-criteria decision making). Finally, Section 7 discusses  
62 a mean to make the relation no longer binary but gradual, building first connec-  
63 tions to fuzzy relations.

## 64 2 Main proposal

65 This section recalls the basic tools that are necessary to understand this paper,  
66 and present our main proposal. The next sections will then focus on studying  
67 its properties and connection with other works.

### 68 2.1 Relations and their properties

69 Given some (here finite) space  $X$ , a relation  $\mathbf{R}$  between subsets of  $X$  (i.e., on  
70 the power set  $2^X$ ) is just a subset  $\mathbf{R} \subseteq 2^X \times 2^X$  that specifies which pair of  
71 subsets are related to each others. For convenience, we will write  $\mathbf{A}\mathbf{R}\mathbf{B}$  whenever  
72  $(A, B) \in \mathbf{R}$ , and  $\neg\mathbf{A}\mathbf{R}\mathbf{B}$  whenever  $(A, B) \notin \mathbf{R}$ .

*Example 1.* As an illustration, let us consider the binary space  $X = \{a, b\}$ , and  
the strict inclusion relation  $\mathbf{R} = \subset$ . Then we have

$$\mathbf{R} = \{(\emptyset, \{a\}), (\emptyset, \{b\}), (\emptyset, \{a, b\}), (\{a\}, \{a, b\}), (\{b\}, \{a, b\})\}$$

73 and the fact that  $(\{a\}, \{a, b\}) \in \mathbf{R}$  can be denoted  $\{a\}\mathbf{R}\{a, b\}$ . The fact that  
74  $(\{a\}, \{b\}) \notin \mathbf{R}$  is denoted  $\neg\{a\}\mathbf{R}\{b\}$ .

Such relations can have many different properties, the main ones that can be found in the literature being the following:

1. Symmetry:  $\mathbf{R}$  is symmetric iff  $\mathbf{A}\mathbf{R}\mathbf{B} \implies \mathbf{B}\mathbf{R}\mathbf{A}$  for all  $A, B \subseteq X$
2. Antisymmetry:  $\mathbf{R}$  is antisymmetric iff  $\mathbf{A}\mathbf{R}\mathbf{B} \wedge \mathbf{B}\mathbf{R}\mathbf{A} \implies A = B$  for all  $A, B \subseteq X$
3. Asymmetry:  $\mathbf{R}$  is asymmetric iff  $\mathbf{A}\mathbf{R}\mathbf{B} \implies \neg(\mathbf{B}\mathbf{R}\mathbf{A})$  for all  $A, B \subseteq X$
4. Reflexivity:  $\mathbf{R}$  is reflexive iff  $\mathbf{A}\mathbf{R}\mathbf{A}$  for all  $A \subseteq X$
5. Irreflexivity:  $\mathbf{R}$  is irreflexive iff  $\neg(\mathbf{A}\mathbf{R}\mathbf{A})$  for all  $A \subseteq X$
6. Transitivity:  $\mathbf{R}$  is transitive iff  $\mathbf{A}\mathbf{R}\mathbf{B} \wedge \mathbf{B}\mathbf{R}\mathbf{C} \implies \mathbf{A}\mathbf{R}\mathbf{C}$  for all  $A, B, C \subseteq X$
7. Completeness:  $\mathbf{R}$  is complete, or total, iff  $\mathbf{A}\mathbf{R}\mathbf{B} \vee \mathbf{B}\mathbf{R}\mathbf{A}$  for all  $A, B \subseteq X$

In addition to those properties, more complex relations have been defined as combination of those properties, that play an important role in many problems. These are, for instance, equivalence relations as well as order relations of different types. They are summarised in Table 1, together with the properties they satisfy.

Name	1	2	3	4	5	6	7
Tolerance	✓		✓				
Partial equivalence	✓					✓	
Equivalence	✓		✓		✓	✓	
Preorder				✓	✓		
Total Preorder				✓	✓	✓	✓
Partial order		✓	✓	✓	✓		
Total order		✓	✓	✓	✓	✓	✓

Table 1. Complex relations

## 2.2 Belief functions

Belief functions or their equivalent representations as mathematical tools can be traced back at least to Choquet [4], but their use as uncertainty representation was popularised first by Dempster [6] and Shafer [27], before being used by Smets [29] in his Transferable Belief Model.

Their mathematical properties makes them interesting uncertainty models, as they generalise a number of uncertainty representations [9] (possibility measures, sets of cumulative distributions, probabilities), while remaining of limited complexity when compared to more complex models such as lower previsions or desirable gambles [7].

Formally, a belief function on a finite space  $X = \{x_1, \dots, x_K\}$  is in one-to-one correspondence with a mass function  $m : 2^X \rightarrow [0, 1]$  that satisfies  $\sum_{A \subseteq X} m(A) = 1$ . From such a mass function, the belief and plausibility of an event  $A \subseteq X$  respectively read

$$Bel(A) = \sum_{\emptyset \neq E \subseteq A} m(E) \text{ and } Pl(A) = \sum_{E \cap A \neq \emptyset} m(E). \quad (1)$$

103 If  $m(\emptyset) = 0$ , they can be interpreted as bounds of the probability  $P(A)$  of  $A$ ,  
 104 inducing the probability set

$$\mathcal{P} = \{P : Bel(A) \leq P(A) \leq Pl(A), \forall A \subseteq X\}. \quad (2)$$

105 Within this latter interpretation and in contrast with the works set within the so-  
 106 called Dempster-Shafer theory, the mass function is not a central tool, but merely  
 107 a possible transformation of the lower envelope of  $\mathcal{P}$  given by the belief function.  
 108 As the mass function  $m$  plays a fundamental role in our proposal, the current  
 109 work is more in-line with the Dempster-Shafer interpretation of belief functions,  
 110 however it does not prevent it to have links with an imprecise probabilistic  
 111 interpretation.

112 We denote by  $\mathcal{B}^X$  the set of all belief functions on  $X$ . A particularly interest-  
 113 ing subclass of belief functions for this study are categorical ones. A categorical  
 114 mass function, denoted  $m_B$ , is such that  $m_B(B) = 1$ .

### 115 2.3 Extending set relations to belief functions

116 Let  $\mathbf{R}$  be a relation on  $2^X$  (equivalently a subset of  $2^X \times 2^X$ ). We then propose  
 117 the following simple definition to extend this relation to belief functions, i.e. into  
 118 a relation on  $\mathcal{B}^X$ :

**Definition 1.** *Given two mass functions  $m_1, m_2$  and a subset relation  $\mathbf{R}$ , we say that  $m_1 \tilde{\mathbf{R}} m_2$  iff there is a (left)<sup>1</sup> stochastic matrix  $S$  such that  $\forall A, B \subseteq X$*

$$m_1(A) = \sum_{B \subseteq X} S(A, B) m_2(B) \quad (3)$$

$$\text{with } S(A, B) > 0 \wedge m_2(B) > 0 \implies \mathbf{R}A, B. \quad (4)$$

Definition 1 states that  $m_1 \tilde{\mathbf{R}} m_2$  iff  $m_1$  can be obtained from  $m_2$  by transferring each mass  $m_2(B)$  to a subset  $A$  such that  $\mathbf{R}A, B$ . It is easily checked that  $\tilde{\mathbf{R}}$  is a generalisation of  $\mathbf{R}$  in the sense that

$$m_A \tilde{\mathbf{R}} m_B \Leftrightarrow \mathbf{R}A, B, \forall A, B \subseteq X. \quad (5)$$

119 Indeed, if  $\mathbf{R}A, B$ , we can choose  $S(E, F) = m_A(E)$  for all  $F \subseteq X$ , and this  
 120 matrix matches the conditions of Definition 1, hence  $m_A \tilde{\mathbf{R}} m_B$ . Conversely, if  
 121  $m_A \tilde{\mathbf{R}} m_B$ , then (3) implies  $S(A, B) = 1$  and (4) then gives  $\mathbf{R}A, B$ .

122 Also, there is only one relation  $\tilde{\mathbf{R}}$  on belief functions spanned by Definition  
 123 1 from a given set relation  $\mathbf{R}$ . To see this, suppose two such belief function  
 124 relations exist. If a matrix matching the conditions of Definition 1 was found for  
 125 the first one then the same matrix also works for the other and the relations are  
 126 equivalent. Likewise, two relations  $\mathbf{R}$  and  $\mathbf{R}'$  defined on sets cannot lead, through  
 127 Definition 1, to the same relation  $\tilde{\mathbf{R}}$  on belief functions. This is an immediate  
 128 consequence of (5). Consequently and by a small abuse of notation, we will use

<sup>1</sup> We use left-stochasticity only throughout the paper.

129 the same notation for a relation  $\mathbf{R}$  on the subset or belief function side in the  
 130 remainder of the paper, as it introduces no ambiguity. However, in general, the  
 131 stochastic matrix  $S$  involved in Definition 1 is not unique when  $m_1 \mathbf{R} m_2$  holds.

132 Definition 1 is inspired from previous works on specificity of belief functions  
 133 [15, 16, 30], as well as on recent proposals dealing with set ordering [24].  
 134 In particular, Definition 1 can be endowed with an interpretation similar to the  
 135 one given in [15], as  $S(A, B)$  can be seen as the ratio of  $m(B)$  that flows from  
 136  $B$  to  $A$ , with the flow being possibly non-null only when  $\mathbf{A} \mathbf{R} B$ .

137 *Remark 1.* Readers that are familiar with the belief function literature may wonder  
 138 why the condition  $m_2(B) > 0$  is necessary in (4), as this condition does not  
 139 appear in related works. This condition is necessary to generalise any relation  
 140 on sets that is not inverse serial, i.e. a relation such that there is a  $B_*$  with  
 141  $\neg(\mathbf{A} \mathbf{R} B_*)$ ,  $\forall A \subseteq X$ . For such sets  $B_*$ , left stochasticity is incompatible with  
 142 the implication  $S(A, B_*) > 0 \implies \mathbf{A} \mathbf{R} B_*$ , and without checking  $m_2(B) > 0$   
 143 in Definition 1 the relation on belief functions of a not inverse serial  $\mathbf{R}$  would  
 144 always be empty. By checking  $m_2(B) > 0$  in Definition 1, we can induce a non  
 145 empty relation on belief functions. When  $B_*$  is a focal element of  $m_2$ , we have  
 146  $\neg(m_1 \mathbf{R} m_2)$ , which makes perfect sense. When  $m(B_*) = 0$ , then a null mass can  
 147 be distributed to any set  $A$  without harm.

148 As the above mentioned related works dealt with directional, or rather asym-  
 149 metric relations, Definition 1 is naturally asymmetric. However, Proposition 1  
 150 shows that it has a somehow symmetric counterpart.

151 **Proposition 1.** *Consider two mass functions  $m_1, m_2$  and a belief function re-*  
 152 *lation  $\mathbf{R}$ . Then the two following conditions are equivalent:*

1. *there is a stochastic matrix  $S(A, B)$  such that*

$$m_1(A) = \sum_{B \subseteq X} S(A, B) m_2(B),$$

*with  $S(A, B) > 0 \wedge m_2(B) > 0 \implies \mathbf{A} \mathbf{R} B$ .*

2. *there is a joint mass function  $m_{12}(A, B)$  on  $2^X \times 2^X$  such that*

$$m_{12}(A, B) > 0 \implies \mathbf{A} \mathbf{R} B, \quad (6)$$

$$m_1(A) = \sum_B m_{12}(A, B), \quad (7)$$

$$m_2(B) = \sum_A m_{12}(A, B). \quad (8)$$

*Proof.* 1.  $\implies$  2. First, consider the matrix  $S(A, B)$ , that we know exists if  
 $m_1 \mathbf{R} m_2$ . Let us now simply define the joint  $m_{12}$  as

$$m_{12}(A, B) = m_2(B) S(A, B) \text{ for any } A, B.$$

We clearly have  $m_{12}(A, B) > 0$  only if  $\mathbf{AR}B$ , since  $S(A, B) > 0$  means that either  $\mathbf{AR}B$  or  $m_2(B) = 0$  (in the other cases it is null), and moreover

$$\sum_B m_{12}(A, B) = \sum_B m_2(B)S(A, B) = m_1(A),$$

$$\sum_A m_{12}(A, B) = \sum_A m_2(B)S(A, B) = m_2(B) \sum_A S(A, B) = m_2(B)$$

153 with the last equality following from  $S$  being stochastic.

2.  $\implies$  1. Again, consider the joint  $m_{12}(A, B)$  satisfying constraints (6)-(8), that we know exists by assumption. If we assume that this implies the existence of matrix  $S$ , we get

$$m_1(A) = \sum_B m_{12}(A, B) = \sum_B m_2(B)S(A, B).$$

154 For any  $B$  s.t.  $m_2(B) > 0$ , we thus define

$$S(A, B) = \frac{m_{12}(A, B)}{m_2(B)}. \quad (9)$$

The other entries of  $S$  are set to arbitrary values provided that these latter are compliant with left stochasticity. For those entries which are set according to (9), i.e. when  $m_2(B) > 0$ , we can now check that  $S(A, B)$  satisfies the required properties, as

$$\sum_A S(A, B) = \frac{\sum_A m_{12}(A, B)}{m_2(B)} = \frac{m_2(B)}{m_2(B)} = 1,$$

$$S(A, B) > 0 \Leftrightarrow \frac{m_{12}(A, B)}{m_2(B)} > 0 \Rightarrow \mathbf{AR}B$$

□

This proposition shows, in particular, that any stochastic matrix  $S$  can be associated to a unique joint mass function  $m_{12}$ , and vice-versa. Also note that, using a transformation similar to the one of the second part of the proof, we can alternatively build a stochastic matrix  $S'$  such that

$$S'(B, A) = \begin{cases} \frac{m_{12}(A, B)}{m_1(A)} & \text{if } m_1(A) > 0 \\ \lambda_B^{(A)} & \text{if } m_1(A) = 0 \end{cases},$$

with  $\sum_B \lambda_B^{(A)} = 1$ .  $S'$  is such that

$$m_2(B) = \sum_{A \subseteq X} S'(B, A)m_1(A).$$

155 Moreover,  $S'(A, B) > 0$  and  $m_1(A) > 0$  imply  $\mathbf{BR}A$  but gives no guarantee on  
156  $\mathbf{AR}B$ .

157 *Remark 2.* Proposition 1 shows that we can view our definition of relations in  
 158 two different ways: as a "transfer" matrix  $S$  allowing to go from  $m_2$  to  $m_1$   
 159 without violating the relation on sets, or as the existence of a joint structure  
 160 consistent with  $m_1, m_2$  and the relation  $\mathbf{R}$ . Although we consider that the joint  
 161 structure is more intuitive and easier to explain, both views have been adopted  
 162 in the past and are in our opinion useful, as:

- 163 – there are settings where one mathematical tool is more natural then the  
 164 other. For instance, Smets' matrix computations [28] make a heavy use of the  
 165 first view, while recent works about consistency adopt the second view [11];
- 166 – mathematically, it may also be more convenient to use one or the other, for  
 167 instance in proofs. For example, most of our negative proofs and examples  
 168 use joint matrices and the second view, but Propositions 7 and 10 are simpler  
 169 to prove using the first view.

170 Finally, let us note that the relation  $\mathbf{R}$  on belief functions can be interpreted  
 171 in exactly the same way as the relation on sets it extends, this interpretation  
 172 varying according to the application and pursued goal. For instance, the relation  
 173  $ARB$  iff  $A \cap B \neq \emptyset$  will often be used when  $A, B$  concern the same object of inter-  
 174 ests but are issued from different sources, and when one wants to check whether  
 175 they are consistent. In contrast, ranking relations between  $A, B$  will often be  
 176 used when  $A, B$  concern different objects or alternatives evaluated on the same  
 177 scale (e.g., movies given a finite number of stars). Generally speaking, mass func-  
 178 tions are random set distributions [26] and relation  $\mathbf{R}$  is one way to propagate  
 179 a relation (and its interpretation) on sets to their random counterparts.

### 180 3 Property preservation

#### 181 3.1 Preservation of simple properties

182 We may now wonder how many of the initial relation  $\mathbf{R}$  properties between sets  
 183 are preserved when extended to belief functions according to Definition 1. We will  
 184 now provide a series of results for common properties, either by providing proofs  
 185 or counter-examples. We will keep the proposition/proof format, to provide a  
 186 uniform presentation.

**Proposition 2 (Preserved symmetry).** *If  $\mathbf{R}$  is symmetric on sets, it is so on belief functions:*

$$m_1 \mathbf{R} m_2 \implies m_2 \mathbf{R} m_1, \forall m_1, m_2.$$

*Proof.* Let us assume that  $S(A, B)$  is a stochastic matrix satisfying Definition 1 for  $m_1 \mathbf{R} m_2$ , and  $m_{12}$  is its associated joint mass. Then we can see that

$$S'(A, B) > 0 \text{ and } m_1(A) > 0 \implies B \mathbf{R} A \Leftrightarrow A \mathbf{R} B,$$

since  $\mathbf{R}$  is symmetric. Matrix  $S'$  satisfies the conditions of Definition 1, hence  $m_2 \mathbf{R} m_1$ .  $\square$



**Proposition 3 (Unpreserved antisymmetry).** *If  $\mathbf{R}$  is antisymmetric on sets, it is not necessarily so on belief functions, as*

$$m_1 \mathbf{R} m_2 \wedge m_2 \mathbf{R} m_1 \not\Rightarrow m_2 = m_1$$

*Proof.* Consider two mass functions that are positive only on subsets  $A, B, C$  and such that

$$\begin{aligned} m_1(A) &= 0.3, m_1(B) = 0.5, m_1(C) = 0.2, \\ m_2(A) &= 0.4, m_2(B) = 0.3, m_2(C) = 0.3, \end{aligned}$$

as well as the antisymmetric relation  $\mathbf{R}$  on those subsets summarised by the matrix

$$\begin{array}{c} A \quad B \quad C \\ \begin{array}{l} A \\ B \\ C \end{array} \left[ \begin{array}{ccc} \mathbf{R}A & \mathbf{R}B & \\ & \mathbf{R}B & \mathbf{R}C \\ \mathbf{R}A & & \mathbf{R}C \end{array} \right]. \end{array}$$

We can then consider the joint mass function

$$\begin{aligned} m_{12}(A, A) &= 0.3, m_{12}(B, B) = 0.3, \\ m_{12}(B, C) &= 0.2, m_{12}(C, A) = 0.1, m_{12}(C, C) = 0.1, \end{aligned}$$

that shows that we have  $m_1 \mathbf{R} m_2$ , while the joint mass function

$$\begin{aligned} m_{12}(A, A) &= 0.2, m_{12}(B, B) = 0.3, \\ m_{12}(A, C) &= 0.1, m_{12}(B, A) = 0.2, m_{12}(C, C) = 0.2, \end{aligned}$$

shows that  $m_2 \mathbf{R} m_1$ , hence we can have both without  $m_1 = m_2$ .  $\square$

**Proposition 4 (Unpreserved asymmetry).** *If  $\mathbf{R}$  is asymmetric on sets, it is not necessarily so on belief functions, as*

$$m_1 \mathbf{R} m_2 \not\Rightarrow \neg(m_2 \mathbf{R} m_1)$$

*Proof.* Simply consider two mass functions  $m_1, m_2$  that are positive only on subsets  $A, B, C, D, E$  and such that

$$\begin{aligned} m_1(A) &= 0.2, m_1(B) = 0.3, m_1(C) = 0.2, m_1(D) = 0.1, m_1(E) = 0.2, \\ m_2(A) &= 0.2, m_2(B) = 0.1, m_2(C) = 0.3, m_2(D) = 0.3, m_2(E) = 0.1 \end{aligned}$$

as well as the asymmetric relation  $\mathbf{R}$  on those subsets summarised by the matrix

$$\begin{array}{c} A \quad B \quad C \quad D \quad E \\ \begin{array}{l} A \\ B \\ C \\ D \\ E \end{array} \left[ \begin{array}{ccccc} & & \mathbf{R}C & \mathbf{R}D & \\ \mathbf{R}A & & & & \mathbf{R}E \\ & \mathbf{R}B & & \mathbf{R}D & \\ \mathbf{R}A & \mathbf{R}B & & & \mathbf{R}E \\ \mathbf{R}A & & \mathbf{R}C & & \end{array} \right] \end{array}$$

We can then consider the joint mass function

$$m_{12}(A, C) = 0.1, m_{12}(A, D) = 0.1, m_{12}(B, A) = 0.2$$

$$m_{12}(B, E) = 0.1, m_{12}(C, D) = 0.2, m_{12}(D, B) = 0.1, m_{12}(E, C) = 0.2$$

that shows that we have  $m_1 \mathbf{R} m_2$ , while the joint mass function

$$m_{12}(A, B) = 0.1, m_{12}(A, E) = 0.1, m_{12}(B, C) = 0.2$$

$$m_{12}(B, D) = 0.1, m_{12}(C, A) = 0.2, m_{12}(D, C) = 0.1, m_{12}(E, D) = 0.2$$

shows that  $m_2 \mathbf{R} m_1$ , hence we can have both. □

**Proposition 5 (Preserved reflexivity).** *If  $\mathbf{R}$  is reflexive on sets, it is so on belief functions:*

$$\forall m, \text{ we have } m \mathbf{R} m$$

*Proof.* Simply observe that, if  $\mathbf{R}$  is reflexive ( $A \mathbf{R} A$  for any subset) and if  $m_1 = m_2 = m$ , we can always define the joint mass function such that for any  $A$  we have  $m_{12}(A, A) = m(A)$ , that satisfies Equations (6)-(8). □

187 **Proposition 6 (Unpreserved irreflexivity).** *If  $\mathbf{R}$  is irreflexive on sets, it is*  
188 *not necessarily so on belief functions, as we may have  $m \mathbf{R} m$  for some  $m \in \mathcal{B}^X$ .*

*Proof.* Consider the following mass function

$$m(A) = 0.5, m(B) = 0.5$$

and the relation  $\mathbf{R}$  summarised in the following matrix

$$\begin{array}{c} A \\ B \end{array} \begin{array}{cc} A & B \\ \left[ \begin{array}{cc} & A \mathbf{R} B \\ B \mathbf{R} A & \end{array} \right] \end{array}$$

which is irreflexive. However, the joint  $m(A, B) = m(B, A) = 0.5$  shows that we have  $m \mathbf{R} m$ , hence  $\mathbf{R}$  may not be irreflexive for belief functions. □

**Proposition 7 (Preserved transitivity).** *If  $\mathbf{R}$  is transitive on sets, it is so on belief functions:*

$$m_1 \mathbf{R} m_2 \wedge m_2 \mathbf{R} m_3 \implies m_1 \mathbf{R} m_3$$

*Proof.* If we have  $m_1 \mathbf{R} m_2 \wedge m_2 \mathbf{R} m_3$ , this means that there are two matrices  $S_{12}$  and  $S_{23}$  satisfying Definition 1 and such that

$$m_1(A) = \sum_B S_{12}(A, B) m_2(B),$$

$$m_2(B) = \sum_C S_{23}(B, C) m_3(C).$$

We therefore have

$$\begin{aligned} m_1(A) &= \sum_B S_{12}(A, B) \sum_C S_{23}(B, C) m_3(C) \\ &= \sum_B \sum_C S_{12}(A, B) S_{23}(B, C) m_3(C) \\ &= \sum_C m_3(C) \sum_B S_{12}(A, B) S_{23}(B, C) \end{aligned}$$

Now, let us define the matrix  $S_{13}$  elements as

$$S_{13}(A, C) = \sum_B S_{12}(A, B) S_{23}(B, C),$$

189 meaning that  $S_{13} = S_{12} \cdot S_{23}$  is the result of a matrix product. One can then  
190 show that  $S_{13}$  satisfies Definition 1 and that  $m_1 \mathbf{R} m_3$  as

- 191 –  $S_{13}$  is stochastic, being the product of stochastic matrices ;  
– we have that

$$\begin{aligned} S_{13}(A, C) > 0 &\Leftrightarrow \exists(A, B) \text{ and } (B, C) \text{ s.t. } S_{12}(A, B) S_{23}(B, C) > 0 \\ &\Rightarrow \mathbf{A} \mathbf{R} \mathbf{B} \wedge \mathbf{B} \mathbf{R} \mathbf{C} \\ &\Rightarrow \mathbf{A} \mathbf{R} \mathbf{C}. \end{aligned}$$

□

192 **Proposition 8 (Unpreserved completeness).** *If  $\mathbf{R}$  is complete (or total) on*  
193 *sets, it is not necessarily so on belief functions: for any two  $m_1, m_2$  we may have*  
194 *neither  $m_1 \mathbf{R} m_2$  nor  $m_2 \mathbf{R} m_1$ .*

*Proof.* Consider the relation  $\mathbf{R}$  = “having a lower cardinality than” on the space  
 $X = \{a, b, c\}$ , meaning that  $\mathbf{A} \mathbf{R} \mathbf{B} \Leftrightarrow |A| \leq |B|$ , which is a complete relation on  
sets. Consider now the two mass functions

$$\begin{aligned} m_1(\{a\}) &= 0.6, & m_1(\{a, b\}) &= 0.4, \\ m_2(\{a\}) &= 0.8, & m_2(X) &= 0.2. \end{aligned}$$

195 Then, we have neither  $m_1 \mathbf{R} m_2$ , nor  $m_2 \mathbf{R} m_1$ , as indeed all stochastic matrices  
196 such that  $m_1 = S \cdot m_2$  or  $m_2 = S \cdot m_1$  must contain non-null value on pairs of  
197 subsets  $A, B$  with  $\neg(\mathbf{A} \mathbf{R} \mathbf{B})$ . Consider for instance the case

$$m_1 : \begin{matrix} \{a\} \\ \{a, b\} \\ X \end{matrix} \begin{pmatrix} 0.6 \\ 0.4 \\ 0 \end{pmatrix} = \begin{pmatrix} 3/4 & 0 & 0 \\ 1/4 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0.8 \\ 0 \\ 0.2 \end{pmatrix} : m_2.$$

198 We only display the submatrix of  $S$  corresponding to focal elements of the mass  
199 functions. Entries in green can be set to either 0 or some positive number. Entries  
200 in red cannot be assigned a positive number. It is clear that at least some non-  
201 null value must be given to  $S(\{a, b\}, \{a\})$ , hence  $\neg m_1 \mathbf{R} m_2$ . A similar observation  
202 can be made for the reverse case.

□

The expected cardinality [17] of a belief function, defined as

$$C(m) = \sum_{A \subseteq X} m(A)|A|,$$

203 yields a complete relation between belief functions that is a generalisation of the  
 204 relation  $\mathbf{R}$  defined in the above proof. This is therefore also an illustration that  
 205 not all binary relations on belief functions can be retrieved via the mechanism  
 206 under study.

207 Table 2 summarises our obtained results so far, and in particular which prop-  
 208 erties existing on subsets is guaranteed to be preserved when considering them  
 209 on the richer language of belief functions. It should be noted that even if a prop-  
 210 erty is not guaranteed to be preserved in general, it may be preserved in specific  
 211 cases: for instance, the inclusion relation is antisymmetric, and its generalisation  
 212 to belief functions, called specialisation [15], is too. The same remark is true for  
 213 strict inclusion, that is asymmetric. We will see later on that all partial orders  
 214 (among which inclusion) are in fact preserved by Definition 1.

$\mathbf{R}$ on $2^X$ is	$\implies$ $\mathbf{R}$ on $\mathcal{B}^X$ is
Symmetric	Yes
Antisymmetric	No
Asymmetric	No
Reflexive	Yes
Irreflexive	No
Transitive	Yes
Complete	No

**Table 2.** Summary of properties preservation

215 Finally, we can also consider two different binary relations  $\mathbf{R}$  and  $\mathbf{R}'$  and  
 216 check whether a property for this pair of relations is preserved. There is mainly  
 217 one such property which is implication.

218 **Proposition 9 (Preserved implication).** *If  $\mathbf{R}$  and  $\mathbf{R}'$  are such that  $ARB \implies$   
 219  $AR'B$  for any subsets  $A$  and  $B$ , it is so on belief functions.*

220 *Proof.* Let  $m_1$  and  $m_2$  denote two mass functions and  $S$  is a stochastic matrix  
 221 compliant with definition 1 for relation  $\mathbf{R}$ . Obviously,  $S$  is also compliant with  
 222 definition 1 for relation  $\mathbf{R}'$  because when  $m_2(B) > 0$ , then  $S(A, B) > 0 \implies$   
 223  $ARB \implies AR'B$ .

### 224 3.2 Preservation of classical relations

225 In this section, we study whether some classical relations composed of multi-  
 226 ple properties are preserved by our definition. We will limit ourselves to the

relations recalled in Table 1, as those are the most common, but clearly others could be studied, such as specific order relations (semi-orders, interval orders, tournaments, ...) [19].

A very first remark is that, if the relation is defined by a set of properties that are all preserved by our definition, then it is immediate that it is preserved when considering it on belief functions. Among other things, this means that

- Tolerance relations
- (Partial) Equivalence relations
- Preorder relations

are preserved when extended to belief functions. However, total preorders are usually not preserved when extended to belief functions. A simple example is given in the proof of Proposition 8. In this proof we examine the relation “having a lower cardinality than” which is a total preorder on sets while the induced relation on belief functions fails to be complete.

Moreover, it is easy to see that total orders are not preserved either, as any refinement of the relation “having a lower cardinality than” into a total order would constrain even more the element on which the stochastic matrix has to be positive. Another very common class of binary relations on sets are partial orders, for which we can show that they are preserved:

**Proposition 10 (Preserved partial order).** *If  $\mathbf{R}$  is a partial order on sets, it is so on belief functions.*

See Appendix A for a proof of the above proposition.

## 4 Related works

### 4.1 Inclusion and consistency

In the case where the relations are either inclusion or consistency, then we retrieve well-known results of the literature:

- in the case of inclusion we have  $\mathbf{A}\mathbf{R}\mathbf{B}$  iff  $A \subseteq B$ , and Definition 1 is then essentially equivalent to that of specialisation [15]. The only difference amounts to checking if  $m_2(B) > 0$  in condition (4), that we need to handle generic relations, but that is not needed in the specific case of specialisation (as in this case, for any  $B$  there is always a subset  $A$  such that  $\mathbf{A}\mathbf{R}\mathbf{B}$ ). Beyond this difference, the notion of specialisation and the extension of inclusion to belief functions proposed in this paper are actually formally equivalent in the sense that for any mass functions  $m_1$  and  $m_2$ ,  $m_1$  is a specialisation of  $m_2$  if and only if  $m_1 \mathbf{R} m_2$ .
- in the case of consistency, we have  $\mathbf{A}\mathbf{R}\mathbf{B}$  iff  $A \cap B \neq \emptyset$ , and one can see that  $m_1 \mathbf{R} m_2$  iff there is a joint mass assigning positive mass to pairs of sets having a non-empty intersection. This is equivalent to require  $\mathcal{P}_1 \cap \mathcal{P}_2 \neq \emptyset$ , with  $\mathcal{P}_i$  the probability set induced by  $m_i$  [3].

266 **4.2 Rankings**

267 When the space  $X = \{x_1, \dots, x_n\}$  is ordered (with  $x_i \leq x_{i+1}$ ) and possibly  
 268 infinite, it makes sense to consider relations of the kind “higher than” in order  
 269 to compare sets. There are many ways to rank two sets  $A, B$ , such as:

- 270 – Single-bound dominance, that can be declined itself into four notions:
  - 271 • loose dominance:  $\mathbf{AR}_{\leq LD} B$  if  $\min A \leq \max B$
  - 272 • lower bound:  $\mathbf{AR}_{\leq LB} B$  if  $\min A \leq \min B$
  - 273 • upper bound:  $\mathbf{AR}_{\leq UB} B$  if  $\max A \leq \max B$
  - 274 • strict dominance:  $\mathbf{AR}_{\leq SD} B$  if  $\max A \leq \min B$
- 275 – Pairwise-bound or lattice dominance:  $\mathbf{AR}_{\leq PD} B$  if  $\min A \leq \min B$  and  $\max A \leq$   
 276  $\max B$ , whose extension to belief functions studied in [24] correspond to our  
 277 proposal.

278 Extensions of this kind of relations to belief functions have already been in-  
 279 vestigated in [24], and are connected to the extensions of stochastic dominance  
 280 explored in [8] for belief functions, and in [25] for the general case of sets of cumu-  
 281 lative distributions. In fact, let us first define the following stochastic dominance  
 282 notions:

- stochastic loose dominance:

$$m_1 \prec_{LD}^{St} m_2 \text{ iff } Pl_1([x_1, \dots, x_i]) \geq Bel_2([x_1, \dots, x_i]), \forall x_i \in X$$

- stochastic lower bound:

$$m_1 \prec_{LB}^{St} m_2 \text{ iff } Pl_1([x_1, \dots, x_i]) \geq Pl_2([x_1, \dots, x_i]), \forall x_i \in X$$

- stochastic upper bound:

$$m_1 \prec_{UB}^{St} m_2 \text{ iff } Bel_1([x_1, \dots, x_i]) \geq Bel_2([x_1, \dots, x_i]), \forall x_i \in X$$

- stochastic strict dominance:

$$m_1 \prec_{SD}^{St} m_2 \text{ iff } Bel_1([x_1, \dots, x_i]) \geq Pl_2([x_1, \dots, x_i]), \forall x_i \in X$$

- stochastic lattice dominance:

$$m_1 \prec_{PD}^{St} m_2 \text{ iff } (m_1 \prec_{LB}^{St} m_2) \wedge (m_1 \prec_{UB}^{St} m_2)$$

283 We then have the following strong relationships between the extensions of  
 284 ranking to belief functions and the stochastic dominance relations:

**Proposition 11.** *For  $y \in \{LD, LB, UB, SD\}$ , we have that*

$$m_1 \mathbf{R}_{\leq y} m_2 \Leftrightarrow m_1 \prec_y^{St} m_2$$

285 *Proof.* We will only demonstrate the relation for one of the  $y$ , that is  $SD$  (the  
286 strongest relation), as proofs for the other cases are analogous.

287  $\Leftarrow$  First, let us remind that if  $m_1 \prec_{SD}^{St} m_2$ , it means that the cumulative  
288 distribution induced by the minimal values of the focal elements of  $m_2$  stochas-  
289 tically dominates the one induced by the maximal values of  $m_1$ . Let us denote  
290  $A_1, \dots, A_n$  and  $B_1, \dots, B_m$  the focal sets of  $m_1, m_2$ , and assume without loss of  
291 generality that they are ordered according to their maximal values for  $m_1$ , and  
292 their minimal values for  $m_2$ , that is  $\max A_i \leq \max A_{i+1}$  for any  $i = 1, \dots, n-1$   
293 and  $\min B_i \leq \min B_{i+1}$  for any  $i = 1, \dots, m-1$ .

Let us denote by  $\alpha_i = \sum_{j=1}^i m_1(A_j)$  and  $\beta_i = \sum_{j=1}^i m_2(B_j)$  the cumulated weights of the first  $i$  elements of  $m_1$  and  $m_2$ , assuming all  $\alpha_i, \beta_i$  are all distinct for easiness. We denote by

$$\gamma_1, \dots, \gamma_{n+m-1} = \{\alpha_1, \dots, \alpha_n\} \cup \{\beta_1, \dots, \beta_m\}$$

the union of all distinct possible cumulative values of masses, assuming that they are also ordered, i.e.,  $\gamma_i \leq \gamma_{i+1}$  (we have  $m+n-1$  distinct values because  $\alpha_n = \beta_m = 1$ ). Let us now define the following joint mass function  $m_{12}$  such that, for any  $i = 1, \dots, m+n-1$ ,

$$m_{12}(A_{\gamma_i}, B_{\gamma_i}) = \gamma_i - \gamma_{i-1}$$

with  $\gamma_0 = 0$ , and the following definitions for the focal sets:

$$A_{\gamma_i} = \{A_i : (\sum_{j=1}^i m_1(A_j) \geq \gamma_i) \wedge (\sum_{j=1}^{i-1} m_1(A_j) < \gamma_i)\},$$

$$B_{\gamma_i} = \{B_i : (\sum_{j=1}^i m_2(B_j) \geq \gamma_i) \wedge (\sum_{j=1}^{i-1} m_2(B_j) < \gamma_i)\},$$

294 that by construction satisfy Equations (7)-(8). The construction is illustrated in  
295 Figure 1 for the case of  $n = 3$  and  $m = 2$ . This comes down to construct the  
296 joint mass in a level-wise manner, and since we also have that  $m_1 \prec_{SD}^{St} m_2$ , we  
297 have that for any  $i$ ,  $\max A_{\gamma_i} \leq \min B_{\gamma_i}$ , hence  $A_{\gamma_i} \mathbf{R}_{\leq SD} B_{\gamma_i}$

$\Rightarrow$  if  $m_1 \mathbf{R}_{\leq SD} m_2$ , this means that there is a joint  $m_{12}(A, B)$  that is positive only if  $\max A \leq \min B$ . Let us now show that this implies, for any  $x_i \in X$ ,

$$Bel_1([x_1, \dots, x_i]) \geq Pl_2([x_1, \dots, x_i]),$$

which is equivalent to

$$\sum_{A: \max A \leq x_i} m_1(A) \geq \sum_{B: \min B \leq x_i} m_2(B).$$

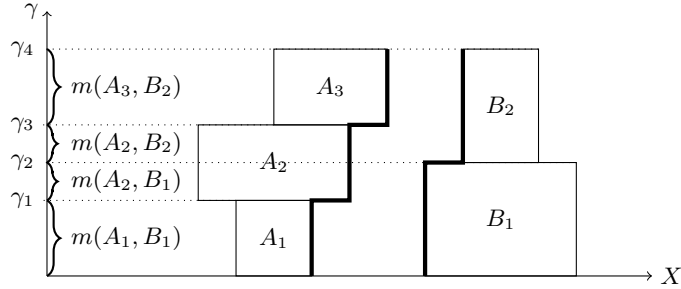
Using the relation between  $m_{12}, m_1$  and  $m_2$ , we get

$$\sum_{A: \max A \leq x_i} \sum_B m_{12}(A, B) \geq \sum_{B: \min B \leq x_i} \sum_A m_{12}(A, B)$$

but since  $m_{12}(A, B) > 0$  only if  $\max A \leq \min B$ , we can write

$$\sum_{A: \max A \leq x_i} \sum_B m_{12}(A, B) \geq \sum_{B: \min B \leq x_i} \sum_{A: \max A \leq x_i} m_{12}(A, B)$$

as all the elements on the right-hand side summation are also in the left-hand side, this latter can only be bigger.  $\square$



**Fig. 1.** Illustration of proof of Proposition 11 (construction of joint mass).

298 The above proposition shows a clear relation between ranking relations (when  
 299 extended according to Definition 1) and the corresponding stochastic dominance  
 300 relation. While this confirms the interest of our proposal and its links with  
 301 existing, more specific works, this also provides an efficient computational way to  
 302 check whether  $m_1, m_2$  are in a ranking relation, as checking stochastic dominance  
 303 is easier than checking whether a relation holds (which can be done by solving  
 304 a linear programming problem, as suggested in Section 7).

305 The next example however shows that the property is not true for pairwise  
 306 bounds, essentially because focal elements are usually not totally ordered with  
 307 respect to pairwise bounds.

*Example 2.* Let us consider the space  $X = \{x_1, \dots, x_{12}\}$  and the two following mass functions

$$\begin{aligned} m_1(\{x_1, \dots, x_7\} = A_1) &= 1/3, & m_2(\{x_2, \dots, x_{12}\} = B_1) &= 1/3, \\ m_1(\{x_3, \dots, x_9\} = A_2) &= 1/3, & m_2(\{x_4, \dots, x_8\} = B_2) &= 1/3, \\ m_1(\{x_5, \dots, x_{11}\} = A_3) &= 1/3, & m_2(\{x_6, \dots, x_{10}\} = B_3) &= 1/3. \end{aligned}$$



The pairs of sets satisfying the relation  $\mathbf{R}_{\leq PD}$  is summarised in the matrix

$$\begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} \begin{array}{ccc} B_1 & B_2 & B_3 \\ \left[ \begin{array}{ccc} \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} \\ & & \mathbf{R}_{\leq PD} \\ & & \mathbf{R}_{\leq PD} \end{array} \right] \end{array}$$

308 which shows that there are no  $m_{12}$  for which  $m_1 \mathbf{R}_{\leq PD} m_2$ , since at least a positive  
 309 number must be put on the third row. However, we do have  $m_1 \prec_{PD}^{St} m_2$ , as the  
 310 bounds of each focal elements, once increasingly re-ordered separately, satisfy  
 311 the pairwise dominance notion.

312 However, that the converse holds (if  $m_1 \mathbf{R}_{\leq PD} m_2$ , then  $m_1 \prec_{PD}^{St} m_2$ ) has  
 313 been shown in [24]. Finally, from Proposition 9 and the existing implications  
 314 between the different rankings, we can easily conclude that:

$$m_1 \mathbf{R}_{\leq SD} m_2 \Rightarrow m_1 \mathbf{R}_{\leq PD} m_2 \Rightarrow \left\{ \begin{array}{l} m_1 \mathbf{R}_{\leq UB} m_2 \\ m_1 \mathbf{R}_{\leq LB} m_2 \end{array} \right\} \Rightarrow m_1 \mathbf{R}_{\leq LD} m_2 \quad (10)$$

## 315 5 Preservation through functional mapping

### 316 5.1 Univariate functions

317 This section investigates whether a function that is compatible (in the sense  
 318 of Definition 2 below) with set relations given respectively on its domain and  
 319 codomain, is also compatible with the extensions of these relations to belief func-  
 320 tions. We consider first the case of univariate functions; multivariate functions  
 321 are handled in Section 5.2.

322 Let  $f$  be some function with domain  $X$  and codomain  $Y$ , i.e.,  $f : X \rightarrow Y$ .  
 323 We recall that the image  $f(A)$  of some subset  $A \subseteq X$  under  $f$  is the subset  
 324  $f(A) = \{f(x) : x \in A\} \subseteq Y$ . More generally, the image  $f(m)$  of some mass  
 325 function  $m \in \mathcal{B}^X$  under  $f$  is the mass function  $f(m) \in \mathcal{B}^Y$  defined, for all  
 326  $B \subseteq Y$ , as

$$f(m)(B) = \sum_{f(A)=B} m(A). \quad (11)$$

**Definition 2.** Let  $f : X \rightarrow Y$ . Let  $\mathbf{R}^X$  and  $\mathbf{R}^Y$  be relations on  $2^X$  and  $2^Y$ , respectively. The function  $f$  is said to be  $(\mathbf{R}^X; \mathbf{R}^Y)$ -compatible if

$$A \mathbf{R}^X B \Rightarrow f(A) \mathbf{R}^Y f(B), \forall A, B \subseteq X.$$

327 *Example 3.* Let  $\mathbf{R}_{\subseteq}^X$  be the relation corresponding to inclusion on  $X$ , i.e.,  $A \mathbf{R}_{\subseteq}^X B$   
 328 iff  $A \subseteq B$ ,  $A, B \subseteq X$ . Similarly, let  $\mathbf{R}_{\subset}^X$  and  $\mathbf{R}_{\cap}^X$  denote the relations correspond-  
 329 ing, respectively, to strict inclusion and consistency on  $X$ , and let  $\mathbf{R}_{\subseteq}^Y$  denote  
 330 inclusion on  $Y$ .

331 Since for any function  $f$  and any  $A, B \subseteq X$  such that  $A \subseteq B$  it holds that  
 332  $f(A) \subseteq f(B)$ , any function  $f$  is  $(\mathbf{R}_{\subseteq}^X; \mathbf{R}_{\subseteq}^Y)$ -compatible. Similarly, any function  $f$   
 333 is  $(\mathbf{R}_{\subseteq}^X; \mathbf{R}_{\subseteq}^Y)$ -compatible.

334 However, not all functions  $f$  are  $(\mathbf{R}_{\cap}^X; \mathbf{R}_{\cap}^Y)$ -compatible. For instance, if  $f$  is  
 335 a constant function, i.e.  $f(x) = y$  for some  $y \in Y$  and all  $x \in X$ , then  $f$  is  
 336  $(\mathbf{R}_{\cap}^X; \mathbf{R}_{\subseteq}^Y)$ -compatible (in this case we have  $f(A) \subseteq f(B)$  for all  $A, B \subseteq X$  such  
 337 that  $A \cap B \neq \emptyset$  since  $f(A) = f(B) = \{y\}$ ). However, if  $f$  is the identity function,  
 338 i.e.  $X = Y$  and  $f(x) = x$  for all  $x \in X$ , then  $f$  is not  $(\mathbf{R}_{\cap}^X; \mathbf{R}_{\subseteq}^Y)$ -compatible  
 339 (in this case  $f(A) = A$  for all  $A \subseteq X$  and in general  $A \cap B \neq \emptyset \not\Rightarrow A \subseteq B$ ,  
 340  $A, B \subseteq X$ ).

341 Similarly, let  $X$  and  $Y$  be two ordered spaces and let  $\mathbf{R}_{\leq PD}^X$  and  $\mathbf{R}_{\leq PD}^Y$  be the  
 342 relations corresponding to pairwise-bound dominance on  $X$  and on  $Y$ , respec-  
 343 tively. Then, not all functions  $f$  are  $(\mathbf{R}_{\leq PD}^X; \mathbf{R}_{\leq PD}^Y)$ -compatible. For instance, if  
 344  $f$  is decreasing, i.e.  $f(x) \leq f(x')$  for all  $x \in X$  and  $x' \in X$  such that  $x \geq x'$ ,  
 345 then we have  $f(A) \geq_{PD} f(B)$  for all  $A, B \subseteq X$  such that  $A \leq_{PD} B$ , and  
 346 thus  $f$  is not  $(\mathbf{R}_{\leq PD}^X; \mathbf{R}_{\leq PD}^Y)$ -compatible since in general we have in this case  
 347  $A \leq_{PD} B \not\Rightarrow f(A) \leq_{PD} f(B)$ . However, if  $f$  is monotonically non-decreasing,  
 348 then it is  $(\mathbf{R}_{\leq PD}^X; \mathbf{R}_{\leq PD}^Y)$ -compatible since if  $f(x) \leq f(x')$  for all  $x \in X$  and  
 349  $x' \in X$  such that  $x \leq x'$  then we have  $A \leq_{PD} B \Rightarrow f(A) \leq_{PD} f(B)$ .

350 **Proposition 12 (Preserved compatibility).** *If  $f$  is  $(\mathbf{R}^X; \mathbf{R}^Y)$ -compatible, it*  
 351 *is so on belief functions:*

$$m_1 \mathbf{R}^X m_2 \Rightarrow f(m_1) \mathbf{R}^Y f(m_2). \quad (12)$$

352 *Proof.* Since  $m_1 \mathbf{R}^X m_2$ , there exists a joint mass function  $m_{12}$  on  $X^2$  satisfy-  
 353 ing (6)-(8) for  $\mathbf{R}^X$ . Consider the joint mass function  $m$  on  $Y^2$  defined as

$$m(A', B') = \sum_{f(A)=A', f(B)=B'} m_{12}(A, B), \quad \forall A', B' \subseteq Y. \quad (13)$$

354 Since  $m_{12}(A, B) > 0 \Rightarrow \mathbf{A} \mathbf{R}^X B$  and  $\mathbf{A} \mathbf{R}^X B \Rightarrow f(A) \mathbf{R}^Y f(B)$ , then  $m(A', B') >$   
 355  $0 \Rightarrow A' \mathbf{R}^Y B'$ . Besides,

$$\begin{aligned} \sum_{B'} m(A', B') &= \sum_{B'} \sum \{m_{12}(A, B) | f(A) = A', f(B) = B'\} \\ &= \sum_{f(A)=A'} \sum_{B'} \sum_{f(B)=B'} m_{12}(A, B) \\ &= \sum_{f(A)=A'} \sum_B m_{12}(A, B) \\ &= \sum_{f(A)=A'} m_1(A) \\ &= (f(m_1))(A') \end{aligned}$$

356 Similarly,  $\sum_{A'} m(A', B') = (f(m_2))(B')$ .

In sum, when  $m_1 \mathbf{R}^X m_2$  and  $f$  is  $(\mathbf{R}^X; \mathbf{R}^Y)$ -compatible, there exists a joint mass function  $m$  on  $Y^2$  such that  $m(A', B') > 0 \Rightarrow A' \mathbf{R}^Y B'$ , for all  $A', B' \subseteq Y$ , and whose marginals are  $f(m_1)$  and  $f(m_2)$ , hence  $f(m_1) \mathbf{R}^Y f(m_2)$ .  $\square$

357 **Corollary 1.** For any function  $f$ ,  $m_1 \mathbf{R}_{\subseteq}^X m_2 \Rightarrow f(m_1) \mathbf{R}_{\subseteq}^Y f(m_2)$ , which follows  
358 from the  $(\mathbf{R}_{\subseteq}^X; \mathbf{R}_{\subseteq}^Y)$ -compatibility of any  $f$ .

359 Corollary 1 was already known [16, Proposition 2]. Proposition 12 is a generali-  
360 sation of this latter result.

## 361 5.2 Multivariate functions

362 These results can be extended to functions having more than one argument:

363 **Definition 3.** Let  $f : X_1 \times X_2 \rightarrow Y$ . Let  $\mathbf{R}^{X_1}$ ,  $\mathbf{R}^{X_2}$  and  $\mathbf{R}^Y$  be relations on  $2^{X_1}$ ,  
364  $2^{X_2}$  and  $2^Y$ , respectively. The function  $f$  is said to be  $(\mathbf{R}^{X_1}, \mathbf{R}^{X_2}; \mathbf{R}^Y)$ -compatible  
365 if, for all  $A_1, B_1 \subseteq X_1$  and all  $A_2, B_2 \subseteq X_2$

$$A_1 \mathbf{R}^{X_1} B_1 \wedge A_2 \mathbf{R}^{X_2} B_2 \Rightarrow f(A_1, A_2) \mathbf{R}^Y f(B_1, B_2). \quad (14)$$

366 *Example 4.* Since for any function  $f$  and any  $A_1, B_1 \subseteq X_1$  and  $A_2, B_2 \subseteq X_2$ ,  
367 such that  $A_1 \subseteq B_1$  and  $A_2 \subseteq B_2$  it holds that  $f(A_1, A_2) \subseteq f(B_1, B_2)$ , any  
368 function is  $(\mathbf{R}_{\subseteq}^{X_1}, \mathbf{R}_{\subseteq}^{X_2}; \mathbf{R}_{\subseteq}^Y)$ -compatible.

Let  $X_1$ ,  $X_2$  and  $Y$  be ordered spaces. If  $f$  is non-decreasing in both its arguments (for short, *non-decreasing*), i.e., for all  $(x_1, x_2), (x'_1, x'_2) \in X_1 \times X_2$ ,

$$x_1 \leq x'_1 \wedge x_2 \leq x'_2 \Rightarrow f(x_1, x_2) \leq f(x'_1, x'_2),$$

369 then for  $y \in \{LD, LB, UB, SD, PD\}$  we have  $f(A_1, A_2) \leq_y f(B_1, B_2)$  for all  
370  $A_1, B_1 \subseteq X_1$  and  $A_2, B_2 \subseteq X_2$  such that  $A_1 \leq_y B_1$  and  $A_2 \leq_y B_2$ , i.e.  $f$  is  
371  $(\mathbf{R}_{\leq_y}^{X_1}, \mathbf{R}_{\leq_y}^{X_2}; \mathbf{R}_{\leq_y}^Y)$ -compatible. This can easily be shown as follows (we provide  
372 only the proof for the case  $y = LD$ , the other cases being similar). Since  $f$  is non-  
373 decreasing, we have  $\min f(A_1, A_2) = f(\min A_1, \min A_2)$  and  $\max f(B_1, B_2) =$   
374  $f(\max B_1, \max B_2)$ . Besides, since  $\min A_1 \leq \max B_1$  and  $\min A_2 \leq \max B_2$  and  
375  $f$  is non-decreasing, we obtain  $\min f(A_1, A_2) \leq \max f(B_1, B_2)$ .

We remind that the image  $f(m_{12})$  of some joint mass function  $m_{12} \in \mathcal{B}^{X_1 \times X_2}$  under  $f$  is the mass function  $f(m_{12}) \in \mathcal{B}^Y$  defined, for all  $B \subseteq Y$ , as [16]:

$$(f(m_{12}))(B) = \sum_{f(A_1, A_2) = B} m_{12}(A_1, A_2).$$

376 Let us also recall that if  $m_{12}$  satisfies  $m_{12}(A_1, A_2) = m_1(A_1)m_2(A_2)$  with  $m_i$   
377 the marginal of  $m_{12}$  on  $X_i$ ,  $i = 1, 2$ , then  $m_1$  and  $m_2$  are said to be independent.  
378 This independence notion is the main one used in evidence theory, but can also be  
379 interpreted and used as an outer approximation within imprecise probability [18,  
380 5].

381 **Proposition 13 (Preserved compatibility, several arguments).** *Let  $m_{12}$*   
 382 *(resp.  $m'_{12}$ ) denote the joint mass function on  $X_1 \times X_2$  obtained from independent*  
 383 *mass functions  $m_1$  and  $m_2$  (resp.  $m'_1$  and  $m'_2$ ) defined on  $X_1$  and  $X_2$ , respectively.*

*If  $f$  is  $(\mathbf{R}^{X_1}, \mathbf{R}^{X_2}; \mathbf{R}^Y)$ -compatible, it is so on belief functions:*

$$m_1 \mathbf{R}^{X_1} m'_1 \wedge m_2 \mathbf{R}^{X_2} m'_2 \Rightarrow f(m_{12}) \mathbf{R}^Y f(m'_{12}).$$

384 *Proof.* Since  $m_i \mathbf{R}_i^X m'_i$ ,  $i = 1, 2$ , there exist joint mass functions  $m_{11'}$  and  $m_{22'}$   
 385 satisfying

$$\begin{aligned} m_{11'}(A_1, B_1) > 0 &\implies A_1 \mathbf{R}_1^X B_1, \\ m_{22'}(A_2, B_2) > 0 &\implies A_2 \mathbf{R}_2^X B_2. \end{aligned}$$

386 Furthermore, let  $m_{11'22'}$  denote the joint mass function on  $X_1 \times X_1 \times X_2 \times$   
 387  $X_2$  obtained from independent marginals  $m_{11'}$  and  $m_{22'}$ . Mass function  $m_{11'22'}$   
 388 satisfies

$$m_{11'22'}(A_1, B_1, A_2, B_2) > 0 \implies A_1 \mathbf{R}_1^X B_1 \wedge A_2 \mathbf{R}_2^X B_2. \quad (15)$$

389 Moreover, we have

$$\begin{aligned} \sum_{B_1, B_2} m_{11'22'}(A_1, B_1, A_2, B_2) &= \sum_{B_1, B_2} m_{11'}(A_1, B_1) m_{22'}(A_2, B_2) \\ &= \sum_{B_1} m_{11'}(A_1, B_1) \cdot \sum_{B_2} m_{22'}(A_2, B_2) \\ &= \sum_{B_1} m_{11'}(A_1, B_1) \cdot m_2(A_2) \\ &= m_1(A_1) \cdot m_2(A_2) \\ &= m_{12}(A_1, A_2) \end{aligned}$$

390 and similarly

$$\sum_{A_1, A_2} m_{11'22'}(A_1, B_1, A_2, B_2) = m'_{12}(B_1, B_2).$$

391 In other words,  $m_{11'22'}$  has  $m_{12}$  and  $m'_{12}$  as marginals.

Consider the joint mass function  $m$  on  $Y^2$  defined as, for any  $A', B' \subseteq Y$ ,

$$m(A', B') = \sum_{f(A_1, A_2)=A', f(B_1, B_2)=B'} m_{11'22'}(A_1, B_1, A_2, B_2).$$

392 Since Eqs. (15) and (14) hold, then  $m(A', B') > 0 \Rightarrow A' \mathbf{R}^Y B'$ . Besides,

$$\begin{aligned}
\sum_{B'} m(A', B') &= \sum_{B'} \sum_{B'} \{m_{11'22'}(A_1, B_1, A_2, B_2) | f(A_1, A_2) = A', f(B_1, B_2) = B'\} \\
&= \sum_{f(A_1, A_2)=A'} \sum_{B'} \sum_{f(B_1, B_2)=B'} m_{11'22'}(A_1, B_1, A_2, B_2) \\
&= \sum_{f(A_1, A_2)=A'} \sum_{B_1, B_2} m_{11'22'}(A_1, B_1, A_2, B_2) \\
&= \sum_{f(A_1, A_2)=A'} m_{12}(A_1, A_2) \\
&= (f(m_{12}))(A')
\end{aligned}$$

Similarly,  $\sum_{A'} m(A', B') = (f(m'_{12}))(B')$ . □

393 **Corollary 2.** For any function  $f$ ,  $m_1 \mathbf{R}_{\subseteq}^{X_1} m'_1 \wedge m_2 \mathbf{R}_{\subseteq}^{X_2} m'_2 \Rightarrow f(m_{12}) \mathbf{R}_{\subseteq}^Y f(m'_{12})$ ,  
394 which follows from any  $f$  being  $(\mathbf{R}_{\subseteq}^{X_1}, \mathbf{R}_{\subseteq}^{X_2}; \mathbf{R}_{\subseteq}^Y)$ -compatible.

395 **Corollary 3.** For  $X_1, X_2$  and  $Y$  ordered spaces and  $f$  any monotonically non  
396 decreasing function,  $m_1 \mathbf{R}_{\leq_y}^{X_1} m'_1 \wedge m_2 \mathbf{R}_{\leq_y}^{X_2} m'_2 \Rightarrow f(m_{12}) \mathbf{R}_{\leq_y}^Y f(m'_{12})$ , for  $y \in$   
397  $\{LD, LB, UB, SD, PD\}$ , which follows from the  $(\mathbf{R}_{\leq_y}^{X_1}, \mathbf{R}_{\leq_y}^{X_2}; \mathbf{R}_{\leq_y}^Y)$ -compatibility  
398 of any such functions for  $y \in \{LD, LB, UB, SD, PD\}$ .

399 Corollary 2 was already known [16, Proposition 3], and Corollary 3 generalises  
400 a result in [24, proof of Proposition 5], where  $f$  is the addition of integers and  
401  $y = PD$ . Proposition 13, which can be readily extended to the case of functions  
402 having more than two arguments, is thus a generalisation of these results of [16,  
403 24].

404 We note that the setting of Corollary 3, *i.e.*, monotonic non-decreasing func-  
405 tions on ordered spaces, is commonly encountered in multi-criteria decision mak-  
406 ing [23, 10], reliability analysis [13], and optimisation problems [21] hence this  
407 corollary may be useful for such problems when function arguments are tainted  
408 with uncertainty. Next section provides such examples.

## 409 6 Illustrative applications

### 410 6.1 System reliability

411 In multi-state system reliability assessment, one main issue is to assess the avail-  
412 ability, or the performance of a whole system, given the performances of its  
413 components. Usually, the system is assumed to have  $n$  components  $x_i$  that take  
414 values on a finite, ordered scale  $X_i$ , that we will denote here by natural num-  
415 bers. The performance of the system then depends on the state of each of its  
416 component, and is usually modelled by structure function  $f(x_1, \dots, x_n)$  that is  
417 non-decreasing, as the system performance can only increase or stay the same if  
418 a component performance increases.

Let us consider a simple communication system made of one source and one receiver, with a channel in between made of  $n$  repeaters  $x_i$ , where each of them can be in different states  $X_i = \{1, 2, \dots, K\}$  which is the maximal number of messages this repeater can store and send. Given this, we may be interested in the global capacity of a channel, which is

$$f(x_1, \dots, x_n) = \min(x_1, \dots, x_n),$$

419 as this is a series system. Now, assume we want to compare the capacity of two  
 420 different channels in order to choose the best one, with repeaters whose state is  
 421 uncertain due to the fact that they have degraded over time. If this uncertainty is  
 422 modelled by belief functions and that  $m_i^j$  models the state of the  $i^{\text{th}}$  repeater of  
 423 the  $j^{\text{th}}$  channel, then if we have  $m_i^1 \mathbf{R}_{\leq y} m_i^2$  for any  $y \in \{LD, LB, UB, SD, PD\}$ ,  
 424 then we know from Corollary 3 that channel 1 achieves at most the same level  
 425 of performance than channel 2, without even computing the propagated mass  
 426 function through  $f$ . Note that this would be true, whatever the function  $f$  is (as  
 427 long as it is non-decreasing).

*Example 5.* Assume we have four identical repeaters  $x_1, x_2, x_3$  and  $x_4$  working independently with  $X_i = \{1, 2, 3\}$ , and are considering two different, partially known technologies at our disposal to build the system. Then, if our knowledge of these two technologies is such that,  $\forall i$

$$\begin{aligned} m_i^1(\{2\}) &= 0.5 & m_i^2(\{3\}) &= 0.5 \\ m_i^1(\{1, 2\}) &= 0.3 & m_i^2(\{2, 3\}) &= 0.3 \\ m_i^1(\{1, 2, 3\}) &= 0.2 & m_i^2(\{1, 2, 3\}) &= 0.2 \end{aligned}$$

428 we can easily conclude that  $f^1 \mathbf{R}_{\leq PD} f^2$  and this without performing any com-  
 429 putation, as  $m_i^1 \mathbf{R}_{\leq PD} m_i^2$ .

## 430 6.2 Multi-criteria decision making

431 In multi-criteria decision making, it is quite common to consider as variables  
 432  $x_i$  the utilities of the criteria. For instance, these could be scores obtained by  
 433 students over different courses, or the evaluation of students regarding the qual-  
 434 ity of courses (e.g., with respect to interest, teaching quality and study time  
 435 required). It could be that, for some reasons, those assessments are uncertain  
 436 (e.g., because students are allowed to provide imprecise assessments in case of  
 437 hesitation).

438 One common function to aggregate utilities is the weighted average, that is  
 439 to have

$$f(x_1, \dots, x_n) = \sum_{i=1}^n w_i x_i \quad (16)$$

440 or one of its extension such as the Choquet integral. Again, if we want to compare  
 441 two courses, and we have  $m_i^1 \mathbf{R}_{\leq PD} m_i^2$  for all  $x_i$ , then we know from Corollary 3,

442 without any computation, that the second course is considered better by the  
 443 students. Besides, if there is a third course such that  $m_i^2 \mathbf{R}_{\leq PD} m_i^3$  for all  $x_i$ ,  
 444 then, without any computation, we know from Corollary 3 that this course is  
 445 preferred over the second one but also over the first one as the set relation  $\mathbf{R}_{\leq PD}$   
 446 is transitive, a property that we know is preserved thanks to Proposition 7.

*Example 6.* Assume three courses evaluated by students against two criteria  $x_1$  and  $x_2$  with  $X_i = \{0, \dots, 10\}$ . Let  $m_i^j$  be the mass function representing the uncertain evaluation of course  $j$  according to criterion  $i$ . Suppose we have for criterion  $x_1$

$$m_1^1(\{2, 3, 4\}) = 1, \quad m_1^2(\{2, 3, 4\}) = 0.6, \quad m_1^3(\{6, 7, 8\}) = 1, \\ m_1^2(\{3, 4\}) = 0.4,$$

and for criterion  $x_2$

$$m_2^1(\{4\}) = 0.13, \quad m_2^2(\{5, 6\}) = 0.65, \quad m_2^3(\{5, 6, 7\}) = 0.3, \\ m_2^1(\{7\}) = 0.07, \quad m_2^2(\{7\}) = 0.35, \quad m_2^3(\{7, 8\}) = 0.2, \\ m_2^1(\{5, 6\}) = 0.8, \quad m_2^3(\{8\}) = 0.5.$$

If the weighted average (16) is used to aggregate these evaluations, then whatever the weights  $w_i$ , we can easily conclude that the overall uncertain score  $f^1$  of the first course will be such that  $f^1 \mathbf{R}_{\leq PD} f^2$ , with  $f^2$  the uncertain score of the second course, since  $m_i^1 \mathbf{R}_{\leq PD} m_i^2$  for all  $x_i$ . Indeed, for the first criterion the only joint mass function obtainable from  $m_1^1$  and  $m_1^2$  is

$$m_1^{12}(\{2, 3, 4\}, \{2, 3, 4\}) = 0.6, \quad m_1^{12}(\{2, 3, 4\}, \{3, 4\}) = 0.4$$

and we can easily see that  $\{2, 3, 4\} \mathbf{R}_{\leq PD} \{2, 3, 4\}$  and  $\{2, 3, 4\} \mathbf{R}_{\leq PD} \{3, 4\}$ . For the second criterion, we can consider the joint mass function

$$m_2^{12}(\{5, 6\}, \{5, 6\}) = 0.65, \quad m_2^{12}(\{5, 6\}, \{7\}) = 0.15, \\ m_2^{12}(\{7\}, \{7\}) = 0.07, \quad m_2^{12}(\{4\}, \{7\}) = 0.13.$$

447 where every pair of sets satisfy the relation  $\mathbf{R}_{\leq PD}$ . From those two facts and  
 448 Corollary 3, we can conclude that  $f^1 \mathbf{R}_{\leq PD} f^2$  for any increasing function of the  
 449 two criteria. Similarly, we obtain  $f^2 \mathbf{R}_{\leq PD} f^3$  and  $f^1 \mathbf{R}_{\leq PD} f^3$ . Since  $\mathbf{R}_{\leq PD}$  is  
 450 transitive, the latter comparison could have been deduced from  $f^1 \mathbf{R}_{\leq PD} f^2$  and  
 451  $f^2 \mathbf{R}_{\leq PD} f^3$ .

As an illustration, let us confirm that for the first two courses and the simple weighted average with  $w_1 = 0.5$  and  $w_2 = 0.5$ , we do have  $f^1 \mathbf{R}_{\leq PD} f^2$ : denoting by  $m_f^j$  the propagated evaluation of course  $j$ , we get

$$m_f^1(\{3, 3.5, 4\} = A_1) = 0.13, \quad m_f^2(\{3.5, 4, 4.5, 5\} = B_1) = 0.39 \\ m_f^1(\{4.5, 5, 5.5\} = A_2) = 0.07, \quad m_f^2(\{4.5, 5, 5.5\} = B_2) = 0.21 \\ m_f^1(\{3.5, 4, 4.5, 5\} = A_3) = 0.8, \quad m_f^2(\{4, 4.5, 5\} = B_3) = 0.26 \\ m_f^2(\{5, 5.5\} = B_4) = 0.14$$

With the following matrix summarising the pairs of sets where the relation  $\mathbf{R}_{\leq PD}$  holds

$$\begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} \begin{bmatrix} B_1 & B_2 & B_3 & B_4 \\ \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} \\ & \mathbf{R}_{\leq PD} & & \mathbf{R}_{\leq PD} \\ \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} & \mathbf{R}_{\leq PD} \end{bmatrix}.$$

452 This means that any joint mass function where  $m(A_2, B_1)$  and  $m(A_2, B_3)$  are  
 453 null satisfy our definition, and it is easy to see that such a mass function exists,  
 454 for example by taking  $m(A_2, B_2) = 0.07$ .

### 455 6.3 Equivalence relations in taxonomies

456 Let us now assume that the elements of space  $X$  are concepts linked together by  
 457 a taxonomy, modelled as a rooted tree. Then, one possible question about two  
 458 uncertain observations of such a taxonomy is whether they belong to the same  
 459 general sub-concept of interest, or in other words whether they belong to the  
 460 same branch of the tree. In practice, this comes down to define a corresponding  
 461 partition  $C_1, \dots, C_K$  of  $X$ , and to say that  $\mathbf{A}\mathbf{R}\mathbf{B}$  iff  $A \cup B \subseteq C_i$  for some  $i$ .

*Example 7.* Assume we have the space  $X = \{(M)otorcycle, (T)ruck, (C)at, (D)og\}$  together with the taxonomy provided by Figure 2. The partition defined by the concepts of the first level (Vehicle and Animal) is  $C_1 = \{M, T\}$  and  $C_2 = \{C, D\}$ . We could then wonder whether two uncertain objects belong to the same category, given this granularity. For instance, consider the three mass functions

$$\begin{aligned} m_1(\{C\}) &= 0.6, & m_2(\{T, C\}) &= 0.2, & m_3(\{D\}) &= 0.4, \\ m_1(\{D, C\}) &= 0.4, & m_2(X) &= 0.8, & m_3(\{D, C\}) &= 0.6. \end{aligned}$$

We do have  $m_1\mathbf{R}m_3$ , as  $\{C\}\mathbf{R}\{D\}\mathbf{R}\{D, C\}$ , but not  $m_1\mathbf{R}m_2$  nor  $m_3\mathbf{R}m_2$ , concluding that while the first and third objects belong to the same category,  $m_2$  does not. To see that  $m_1\mathbf{R}m_3$ , one can consider the joint mass function

$$m_{13}(\{D, C\}, \{D\}) = 0.4, \quad m_{13}(\{C\}, \{D, C\}) = 0.6,$$

462 and to see that  $\neg m_1\mathbf{R}m_2$  and  $\neg m_3\mathbf{R}m_2$ , it is sufficient to observe that  $\neg X\mathbf{R}A$   
 463 for any  $A$ , and that the mass  $m_2(X)$  is strictly positive, hence that some joint  
 464 mass must be given to it. Since we know that  $\mathbf{R}$  is an equivalence relation,  
 465 that is preserved when considering belief functions, we could have deduced from  
 466  $\neg m_1\mathbf{R}m_2$  that  $\neg m_3\mathbf{R}m_2$ .

467 Such notions could be used for example in formal concept analysis [22], where  
 468 we may want to know the most specific common concept to which two uncertain  
 469 objects belong. Another possibility includes for instance hierarchical classifica-  
 470 tion [1], in which a usually very large number of classes are structured according  
 471 to a taxonomy that is used to find the right class (leaf) of an object. Being able to  
 472 tell whether two classifiers agree that a particular instance belong to a common  
 473 sub-tree may then be a helpful item of information.



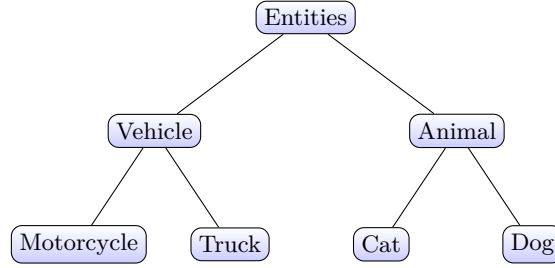


Fig. 2. Taxonomy of Example 7

## 474 7 From binary to gradual relations

475 So far, we have been concerned with the problem of assessing whether or not  
 476 two mass functions  $m_1, m_2$  were in relation, viewing this as a binary value that  
 477 could only takes values 0 ( $\neg m_1 \mathbf{R} m_2$ ) or 1 ( $m_1 \mathbf{R} m_2$ ). It can be interesting to  
 478 relax this assumption by allowing the relation to be gradual, that is to take any  
 479 value between 0 and 1.

An easy way to do that is to follow an optimistic principle and to say that  
 for any two mass functions  $m_1, m_2$ , the degree  $\alpha_{\mathbf{R}}$  to which  $m_1$  is in relation  $\mathbf{R}$   
 with  $m_2$  is the solution of the optimisation problem

$$1 - \alpha_{\mathbf{R}} = \min \sum_{A, B: \neg(ARB)} m_{12}(A, B)$$

$$m_1(A) = \sum_B m_{12}(A, B),$$

$$m_2(B) = \sum_A m_{12}(A, B).$$

480 This generalises the approach taken so far, as we will have  $m_1 \mathbf{R} m_2$  iff  $\alpha_{\mathbf{R}} = 1$ ,  
 481 that is if the degree to which they are in relation is maximal. Conversely,  $\alpha_{\mathbf{R}} = 0$   
 482 iff there is no pair of subsets  $A, B$  having positive mass such that  $ARB$ .

483 For instance, if we consider again the “having a lower cardinality than”  
 484 example from the proof of Proposition 8, we would have that  $m_1 \alpha_{\mathbf{R}} m_2$  with  
 485  $\alpha_{\mathbf{R}} = 1 - 0.2 = 0.8$ , a quite high value. Such gradual relations have been pro-  
 486 posed in the past, for example the conflict measure  $\kappa_m^2$  proposed in [11] is nothing  
 487 else but the solution of the optimisation problem applied to the relation  $ARB$   
 488 iff  $A \cap B \neq \emptyset$ .

489 Studying the properties and implications of using such gradual relations in  
 490 detail goes out of the scope of the current paper, yet a clear first step would be  
 491 to relate such a gradual view to the large literature concerning fuzzy relations.  
 492 Indeed, fuzzy relations are also  $[0, 1]$ -valued, and researchers of this field have  
 493 come with various proposals of how classical properties can be extended to this  
 494 case, e.g., to deal with fuzzy preferences [20, 14] or fuzzy equivalence relations [2].

## 495 8 Conclusions

496 In this paper, a universal way to generalise a binary relation from sets to belief  
 497 functions is introduced. Several results are provided showing which properties of  
 498 the relation are preserved through this mechanism, including its compatibility  
 499 with functions. Our proposal is also connected to more specific generalisation of  
 500 binary relations such as the notion of specialisation. Consequently, our results  
 501 are also a generalisation of pre-existing ones for specific relations.

502 There are however several questions that remain to address. A first one is to  
 503 consider not relations on the same space, but more general relations on different  
 504 spaces, including compositions of such relations.

505 Finally, we have also proposed a way to transform the initial binary relation  
 506 on sets into a gradual relation on belief functions. This also opens up a whole  
 507 avenue of research, as this directly connect our proposal to the various notions of  
 508 fuzzy relations, that consist in providing a number in the unit interval reflecting  
 509 how much a relation holds. Performing such a study goes beyond the actual  
 510 scope of the present paper.

## 511 acknowledgment

512 This work was carried out in the framework of the Labex MS2T, funded by  
 513 the French Government, through the National Agency for Research (Reference  
 514 ANR-11-IDEX-0004-02). We also want thank the attendants of BELIEF 2018,  
 515 in particular Thierry Denoeux, Prakash Shenoy and Didier Dubois, for their  
 516 insightful questions and comments that have helped to improve this paper.

## 517 A Proof of Proposition 10

518 In this appendix, we give a proof that if  $\mathbf{R}$  is a partial order on sets then the  
 519 mechanism described in Definition 1 yields a partial order on belief functions. We  
 520 already know that the induced relation on belief functions inherits the reflexivity  
 521 and transitivity properties (c.f. Propositions 5 and 7). So we only need to prove  
 522 that antisymmetry holds.

Suppose there exist two mass functions  $m_1$  and  $m_2$  such that  $m_1 \mathbf{R} m_2$  and  
 $m_2 \mathbf{R} m_1$ . This means that there are two matrices  $S_1$  and  $S_2$  compliant with  
 Definition 1 and such that  $m_1 = S_1 \cdot m_2$  and  $m_2 = S_2 \cdot m_1$  (mass functions are  
 seen as vectors here). By plugging these two equations together, we obtain

$$m_2 = S_2 \cdot S_1 \cdot m_2.$$

523 To complete the proof, we will be needing the following intermediate result:

524 **Lemma 1.** *If  $\mathbf{R}$  is a partial order on sets, then there is an indexation of sub-*  
 525 *sets of  $X$  such that for any pair of mass functions  $(m_1; m_2)$  with  $m_1 \mathbf{R} m_2$ , the*  
 526 *stochastic matrix  $S$  satisfying Definition 1 is upper triangular.*

*Proof.* Define a refinement  $\mathbf{R}_*$  of  $\mathbf{R}$  such that  $\mathbf{R}_*$  is a total order (subsets that cannot be ordered using  $\mathbf{R}$  can be ordered in an arbitrary way). Let  $N$  denote the cardinality of  $2^X$ . Let  $(A_i)_{i=1}^N$  denote all subsets of  $X$  indexed using  $\mathbf{R}_*$  so that

$$A_i \mathbf{R}_* A_j \Leftrightarrow i \leq j.$$

527 Now suppose  $S_*$  is a stochastic matrix compliant with Definition 1 for relation  
528  $\mathbf{R}_*$  for some pair of mass functions  $m_1$  and  $m_2$ . If  $m_2(A_j) > 0$ , we need to have  
529  $S_*(A_i, A_j) = 0$  whenever  $\neg(A_i \mathbf{R}_* A_j) \Leftrightarrow i > j$ .

Suppose  $S$  is a stochastic matrix compliant with Definition 1 for relation  $\mathbf{R}$  for the same pair of mass functions  $m_1$  and  $m_2$ . If  $m_2(A_j) = 0$ , we can reassign entries of the  $j^{\text{th}}$  column of  $S$  as we want while remaining compliant with Definition 1. For instance, we can set

$$S(A_i, A_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Provided that the above reassignment is completed, the matrix  $S$  is upper triangular because when  $m_2(A_j) > 0$ ,

$$i > j \Leftrightarrow \neg(A_i \mathbf{R}_* A_j) \Rightarrow \neg(A_i \mathbf{R} A_j) \Rightarrow S(A_i, A_j) = 0.$$

□

530 Based on the above lemma, we can require that  $S_1$  and  $S_2$  are upper triangu-  
531 lar and consequently  $S_{21} = S_2 \cdot S_1$  as well. Since left stochasticity is preserved by  
532 matrix product, we know that  $S_{21}$  is also left stochastic. Consequently, we have  
533 that  $S_{21}$  coincides with the identity matrix  $I$  on every column corresponding to  
534 a focal element of  $m_2$ . In other words, if  $m_2(A_j) > 0$ , then

$$S_{21}(A_i, A_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}, \quad (17)$$

where  $(A_i)_{i=1}^N$  are all subsets of  $X$  ordered in the way arising from the definitions of matrices  $S_1$  and  $S_2$  and their upper triangularities. To prove this, observe that

$$m_2(A_i) = \sum_{j=1}^N S_{21}(A_i, A_j) m_2(A_j), \quad (18)$$

$$= \sum_{j=i}^N S_{21}(A_i, A_j) m_2(A_j) \quad (19)$$

Let  $A_{k_1}$  denote the focal element of  $m_2$  with maximal index, i.e.  $A_i$  is not a focal element of  $m_2$  if  $i > k_1$ . We necessarily have that

$$m_2(A_{k_1}) = S_{21}(A_{k_1}, A_{k_1}) m_2(A_{k_1}).$$

535 This obviously implies that  $S_{21}(A_{k_1}, A_{k_1}) = 1$  and that all other entries of the  
 536 column corresponding to  $A_{k_1}$  are null because  $S_{21}$  is left stochastic. Now, let  $A_{k_2}$   
 537 be the focal element of  $m_2$  with the second maximal index, i.e. if  $i > k_2$ , then  
 538  $A_i$  is either  $A_{k_1}$  or not a focal element of  $m_2$ . We have now

$$m_2(A_{k_2}) = S_{21}(A_{k_2}, A_{k_2})m_2(A_{k_2}) + S_{21}(A_{k_2}, A_{k_1})m_2(A_{k_1}).$$

539  $S_{21}(A_{k_2}, A_{k_1})$  is in the  $k_1^{\text{th}}$  column of  $S_{21}$  therefore we deduce that  $S_{21}(A_{k_2}, A_{k_2}) =$   
 540 1 and that all other entries of the corresponding column are null. We can continue  
 541 to iterate on the focal elements of  $m_2$  to obtain (17).

Furthermore, the upper triangularities of  $S_1$  and  $S_2$  give that  $S_{21}(A_i, A_i) =$   
 $S_2(A_i, A_i) \times S_1(A_i, A_i), \forall i$ . When  $m_2(A_i) > 0$ , we know that  $S_{21}(A_i, A_i) =$   
 1 and we deduce that  $S_1(A_i, A_i) > 0$  and  $S_2(A_i, A_i) > 0$ . From the upper  
 triangularity of the matrices, we also have

$$S_{21}(A_{i-1}, A_i) = S_2(A_{i-1}, A_{i-1})S_1(A_{i-1}, A_i) + S_2(A_{i-1}, A_i)S_1(A_i, A_i).$$

542 When  $m_2(A_i) > 0$ , we know that  $S_{21}(A_{i-1}, A_i) = 0$  and we deduce that  
 543  $S_1(A_{i-1}, A_i) = 0$ . We can iterate and show that  $S_1(A_{i-k}, A_i) = 0$  for any  
 544  $k \in \{1; \dots; i-1\}$  which in turn implies that  $S_1(A_i, A_i) = 1$  because  $S_1$  is  
 545 stochastic. Since  $m_1 = S_1 \cdot m_2$ , we see that, for any focal element  $A$  of  $m_2$ , we  
 546 have  $m_1(A) \geq m_2(A)$ . Finally, we necessarily have that  $m_1 = m_2$  because the  
 547 masses of  $m_1$  need to sum to one.

## 548 References

- 549 1. Bengio, S., Weston, J., Grangier, D.: Label embedding trees for large multi-class  
 550 tasks. In: Advances in Neural Information Processing Systems. pp. 163–171 (2010)
- 551 2. Boixader, D., Jacas, J., Recasens, J.: Fuzzy equivalence relations: advanced mater-  
 552 ial. In: Fundamentals of Fuzzy Sets, pp. 261–290. Springer (2000)
- 553 3. Chateauneuf, A.: Combination of compatible belief functions and relation of speci-  
 554 ficity. In: Advances in the Dempster-Shafer theory of evidence. pp. 97–114. John  
 555 Wiley & Sons, Inc. (1994)
- 556 4. Choquet, G.: Theory of capacities. In: Annales de l’institut Fourier. vol. 5, pp.  
 557 131–295 (1954)
- 558 5. Couso, I., Moral, S., Walley, P.: A survey of concepts of independence for imprecise  
 559 probabilities. Risk, Decision and Policy 5(2), 165–181 (2000)
- 560 6. Dempster, A.: Upper and lower probabilities induced by a multiple valued mapping.  
 561 Annals of Mathematical Statistics 38, 325 – 339 (1967)
- 562 7. Denoeux, T., Dubois, D., Prade, H.: Representations of uncertainty in artificial  
 563 intelligence: Beyond probability and possibility. In: Marquis, P., Papini, O., Prade,  
 564 H. (eds.) A Guided Tour of Artificial Intelligence Research. Springer Verlag, to  
 565 appear (2018)
- 566 8. Denœux, T.: Extending stochastic ordering to belief functions on the real line.  
 567 Information Sciences 179(9), 1362–1376 (2009)
- 568 9. Destercke, S., Dubois, D., Chojnacki, E.: Unifying practical uncertainty represen-  
 569 tations: I. generalized p-boxes. International Journal of Approximate Reasoning  
 570 49, 649–663 (2008)

- 571 10. Destercke, S.: A generic framework to include belief functions in preference handling and multi-criteria decision. *International Journal of Approximate Reasoning* 98, 62–77 (2018)
- 572
- 573
- 574 11. Destercke, S., Burger, T.: Toward an axiomatic definition of conflict between belief functions. *IEEE transactions on cybernetics* 43(2), 585–596 (2013)
- 575
- 576 12. Destercke, S., Dubois, D.: Other uncertainty theories based on capacities. In: Augustin, T., Coolen, F., Cooman, G.D., Troffaes, M. (eds.) *Introduction to Imprecise Probabilities*, pp. 93–112. Wiley series in probability and statistics, Wiley (May 2014), <https://hal.archives-ouvertes.fr/hal-00999503>
- 577
- 578
- 579
- 580 13. Destercke, S., Sallak, M.: An extension of universal generating function in multi-state systems considering epistemic uncertainties. *IEEE Transactions on reliability* 62(2), 504–514 (2013)
- 581
- 582
- 583 14. Díaz, S., De Baets, B., Montes, S.: General results on the decomposition of transitive fuzzy relations. *Fuzzy Optimization and Decision Making* 9(1), 1–29 (2010)
- 584
- 585 15. Dubois, D., Prade, H.: A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems* 12(3), 193–226 (1986)
- 586
- 587
- 588 16. Dubois, D., Prade, H.: Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems* 42, 87–101 (1991)
- 589
- 590 17. Dubois, D., Prade, H.: Consonant approximations of belief functions. *International Journal of Approximate Reasoning* 4, 419 – 449 (1990)
- 591
- 592 18. Fetz, T., Oberguggenberger, M.: Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering & System Safety* 85(1-3), 73–87 (2004)
- 593
- 594
- 595 19. Fishburn, P.C.: Interval graphs and interval orders. *Discrete mathematics* 55(2), 135–149 (1985)
- 596
- 597 20. Fodor, J., De Baets, B.: Fuzzy preference modelling: fundamentals and recent advances. In: *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, pp. 207–217. Springer (2008)
- 598
- 599
- 600 21. Fortin, J., Zieliński, P., Dubois, D., Fargier, H.: Criticality analysis of activity networks under interval uncertainty. *Journal of Scheduling* 13(6), 609–627 (2010)
- 601
- 602 22. Ganter, B., Wille, R.: *Formal concept analysis: mathematical foundations*. Springer Science & Business Media (2012)
- 603
- 604 23. Grabisch, M., Labreuche, C.: A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Annals of Operations Research* 175(1), 247–286 (2010)
- 605
- 606
- 607 24. Helal, N., Pichon, F., Porumbel, D., Mercier, D., Lefèvre, E.: The capacitated vehicle routing problem with evidential demands. *International Journal of Approximate Reasoning* 95, 124–151 (2018)
- 608
- 609
- 610 25. Montes, I., Miranda, E., Montes, S.: Stochastic dominance with imprecise information. *Computational Statistics & Data Analysis* 71, 868–886 (2014)
- 611
- 612 26. Nguyen, H.T.: On random sets and belief functions. *Journal of Mathematical Analysis and Applications* 65(3), 531 – 542 (1978)
- 613
- 614 27. Shafer, G.: *A mathematical theory of evidence*. Princeton University Press (1976)
- 615 28. Smets, P.: The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning* 31, 1–30 (2002)
- 616
- 617 29. Smets, P., Kennes, R.: The transferable belief model. *Artificial Intelligence* 66(2), 191 – 234 (1994)
- 618
- 619 30. Yager, R.R.: The entailment principle for Dempster-Shafer granules. *International Journal of Intelligent Systems* 1(4), 247–262 (1986)
- 620