Belief functions on MV-algebras of fuzzy sets: an overview

Tommaso Flaminio[1], Lluís Godo[2], Tomáš Kroupa[3]

[1] Department of Theoretical and Applied Science (DiSTA)

University of Insubria, Via Mazzini 5 - 21100 Varese, Italy

tommaso.flaminio@uninsubria.it

[2] IIIA Artificial Intelligence Research Institute (CSIC)

Campus UAB s/n, Bellaterra 08193, Spain

godo@iiia.csic.es

[3] Institute of Information Theory and Automation of the ASCR

Pod Vodárenskou věží 4, 182 08 Prague, Czech Republic

kroupa@utia.cas.cz

Abstract

Belief functions are the measure theoretical objects Dempster-Shafer evidence theory is based on. They are in fact totally monotone capacities, and can be regarded as a special class of measures of uncertainty used to model an agent's degrees of belief in the occurrence of a set of events by taking into account different bodies of evidence that support those beliefs. In this chapter we present two main approaches to extending belief functions on Boolean algebras of events to MV-algebras of events, modelled as fuzzy sets, and we discuss several properties of these generalized measures. In particular we deal with the normalization and soft-normalization problems, and on a generalization of Dempster's rule of combination.

1   Introduction and motivations

Dempster-Shafer theory of evidence [13, 50] is a generalization of Bayesian probability theory in which degrees of uncertainty are evaluated by belief functions, rather than by probability measures. Belief functions [50, 52] can be regarded as a special class of measures of uncertainty used to represent an agent's degrees of confidence in the occurrence of events of interest by taking into account different bodies of evidence that support these beliefs [50]. Such evidence plays a pivotal role in determining the agent's beliefs. Indeed, as we will recall in a while, although any belief function on the Boolean algebra 2^X of subsets of a finite set X might be seen as a particular probability, its associated distribution (called *mass* in Dempster-Shafer theory) maps the whole algebra 2^X into

1

$[0, 1]$, and not only its atoms. Every set $Y \subseteq X$ with a strictly positive mass represents a particular body of evidence and is called a *focal element*.

Given the relevance that Dempster-Shafer theory has in real-world situations, we may argue that usually, when a person is asked to provide an evidence about a fact she were witness of, her description of the facts would be affected not only by uncertainty regarding the statements, but also by a possible *imprecision* in the statements themselves. Therefore the classical framework would be insufficient to model the analysis provided by a witness.

Fuzzy sets were introduced by Zadeh [57] as an extension of classical sets: given a referential set X, a fuzzy subset of X can be identified with a function f from X into the real unit interval $[0, 1]$. Given a fuzzy sets f, the idea is to interpret, for every $x \in X$, the value $f(x) \in [0, 1]$ as the *degree of membership* of x to f. Adopting this interpretation, fuzzy set theory has become a basic mathematical model to represent *imprecision* and *vagueness*, but of course many other different interpretations are also possible. A typical example which explains how fuzzy sets can be used in order to deal with imprecision is about the *height* of a person (we will turn back to this example in the last section of this chapter). Indeed, when we are asked whether an individual x belongs to the set of *tall* persons, the classical truth values 1 (true) and 0 (false) might be insufficient. On the other hand, values in the real unit interval $[0, 1]$ provide a wider spectrum with which one can evaluate to what extent an individual can be considered as tall. In this prospective the *fuzzy* set of tall persons becomes a function $\mu_{tall} : X \to [0, 1]$ from the set X of individuals to $[0, 1]$, assigning to each individual x its degree $\mu_{tall}(x)$ of being tall. We refer the reader e.g. to [20, 32, 39] for monographs on the topic.

In the literature several attempts to extend belief functions on fuzzy events can be found. The first extensions of Dempster-Shafer theory to the general framework of fuzzy set theory was proposed by Zadeh in the context of information granularity and possibility theory [59] in the form of an expected conditional necessity, and by Smets who proposed in [51] to extend a classical belief function Bel on 2^X to fuzzy subsets A of X as the lower expectation of the characteristic function of A with respect to the class of probability measures lower bounded by Bel. After Zadeh and Smets, several further generalizations were proposed depending on the way a measure of *inclusion among fuzzy sets* is used to define the belief functions of fuzzy events based on fuzzy evidence. Indeed, given a mass assignment m for the bodies of evidence $\{A_1, A_2, \ldots\}$, and a measure $I(A \subseteq B)$ of inclusion among fuzzy sets, the belief of a fuzzy set B can be defined in general by the value: $Bel(B) = \sum_i I(A_i \subseteq B) \cdot m(A_i)$. We refer the reader to [37, 55, 56] for exhaustive surveys, and to [2] for another approach through fuzzy subsethood.

Belief functions on distributive lattices were studied in [33] and [61] where the authors define, starting from a given mass assignment $m : L \to [0, 1]$, the belief degree $Bel(a)$ of an element a of a distributive lattice L to be $Bel(a) = \sum_{x \in I_L, x \leq a} m(x)$, where I_L denotes the set of all join-irreducible elements of L. Notice that, although the framework of distributive lattices is much more general than the framework we are going to discuss in this chapter (we invite

the reader to consult Section 3), the inclusion operator used in [33, 61] is crisp and hence it does not take into account a graded notion of inclusion.

Different definitions were also introduced by Dubois and Prade [19] and by Denœux [15, 16] to deal with belief functions ranging over intervals or fuzzy numbers.

Of course, moving from classical sets to fuzzy sets conveys non trivial complications in the description of the algebraic model aimed at representing the available evidence. In particular, although we start with a finite set X, while the Boolean algebra 2^X is finite and hence atomic, the class $[0,1]^X$ of the fuzzy subsets $f : X \to [0,1]$ of X contains uncountably many elements and hence it is not always trivial to define a mass m over them. Moreover there are several different ways to generalize Boolean algebras to algebras of fuzzy sets. Usually the generalization of belief functions to this frame, is done in the algebraic context of the so called *De Morgan triples* (or *Zadeh-algebras*) over classes of fuzzy sets, and where intersection, union, and complementation, are replaced in $[0,1]^X$ by the pointwise extensions of the operations in $[0,1]$ of min, max, and standard negation $\neg : x \mapsto 1 - x$ respectively (see for instance [60, §2.1]).

In [25, 41, 42], to further generalize belief functions on fuzzy sets, the authors frame their investigation in the algebraic setting of MV-algebras [7, 45] (in fact in every MV-algebra a Zadeh-algebra is obtainable as a reduct) and, since a belief function can be equivalently represented by a probability measure $P_m : 2^{2^X} \to [0,1]$ such that $P(\{\emptyset\}) = 0$, they replace the usual probability measures by the notion of *state* on MV-algebras [44]. Indeed MV-algebras are the algebraic structures for fuzzy sets enabling the most natural treatment of many-valued probability theory. The reason is in the formula for a probability of a fuzzy event proposed already by Zadeh [58]. His definition — the expected value of the membership function of the fuzzy set w.r.t. a probability measure on its domain — later turned out to be a consequence of the axiomatic treatment of MV-probability. We will provide a more detailed introduction about these topics in Subsection 3.1.

In this chapter we will survey recent developments on belief functions on MV-algebras of fuzzy sets, mainly following the lines of the already above cited papers [25, 28, 41, 42]. The paper is organized as follows. In Section 2 we recall how belief functions are defined on Boolean algebras and in particular we will present a first definition based on mass assignments, and a second (equivalent) one based on probability measures. Then, in Section 3 we introduce the main algebraic structures we will need along the paper, namely MV-algebras. In the same section we also introduce states on MV-algebra (in Subsection 3.1) and we recall some basic results we are going to use later. Section 4 contains two main approaches to define belief functions on MV-algebras: we will deal with a first approach in which belief functions on fuzzy sets are built up over crisp, Boolean focal elements (Subsection 4.1), and a second, more general approach, in which belief functions on fuzzy sets are defined in a way to allow for focal elements to be fuzzy as well (Subsection 4.2). Belief functions on MV-algebras are not necessarily *normalized* measures, in the sense of the belief of the empty

set being zero. We will discuss the normalization problem in Section 5 and the case of a *soft-normalization* of mass assignments, and hence of belief functions, in Subsection 6. Then in Section 7 we present a generalization of Dempster rule of combination, we discuss some particular cases and we provide an example aimed at clarifying the use of such a construction in the general frame of fuzzy sets. We will end this chapter with Section 8 where we present some concluding remarks and we also suggest alternative readings about the topic.

2   Belief functions on Boolean algebras

Consider a finite set X whose elements can be regarded as mutually exclusive (and exhaustive) propositions of interest, and whose powerset 2^X represents all combinations of such propositions. The set X is usually called the *frame of discernment*, and every element $x \in X$ represents the lowest level of discernible information we can deal with.

A map $m : 2^X \to [0,1]$ is said to be a *basic belief assignment*, or a *mass assignment* whenever

$$m(\emptyset) = 0 \text{ and } \sum_{A \in 2^X} m(A) = 1.$$

Given such a mass assignment m on 2^X, for every $A \in 2^X$, the *belief of A* is defined as

$$Bel_m(A) = \sum_{B \subseteq A} m(B). \tag{1}$$

Every mass assignment m on 2^X is in fact a probability distribution on 2^X that naturally induces a probability measure P_m on 2^{2^X}. Consequently, the belief function Bel_m corresponding to m can be equivalently described as follows: for every $A \in 2^X$,

$$Bel_m(A) = P_m(\{B \in 2^X : B \subseteq A\}). \tag{2}$$

Therefore, identifying the set $\{B \in 2^X : B \subseteq A\}$ with its characteristic function on 2^{2^X} defined by

$$\beta_A : B \in 2^X \mapsto \begin{cases} 1 & \text{if } B \subseteq A, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

it is easy to see that, for every $A \in 2^X$, and for every mass assignment $m : 2^X \to [0,1]$, we have

$$Bel_m(A) = P_m(\beta_A). \tag{4}$$

This easy characterization will be important when we discuss the extensions of belief functions on MV-algebras. The following is a trivial observation about the map β_A that can be useful to understand our generalization: for every $A \in 2^X$, β_A can be regarded as a map evaluating the (Boolean) inclusion of B into A, for every subset B of X.

A subset A of X such that $m(A) > 0$ is said to be a *focal element*. Every belief function is characterized by the value that m takes over its focal elements, and therefore, the focal elements of a belief function Bel_m contain the pieces of evidence that characterize Bel_m itself. For every set X and for every mass assignment m, call $\mathfrak{F}_m$ the set of focal elements of 2^X with respect to m. It is well known that several subclasses of belief functions can be characterized just by the structure of their focal elements. In particular, when $\mathfrak{F}_m \subseteq \{\{x\} : x \in X\}$, it is clear that Bel_m is indeed a probability measure. Moreover, if the focal elements are nested subsets of X, i.e. $\mathfrak{F}_m$ is a chain with respect to the inclusion relation between sets, then Bel_m is a *necessity measure* [19, 50]; this means e.g. that in such a case, it holds that $Bel_m(A_1 \cap A_2) = \min\{Bel_m(A_1), Bel_m(A_2)\}$ for eveary $A_1, A_2 \in 2^X$.

The whole class of belief functions on Boolean algebras is characterized by the property of non-decreasing differences of all possible orders. This property can be formulated for any function $v : L \to \mathbb{R}$ defined on a distributive lattice L [33, 61]. We say that v is *totally monotone* if, for every $n \geq 2$ and every $a_1, \ldots, a_n \in L$, we have

$$v \left(\bigvee_{i=1}^{n} a_i \right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} v \left(\bigwedge_{i \in I} a_i \right).$$

Shafer [50] has shown that the following assertions are equivalent for a function $v : 2^X \to [0,1]$ such that $v(\emptyset) = 0$ and $v(X) = 1$:

- v is a belief function,

- v is a totally monotone function on the distributive lattice 2^X.

As we will see in the following sections, this property does not characterize, in general, belief functions on fuzzy sets.

3   MV-algebras: an algebraic frame for many-valued events

In the same way Boolean algebras are the algebraic structures related to classical logic, MV-algebras are the algebras naturally associated to infinitely-valued Łukasiewicz logic.

The language of Łukasiewicz logic Ł (cf. [7, 35]), consists of a countable set of propositional variables $\{p_1, p_2, \ldots\}$, the binary connective $\to$, and the truth constant $\bot$. Formulas are defined by the usual inductive clauses. The following formulas provide an axiomatization for Ł:

(Ł1) $\varphi \to (\psi \to \varphi)$

(Ł2) $(\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))$

(Ł3) $((\varphi \to \bot) \to (\psi \to \bot)) \to (\psi \to \varphi)$

(**Ł4**) $((\varphi \to \psi) \to \psi) \to ((\psi \to \varphi) \to \varphi)$

The rule of inference of Ł is *modus ponens*: from φ and $\varphi \to \psi$, deduce ψ.

Further connectives in Ł are definable as follows: $\neg\varphi = \varphi \to \bot$; $\varphi \oplus \psi = \neg\varphi \to \psi$; $\varphi \odot \psi = \neg(\varphi \to \neg\psi)$; $\varphi \vee \psi = (\varphi \to \psi) \to \psi$; $\varphi \wedge \psi = \neg(\neg\varphi \vee \neg\psi)$; $\top = \neg\bot$.

Łukasiewicz logic is an algebraizable logic in the sense of Blok and Pigozzi [3], and its equivalent algebraic semantics is constituted by the class of MV-algebras [6, 7]. In algebraic terms, an MV-algebra is a structure $\mathbf{A} = (A, \oplus, \neg, 0)$ of type $(2, 1, 0)$ satisfying the following equations, for every $a, b, c \in A$:

(**MV1**) $a \oplus (b \oplus c) = (a \oplus b) \oplus c$,

(**MV2**) $a \oplus b = b \oplus a$,

(**MV3**) $a \oplus 0 = a$,

(**MV4**) $\neg\neg a = a$,

(**MV5**) $a \oplus \neg 0 = \neg 0$,

(**MV6**) $\neg(\neg a \oplus b) \oplus b = \neg(\neg b \oplus a) \oplus a$.

Further (definable) operations can be defined from $\oplus, \neg$ and 0 in a similar way as for the logical connectives above. In particular: $a \to b = \neg a \oplus b$, $a \odot b = \neg(\neg a \oplus \neg b)$; $a \vee b = \neg(\neg a \oplus b) \oplus b$; $a \wedge b = \neg(\neg a \vee \neg b)$; $1 = \neg 0$.

For every MV-algebra $\mathbf{A} = (A, \oplus, \neg, 0, 1)$, the structure $\mathbf{L}(\mathbf{A}) = (A, \wedge, \vee, 0, 1)$, where $\wedge$ and $\vee$ are defined as above, is a bounded distributive lattice and moreover the order relation $\leq$ defined by the stipulation: for all $a, b \in A$

$$x \leq y \text{ iff } x \to y = 1,$$

coincides with the lattice order of $\mathbf{L}(\mathbf{A})$. An MV-algebra whose order $\leq$ is linear is called an *MV-chain*. The class of MV-algebras forms a variety that we denote by $\mathbb{MV}$.

Let $\mathbf{A}$ be an MV-algebra. Then a non empty subset $\mathfrak{f}$ of A is said to be a *filter* of $\mathbf{A}$ iff: (i) $1 \in \mathfrak{f}$, (ii) if $a, b \in \mathfrak{f}$, then $a \odot b \in \mathfrak{f}$, and (iii) if $a \in \mathfrak{f}$ and $b \geq a$, then $b \in \mathfrak{f}$. A filter $\mathfrak{f}$ of an MV-algebra $\mathbf{A}$ is said to be *proper*, if $\mathfrak{f} \neq A$. A filter $\mathfrak{m}$ is said to be a *maximal filter* (or an *ultrafilter*) whenever for any proper filter $\mathfrak{f}$ such that $\mathfrak{f} \supseteq \mathfrak{m}$, either $\mathfrak{f} = A$, or $\mathfrak{f} = \mathfrak{m}$. The set of all ultrafilters of an MV-algebra $\mathbf{A}$ will be henceforth denoted by $\mathfrak{M}(\mathbf{A})$, or, when there is no danger of confusion, simply by $\mathfrak{M}$. For every MV-algebra $\mathbf{A}$, the set $\mathfrak{M}(\mathbf{A})$ is non-empty and it can be endowed with a compact Hausdorff topology, the so-called *spectral topology*: for an arbitrary filter $\mathfrak{f}$ of A, any set of the form $O_{\mathfrak{f}} = \{\mathfrak{m} \in \mathfrak{M}(\mathbf{A}) : \mathfrak{m} \not\supseteq \mathfrak{f}\}$ is open in this topology.

Observe that an intersection of a family of filters is a filter. The intersection of the family of all maximal filters of an MV-algebra $\mathbf{A}$ is called the *radical* of $\mathbf{A}$ and it is usually written $Rad(\mathbf{A})$. An MV-algebra $\mathbf{A}$ is *semisimple* whenever

$Rad(\mathbf{A}) = \{1\}$. It is well-known (see [7] for instance) that the congruences lattice and the filters lattice of any MV-algebra $\mathbf{A}$ are mutually isomorphic, via the isomorphism which associates to every congruence[1] θ the filter $\mathfrak{f}_\theta = \{a \in A \mid (a, 1) \in \theta\}$.

Example 3.1. The following are four relevant examples of MV-algebras:

(1) Every Boolean algebra is an MV-algebra, and moreover for every MV-algebra $\mathbf{A}$, the set $B(\mathbf{A}) = \{a \in A : a \oplus a = a\}$ of its idempotent elements is the domain of the largest Boolean subalgebra of $\mathbf{A}$. The algebra having $B(\mathbf{A})$ as universe is usually called the *Boolean skeleton* of $\mathbf{A}$.

(2) Define on the real unit interval $[0, 1]$ the operations $\oplus$ and $\neg$ as follows: for all $a, b \in [0, 1]$,

$$a \oplus b = \min\{1, a + b\}, \text{ and } \neg a = 1 - a.$$

Then the structure $[0, 1]_{MV} = ([0, 1], \oplus, \neg, 0)$ is an MV-algebra. The MV-algebra $[0, 1]_{MV}$ is generic for the variety of MV-algebras (i.e. it generates the whole variety $\mathbb{MV}$) and it is usually called the *standard* MV-algebra. In equivalent terms, Łukasiewicz logic is complete with respect to the semantics defined by the standard MV-algebra.

(3) Fix $k \in \mathbb{N}$, and let $F(k)$ be the set of all the McNaughton functions (cf. [7]) from the hypercube $[0, 1]^k$ into $[0, 1]$. In other words, $F(k)$ is the set of all the functions $f : [0, 1]^k \to [0, 1]$ which are continuous, piecewise linear and such that each linear piece has integer coefficients only. The following pointwise operations defined on $F(k)$,

$$(f \oplus g)(x) = \min\{1, f(x) + g(x)\}, \text{ and } (\neg f)(x) = 1 - f(x),$$

make the structure $\mathbf{F}(k) = (F(k), \oplus, \neg, \overline{0})$ into an MV-algebra, where $\overline{0}$ clearly denotes the function constantly equal to 0. Actually, $\mathbf{F}(k)$ is the free MV-algebra over k generators [7].

(4) Let X be a non-empty set, and let $\mathbf{A} = [0, 1]^X$ the set of all functions from X into $[0, 1]$, endowed with operations defined by the pointwise application of those in $[0, 1]_{MV}$. The structure $[0, 1]^X$ is clearly an MV-algebra, which we will henceforth call *MV-algebra of fuzzy sets* in order to point out that every fuzzy subset of X is indeed included into A. Every MV-subalgebra of $[0, 1]^X$ is called an *MV-clan* or simply a *clan* (cf. [4, 46]). Notice that, for a finite non-empty set X, the Boolean skeleton of the MV-algebra of fuzzy sets $[0, 1]^X$ coincides with the power set 2^X of X.

---

[1]A congruence θ in a MV-algebra $\mathbf{A}$ is an equivalence relation on A respecting the operations, i.e. if $(x, y) \in \theta$ then $(\neg x, \neg y) \in \theta$, and if $(x, y), (x', y') \in \theta$ then $(x \oplus x', y \oplus y') \in \theta$.

Notation 1. As already recalled in the introduction, MV-algebras, and MV-clans in particular, constitute the algebraic framework on top of which we will define belief functions. Indeed, elements of an MV-algebra $\mathbf{A}$ will be always intended to be the *fuzzy sets* we will work with. Therefore, although in the previous sections we used the notation $f, g, \ldots$ to indicate fuzzy sets, we will henceforth denote them by $a, b, c, \ldots$ without danger of confusion. At the same time, the notation $f, g, \ldots$ will be reserved to indicate functions in general. Moreover, in order to distinguish fuzzy sets from crisp sets, the latter will be indicated using capitals letters. So, for example, for any MV-algebra $\mathbf{A}$ we will denote the elements of $B(\mathbf{A})$ by $C, D, \ldots$, while for generic elements of $\mathbf{A}$ we will use the notation $a, b, c, \ldots$.

It is worth noticing that in $[0, 1]_{MV}$, the interpretation of the lattice operations of $\wedge$ and $\vee$ is, respectively, in terms of the min and max operators. Therefore, we will henceforth use both the notations $\wedge$ and min, and $\vee$ and max, without danger of confusion.

Roughly speaking the class of MV-algebras can be divided into semisimple and non-semisimple MV-algebras. This is, in particular, a key point of distinction between MV and Boolean algebras. Remember, in fact, that every Boolean algebra is semisimple, and that all MV-algebras of fuzzy sets $\mathbf{A} = [0, 1]^X$ are semisimple as well.

A semisimple MV-algebra $\mathbf{A} \subseteq [0, 1]^X$ is said to be *separating* provided that for each $x_1 \neq x_2 \in X$, there is a $a \in A$ such that $a(x_1) \neq a(x_2)$. Hence, the MV-algebras of fuzzy sets $\mathbf{A} = [0, 1]^X$ are both separating and semisimple. The following theorem provides a representation of semisimple MV-algebras by algebras of continuous functions.

Theorem 3.2 (Chang [6], Belluce [1]). *Up to isomorphism, every semisimple MV-algebra $\mathbf{A}$ is a separating algebra of continuous $[0, 1]$-valued functions over the compact Hausdorff space $\mathfrak{M}(\mathbf{A})$ of ultrafilters of $\mathbf{A}$.*

The following result, which we state for the particular case of MV-algebras of fuzzy sets, holds in a much more general setting. We invite the interested reader to consult [7] for a more exhaustive treatment.

Theorem 3.3. *For every MV-algebra $\mathbf{A} = [0, 1]^X$, there exists a one-to-one correspondence between the points of X and the class $Hom(\mathbf{A}, [0, 1]_{MV})$ of homomorphisms of $\mathbf{A}$ into the standard MV-algebra $[0, 1]_{MV}$.*

Thanks to the above Theorem 3.3 we will henceforth identify points in X with homomorphisms of $\mathbf{A}$ in the standard MV-algebra $[0, 1]_{MV}$ without loss of generality. Moreover, the following holds:

Corollary 3.4. *Let $\{\tau_1, \ldots, \tau_s\}$ be a finite subset of an MV-algebra $\mathbf{A} = [0, 1]^X$. Then*

$$\{\langle h(\tau_1), \ldots, h(\tau_s) \rangle \in [0, 1]^s : h \in Hom(\mathbf{A}, [0, 1]_{MV})\} =$$
$$\{\langle \tau_1(x), \ldots, \tau_s(x) \rangle : x \in X\}.$$

In this paper we will mainly concentrate on MV-algebras which are MV-clans $[0,1]^X$ defined over a *finite* set X. Such MV-algebras can be identified with finite direct products of copies of $[0,1]_{MV}$.

3.1 States on MV-algebras

States on MV-algebras have been introduced by Mundici in [44] as averaging processes for the infinitely-valued Łukasiewicz calculus.

Definition 3.5. Let $\mathbf{A}$ be an MV-algebra. A *state* on $\mathbf{A}$ is a map $s : A \to [0,1]$ such that:

(s1) $s(1) = 1$,

(s2) For all $a, b \in A$ such that $a \odot b = 0$, $s(a \oplus b) = s(a) + s(b)$.

A state s on an MV-algebra $\mathbf{A}$ is said to be *faithful* if $s(x) = 0$ implies $x = 0$.

For a given MV-algebra $\mathbf{A}$, the class of all states on $\mathbf{A}$ is denoted by $\mathcal{S}(\mathbf{A})$. States play the same role on MV-algebra as probability measures do on Boolean algebras: indeed, the two properties (s1) and (s2) characterize each state on $\mathbf{A}$ as a $[0,1]$-valued map that is normalized (s1) and additive (s2) with respect to the MV-algebraic operations. Moreover, it is easy to see that, for every MV-algebra $\mathbf{A}$ and for every $s \in \mathcal{S}(\mathbf{A})$, the restriction of s to the Boolean skeleton $B(\mathbf{A})$ of $\mathbf{A}$ is a finitely additive probability measure. The following theorem, independently proved in [40] and [47], shows an integral representation of states by Borel probability measures defined on the σ-algebra $\mathfrak{B}(X)$ of Borel subsets of X, where X is any compact Hausdorff topological space.

Theorem 3.6. *Let $\mathbf{A} \subseteq [0,1]^X$ be a separating clan of continuous functions over a compact Hausdorff space X. Then there is a one-to-one correspondence between the class $\mathcal{S}(\mathbf{A})$ of states on $\mathbf{A}$, and the regular Borel probability measures on $\mathfrak{B}(X)$. In particular, for every state s on $\mathbf{A}$, there exists a unique regular Borel probability measure μ on $\mathfrak{B}(X)$ such that for every $a \in A$,*

$$s(a) = \int_X a \, \mathrm{d}\mu. \tag{5}$$

In the next example we consider a particular case of states on MV-algebras of fuzzy sets, focusing on the integral representation presented above.

Example 3.7. Let X be a finite non empty set. Let $\mathbf{A} = [0,1]^{(2^X)}$ be the MV-algebra of fuzzy sets consisting of all functions from 2^X to $[0,1]$ (i.e. $\mathbf{A}$ is the MV-algebra of all fuzzy subsets of the powerset 2^X of X). We will henceforth deal with those states on $[0,1]^{(2^X)}$ satisfying $s(\chi_{\{\emptyset\}}) = 0$ (where $\chi_{\{\emptyset\}}$ denotes the characteristic function of $\emptyset \in 2^X$). The above Theorem 3.6 ensures that for each such state s, there exists a unique finitely additive probability measure $\mu : 2^{(2^X)} \to [0,1]$ such that, for every $a \in [0,1]^{(2^X)}$,

$$s(a) = \sum_{A \subseteq X} a(A) \cdot \mu(\{A\}),$$

and $\mu(\{\emptyset\}) = 0$.

Obviously the class $\mathcal{S}(\mathbf{A})$ is non-empty since $Hom(\mathbf{A}, [0, 1]_{MV}) \subseteq \mathcal{S}(\mathbf{A})$. Moreover $\mathcal{S}(\mathbf{A})$ is a convex subset of the compact Hausdorff space $[0, 1]^A$ whose set of extremal points coincides with $Hom(\mathbf{A}, [0, 1]_{MV})$. For every subset X of a topological vector space, let us write $\overline{\text{co}}(X)$ to denote the closure of the convex hull of the set X [21]. Then, Krein-Mil'man Theorem [31] gives the following result.

Theorem 3.8. *For every MV-algebra* $\mathbf{A}$*,* $\mathcal{S}(\mathbf{A}) = \overline{\text{co}}(Hom(\mathbf{A}, [0, 1]_{MV}))$*.*

The following example is obtained by applying the above Theorem 3.8 to the particular case of MV-algebras of fuzzy sets, and it will be needed in the remaining of part of this chapter.

Example 3.9. As in Example 3.7, let X be finite and let $\mathcal{S}_0$ be the subset of $\mathcal{S}([0, 1]^{(2^X)})$ of those states $s : [0, 1]^{(2^X)} \to [0, 1]$ further satisfying $s(\chi_{\{\emptyset\}}) = 0$. The set $\mathcal{S}_0$ is a convex subset of the $(2^{|X|} - 1)$-dimensional Euclidean space. Since the correspondence between $\mathcal{S}_0$ and the set of all probabilities μ on $2^{(2^X)}$ satisfying $\mu(\{\emptyset\}) = 0$ is a one-to-one affine mapping, the set $\mathcal{S}_0$ is a $(2^{|X|} - 2)$-simplex as well. Regarding the extreme points of $\mathcal{S}_0$, we can observe that they are in one-to-one correspondence with the non-empty subsets of X, and hence every state s_A, with $A \in 2^X \setminus \{\emptyset\}$, such that $s_A(a) = a(A)$ for each $a \in [0, 1]^{(2^X)}$ is an extreme point of $\mathcal{S}_0$.

4 Belief functions on MV-algebras of fuzzy sets

In this section we are going to discuss two main MV-algebraic generalizations of belief functions. Our approach is to generalize the definition (4), where Bel_m is derived from a probability measure P_m on 2^{2^X}. Therefore, we need to generalize both the inclusion map β_A and the probability measure P_m. In the following subsections we investigate two directions in which these notions can be generalized.

4.1 The case of crisp focal elements

Let X be a finite nonempty set, and let, for each element a in the MV-algebra $[0, 1]^X$, the map $\hat{\rho}_a : 2^X \to [0, 1]$ be defined by the following stipulation: for all $B \in 2^X$,

$$\hat{\rho}_a(B) = \begin{cases} \min_{x \in B} a(x) & \text{if } B \neq \emptyset, \\ 1 & \text{if } B = \emptyset. \end{cases} \tag{6}$$

Remark 4.1. $\hat{\rho}_a$ generalizes the map β_A we discussed in Section 2 in the following sense: whenever $A \in B(\mathbf{A}) = 2^X$, then $\hat{\rho}_A = \beta_A$. Indeed, for every $A \in B(\mathbf{A})$, $\hat{\rho}_A(B) = 1$ if $B \subseteq A$, and $\hat{\rho}_A(B) = 0$, otherwise.

Definition 4.2 (Crisp Focal Elements)**.** Let X be a finite nonempty set. Then a map $\hat{\mathbf{b}} : [0,1]^X \to [0,1]$ is a *crisp-focal element belief function*, if there exists a state $s : [0,1]^{(2^X)} \to [0,1]$ such that, for all $a \in [0,1]^X$

$$\hat{\mathbf{b}}(a) = s(\hat{\rho}_a).$$

The state s defining $\hat{\mathbf{b}}$ will be henceforth called the *state assignment* of $\hat{\mathbf{b}}$.

The integral representation theorem for states (Theorem 3.6) can be generalized to crisp-focal belief functions. This requires the introduction of the Choquet integral (cf. [14]). Let f be any function from a finite nonempty set X to $[0,1]$, and let σ be a set function $\sigma : 2^X \to [0,1]$ such that $\sigma(\emptyset) = 0$. Then the *Choquet integral* of f with respect to σ is defined as

$$\oint f \, \mathrm{d}\sigma = \int_0^1 \sigma(f^{-1}([t,1])) \, \mathrm{d}t.$$

Since X is finite, the integral $\oint f \, \mathrm{d}\sigma$ exists and takes the form of a finite sum. In fact, without loss of generality let $X = \{x_1, \ldots, x_n\}$, where the numbers $y_i = f(x_i)$ satisfy $y_1 \geq \cdots \geq y_n$. Put $y_{n+1} = 0$ and for each $i = 1, \ldots, n$, $S_i = \{x_1, \ldots, x_i\}$, then

$$\oint f \, \mathrm{d}\sigma = \sum_{i=1}^n (y_i - y_{i+1})\sigma(S_i).$$

Proposition 4.3. *For every crisp-focal belief function* $\hat{\mathbf{b}} : [0,1]^X \to [0,1]$ *there exists a unique belief function* $Bel : 2^X \to [0,1]$ *such that, for each* $a \in [0,1]^X$,

$$\hat{\mathbf{b}}(a) = \oint a \, \mathrm{d}(Bel).$$

Proof. Let s be the state assignment on $[0,1]^{(2^X)}$ which defines $\hat{\mathbf{b}}$. According to Example 3.7 there is a unique finitely additive probability measure μ on $2^{(2^X)}$ such that, for each $f \in [0,1]^{(2^X)}$, one has $s(f) = \sum_{A \subseteq X} f(A) \cdot \mu(\{A\})$ and $\mu(\{\emptyset\}) = 0$. Therefore, the crisp-focal belief function $\hat{\mathbf{b}}$ is expressed as follows: for every $a \in [0,1]^X$,

$$\hat{\mathbf{b}}(a) = s(\hat{\rho}_a) = \sum_{A \subseteq X} \hat{\rho}_a(A) \cdot \mu(\{A\}). \tag{7}$$

Recalling the definition (3) of the map β_A, we have $\hat{\rho}_a(A) = \min\{a(x) : x \in A\} = \oint a \, \mathrm{d}\beta_A$, for every $a \in [0,1]^X$ and for every $A \subseteq X$. Equation (7) together with the linearity of Choquet integral with respect to the integrating set function β_A yields

$$\hat{\mathbf{b}}(a) = \sum_{A \in 2^X \setminus \{\emptyset\}} \mu(\{A\}) \cdot \oint a \, \mathrm{d}\beta_A = \oint a \, \mathrm{d}\left(\sum_{A \in 2^X \setminus \{\emptyset\}} \mu(\{A\}) \cdot \beta_A \right).$$

The claim then follows noticing that the function $Bel : 2^X \to [0, 1]$ such that for each $B \subseteq X$,

$$Bel(B) = \sum_{A \in 2^X \setminus \{\emptyset\}} \mu(\{A\}) \cdot \beta_A(B) = \mu(\{A \subseteq X \mid A \subseteq B\})$$

is a belief function on 2^X. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

With X being finite, despite the previous representation theorem for crisp-focal belief functions in terms of Choquet integral, Theorem 3.6 and Example 3.7 yields a unique probability measure $\mu : 2^{(2^X)} \to [0, 1]$ such that for every $a \in [0, 1]^X$

$$s(\hat\rho_a) = \sum_{C \in 2^X} \hat\rho_a(C) \cdot \mu(\{C\}). \qquad\qquad (8)$$

Moreover, it is easy to see that, for every $C \subseteq 2^X$, $\mu(\{C\}) = s(\{C\})$, where $s(\{C\})$ is a succint expression for $s(\chi_{\{C\}})$. Since $\mu(\{\emptyset\}) = 0$, the probability measure μ induces a mass assignment m such that $m(C) = \mu(\{C\})$. This remark explains the name *crisp-focal* for the belief functions as in Definition 4.2. In fact, from (8), each crisp-focal belief function $\hat{\mathbf{b}}$ assigns, to each element $a \in [0, 1]^X$, the value $\hat{\mathbf{b}}(a) = \sum_{C \subseteq X} \hat\rho_a(C) \cdot m(C)$. Therefore $\hat{\mathbf{b}}(a)$ is only determined by the *crisp* elements $C \subseteq X$ for which $m(C) > 0$, i.e. Boolean (crisp) focal elements.

In Dempster-Shafer theory, given a belief function $Bel : 2^X \to [0, 1]$, the mass m that defines Bel can be recovered from Bel by the *Möbius transform* [50] of Bel:

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} Bel(B).$$

In case of crisp-focal belief functions, the situation is analogous.

Proposition 4.4. *Let* $\hat{\mathbf{b}} : [0, 1]^X \to [0, 1]$ *be a crisp-focal belief function, defined as* $\hat{\mathbf{b}}(a) = s(\rho_a)$ *for some state s on* $[0, 1]^{(2^X)}$ *such that $s(\{\emptyset\}) = 0$ and $s(\{C\}) > 0$ iff $C(x) \in \{0, 1\}$, where $C \neq \emptyset$. Then*

$$s(\{A\}) = m(A) = \textstyle\sum_{B \subseteq A}(-1)^{|A \setminus B|}\hat{\mathbf{b}}(B)$$

for each $A \subseteq X$.

Proof. Definition (4.2) directly gives that $\hat\rho_A(C) \in \{0, 1\}$ for each pair of crisp sets $A, C \subseteq X$ and thus

$$\hat{\mathbf{b}}(A) = \textstyle\sum_{C \in 2^X} \hat\rho_A(C) \cdot s(\{C\}) = \sum_{B \subseteq A} s(\{B\}) = \sum_{B \subseteq A} m(B).$$

This implies that the restriction of $\hat{\mathbf{b}}$ to 2^X is a classical belief function. See [42] for further details. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

As a corollary, observe that, in the hypothesis of the above proposition, the values $\hat{\mathbf{b}}(a)$ for non-crisp $a \in [0, 1]^X$ are fully determined by the values of $\hat{\mathbf{b}}$ over the crisp sets of 2^X. Indeed, Proposition 4.3 proves that, for any $a \in [0, 1]^X$, $\hat{\mathbf{b}}(a)$ is the Choquet integral of a with respect to the restriction of $\hat{\mathbf{b}}$ over 2^X. In this way we arrive at another characterization of crisp-focal belief functions.

Theorem 4.5. *A function $\hat{\mathbf{b}} : [0,1]^X \to [0,1]$ is a crisp-focal belief function iff its restriction onto 2^X is a totally monotone function, i.e., for every natural n and every $A_1, \ldots, A_n \in 2^X$, the following inequality holds:*

$$\hat{\mathbf{b}}\left(\bigvee_{i=1}^{n} A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1, \ldots, n\}} (-1)^{|I|+1} \cdot \hat{\mathbf{b}}\left(\bigwedge_{k \in I} A_k\right).$$

The geometrical structure of the set of all crisp-focal belief functions on $[0,1]^X$ is completely determined by the associated simplex of state assignments on $[0,1]^{(2^X)}$. For each $A \subseteq X$, a crisp focal belief function $\hat{\mathbf{b}}_A(a) = \min\{a(x) : x \in A\}$ for $a \in [0,1]^X$ corresponds to the state assignment s_A (see Example 3.9). Consequently, we obtain the following characterization of the set of all crisp focal belief functions.

Proposition 4.6. *The set of all crisp focal belief functions on $[0,1]^X$ is a $(2^{|X|} - 2)$-simplex whose set of extreme points is $\{\hat{\mathbf{b}}_A \mid A \in 2^X \setminus \{\emptyset\}\}$.*

4.2 The case of fuzzy-focal elements

The notion of belief function on MV-algebras we are going to introduce in this section (cf. [25]) generalizes crisp-focal belief function by introducing, for every $a \in \mathbf{A} = [0,1]^X$, a more general inclusion map ρ_a associating with each *fuzzy set $b \in A$* the degree of inclusion of the fuzzy set b into the fuzzy set a as follows:

$$\rho_a(b) = \min\{b(x) \Rightarrow a(x) : x \in X\}, \tag{9}$$

where $\Rightarrow$ is Lukasiewicz implication in the standard algebra $[0,1]_{MV}$ defined as $u \Rightarrow v = (\neg u) \oplus v = \min(1, 1 - u + v)$, for all $u, v \in [0,1]$. The choice of $\Rightarrow$ in the above definition is clearly due to the MV-algebraic setting, but different choices could be made in other fuzzy logic settings.

The mapping ρ_a can be indeed regarded as a *generalized inclusion operator* between fuzzy sets (cf. [25] for further details) since the following intuitive properties are satisfied by such mappings:

- $\rho_a(b) = 1$ iff $b(x) \leq a(x)$, for all $x \in X$;

- $\rho_a(b) \geq \rho_a(b')$, whenever $b(x) \leq b'(x)$, for all $x \in X$;

- $\rho_a(b) = 0$ iff there is $x \in X$ such that $b(x) = 1$ and $a(x) = 0$.

Next proposition shows that the mapping ρ_a generalizes both the previously introduced mappings β_A and $\hat{\rho}_a$.

Proposition 4.7. (i) *For all $a, a' \in \mathbf{A}$, $\rho_{a \wedge a'} = \min\{\rho_a, \rho_{a'}\}$, and $\rho_{a \vee a'} \geq \max\{\rho_a, \rho_{a'}\}$.*

(ii) *For every $a \in \mathbf{A}$, the restriction of ρ_a to $B(\mathbf{A})$ coincides with the transformation $\hat{\rho}_a$ defined by (6).*

(iii) *For every $A \in B(\mathbf{A})$, the restriction of ρ_A to $B(\mathbf{A})$ coincides with the transformation β_A defined by (3).*

Proof. (i) In every MV-chain, and in particular in the standard chain $[0,1]_{\mathrm{MV}}$ the equation $\neg c \oplus (a \wedge b) = (\neg c \oplus a) \wedge (\neg c \oplus b)$ holds:, i.e. $(c \Rightarrow (a \wedge b)) = (c \Rightarrow a) \wedge (c \Rightarrow b)$. Therefore, for every $a, a', b \in \mathbf{A}$,

$$
\begin{aligned}
\rho_{a \wedge a'}(b) &= \min\{b(x) \Rightarrow (a \wedge a')(x) : x \in X\} \\
&= \min\{b(x) \Rightarrow (a(x) \wedge a'(x)) : x \in X\} \\
&= \min\{(b(x) \Rightarrow a(x)) \wedge (b(x) \Rightarrow a'(x)) : x \in X\} \\
&= \min\{\rho_a(b), \rho_{a'}(b)\}.
\end{aligned}
$$

An easy computation shows that $\rho_{a \vee a'} \geq \max\{\rho_a, \rho_{a'}\}$.

(ii) For every $B \in B(\mathbf{A})$, $\rho_a(B) = \min\{B(x) \Rightarrow a(x) : x \in X\}$. Whenever $x \notin B$, $B(x) = 0$, and hence $B(x) \Rightarrow a(x) = 1$ for all those $x \notin B$. On the other hand for all $x \in B$, $B(x) = 1$, and so $B(x) \Rightarrow a(x) = 1 \Rightarrow a(x) = a(x)$ for all $x \in B$. Consequently, $\rho_a(B) = \min\{a(x) : x \in B\}$.

(iii) It follows immediately from (ii) together with Remark 4.1.　□

For every $A \in 2^X$ (i.e. whenever A is identified with a vector in $[0,1]^X$ having *integer* coordinates), the map $\rho_A : [0,1]^X \to [0,1]$ is a pointwise minimum of finitely many linear functions with integer coefficients, and hence ρ_A is a non-increasing McNaughton function [7].

Lemma 4.8. *The MV-algebra $\mathbf{R_2}$ generated by the set $\varrho_2 = \{\rho_A : A \in 2^X\}$ coincides with the free MV-algebra over n generators $\mathbf{F}(n)$, where n is the cardinality of X.*

Proof. By [8, Theorem 3.13], if a variety $\mathbb{V}$ of algebras is generated by an algebra $\mathbf{A}$, then the free algebra over a cardinal $n > 0$ is, up to isomorphisms, the subalgebra of $\mathbf{A}^{\mathbf{A}^X}$ generated by the projection functions $\theta_i : \mathbf{A}^X \to \mathbf{A}$. Therefore, in order to prove our claim it suffices to show that the projection functions $\theta_1, \ldots, \theta_n$ belong to ϱ_2.

For every $i = 1, \ldots, n$, let the vector $\bar{i} \in \{0,1\}^X$ be defined as

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

Then $\rho_{\bar{i}} = 1 - \theta_i$. In fact, for every $b \in [0,1]^X$, and for every $i, j \in X$ such that $j \neq i$, we have $b(j) \to \bar{i}(j) = 1$, and $b(i) \to \bar{i}(i) = 1 - b(i)$, so that $1 - \rho_{\bar{i}(b)} = \theta_i(b) = b(i)$. This actually shows that the MV-algebra $\mathbf{R_2^{\neg}}$ generated by the set $\neg \varrho_2 = \{1 - \rho_A : A \in 2^X\}$ is isomorphic to $\mathbf{F}(n)$. Clearly $\mathbf{R_2}$ and $\mathbf{R_2^{\neg}}$ are isomorphic through the map $g : a \in \mathbf{R_2} \mapsto 1 - a \in \mathbf{R_2^{\neg}}$.　□

Therefore, if we consider the MV-algebra $\mathbf{R}_X$ generated by the set $\varrho = \{\rho_a : a \in [0,1]^X\}$ we obtain a semisimple MV-algebra that properly extends $\mathbf{F}(n)$, and whose elements are continuous functions from $[0,1]^X$ into $[0,1]$. This implies, in particular, that $\mathbf{R}_X$ is separating.

Definition 4.9 (Fuzzy-Focal Belief Function)**.** Let X be a finite set and let $\mathbf{A} = [0,1]^X$. A map $\mathbf{b} : \mathbf{A} \to [0,1]$ will be called a *(fuzzy-focal) belief function* on the finite domain MV-clan $\mathbf{A}$ provided there exists a state $s : \mathbf{R}_X \to [0,1]$ such that for every $a \in A$,

$$\mathbf{b}(a) = s(\rho_a). \tag{10}$$

We will denote by $Bel(\mathbf{A})$ the class of all the (fuzzy-focal) belief functions over the finite domain MV-clan $\mathbf{A}$.

In analogy with the case of crisp-focal belief functions, the state s defining $\mathbf{b}$ will be henceforth called the *state assignment* of $\mathbf{b}$.

As in the previous section, we will identify the mass of a belief function $\mathbf{b}$ with the unique Borel regular probability measure μ over $\mathfrak{B}([0,1]^X)$ that represents the state s via Theorem 3.6.

Remark 4.10. Note that if the set $\{b \in [0,1]^X \mid \mu(\{b\}) > 0\}$ is countable then the above Definition 4.9 yields

$$\mathbf{b}(a) = \sum_{b \in A} \rho_a(b) \cdot \mu(\{b\}). \tag{11}$$

In this case, a *focal element* is any $b \in A$ such that $\mu(\{b\}) > 0$ and hence, in contrast with the case of crisp-focal belief functions, it is clear that focal elements of $\mathbf{b}$ can be proper fuzzy sets.

We showed in Theorem 4.5 that the property of total monotonicity characterizes crisp focal belief functions on MV-algebras. As for the case of fuzzy-focal belief functions, the problem of characterizing those belief functions in terms of (a variant of) total monotonicity is open, but the following implication holds.

Proposition 4.11. *For every finite-domain MV-clan $\mathbf{A}$ and for every $\mathbf{b} \in Bel(M)$, $\mathbf{b}$ is totally monotone on the lattice reduct of $\mathbf{A}$.*

Proof. Since for every $a \in \mathbf{A}$, ρ_a is monotone, and every state s is monotone, $\mathbf{b}$ is monotone as well. Moreover, for every n and for every $a_1, \ldots, a_n \in \mathbf{A}$, from (10) and Proposition 4.7 (i) we have the following chain of inequalities:

$$
\begin{aligned}
\mathbf{b}\left(\bigvee_{i=1}^n a_i\right) &= s(\rho_{a_1 \vee \ldots \vee a_n}) \\
&\geq s(\rho_{a_1} \vee \ldots \vee \rho_{a_n}) \\
&= \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \cdot s\left(\bigwedge_{k \in I} \rho_{a_k}\right) \\
&= \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \cdot s\left(\rho_{(\bigwedge_{k \in I} a_k)}\right) \\
&= \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \cdot \mathbf{b}\left(\bigwedge_{k \in I} a_k\right).
\end{aligned}
$$

$\square$

Since belief functions on $[0,1]^X$ are defined by states on $\mathbf{R}_X$ and different states s_1 and s_2 determine different belief functions $\mathbf{b}_1$ and $\mathbf{b}_2$, the set $Bel([0,1]^X)$ of belief functions on $[0,1]^X$ is in 1-1 correspondence with the set

$\mathcal{S}(\mathbf{R}_X)$ of all states on $\mathbf{R}_X$. Moreover, this correspondence is an affine mapping. Hence $Bel([0,1]^X)$ is a compact convex subset of $[0,1]^{([0,1]^X)}$. Therefore Krein-Mil'man theorem shows that $Bel([0,1]^X)$ is in the closed convex hull of its extremal points. The following result characterizes the extremal points of $Bel([0,1]^X)$.

Proposition 4.12. *For every $x \in [0,1]^X$, the belief function $\mathbf{b}_x$ defined by*

$$\mathbf{b}_x(a) = s_x(\rho_a) = \rho_a(x), \quad a \in [0,1]^X, \tag{12}$$

is an extremal point of $Bel([0,1]^X)$.

Proof. A belief function $\mathbf{b} \in Bel([0,1]^X)$ is extremal iff its state assignment is extremal in $\mathcal{S}(\mathbf{R}_X)$. In fact s is not extremal iff there exist $s_1, s_2 \in \mathcal{S}(\mathbf{R}_X)$ and a real number $\lambda \in (0,1)$ such that $s = \lambda s_1 + (1-\lambda)s_2$. In particular, for every $a \in [0,1]^X$,

$$\mathbf{b}(a) = s(\rho_a) = \lambda s_1(\rho_a) + (1-\lambda)s_2(\rho_a) = \lambda \mathbf{b}_1(a) + (1-\lambda)\mathbf{b}_2(a),$$

whence $\mathbf{b}$ would not be extremal as well. $\qquad\square$

As recalled above, $\mathbf{R}_X$ is separating. Therefore from Proposition 4.12 the extreme points of its state space are MV-homomorphisms s_x, for each $x \in [0,1]^X$. Hence the following holds due to (12).

Theorem 4.13. *Every belief function $\mathbf{b}$ is a pointwise limit of a convex combination of some functions $\rho_.(a^1), \ldots, \rho_.(a^k)$, where $a^1, \ldots, a^k \in [0,1]^X$.*

Remark 4.14. Consider the restriction $\mathbf{b}^-$ of a fuzzy-focal belief function $\mathbf{b}$ to the Boolean skeleton 2^X of its domain $[0,1]^X$. Then, although it has fuzzy-focal elements, the map $\mathbf{b}^-$ actually is a classical belief function since, by Proposition 4.11, $\mathbf{b}^-$ keeps being total monotone. Therefore, there exists a mass assignment on crisp subsets of X giving the same $\mathbf{b}^-$. In other words there exists a mass assignment $m^- : 2^X \to [0,1]$ such that, for every $A \in 2^X$ one has:

$$\mathbf{b}(A) = \mathbf{b}^-(A) = \sum_{B \subseteq A} m^-(B)$$

In the framework of finitely-valued fuzzy sets on the scale $S_k = \{0, 1/k, \ldots, (k-1)/k, 1\}$, an interesting question is how to compute the mass m^- from the mass μ giving $\mathbf{b}$, that is, what is the map $m^- : 2^X \to [0,1]$ fulfilling the constraints

$$\sum_{a \in (S_k)^X} \rho_A(a) \cdot \mu(a) = \sum_{B \subseteq A} m^-(B)$$

for each $A \in 2^X$. In fact, following [18, 55], to find the solution the idea is to decompose the mass $\mu(a)$ of each fuzzy-focal element a into its level cuts a_{α_i}, with $\alpha_i \in S_k$, as follows:

$$m_a(a_{\alpha_i}) = \mu(a) \cdot (\alpha_i - \alpha_{i-1})$$

where $\alpha_i = i/k$, for $i = 0, 1, \ldots, k$. Finally, since it may be the case that two (or more) level sets of different fuzzy-focal elements coincide, we define for each $A \in 2^X$:

$$m^-(A) = \sum \{m_a(a_\alpha) \mid a \in (S_k)^X, \alpha \in S_k \text{ such that } a_\alpha = A\}.$$

5   On normalized belief functions

The *height* of a fuzzy set $a \in [0,1]^X$ is defined in the literature as

$$h(a) = \max\{a(x) : x \in X\}. \tag{13}$$

The value $h(a)$ can be interpreted as the degree of normalization of a. As a matter of fact, a fuzzy set a is called *normalized* whenever $h(a) = 1$, otherwise it is called *non-normalized*. A non-normalized fuzzy set represents a partially inconsistent information.

The map $\rho_{\overline{0}}$ evaluating the degree of inclusion of any fuzzy set $b \in [0,1]^X$ in the empty fuzzy set $\overline{0}$ (constantly zero function) does not coincide, in general, with $\overline{0}$ itself. In fact, whenever b is a non-normalized fuzzy set (i.e. $h(b) < 1$), $\rho_{\overline{0}}(b) > 0$. Therefore, if s is a faithful state on $\mathbf{R}_X$, the fuzzy-focal belief function defined through s satisfies $\mathbf{b}(\overline{0}) > 0$.

Definition 5.1. A (fuzzy-focal) belief function $\mathbf{b} : [0,1]^X \to [0,1]$ is said to be *normalized* provided that

$$\mathbf{b}(\overline{0}) = s(\rho_{\overline{0}}) = 0. \tag{14}$$

In this section we will focus on normalized fuzzy-focal belief functions. Indeed it is worth noticing that crisp-focal belief functions are normalized, i.e. they always satisfy (14).

In classical Dempster-Shafer theory, the notion of focal element is crucial for classifying belief functions. Whenever $X = \{1, \ldots, n\}$ is a finite set, the Boolean algebra 2^X is finite, and hence the mass assignment $m : 2^X \to [0,1]$ obviously defines only finitely many focal elements. On the other hand, the MV-algebra $[0,1]^X$ has uncountably many elements, and hence we cannot find in general a mass assignment μ defined over $\mathfrak{B}([0,1]^X)$ and supported by an at most countable set only. This observation leads to the following definition.

Definition 5.2. Let $\mathcal{K}$ be the set of all compact subsets of an MV-algebra of fuzzy sets $[0,1]^X$. For every regular Borel probability measure μ defined on $\mathfrak{B}([0,1]^X)$, we call the set

$$\operatorname{spt} \mu = \bigcap \{K \mid K \in \mathcal{K}, \mu(K) = 1\}$$

the *support of* μ.

By Theorem 3.6 we can regard spt μ as the support of the state s defined from μ via (5). In particular, the following holds:

$$\mathbf{b}(a) = \int_{[0,1]^X} \rho_a \, \mathrm{d}\mu = \int_{\mathrm{spt}\, \mu} \rho_a \, \mathrm{d}\mu. \tag{15}$$

Therefore, for a belief function $\mathbf{b}$ on $[0,1]^X$ whose state assignment s is represented by a regular Borel probability measure μ, we will henceforth refer to spt μ as the set of focal elements of $\mathbf{b}$.

Proposition 5.3. *The set $\mathcal{S}_0$ of all states on $\mathbf{R}_X$ satisfying (14) is a nonempty compact convex subset of $[0,1]^{\mathbf{R}_X}$ considered with its product topology.*

Proof. $\mathcal{S}_0$ is nonempty: let s_1 be defined by

$$s_1(\rho) = \rho(\overline{1}),$$

for every $\rho \in \mathbf{R}_X$, where $\overline{1} : X \to [0,1]$ is the constant function of value 1. This gives in particular $s_1(\rho_{\overline{0}}) = \rho_{\overline{0}}(\overline{1}) = 0$ and thus $s_1 \in \mathcal{S}_0$. Let $s, s' \in \mathcal{S}_0$ and $\alpha \in (0,1)$. Then the function given by

$$\alpha s + (1 - \alpha)s'$$

is a state on $\mathbf{R}_X$ which clearly satisfies (14). Hence $\mathcal{S}_0$ is a convex subset of the product space $[0,1]^{\mathbf{R}_X}$. Since the space $[0,1]^{\mathbf{R}_X}$ is compact, we only need to show that $\mathcal{S}_0$ is closed (in its subspace product topology). To this end, consider a convergent sequence $(s_m)_{m \in \mathbb{N}}$ in $\mathcal{S}_0$ whose limit is s. As the set of all states on $\mathbf{R}_X$ is closed, s is a state. That s satisfies (14) follows from the fact that $s(\rho_{\overline{0}}) = \lim_{m \to \infty} s_m(\rho_{\overline{0}}) = 0$. $\square$

The family of states $\mathcal{S}_0$ can be characterized by employing the integral representation of states. Namely, we will show that a state assignment $s \in \mathcal{S}_0$ iff s is "supported" by normal fuzzy sets in $[0,1]^X$, i.e. fuzzy sets $a \in [0,1]^X$ such that $a(x) = 1$ for some $x \in X$. We will denote by $\mathcal{NF}(X)$ the set of normalized fuzzy sets from $[0,1]^X$, i.e.

$$\mathcal{NF}(X) = \{a \in [0,1]^X \mid a(x) = 1 \text{ for some } x \in X\}.$$

The following result characterizes normalized fuzzy-focal belief functions in terms of the support of their state assignment.

Theorem 5.4. *Let s be a state assignment on $\mathbf{R}_X$ and μ be the regular Borel probability measure associated with s. Then spt $\mu \subseteq \mathcal{NF}(X)$ if and only if $s \in \mathcal{S}_0$.*

Proof. Let μ be a probability measure on Borel subsets of $[0,1]^X$ such that spt $\mu \subseteq \mathcal{NF}(X)$. Put

$$s(a) = \int_{[0,1]^X} a \, \mathrm{d}\mu, \quad a \in \mathbf{R}_X. \tag{16}$$

Since $\rho_{\bar{0}}(a) = 0$ for each $a \in$ spt μ, it follows that

$$s(\rho_{\bar{0}}) = \int_{\text{spt } \mu} \rho_{\bar{0}} \, d\mu = 0,$$

hence $s \in \mathcal{S}_0$. Conversely, assume that

$$s(\rho_{\bar{0}}) = \int_{[0,1]^X} \rho_{\bar{0}} \, d\mu = 0,$$

which implies $\rho_{\bar{0}} = 0$ μ-almost everywhere over $[0,1]^X$. Since $\rho_{\bar{0}}(a) = 0$ iff $a \in [0,1]^X$ is such that $a(x) = 1$, for some $x \in X$, we obtain $\mu(\mathcal{NF}(X)) = 1$. $\quad\square$

In particular, every state assignment of a crisp-focal belief function belongs to $\mathcal{S}_0$.

6 Soft-normalization for fuzzy-focal belief functions

Throughout the rest of the paper, we stipulate the following:

We always assume a mass μ such that its support spt μ *is countable.*

Consider a belief function $\mathbf{b}$ with a state assignment s supported by spt μ. Assume that there exists a focal element $a \in$ spt μ that is a non-normalized fuzzy set. Since spt μ is countable, we have neccesarily $\mu(\{a\}) > 0$,[2] and $\mathbf{b}$ is associating a positive degree of evidence to a (partially) inconsistent information, which is reflected on the value that $\mathbf{b}$ assigns to the empty fuzzy set $\bar{0}$. Indeed, in this case we have $\rho_{\bar{0}}(a) > 0$, and hence

$$\mathbf{b}(\bar{0}) = s(\rho_{\bar{0}}) = \sum_{b \in \text{spt } \mu} \rho_{\bar{0}}(b)\mu(\{b\}) \geq \rho_{\bar{0}}(a) \cdot \mu(\{a\}) > 0.$$

Notice that the more inconsistent the focal elements of $\mathbf{b}$ are, the greater is the value $\mathbf{b}(\bar{0})$. When events and focal elements are crisp sets (and hence the unique possible non-normalized focal element is $\bar{0}$), normalization consists in redistributing the mass that μ assigns to $\bar{0}$ to the other focal elements of μ (if any).

Dealing with fuzzy-focal elements makes it possible to introduce a notion of *soft normalization* for belief functions. In particular, this construction allows for a finer redistribution of the masses, which depends on two thresholds. Recall the notion of height $h(a)$ of a fuzzy set a introduced in (13). Then we introduce the following definition of α-normalization.

Definition 6.1. A mass assignment $\mu : [0,1]^X \rightarrow [0,1]$ is said to be α-*normalized* provided that $\inf\{h(a) : a \in$ spt $\mu\} = \alpha$.

---

[2]If spt μ is not countable, the condition $a \in$ spt μ does not guarantee $\mu(\{a\}) > 0$.

In other words, a mass is α-normalized provided that each focal element of μ has at least height α. In particular, for a belief function $\mathbf{b}$ we define the *degree of normalization* of $\mathbf{b}$ as the value

$$\inf\{h(a) : a \in \mathrm{spt}\ (\mu)\},$$

where μ is the mass associated to $\mathbf{b}$.

Let now $\mu : [0,1]^X \to [0,1]$ be an α-normalized mass assignment, and assume that there exists a focal element b for m such that $h(b) = \beta > \alpha$.

The mass μ can be renormalized to the higher degree β by defining a new mass μ^β as follows: for every $a \in [0,1]^X$,

$$\mu^\beta(\{a\}) = \begin{cases} 0, & \text{if } h(a) < \beta \\ \frac{\mu(\{b\})}{1-K}, & \text{otherwise} \end{cases} \tag{17}$$

where $K = \sum_{h(l)<\beta} \mu(\{l\})$.

The idea of this β-*normalization*, similarly to the classical normalization, consists in fixing the value β as a new level of consistency for the mass we are considering. Since $\alpha < \beta \le 1$, the class of focal elements of height lower then β is not empty. Then the process of β-normalization consists in redistributing all the mass which μ assigns to the fuzzy sets of height lower than β, i.e. $K = \sum_{h(l)<\beta} \mu(\{l\})$, to those focal elements of height greater of (or equal to) β.

Clearly, every mass μ can be renormalized up to a maximum value given by

$$\beta_{max} = \sup\{h(a) : a \in \mathrm{spt}\ (\mu)\}.$$

We will make use of β-normalization in the next section where we discuss a generalization of the Dempster rule of combination.

7   Generalized Dempster's rule of combination

The power of Dempster-Shafer theory is in the possibility of combining all the available evidences about an event. In order to describe this aggregation process, Dempster introduced in [13] the so called *Dempster rule of combination*, briefly recalled next. Given a frame of discernment X, consider two masses m_1 and m_2 on 2^X encoding the beliefs about evidences coming from two (possibly different) sources. Then the new mass assignment m on 2^X is obtained from m_1 and m_2 according to the Dempster rule is as follows: for every $Y \subseteq X$,

$$m(Y) = \sum_{A\cap B=Y} m_1(A) \cdot m_2(B) \tag{18}$$

This rule may result in a non-normalized mass assignment as soon as there exist focal elements A and B for m_1 and m_2 respectively such that $A \cap B = \emptyset$. The normalized version of the rule yields the mass assignment m' defined as $m'(\emptyset) = 0$ and for every $\emptyset \ne Y \subseteq X$,

$$m'(Y) = \frac{m(Y)}{1 - \left(\displaystyle\sum_{C \cap D = \emptyset} m_1(C) \cdot m_2(D) \right)} \ . \tag{19}$$

Let us illustrate this situation with the following well known example due to Smets [52] (see also [49]): Mr Jones has been murdered, and we know the murderer was in the set $X = \{Peter, Paul, Mary\}$. The only evidence we have is that Mrs Jones, who saw the killer leaving the scene of the murder, is 80% sure that the murderer is a man. Hence all the available evidence is expressed as $Prob(Man) = 0.8$.

As recalled, within Dempster-Shafer theory each piece of evidence is encoded by a mass assignment $m : 2^X \to [0, 1]$, assigning a value to each subset of X such that $\sum_{A \in 2^X} m(A) = 1$ and $m(\emptyset) = 0$ (i.e. there is no belief on the empty set).

Turning back to the case of Mr. Jones' murder, and since we know that $Prob(Man) = 0.8$, the set $\{Paul, Peter\}$ is a focal element of mass m_1, and in particular we assign $m_1(\{Paul, Peter\}) = 0.8$. We know nothing about the remaining probabilities, so we allocate the remaining mass 0.2 to the whole frame of discernment X (i.e. $m_1(X) = 0.2$). Therefore, the only focal elements in this example are $X = \{Peter, Paul, Mary\}$ and $\{Paul, Peter\}$.

Consider a possible second piece of evidence providing an alibi for Peter with confidence 0.6. This new information is hence encoded in the model with a new mass assignment m_2 such that $m_2(\{Paul, Mary\}) = 0.6$ and $m_2(X) = 0.4$.

The Dempster rule of combination given by (18) then provides a new mass assignment m resulting from the combination of m_1 and m_2:

$$
\begin{aligned}
m(\{\emptyset\}) &= 0; \\
m(\{Paul\}) &= 0.48; \\
m(\{Paul, Mary\}) &= 0.12; \\
m(\{Peter, Paul\}) &= 0.32; \\
m(\{X\}) &= 0.08.
\end{aligned}
$$

Notice that this mass is normalized and hence it coincides with the mass resulting from (19). From the combined mass assignment m we can compute the resulting belief function $\mathbf{b}_m : 2^X \to [0, 1]$ as follows: for every $Y \subseteq X$

$$\mathbf{b}_m(Y) = \sum_{Z \subseteq Y} m(Z). \tag{20}$$

The previous formula yields for instance:

$$
\begin{aligned}
\mathbf{b}_m(\{Paul\}) &= 0.48; \\
\mathbf{b}_m(\{Paul, Mary\}) &= 0.6; \\
\mathbf{b}_m(\{Paul, Peter\}) &= 0.8; \\
\mathbf{b}_m(X) &= 1.
\end{aligned}
$$

In the example of Mr Jones' murder, the masses m_1 and m_2 were assigned to statements expressing precise properties regarding the individuals in the set

of hypothesis X. On the other hand, we may argue that usually a description of the witness would be affected not only by uncertainty regarding the statements, but also by the imprecision of the statements. Therefore the classical framework would be insufficient to analyse the facts provided by the witness.

In [25] the authors present a generalization of the Dempster rule in order to combine the information carried by two belief functions $\mathbf{b}_1, \mathbf{b}_2 \in Bel([0,1]^X)$ into a single belief function $\mathbf{b}_{1,2} \in Bel([0,1]^X)$. In the rest of this section we will recall the basic steps of that construction and we will conclude with some remarks about the procedure. We start with an easy result about the definition of states in a product space needed in the construction of the generalized Dempster rule.

Proposition 7.1. *For every MV-algebra $\mathbf{A}$, and for every pair of states $s_1, s_2 :$ $\mathbf{A} \to [0,1]$, there exists a state $s^{1,2}$ defined on the direct product $\mathbf{A} \times \mathbf{A}$ such that for every $(b,c) \in \mathbf{A} \times \mathbf{A}$, $s^{1,2}(b,c) = s_1(b) \cdot s_2(c)$.*

Let now $\mathbf{A} = [0,1]^X$, and let $\mathbf{R}_X$ be the MV-algebra defined in Section 4.2. Further, let s_1, s_2 be two state assignments on $\mathbf{R}_X$ such that $\mathbf{b}_1(a) = s_1(\rho_a)$ and $\mathbf{b}_2(a) = s_2(\rho_a)$ for all $a \in \mathbf{A}$. Assume that $\mu_1, \mu_2 : \mathfrak{B}(\mathbf{A}) \to [0,1]$ are the two regular probability measures of support spt μ_1 and spt μ_2, respectively, such that for $i = 1,2$,

$$s_i(f) = \int_{\text{spt } \mu_i} f \, \mathrm{d}\mu_i, \quad f \in \mathbf{R}_X.$$

Let

$$\mu_{1,2} : \mathfrak{B}(\mathbf{A} \times \mathbf{A}) \to [0,1]$$

be the product measure on Borel subsets generated by $\mathbf{A} \times \mathbf{A}$ and $s_{1,2}$ be the state on the MV-algebra of all measurable functions $\mathbf{A} \times \mathbf{A} \to [0,1]$ that is defined as an integral with respect to $\mu_{1,2}$.

Notice that $s_{1,2}$ actually coincides with the state $s^{1,2}$ on $\mathbf{R}_X \times \mathbf{R}_X$ defined as in Proposition 7.1 through the identification

$$s^{1,2}(g,h) = \int_{\mathfrak{B}(\mathbf{A})} g \, \mathrm{d}\mu_1 \cdot \int_{\mathfrak{B}(\mathbf{A})} h \, \mathrm{d}\mu_2 = \int_{\mathfrak{B}(\mathbf{A} \times \mathbf{A})} g \cdot h \, \mathrm{d}\mu_{1,2} = s_{1,2}(g \cdot h).$$

Hence, in particular, for any $g, h : \mathbf{A} \to [0,1]$ and f such that $f(x,y) = g(x) \cdot h(y)$, then Proposition 7.1 yields $s_{1,2}(f) = s_1(g) \cdot s_2(h)$.

Finally, for every $a \in \mathbf{A}$, consider the map $\rho_a^\wedge : \mathbf{A} \times \mathbf{A} \to [0,1]$ defined by

$$\rho_a^\wedge(b,c) = \rho_a(b \wedge c).$$

Then we are ready to define the following combination of belief functions.

Definition 7.2 (Generalized Dempster rule)**.** Given $\mathbf{b}_1, \mathbf{b}_2 \in Bel(\mathbf{A})$ as above, define its min-*conjunctive combination* $\mathbf{b}_{1,2} : \mathbf{A} \to [0,1]$ as follows: for all $a \in \mathbf{A}$,

$$\mathbf{b}_{1,2}(a) = s_{1,2}(\rho_a^\wedge). \tag{21}$$

Regarding the support of the combined measure, it is worth noticing that by [29, Theorem 417C (v)], spt $\mu_{1,2} = $ spt $\mu_1 \times$ spt μ_2, and hence, if μ_1 and μ_2 are normalized in the sense that their support is included into $\mathcal{NF}(X)$, then spt $\mu_{1,2} \subseteq \mathcal{NF}(X)$ as well. Therefore, by Proposition 5.4 one might deduce that, if $\mathbf{b}_1$ and $\mathbf{b}_2$ are normalized belief functions, then $\mathbf{b}_{1,2}$ is normalized as well. The following example shows that this is not the case, since in the definition of $\mathbf{b}_{1,2}$, together with the product measure $\mu_{1,2}$, we also use the map $\rho^\wedge$ which, in fact, is not a genuine fuzzy-inclusion operator.

Example 7.3. Consider two belief functions $\mathbf{b}_1$ and $\mathbf{b}_2$ on $[0,1]^2$ with masses concentrated as follows:

$$\mu_1(1,0) = 1/4; \, \mu_1(1,1) = 3/4; \quad \mu_2(0,1) = 1/2; \, \mu_2(1,1) = 1/2.$$

Then, the product measure $\mu_{1,2}$ has support in the cartesian product of the supports of the two masses:

$$\{((1,0),(0,1)),((1,0),(1,1)),((1,1),(0,1)),((1,1),(1,1))\},$$

and it takes the following values:

$$\begin{aligned}
\mu_{1,2}((1,0),(0,1)) &= 1/8, \\
\mu_{1,2}((1,0),(1,1)) &= 1/8, \\
\mu_{1,2}((1,1),(0,1)) &= 3/8, \\
\mu_{1,2}((1,1),(1,1)) &= 3/8.
\end{aligned}$$

So, $\mu_{1,2}$ is normalized in the sense that each of its focal elements can be regarded as a normal fuzzy set in $[0,1]^2 \times [0,1]^2$. On the other hand, $\mathbf{b}_{1,2}$ is non-normalized: indeed, since $(0,0) = (1,0) \wedge (0,1)$, $\rho_{(0,0)}(0,0) = 1$ and $\rho_{(0,0)}(b,c) = 0$ for focal $(b,c) \neq (0,0)$, we have

$$\mathbf{b}_{1,2}(0,0) = \sum_{b,c:b\wedge c=(0,0)} \rho_{(0,0)}(b \wedge c)) \cdot \mu_{1,2}(b,c) = \mu_{1,2}((1,0),(0,1)) = 1/8 > 0.$$

The example we presented at the beginning of this section —the murder of Mrs Jones [52]— can be adapted to the context of belief functions on fuzzy sets as follows.

Example 7.4. Recall the 3 suspects of Mr. Jones' murder: Peter, Paul, and Mary. Consider the information provided by Mrs. Jones, she heard his husband yelling and the person she saw running was *a man* . It turns out that Mary has short hair, so she may be mistaken for a man at first sight, and hence the set of suspects looking like a man can be considered fuzzy as well, with membership function:

$$\mu_{man\text{-}like}(Peter) = 1, \, \mu_{man\text{-}like}(Paul) = 1, \, \mu_{man\text{-}like}(Mary) = 0.5.$$

The evidence supplied by Mrs Jones may be represented by a mass assignment $m_1 : [0,1]^X \to [0,1]$ such that $m_1(man\text{-}like) = \alpha > 0$, $m_1(X) = 1 - \alpha$ and $m_1(a) = 0$ for any other $a \in [0,1]^X$.

A second piece of evidence is provided by the janitor living in the same house, who reports that he saw in the darkness a small person quickly leaving the scene of the crime. Paul and Mary are not tall while Peter is taller (Paul is 1.65 m tall, Mary is 1.60 m tall and Peter is 1.85 m). So, actually, the subset of small suspects of $X = \{Peter, Paul, Mary\}$ can be also considered as a fuzzy set, with membership function μ_{small} given by, say,

$$\mu_{small}(Peter) = 0, \ \mu_{small}(Paul) = 0.7, \ \mu_{small}(Mary) = 0.9.$$

The evidence supplied by the janitor may be represented by a second mass assignment $m_2 : [0,1]^X \to [0,1]$ such that $m_2(small) = \beta > 0$, $m_2(X) = 1 - \beta$ and $m_2(a) = 0$ for any other $a \in [0,1]^X$.

Let us compute the resulting mass by combining m_1 and m_2 by means of the generalized Dempster rule according to Definition 7.2:

$$\mathbf{b}_{1,2}^{\wedge}(a) \quad = \quad \sum_{b,c \in \{man\text{-}like, small, X\}} \rho_a(b \wedge c) \cdot m_1(b) \cdot m_2(c) \ .$$

Here, the membership function of $small \wedge man\text{-}like$ (interpreting $\wedge$ by the min) is given by

$$\begin{aligned}
\mu_{small \wedge man\text{-}like}(Peter) &= & 0; \\
\mu_{small \wedge man\text{-}like}(Paul) &= & 0.7; \\
\mu_{small \wedge man\text{-}like}(Mary) &= & 0.5.
\end{aligned}$$

Suppose we are interested in computing the belief that the suspect is Paul. We then need to compute:

$$\begin{aligned}
\rho_{\{Paul\}}^{\wedge}(small \wedge man\text{-}like) &= & \min_{x \in X}\{\mu_{small \wedge man\text{-}like}(x) \Rightarrow \mu_{Paul}(x)\} \\
&= & \min\{0 \Rightarrow 0, 1 \Rightarrow 1, 0.5 \Rightarrow 0\} \\
&= & \min\{1, 0.5\} = 0.5
\end{aligned}$$

$$\begin{aligned}
\rho_{\{Paul\}}^{\wedge}(small \wedge X) &= & \min_{x \in X}\{\mu_{small}(x) \Rightarrow \mu_{Paul}(x)\} \\
&= & \min\{0 \Rightarrow 0, 0.7 \Rightarrow 1, 0.9 \Rightarrow 0\} \\
&= & \min\{1, 0.1\} = 0.1
\end{aligned}$$

$$\begin{aligned}
\rho_{\{Paul\}}^{\wedge}(X \wedge man\text{-}like) &= & \min_{x \in X}\{\mu_{man\text{-}like}(x) \Rightarrow \mu_{Paul}(x)\} \\
&= & \min\{1 \Rightarrow 0, 1 \Rightarrow 1, 0.5 \Rightarrow 0\} \\
&= & \min\{0, 1, 0.5\} = 0
\end{aligned}$$

and $\rho^{\wedge}_{\{Paul\}}(X) = 0$. Finally, we have

$$
\begin{aligned}
\mathbf{b}^{\wedge}_{1,2}(\{Paul\}) &= \sum_{a \in [0,1]^X} \rho^{\wedge}_{\{Paul\}}(b \wedge c) \cdot m_1(b) \cdot m_2(c) \\
&= \rho^{\wedge}_{\{Paul\}}(small \wedge man\text{-}like) \cdot m_1(man\text{-}like) \cdot m_2(small) + \\
&\quad\ \rho^{\wedge}_{\{Paul\}}(small) \cdot m_1(X) \cdot m_2(small) \\
&= 0.5 \cdot \alpha \cdot \beta + 0.1 \cdot (1 - \alpha) \cdot \beta > 0.
\end{aligned}
$$

Hence, we get a positive belief degree of Paul being the murderer. This is in contrast with the results we would obtain, in case we assume Mary can be mistaken for a man, with both the classical model and the crisp-focal model, where focal elements are only allowed to be classical subsets of X. Indeed, in that case, we would be forced to take as focal element for m_1, besides X itself, the set $man\text{-}like = \{Paul, Mary\}$, and since there would be no focal element included in $\{Paul\}$, we would get $\mathbf{b}^{\wedge}_{1,2}(\{Paul\}) = 0$. $\qquad\square$

The above min-conjunctive combination can easily be extended to well-known MV-operations on fuzzy sets, such as max-disjunction $\vee$, strong conjunction $\odot$ and strong disjunction $\oplus$, by defining

$$
(b_1 \circledast b_2)(a) = s_{1,2}(\rho^{\circledast}_a),
$$

for $\circledast$ being one of these operations, and defining

$$
\rho^{\circledast}_a(b, c) = \rho_a(b \circledast c).
$$

In this generalized case, the map $\mathbf{b}^{\circledast}_{1,2}$ resulting from the respective combination rule will be called the $\circledast$-*combination of* $\mathbf{b}_1$ *and* $\mathbf{b}_2$.

Whenever the supports of μ_1 and μ_2 are countable, it is easy to prove that $\mathbf{b}^{\circledast}_{1,2}$ is a belief function in the sense of Definition 4.9. In fact, in this case Definition 7.2 yields

$$
\mathbf{b}^{\circledast}_{1,2}(a) = \sum_{b,c \in \mathbf{A}} \rho_a(b \circledast c) \cdot \mu_1(\{b\}) \cdot \mu_2(\{c\}). \tag{22}
$$

Notice that (22) reduces to

$$
\mathbf{b}^{\circledast}_{1,2}(a) = \sum_{d \in \mathbf{A}} \sum_{\substack{b,c \in \mathbf{A} \\ b \circledast c = d}} \rho_a(d) \cdot (\mu_1(\{b\}) \cdot \mu_2(\{c\})) = \sum_{d \in \mathbf{A}} \rho_a(d) \cdot \mu^*(\{d\}),
$$

where

$$
\mu^*(\{d\}) = \sum_{\substack{b,c \in \mathbf{A} \\ b \circledast c = d}} \mu_1(\{b\}) \cdot \mu_2(\{c\})
$$

is indeed a mass assignment and hence $\mathbf{b}^{\circledast}_{1,2} \in Bel([0,1]^X)$. Therefore, turning back to the above Example 7.3 and Proposition 5.4, there exists a mass $\mu \neq \mu_{1,2}$ for $\mathbf{b}^{\circledast}_{1,2}$ such that spt $\mu \not\subseteq \mathcal{NF}(X)$.

8   Conclusions and further reading

In this chapter we have discussed several ways to define belief functions on MV-algebras of fuzzy sets based on some previous results by the authors [41, 42, 25, 28]. In particular we have surveyed two main frames in which belief functions on fuzzy sets are characterized by the fact that focal elements are either crisp or fuzzy sets. We have then studied the normalization and soft-normalization problem together with a generalization of Dempster's rule of combination.

Another logical-based approach, extending the one in [30] for classical events, has been introduced in [27], where a modal logic for belief functions on an MV-algebra has been presented.

Belief functions on Boolean algebras can also be described in geometrical terms: in [9, 10] the author presents several results in this directions. In a similar way, the problem of extending a partial assignment over formulas of Łukasiewicz logic can be characterized in geometrical terms by combining tropical-idempotent convex geometry and classical Euclidean convex geometry as well. The paper [22] studies this geometrical foundation for belief functions on MV-algebras.

The problem of extending a partial assignment to a probability measure is well known in the literature as de Finetti's coherence criterion [11, 12]. As for belief functions on Boolean events, a similar interpretation of belief functions in terms of betting scheme has been presented in [38] and there is some ongoing work for the case of fuzzy events [23].

Acknowledgements.

References

[1] L. P. Belluce. Semisimple Algebras of Infinite Valued Logic and Bold Fuzzy Set Theory. *Canad. J. Math.* 38(6): 1356–1379, 1986.

[2] L. Biacino. Fuzzy subsethood and belief functions of fuzzy events. *Fuzzy Sets and Systems* 158(1): 38–49, 2007.

[3] W. J. Blok and D. Pigozzi. Algebraizable logics. *Memoirs of the American Mathematical Society*, 396, Vol. 77, 1989.

[4] D. Butnariu and E. P. Klement. *Triangular Norm Based Measures and Games with Fuzzy Coalitions.* Kluwer, Dordrecht, 1993.

[5] T. Calvo, G. Mayor, and R. Mesiar. (Editors) *Aggregation Operators – New Trends and Applications*, volume 97 of Studies in Fuzziness and Soft Computing. Springer Verlag, 2002.

[6] C. C. Chang. Algebraic Analysis of Many-valued Logics. *Trans. Am. Math. Soc.* 88: 467–490, 1958.

[7] R. Cignoli, I.M.L. D'Ottaviano and D. Mundici. *Algebraic Foundations of Many-valued Reasoning.* Kluwer, Dordrecht, 2000.

[8] P.M. Cohn. *Universal Algebra.* Revisited Edition, D. Reidel Pub. Co., Dordrecht, 1981.

[9] F. Cuzzolin. A Geometric Approach to the Theory of Evidence. *IEEE Transactions on Systems, Man, and Cybernetics*, Part C 38(4): 522-534, 2008.

[10] F. Cuzzolin. *The geometry of uncertainty.* Springer-Verlag, Information Science and Statistics Series, 2012.

[11] B. de Finetti. Sul significato soggettivo della probabilità. *Fundamenta Mathematicae* 17:298–329, 1931. Translated into English as "On the subjective meaning of probability", in: Paola Monari and Daniela Cocchi (Eds.), Probabilità e Induzione, Clueb, Bologna, 291–321, 1993.

[12] B. de Finetti. *Theory of Probability*, Vol.1, Wiley, New York, 1974.

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

[14] D. Denneberg. *Non-additive measure and integral.* Theory and Decision Library. Series B: Mathematical and Statistical Methods, vol 27. Kluwer Academic Publishers Group, Dordrecht, 1994.

[15] T. Denœux. Reasoning with imprecise belief structures. *Int. J. Approx. Reasoning* 20(1): 79–111, 1999.

[16] T. Denœux. Modeling vague beliefs using fuzzy-valued belief structures. *Fuzzy Sets and Systems* 116(2): 167–199, 2000.

[17] M. Develin and B. Sturmfels. Tropical convexity. *Doc. Math.*, 9: 1–27, 2004.

[18] D. Dubois and H. Prade. On several representations of an uncertain body of evidence, in *Fuzzy Information and Decision Processes*, M. M. Gupta and E. Sanchez, Ed. New York: North- Holland, 167–181, 1982.

[19] D. Dubois and H. Prade. Evidence Measures Based on Fuzzy Information. *Automatica* 21(5): 547–562, 1985.

[20] D. Dubois and H. Prade. (Editors.) *Fundamentals of Fuzzy Sets*, The Handbooks of Fuzzy Sets Series, Kluwer, 2000.

[21] G. Edwald. *Combinatorial Convexity and Algebraic Geometry.* Springer, New York, 1996.

[22] T. Flaminio and L. Godo, A note on the convex structure of uncertainty measures on MV-algebras. In *Advances in Intelligent and Soft Computing* 190 (Springer), R. Kruse et al. (Eds.): Synergies of Soft Computing and Statistics for Intelligence Data Analysis, 73–82, 2013.

[23] T. Flaminio and L. Godo. Towards a betting interpretation for belief functions on MV-algebras. In *Non-Classical Measures and Integrals: Abstracts of the 34th Linz Seminar on Fuzzy Set Theory* (Linz 2013), R. Mesiar, E. Pap and E.P. Klement (eds.), Linz, Austria, 51-53, 2013.

[24] T. Flaminio, L. Godo and E. Marchioni. On the logical formalization of possibilistic counterpart of states over n-valued Łukasiewicz events. *J. Logic Comput.* 21(3): 447–464, 2011.

[25] T. Flaminio, L. Godo and E. Marchioni. Belief Functions on MV-Algebras of Fuzzy Events Based on Fuzzy Evidence. In *Proceedings of ECSQARU 2011*, Lecture Notes in Artificial Intelligence 6717, Weiru Liu (Ed.): 628–639, 2011.

[26] T. Flaminio, L. Godo and E. Marchioni. Geometrical aspects of possibility measures on finite domain MV-clans. *Soft Computing.* 16(11): 1863–1873, 2012.

[27] T. Flaminio, L. Godo and E. Marchioni. Logics for belief functions on MV-algebras. *International Journal of Approximate Reasoning*, 54(4): 491–512, 2013.

[28] T. Flaminio, L. Godo and T. Kroupa. Combination and Soft-Normalization of Belief Functions on MV-Algebras. In *Proceedings of MDAI 2012*, V. Torra et al. (Eds.), Lecture Notes in Computer Science 7647, 23–34, 2012.

[29] D. H. Fremlin. *Measure theory. Vol. 4.* Torres Fremlin, Colchester, 2006. Topological measure spaces. Part I, II, Corrected second printing of the 2003 original.

[30] L. Godo, P. Hájek and F. Esteva. A Fuzzy Modal Logic for Belief Functions. *Fundamenta Informaticae* 57(2-4): 127–146, 2003.

[31] K. R. Goodearl. *Partially Ordered Abelian Group with Interpolation.* AMS Math. Survey and Monographs 20, 1986.

[32] S. Gottwald. *Fuzzy sets and fuzzy logic: the foundations of application–from a mathematical point of view.* Teknea, 1993.

[33] M. Grabisch. Belief Functions on Lattices. *Int. J. Intell. Syst.* 24(1): 76–95, 2009.

[34] M. Grabisch, T. Murofushi and M. Sugeno. Fuzzy Measure of Fuzzy Events Defined by Fuzzy Integrals. *Fuzzy Sets and Systems* 50: 293-313, 1992.

[35] P. Hájek. *Metamathematics of fuzzy logic*, volume 4 of *Trends in Logic— Studia Logica Library*. Kluwer Academic Publishers, Dordrecht, 1998.

[36] J. H. Halpern. *Reasoning about Uncertainty*. MIT Press, 2003.

[37] C. Hwang and M. Yang. Generalization of Belief and Plausibility Functions to Fuzzy Sets Based on the Sugeno Integral. *International Journal of Intelligent Systems* 22: 1215–1228, 2007.

[38] J.-Y. Jaffray. Coherent bets under partially resolving uncertainty and belief functions. *Theory and Decision*, 26: 99–105, 1989.

[39] G. J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, 1995.

[40] T. Kroupa. Every state on semisimple MV-algebra is integral. *Fuzzy Sets and Systems*, 157: 2771–2782, 2006.

[41] T. Kroupa. From Probabilities to Belief Functions on MV-algebras, in *Combining Soft Computing and Statistical Methods in Data Analysis*, volume 77 of Advances in Soft Computing: 387–394. Springer-Verlag, 2010.

[42] T. Kroupa. Extension of Belief Functions to Infinite-valued Events. *Soft Computing*, 16(11): 1851–1861, 2012.

[43] G. L. Litvinov. Maslov Dequantization, Idempotent and Tropical Mathematics: a Brief Introduction. *Journal of Mathematical Science* 140 (3): 426–444, 2007.

[44] D. Mundici. Averaging the truth-value in Łukasiewicz logic. *Studia Logica*, 55(1): 113–127, 1995.

[45] D. Mundici. *Advanced Łukasiewicz calculus and MV-algebras*. Trends in Logic, vol 35. Springer, 2011.

[46] M. Navara. Triangular norms and measures of fuzzy sets. In E.P. Klement and R. Mesiar, editors, *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*, 345–390. Elsevier, 2005.

[47] G. Panti. Invariant measures on free MV-algebras. *Communications in Algebra* 36(8): 2849–2861, 2008.

[48] J. B. Paris. A note on the Dutch Book method, Revised version of a paper of the same title which appeared in *Proceedings of the Second Internat. Symp. on Imprecise Probabilities and their Applications*, ISIPTA 2001, Ithaca, New York, 2001.

[49] S. Parsons. Some qualitative approaches to applying the Dempster-Shafer theory. *Information and Decision Technologies* 19: 321–337, 1994.

[50] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton 1976.

[51] P. Smets. The Degree of Belief in a Fuzzy Event. *Information Sciences*, 25: 1–19, 1981.

[52] P. Smets. Belief Functions. In *Nonstandard Logics for Automated Reasoning*, P. Smets et al. (eds.), Academic Press, London, 253–277, 1988.

[53] M. Sugeno. *Theory of fuzzy integrals and its applications*. Phd. Dissertation, Tokyo Institute of Technology, Tokyo, Japan, 1974.

[54] S. Weber. ⊥-decomposable measures integrals for Archimedean t-conorms ⊥. *J. Math. Anal. Appl.* 101: 114–138, 1984.

[55] J. Yen. Generalizing the Dempster-Shafer theory to fuzzy sets. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(3): 559–570, 1990.

[56] J. Yen. Computing Generalized Belief Functions for Continuous Fuzzy Sets. *Int. J. Approx. Reasoning* 6: 1–31, 1992.

[57] L. A. Zadeh. Fuzzy sets. *Information and Control* 8: 338–353, 1968.

[58] L. A. Zadeh. Probability measures of fuzzy events, *J. Math. Anal. Appl.* 23: 421–427, 1968.

[59] L. A. Zadeh. Fuzzy sets and information granularity. In *Advances in Fuzzy Sets Theory and Applications*, (M. Gupta et al. eds), North Holland: 3–18, 1979.

[60] H. J. Zimmermann. Fuzzy sets in operational research. *European Journal of Operational Research* 13: 201–216, 1983.

[61] C. Zhou. Belief functions on Distributive Lattices, *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*: 1968–1974, 2012.