Probabilistic Analysis of Dempster-Shafer Theory. Part Three

Kramosil, Ivan
1999

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Probabilistic Analysis of Dempster–Shafer Theory
Part Three

Ivan Kramosil

Technical report No. 777

April, 1999

Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic
phone: (+4202) 6605 3940 fax: (+4202) 8585789
e-mail: kramosil@uivt.cas.cz

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Probabilistic Analysis of Dempster–Shafer Theory
Part Three

Ivan Kramosil[1]

Technical report No. 777
April, 1999

Abstract

Dempster–Shafer theory is an interesting and useful mathematical tool for uncertainty quantification and processing. From one point of view it can be seen as an alternative apparatus to probability theory and mathematical statistics based on this probability calculus, as D.–S. theory can be developed in a way quite independent of probability theory, beginning with a collection of more or less intuitive demands which an uncertainty degree calculus should meet. On the other side, however, D.–S. theory can be developed also as a particular sophisticated application of probability theory, using the notion of non–numerical, in particular, set–valued random variables (random sets) and their numerical characteristics. This later aspect enables to generalize D.–S. theory beyond its classical scopes using appropriately the apparatus of probability theory and measure theory.

This report is the third part of a surveyal work cumulating, and presenting in a systematic way, some former author's ideas and achievements dealing with applications of probability theory and mathematical statistics when defining, developing, and generalizing various parts of D.–S. theory. The more detailed contents of this report can be understood from the list of the titles of the particular chapters presented just below.

Keywords

Dempster–Shafer theory, probability theory, belief function, random variable, random set

[1]This work has been sponsored by the grant no. A1030504 of the GA AS CR.

Contents

11 Signed Belief Functions and Belief Functions with Nonstandard Values

Both signed belief functions and belief functions with nonstandard values generalize the notion of belief function in the sense that the domain of this function is the same as in the classical case, i.e., the field of all subset of a nonempty set S (as a rule, we shall limit ourselves to finite sets S), but the values are either real numbers including those beyond the scope of the unit interval $\langle 0, 1\rangle$, or even some objects from a more complicated structure. A theoretical motivation for such generalization can be given by our attempt to define an operation inverse to the Dempster combination rule $\oplus$, i.e., to define an operation $\ominus$ such that, given basic probability assignments (b.p.a.'s) m_1 and m_2 on S, the equality $((m_1 \oplus m_2) \ominus m_2)(A) = m_1(A)$ would hold for all $A \subset S$. Although the problem is stated at a purely theoretical and algebraical level, it possesses an intuitive interpretation which is perhaps worth being discussed in more detail. Let us consider a subject whose degrees of belief concerning the membership of the actual state of the investigated system in particular subsets of the set S of all states are quantified by a basic probability assignment (b.p.a.) m_1 and by the corresponding belief function bel_{m_1}. The subject combines her/his beliefs with the beliefs of her/his colleague quantified by a b.p.a. m_2 and by bel_{m_2}, so that she/he obtains the actualized beliefs quantified by $m_1 \oplus m_2$ and by $bel_{m_1} \oplus bel_{m_2}$, and completely forgets the original beliefs m_1 and m_2, erasing them totally from her/his memory. Later, however, she/he obtains a new piece of information saying that the former information given by the second subject was completely irrelevant, unreliable and *in this sense* wrong. As the first subject takes, for no matter which reason, this new information as more reliable than that offered sooner by the second subject, the first subject wants to cancel, somehow, the impact of m_2 to her/his beliefs and to turn back to the original beliefs m_1. If she/he wants to realize such a cancellation by an application of the Dempster combination rule, she/he must express the reliable information claiming the nonreliability of the second subject by the means of a b.p.a. m_3 such that $(m_1 \oplus m_2) \oplus m_3 \equiv m_1$. As the operation $\oplus$ is associative, $(m_1 \oplus m_2) \oplus m_3 \equiv m_1 \oplus (m_2 \oplus m_3) \equiv m_1 \oplus m_S$ should be valid, so that the problem converts into that to find m_3 such that $m_2 \oplus m_3 \equiv m_S$ (let us recall that $m_S(S) = 1$, $m_S(A) = 0$ for each $A \subset S$, $A \neq S$). As can be easily seen, up to trivial cases when $m_2 \equiv m_S$ holds (and in this case also $m_3 \equiv m_S$ holds), this problem is unsolvable at least within the space of b.p.a.'s and belief functions as defined above.

Let us illustrate the problem of "deconditionalization" or "de-actualization" by the most simple case. Let the first subject know nothing about the problem in question which she/he is to solve so that her/his beliefs are quantified by the vacuous b.p.a. m_S expressing just the assumption of closed world. Then the subject obtains a new information saying that the actual state of the system is in a proper subset A of S. Hence, she/he combines m_S and m_A ($m_A(A) = 1$, $m_A(B) = 0$ for all $B \subset S$, $B \neq A$) by the Dempster rule and obtains $m_S \oplus m_A = m_A$. Later, a new piece of information arrives saying that the last information was not true. This can be taken, however, in the two different ways.

Either, it is taken as the information saying that the *negation of the former information holds true*, i. e., as the information saying that the actual state of the system is not in A, consequently, due to the closed world assumption, that it is in $S - A$. This information is quantified by the b.p.a. m_{S-A}, however, combining $m_S \oplus m_A \equiv m_A$ with m_{S-A} we obtain the contradiction and we cannot escape from this contradiction if no matter which further information is combined with m_S, m_A and m_{S-A}.

Another and more acceptable, in our context, interpretation of the meta-information "the information contained in m_A is not true" is to take this information as *completely irrelevant* in the sense that all consequences possibly drawn from the information that the actual state of the system is in A should be cancelled and the state of the subject's beliefs should turn back to the previous state, i. e., to the state expressed by the belief function m_S in our most simplified exmaple. This is just the case which should be expressed formally by combining m_A with its inverse element m_A^{-1} in such a way that the result should be m_S, and to this case we shall orient our effort in the rest of this chapter. Before going on with an explanation of the solution proposed in this chapter let us refer the reader to [44] or [45], where the algebraic properties of b.p.a.'s, belief functions and Dempster operations are discussed in more detail. Besides the mathematical and methodological motivations for an operation inverse to the Dempster combination rule there are also motivations of a very practical nature connected with the HUGIN expert system (cf. [1] for more detail). This field of investigation seems to be very interesting and useful, but because of a limited extent of this chapter and because of the declared theoretical and formally mathematical nature of this work as a whole we have to postpone such a research till another occasion.

When facing the negative result of our effort to define m^{-1} for non-vacuous b.p.a.'s m, our solution will follow the clasical paradigma applied already many times in mathematics: if some operation can be defined only partially within some structure, we shall extend the support of this structure by new objects in order to make the operation in question totally definable. The new problem is then to find an interpretation for the new objects as close as possible to the interpretation introduced before for the objects of the original structure. An alternative motivation can read as follows: to solve a problem it is often advantageous to embed it in a larger context, and to solve it here, if it is guaranteed that the solution itself belongs to the original domain (remember the use of complex numbers in physics, e. g.). For example, negative integers were defined in order to be able to define the operation of substraction as a total operation, and then these objects were interpreted as quantities for debts, altitudes below the sea surface, etc. The same was the story with rational, real and complex numbers. Namely, in our case of belief functions we shall take profit of the definition of b.p.a.'s and belief functions through set-valued random variables replacing, in this definition, the underlying notion of probability space by a more general space with signed measure.

From one side it is quite natural and legitimate to consider probability measures as functions taking their values in the unit interval of real numbers. Or, probabilities have been always conceived as idealized (from the philosophical point of view) or limit (from the mathematical point of view expressed by various laws of large numbers) values of relative frequences which are trivially embedded within the unit interval by definition (cf. [12] for a more detailed philosophical discussion). From the other side, however, probabilities are defined by measures, i. e., by functions quantifying numerically the sizes of some sets and obeying the common laws of such quantifications postulated as soon as in the antic Greece (cf. [49]). But these rules and laws allow to consider also size quantifications taking values outside the unit interval. Consequently, at least from the purely mathematical point of view taken as primary in this work, generalizations of probability measures which extend the scope of their possible values are worth considering. Let us introduce the formal definitions.

Definition 11.1. Let $\langle \Omega, \mathcal{A} \rangle$ be a measurable space, i. e., Ω is a nonempty set and $\mathcal{A}$ is a nonempty σ-field of subsets of Ω.

(i) A mapping $P : \mathcal{A} \rightarrow \langle 0, 1 \rangle$ is called *probability measure*, if it is σ-additive, i. e., $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ holds for each infinite sequence $A_1, A_2, \ldots$ of mutually disjoint sets from $\mathcal{A}$, and $P(\Omega) = 1$;

(ii) a mapping $\mu : \mathcal{A} \rightarrow R^+ = \langle 0, \infty) \cup \{\infty\}$ is called *measure*, if it is σ-additive and $\mu(\emptyset) = 0$ for the empty subset $\emptyset$ of Ω. Obviously, $\{\emptyset, \Omega\} \subset \mathcal{A}$ holds for each nonempty σ-field of subsets of Ω. The usual conventions concerning the arithmetical operations with the value ∞ are supposed to be adopted;

(iii) a mapping $\mu : \mathcal{A} \rightarrow R^* = (-\infty, \infty) \cup \{-\infty\} \cup \{\infty\}$ is called *signed measure*, if it is σ-additive, if $\mu(\emptyset) = 0$, and if it takes at most one of the values $-\infty$, ∞, i. e., if there are no sets A, $B \in \mathcal{A}$ such that $\mu(A) = \infty$ and $\mu(B) = -\infty$ in order to avoid expressions like $\infty - \infty$, cf. [18].

Let $\langle \Omega, \mathcal{A} \rangle$ be a measurable space. A triple $\langle \Omega, \mathcal{A}, \mu \rangle$ is called *probability space*, if μ is a probability measure on $\mathcal{A}$, it is called *space with (signed) measure*, if μ is a (signed) measure. A measurable mapping f which takes a probability space or a space with (signed) measure $\langle \Omega, \mathcal{A}, \mu \rangle$ into a measurable space $\langle X, \mathcal{X} \rangle$ is called *random variable*. Sometimes the term *generalized random variable* is used in this case preserving the expression "random variable" for the particular case of real-valued mappings measurable in the Borel sense, i. e., for the case when $\langle X, \mathcal{X} \rangle = \langle (-\infty, \infty), \mathcal{B} \rangle$.

The way to the notions of basic signed measure assignment and signed belief function mostly copies that one presented above for the probabilistic case just with probability space replaced by a space with signed measure and with a stronger simplifying consistence condition than in the probabilistic case. So, let S be a finite nonempty set of states of a system, let E be a space of possible values of empirical data and observations concerning the system and equipped by a nonempty σ-field of subsets of E denoted by $\mathcal{E}$, let $\rho : S \times E \to \{0,1\}$ be a compatibility relation, let $\langle \Omega, \mathcal{A}, \mu \rangle$ be a measurable space with a signed measure μ, and let $X : \langle \Omega, \mathcal{A}, \mu \rangle \to \langle E, \mathcal{E} \rangle$ be a random variable such that the composed mapping $U : \langle \Omega, \mathcal{A}, \mu \rangle \to \langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$ defined by $U(\omega) = U(X(\omega)) = \{s \in S : \rho(s, X(\omega)) = 1\}$, for each $\omega \in \Omega$, is a set-valued random variable taking as its values subsets of S. Besides the already presented simplifying assumption that S is finite we shall also suppose that signed belief functions are defined only if the composed mapping $U : \Omega \to \mathcal{P}(S)$ is *strongly consistent* in the sense that, for all $\omega \in \Omega$, $U(\omega) \neq \emptyset$. The reason is that we want to escape from difficulties connected with the necessity to introduce conditional signed measures and to investigate their properties which do not need copy the properties of conditional properties in the extent necessary for our purposes. The introduced stronger form of the consistence condition ($\{\omega \in \Omega : U(\omega) = \emptyset\} = \emptyset$ instead of $\mu(\{\omega \in \Omega : U(\omega) = \emptyset\}) = 0$) is involved by the fact that sets of zero measure are of different nature in spaces with signed measures and in probability spaces, e. g., a set of zero signed measure can be a union of two sets of non-zero measures.

Since now, we shall suppose that we have a space $\langle \Omega, \mathcal{A}, \mu \rangle$ with signed measure, a finite set S, and a strongly consistent set-valued random variable U taking $\langle \Omega, \mathcal{A}, \mu \rangle$ into $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$.

Definition 11.2. Let S be a nonempty finite set. *Basic signed measure assignment* (b.s.m.a.) defined on S is a mapping $m : \mathcal{P}(S) \to R^* = (-\infty, \infty) \cup \{-\infty\} \cup \{\infty\}$ such that m takes at most one of the infinite values $-\infty$, ∞. A b.s.m.a. m on S is *induced by a set-valued random variable* U defined on a space $\langle \Omega, \mathcal{A}, \mu \rangle$ with signed measure and taking its values in $\mathcal{P}(S)$, if $m(A) = \mu(\{\omega \in \Omega : U(\omega) = A\})$ for each $A \subset S$. If m is a b.s.m.a. on S induced by a strongly consistent U, then the *signed belief function induced by* m is the mapping $bel_m : \mathcal{P}(S) \to R^*$ defined by $bel_m(A) = \sum_{B \subset A} m(B)$ for each $A \subset S$. If m is defined by a strongly consistent set-valued random variable U, then obviously $bel_m(A) = \mu(\{\omega \in \Omega : U(\omega) \subset A\})$ for each $A \subset S$. A b.s.m.a. m (signed belief function bel_m, resp.) is called *finite*, if $-\infty < m(A) < \infty$ ($-\infty < bel_m(A) < \infty$, resp.) holds for each $A \subset S$. □

Lemma 11.1. Let m be a b.s.m.a. on a finite set S. Let $m^*(\mathcal{B}) = \sum_{A \in \mathcal{B}} m(A)$ for each $\emptyset \neq \mathcal{B} \subset \mathcal{P}(S)$, let $m^*(\emptyset) = 0$ for the empty subset of $\mathcal{P}(S)$. Then m^* is a signed measure on the measurable space $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$. □

Proof. As S and $\mathcal{P}(S)$ are finite spaces, σ-additivity of m^* coincides with finite additivity and this property follows immediately from the definition, as well as the equality $m^*(\emptyset) = 0$. As each $\mathcal{B} \subset \mathcal{P}(S)$ is finite, $m^*(\mathcal{B}) = \infty$ ($= -\infty$, resp.) can hold iff $m(A) = \infty$ ($= -\infty$, resp.) holds for at least one $A \in \mathcal{B}$. So, m^* can take only this infinite value which is taken by m, consequently, m^* can take at most one infinite value. Hence, m^* is a signed measure. □

The following statement proves that in the case of finite b.s.m.a.'s there exists a one-to-one relation between b.s.m.a.'s and signed belief functions like as in the case of basic probability assignments investigated above. On the other side, however, if m is a b.s.m.a. ascribing an infinite value to a proper subset of S, then there exists another b.s.m.a. m', $m' \not\equiv m$, generating the same belief function as m.

Lemma 11.2. Let m_1, m_2 be finite b.s.m.a.'s on a finite set S such that bel_{m_1} and bel_{m_2} are defined. If $m_1 \not\equiv m_2$, then $bel_{m_1} \not\equiv bel_{m_2}$, hence, if there exists $A \subset S$ such that $m_1(A) \neq m_2(A)$, then there exists $B \subset S$ such that $bel_{m_1}(B) \neq bel_{m_2}(B)$. If m is a b.s.m.a. on a finite set S such that bel_m is defined and $m(A) = \infty$ for a subset $A \subset S$, $A \neq S$, then there exists b.s.m.a. m_1 on S such that $m(B) \neq c\, m_1(B)$ for some $B \subset S$ and for all $-\infty < c < \infty$, but $bel_m(C) = bel_{m_1}(C)$ for all $C \subset S$. □

Proof. The proof for the case of finite m_1, m_2 over a finite set S is by induction on the cardinality of S like as in the case of basic probability assignments investigated above. Let

$$n_0 = \min\left\{n \in \mathcal{N}^+ = \{1, 2, \dots\} : (\exists\, A \subset S)\,(\text{card}A = n \,\&\, m_1(A) \neq m_2(A))\right\}; \qquad (11.1)$$

by the conditions imposed on m_1 and m_2 such an n_0, $1 \leq n_0 \leq \text{card}S$, is uniquely defined. Let $A \subset S$ be such that $\text{card}A = n_0$ and $m_1(A) \neq m_2(A)$. Then

$$\begin{aligned} bel_{m_1}(A) &= \sum\nolimits_{B \subset A} m_1(B) = \sum\nolimits_{B \subset A,\, B \neq A} m_1(B) + m_1(A) \neq \\ &\neq \sum\nolimits_{B \subset A,\, B \neq A} m_2(B) + m_2(A) = bel_{m_2}(A), \end{aligned} \qquad (11.2)$$

as $m_1(B) = m_2(B)$ for all $B \subset S$, $\text{card}B < n_0$, in particular, for all $B \subset A$, $B \neq A$, and $m_1(A) \neq m_2(A)$.

Let m be a b.s.m.a. on a finite set S, let $A \subset S$, $A \neq S$, be such that $m(A) = \infty$ (consequently, $-\infty < m(A)$ holds for each $A \subset S$). Let B, $A \subset B \subset S$ be such that $A \neq B$, as $A \neq S$, such a B always exists. Let $m_1(B) = \infty$, if $m(B) < \infty$, let $m_1(B) < \infty$ be chosen arbitrarily, if $m(B) = \infty$. Let $m_1(C) = m(C)$ for all $C \subset S$, $C \neq B$, so that $m(B) \neq c\, m_1(B)$ holds for all $-\infty < c < \infty$.

Let $C \subset S$ be such that $B \not\subset C$. Then

$$bel_m(C) = \sum\nolimits_{D \subset C} m(D) = \sum\nolimits_{D \subset C} m_1(D) = bel_{m_1}(C), \qquad (11.3)$$

as $B \not\subset C$ implies that $m(D) = m_1(D)$ for all $D \subset C$. Let $C \subset S$ be such that $B \subset C$. Then also $A \subset C$ holds and

$$\begin{aligned} bel_m(C) &= \sum\nolimits_{D \subset C,\, D \neq A} m(D) + m(A) = \infty = \\ &= \sum\nolimits_{D \subset C,\, D \neq A} m_1(D) + m_1(A) = bel_{m_1}(C), \end{aligned} \qquad (11.4)$$

as $m(A) = m_1(A) = \infty$. So, $bel_m \equiv bel_{m_1}$ and the lemma is proved. □

Definition 11.3. Two random variables U_1, $U_2 : \langle \Omega, \mathcal{A}, \mu \rangle \to \langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$, defined on a measurable space $\langle \Omega, \mathcal{A} \rangle$ with a signed measure μ, taking as their values subsets of S and such that the values $\mu(\{\omega \in \Omega : U_i(\omega) = A\}) = m_i(A)$ are finite for all $A \subset S$ and for both $i = 1, 2$, are called *statistically (stochastically) independent*, if the equality

$$\mu(\{\omega \in \Omega : U_1(\omega) = A,\, U_2(\omega) = B\}) = m_1(A)\, m_2(B) \tag{11.5}$$

holds true for all A, $B \subset S$. The generalization to the case of a finite sequence $U_1, U_2, \ldots, U_n$ of random variables is straightforward. □

Not so straightforward, however, is a generalization of this definition to the case when the b.s.m.a. m_i generated by U_1 or U_2 (or both) on $\mathcal{P}(S)$ can take also infinite values, as in such a case we have to adopt some conventions concerning expressions like $0 \cdot \infty$, $0 \cdot (-\infty)$, $\infty \cdot 0$, $(-\infty) \cdot 0$, $\infty \cdot \infty$, $\infty \cdot (-\infty)$, etc. The notion of statistical independence will then substantially depend on the convention adopted. E. g., random events $\{\omega \in \Omega : U_1(\omega) = A\}$ and $\{\omega \in \Omega : U_2(\omega) = B\}$, A, $B \subset S$, such that

$$\begin{aligned} &\mu(\{\omega \in \Omega : U_1(\omega) = A\}) = \mu(\{\omega \in \Omega : U_1(\omega) = A,\, U_2(\omega) = B\}) = 0, \\ &\mu(\{\omega \in \Omega : U_2(\omega) = B\}) = \infty, \end{aligned} \tag{11.6}$$

are statistically independent, if $0 \cdot \infty = 0$, but they are not statistically independent, if $0 \cdot \infty = 1$. In order to avoid, in the most possible degree, the influence of certain arbitrariness connected with the conventions of this kind, we shall investigate, at least now, just the case of such random variables which generate finite b.s.m.a.'s on S.

Let U_1, U_2 be two stochastically independent set-valued random variables defined on a measurable space $\langle \Omega, \mathcal{A} \rangle$ with a signed measure μ, taking as their values subsets of a finite set S, and such that both the b.s.m.a.'s m_1, m_2 generated by U_1, U_2 are finite. Let $A \subset S$. Then

$$\begin{aligned} &\mu(\{\omega \in \Omega : U_1(\omega) \cap U_2(\omega) = A\}) = \\ = &\sum\nolimits_{\langle B,C\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), B\cap C = A} \mu(\{\omega \in \Omega : U_1(\omega) = B,\, U_2(\omega) = C\}) = \\ = &\sum\nolimits_{\langle B,C\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), B\cap C = A} \mu(\{\omega \in \Omega : U_1(\omega) = B\}) \cdot \mu(\{\omega \in \Omega : U_2(\omega) = C\}) = \\ = &\sum\nolimits_{\langle B,C\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), B\cap C = A} m_1(B)\, m_2(C). \end{aligned} \tag{11.7}$$

So, denoting by m_3 the b.s.m.a. generated by the random variable $U_1 \cap U_2$, i. e., $m_3(A) = \mu(\{\omega \in \Omega : U_1(\omega) \cap U_2(\omega) = A\})$, the relation between the b.s.m.a. m_3 and the pair $\langle m_1, m_2 \rangle$ of b.s.m.a.'s is close to that between basic probability assignments m_1, m_2, and $m_1 \oplus m_2$, for the non-normalized Dempster combination rule. This analogy motivates the following definition.

Definition 11.4. Let m_1, m_2 be finite basic signed measure assignments over a finite set S. Let m_3 be the b.s.m.a. on S defined, for each $A \subset S$, by the relation

$$m_3(A) = \sum\nolimits_{\langle B,C\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), B\cap C = A} m_1(B)\, m_2(C). \tag{11.8}$$

Then m_3 is called the *Dempster product* of the b.s.m.a.'s m_1 and m_2 and denoted by $m_1 \boxplus m_2$. The operation $\boxplus$ which transforms pairs of b.s.m.a.'s into a new b.s.m.a. is called the *Dempster combination rule* for finite b.s.m.a.'s. □

In the same way as in the case of b.p.a.'s we can prove that also the operation $\boxplus$ over the pairs of finite b.s.m.a.'s is commutative and associative, i.e., $m_1 \boxplus m_2 \equiv m_2 \boxplus m_1$, and $m_1 \boxplus (m_2 \boxplus m_3) \equiv (m_1 \boxplus m_2) \boxplus m_3$ holds for all finite b.s.m.a.'s m_1, m_2, m_3 with $\equiv$ denoting, as above, the equality of the corresponding values for all subsets of S. Given $A \subset S$ and $a \in (-\infty, \infty)$, denote by $m_{A,a}$ the b.s.m.a. for which $m_{A,a}(A) = a$ and $m_{A,a}(B) = 0$ for each $B \subset S$, $B \neq A$. In particular, we write $\mathbf{1}_S$ for $m_{S,1}$ and $\mathbf{0}_S$ for $m_{S,0}$ ($\equiv m_{A,0}$ for all $A \subset S$), hence $\mathbf{1}_S(S) = 1$, $\mathbf{1}_S(A) = 0$ for all $A \subset S$, $A \neq S$, and $\mathbf{0}_S = 0$ for all $A \subset S$. The index S in $\mathbf{1}_S$ and $\mathbf{0}_S$ will be omitted supposing that S is fixed and no misunderstanding menaces.

Lemma 11.3. The b.s.m.a. $\mathbf{0}_S$ is a zero element and the b.s.m.a. $\mathbf{1}_S$ is a unit element in the space of all finite b.s.m.a.'s over a finite set S and with respect to the Dempster combination rule $\boxplus$. □

Proof. An easy calculation yields that for each finite b.s.m.a. m over a finite S, and for each $A \subset S$,

$$\begin{aligned} (m \boxplus \mathbf{1})(A) &= (\mathbf{1} \boxplus m)(A) = \textstyle\sum_{\langle B,C\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), B\cap C = A} \mathbf{1}(B)\, m(C) \\ &= \textstyle\sum_{C \subset S,\, S\cap C = A} \mathbf{1}(S)\, m(C) = m(A), \end{aligned} \tag{11.9}$$

and

$$(m \boxplus \mathbf{0})(A) = (\mathbf{0} \boxplus m)(A) = \textstyle\sum_{\langle B,C\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), B\cap C = A} \mathbf{0}(B)\, m(C) = 0, \tag{11.10}$$

as $\mathbf{0}(B) = 0$ for all $B \subset S$. □

As can be easily proved, $\mathbf{0}_S$ and $\mathbf{1}_S$ are the only zero and unit elements with respect to the Dempster combination rule $\boxplus$.

It is perhaps worth recalling explicitly that the mapping $q_m : \mathcal{P}(S) \to (-\infty, \infty)$ defined, for each b.s.m.a. m on S and each $A \subset S$, by $q_m(A) = \sum_{B \supset A} m(B)$, is nothing else than a straightforward generalization of the so called *commonality function* defined in this way when m is a b.p.a. on S. This function possesses a number of properties dual, in an intuitive sense, to those of belief functions. Moreover, some formulas describing the basic properties of the Dempster–Shafer theory take a more simple syntactical pattern when using commonality degrees (values of commonality functions) as the basic numerical degrees of uncertainty instead of the degrees of belief. On the other side, at least in the author's subjective opinion, the possible intuitions and interpretations behind the degrees of belief ("probability of provability", say) seem to be more transparent and lucid, so that the use of the degrees of belief as the basic stones in our constructions and reasonings is perhaps at least partially justified.

Definition 11.5. A finite signed measure assignment m over a finite set S is called *invertible*, if the inequality $\sum_{B \supset A} m(B) \neq 0$ holds for each $A \subset S$. □

The following statement illustrates the motivation for the adjective "invertible" just introduced. Let us postpone a discussion on this notion till an appropriate place below.

Theorem 11.1. Let m be an invertible finite b.s.m.a. over a finite set S. Let m^{-1} be the b.s.m.a. over S defined recurrently in this way:

$$m^{-1}(A) = (m(S))^{-1}, \tag{11.11}$$

$$m^{-1}(A) = \frac{\sum_{B \subset S, C \subset S, B \cap C = A, B \neq A} m^{-1}(B)\, m(C)}{\sum_{B \subset S, B \supset A} m(B)}, \tag{11.12}$$

if $A \subset S,\ A \neq S$. Then $m \boxed{+} m^{-1} \equiv \mathbf{1}$. □

Remark. An open question arises, whether there is some relation between the transformation just defined and the so called Möbius transformation, but we shall not analyze this problem in more detail now.

Proof. First of all, we have to prove that the definitions (11.11) and (11.12) are correct. Applying the condition $\sum_{B \supset A} m(B) \neq 0$ to the case $A = S$, we obtain that $m(S) \neq 0$ for invertible b.s.m.a.s. Hence, $(m(S))^{-1}$ is defined. The summation on the right-hand side of (11.12) goes over the sets B such that $B \cap C$=A and $B \neq A$, hence, over the sets $B \subset S$ such that $B \supset A$ and $\text{card}(B) > \text{card}(A)$. Consequently, $m^{-1}(A)$ is uniquely defined in the recurrent way according to the decreasing cardinality of A. I.e., first of all $m^{-1}(S)$ is defined by (11.11), in the definition of $m^{-1}(A)$ for A such that $\text{card}(A) = \text{card}(S) - 1$ only $m^{-1}(S)$ occurs, so that the definition is correct, and so on till

$$m^{-1}(\emptyset) = \frac{\sum_{B \subset S, C \subset S, B \cap C = \emptyset, B \neq \emptyset} m^{-1}(B)\, m(C)}{\sum_{B \subset S} m(B)}. \tag{11.13}$$

For the whole set S we obtain

$$(m \boxed{+} m^{-1})(S) = \sum_{B \subset S, C \subset S, B \cap C = S} m(B)\, m^{-1}(C) = m(S)\, m^{-1}(S) = 1. \tag{11.14}$$

Let $A \subset S,\ A \neq S$. Then

$$(m \boxed{+} m^{-1})(A) = \sum_{B \subset S, C \subset S, B \cap C = A} m(B)\, m^{-1}(C) \tag{11.15}$$

by definition. The set $\{\langle B, C \rangle \in \mathcal{P}(S) \times \mathcal{P}(S),\ B \cap C = A\}$ can be decomposed into four disjoint subsets:

$$\begin{aligned} &\{\langle A, A \rangle\}, \\ &\{\langle B, A \rangle : B \subset S,\ B \supset A,\ \text{so that } B \cap A = A,\ B \neq A\}, \\ &\{\langle A, B \rangle : B \subset S,\ B \supset A,\ \text{so that } B \cap A = A,\ B \neq A\}, \\ &\{\langle B, C \rangle : B \subset S,\ C \subset S,\ B \neq A,\ C \neq A,\ B \cap C = A\}. \end{aligned} \tag{11.16}$$

So, omitting the expression "$\ldots \in \mathcal{P}(S) \times \mathcal{P}(S)$" for the sake of simplicity we obtain that

$$\begin{aligned}(m \boxplus m^{-1})(A) &= \textstyle\sum_{B,C,\,B=A,\,C=A} m(B)\,m^{-1}(C) + \\ &\quad + \textstyle\sum_{B,C,\,B\supset A,\,B\neq A,\,C=A} m(B)\,m^{-1}(C) + \\ &\quad + \textstyle\sum_{B,C,\,B\supset A,\,C=A} m(B)\,m^{-1}(C) + \\ &\quad + \textstyle\sum_{B,C,\,B\neq A,\,C\neq A,\,B\cap C=A} m(B)\,m^{-1}(C) = \\ &= m(A)\,m^{-1}(A) + \textstyle\sum_{B\supset A,\,B\neq A} m(B)\,m^{-1}(A) + \\ &\quad + \textstyle\sum_{B,\,B\supset A,\,B\neq A} m(A)\,m^{-1}(B) + \\ &\quad + \textstyle\sum_{B,C,\,B\neq A,\,C\neq A,\,B\cap C=A} m(B)\,m^{-1}(C) = \\ &= m^{-1}(A)\left[m(A) + \textstyle\sum_{B,\,B\supset A,\,B\neq A} m(B)\right] + \\ &\quad + \textstyle\sum_{B,C,\,B\supset A,\,B\neq A,\,C\neq A} m(C)\,m^{-1}(B) + \\ &\quad + \textstyle\sum_{B,C,\,B\neq A,\,C\neq A,\,B\cap C=A} m^{-1}(B)\,m(C),\end{aligned} \tag{11.17}$$

as the last sum contains, with each product $m(B)\,m^{-1}(C)$, also the product $m(C)\,m^{-1}(B)$. Or, if B, C are such that $B \neq A$, $C \neq A$, and $B \cap C = A$, the same holds for the pair $\langle C, B\rangle$. Consequently,

$$\begin{aligned}(m \boxplus m^{-1})(A) &= m^{-1}(A)\left[\textstyle\sum_{B,\,B\supset A} m(B)\right] + \\ &\quad + \textstyle\sum_{B,C,\,B\supset A,\,B\neq A,\,C\supset A,\,B\cap C=A} m^{-1}(B)\,m(C).\end{aligned} \tag{11.18}$$

Combining (11.18) and (11.12) we obtain that $(m \boxplus m^{-1})(A) = 0$ for each $A \subset S,\ A \neq S$, so that $m \boxplus m^{-1} \equiv \mathbf{1}$ holds. The theorem is proved. $\square$

Let S be a finite set. As basic probability assignments (b.p.a.'s) over S are particular cases of finite b.s.m.a.'s, Theorem 11.1 holds for b.p.a.'s as well. As can be easily seen, a b.p.a. m is invertible iff $m(S)$ is positive. Or, if $m(S) > 0$, then $\sum_{B\supset A} m(B) \geq m(S) > 0$ holds for each $A \subset S$, if $m(S) = 0$, then $\sum_{B\supset S} m(B) = m(S) = 0$ and m is not invertible. It follows immediately that the A-conditioning b.p.a. m_A defined, for $\emptyset \neq A \subset S,\ A \neq S$, by $m_A(A) = 1$, so that $m_A(B) = 0$ for each $B \subset S,\ B \neq A$, cannot be inverted. Let us consider an (ε, A)-conditioning b.p.a. $m_{\varepsilon,A}$, where $\varepsilon \in (0,1)$ is a real number, defined in this way ($\emptyset \neq A \subset S,\ A \neq S$):

$$m_{\varepsilon,A}(A) = 1 - \varepsilon, \quad m_{\varepsilon,A}(S) = \varepsilon, \quad m_{\varepsilon,A}(B) = 0, \quad B \subset S,\ B \neq A,\ B \neq S. \tag{11.19}$$

Hence, the result of (ε, A)-conditioning applied to a b.p.a. m is defined by the Dempster product $m \boxplus m_{\varepsilon,A}$ ($= m \oplus m_{\varepsilon,A}$ in this particular case).

Lemma 11.4. Let $m_{\varepsilon,A}$, $\varepsilon \in (0,1)$, $\emptyset \neq A \neq S$, $A \subset S$, be defined by (11.19). Then

$$\begin{aligned} m^{-1}_{\varepsilon,A}(S) &= 1/\varepsilon, \quad m^{-1}_{\varepsilon,A}(A) = \\ &= -(1-\varepsilon)/\varepsilon, \quad m^{-1}_{\varepsilon,A}(B) = 0, \quad B \subset S,\ B \neq A,\ B \neq S. \end{aligned} \tag{11.20}$$

Proof. Let us omit the indices ε and A throughout this proof. An easy calculation yields that

$$(m \boxplus m^{-1})(S) = m_{\varepsilon,A}(S)\, m_{\varepsilon,A}^{-1}(S) = 1, \tag{11.21}$$

$$\begin{aligned}(m \boxplus m^{-1})(A) &= \textstyle\sum_{B,C\subset S,\, B\cap C=A} m(B)\, m^{-1}(C) = \\ &= \textstyle\sum_{B,C\subset S,\, B\cap C=A,\, m(B)\neq 0,\, m^{-1}(C)\neq 0} m(B)\, m^{-1}(C) = \\ &= \textstyle\sum_{\langle B,C\rangle\in\{\langle A,A\rangle,\langle S,A\rangle,\langle A,S\rangle\}} m(B)\, m^{-1}(C) = \\ &= m(A)\, m^{-1}(A) + m(S)\, m^{-1}(A) + m(A)\, m^{-1}(S) = \\ &= -\varepsilon^{-1}(1-\varepsilon)^2 - \varepsilon^{-1}(\varepsilon(1-\varepsilon)) + \varepsilon^{-1}(1-\varepsilon) = \\ &= \varepsilon^{-1}(-1+2\varepsilon-\varepsilon^2-\varepsilon+\varepsilon^2+1-\varepsilon) = 0.\end{aligned} \tag{11.22}$$

Let $D \subset S$, $D \neq A$, $D \neq S$. Then $A\cap A \neq D$, $A \cap S \neq D$, $S \cap A \neq D$, $S \cap S \neq D$, but for each $\langle B, C\rangle \in \mathcal{P}(S)\times\mathcal{P}(S)$ such that $\langle B,C\rangle \notin \{\langle A,A\rangle, \langle A,S\rangle, \langle S,A\rangle, \langle S,S\rangle\}$, either $m(B) = 0$ or $m^{-1}(C) = 0$, so that

$$(m \boxplus m^{-1})(D) = \textstyle\sum_{B,C\subset S,\, B\cap C=D} m(B)\, m^{-1}(C) = 0. \tag{11.23}$$

Consequently, $(m \boxplus m^{-1})(B) = 0$ for all $B \subset S$, $B \neq S$, so that $m_{\varepsilon,A} \boxplus m_{\varepsilon,A}^{-1} \equiv \mathbf{1}$. □

Definition 11.6. A nonempty set $\mathcal{R}$ of basic signed measure assignments over a finite set S is called *coherent*, if there are no $m_1, m_2 \in \mathcal{R}$ and no $A \subset S$ such that $m_1(A) = \infty$ and $m_2(A) = -\infty$. □

Lemma 11.5. Let $\mathcal{R}$ be a coherent set of b.s.m.a.'s over a finite set S, let $\rho : \mathcal{R}\times\mathcal{R} \to R^* = (-\infty,\infty)\cup\{\infty\}\cup\{-\infty\}$ be defined by

$$\rho(m_1, m_2) = \max\{|m_1(A) - m_2(A)| : A \subset S\}, \quad m_1,\, m_2 \in \mathcal{R}, \tag{11.24}$$

where $\infty - a = \infty$, $-\infty - a = -\infty$, $a - (-\infty) = \infty$, $a - \infty = -\infty$, $\infty - \infty = (-\infty) - (-\infty) = 0$ for each $a \in (-\infty,\infty)$. Then ρ is a metric on $\mathcal{R}$. □

Proof. Obviously $\rho(m_1, m_1) = 0$ and $\rho(m_1,m_2) = \rho(m_2,m_1)$ for all $m_1,\, m_2 \in \mathcal{R}$. Let $m_1,\, m_2,\, m_3 \in \mathcal{R}$. Then

$$\begin{aligned}&\rho(m_1,m_3) = \max\{|m_1(A) - m_3(A)| : A \subset S\} = \\ =\ &\max\{|m_1(A) - m_2(A) + m_2(A) - m_3(A)| : A \subset S\} \leq \\ \leq\ &\max\{|m_1(A) - m_2(A)| + |m_2(A) - m_3(A)| : A \subset S\} \leq \\ \leq\ &\max\{|m_1(A) - m_2(A)| : A \subset S\} + \max\{|m_2(A) - m_3(A)| : A \subset S\} = \\ =\ &\rho(m_1,m_2) + \rho(m_2,m_3),\end{aligned} \tag{11.25}$$

all the inequalities in (11.25) being evidently valid also when some of the values $m_i(A)$, $i = 1,2,3$, are infinite. Hence, the triangular inequality for ρ and the lemma as a whole are proved. □

Every b.s.m.a. over a finite set S is almost invertible in the sense defined in the following assertion.

Theorem 11.2. For each b.s.m.a. m over a finite set S and for each $\varepsilon > 0$ there exists a b.s.m.a. m_0 over S such that $\rho(m, m_0) < \varepsilon$ holds and m_0^{-1} is defined. □

Proof. Let m be a b.s.m.a. over a finite set S, let $\varepsilon > 0$ be given, let $0 < \varepsilon_1 < \varepsilon$. If the inequality $\sum_{B \supset A} m(B) \neq 0$ holds for each $A \subset S$, then m^{-1} is defined, hence, for $m = m_0$ the assertion trivially holds, as $\rho(m, m) = 0 < \varepsilon$. If this is not the case, set

$$i_1 = \max\{n : \exists\, A \subset S,\ \text{card}A = n, \textstyle\sum_{B \supset A} m(B) = 0\}. \tag{11.26}$$

Set also, for each $A \subset S$,

$$\begin{aligned} m_1(A) &= m(A) + \varepsilon_1, \text{ if card}A = i, \text{ and } \textstyle\sum_{B \supset A} m(B) = 0, \\ m_1(A) &= m(A) \text{ otherwise.} \end{aligned} \tag{11.27}$$

In particular, (11.27) yields that $m_1(A) = m(A)$ for all $A \subset S$ such that $\text{card}A > i_1$ or $\text{card}A < i_1$ holds.

Let $A \subset S$ be such that $\text{card}A > i_1$. Then $\sum_{B \supset A} m_1(B) = \sum_{B \supset A} m(B) \neq 0$ holds due to the definition of i_1, as for each $B \supset A$ the relations $\text{card}B > i_1$ and $m_1(B) = m(B)$ follow. Let $A \subset S$ be such that $\text{card}A = i_1$ and $\sum_{B \supset A} m(B) \neq 0$. Then, again, $m_1(B) = m(B)$ for all $B \supset A$, $B \neq A$, but also $m_1(A) = m(A)$ due to (11.27), so that $\sum_{B \supset A} m_1(B) = \sum_{B \supset A} m(B) \neq 0$. Finally, let $A \subset S$ be such that $\text{card}A = i_1$ and $\sum_{B \supset A} m(B) = 0$. Then $m_1(B) = m(B)$ for each $B \supset A$, $B \neq A$, but $m_1(A) = m(A) + \varepsilon_1$, so that

$$\begin{aligned} \textstyle\sum_{B \supset A} m_1(B) &= \textstyle\sum_{B \supset A, B \neq A} m_1(B) + m_1(A) = \\ &= \textstyle\sum_{B \supset A} m(B) + \varepsilon_1 \neq 0. \end{aligned} \tag{11.28}$$

Consequently, for each $A \subset S$, $\text{card}A \geq i_1$, the inequality $\sum_{B \supset A} m_1(B) \neq 0$ holds.

By induction, let us apply the same modification to m_1. Set

$$i_2 = \max\{n : \exists\, A \subset S,\ \text{card}A = n, \textstyle\sum_{B \supset A} m_1(B) = 0\}, \tag{11.29}$$

and define m_2 by (11.27), just with m replaced by m_1. Evidently, $i_2 < i_1$ and $\sum_{B \supset A} m_2(B) \neq 0$ for all $A \subset S$, $\text{card}A \geq i_2$, by the same way of reasoning as above. Moreover, $m_2(B) = m_1(B)$ for all $B \subset S$ such that $\text{card}B > i_2$, so that $m_2(B) = m(B)$ for all $B \subset S$ with $\text{card}B > i_1$. Hence, repeating this induction step n_0-times for an appropriate $n_0 \leq \text{card}S$, we arrive at a b.s.m.a. m_{n_0} such that $\sum_{B \supset A} m_{n_0}(B) \neq 0$ holds for each $A \subset S$, consequently, $m_{n_0}^{-1}$ is defined. For each particular $A \subset S$ the original value $m(A)$ is changed at most once during the procedure leading from m to m_{n_0}, so that either $m_{n_0}(A) = m(A)$, or $m_{n_0}(A) = m(A) + \varepsilon_1$ for each $A \subset S$. Hence, $\rho(m, m_{n_0}) = \max\{|m(A) - m_{n_0}(A)| : A \subset S\} \leq \varepsilon_1 < \varepsilon$ obviously follows, so that, setting $m_0 = m_{n_0}$, we can conclude the proof. □

Corollary 11.1. Let $\mathcal{R}$ be a coherent set of b.s.m.a.'s over a finite set S, let $\rho^1 : \mathcal{R} \times \mathcal{R} \to R \cup \{\infty\} \cup \{-\infty\}$ be defined, for each $m_1, m_2 \in \mathcal{R}$, by

$$\rho^1(m_1, m_2) = \textstyle\sum_{B \subset S} |m_1(B) - m_2(B)|. \tag{11.30}$$

Then Theorem 11.2 holds for ρ^1 instead of ρ. □

Proof. Given $\varepsilon > 0$, take $\varepsilon_1 = \varepsilon/(2^{\mathrm{card} S}+1)$ and apply the same inductive process of modification as in the proof of Theorem 11.2 in order to obtain the b.s.m.a. m_0 such that $\sum_{B \supset A} m_0(B) \neq 0$ holds for each $A \subset S$. Again, for each $B \subset S$ either $m_0(B) = m(B)$ or $m_0(B) = m(B) + \varepsilon_1$ holds, so that $|m_0(B) - m(B)| \leq \varepsilon_1$. Consequently,

$$\begin{aligned} \rho^1(m, m_0) &= \textstyle\sum_{B \subset S} |m(B) - m_0(B)| \leq \varepsilon_1 \mathrm{card} \mathcal{P}(S) = \\ &= \left[\varepsilon/(2^{\mathrm{card} S} + 1)\right] 2^{\mathrm{card} S} < \varepsilon, \end{aligned} \tag{11.31}$$

so that the corollary is proved. □

Let us generalize the approach explained in this chapter using also signed measures with "almost zero" and "almost infinite" values. We refer the reader to [32] for more detail, but we consider as useful to present a brief sketch of this generalization also here, so that the connections can be easily seen.

We shall generalize the approximative solution to the invertibility problem for b.p.a.'s illustrated above by the ε-quasiconditioning approach. The weak point of this approximation consists in the fact that it introduces a new and ontologically independent parameter ε, the actual value of which cannot be justified only within the framework of the used mathematical formalism. Consequently, the subject (user) must choose some value on the grounds of her/his subjective opinion taking into consideration, e. g., the intended field of application and other extra-mathematical circumstances. Below, we shall present a model which enables to invert also b.p.a.'s ascribing to the whole space a value "greater than 0 but smaller than any positive ε", in other words, a "quasi-zero value", both these notions being given a correct mathematical sense. The corresponding inverse "generalized b.p.a." then will take "quasi-infinite values" smaller than ∞ but greater than any finite real number.

Let $\mathcal{R} = \mathsf{X}_{i=1}^{\infty} R_i$, $R_i = (-\infty, \infty)$ for each $i \in \mathcal{N}^+ = \{1, 2, \ldots\}$ be the space of all infinite sequences of real numbers. For each $x \in \mathcal{R}$, x_i denotes the i-th member of x, so that $x = \langle x_i \rangle_{i=1}^{\infty}$. Given $x \in \mathcal{R}$, set $w(x) = \lim_{i \to \infty} x_i$ supposing that this limit value is defined and including the case when $w(x) = \pm\infty$. Let $\mathcal{R}_c = \{x \in \mathcal{R} : w(x) \text{ is defined}\}$ be the space of all convergent infinite sequences of real numbers, let $\mathcal{R}_{cf} = \{x \in \mathcal{R}_c : -\infty < w(x) < \infty\}$ be the subspace of convergent infinite sequences with finite limit values. Let us define three following binary relations in $\mathcal{R}$:

(i) *identity*: given $x, y \in \mathcal{R}$, $x = y$ iff $x_i = y_i$ for each $i \in \mathcal{N}^+$;

(ii) *strong equivalence* (s.e.): given $x, y \in R$, $x \approx y$ if there exists $i_0 \in \mathcal{N}^+$ such that $x_i = y_i$ for each $i \geq i_0$;

(iii) *weak equivalence* (w.e.): given $x, y \in R_c$, $x \sim y$ iff $w(x) = w(y)$.

Arithmetical operations in $\mathcal{R}$ will be defined in the pointwise way, so that $x + y = \langle x_i + y_i \rangle_{i=1}^{\infty}$ and $xy = \langle x_i y_i \rangle_{i=1}^{\infty}$ for each $x, y \in \mathcal{R}$. It follows easily that $\sum_{j=1}^{n} x^j = \left\langle \sum_{j=1}^{n} x_i^j \right\rangle_{i=1}^{\infty}$ and $\prod_{j=1}^{n} x^j = \left\langle \prod_{j=1}^{n} x_i^j \right\rangle_{i=1}^{\infty}$ holds for each finite sequence $x^1, x^2, \ldots, x^n$ of sequences from $\mathcal{R}$. Both the addition and multiplication operations can be easily extended to equivalence classes from $\mathcal{R}_{\approx}$ or $\mathcal{R}_{cf\sim}$. Let $[x]_{\approx} = \{y \in \mathcal{R} : y \approx x\}$, let $[x]_{\sim} = \{y \in \mathcal{R}_{cf} : y \sim x\}$. Setting $[x]_{\approx} + [y]_{\approx} = [x + y]_{\approx}$ for all $x, y \in \mathcal{R}$, and $[x]_{\sim} + [y]_{\sim} = [x + y]_{\sim}$ for all $x, y \in \mathcal{R}_{cf}$, we obtain a correct extension of both the arithmetical operations to $\mathcal{R}_\approx$ and $\mathcal{R}_{cf\sim}$, as the definitions do not depend on the chosen respresentants of the classes $[x]_\approx$, $[y]_\approx$, $[x]_\sim$, and $[y]_\sim$. For $[x]_\approx \cdot [y]_\approx$ defined by $[xy]_\approx$, if $x,\, y \in \mathcal{R}$, and for $[x]_\sim \cdot [y]_\sim$ defined by $[xy]_\sim$, if $x,\, y \in \mathcal{R}_{cf}$, as well as for finite products in general, the situation is analogous. However, if $w(x) = \pm\infty$ or $w(y) = \pm\infty$, the definitions of $[x]_\sim + [y]_\sim$ and $[x]_\sim \cdot [y]_\sim$ are evidently not correct. Also the extension of both the operations to infinite sums and products is impossible, take $x^j = \langle x_i^j \rangle_{i=1}^\infty$ such that $x_i^i = 1$, $x_i^j = 0$ for each $i \neq j$. Then $w(x^j) = 0$ for each $j \in \mathcal{N}^+$, but $w\left(\sum_{j=1}^\infty x^j\right) = 1$. We use intentionally the same symbols $+$ and $\cdot$ for operations in R, $\mathcal{R}$, $\mathcal{R}_\approx$ and $\mathcal{R}_{cf\sim}$, to emphasize their analogous role in all the cases. It should be always clear from the context, in which space these operations work.

Also the ordering relation can be extended, even if only as a partial ordering, from $R = (-\infty, \infty)$ to other spaces under consideration. If $x,\, y \in \mathcal{R}$, we write $x > y$, if $x_i > y_i$ holds for each $i \in \mathcal{N}^+$, and we write $x >. y$, if there exists $i_0 \in \mathcal{N}^+$ such that $x_i > y_i$ holds for each $i \geq i_0$, if $x,\, y \in \mathcal{R}_c$, we write $x \succ y$, if $w(x) > w(y)$ holds. The inequality $\geq$ and $\geq.$ on $\mathcal{R}$, and $\succeq$ on $\mathcal{R}_c$, are defined analogously, just with $x_i > y_i$ replaced by $x_i \geq y_i$, and $w(x) > w(y)$, by $w(x) \geq w(y)$, also the inequalities $<$, $\leq$, $<.$, $\prec$ and $\preceq$ are obvious. The relations $>.$ and $\geq.$ can be extended to $\mathcal{R}_\approx$, and $\succ$, $\succeq$ to $\mathcal{R}_{cf\sim}$, setting $[x]_\approx >. [y]_\approx$, if $x >. y$, and setting $[x]_\sim \succ [y]_\sim$, if $x \succ y$. Both these definitions are correct, i. e., independent of the representants of the equivalence classes in question. Inequalities $\leq$ and $\preceq$ are extended to $\mathcal{R}_\approx$ and $\mathcal{R}_{cf\sim}$ in the same way.

Let $a \in R$ be a real number, let $a^* = \langle a_i \rangle_{i=1}^\infty$, $a_i = a$ for each $i \in \mathcal{N}^+$ be the corresponding constant sequence from R_{cf}. The real line can be embedded into $\mathcal{R}_\approx$, when identifying each $a \in R$ with the equivalence class $[a^*]_\approx$, and R can be embedded into $\mathcal{R}_{cf\sim}$, when identifying a with $[a^*]_\sim$. However, there is an important difference. The second mapping takes R *onto* $\mathcal{R}_{cf\sim}$, as for each $x \in \mathcal{R}_{cf}$ there is a real number a, namely $a = w(x)$, which is mapped on $[x] \in \mathcal{R}_{cf\sim}$. Contrary to this fact, there are many $x \in \mathcal{R}$ such that no $a \in R$ is mapped onto $[x]$, take, e. g., $x = \langle n \rangle_{n=1}^\infty$, or $y = \langle 1/n \rangle_{n=1}^\infty$. Denoting the classes from $\mathcal{R}_\approx$ which are images of real numbers from R, by *standard real numbers*, $[x]_\approx$ and $[y]_\approx$ just defined are examples of *nonstandard real numbers*. In particular, $[x]_\approx$ is an example of *quasi-infinite nonstandard real number*, as $[x]_\approx >. [a^*]$ obviously holds for each $a \in R$, and $[y]_\approx$ can serve as an example of *quasi-zero nonstandard real number*, as $[y]_\approx >. [0^*]_\approx$, but also $[a^*]_\approx >. [y]_\approx$ for each $a \in R$, $a > 0$, can be easily verified.

Definition 11.7. Let S be a finite nonempty set. *Basic dynamic assignment* (b.d.a.) on S is a mapping $m : \mathcal{R}(S) \to \mathcal{R} = \mathsf{X}_{i=1}^\infty R_i$, $R_i = (-\infty, \infty)$, $i = 1, 2, \ldots$. If $m(\emptyset) = 0^*$ $(= \langle 0, 0, \ldots \rangle)$, then *dynamic belief function* induced by b.d.a. m on S is a mapping $bel_n : \mathcal{P}(S) \to \mathcal{R}$ such that $bel_m(A) = \sum_{B \subset A} m(B)$ for each $A \subset S$. If m is a finite b.s.m.a. on S, then m^* is the b.d.a. on S defined by $m^*(A) = (m(A))^*$ for each $A \subset S$, if m is a b.d.a. on S, then $m^* \equiv m$, i. e. $m^*(A) = m(A)$ for each $A \subset S$. Let us recall that $a^* = \langle a, a, \ldots \rangle$ for each $a \in R = (-\infty, \infty)$. Let m_1, m_2 be b.p.a.'s, b.s.m.a.'s or b.d.a.'s on S. Then m_1 is *strongly equivalent to* m_2, $m_1 \approx m_2$ is symbols, if $m_1^*(A) \approx m_2^*(A)$ holds for each $A \subset S$, m_1 is *weakly equivalent to* m_2, $m_1 \sim m_2$ in symbols, if $m_1^*(A) \sim m_2^*(A)$ holds for each $A \subset S$. □

Definition 11.8. Let m_1, m_2 be b.d.a.'s on a finite set S. Their *Dempster product* $m_1 \boxplus m_2$ is the b.d.a. on S defined by the relation

$$(m_1 \boxplus m_2)(A) = \sum_{B,C \subset S,\, B \cap C = A} m_1(B)\, m_2(C) \tag{11.32}$$

for each $A \subset S$. If m is a b.d.a. on S, then m_i denotes the finite b.s.m.a. on S defined by $(m_i)(A) = (m(A))_i$ for each $i \in \mathcal{N}^+$ and $A \subset S$. Let us recall that x_i is the i-th member of the sequence $x = \langle x_i \rangle_{i=1}^{\infty} \in \mathcal{R}$. *Vacuous* b.d.a. on S is such m_V that $m_V(S) = 1^*$ and $m_V(A) = 0^*$ for each $A \subset S$, $A \neq S$. Consequently, $(m_V)_i$ is the vacuous b.s.m.a. (and the vacuous b.p.a.) on S for each $i \in \mathcal{N}^+$. The *dynamic commonality function* induced by a b.d.a. m on a finite set S is the mapping $q_m : \mathcal{P}(S) \rightarrow \mathcal{R}$ defined by $q_m(A) = \sum_{B \supset A} m(B)$ for each $A \subset S$. □

Theorem 11.3. Let m be a convergent b.d.a. on a finite set S, i.e., let $m(A) \in \mathcal{R}_c$ for each $A \subset S$. Then there exist b.d.a.'s m_1 and m_1^{-1} on S such that $m_1 \sim m$ and $m_1 \boxplus m_1^{-1} \equiv m_V$. In other words: for each convergent b.d.a. m there exists an invertible b.d.a. weakly equivalent to m. □

Proof. Let m be a convergent b.d.a. on a finite set S. For each $i \in \mathcal{N}^+$, let m_i^0 be such a b.s.m.a. on S, that $|m_i(A) - m_i^0(A)| < 1/i$ and $q_i(A) \neq 0$ holds for each $A \subset S$, such m_i^0 exists due to Theorem 11.2. Let m^0 be the b.d.a. on S such that $m^0(A) = \langle m_i^0(A) \rangle_{i=1}^{\infty}$ for each $A \subset S$. As $w(m(A)) = \lim_{i \rightarrow \infty} m_i(A)$ exists for each $A \subset S$, $w(m^0(A)) = \lim_{i \rightarrow \infty} m_i^0(A) = w(m(A))$ exists as well, so that $m \sim m_0$ holds. Let $(m_i^0)^{-1}$ be the b.s.m.a. on S defined by (11.11) and (11.12) above (cf. Theorem 11.1) for the b.s.m.a. m_i^0. Then $(m_i^0 \boxplus (m_i^0)^{-1}) \equiv m_S$ (the vacuous b.s.m.a. on S), so that, setting $(m^0)^{-1}(A) = \langle (m_i^0)^{-1}(A) \rangle_{i=1}^{\infty}$ we obtain, that $m^0 \boxplus (m^0)^{-1} \equiv m_V$, here m_V is the (obviously convergent) vacuous b.d.a. on S defined above. The assertion is proved. □

Let us illustrate this statement by the example of the single support b.p.a. m_A defined by $m_A(A) = 1$ for a given $A \subset S$, hence $m_A(B) = 0$ for each $B \subset S$, $B \neq A$, this b.p.a. is used, if $\emptyset \neq A \neq S$, in the conditioning operation. Here we can define $m_A^0 \sim m_A$ in such a way that $(m_A^0)_i(A) = 1 - (1/i)$, $(m_A^0)_i(S) = 1/i$, $(m_A^0)_i(B) = 0$ for each $B \subset S$, $B \neq S$, $B \neq A$. Let $(m_A^0)^{-1}$ be the b.d.a. defined by $((m_A^0)^{-1})_i(S) = i$, $((m_A^0)^{-1})_i(A) = -(i-1)$, $((m_A^0)^{-1})_i(B) = 0$ for all $B \subset S$, $B \neq S$, $B \neq A$ and for all $i \in \mathcal{N}^+$. An easy calculation then yields that $m_A \sim m_A^0$ and $m_A^0 \boxplus (m_A^0)^{-1} \equiv m_V$ hold.

12 Jordan Decomposition of Signed Belief Functions

In this chapter we shall try to arrive at some decompositions into generalized or even into classical probabilistic belief functions inspired by, and similar to, the Jordan decomposition of signed measure. Here *generalized belief function* is a particular case of signed belief function supposing that the signed measure μ used when defining the belief function in question takes only non-negative values (possibly including $+\infty$), so that μ is just what is called simply (σ-additive) measure in [18]. We shall also prove that generalizations of basic probability assignments, generated on $\mathcal{P}(S)$ with a finite S by signed belief functions, are also signed measures on the measurable space $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$. The following well-known theorem will play the key role of an inspiration, but also as a technical tool for all our reasonings and constructions throughout this chapter.

Theorem 12.1. (Hahn decomposition theorem.) Let $\langle \Omega, \mathcal{A} \rangle$ be a measurable space, let μ be a signed measure defined on $\langle \Omega, \mathcal{A} \rangle$. Then there exist disjoint subsets Ω^+, Ω^- of Ω such that $\Omega^+ \cup \Omega^- = \Omega$, $\Omega^+ \in \mathcal{A}$, hence, $\Omega^- \in \mathcal{A}$ as well, and both the mappings μ_1, $\mu_2 : \mathcal{A} \rightarrow (-\infty, \infty)$ defined by $\mu_1(A) = \mu(A \cap \Omega^+)$ and $\mu_2(A) = -\mu(A \cap \Omega^-)$ for each $A \in \mathcal{A}$, are measures (i. e., non-negative signed measures), so that $\mu_i(A) \in \langle 0, \infty \rangle$ for both $i = 1, 2$. So, for each signed measure μ defined on $\langle \Omega, \mathcal{A} \rangle$ and for each $A \in \mathcal{A}$ the value $\mu(A) = \mu_1(A) - \mu_2(A)$ is the difference of the values ascribed to A by two (non-negative) measures defined on $\langle\Omega,\mathcal{A}\rangle$. The pair $\langle\Omega^+,\Omega^-\rangle$ is called the *Hahn decomposition of* Ω and it need not be defined uniquely, the pair $\langle\mu_1,\mu_2\rangle$ is called the *Jordan decomposition of the signed measure* μ and it is defined uniquely. □

Proof. Cf. [18] for a more general case of signed measures defined on σ-rings of subsets of Ω. □

The following statement is an almost obvious application of the Hahn decomposition theorem to the case of basic signed measure assignments.

Theorem 12.2. Let $\langle\Omega,\mathcal{A}\rangle$ be a measurable space, let μ be a signed measure on $\langle\Omega,\mathcal{A}\rangle$, let S be a finite nonempty set, let U be a measurable mapping which takes the measurable space $\langle\Omega,\mathcal{A}\rangle$ into the measurable space $\langle\mathcal{P}(S),\mathcal{P}(\mathcal{P}(S))\rangle$. Set, for each $A\subset S$, $m(A)=\mu\ (\{\omega\in\Omega:U(\omega)=A\})$ (i.e., $\mu(\{\omega\in\Omega:U(\omega):U(\omega)\in\{A\}\in\mathcal{P}(\mathcal{P}(S))\}))$ and set, for each $\mathcal{B}\subset\mathcal{P}(S)$, $m^*(\mathcal{B})=\sum_{A\in\mathcal{B}}m(A)$ with the convention that $\mu^*(\emptyset)=0$ for the empty subset $\emptyset$ of $\mathcal{P}(\mathcal{P}(S))$. Then $m^*:\mathcal{P}(\mathcal{P}(S))\rightarrow(-\infty,\infty)$ is a signed measure on $\langle\mathcal{P}(S),\mathcal{P}(\mathcal{P}(S))\rangle$. Set $\mathcal{P}_1(S)=\{A\subset S:m(A)>0\}$, $\mathcal{P}_2(S)=\{A\subset S:m(A)<0\}$. Let $P^+(S)\subset\mathcal{P}(S)$, $p^-\subset\mathcal{P}(S)$ be such that $\mathcal{P}_1(S)\subset\mathcal{P}^+(S)$, $\mathcal{P}_2(S)\subset\mathcal{P}^-(S)$, $\mathcal{P}^+(S)\cap\mathcal{P}^-(S)=\emptyset$, $\mathcal{P}^+(S)\cup\mathcal{P}^-(S)=\mathcal{P}(S)$. Then $\langle\mathcal{P}^+(S),\mathcal{P}^-(S)\rangle$ is a Hahn decomposition of $\mathcal{P}(S)$ with respect to the signed measure m^*. □

Remark. Obviously, this Hahn decomposition is not uniquely defined supposing that $m(\emptyset)=0$, as in this case we can take either $\emptyset\in\mathcal{P}^+(S)$ or $\emptyset\in\mathcal{P}^-(S)$.

Proof. The equality $m^*(\emptyset)=0$ follows immediately from the definition of m^*. Due to the finiteness of S and, consequently, of $\mathcal{P}(S)$, σ-additivity reduces to finite additivity and this property obviously holds for m^*. Or, if $\mathcal{B}_1$, $\mathcal{B}_2$ are disjoint subsets of $\mathcal{P}(S)$, then

$$\begin{aligned} m^*(\mathcal{B}_1\cup\mathcal{B}_2) &= \textstyle\sum_{A\in\mathcal{B}_1\cup\mathcal{B}_2}m(A)=\sum_{A\in\mathcal{B}_1}m(A)+\sum_{A\in\mathcal{B}_2}m(A)= \\ &= m^*(\mathcal{B}_1)+m^*(\mathcal{B}_2). \end{aligned} \tag{12.1}$$

Let the signed measure μ be such that $-\infty<\mu(E)\leq\infty$ holds for each $E\in\mathcal{A}$ (the case when $-\infty\leq\mu(E)<\infty$ holds is processed analogously). Hence, for each $A\subset S$ the relation $-\infty<m(A)=\mu(\{\omega\in\Omega:U(\omega)=A\})\leq\infty$ holds so that, as S and $\mathcal{P}(S)$ are finite sets, the same relation

$$-\infty<m^*(\mathcal{B})=\textstyle\sum_{A\in\mathcal{B}}m(A)\leq\infty \tag{12.2}$$

holds for each $\mathcal{B}\subset\mathcal{P}(S)$. Hence, m^* is signed measure on the measurable space $\langle\mathcal{P}(S),\mathcal{P}(\mathcal{P}(S))\rangle$.

Let $\langle\mathcal{P}^+(S),\mathcal{P}^-(S)\rangle$ be a decomposition of $\mathcal{P}(S)$ such that, for all $A\in\mathcal{P}^+(S)$ ($A\in\mathcal{P}^-(S)$, resp.), the relation $m(A)\geq0$ ($m(A)\leq0$, resp.) holds. Then, evidently, the inclusions $\mathcal{P}_1(S)=\{A\subset S:m(A)>0\}\subset\mathcal{P}^+(S)$ and $\mathcal{P}_2(S)=\{A\subset S:m(A)<0\}\subset\mathcal{P}^+(S)$ are valid. Let $\mathcal{B}\subset\mathcal{P}(S)$. Then the inequalities

$$m^*(\mathcal{B}\cap\mathcal{P}^+(S))=\textstyle\sum_{A\in\mathcal{B}\cap\mathcal{P}^+(S)}m(A)\geq0 \tag{12.3}$$

and

$$m^*(\mathcal{B} \cap \mathcal{P}^-(S)) = \sum_{A \in \mathcal{B} \cap \mathcal{P}^-(S)} m(A) \leq 0 \tag{12.4}$$

obviously hold, as $m(A) \geq 0$ for each $A \in \mathcal{P}^+(S)$ and $m(A) \leq 0$ for each $A \in \mathcal{P}^-(S)$. Hence, $\langle \mathcal{P}^+(S), \mathcal{P}^-(S) \rangle$ is a Hahn decomposition of $\mathcal{P}(S)$ with respect to the signed measure m^* and the (nonnegative and σ-additive) measures $m_1^*(\cdot) = m^*(\cdot \cap \mathcal{P}^+(S))$ and $m_2^*(\cdot) = -m^*(\cdot \cap \mathcal{P}^-(S))$ on $\mathcal{P}(\mathcal{P}(S))$ represent the (obviously uniquely defined) Jordan decomposition of m^*. □

Obviously, *each* Hahn decomposition $\langle \mathcal{P}^+(S), \mathcal{P}^-(S) \rangle$ of $\mathcal{P}(S)$ is such that $\mathcal{P}_1(S) \subset \mathcal{P}^+(S) \subset \{A \subset S : m(A) \geq 0\}$ and $\mathcal{P}_2(S) \subset \mathcal{P}^-(S) \subset \{A \subset S : m(A) \leq 0\}$ hold. Or, if there exists $A \subset S$ such that $m(A) < 0$ and $A \in \mathcal{P}^+(S)$ hold simultaneously, then for $\mathcal{B} = \{S\} \subset \mathcal{P}(S)$ we obtain that $m^*(\mathcal{B} \cap \mathcal{P}^+(S)) = m^*(\{A\}) < 0$. Hence, $\mathcal{P}^+(S)$ is not a positive set and this fact contradicts the assumption that $\langle \mathcal{P}^+(S), \mathcal{P}^-(S) \rangle$ is a Hahn decomposition of $\mathcal{P}(S)$. The case with $A \in \mathcal{P}^-(S)$ such that $m(A) > 0$ is treated analogously.

Theorem 12.3. Let the notations and conditions of Theorem 12.2 hold. Let $\mathcal{A}_0 \subset \mathcal{A}$ be the minimal σ-field of subsets of Ω containing all the sets $\{\omega \in \Omega : U(\omega) = A\}$, $A \subset S$. Let x be an object different from all subsets of S. Then there exist probability measures P_1, P_2, defined on the measurable space $\langle \Omega, \mathcal{A}_0 \rangle$, two random variables U_1, U_2 taking $\langle \Omega, \mathcal{A}_0 \rangle$ into $\langle \mathcal{P}(S) \cup \{x\}, \mathcal{P}(\mathcal{P}(S) \cup \{x\}) \rangle$, and two finite nonnegative real numbers α, β such that, for all $A \subset S$ with $-\infty < bel(U, \mu)(A) < \infty$,

$$bel(U, \mu)(A) = \alpha\, bel(U_1, P_1)(A) - \beta\, bel(U_2, P_2)(A). \tag{12.5}$$

Remark. The values α and β are independent of A. Hence, the value ascribed to $A \subset S$ by the signed belief function $bel(U, \mu)$ can be obtained as a linear combination of the values ascribed to the same set A by two (classical probabilistic) belief functions $bel(U_1, P_1)$ and $bel(U_2, P_2)$. The relation (12.5) can be taken as something like a Jordan decomposition of signed belief functions. If $bel(U, \mu)(A)$ is infinite, (12.5) obviously cannot hold for finite α, β, as $bel(U_i, P_i)(A)$, $i = 1, 2$, are probability values, hence, values embedded within the unit interval of reals.

Proof of Theorem 12.3. Let $\langle \mathcal{P}^+(S), \mathcal{P}^-(S) \rangle$ be a Hahn decomposition of $\mathcal{P}(S)$ with respect to m^*. Define the mappings $U_i : \Omega \to \mathcal{P}(S) \cup \{x\}$, $i = 1, 2$, in this way: $U_1(\omega) = A \subset S$ iff $U(\omega) = A$ and $A \in \mathcal{P}^+(S)$, $U_1(\omega) = x$ otherwise, $U_2(\omega) = A$ iff $U(\omega) = A$ and $A \in \mathcal{P}^-(S)$, $U_2(\omega) = x$ otherwise. Evidently, both U_1, U_2 are measurable mappings, i. e. (generalized, non-numerical) random variables, as $\{\omega \in \Omega : U_i(\omega) = A\} \in \mathcal{A}_0$ holds for both $i = 1, 2$ and for all $A \subset S$ or $A = x$. For $A \subset S$ it is clear, for $A = x$ we obtain that $\{\omega \in \Omega : U_1(\omega) = x\} = \bigcup_{A \in \mathcal{P}^-(S)} \{\omega \in \Omega : U(\omega) = A\}$ and $\{\omega \in \Omega : U_2(\omega) = x\} = \bigcup_{A \in \mathcal{P}^+(S)} \{\omega \in \Omega : U(\omega) = A\}$ and both these sets are in $\mathcal{A}_0$ due to the fact that $\mathcal{P}^+(S)$ and $\mathcal{P}^-(S)$ are finite systems of sets. Moreover, if $U_1(\omega) = A$ for some $A \subset S$, then $m(A) = \mu(\{\omega \in \Omega : U(\omega) = A\}) \geq 0$, and if $U_2(\omega) = A \subset S$, then $m(A) \leq 0$ due to the definitions of U_1 and U_2. As $\{\omega \in \Omega : U(\omega) = \emptyset\} = \emptyset$, also $\{\omega \in \Omega : U_i(\omega) = \emptyset\} = \emptyset$ for both $i = 1, 2$.

Let us suppose that there exist A, $B \subset S$ such that $0 < m(A) < \infty$ and $0 > m(B) > -\infty$ hold, in other words, suppose that $\mathcal{P}_1(S) \cap Fin^+(m,S) \neq \emptyset$ and $\mathcal{P}_2(S) \cap Fin^-(m,S) \neq \emptyset$, where $Fin^+(m,S) = \{A \subset S : m(A) < \infty\}$ and $Fin^-(m,S) = \{A \subset S : m(A) > -\infty\}$; let us recall that $\mathcal{P}_1(S) = \{A \subset S : m(A) > 0\}$ and $\mathcal{P}_2(S) = \{A \subset S : m(A) < 0\}$. Set

$$\begin{aligned} \alpha &= \sum\nolimits_{A \in \mathcal{P}^+(S) \cap Fin^+(m,S)} m(A), \\ \beta &= -\sum\nolimits_{A \in \mathcal{P}^-(S) \cap Fin^-(m,S)} m(A). \end{aligned} \tag{12.6}$$

By assumptions ($\mathcal{P}(S)$ finite, $\mathcal{P}_1(S)$ and $\mathcal{P}_2(S)$ nonempty) we obtain that $0 < \alpha < \infty$ and $0 < \beta < \infty$ hold.

Define $P_i : \mathcal{A}_0 \to \langle 0, 1 \rangle$, $i = 1, 2$, in this way:

$$\begin{aligned} P_1(\{\omega \in \Omega : U(\omega) = A\}) &= m(A)/\alpha, \text{ if } A \in \mathcal{P}^+(S),\ m(A) < \infty, \\ &= 0 \text{ for other } A \subset S, \end{aligned} \tag{12.7}$$

$$\begin{aligned} P_2(\{\omega \in \Omega : U(\omega) = A\}) &= -m(A)/\beta, \text{ if } A \in \mathcal{P}^-(S),\ m(A) > -\infty, \\ &= 0 \text{ for other } A \subset S. \end{aligned} \tag{12.8}$$

Both the mappings P_1, P_2 can be uniquely extended to σ-additive probability measures on the σ-field $\mathcal{A}_0$, as non-negativity is clear and

$$\sum\nolimits_{A \subset S} P_1(\{\omega \in \Omega : U_1(\omega) = A\}) = \sum\nolimits_{A \in \mathcal{P}^+(S) \cap Fin^+(m,S)} m(A) = 1 \tag{12.9}$$

(similarly for P_2 and U_2). Hence, for both $i = 1, 2$, $\langle \Omega, \mathcal{A}_0, P_i \rangle$ is a probability space and $U_i : \Omega \to \mathcal{P}(S) \cup \{x\}$ is a random variable (measurable mapping with respect to the σ-field $\mathcal{P}(\mathcal{P}(S) \cup \{x\})$). Consequently, we can define the classical probabilistic belief functions $bel(U_1, P_1)$ and $bel(U_2, P_2)$ on $\mathcal{P}(S \cup \{x\})$, setting $bel(U_i, P_i)(A) = P_i(\{\omega \in \Omega : U_i(\omega) \subset A\})$ for every $A \subset S \cup \{x\}$. In particular, for $A \subset S$ an easy calculation yields that

$$\begin{aligned} & bel(U_1, P_1)(A) = P_1(\{\omega \in \Omega : U_1(\omega) \subset A\}) = \\ = & \sum\nolimits_{B \subset A} P_1(\{\omega \in \Omega : U_1(\omega) = B\}) = \\ = & \sum\nolimits_{B \subset A,\, B \in \mathcal{P}^+(S),\, m(B) < \infty} P(\{\omega \in \Omega : U_1(\omega) = B\}) = \\ & \sum\nolimits_{B \subset A,\, B \in \mathcal{P}^+(S), m(B) < \infty} m(B)/\alpha = \\ = & (1/\alpha)\, \mu \left(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap \mathcal{P}^+(S) \cap Fin^+(m,S)\} \right), \end{aligned} \tag{12.10}$$

so that

$$\begin{aligned} & \mu(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap \mathcal{P}^+(S) \cap Fin^+(m,S)\}) = \\ = & \alpha\, bel(U_1, P_1)(A). \end{aligned} \tag{12.11}$$

Analogously,

$$\begin{aligned} & bel(U_2, P_2)(A) = \sum\nolimits_{B \subset A,\, B \in \mathcal{P}^+(S),\, m(B) > -\infty} (-m(B)/\beta) = \\ = & -(1/\beta)\, \mu \left(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap \mathcal{P}^-(S) \cap Fin^-(m,S)\} \right) \end{aligned} \tag{12.12}$$

and

$$\begin{aligned} & \mu(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap \mathcal{P}^-(S) \cap Fin^-(m,S)\}) = \\ = & -\beta bel(U_2, P_2)(A). \end{aligned} \tag{12.13}$$

Let $A \subset S$ be such that

$$bel(U,\mu)(A) = \mu(\{\omega \in \Omega : U(\omega) \subset A\}) = \sum_{B \subset A} m(B) < \infty. \tag{12.14}$$

If there exists $B \subset A$ such that $m(B) = \infty$, then it cannot exist $B' \subset S$ such that $m(B') = -\infty$, as in this case $m^*(\{B\}) = \infty$ and $m^*(\{B'\}) = -\infty$, what contradicts the proved fact that m^* is a signed measure on $\mathcal{P}(\mathcal{P}(S))$. Hence, $m(B) = \infty$ for some $B \subset A$ implies that $m(A) = \infty$ as well, in other words, $bel(U,\mu)(A) < \infty$ implies that $m(B) < \infty$ for all $B \subset A$, in symbols, (12.14) yields that $\mathcal{P}(A) \subset Fin^+(m,S)$. Analogously, if $bel(U,\mu)(A) > -\infty$, then $m(B) > -\infty$ for all $B \subset A$, hence, $\mathcal{P}(A) \subset Fin^-(m,S)$. So, if $bel(U,\mu)(A)$ is finite, then $\mathcal{P}(A) \subset Fin^+(m,S) \cap Fin^-(m,S)$ and

$$\begin{aligned} & bel(U,\mu)(A) = \mu(\{\omega \in \Omega : U(\omega) \subset A\}) = \\ = & \mu(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A)\}) = \\ = & \mu(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap \mathcal{P}^+(S) \cap Fin^+(m,S)\}) + \\ & + \mu(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap \mathcal{P}^-(S) \cap Fin^-(m,S)\}) = \\ = & \alpha bel(U_1, P_1)(A) - \beta bel(U_2 P_2)(A) \end{aligned} \tag{12.15}$$

by (12.11) and (12.13), so that (12.5) holds.

Let us consider, now, the case when $\mathcal{P}_1(S)$ or $\mathcal{P}_2(S)$ is empty. If $\mathcal{P}_1(S) = \mathcal{P}_2(S) = \emptyset$, then $m(A) = 0$ for all $A \subset S$ so that $bel(U,\mu)(A) = 0$ for all $A \subset S$. Consequently,

$$bel(U,\mu)(A) = 0 \cdot bel(U_1,P_1)(A) - 0 \cdot bel(U_2,P_2)(A) \tag{12.16}$$

for no matter which probability measures P_1, P_2 on $\mathcal{A}_0$. Let $\mathcal{P}_1(S) \neq \emptyset$, $\mathcal{P}_2(S) = \emptyset$, so that $m(A) \geq 0$ holds for each $A \subset S$. Then $\langle \mathcal{P}(S), \emptyset \rangle$ is also a Hahn decomposition of $\mathcal{P}(S)$ with respect to m^* (here $\emptyset$ is the empty subset of $\mathcal{P}(S)$, i.e., the empty system of subsets of S). As the Jordan decomposition of m^* is independent of the Hahn decomposition of $\mathcal{P}(S)$, we can replace $\langle \mathcal{P}^+(S), \mathcal{P}^+(S) \rangle$ by $\langle \mathcal{P}(S), \emptyset \rangle$ and apply the same way of reasoning as above. Let there exist at least one $A \subset S$ such that $0 < m(A) < \infty$, i.e. let $\mathcal{P}_1(S) \cap Fin^+(m,S) \neq \emptyset$. Then

$$0 < \alpha = \sum_{A \subset Fin^+(m,S)} m(A) < \infty \tag{12.17}$$

holds, so that, for P_1 defined as above with $\mathcal{P}^+(S) = \mathcal{P}(S)$ and for P_2 defined arbitrarily on $\mathcal{A}_0$, we obtain that

$$bel(U_1,P_1)(A) = (1/\alpha)\,\mu(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap Fin^+(m,S)\}). \tag{12.18}$$

In the same way as in (12.7) we prove that if $bel(U,\mu)(A) < \infty$, then $\mathcal{P}(A) \subset Fin^+(m,S)$, so that

$$\begin{aligned} & bel(U,\mu) = \mu(\{\omega \in \Omega : U(\omega) \subset A\}) = \\ = & \mu(\{\omega \in \Omega : U(\omega) \in \mathcal{P}(A) \cap Fin^+(m,S)\}) = \\ = & \alpha bel(U_1,P_1)(A) - 0 \cdot bel(U_2,P_2)(A), \end{aligned} \tag{12.19}$$

so that (12.7) holds again with $\beta = 0$. The case when $\mathcal{P}_1(S) = \emptyset$ and $\mathcal{P}_2(S) \cap Fin^{-}(m, S) \neq \emptyset$ is processed analogously. Finally, if $m(A) = \infty$ for all $A \subset S$ (or if $m(A) = -\infty$ for all $A \subset S$), then there is no $A \subset S$ with finite and nonozero value $bel(U, \mu)\,(A)$, so that (12.7) holds either trivially, or with $\alpha = \beta = 0$. The theorem is proved. □

Let us extend the definition of conditioned belief functions to the generalized versions of belief functions as defined above. The assertion following this definition proves that if the basic space S is finite, the well-known combinatorial expressions for the conditioned belief functions are valid in the generalized case as well. As elsewhere in this text, we shall limit ourselves to the purely formalized and mathematical aspects of conditioning operations for belief functions, leaving aside the motivation and interpretation (the reader is kindly invited to consult [9] or other sources for this sake.)

Definition 12.1. Let $\langle \Omega, \mathcal{A} \rangle$ be a measurable space, let μ be a probability measure (a measure, a signed measure, resp.) defined on $\langle \Omega, \mathcal{A} \rangle$, let S be a nonempty finite set, let $U : \langle \Omega, \mathcal{A} \rangle \to \langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$ be a measurable mapping which takes Ω into $\mathcal{P}(S)$, let T be a nonempty subset of S. The *conditioned classical probabilistic (generalized, signed, resp.) belief function* $bel(U, \mu)\,(\cdot/T)$ is the mapping of $\mathcal{P}(S)$ into $\langle 0, 1 \rangle$ defined, for each $A \subset S$, by the relation: $bel(U, \mu)\,(A/T) = 0$, if $A \not\subset T$,

$$bel(U, \mu)\,(A/T) = \mu(\{\omega \in \Omega : \emptyset \neq U(\omega) \cap T \subset A \cap T\}) \tag{12.20}$$

otherwise. Evidently, $bel(U, \mu)\,(\cdot) = bel(U, \mu)\,(\cdot/S)$ for all the three kinds of belief functions.

Theorem 12.4. Let the notations and conditions of Definition 12.1 hold. Then for the conditioned classical probabilistic (generalized, signed, resp.) belief function $bel(U, \mu)\,(\cdot/T)$ the relation

$$\begin{aligned} & bel(U, \mu)\,(A, T) = bel(U, \mu)\,(A \cup (S - T)) - bel(U, \mu)\,(S - T) = \\ = \; & \textstyle\sum_{\emptyset \neq B \subset A \cap T} \sum_{X \subset S - T} m(B \cup X) \end{aligned} \tag{12.21}$$

holds for each $A \subset T$, where $m(C) = \mu(\{\omega \in \Omega : U(\omega) = C\})$ for all $C \subset S$. □

Proof. Because of the evident fact that probability measures are a special case of measures and the later ones are a special case of signed measures, it is sufficient to prove the assertion for the case when μ is a signed measure. For the sake of simplicity we omit the parameters U and μ throughout this proof.

As S and, consequently, $\mathcal{P}(S)$ are finite sets, the σ-additivity of μ immediately yields that, for $A \subset T$,

$$bel(A/T) = \textstyle\sum_{B \subset S,\, \emptyset \neq B \cap T \subset A \cap T} \mu(\{\omega \in \Omega : U(\omega) = B\}). \tag{12.22}$$

Each $B \subset S$ such that $\emptyset \neq B \cap T \subset A \cap T$ can be uniquely decomposed into disjoint subsets $\emptyset \neq B_1 = B \cap T$ and $B_2 = B \cap (S - T)$, conversely, for each $\emptyset \neq B_1 \subset A \cap T$ and

$B_2 \subset S-T$ the set $B = B_1 \cup B_2$ belongs to the class of sets over which the summation in (12.22) is defined. As the mapping $B \leftrightarrow \langle B_1, B_2 \rangle$ is one-to-one, (12.22) immediately yields that

$$bel(A/T) = \sum_{\emptyset \neq B_1 \subset A \cap T} \sum_{B_2 \subset S-T} m(B_1 \cup B_2) \tag{12.23}$$

and the equality between the first and the third item in (12.21) is proved. In fact, it is just this equality which serves, as a rule, as the definition of conditioned belief function in the works dealing with belief functions over finite spaces S.

By the definition of the (unconditioned) signed belief function we obtain that

$$\begin{aligned} & bel(A \cup (S-T)) - bel(S-T) = \sum_{B \subset A \cup (S-T)} m(B) - \sum_{B \subset S-T} m(B) \qquad (12.24) \\ = \; & \sum_{B \subset A \cup (S-T),\, B \not\subset S-T} m(B) = \\ = \; & \sum_{B \in \{C \subset S : C \cap A \cap T \neq \emptyset,\, C \subset (A \cap T) \cup (S-T)\}} m(B) = \\ = \; & \sum_{B = B_1 \cup B_2,\, \emptyset \neq B_1 \subset A \cap T,\, B_2 \subset S-T} m(B_1 \cup B_2) = \\ = \; & \sum_{\emptyset \neq B_1 \subset A \cap T} \sum_{B_2 \subset S-T} m(B_1 \cup B_2). \end{aligned}$$

The assertion is proved. □

Theorem 12.5. Let the notations and conditions of Theorem 12.2 hold, let $\mathcal{A}_0$ and x be as in Theorem 12.3, let $bel(U,\mu)\,(A/T)$ be defined as in Definition 12.1. Then there exist probability measures P_1, P_2, random variables U_1, U_2, and finite nonnegative real numbers α, β with the properties claimed in Theorem 12.3 and such that, for all $A \subset T \subset S$ with the property that $bel(U,\mu)\,(A \cup (S-T))$ and $bel(U,\mu)\,(S-T)$ are finite numbers, the equality

$$bel(U,\mu)\,(A/T) = \alpha bel(U_1,P_1)\,(A/T) - \beta bel(U_2,P_2)\,(A/T) \tag{12.25}$$

holds. □

Proof. Applying (12.5) to the sets $A \cup (S-T) \subset S$ and $S-T \subset S$ we obtain, by an easy calculation, that

$$\begin{aligned} & bel(U,\mu)\,(A/T) = bel(U,\mu)\,(A \cup (S-T)) - bel(U,\mu)\,(S-T) = \quad (12.26) \\ = \; & [\alpha bel(U_1,P_1)\,(A \cup (S-T)) - \beta bel(U_2,P_2)\,(A \cup (S-T))] - \\ & - [\alpha bel(U_1,P_1)\,(S-T) - \beta bel(U_2,P_2)\,(S-T)] = \\ & \alpha\,[bel(U_1,P_1)\,(A \cup (S-T)) - bel(U_1,P_1)\,(S-T)] - \\ & -\beta\,[bel(U_2,P_2)\,(A \cup (S-T)) - bel(U_2,P_2)\,(S-T)] = \\ = \; & \alpha bel(U_1,P_1)\,(A/T) - \beta bel(U_2,P_2)\,(A/T). \end{aligned}$$

□

Hence, Jordan decomposition of a conditioned signed belief function can be obtained as the linear combination of the corresponding conditioned classical probabilistic belief functions with the same coefficients as in the unconditioned case. The last fact is quite intuitive, as in the case when $T = S$ (12.25) immediately converts into (12.5).

13 Monte–Carlo Estimations for Belief Functions

Let S be a finite set, let E be an empirical space, let $\rho : S \times E \to \{0,1\}$ be a compatibility relation, let $X : \langle \Omega, \mathcal{A}, P \rangle \to \langle E, \mathcal{E} \rangle$ be a random variable, let $bel_{m_{\rho,X}}$ be the belief function defined, for each $A \subset S$, by

$$bel_{m_{\rho,X}}(A) = P(\{\omega \in \Omega : U_\rho(X(\omega)) \subset A\}). \tag{13.1}$$

Let us suppose that there exists, for each $x \in E$, a value $s \in S$ such that $\rho(s,x) = 1$ so that $U_\rho(X(\omega)) \neq \emptyset$ holds for every $\omega \in \Omega$, hence, $m_{\rho,X}(\emptyset) = 0$.

Taken as a mapping from $\mathcal{P}(S)$ into $\langle 0, 1 \rangle$, belief function does not meet, in general, the demands imposed to probability measure, as a matter of fact, $bel_{m_{\rho,X}}$ is additive iff $bel_{m_{\rho,X}}(A) = 0$ for all $A \subset S$ such that $\text{card}(A) \neq 1$. However, as the particular values of belief functions can be defined, due to (13.1), by values of probability measures ascribed to certain random events, they can be also estimated using the well-known Monte-Carlo methods (or probabilistic algorithms, in other terms). The basic idea behind is very simple: to take arithmetical means, defined by a large enough finite number of stochastically independent and identically distributed (i.i.d.) random samples, as reasonable and good enough estimations of the corresponding expected values. The theoretical soundness of the basic idea of Monte-Carlo methods as just outlined is based on two elementary assertions of the axiomatic probability theory: *the strong law of large number* (SLLN) and the *Chebyshev inequality*. Both of these statements will be used below in the form presented in Chapter 2 (Part One of this Research Report) so that it is not necessary to recall them here in more detail.

The most obvious and immediate application of the SLLN to the case of belief functions reads as follows.

Theorem 13.1. Let $\langle \Omega, \mathcal{A}, P \rangle$ be a probability space, let $\langle \mathcal{P}(S), \mathcal{S} \rangle$ be a measurable space over a nonempty set S such that $\mathcal{P}(T) \in \mathcal{S}$ holds for each $T \subset S$, let U be a (generalized set-valued) random variable defined on $\langle \Omega, \mathcal{A}, P \rangle$, taking its values in $\langle \mathcal{P}(S), \mathcal{S} \rangle$, and such that $U(\omega) \neq \emptyset$ holds for each $\omega \in \Omega$. Let $U_1, U_2, \ldots$ be a sequence of statistically independent random variables defined on $\langle \Omega, \mathcal{A}, P \rangle$, taking their values in $\langle \mathcal{P}(S), \mathcal{S} \rangle$, and possessing the same probability distribution as U, so that for each $k \in \mathcal{N}^+ = \{1, 2, \ldots\}$, each $n_1 < n_2 < \cdots < n_k$, and $\mathcal{V}_1, \mathcal{V}_2, \ldots, V_k \in \mathcal{S}$ the equality

$$P\left(\bigcap_{i=1}^{k} \{\omega \in \Omega : U_{n_i}(\omega) \in \mathcal{V}_i\}\right) = \prod_{i=1}^{k} P(\{\omega \in \Omega : U_{n_i}(\omega) \in \mathcal{V}_i\}) \tag{13.2}$$

holds. Moreover, for each $i \in \mathcal{N}^+$ and each $\mathcal{V} \in \mathcal{S}$

$$P\left(\{\omega \in \Omega : U_i(\omega) \in \mathcal{V}\}\right) = P\left(\{\omega \in \Omega : U(\omega) \in \mathcal{V}\}\right). \tag{13.3}$$

Let $T \subset S$, let $\chi_{\mathcal{P}(T)}$ be the characteristic function (identifier) of the system $\mathcal{P}(T)$ of all subsets of T, so that $\chi_{\mathcal{P}(T)}(A) = 1$, if $A \in \mathcal{P}(T)$, i.e., if $A \subset T$, $\chi_{\mathcal{P}(T)}(A) = 0$ otherwise. Let

$$bel_U(T) = P\left(\{\omega \in \Omega : U(\omega) \subset T\}\right) \tag{13.4}$$

denote the belief function induced by the random variable U. Then

$$P\left(\left\{\omega \in \Omega : \lim_{n \to \infty}\left[n^{-1} \sum_{i=1}^{n} \chi_{\mathcal{P}(T)}(U_i(\omega))\right] = bel_U(T)\right\}\right) = 1. \tag{13.5}$$

□

Proof. An immediate consequence of SLLN. Setting

$$A_i = \{\omega \in \Omega : U_i(\omega) \in \mathcal{P}(T)\} \quad (= \{\omega \in \Omega : U_i(\omega) \subset T\}), \tag{13.6}$$

we can easily observe that $\langle A_1, A_2, \ldots \rangle$ is a sequence of statistically independent random events such that $P(A_i) = bel_U(T)$ holds for each $i \in \mathcal{N}^+$. Hence, $n^{-1} \sum_{i=1}^{n} \chi_{\mathcal{P}(T)}(U_i(\omega))$ is the relative frequency of occurences of a random event with the probability $bel_U(T)$ in a sequence of n statistically independent realizations, and this relative frequency tends to $bel_U(T)$ almost surely. □

Before focusing our attention to more sophisticated and complex Monte-Carlo algorithms for belief functions we should mention, at least very briefly, our motivation for such an effort. In order to define and to compute the value $bel_U(T)$ we need not only the notion "state compatible with an observation", defined by a binary compatibility relation between states and empirical values at the *intensional* level, but also the *extension* of this notion in the form of a subset of the state space S. If S is finite, such an extension can be obtained by testing, sequentially for each $s \in S$, whether s is compatible with the observation in question, if neglecting, for the moment, the problems with high computational complexity of such a procedure for large S. If S

is infinite, this approach is impossible and problems arise with the definability of the set of compatible states in terms acceptable by the used formal apparatus. So, our aim will be to suggest a Monte-Carlo algorithm for belief functions using the binary compatibility relation between states and empirical values just at the level of a black box outputing the answers YES (compatible) of NO (incompatible) for each input pair "state-empirical value", but incapable to generate sets of states compatible with a given empirical observation.

Let S, $\langle E, \mathcal{E} \rangle$ and $\rho : S \times E \rightarrow \{0,1\}$ be as above. A random variable X defined on the probability space $\langle \Omega, \mathcal{A}, P \rangle$ and taking its values in the measurable space $\langle E, \mathcal{E} \rangle$ is called *regular with respect to the compatibility relation* ρ, if the mapping $U_{X,\rho} : \Omega \rightarrow \mathcal{P}(S)$, defined by

$$U_{X,\rho}(\omega) = U_{m_\rho, X}(\omega) = \{s \in S : \rho(s, X(\omega)) = 1\} \tag{13.7}$$

and each $\omega \in \Omega$, is a random variable defined on $\langle \Omega, \mathcal{A}, P \rangle$ and taking its values in the measurable space $\langle \mathcal{P}(S), \mathcal{S} \rangle$, where $\mathcal{S}$ is the minimal σ-field over the system $\{\mathcal{P}(T) : T \subset S\} \subset \mathcal{P}(\mathcal{P}(S))$. Hence, if X is regular with respect to ρ, then $\{\omega \in \Omega : U_{X,\rho}(X(\omega)) \subset T\} \in \mathcal{A}$ holds for each $T \subset S$, so that the value $bel_{U_{X,\rho}}(T) = P(\{\omega \in \Omega : U_{X,\rho}(\omega) \subset T\})$ is defined for all $T \subset S$. Let $X_1, X_2, \ldots$ be a sequence of statistically independent and identically distributed random variables defined on $\langle \Omega, \mathcal{A}, P \rangle$, taking their values in $\langle E, \mathcal{E} \rangle$ and regular with respect to ρ.

Theorem 13.2. Let $U_i = U_{X_i,\rho}$ for each $i = 1, 2, \ldots$. Then $U_1, U_2, \ldots$ is a sequence of statistically independent and identically distributed random variables defined on $\langle \Omega, \mathcal{A}, P \rangle$ and taking their values in $\langle \mathcal{P}(S), \mathcal{S} \rangle$. □

Remark. In the rest of this chapter we shall omit the symbols "$\ldots \omega \in \Omega : \ldots$" in the expressing defining particular subsets of Ω, if no misunderstanding menaces.

Proof. Let $\mathcal{V} \in \mathcal{S}$, let $i \in \mathcal{N}^+$. Then

$$\begin{aligned} & P(\{U_i(\omega) \in \mathcal{V}\}) = P(\{U(X_i(\omega)) \in \mathcal{V}\}) = P(\{X_i(\omega) \in U^{-1}(\mathcal{V})\}) = \\ = \;& P(\{X_i(\omega) \in \{x \in E : U(x) \in \mathcal{V}\}\}) = P(\{X_1(\omega) \in \{x \in E : U(x) \in \mathcal{V}\}\}) = \\ = \;& P(\{U(X_1(\omega)) \in \mathcal{V}\}) = P(\{U_1(\omega) \in \mathcal{V}\}) \end{aligned} \tag{13.8}$$

due to the identical distribution of X_i and X_1, so that the random variables U_i and U_1 are also identically distributed. Let $k \in \mathcal{N}^+$, let $n_1 < n_2 < \cdots < n_k$ be positive integers, let $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k \in \mathcal{S}$. Then

$$\begin{aligned} & P\left(\bigcap_{i=1}^k \{U_{n_i}(\omega) \in \mathcal{V}_i\}\right) = P\left(\bigcap_{i=1}^k \{U(X_{n_i}(\omega)) \in \mathcal{V}_i\}\right) = \\ = \;& P\left(\bigcap_{o=1}^k \{X_{n_i}(\omega) \in U^{-1}(\mathcal{V}_i)\}\right) = \prod_{i=1}^k P(\{X_{n_i}(\omega) \in U^{-1}(\mathcal{V}_i)\}) = \\ = \;& \prod_{i=1}^k P(\{U(X_{n_i}(\omega)) \in \mathcal{V}_i\}) = \prod_{i=1}^k P(\{U_{n_i}(\omega) \in \mathcal{V}_i\}) \end{aligned} \tag{13.9}$$

due to the statistical independence of the random variables $X_1, X_2, \ldots$. So, $U_1, U_2, \ldots$ is a sequence of statistically independent and identically distributed (generalized set-valued) random variables. □

Suppose, since now on, that the state space S is finite or countable. Let $\{Y_{ij}\}_{i,j=1}^{\infty}$ be a system of statistically independent and identically distributed random variables defined on $\langle \Omega, \mathcal{A}, P \rangle$, taking their values in the measurable space $\langle S, \mathcal{P}(S) \rangle$ and such that, for each $s \in S$, $P(\{Y_{11}(\omega) = s\}) > 0$ holds. What we have at our disposal, it is the system $\{\rho(Y_{ij}(\omega))\}_{i=1}^{n}{}_{j=1}^{m(i)}$ for some n, $m(i) \in \mathcal{N}^{+}$ of binary values and our aim will be to estimate the value $bel_{U_{X_1,\rho}}(T)$, for a given $T \subset S$, on the ground of these data.

Informally, if for some $i \in \mathcal{N}^{+}$, for $m(i)$ large enough, and for all $j \leq m(i)$ such that $Y_{ij}(\omega)$ is compatible with $X_i(\omega)$ also the relation $Y_{ij}(\omega) \in T$ holds, we are temptated to believe that *all* $s \in S$ compatible with $X_i(\omega)$ belong to T, in other words, that the inclusion $U(X_i(\omega)) \subset T$ holds. So, setting $\lambda_i(T,\omega) = 1$, if for all $j \leq m(i)$ such that $\rho(Y_{ij}(\omega), X_i(\omega)) = 1$ also $\chi_T(Y_{ij}(\omega)) = 1$, and setting $\lambda_i(T,\omega) = 0$ otherwise, we feel that $\lambda_i(T,\omega)$ is a reasonable and relatively good approximation of the value $\chi_{\mathcal{P}(T)}$ $(U_{X_i,\rho})$. Defining $\lambda_i(T,\omega)$ more formally we can write that

$$\begin{aligned} \lambda_i(T,\omega) &= 1, \quad \text{if } (\forall\, j \leq m(i))\, (\rho(Y_{ij}(\omega), X_i(\omega)) \leq \chi_T(Y_{ij}(\omega))), \\ \lambda_i(T,\omega) &= 0 \quad \text{otherwise,} \end{aligned} \tag{13.10}$$

hence

$$\lambda_i(T,\omega) = \prod\nolimits_{j=1}^{m(i)} \varphi_{ij}(T,\omega), \tag{13.11}$$

where

$$\begin{aligned} \varphi_{ij}(T,\omega) &= 1, \quad \text{if } \rho(Y_{ij}(\omega), X_i(\omega)) \leq \chi_T(Y_{ij}(\omega)), \\ \varphi_{ij}(T,\omega) &= 0 \quad \text{otherwise.} \end{aligned} \tag{13.12}$$

Accepting the convention that $0^0 = 1$ we obtain easily that

$$\begin{aligned} \varphi_{ij}(T,\omega) &= \min\left[1,\, 1 - \rho(Y_{ij}(\omega), X_i(\omega)) + \chi_T(Y_{ij}(\omega))\right] = \\ &= \left[\chi_T(Y_{ij}(\omega))\right]^{\rho(Y_{ij}(\omega), X_i(\omega))}, \end{aligned} \tag{13.13}$$

and

$$\begin{aligned} \lambda_i(T,\omega) &= \prod\nolimits_{j=1}^{m(i)} \min\left(1,\, 1 - \rho(Y_{ij}(\omega), X_i(\omega)) + \chi_T(Y_{ij}(\omega))\right) = \\ &= \prod\nolimits_{j=1}^{m(i)} \left[\chi_T(Y_{ij}(\omega))\right]^{\rho(Y_{ij}(\omega), X_i(\omega))}. \end{aligned} \tag{13.14}$$

Using the values $\lambda_i(T,\omega)$ as approximations of $\chi_{\mathcal{P}(T)}(U_{X_i,\rho}(\omega))$, like as in Theorem 13.2, we arrive at the following random variable.

$$\begin{aligned} & B_{X_1,\rho}\left(T, n, \langle m(i) \rangle_{i=1}^{n}, \omega\right) = n^{-1} \sum\nolimits_{i=1}^{n} \lambda_i(T,\omega) = \\ &= n^{-1} \sum\nolimits_{i=1}^{n} \prod\nolimits_{j=1}^{m(i)} \min\left(1,\, 1 - \rho(Y_{ij}(\omega), X_i(\omega)) + \chi_T(Y_{ij}(\omega))\right) = \\ &= n^{-1} \sum\nolimits_{i=1}^{n} \prod\nolimits_{j=1}^{m(i)} \left[\chi_T(Y_{ij}(\omega))\right]^{\rho(Y_{ij}(\omega), X_i(\omega))}. \end{aligned} \tag{13.15}$$

We shall try to demonstrate, in the rest of this chapter, that, and in which sense and degree, the value $B_{X_1,\rho}(T, n, \langle m(i) \rangle_{i=1}^{n}, \omega)$ is a reasonable and good estimation of the value $bel_{U_{X_1,\rho}}(T)$.

Theorem 13.3. Let the systems $\{X_i\}_{i=1}^n$ and $\{Y_{ij}\}_{i=1,\,j=1}^{n\quad m(i)}$ of random variables defined above be mutually statistically independent, let $\langle E, \mathcal{E}\rangle$ be such that E is finite or countable and $\mathcal{E} = \mathcal{P}(E)$, let $\varepsilon > 0$ be given. Then there exists a value $\lambda^* \in \langle bel_{U_{X_1,\rho}}(T), bel_{U_{X_1,\rho}}(T) + \varepsilon\rangle$ and a value $m^* \in \mathcal{N}^+$ such that, if $m(i) \geq m^*$ for each $i \in \mathcal{N}^+$, then the relation

$$P\left(\left\{\omega \in \Omega : \lim_{n\to\infty} B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) = \lambda^*\right\}\right) = 1 \tag{13.16}$$

holds. □

Proof. First, let us prove that, given $T \subset S$, $\langle \lambda_i(T,\cdot)\rangle_{i=1}^n$ is a sequence of statistically independent random variables. So, we have to prove that the equality

$$P\left(\bigcap_{i=1}^n \{\lambda_i(T,\omega) = \kappa_i\}\right) = \prod_{i=1}^n P(\{\lambda_i(T,\omega) = \kappa_i\}) \tag{13.17}$$

holds for each n-tuple $\langle \kappa_1, \kappa_2, \ldots, \kappa_n\rangle \in \{0,1\}^n$. The proof will be done by induction on the number of occurrences of zeros in $\langle \kappa_1, \kappa_2, \ldots, \kappa_n\rangle$, so that the first step deals with the case when $k = 0$, i. e., when $\kappa_i = 1$ for each $i \leq n$.

Let $F_T(s) = \{x \in E : \rho(s,x) = 0\}$, if $s \in S - T$, set $F_T(s) = E$ otherwise. An easy calculation using the supposed statistical independences among particular random variables $X_i,\ i \leq n,\ Y_{ij},\ i \leq n,\ j \leq m(i)$, as well as between the systems $\{X_i\}_{i=1}^n$ and $\{Y_{ij}\}_{i=1,\,j=1}^{n\quad m(i)}$ of random variables yield that

$$P\Big(\bigcap_{i=1}^n \{\lambda_i(T,\omega) = 1\}\Big) = P\Big(\bigcap_{i=1}^n \bigcap_{j=1}^{m(i)} \{\rho(Y_{ij}(\omega), X_i(\omega)) \leq \chi_T(Y_{ij}(\omega))\}\Big) \tag{13.18}$$

$$= P\Big(\bigcup_{\langle\langle s_1^1,\ldots,s_{m(1)}^1\rangle,\ldots,\langle s_1^n,\ldots,s_{m(n)}^n\rangle\rangle \in \mathsf{X}_{i=1}^n S^{m(i)}} \bigcap_{\langle j_1,\ldots,j_n\rangle = \langle 1,1,\ldots,1\rangle}^{\langle m(1),m(2),\ldots,m(n)\rangle} \{\rho(s_{j_i}^i, X_i(\omega)) \leq \chi_T(s_{j_i}^i),\ Y_{ij}(\omega) = s_{j_i}^i\}\Big) =$$

$$= \sum_{\langle\langle s_1^1,\ldots,s_{m(1)}^1\rangle,\ldots,\langle s_1^n,\ldots,s_{m(n)}^n\rangle\rangle \in \mathsf{X}_{i=1}^n S^{m(i)}} P\Big(\bigcap_{i=1}^n \bigcap_{j=1}^{m(i)} \{X_i(\omega) \in F_T(s_j^i), Y_{ij}(\omega) = s_j^i\}\Big) =$$

$$= \sum_{\langle\langle s_1^1,\ldots,s_{m(1)}^1\rangle,\ldots,\langle s_1^n,\ldots,s_{m(n)}^n\rangle\rangle \in \mathsf{X}_{i=1}^n S^{m(i)}} \prod_{i=1}^n \prod_{j=1}^{m(i)} P(\{X_i(\omega) \in F_T(s_j^i), Y_{ij}(\omega) = s_j^i\}) =$$

$$= \sum_{\langle s_1^n,\ldots,s_{m(n)}^n\rangle \in S^{m(n)}} \sum_{\langle\langle s_1^1,\ldots,s_{m(1)}^1\rangle,\ldots,\langle s_1^{n-1},\ldots,s_{m(n-1)}^{n-1}\rangle\rangle \in \mathsf{X}_{i=1}^{n-1} S^{m(i)}} \left[\prod_{i=1}^{n-1} \prod_{j=1}^{m(i)} P(\{X_i(\omega) \in F_T(s_j^i), Y_{ij}(\omega) = s_j^i\})\right] \cdot \left[\prod_{k=1}^{m(n)} P(\{X_n(\omega) \in F_T(s_k^n), Y_{nk}(\omega) = s_k^n\})\right] =$$

$$= \left[\sum_{\langle s_1^n,\ldots,s_{m(n)}^n\rangle \in S^{m(n)}} \prod_{k=1}^{m(n)} P(\{X_n(\omega) \in F_T(s_k^n), Y_{nk}(\omega) = s_k^n\})\right] \cdot \Big[\sum_{\langle\langle s_1^1,\ldots,s_{m(1)}^1\rangle,\ldots,\langle s_1^{n-1},\ldots,s_{m(n-1)}^{n-1}\rangle\rangle \in \mathsf{X}_{i=1}^{n-1} S^{m(i)}} \prod_{i=1}^n \prod_{j=1}^{m(i)} P(\{X_i(\omega) \in F_T(s_j^i), Y_{ij}(\omega) = s_j^i\})\Big].$$

Applying the same calculation to the last component in the last right-hand side expression in (13.18) we obtain, after $n-1$ further steps, that

$$\begin{aligned} & P\left(\bigcap_{i=1}^n \{\lambda_i(T,\omega)=1\}\right) = \qquad (13.19)\\ = & \prod_{i=1}^n \left[\sum_{\langle s_1^j,\dots,s_{m(i)}^i\rangle\in S^{m(i)}} \prod_{k=1}^{m(i)} P(\{X_i(\omega)\in F_T(s_k^i), Y_{ik}(\omega)=s_k^i\})\right] = \\ = & \prod_{i=1}^n \left[\sum_{\langle s_1^j,\dots,s_{m(i)}^i\rangle\in S^{m(i)}\rangle} \prod_{k=1}^{m(i)} P(\{\rho(s_k^i, X_i(\omega)\le \chi_T(s_k^i), Y_{ik}(\omega)=s_k^i\})\right] = \\ = & \prod_{i=1}^n P\left(\bigcap_{k=1}^{m(i)} \{\rho(X_i(\omega),Y_{ik}(\omega))\le\chi_T(Y_{ik}(\omega))\}\right) \\ = & \prod_{i=1}^n ((\{\lambda_i(T,\omega)=1\})). \end{aligned}$$

Hence, the first step of our induction, i.e., the validity of the relation (13.17) for $\langle\kappa_1,\kappa_2,\dots,$
$\dots,\kappa_n\rangle = \langle 1,1,\dots,1\rangle$, is proved.

Let (13.17) hold for $k\le K$,, i.e., for each $\langle\kappa_1,\kappa_2,\dots,\kappa_n\rangle\in\{0,1\}^n$ such that $\sum_{i=1}^n\kappa_i\ge n-K$, let $\langle\kappa_1,\kappa_2,\dots,\kappa_n\rangle$ be such that $\sum_{i=1}^n\kappa_i = n-K-1$. Take arbitrarily $i_0\le n$ such that $\kappa_{i_0}=0$, and set $\kappa_i^*=\kappa_i$ for all $i\ne i_0$, $\kappa_{i_0}^*=1$, hence, $\sum_{i=1}^n\kappa_i^*=n-K$. Then

$$\begin{aligned} & P\left(\bigcap_{i=1}^n\{\lambda_i(T,\omega)=\kappa_i\}\right) + P\left(\bigcap_{i=1}^n\{\lambda_i(T,\omega)=\kappa_i^*\}\right) = \qquad (13.20)\\ = & P\left(\bigcap_{i=1}^n\{\lambda_i(T,\omega)=\kappa_i\}\right) + \prod_{i=1}^n P\left(\{\lambda_i(T,\omega)=\kappa_i\}\right) = \\ = & P\left(\bigcap_{i=1}^n\{\lambda_i(T,\omega)=\kappa_i^*\}\right) = \prod_{i=1,\, i\ne i_0}^n P\left(\{\lambda_i(T,\omega)=\kappa_i^*\}\right), \end{aligned}$$

so that

$$\begin{aligned} & P\left(\bigcap_{i=1}^n\{\lambda_i(T,\omega)=\kappa_i^*\}\right) = \prod_{i=1,\, i\ne i_0}^n P\left(\{\lambda_i(T,\omega)=\kappa_i^*\}\right) \qquad (13.21)\\ & \quad -\prod_{i=1}^n P\left(\{\lambda_i(T,\omega)=\kappa_i^*\}\right) = \\ = & \prod_{i=1,\, i\ne i_0}^n P\left(\{\lambda_i(T,\omega)=\kappa_i^*\}\right)[1-P(\{\lambda_{i_0}(T,\omega)=\kappa_i^*\})] = \\ = & \prod_{i=1,\, i\ne i_0}^n P\left(\{\lambda_i(T,\omega)=\kappa_i^*\}\right)[1-P(\{\lambda_{i_0}(T,\omega)=1\})] = \\ = & \left[\prod_{i=1,\, i\ne i_0}^n P\left(\{\lambda_i(T,\omega)=\kappa_i^*\}\right)\right] P(\{\lambda_{i_0}(T,\omega)=0\}) = \\ = & \left[\prod_{i=1,\, i\ne i_0}^n P\left(\{\lambda_i(T,\omega)=\kappa_i^*\}\right)\right] P(\{\lambda_{i_0}(T,\omega)=\kappa_{i_0}\}) = \\ = & \prod_{i=1}^n P(\{\lambda_i(T,\omega)=\kappa_i\}). \end{aligned}$$

So, (13.17) holds also for all $\langle\kappa_1,\kappa_2,\dots,\kappa_n\rangle$ with $K+1$ occurrences of zero, hence, by induction, (13.17) holds for each $\langle\kappa_1,\kappa_2,\dots,\kappa_n\rangle\in\{0,1\}^n$, and the random variables $\{\lambda_i(T,\omega)\}_{i=1}^n$ are proved to be statistically independent. Our further goal will be to compute the value $P(\{\lambda_i(T,\omega)=1\})$, which is evidently identical with the expected value $E\,\lambda_i(T,\cdot)$. An easy calculation yields that

$$P(\{\lambda_i(T,\omega)=1\}) = P\left(\bigcap_{j=1}^{m(i)}\{\rho(Y_{ij}(\omega),X_i(\omega))\le\chi_T(Y_{ij}(\omega))\}\right) = \qquad (13.22)$$

$$= \; P\left(\left[\bigcap_{j=1}^{m(i)}\{\rho(Y_{ij}(\omega),X_i(\omega))\le\chi_T(Y_{ij}(\omega))\}\right]\cap\{U(X_i(\omega))\subset T\}\right)+$$
$$+P\left(\left[\bigcap_{j=1}^{m(i)}\{\rho(Y_{ij}(\omega),X_i(\omega))\le\chi_T(Y_{ij}(\omega))\}\right]\cap\{U(X_i(\omega))\not\subset T\}\right)=$$
$$= \; P(\{U(X_i(\omega))\subset T\})+$$
$$+P\left(\left[\bigcap_{j=1}^{m(i)}\{\rho(Y_{ij}(\omega),X_i(\omega))\le\chi_T(Y_{ij}(\omega))\}\right]\cap\{U(X_i(\omega))\not\subset T\}\right),$$

as $U(X_i(\omega))\subset T$ implies that if $Y_{ij}(\omega)\in U(X_i(\omega))$, i.e., if $\rho(Y_{ij}(\omega),X_i(\omega))=1$, then $Y_{ij}(\omega)\in T$, hence, $\chi_T(Y_{ij}(\omega))=1$, consequently, for all $j\le m(i)$ the inequality $\rho(Y_{ij}(\omega),X_i(\omega))\le\chi_T(Y_{ij}(\omega))$ holds. Another easy calculation yields that

$$P\left(\left[\bigcap_{j=1}^{m(i)}\{\rho(Y_{ij}(\omega),X_i(\omega))\le\chi_T(Y_{ij}(\omega))\}\right]\cap\{U(X_i(\omega))\not\subset T\}\right)= \tag{13.23}$$
$$= \; P\left(\left[\bigcap_{j=1}^{m(i)}\{Y_{ij}(\omega)\notin(U(X_i(\omega))-T)\}\right]\cap\{U(X_i(\omega))\not\subset T\}\right)=$$
$$= \; P\left(\left[\bigcap_{j=1}^{m(i)}\{Y_{ij}(\omega)\in S-(U(X_i(\omega))-T)\}\right]\cap\{U(X_i(\omega))\not\subset T\}\right)=$$
$$= \; \sum_{x\in E,U(x)\not\subset T}P\left(\bigcap_{j=1}^{m(i)}\{Y_{ij}(\omega)\in S-(U(x)-T)\}\cap\{X_i(\omega)=x\}\right)=$$
$$= \; \sum_{x\in E,U(x)\not\subset T}\left\{\left[\prod_{j=1}^{m(i)}[1-P(\{Y_{ij}(\omega)\in U(x)-T\})]\right]\right\}P(\{X_i(\omega)=x\})=$$
$$= \; \sum_{x\in E,U(x)\not\subset T}\{[1-P(\{Y_{11}(\omega)\in U(x)-T\})]^{m(i)}\}\,P(\{X_i(\omega)=x\}),$$

as the random variables Y_{ij} are supposed to be statistically independent and identically distributed. Random variables X_i are also identically distributed, so that we can replace X_i by X_1 in (13.21).

As the set E is finite of countable, there exists, given $\varepsilon>0$, a finite subset $E_0(\varepsilon)\subset E$ such that $P(\{X_1(\omega)\in E_0\})>1-(\varepsilon/2)$ holds. If $x\in E_0$ is such that $U(x)\not\subset T$, i.e., such that $U(x)-T\ne\emptyset$, then $P(\{Y_{11}(\omega)\in U(x)-T\})=\sum_{s\in U(x)-T}P(\{Y_{11}(\omega)=s\})>0$ holds due to the conditions imposed to random variables Y_{ij}. Hence, the value

$$m^*(x)=\min\{m\in\mathcal{N}^+:(1-P(\{Y_{11}(\omega)\in U(x)-T\}))^m<\varepsilon/2\}<\infty \tag{13.24}$$

is uniquely defined for each $x\in E_0$ such that $U(x)\not\subset T$. Set

$$m^*=\max\{m^*(x):x\in E_0\} \tag{13.25}$$

and define

$$\lambda^{**}=\sum_{x\in E,U(x)\not\subset T}\left\{[1-P(\{Y_{11}(\omega)\in U(x)-T\})]^{m^*}P(\{X_1(\omega)=x\})\right\}. \tag{13.26}$$

Obviously,

$$\lambda^{**}=\sum_{x\in E_0,U(x)\not\subset T}\left\{[1-P(\{Y_{11}(\omega)\in U(x)-T\})]^{m^*}P(\{X_1(\omega)=x\})\right\}+ \tag{13.27}$$
$$+\sum_{x\in E-E_0,U(x)\not\subset T}\left\{[1-P(\{Y_{11}(\omega)\in U(x)-T\})]^{m^*}P(\{X_1(\omega)=x\})\right\}\le$$
$$\le \; \sum_{x\in E_0,U(x)\not\subset T}\left\{[1-P(\{Y_{11}(\omega)\in U(x)-T\})]^{m^*}P(\{X_1(\omega)=x\})\right\}+$$
$$+\sum_{x\in E-E_0}P(\{X_1(\omega)=x\})<$$
$$< \; \sum_{x\in E_0}(\varepsilon/2)P(\{X_1(\omega)=x\})+P(\{X_1(\omega)\in E-E_0\})\le$$
$$\le \; (\varepsilon/2)\,P(\{X_1(\omega)\in E_0\})+(\varepsilon/2)\le\varepsilon.$$

Consequently, setting $m(i) = m^*$ for each $i \in \mathcal{N}^+$, setting

$$\lambda^* = P(\{U(X_1(\omega)) \subset T\}) + \lambda^{**} = bel_{U_{X_1,\rho}}(T) + \lambda^{**}, \tag{13.28}$$

and combining together all the relations from (13.22) to (13.28) we obtain that, for each $i \in \mathcal{N}^+$,

$$\begin{aligned} P(\{\lambda_i(T,\omega) = 1\}) = P(\{U(X_1(\omega)) \subset T\}) + \lambda^{**} = \lambda^* \in \\ \in \langle P(\{U(X_1(\omega)) \subset T\}),\, P(\{U(X_1(\omega)) \subset T\}) + \varepsilon) = \\ = \ \langle bel_{U_{X_1,\rho}}(T), bel_{U_{X_1,\rho}}(T) + \varepsilon). \end{aligned} \tag{13.29}$$

Hence, $\{\lambda_i(T,\cdot)\}_{i=1}^{\infty}$ is a sequence of statistically independent and identically distributed binary-valued random variables so that, due to the strong law of large numbers, the relative frequency of occurrences of unit values taken by $\lambda_i(T,\omega)$, i. e., the value $B_{X,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega)$ according to (13.15), tends, with n increasing, almost surely to the probability λ^* with which each $\lambda_i(T,\cdot)$ takes the unit value. So, the assertion (13.16) is proved. □

Theorem 13.4. Let the notations and conditions of Theorem 13.3 hold, let $\lim_{i\to\infty} m(i) = \infty$. Then the equality

$$P\left(\left\{\omega \in \Omega : \lim_{n\to\infty} B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) = bel_{U_{X_1,\rho}}(T)\right\}\right) = 1 \tag{13.30}$$

holds. □

Proof. Let $m(1), m(2), \ldots \in \mathcal{N}^+$ be such that $\lim_{i\to\infty} m(i) = \infty$. Set, for each $i,\, k \in \mathcal{N}^+$,

$$\tilde{m}^{(k)}(i) = m(k), \text{ if } i \le k,\ \tilde{m}^{(k)}(i) = m(i), \text{ if } i \ge k. \tag{13.31}$$

Define $\tilde{\lambda}_i^{(k)}(T,\omega)$ analogously to (13.28), but with $m(i)$ replaced by $\tilde{m}^{(k)}(i)$, so that

$$\tilde{\lambda}_i^{(k)}(T,\omega) = \prod_{j=1}^{\tilde{m}^{(k)}(i)} \min\left\{1, 1 - \rho(Y_{ij}(\omega), X_i(\omega)) + \chi_T(Y_{ij}(\omega))\right\}. \tag{13.32}$$

Consequently, for $i \ge k$ the equality $\tilde{\lambda}_i^{(k)}(T,\omega) = \lambda_i(T,\omega)$ holds for each $\omega \in \Omega$. The values taken by the random variables $\tilde{\lambda}_i^{(k)}(T,\cdot)$ are either 0 or 1, so that the limit values of the relative frequencies of both the outcomes, if defined, do not depend on any initial segment of a sequence of such values. Hence,

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^n \tilde{\lambda}_i^{(k)}(T,\omega) = \lim_{n\to\infty} n^{-1} \sum_{i=1}^n \lambda_i(T,\omega) \tag{13.33}$$

for each $k \in \mathcal{N}^+$ each and $\omega \in \Omega$ for which one of these limit values is defined. So, setting

$$\tilde{B}_{X_1,\rho}^{(k)}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) = n^{-1} \sum_{i=1}^n \tilde{\lambda}_i^{(k)}(T,\omega) \tag{13.34}$$

for each $k \in \mathcal{N}^+$ and each $\omega \in \Omega$, we obtain that

$$\begin{aligned} & \lim_{n\to\infty} \tilde{B}^{(n)}_{X_1,\rho}\,(T, n, \langle m(i)\rangle_{i=1}^n, \omega) = \\ = \; & \lim_{n\to\infty} \tilde{B}^{(n)}_{X_1,\rho}\left(T, n, \langle \tilde{m}^{(k)}(i)\rangle_{i=1}^n, \omega\right) = \\ = \; & \lim_{n\to\infty} n^{-1} \sum\nolimits_{i=1}^n \tilde{\lambda}_i^{(k)}(T,\omega) = \\ = \; & \lim_{n\to\infty} n^{-1} \sum\nolimits_{i=1}^n \lambda_i(T,\omega) = \\ = \; & \lim_{n\to\infty} B_{X_1,\rho}\,(T, n, \langle m(i)\rangle_{i=1}^n, \omega) = \\ = \; & \lim_{n\to\infty} B_{X_1,\rho}\left(T, n, \langle \tilde{m}^{(k)}(i)\rangle_{i=1}^n, \omega\right) \end{aligned} \tag{13.35}$$

by (13.15) for each $\omega \in \Omega$ for which at least one of the limit values in (13.35) is defined, and for each $k \in \mathcal{N}^+$.

Take arbitrarily $\varepsilon > 0$ and define, for $x \in E$, $m^*_\varepsilon(x)$ by (13.24); define also m^*_ε by (13.25). As $\lim_{k\to\infty} m(k) = \infty$, there exists $k_0 \in \mathcal{N}^+$ such that $m(k) \geq m^*_\varepsilon$ holds for every $k \geq k_0$, hence, $\tilde{m}^{(k)}(j) \geq m^*_\varepsilon$ holds for each such k and each $j \in \mathcal{N}^+$. Theorem 13.3 then yields that, for each $k \geq k_0$, the relation

$$\lim_{n\to\infty} B_{X_1,\rho}\left(T, n, \langle \tilde{m}^{(k)}(i)\rangle_{i=1}^n, \omega\right) \in \langle bel_{U_{X_1,\rho}}(T), bel_{U_{X_1,\rho}}(T) + \varepsilon), \tag{13.36}$$

so, due to (13.35), also the relation

$$\lim_{n\to\infty} B_{X_1,\rho}\,(T, m, \langle m(i)\rangle_{i=1}^n, \omega) \in \langle bel_{U_{X_1,\rho}}(T), bel_{U_{X_1,\rho}}(T) + \varepsilon) \tag{13.37}$$

holds for each $\omega \in \Omega_0 \subset \Omega$, where $\Omega_0 \in \mathcal{A}$ and $P(\Omega_0) = 1$. As $\varepsilon > 0$ was taken arbitrarily, (13.37) immediately implies that

$$P\left(\left\{\omega \in \Omega : \lim_{n\to\infty} B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) = bel_{U_{X_1,\rho}}(T)\right\}\right) = 1, \tag{13.38}$$

and the assertion is proved. □

Let us sketch, very briefly, a Monte-Carlo algorithm based on the limit assertions presented and proved above and offering numerical values which can be taken, in the sense specified by Theorem 13.5 below, as reasonable estimations of the values of the belief function in question. Let $S = \{s_1, s_2, \ldots, s_N\}$ be a nonempty finite state space of an investigated system, let $E = \{e_1, e_2, \ldots, e_K\}$ be a nonempty finite space of empirical or observational values, let $\rho : S \times E \to \{0,1\}$ be a compatibility relation such that $\rho(s_j, e_i) = 1$ iff s_j cannot be eliminated from the set of possible internal states of the investigated system supposing that the empirical value e_i was obtained or observed as the result of some measurements, experiments or observations concerning the system in question and/or the environment where it is situated.

The role of the input of the proposed algorithm plays a finite sequence $Y = \langle x_i, y_i, z_i\rangle_{i=1}^R$ of triples where $x_i \in E$, $y_i \in S^* = \bigcup_{k=0}^\infty S^k$, and $z_i \in \{0,1\}^* = \bigcup_{k=0}^\infty \{0,1\}^k$. Namely, for each $1 \leq i \leq R$, $x_i \in E$ is an empirical value, i.e. $x_i = e_j$ for some $j \leq N$, $y_i = \langle s_{i1}, s_{i2}, \ldots, s_{i_{m(i)}}\rangle$ is a finite nonempty string of elements of the state space S, and $z_i = \langle \delta_{i1}, \delta_{i2}, \ldots, \delta_{i\ell(i)}\rangle$ is a finite binary string such that $\ell(i) = m(i)$ and $\delta_{ij} = \rho(s_{ij}, x_i\rangle$ for each $1 \leq j \leq m$. Informally, x_i is an element of E sampled at random by the random variable X_i, s_i are the elements of the state space S sampled at random by the random variables Y_i and tested, whether there are compatible with x_i or not, and the results of these tests are encoded by z_i. Evidently, x_i maybe be the same for different triples and also the possibility that two or more triples in Y are identical is not excluded.

Set, for each $1 \leq i \leq R$ and for $\langle x_i, y_i, z_i \rangle$, $y_i = \langle s_{i1}, s_{i2}, \ldots, s_{i_{m(i)}} \rangle$, $z_i = \langle \delta_{i1}, \delta_{i2}, \ldots, \delta_{im_{(i)}} \rangle$,

$$u(x_i) = \{s_{ij} : 1 \leq j \leq m(i),\, \delta_{ij} = 1\}, \tag{13.39}$$

set also, for each $1 \leq j \leq K$,

$$\mathcal{U}(e_j) = \bigcup_{1 \leq k \leq R,\, x_k = e_j} u(x_k). \tag{13.40}$$

The set $\mathcal{U}(e_j)$ approximates, in the statistical sense, the set $U_\rho(e_j)$ of states compatible with the empirical value $e_j \in E$. Setting, moreover,

$$p(e_j) = R^{-1}\text{card}\{k : 1 \leq k \leq R,\, x_k = e_j\}, \tag{13.41}$$

we can see that $p(e_j)$ approximates the probability with which $X_i(\omega) = e_j$. Combining both those approximations together, we can set

$$bel^0(T) = \sum_{1 \leq j \leq K,\, \mathcal{U}(e_j) \subset T} p(e_j) \tag{13.42}$$

and we obtain that $bel^0(T)$ is a reasonable statistical estimation of the value $bel_{U_{X_1,\rho}}(T)$. The quality of this estimation is stated by the following assertion

Theorem 13.5. Let the notations and conditions of Theorem 13.3 hold, let $\varepsilon > 0$ be such that the inequality $0 < \varepsilon < E\,\lambda_i(T, \cdot) - bel_{U_{X_1,\rho}}(T)$ holds for a given subset $T \subset S$, here $\lambda_i(T, \cdot)$ is defined by (13.10). Then the inequality

$$\begin{aligned} & P\left(\left\{\omega \in \Omega : |B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) - bel_{U_{X_1,\rho}}(T)| \geq \varepsilon\right\}\right) \leq \\ \leq \; & 1/4n(\varepsilon - (E\lambda_i(T, \cdot) - bel_{U_{X_1,\rho}}(T)))^2 \end{aligned} \tag{13.43}$$

holds. □

Proof. Due to the Chebyshev inequality for binary random variables we obtain that

$$P\left(\{\omega \in \Omega : |B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) - E\,\lambda_i(T, \cdot)| \geq \varepsilon\}\right) \leq \frac{1}{4n\varepsilon^2} \tag{13.44}$$

holds due to Theorem 13.3. Evidently, if

$$\left|B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) - bel_{U_{X_1,\rho}}(T)\right| \geq \varepsilon \tag{13.45}$$

holds, then

$$\begin{aligned} & |B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) - E\,\lambda_i(T, \cdot)| \geq \\ \geq \; & \varepsilon - \left|E\,\lambda_i(T, \cdot) - bel_{U_{X_1,\rho}}(T)\right| \end{aligned} \tag{13.46}$$

must hold as well. So

$$\begin{aligned} & P\left(\left\{\omega \in \Omega : |B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) - bel_{U_{X_1,\rho}}(T)| \geq \varepsilon\right\}\right) \leq \\ \leq\ & P\Big(\Big\{\omega \in \Omega : |B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) - E\,\lambda_i(T, \cdot)| \geq \\ & \quad \varepsilon - \Big(E\,\lambda_i(T, \cdot) - bel_{U_{1,\rho}}(T)\Big)\Big\}\Big) \leq \\ \leq\ & 1/4n(\varepsilon - (E\,\lambda_i(T, \cdot) - bel_{U_{X_n,\rho}}(T)))^2 \end{aligned} \tag{13.47}$$

follows immediately from (13.41) when replacing ε by $\varepsilon - |E\,\lambda_i(T, \cdot) - bel_{U_{X_1,\rho}}(T)|$ and taking into consideration that $E\,\lambda_i(T, \cdot) \geq bel_{U_{X_1,\rho}}(T)$ always holds (cf. the proof of Theorem 13.3). □

The dependence of the upper bound on the right-hand side of (13.43) on the value $bel_{U_{X_1,\rho}}(T)$ can be eliminated, using Theorem 13.3, when choosing $m(i)$ uniformly large enough.

Theorem 13.6. Let the notations and conditions of Theorem 13.3 hold, let $\varepsilon > 0$ be given. Then there exists $m^*(\varepsilon) \in \mathcal{N}^+$ such that, if $m(i) \geq m^*(\varepsilon)$ holds for each $i \leq n$, then

$$P\left(\left\{\omega \in \Omega : |B_{X_1,\rho}(T, n, \langle m(i)\rangle_{i=1}^n, \omega) - bel_{U_{X_1,\rho}}(T)| \geq \varepsilon\right\}\right) \leq 1/n\varepsilon^2 \tag{13.48}$$

holds. □

Proof. Theorem 13.3 yields that there exists $m^*(\varepsilon)$ such that

$$\left|E\,\lambda_i(T, \cdot) - bel_{U_{X_1,\rho}}(T)\right| \leq \varepsilon/2 \tag{13.49}$$

holds supposing that $m(i) \geq m^*(\varepsilon)$ holds for each $i \leq n$. Combining (13.43) and (13.49) we obtain immediately that (13.48) is valid. □

14 Boolean–Valued and Boolean–Like Processed Real–Valued Belief Functions

The reasons for which it may seem useful to reconsider the Dempster–Shafer model of uncertainty quantification and processing from the point of view of possible non-numerical quantification of occurring uncertainty degrees can be divided into two groups: why to refuse the numerical real-valued degrees, and why to choose just this or that set of values and structure over this set as an adequate alternative to the original numerical evaluation. First, there are some general arguments in favour of the claim that structures over sets of abstract objects of non-numerical nature can be sometimes more close to the spaces of uncertain events and structures over them than the space of real numbers with all the riches of notions, relations and operations over these numbers (overspecification of the degrees of uncertainty by real numbers, these degrees need not be dichotomic, a danger of an ontological shift from structures over real numbers to structures over uncertainties, and so on). A more detailed discussion in this direction can be found in [8] and [37], as far as fuzzy sets are concerned, in [5] for set-valued probability measures, and in [26], [27] for applications of such probabilities in uncertain data processing expert (knowledge) systems; we shall not repeat this discussion here and refer to these sources. The reason for our particular alternative choice taken in this chapter is twofold: boolean algebras, which generalize the set-theoretical structures over the power-set (set of all subsets) of a fixed basic space, are perhaps the most developed non-numerical abstract mathematical structures. Moreover, just because of their just mentioned near relation to the set-theoretical structures boolean-valued uncertainty degrees in D.-S. theory seem to be easy compatible with the set-theoretical operations (e. g., joints) and relations (e. g., inclusions), often occurring above when defining belief functions and other numerical characteristics typical for the D.-S. theory.

Definition 14.1. *Boolean algebra* is a quadruple $\mathcal{B} = \langle B, \vee, \wedge, \neg \rangle$, where B is a nonempty set called the *support* of $\mathcal{B}$, $\vee$ is a binary operation taking $B \times B$ into B and called *supremum*, $\wedge$ is a binary operation taking $B \times B$ into B and called *infimum*, and $\neg$ is a unary operation taking B into B and called *complement*; these operations are supposed to satisfy, for each x, y, $z \in B$, the following five axioms (cf. [42])

(A1) $x \vee y = y \vee z, \ x \wedge y = y \wedge x,$

(A2) $x \vee (y \vee z) = (x \vee y) \vee z, \ x \wedge (y \wedge z) = (x \wedge y) \wedge z,$

(A3) $(x \wedge y) \vee y = y, \ (x \vee y) \wedge y = y,$

(A4) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \ x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z),$

(A5) $(x \wedge (\neg x)) \vee y = y, \ (x \vee (\neg x)) \wedge y = y,$

where $=$ denotes the identity relation on B. The partial ordering relation $\prec_{\mathcal{B}}$ (or simply $\prec$, if no misunderstanding menaces) is defined by $x \prec y \Leftrightarrow_{\mathrm{df}} x \vee y = y$ or, what turns to be the same, $x \prec y \leftrightarrow_{\mathrm{df}} x \wedge y = x$. As can be easily proved, there exists just one element $\mathbf{0}_{\mathcal{B}} \in B$ such that $\mathbf{0}_{\mathcal{B}} \prec x$ holds for each $x \in B$, this element is called the *zero* or the *minimum* element of the Boolean algebra $\mathcal{B}$. Dually, there exists just one element $\mathbf{1}_{\mathcal{B}} \in B$ such that $x \prec \mathbf{1}_{\mathcal{B}}$ holds for each $x \in B$, this element is called the *unit* or the *maximum* element of $\mathcal{B}$, also in these cases the indices $\mathcal{B}$ are omitted if no misunderstanding menaces. The operations $\vee$ ($\wedge$, resp.) can be easily proved to possess all the properties of supremum (infimum, resp.) operation with respect to the partial ordering relation $\prec_{\mathcal{B}}$. The Boolean algebra $\mathcal{B}$ is called *nontrivial*, if $\mathrm{card}(B) \geq 2$; if this is the case, then $\mathbf{0}_{\mathcal{B}} \neq \mathbf{1}_{\mathcal{B}}$ and only such Boolean algebras will be considered in what follows. □

We refer to [10] and [42] as far as some elementary facts concerning Boolean algebras and used below are concerned.

Definition 14.2. Let Ω be a nonempty set, let $\mathcal{A}$ be a nonempty σ-field of subsets of Ω, let $\mathcal{B} = \langle B, \vee, \wedge, \neg \rangle$ be a nontrivial Boolean algebra. *Conditional boolean-valued probability measure* (c.b.v.p.m.) defined on the measurable space $\langle \Omega, \mathcal{A} \rangle$ and taking its values in $\mathcal{B}$ is a mapping $P : \mathcal{A} \rightarrow B \times B$ such that, setting for each $A \in \mathcal{A}$, $P(A) = \langle P_1(A), P_2(A) \rangle$, $P_i(A) \in B$ for both $i = 1, 2$, the following conditions hold:

$$\begin{array}{ll} \text{(i)} & P_1(A) \prec P_2(A) \text{ for each } A \in \mathcal{A}, \\ \text{(ii)} & \text{if } A = \Omega, \text{ then } P_1(A) = P_2(A) \end{array} \tag{14.1}$$

(iii) $P(\Omega - A) = \langle \neg P_1(A), P_2(A) \rangle$ for each $A \in \mathcal{A}$,

(iv) if $\mathcal{A}_0 = \{A_1, A_2, \ldots\} \subset \mathcal{A}$, and if $P(A_i) = \langle P_1(A_i), P_2(A_i) \rangle$ with the same $P_2(A_i) = C \in B$ independent of i, then $\bigvee_{i=1}^{\infty} P_1(A_i)$ is defined and $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \left\langle \bigvee_{i=1}^{\infty} P_1(A_i), C \right\rangle$.

□

Definition 14.3. Let P be a c.b.v.p.m. over $\langle \Omega, \mathcal{A} \rangle$ taking its values in the Boolean algebra $\mathcal{B}$, let $A, C \in \mathcal{A}$ be two measurable sets (random events), then the conditional boolean-valued probability of A given C (or: under the condition that C holds) generated by P will be denoted by $P(A/C)$ and defined by

$$P(A/C) = \langle P_1(A) \wedge P_1(C),\ P_2(A) \wedge P_1(C) \rangle. \tag{14.2}$$

□

Definition 14.4. C.b.v.p.m. P over Ω, $\mathcal{A}$, and $\mathcal{B}$ is called *regular*, if $P_2(A) \neq \mathbf{0}_{\mathcal{B}}$ is the same for all $A \in \mathcal{A}$; P is called *unconditional, a priori* or *simply boolean-valued probability measure*, if $P_2(A) = \mathbf{0}_{\mathcal{B}}$ for each $A \in \mathcal{A}$. C.b.v.p.m. P over Ω, $\mathcal{A}$, and $\mathcal{B}$ is called *complete*, if $\mathcal{A} = \mathcal{P}(\Omega)$ and if (iv) of Definition 14.2 holds for each $\emptyset \neq \mathcal{A}_0 \subset \mathcal{P}(\Omega)$ $(= \{A : A \subset \Omega\})$.

□

Lemma 14.1. Let P be a.c.b.v.p.m. over Ω, $\mathcal{A}$, and $\mathcal{B}$, let $\mathcal{A}_0 = \{A_1, A_2, \ldots\}$ satisfy the conditions of (iv), Definition 14.2. Then $\bigwedge_{i=1}^{\infty} P_1(A_i)$ is defined and $P\left(\bigcap_{i=1}^{\infty} A_i\right) = \langle \bigwedge_{i=1}^{\infty} P_1(A_i), C \rangle$. □

Proof. As $A_i \in \mathcal{A}$ for each $i = 1, 2, \ldots$, then $\bigcap_{i=1}^{\infty} A_i = \Omega - \bigcup_{i=1}^{\infty}(\Omega - A_i) \in \mathcal{A}$ as well, hence, $P\left(\bigcap_{i=1}^{\infty} A_i\right)$ is defined and

$$\begin{aligned} & P\left(\textstyle\bigcap_{i=1}^{\infty} A_i\right) = \left\langle P_1\left(\textstyle\bigcap_{i=1}^{\infty} A_i\right), P_2\left(\textstyle\bigcap_{i=1}^{\infty} A_i\right)\right\rangle = \\ = \ & \left\langle P_1\left(\Omega - \textstyle\bigcup_{i=1}^{\infty}(\Omega - A_i)\right), P_2\left(\Omega - \textstyle\bigcup_{i=1}^{\infty}(\Omega - A_i)\right)\right\rangle = \\ = \ & \left\langle P_1\left(\Omega - \textstyle\bigcup_{i=1}^{\infty}(\Omega - A_i)\right), C\right\rangle = \left\langle \neg P_1\left(\textstyle\bigcup_{i=1}^{\infty}(\Omega - A_i)\right), C\right\rangle = \\ = \ & \langle \neg \textstyle\bigvee_{i=1}^{\infty} P_1(\Omega - A_i), C \rangle = \langle \neg \textstyle\bigvee_{i=1}^{\infty}(\neg P_1(\Omega - A_i)), C \rangle = \\ = \ & \left\langle \textstyle\bigwedge_{i=1}^{\infty} P_1(A_i), C \right\rangle, \end{aligned} \tag{14.3}$$

using appropriate points of Definition 14.2 in particular steps. □

Definition 14.5. Let $x \in B$, let $\emptyset \neq C_0 \subset B$. C_0 is called a *decomposition* of x, if for each $y, z \in C_0$, $y \neq z$ implies that $y \wedge z = \mathbf{0}_{\mathcal{B}}$ and $\bigvee_{y \in C_0} y = x$. The set of all decompositions of x will be denoted by $Dcp(x)$, obviously, $Dcp(x) \subset \mathcal{P}(\mathcal{P}(B))$. A decomposition C of x is called *strict*, if $\mathbf{0}_{\mathcal{B}} \notin C$, only $x \neq \mathbf{0}_{\mathcal{B}}$ possesses strict decomposition(s). Let $\langle \Omega, \mathcal{A} \rangle$ be a measurable space such that $\{\omega\} \in \mathcal{A}$ holds for each $\omega \in \Omega$, let P be a complete a priori b.v.p.m. over Ω, $\mathcal{A}$ and $\mathcal{B}$. Then the system $\{P_1(\{\omega\}) : \omega \in \Omega\}$ of elements of B is called *a priori boolean-valued probability distribution* (a priori b.v.p.d.) over Ω defined by the a priori b.v.p.m. P. □

Lemma 14.2. If P is a complete a priori b.v.p.m. over Ω, $\mathcal{A}$ and $\mathcal{B}$ such that $\{\omega\} \in \mathcal{A}$ holds for each $\omega \in \Omega$, then $\{P_1(\{\omega\}) : \omega \in \Omega\} \in Dcp(\mathbf{1}_{\mathcal{B}})$. □

Proof. Let $A_\omega = \{\omega\}$ for each $\omega \in \Omega$, so that $\bigcup_{\omega\in\Omega} A_\omega = \Omega$. As P is complete, (ii) and the extended version of (iv) in Definition 14.2 yield that

$$\begin{aligned} & P\left(\textstyle\bigcup_{\omega\in\Omega} A_\omega\right) = P(\Omega) = \langle P_1(\Omega), P_2(\Omega)\rangle = \langle \mathbf{1}_{\mathcal{B}}, \mathbf{1}_{\mathcal{B}}\rangle = \\ = \; & \langle \textstyle\bigvee_{\omega\in\Omega} P_1(A_\omega), \mathbf{1}_{\mathcal{B}}\rangle \end{aligned} \tag{14.4}$$

so that

$$\textstyle\bigvee_{\omega\in\Omega} P_1(A_\omega) = \bigvee_{\omega\in\Omega} P_1(\{\omega\}) = \mathbf{1}_{\mathcal{B}}. \tag{14.5}$$

Let $\omega_1, \omega_2 \in \Omega$, $\omega_1 \neq \omega_2$. Then, for $\mathcal{A} = \{A_1, A_2, \ldots\} \subset \mathcal{A}$ such that $A_1 = \{\omega_1\}$, $A_2 = \{\omega_2\}$, $A_k = \Omega$ for every $k \geq 3$, Lemma 14.1 yields that

$$\begin{aligned} & P_1\left(\textstyle\bigcap_{i=1}^\infty A_i\right) = P_1(\emptyset) = \neg P_1(\Omega) = \mathbf{0}_{\mathcal{B}} = \textstyle\bigwedge_{i=1}^\infty P_1(A_i) = \\ = \; & P_1(\{\omega_1\}) \wedge P_1(\{\omega_2\}) \wedge P_1(\Omega) = P_1(\{\omega_1\}) \wedge P_1(\{\omega_2\}) \wedge \mathbf{1}_{\mathcal{B}} = \\ = \; & P_1(\{\omega_1\}) \wedge P_1(\{\omega_2\}), \end{aligned} \tag{14.6}$$

hence, $\{P_1(\{\omega\}) : \omega \in \Omega\} \in Dcp(\mathbf{1}_{\mathcal{B}})$. □

Let S be a finite set, let P be an a priori b.v.p.m. defined on the measurable space $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S))\rangle$, let $\{P_1(E) : E \subset S\}$ be the a priori b.v.p.d. defined by P and such that $P_1(\emptyset) = \mathbf{0}_{\mathcal{B}}$. Then the *boolean-valued belief function* $bel^{\mathcal{B}}(T)$ and *b.v. plausibility function* $pl^{\mathcal{B}}(T)$ can be defined, in an abstract way following the pattern of algebraic definition of numerical belief and plausibility functions, for each $T \subset S$ as follows:

$$bel^{\mathcal{B}}(T) = \langle \textstyle\bigvee_{A\subset T} P_1(A), \mathbf{1}_{\mathcal{B}}\rangle \tag{14.7}$$

$$pl^{\mathcal{B}}(T) = \left\langle \textstyle\bigvee_{A, A\cap T\neq\emptyset} P_1(A), \mathbf{1}_{\mathcal{B}}\right\rangle. \tag{14.8}$$

If $P_1(A) \neq \mathbf{0}_{\mathcal{B}}$, (14.7) and (14.8) are replaced by

$$bel^{\mathcal{B}}(T) = \left\langle \textstyle\bigvee_{\emptyset\neq A\subset T} P_1(A), \bigvee_{\emptyset\neq A\subset S} P_1(A)\right\rangle, \tag{14.9}$$

$$pl^{\mathcal{B}}(T) = \left\langle \textstyle\bigvee_{A, A\cap T\neq\emptyset} P_1(A), \bigvee_{\emptyset\neq A\subset S} P_1(A)\right\rangle. \tag{14.10}$$

Let U be a measurable mapping defined on an abstract measurable space $\langle \Omega, \mathcal{A}\rangle$ and taking its values in the measurable space $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S))\rangle$, let P^* be an a priori b.v.p.m. on $\langle \Omega, \mathcal{A}\rangle$ such that, for each $A \subset S$

$$P_1^*(\{\omega \in \Omega : U(\omega) = A\}) = P_1(A). \tag{14.11}$$

Then

$$\begin{aligned} & \textstyle\bigvee_{\emptyset\neq A\subset T} P_1(A) = \bigvee_{\emptyset\neq A\subset T} P_1(\{\omega \in \Omega : U(\omega) = A\}) = \\ = \; & P_1^*\left(\textstyle\bigcup_{\emptyset\neq A\subset T} \{\omega \in \Omega : U(\omega) = A\}\right) = P_1^*(\{\omega \in \Omega : \emptyset \neq U(\omega) \subset T\}). \end{aligned} \tag{14.12}$$

In the same way,

$$\bigvee_{\emptyset \neq A \subset S} P_1(A) = P(\{\omega \in \Omega : U(\omega) \neq \emptyset\}). \tag{14.13}$$

So, (14.9) implies that

$$\begin{aligned} bel^{\mathcal{B}}(T) &= \langle P_1^*(\{\omega \in \Omega : \emptyset \neq U(\omega) \subset T\}), P_2^*(\{\omega \in \Omega : \emptyset \neq U(\omega) \subset T\})\rangle = \\ &= P^*(\{\omega \in \Omega : \emptyset \neq U(\omega) \subset T\} / \{\omega \in \Omega : U(\omega) \neq \emptyset\}) \end{aligned} \tag{14.14}$$

according to the definition of conditional boolean-valued probability given by Definition 14.3. Let us recall that P^* is an a priori b.v.p.m., so that $P_2^*(\{\emptyset \neq U(\omega) \subset T\}) = \mathbf{1}_{\mathcal{B}}$. An analogous relation can be easily proved also for $pl^{\mathcal{B}}$, consequently, also boolean-valued belief and plausibility functions can be equivalently defined, using the set-valued random variable U, in a way similar to that in the case of numerical belief functions.

Let us investigate, now, a variant of Dempster combination rule fitted for boolean-valued belief functions. First of all, let us briefly return to the numerical belief functions as developed above. Let bel_1, bel_2 be numerical belief functions. A mapping (or: combination rule) $\oplus$ ascribing the belief function $bel_1 \oplus bel_2$, i.e., the value $(bel_1 \oplus bel_2)(T) \in \langle 0, 1\rangle$ to each $T \subset S$, is called *extensional*, if there exists a function $F : \langle 0,1\rangle \times \langle 0,1\rangle \to \langle 0,1\rangle$ such that

$$(bel_1 \oplus bel_2)(T) = F(bel_1(T), bel_2(T)) \tag{14.15}$$

for each bel_1, bel_2 and each $T \subset S$. The validity of an analogous relation for $pl_1 \oplus pl_2$ (with a different function G instead of F, of course) then immediately follows supposing that $pl_1 \oplus pl_2$ is the plausibility function corresponding to $bel_1 \oplus bel_2$. Neither the Dempster combination rule defined in the abstract algebraic way, nor its alternative definition based on the set-valued random variable $U_1(\omega) \cap U_2(\omega)$, are extensional in this sense, as can be easily seen.

Let m_1, m_2 be two (numerical) b.p.a.'s such that $m_1(\emptyset) = m_2(\emptyset) = 0$, let m_3 be a mapping which takes $\mathcal{P}(S)$ into $\langle 0,1\rangle$ such a way that

$$m_3(E) = \frac{\sum_{A,B \subset S, A \cap B = E} m_1(A)\, m_2(B)}{\sum_{A,B \subset S, A \cap B \neq \emptyset} m_1(A)\, m_2(B)} \tag{14.16}$$

supposing that this value is defined, i.e., supposing that there exist A, $B \subset S$ such that $m(A) > 0$, $m(B) > 0$ and $A \cap B \neq \emptyset$ hold. Then m_3 is also b.p.a. on S and Dempster combination rule reduces to

$$(bel_1 \oplus bel_2)(T) = \sum_{\emptyset \neq E \subset T} m_3(E). \tag{14.17}$$

Hence, the operation $\oplus$ is *weakly extensional* or *quasi-extensional* in the sense that $bel_1 \oplus bel_2$ is defined by a b.p.a. m_3 which is given as an extensional function $h(m_1, m_2)$ of the b.p.a.'s m_1, m_2 defining bel_1 and bel_2.

In order to arrive at the boolean-valued case of the Dempster combination rule, let us consider the case when S is finite, U_i, $i = 1, 2$, are random variables defined on the probability space $\langle \Omega, \mathcal{A}, P\rangle$ and taking their values in $\mathcal{P}(S)$, let bel_i, $i = 1, 2$, be the corresponding belief functions. Then

$$\begin{aligned} & (bel_1 \oplus bel_2)(T) = \\ = \;& P\left(\{\omega \in \Omega : U_1(\omega) \cap U_2(\omega) \subset T\} / \{\omega \in \Omega : U_1(\omega) \cap U_2(\omega) \neq \emptyset\}\right). \end{aligned} \tag{14.18}$$

If P is a numerical probability measure, then (14.18) can be converted into the combinatoric case (14.16) only under some further conditions including the statistical independence of the set-valued random variables U_1, U_2 and analyzed, in more detail, in Section 6 above. If P is an a priori b.v.p.m., we obtain that, for each $T \subset S$

$$\begin{aligned} & (bel_1^{\mathcal{B}} \oplus bel_2^{\mathcal{B}})(T) = \\ = \;& \langle P_1(\{\omega \in \Omega : \emptyset \neq U_1(\omega) \cap U_2(\omega) \subset T\}),\, P_2(\{\omega \in \Omega : \emptyset \neq U_1(\omega) \cap U_2(\omega)\})\rangle . \end{aligned} \tag{14.19}$$

However, due to the extensionality of boolean-valued probability measures we obtain that

$$\begin{aligned} & P(\{\omega \in \Omega : \emptyset \neq U_1(\omega) \cap U_2(\omega) \subset T\}) = \\ = \;& \langle P_1(\{\omega \in \Omega : \emptyset \neq U_1(\omega) \cap U_2(\omega) \subset T\}), \mathbf{1}_{\mathcal{B}}\rangle = \\ = \;& \Big\langle P_1\Big(\textstyle\bigcup_{\emptyset \neq E \subset T} \bigcup_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), A\cap B = E}(\{\omega \in \Omega : U_1(\omega) = A\} \cap \\ & \qquad \cap \{\omega \in \Omega : U_2(\omega) = B\})\Big), \mathbf{1}_{\mathcal{B}}\Big\rangle = \\ = \;& \Big\langle P_1\Big(\textstyle\bigcup_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), A\cap B \subset T, A\cap B \neq \emptyset}(\{\omega \in \Omega : U_1(\omega) = A\} \cap \\ & \qquad \cap \{\omega \in \Omega : U_2(\omega) = B\})\Big), \mathbf{1}_{\mathcal{B}}\Big\rangle = \\ = \;& \Big\langle \textstyle\bigvee_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), \emptyset \neq A\cap B \subset T}(P_1(\{\omega \in \Omega : U_1(\omega) = A\}) \wedge \\ & \qquad \wedge P_1(\{\omega \in \Omega : U_2(\omega) = B\})), \mathbf{1}_{\mathcal{B}}\Big\rangle . \end{aligned} \tag{14.20}$$

In an analogous way we obtain that

$$\begin{aligned} & P(\{\omega \in \Omega : U_1(\omega) \cap U_2(\omega) \neq \emptyset\}) = \\ = \;& \Big\langle \textstyle\bigvee_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), \emptyset \neq A\cap B}(P_1(\{\omega \in \Omega : U_1(\omega) = A\}) \wedge \\ & \qquad \wedge P_1(\{\omega \in \Omega : U_2(\omega) = B\})), \mathbf{1}_{\mathcal{B}}\Big\rangle , \end{aligned} \tag{14.21}$$

so that

$$\begin{aligned} & (bel_1^{\mathcal{B}}) \oplus bel_2^{\mathcal{B}}(T) = \\ = \;& \Big\langle \textstyle\bigvee_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), \emptyset \neq A\cap B \subset T}(P_1(\{\omega \in \Omega : U_1(\omega) = A\}) \wedge \\ & \qquad \wedge P_1(\{\omega \in \Omega : U_2(\omega) = B\})), \\ & \textstyle\bigvee_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), \emptyset \neq A\cap B}(P_1(\{\omega \in \Omega : U_1(\omega) = A\}) \wedge \\ & \qquad \wedge P_1(\{\omega \in \Omega : U_2(\omega) = B\}))\Big\rangle . \end{aligned} \tag{14.22}$$

Hence, setting, for $i = 1, 2$,

$$m_i^{\mathcal{B}} = \{\langle P_1(\{\omega \in \Omega : U_i(\omega)\} = A),\, P_1(\{\omega \in \Omega : U_i(\omega) \neq \emptyset\})\rangle\}, \; \emptyset \neq A \subset S, \tag{14.23}$$

as two boolean-valued basic probability assignments defining the boolean-valued belief functions $bel_i^{\mathcal{B}}$, $i = 1, 2$, and considering the system

$$\begin{aligned} m_3^{\mathcal{B}} \;=\; & \Big\langle \bigvee\nolimits_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), A\cap B = E, A\cap B \neq \emptyset} (P_1(\{\omega \in \Omega : U_1(\omega) = A\})\wedge \\ & \quad \wedge P_1(\{\omega \in \Omega : U_2(\omega) = B\}))\;, \\ & \bigvee\nolimits_{\langle A,B\rangle \in \mathcal{P}(S)\times\mathcal{P}(S), A\cap B \neq \emptyset} (P_1(\{\omega \in \Omega : U_1(\omega) = A\})\wedge \\ & \quad \wedge P_1(\{\omega \in \Omega : U_2(\omega) = B\}))\rangle_{\emptyset \neq E \subset S}\;, \end{aligned} \tag{14.24}$$

we can easily observe that m_3 is also a boolean-valued b.p.a. which defines $bel_1^{\mathcal{B}} \oplus bel_2^{\mathcal{B}}$. Hence, for boolean-valued belief functions the combination rule $\oplus$ is weakly extensional or quasi-extensional as in the case of the abstract combinatoric definition of the Dempster combination rule for numerical b.p.a.'s and the corresponding belief functions. The fact that in the boolean-valued case no supplementary conditions of statistical independence imposed on the set-valued random variables U_1 and U_2 are necessary, is a trivial consequence of the extensional nature of boolean-valued probability measures with respect to the set-theoretic operations over random events.

Let us close this chapter by a short and rather sketched reasoning which shows that the weak or quasi-extensionality of the combination rule for belief functions is achievable, without any conditions of statistical independence of the random variables U_1 and U_2 also for certain numerical-valued probability measures supposing that they are defined in a rather nonstandard way conserving the extensionality of boolean operations over a particular Boolean algebra. Let us focus our attention to the three following particular Boolean algebras which are obviously isomorphic with each other.

Let $\mathcal{N}^+ = \{1, 2, \ldots\}$ be the set of all (standard) positive integers, let $\mathcal{P}(\mathcal{N}^+)$ be the power-set of all subsets of $\mathcal{N}^+$, let $\cup$, $\cap$ and $\mathcal{N}^+ - \cdot$ be the set-theoretic operations of union, intersection and complement. Then the quadruple $\mathcal{B}_0 = \langle \mathcal{P}(\mathcal{N}^+), \cup, \cap, \mathcal{N}^+ - \cdot\rangle$ is evidently a complete Boolean algebra with the empty subset $\emptyset$ of $\mathcal{N}^+$ as the zero element $\mathbf{0}_{\mathcal{B}_0}$, the set $\mathcal{N}^+$ as the unit element $\mathbf{1}_{\mathcal{B}_0}$, and with the set-theoretic relation of inclusion $\subset$ playing the role of the partial ordering $\prec_{\mathcal{B}_0}$.

Let $B_1 = \{0, 1\}^\infty$ be the space of all infinite binary sequences, let $\boldsymbol{x} = \langle x_1, x_2, \ldots\rangle$ or $\boldsymbol{x} = \langle x_i\rangle_{i=1}^\infty$, $x_i \in \{0, 1\}$ for all $i \in \mathcal{N}^+$, denote an element of B_1 (and similarly for $\boldsymbol{y}$, $\boldsymbol{z}, \ldots$). Let $0^\infty = \langle 0, 0, 0, \ldots\rangle \in B_1$ and $1^\infty = \langle 1, 1, 1, \ldots\rangle \in B_1$ denote the two constant sequences, let $\vee_1$ and $\wedge_1$ be binary operations taking $B_1 \times B_1$ into B_1 in such a way that $\boldsymbol{x} \vee_1 \boldsymbol{y} = \langle \sup\{x_i, y_i\}\rangle_{i=1}^\infty$ and $\boldsymbol{x}_1 \wedge_1 \boldsymbol{y} = \langle \inf\{x_i, y_i\}\rangle_{i=1}^\infty$ for each $\boldsymbol{x}$, $\boldsymbol{y} \in B_1$; here sup and inf are the usual supremum and infimum operations in $\{0, 1\}$, so that $\sup\{x_i, y_i\} = 0$ iff $x_i = y_i = 0$ and $\inf\{x_i, y_i\} = 1$ iff $x_i = y_i = 1$. Let $1^\infty - \cdot$ be the unary operation taking B_1 into B_1 in such a way that $1^\infty - \boldsymbol{x} = \langle 1 - x_i\rangle_{i=1}^\infty$ for all $\boldsymbol{x} \in B_1$. Then the quadruple $\mathcal{B}_1 = \langle \{0, 1\}^\infty, \vee_1, \wedge_1, 1^\infty - \cdot\rangle$ is a complete Boolean algebra with the zero element $\mathbf{1}_{\mathcal{B}_1} = O^\infty$ and the unit element $\mathbf{1}_{\mathcal{B}_1} = 1^\infty$. The Boolean algebras $\mathcal{B}_0$ and $\mathcal{B}_1$ are isomorphic, their isomorphism being established by the $1-1$ mapping $\chi : \mathcal{P}(\mathcal{N}^+) \to \{0, 1\}^\infty$ which ascribes to each $A \subset \mathcal{N}^+$ its characteristic function (sequence, in this particular case) $\chi(A) = \langle \chi(A)_i\rangle_{i=1}^\infty \in \{0, 1\}^\infty$, defined for each $i \in \mathcal{N}^+$ by $\chi(A)_i = 1$, if $i \in A$, $\chi(A)_i = 0$ otherwise.

The third Boolean algebra will be obtained by a particular $1-1$ encoding of sets of positive integers and infinite binary sequences by real numbers from (a certain subset of) the unit interval of (standard) real numbers. Let $\mathcal{C}$ be the well-known Cantor subset of $\langle 0,1\rangle$. Informally, $\mathcal{C}$ is defined by erasing the open interval $(1/3,2/3)$ from $\langle 0,1\rangle$, and by repeated applications of the same operation to the remaining closed intervals. I. e., the open interval $(1/9,2/9)$ is reased from $\langle 0,1/3\rangle$ and $(7/9,8/9)$ from $\langle 2/3,1\rangle$ and so on *ad infinitum*, what remains is just the Cantor set $\mathcal{C}$. Formally, $\mathcal{C}$ is the set of all real numbers from the unit interval for which there exists its ternary decomposition (decomposition to the base 3) which does not contain any occurrence of the numeral 1. It follows immediately that if a decomposition satisfying this property exists, it is defined uniquely (the ternary decomposition $0,\alpha_1,\alpha_2,\ldots,\alpha_n,0,1,1,1,\ldots$, alternative to $0,\alpha_1,\ldots,\alpha_n,2,0,0,0,\ldots$ does not meet the constraint not to contain any occurrence of 1). Hence, the mapping $\varphi_0=\{0,1\}^\infty\rightarrow\mathcal{C}$ ascribing to each $\boldsymbol{x}=\langle x_1,x_2,\ldots\rangle\in\{0,1\}^\infty$ the real number $\sum_{i=1}^\infty 2x_i\,3^{-i}$ is a $1-1$ mapping as well as the composed mapping $\varphi:\mathcal{P}(\mathcal{N}^+)\rightarrow\mathcal{C}$ defined by

$$\varphi(A)=\varphi_0(\chi(A))=\sum\nolimits_{i=1}^\infty 2\chi(A)_i\,3^{-i} \tag{14.25}$$

for each $A\subset\mathcal{N}^+$.

The mapping φ induces binary operations $\vee_2$ and $\wedge_2$, and unary operation $1\dot{-}\cdot$ in $\mathcal{C}$ as follows. Set, for each $\alpha,\ \beta\in\mathcal{C}$,

$$\begin{aligned}\alpha\vee_2\beta &= \varphi\left(\varphi^{-1}(\alpha)\cup\varphi^{-1}(\beta)\right),\\ \alpha\wedge_2\beta &= \varphi\left(\varphi^{-1}(\alpha)\cap\varphi^{-1}(\beta)\right),\\ 1\dot{-}\alpha &= \varphi\left(\mathcal{N}^+-\varphi^{-1}(\alpha)\right),\end{aligned} \tag{14.26}$$

in the last row "$-$" denotes the set-theoretic operation of complement. All the three operations are evidently correctly and unambiguously defined, moreover, an easy calculation yields that $1\dot{-}\alpha=1-\alpha$ holds for each $\alpha\in\mathcal{C}$, where $1-\alpha$ denotes the usual operation of substraction in $\langle 0,1\rangle$. The quadruple $\mathcal{B}_2=\langle\mathcal{C},\vee_2,\wedge_2,1-\cdot\rangle$ is a complete Boolean algebra, $\mathbf{1}_{\mathcal{B}_2}=0$ and $\mathbf{1}_{\mathcal{B}_2}=1$, and $\mathcal{B}_2$ is obviously isomorphic with the Boolean algebras $\mathcal{B}_0$ and $\mathcal{B}_1$ due to the mappings φ_0 and φ_1 defined above.

The following *partial* operation $\sum^*:\mathcal{C}^\infty\rightarrow\mathcal{C}$ ascribing to (some) infinite sequences of real numbers from the Cantor set $\mathcal{C}$ a number from $\mathcal{C}$ will be defined as follows. Let $\langle\alpha_1,\alpha_2,\ldots\rangle$ be a sequence of numbers from $\mathcal{C}$ such that the subsets $\varphi^{-1}(\alpha_i)$ of $\mathcal{N}^+$, $i=1,2,\ldots$, are mutually disjoint. Then $\sum_{i=1}^{\infty*}\alpha_i$ is defined by $\varphi\left(\bigcup_{i=1}^\infty\varphi^{-1}(\alpha_i)\right)$, $\sum_{i=1}^{\infty*}\alpha_i$ being undefined otherwise. As can be easily seen, for each sequence $\langle\alpha_1,\alpha_2,\ldots\rangle\in\mathcal{C}^\infty$ the following implication holds: if $\sum_{i=1}^{\infty*}\alpha_i$ is defined, then $\sum_{i=1}^{\infty*}\alpha_i=\sum_{i=1}^\infty\alpha_i$, where the last expression denotes the usual operation of summation in $\langle 0,1\rangle$. The operation $\sum_{i=1}^{\infty*}$ is commutative in the sense that if $\sum_{i=1}^{\infty*}\alpha_i$ is defined, then $\sum_{i=1}^{\infty*}\alpha_{\pi(i)}$ is also defined and, consequently, equal to $\sum_{i=1}^{\infty*}\alpha_i$, $\sum_{i=1}^\infty\alpha_i$, and $\sum_{i=1}^\infty\alpha_{\pi(i)}$, for each $1-1$ mapping $\pi:\mathcal{N}^+\rightarrow\mathcal{N}^+$. The definition of $\sum_{i=1}^{\infty*}\alpha_i$ can be easily extended, contrary to the classical definition of $\sum_{i=1}^\infty\alpha_i$, to that of $\sum_{\alpha\in D}^*\alpha$ for each nonempty subset $D\subset\mathcal{C}$, but we shall not need and use this generalization in what follows.

It is perhaps worth mentioning explicitly, that the partial operation $\sum_{i=1}^{\infty*}$ can be defined also in an alternative, direct product like way based on the classical (standard) operations over particular items of infinite binary sequences. Namely, given

$\langle \alpha_1, \alpha_2, \ldots \rangle \in \mathcal{C}^\infty$, let $d(\alpha_i) = \langle d_i^1, d_i^2, \ldots \rangle$ be the unique ternary decomposition of α_i not containing any occurrence of 1, let $\overline{d}_i^j = d_i^j/2$ for each $i, j \in \mathcal{N}^+$, so that $\overline{d}(\alpha_i) = \langle \overline{d}_i^1, \overline{d}_i^2, \ldots \rangle$ is an infinite binary sequence. Let $\sum_{i=0}^{\infty 0} \overline{d}(\alpha_i) = \left\langle \sum_{i=1}^{\infty} \overline{d}_i^1, \sum_{i=1}^{\infty} \overline{d}_i^2, \ldots \right\rangle$ be the infinite sequence from $\{\mathcal{N}^+ \cup \{0\} \cup \{\infty\}\}^\infty$ resulting when summing, in the standard sense, the particular items of the corresponding binary sequences $\overline{d}(\alpha_1), \overline{d}(\alpha_2), \ldots$. Now, $\sum_{i=1}^{\infty^*} \alpha_i$ is defined iff $\sum_{i=1}^{\infty 0} \overline{d}(\alpha_i) \in \{0,1\}^\infty$ holds, and in this case

$$\sum\nolimits_{i=1}^{\infty *} \overline{d}(\alpha_i) = \sum_{j=1}^{\infty} 2 \left(\sum\nolimits_{i=1}^{\infty} \overline{d}(\alpha_i^j) \right) 3^{-j}. \tag{14.27}$$

Both the definitions of $\sum_{i=1}^{\infty *}$ can be easily proved to be equivalent.

The basic structure, enabling to formalize, at the most abstract level, the notion of probability and random event, is that of probability space. Let us recall, for the sake of reader's convenience, its usual (standard) definition, immediately followed by its nonstandard modification.

Definition 14.6. (i) Let Ω be a nonempty set, let $\mathcal{A}$ be a σ-field of subsets of Ω, i. e., $\mathcal{A}$ is nonempty and, for each $A, A_1, A_2, \ldots \in \mathcal{A}$ also $\Omega - A \in \mathcal{A}$ and $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ hold. The pair $\langle \Omega, \mathcal{A} \rangle$ is called *measurable space* (generated in Ω or over Ω by the σ-field $\mathcal{A}$) and elements of $\mathcal{A}$ are called *measurable sets.*

(ii) A mapping $P : \mathcal{A} \to \langle 0, 1 \rangle$ ascribing to each $A \in \mathcal{A}$ a real number $P(A)$ from the unit interval of reals is called *(standard) probability measures* (p.m., abbreviately) on $\langle \Omega, \mathcal{A} \rangle$, if (a) $P(\Omega) = 1$ ($\Omega \in \mathcal{A}$ and $\emptyset \in \mathcal{A}$ obviously hold for each σ-field $\mathcal{A} \subset \mathcal{P}(\Omega)$) and (b) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ holds for each sequence $\langle A_1, A_2, \ldots \rangle$ of mutually disjoint sets from $\mathcal{A}$.

(iii) A mapping $\mu : \mathcal{A} \to \mathcal{C}$ (Cantor subset of $\langle 0, 1 \rangle$) is called *non-standard (Cantor-valued probability measure* (n.s.p.m., abbreviately) on $\langle \Omega, \mathcal{A} \rangle$, if (a) $\mu(\Omega) = 1$ and (b) for each sequence $\langle A_1, A_2, \ldots \rangle$ of mutually disjoint sets from $\mathcal{A}$ *the series* $\sum_{i=1}^{\infty *} \mu(A_i)$ *is defined* and $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty *} \mu(A_i)$.

(iv) A triple $\langle \Omega, \mathcal{A}, P \rangle$ ($\langle \Omega, \mathcal{A}, \mu \rangle$, resp.), where $\langle \Omega, \mathcal{A} \rangle$ is a measurable space and P is a probability measure (μ is a nonstandard probability measure, resp.) on $\langle \Omega, \mathcal{A} \rangle$ is called *(standard) probability space* (*nonstandard* or *ns-probability space*, resp.). In both the cases, measurable sets, i. e., elements of $\mathcal{A}$, are called *random events.* For each $A \in \mathcal{A}$, the value $P(A)$ ($\mu(A)$, resp.) is called the *probability (nonstandard* or *ns-probability*, resp.) of the random event A. □

It follows immediately from what we told above, that if $\sum_{i=1}^{\infty *} \mu(A_i)$ is defined, then $\sum_{i=1}^{\infty *} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A_i)$, hence, *every ns-probability measure on* $\langle \Omega, \mathcal{A} \rangle$ *is a (special case of) standard probability measure on the same measurable space.* Consequently, perhaps the adjective "strong" or "special" probability measure would better express the relation between the usual operations over the space of values of probability measures, i. e., over the unit interval $\langle 0, 1 \rangle$, and their alternatives introduced above in this chapter. On the other side, not every standard probability measure is also a nonstandard one. Take, e. g., the probability measure defining a regular coin tossing, where

$P(\text{HEAD}) = P(\text{TAIL}) = 1/2$; as $1/2 \notin \mathcal{C}$, P is obviously not nonstandard, in general, this is the case for every probability measure P which takes a value $P(A) \notin \mathcal{C}$ for at least one random event A.

The formal difference between (ii) and (iii) in Definition 14.6, consisting in the fact that the condition of convergence of the series $\sum_{i=1}^{\infty *} \mu(A_i)$ is explicitly requested to be satisfied, in spite of the case when standard probability measure P is defined, could be avoided by joining the condition that $\sum_{i=1}^{\infty} P(A_i)$ is defined to (ii), Definition 14.6. As a matter of fact, if the operation $\sum_{i=1}^{\infty}$ is taken in its usual sense and as an operation over the extended real line including both the infinite values $+\infty$ and $-\infty$, then any series of nonnegative real numbers from the unit interval always converge. However, if $\sum_{i=1}^{\infty}$ is defined as a *partial* operation on the non-extended real line $(-\infty, \infty)$, i. e., not admitting the infinite values, or even as a *partial* operation taking the space $\langle 0, 1\rangle^{\infty}$ of infinite sequences of real numbers from the unit interval into $\langle 0, 1\rangle$, the condition of convergence of the series $\sum_{i=1}^{\infty} P(A_i)$ would become non-trivial and its satisfaction must be explicitly demanded. In the case of nonstandard probability measure the demand that $\sum_{i=1}^{\infty *} \mu(A_i)$ is defined is obviously non-trivial and, as will be seen below, it will play the key role in what follows.

In our context, the most important property of nonstandard probability measures consists in the fact that they are *extensional* in the sense that nonstandard probabilities of random events combined from some "elementary" random events by the set-theoretic operations of union, intersection and complement can be defined and computed as real-valued (vector) functions of the nonstandard probabilities of these "elementary" random events. The corresponding formalized statement reads as follows.

Theorem 14.1. Let $\langle \Omega, \mathcal{A}, \mu\rangle$ be a nonstandard probability space. Then, for all A, $B \in \mathcal{A}$,

$$\begin{aligned} \mu(\Omega - A) &= 1 - \mu(A),\ \mu(A \cup B) = \\ &= \mu(A) \vee_2 \mu(B),\ \mu(A \cap B) = \mu(A) \wedge_2 \mu(B), \end{aligned} \tag{14.28}$$

where $\vee_2$ and $\wedge_2$ are the binary operations taking $\mathcal{C} \times \mathcal{C}$ into $\mathcal{C}$ defined by (14.26). □

Proof. The following relation between the operations $\vee_2$ and $\sum_{i=1}^{\infty *}$ is almost evident. If $\langle \alpha_i\rangle_{i=1}^{\infty}$ is a sequence of real numbers from $\mathcal{C}$ such that $\alpha_i = 0$ for all $i > n$ and $\sum_{i=1}^{\infty *} \alpha_i$ is defined, then

$$\sum_{i=1}^{\infty *} \alpha_i = \alpha_1 \vee_2 \alpha_2 \vee_2 \cdots \vee_2 \alpha_n \tag{14.29}$$

holds and we shall use the notation $\sum_{i=1}^{n *} \alpha_i$ to abbreviate the right-hand side expression in (14.29). Or, (14.29) yields, by an easy induction, that

$$\alpha_1 \vee_2 \alpha_2 \vee_2 \cdots \vee_2 \alpha_n = \varphi\left(\bigcup_{i=1}^{n} \varphi^{-1}(\alpha_i)\right), \tag{14.30}$$

from what also the associativity of the operation $\vee_2$ and, consequently, the possibility to avoid bracketing in the left-hand side expression in (14.30) follows. If $\alpha_i = 0$, then

$\varphi^{-1}(\alpha_i) = \emptyset \subset \mathcal{N}^+$, hence, if $\alpha_i = 0$ for all $i > n$, the equality

$$\sum\nolimits_{i=1}^{\infty *} \alpha_i = \varphi\left(\bigcup\nolimits_{i=1}^{\infty}\varphi^{-1}(\alpha_i)\right) = \varphi\left(\bigcup\nolimits_{i=1}^{n}\varphi^{-1}(\alpha_i)\right) = \tag{14.31}$$
$$= \alpha_1 \vee_2 \alpha_2 \vee_2 \cdots \vee_2 \alpha_n = \sum\nolimits_{i=1}^{n*} \alpha_i$$

holds.

Let A, $B \in \mathcal{A}$. Setting $E_1 = A - B$, $E_2 = A \cap B$, $E_3 = B - A$, and $E_i = \emptyset \subset \Omega$ for each $i > 3$, we obtain a sequence of mutually disjoint measurable sets from $\mathcal{A}$, so that $\sum_{i=1}^{\infty *} \mu(E_i)$ is defined and $\bigcup_{i=1}^{\infty} E_i = A \cup B$. So, (14.26) and (14.31) yield that

$$\mu(A \cup B) = \sum\nolimits_{i=1}^{\infty *} \mu(E_i) = \mu(E_1) \vee_2 \mu(E_2) \vee_2 \mu(E_3) = \tag{14.32}$$
$$= \mu(A - B) \vee_2 \mu(A \cap B) \vee_2 \mu(B - A) =$$
$$= (\mu(A - B) \vee_2 \mu(A \cap B)) \vee_2 \mu(B - A) =$$
$$= \varphi\left(\varphi^{-1}(\mu(A - B)) \cup \varphi^{-1}(\mu(A \cap B))\right) \vee_2 \mu(B - A) =$$
$$= \varphi\left[\varphi^{-1}\left(\varphi(\varphi^{-1}(\mu(A - B)) \cup \varphi^{-1}(\mu(A \cap B)))\right) \cup \varphi^{-1}(\mu(B - A))\right] =$$
$$= \varphi\left[\varphi^{-1}(\mu(A - B)) \cup \varphi^{-1}(\mu(A \cap B)) \cup \varphi^{-1}(\mu(B - A))\right] =$$
$$= \varphi\left[\left(\varphi^{-1}(\mu(A - B)) \cup \varphi^{-1}(\mu(A \cap B)) \cup \varphi^{-1}(\mu(A \cap B)) \cup \varphi^{-1}(\mu(B - A))\right)\right] =$$
$$= \varphi\left[\varphi^{-1}(\mu(A)) \cup \varphi^{-1}(\mu(B))\right] = \mu(A) \vee_2 \mu(B).$$

As each nonstandard probability measure is also a classical probability measure, $\mu(\Omega - A) = 1 - \mu(A)$ holds for each $A \in \mathcal{A}$. De Morgan rules then yield that

$$\mu(A \cap B) = \mu(\Omega - ((\Omega - A) \cup (\Omega - B))) = \tag{14.33}$$
$$= 1 - (\mu(\Omega - A) \vee_2 \mu(\Omega - B)) =$$
$$= \varphi\left(\mathcal{N}^+ - \varphi^{-1}(\mu(\Omega - A) \vee_2 \mu(\Omega - B))\right) =$$
$$= \varphi\left(\mathcal{N}^+ - \varphi^{-1}[\varphi[\varphi^{-1}(\mu(\Omega - A)) \cup \varphi^{-1}(\mu(\Omega - B))]]\right) =$$
$$= \varphi\left(\mathcal{N}^+ - [\varphi^{-1}(\mu(\Omega - A)) \cup \varphi^{-1}(\mu(\Omega - B))]\right) =$$
$$= \varphi\left(\mathcal{N}^+ - [\mathcal{N}^+ - (\varphi^{-1}(\mu(A))) \cap \varphi^{-1}(\mu(B))]\right) =$$
$$= \varphi\left(\varphi^{-1}(A) \cap \varphi^{-1}(B)\right) = \mu(A) \wedge_2 \mu(B),$$

as for each $A \in \mathcal{A}$,

$$\varphi^{-1}(\mu(\Omega - A)) = \varphi^{-1}(1 - \mu(A)) = \mathcal{N}^+ - \varphi^{-1}(\mu(A)) \tag{14.34}$$

holds. The theorem is proved. □

Let us recall explicitly that $\sum_{i=1}^{n*} \alpha_i$ is defined, and if this is the case, equal to $\alpha_1 \vee_2 \alpha_2 \vee_2 \cdots \vee_2 \alpha_n$, iff $\sum_{i=1}^{\infty *} \alpha_i$ is defined for $\alpha_i = 0$ for all $i > n$. Hence, when defining $\bigvee_{\alpha \in D}^{*}$ ($\bigvee^* D$, abbreviately) and $\bigwedge_{\alpha \in D}^{*}$ ($\bigwedge^* D$, abbreviately) for each $\emptyset \neq D \subset \mathcal{C}$ by

$$\bigvee\nolimits^* D = \varphi\left(\bigcup\nolimits_{\alpha \in D}\varphi^{-1}(\alpha)\right), \quad \bigwedge\nolimits^* D = \varphi\left(\bigcap\nolimits_{\alpha \in D}\varphi^{-1}(\alpha)\right), \tag{14.35}$$

then $\bigvee^* D$ is a conservative extension of $\sum_{i=1}^{\infty *} \alpha_i$ in the case when $D = \{\alpha_1, \alpha_2, \ldots\}$ and $\sum_{i=1}^{\infty *} \alpha_i$ is defined. However, $\bigvee^* D$ (and $\bigwedge^* D$) is defined for all $\emptyset \neq D \subset \mathcal{C}$, but $\sum_{i=1}^{\infty *} \alpha_i$ is defined only when $\varphi^{-1}(\alpha_i) \cap \varphi^{-1}(\alpha_j) = \emptyset \subset \mathcal{N}^+$ holds for each i, $j = 1, 2, \ldots,$ $i \neq j$.

Definition 14.7. Let S be a finite nonempty set. *Basic nonstandard probability assignment (b.ns-p.a.) on* S (or: *over* S) is a mapping $m^* : \mathcal{P}(S) \to \mathcal{C}$ such that $\sum^*_{A \subset S} m^*(A)$ is defined and $\sum^*_{A \subset S} m^*(A) = 1$. □

Remark. The value $\sum^*_{A \subset S} m^*(A)$ is defined by $\sum_{i=1}^{\infty^*} \alpha_i$, where $\langle A_1, A_2, \ldots, A_s \rangle$, $s = \text{card}(\mathcal{P}(S)) = 2^{\text{card}(S)}$, is an ordering (without repetitions) of all subsets of $S, \alpha_i = m^*(A_i)$ for $i \leq s$, and $\alpha_i = 0$ for all $i \in \mathcal{N}^+$, $i > s$. If this is the case, i.e., if $\sum^*_{A \subset S} m^*(A)$ is defined, then obviously $\sum_{i=1}^{\infty *} \alpha_i = \sum_{i=1}^{s*} \alpha_i = \sum_{i=1}^{s*} m^*(A_i) = m^*(A_1) \vee_2 m^*(A_2) \vee_2 \cdots \vee_2 m^*(A_s)$. As the operation $\vee_2$ is commutative and associative, the value $\sum^*_{A \subset S} m^*(A)$ is defined unambiguously, i.e., it does not depend on the particular ordering $\langle A_1, A_2, \ldots, A_s \rangle$ of all subsets of S.

As in the case of usual b.p.a.'s, every ns-b.p.a. on a finite set S can be induced by a set – valued random variable defined on a nonstandard probability space, as the following statement claims and proves.

Theorem 14.2. There exists a nonstandard probability space $\langle \Omega, \mathcal{A}, \mu \rangle$ such that, for each finite nonempty set S and each ns.-b.p.a. m^* on S, there exists a measurable mapping (set-valued random variable, in other terms) $U_{m^*} : \langle \Omega, \mathcal{A}, \mu \rangle \to \langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$ such that, for each $A \subset S$,

$$m^*(A) = \mu\left(\{\omega \in \Omega : U_{m^*}(\omega) = A\}\right). \tag{14.36}$$

□

Proof. Let $\Omega = \mathcal{N}^+ = \{1, 2, \ldots\}$ be the set of all positive integers, let $\mathcal{A} = \mathcal{P}(\mathcal{N}^+)$ be the system of all sets of positive integers, let $\mu(\{i\}) = 2 \cdot 3^{-i}$ for all $i \in \mathcal{N}^+$. Hence, $\mu(\Omega) = \mu(\mathcal{N}^+) = \sum_{i=1}^{\infty} 2 \cdot 3^{-i} = (2/3)\,(1 - (1/3))^{-1} = 1$ and $\mu(A) = \sum_{i \in A} \mu(\{i\}) = \sum_{i \in A} 2 \cdot 3^{-i} \in \mathcal{C}$ holds for each $A \subset \mathcal{N}^+$. Consequently, $\langle \Omega, \mathcal{A}, \mu \rangle = \langle \mathcal{N}^+, \mathcal{P}(\mathcal{N}^+), \mu \rangle$ is a ns. probability space. Or, let $A_1, A_2, \ldots \subset \mathcal{N}^+$ be an infinite sequence of mutually disjoint subsets of $\mathcal{N}^+$. Then

$$\begin{aligned} \mu\left(\textstyle\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{i \in \bigcup_{i=1}^{\infty} A_i} 2 \cdot 3^{-j} = \sum_{i=1}^{\infty} \sum_{j \in A_i} 2 \cdot 3^{-j} = \\ &= \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty *} \mu(A_i). \end{aligned} \tag{14.37}$$

Let us recall the $1-1$ mapping $\varphi : \mathcal{P}(\mathcal{N}^+) \to \mathcal{C}$ (the Cantor set) defined above. For each $\boldsymbol{x} \in \mathcal{C}$, $\boldsymbol{x} \leftrightarrow \langle x_1, x_2, \ldots \rangle \in \{0, 2\}^{\infty}$,

$$\varphi^{-1}(\boldsymbol{x}) = \{i \in \mathcal{N}^+ : x_i = 2\}, \tag{14.38}$$

so that $\boldsymbol{x} \in \sum_{i \in \varphi^{-1}(\boldsymbol{x})} 2 \cdot 3^{-i}$. Set, for each $\omega \in \Omega = \mathcal{N}^+$,

$$U_{m^*}(\omega) = A \subset \mathcal{N}^+ \quad \text{iff} \quad \omega \in \varphi^{-1}(m^*(A)). \tag{14.39}$$

Consequently, for each $A \subset \mathcal{N}^+$,

$$\begin{aligned} &\mu\left(\{\omega \in \Omega : U_{m^*}(\omega) = A\}\right) = \mu\left(\{i \in \mathcal{N}^+ : i \in \varphi^{-1}(m^*(A))\}\right) = \\ &= \mu\left(\varphi^{-1}(m^*(A))\right) = \sum_{i \in \varphi^{-1}(m^*(A))} 2 \cdot 3^{-i} = m^*(A), \end{aligned} \tag{14.40}$$

and the assertion is proved. □

The following theorem deduces and presents a boolean-like modification of Dempster combination rule which can be obtained within the framework of our nonstandard model. The obtained combination rule conserves the semi-extensional nature of the classical Dempster combination rule in the sense that the values of the combined ns. basic probability assignment are defined by, and can be computed from, the values of the two particular ns. b.p.a.'s which are to be combined together, but no assumption concerning the statistical independence (or a special kind and/or degree of depencendence) of the random variables in question is needed.

Theorem 14.3. Let $\langle \Omega, \mathcal{A}, \mu \rangle$ be a nonstandard probability space, let S be a nonempty finite set, let $E_i,\ i = 1,2$, be nonempty empirical spaces, let $\mathcal{E}_i,\ i = 1,2,\ \mathcal{E}_i \subset \mathcal{P}(E_i)$, be nonempty σ-fields of subsets of these empirical spaces. Let $X_i : \langle \Omega, \mathcal{A}, \mu \rangle \rightarrow \langle E_i, \mathcal{E}_i \rangle$, $i = 1,2$, be measurable mappings (generalized random variables), let $\rho_i : S \times E_i \rightarrow \{0,1\},\ i = 1,2$, be compatibility relations over the corresponding spaces. Let the mappings $U_i : \Omega \rightarrow \mathcal{P}(S)$ defined, for each $\omega \in \Omega$ and for both $i = 1,2$, by

$$U_i(\omega) = \{s \in S : \rho_i(s, X_i(\omega)) = 1\} \tag{14.41}$$

be measurable mappings taking the ns. probability space $\langle \Omega, \mathcal{A}, \mu \rangle$ into the measurable space $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$.

Let $\rho_{12} : S \times (E_1 \times E_2) \rightarrow \{0,1\}$ be the compatibility relation over the space S and the Cartesian product $E_1 \times E_2$ of the empirical spaces E_1, E_2, defined by

$$\rho_{12}(s, \langle x_1, x_2 \rangle) = \min\{\rho_1(s, x_1),\ \rho_2(s, x_2)\} \tag{14.42}$$

for every $s \in S,\ x_1 \in E_1,\ x_2 \in E_2$. Set

$$U_{12}(\omega) = \{s \in S : \rho_{12}(s, \langle X_1(\omega), X_2(\omega) \rangle) = 1\} \tag{14.43}$$

and denote by $m_i^*(A),\ i = 1,2,12,\ A \subset S$, the value

$$m_i^*(A) = \mu(\{\omega \in \Omega : U_i(\omega) = A\}). \tag{14.44}$$

Then m_i^* is a ns.b.p.a. on S for each $i = 1,2,12$, and

$$m_{12}^*(A) = {\sum}^*_{B,C \subset S, B \cap C = A} m_1^*(B) \wedge_2 m_2^*(C) \tag{14.45}$$

holds for each $A \subset S$, where $\wedge_2$ is the nonstandard infimum operation defined by (14.26). □

Proof. For $i = 1,2$, both the mappings $U_i : \Omega \rightarrow \mathcal{P}(S)$ defined by (14.41) are supposed to be measurable with respect to $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$, so that $\{\omega \in \Omega : U_i(\omega) = A\} \in \mathcal{A}$ holds for each $A \subset S$ and both $i = 1,2$. Consequently, $\mu(\{\omega \in \Omega : U_i(\omega) = A\})$ is defined. If $A_1,\ A_2 \subset S,\ A_1 \neq A_1$, then

$$\{\omega \in \Omega : U_i(\omega) = A_1\} \cap \{\omega \in \Omega : U_i(\omega) = A_2\} = \emptyset \tag{14.46}$$

is valid for both $i = 1, 2$, so that $\{\{\omega \in \Omega : U_i(\omega) = A\} : A \subset S\}$ is a system of mutually disjoint subsets of Ω (a decomposition of Ω to subsets from $\mathcal{A}$, in fact), and for such systems $\sum^*_{A \subset S} \mu(\{\omega \in \Omega : U_i(\omega) = A\})$ is defined and equals to 1 for $i = 1, 2$, as $\langle \Omega, \mathcal{A}, \mu \rangle$ is a nonstandard probability space. Hence, both m_1^* and m_2^* defined by (14.44) are ns.b.p.a.'s over S.

As in the standard case, (14.42) and (14.43) yield that

$$
\begin{aligned}
U_{12}(\omega) &= \{s \in S : \min\{\rho_1(s, X_1(\omega)), \rho_2(s, X_2(\omega))\} = 1\} = \\
&= \{s \in S : \rho_1(s, X_1(\omega)) = \rho_2(s, X_2(\omega)) = 1\} = \\
&= \{s \in S : \rho_1(s, X_1(\omega)) = 1\} \cap \{s \in S : \rho_2(s, X_2(\omega)) = 1\} = \\
&= U_1(\omega) \cap U_2(\omega).
\end{aligned} \tag{14.47}
$$

For each $A \subset S$,

$$
\begin{aligned}
&\{\omega \in \Omega : U_{12}(\omega) = A\} = \{\omega \in \Omega : U_1(\omega) \cap U_2(\omega) = A\} = \\
= &\bigcup_{B,C,\subset S, B \cap C = A} (\{\omega \in \Omega : U_1(\omega) = B\} \cap \{\omega \in \Omega : U_2(\omega) = C\}).
\end{aligned} \tag{14.48}
$$

This subset of Ω belongs to $\mathcal{A}$, as due to the finiteness of S and, consequently, also of $\mathcal{P}(S)$, there is only a finite number of pairs $\langle B, C \rangle \in \mathcal{P}(S) \times \mathcal{P}(S)$ such that $B \cap C = A$. Being a σ-field, $\mathcal{A}$ is closed with respect to finite intersections and finite unions. Hence, U_{12} is also a measurable mapping which take $\langle \Omega, \mathcal{A}, \mu \rangle$ into $\langle \mathcal{P}(S), \mathcal{P}(\mathcal{P}(S)) \rangle$, and the relation

$$
\sum\nolimits^*_{A \subset S} \mu(\{\omega \in \Omega : U_{12}(\omega) = A\}) = \sum\nolimits^*_{A \subset S} m_{12}^*(A) = 1 \tag{14.49}
$$

can be proved in the same way as in the case of m_1^* and m_2^*.

Let $\langle B_1, C_1 \rangle$, $\langle B_2, C_2 \rangle$ be two different pairs of subsets of S, so that either $B_1 \neq B_2$ or $C_1 \neq C_2$. Then, obviously,

$$
\{\omega \in \Omega : U_1(\omega) = B_1,\ U_2(\omega) = C_1\} \cap \{\omega \in \Omega : U_1(\omega) = B_2,\ U_2(\omega) = C_2\} = \emptyset, \tag{14.50}
$$

so that $m_{12}^*(A)$ can be written as

$$
\begin{aligned}
&m_{12}^*(A) = \mu(\{\omega \in \Omega : U_{12}(\omega) = A\}) = \\
= &\sum\nolimits^*_{B,C \subset S, B \cap C = A} \mu(\{\omega \in \Omega : U_1(\omega) = B\} \cap \{\omega \in \Omega : U_2(\omega) = C\}) = \\
&\sum\nolimits^*_{B,C \subset S, B \cap C = A} \mu(\{\omega \in \Omega : U_1(\omega) = B\}) \wedge_2 \mu(\{\omega \in \Omega : U_2(\omega) = C\}) = \\
= &\sum\nolimits^*_{B,C \subset S, B \cap C = A} m_1^*(B) \wedge m_2^*(C)
\end{aligned} \tag{14.51}
$$

due to Theorem 14.1 and due to the definition of $m_1^*(B)$ and $m_2^*(C)$ by (14.44). The theorem is proved. □

It is perhaps worth noting explicitly, that the formula (14.45), which defines the modified form of the Dempster combination rule, can be rewritten in a way still more close to the standard one. As a matter of fact, the binary operation $\wedge_2$ on $\mathcal{C}$ can be seen also as a natural extension of the standard multiplication (product) operation in $\langle 0, 1 \rangle$ extended in the same boolean-like pointwise way as $\sum^*$ extends addition. This idea follows from the trivial fact that the infimum operation $\wedge_1$ on $B_1 = \{0,1\}^\infty$, defined, for each $\boldsymbol{x} = \langle x_1, x_2, \ldots \rangle$, $\boldsymbol{y} = \langle y_1, y_2, \ldots \rangle \in \{0,1\}^\infty$ by $\boldsymbol{x} \wedge_1 \boldsymbol{y} = \langle \inf\{x_i, x_i\} \rangle \{i = 1\}^\infty$, is identical with the binary operation $\odot_1$, defined by $\boldsymbol{x} \odot_1 \boldsymbol{y} = \langle x_i y_i \rangle_{i=1}^\infty$. Hence, we can write also $\odot_2$ instead of $\wedge_2$ and (14.45) can be converted into

$$m_{12}^*(A) = \sum\nolimits_{B,C,\subset S, B\cap C = A}^{*} m_1^*(B) \odot_2 m_2^*(C). \tag{14.52}$$

An open, and perhaps interesting question for a further research reads, whether the Boolean algebra $\mathcal{B}_0$ of all subsets of the set $\mathcal{N}^+$ of positive integers can be replaced by another and perhaps more general and abstract Boolean algebra $\mathcal{B}$, and which restrictions should be imposed on $\mathcal{B}$ in order to be able to encode its elements by (some) real numbers from the unit interval in a reasonable and one-to-one way, or at least to assure that such a mapping exists. However, let us postpone such an investigation till another occasion.

Bibliography

[1] S.K. Andersen, K.G. Olesen, F.V. Jensen and F. Jensen: HUGIN – A shell for building Bayesian belief universes for expert systems. In: Proceedings of the 11th International Joint Conference on Artificial Intelligence, 1989, pp. 1080–1085.

[2] B. Balcar, P. Štěpánek: Teorie množin (Set Theory – in Czech). Academia, Prague, 1986.

[3] D. Blackwell, M.A. Girshick: Theory of Games and Statistical Decisions. John Wiley, New York, 1954.

[4] W. Brachinger, J. Kohlas: Argumentation systems and evidence theory. In: Advances in Intelligent Computing – IPMU 94 (Bouchon–Meunier and Yager, Eds.). Lecture Notes in Computer Science 945, Springer Verlag, Berlin – Heidelberg – New York, 1994, pp. 41–50.

[5] A. Bundy: Incidence calculus – a mechanism for probabilistic reasoning. Journal of Automated Reasoning 1 (1985), pp. 263–283.

[6] G. Choquet: Theory of capacities. Annales de l'Institut Fourier 5 (1953–4), pp. 131–295.

[7] A.P. Dempster: Upper and lower probabilities induced by a multi-valued mapping. Annals of Mathematical Statistics 38 (1967), no. 2, pp. 325–339.

[8] C.A. Drossos: Foundation of fuzzy sets – a nonstandard approach. Fuzzy Sets and Systems 37 (1990), pp. 287–307.

[9] D. Dubois, H. Prade: Théorie de Possibilités – Applications à le Représentation de Connaissances en Informatique. Mason, Paris, 1985.

[10] R. Faure, E. Heurgon: Structures Ordonnées et Algèbres de Boole. Gauthier–Villars, Paris, 1971.

[11] W. Feller: An Introduction to Probability Theory and its Applications I, II. J. Wiley and Sons, New York – London, Vol. I 1957, Vol. II 1965.

[12] T.L. Fine: Theories of Probability – an Examination of Foundations. Academic Press, New York, 1973.

[13] A. A. Fraenkel, Y. Bar–Hillel: Foundations of Set Theory. North Holland, Amsterdam, 1959.

[14] B. V. Gnedenko: Kurs tĕoriji vĕrojatnostĕj (A Course in Probability Theory – in Russian). Nauka, Moscow, 1965.

[15] J. A. Goguen: L–fuzzy sets. J. of Math. Analysis and Applications 18 (1967), pp. 145–174.

[16] S. Guiasu: Information Theory with Applications. McGraw–Hill, New York, 1977.

[17] P. Hájek, T. Havránek, R. Jiroušek: Uncertain Information Processing in Expert Systems, CRC Press, 1992.

[18] P. R. Halmos: Measure Theory. D. van Nostrand, New York – Toronto – London, 1950.

[19] D. Harmanec: Uncertainty in Dempster–Shafer Theory. Ph. D.–dissertation, Univ. of New York, Binghamton, 1996.

[20] S. J. Henkind, M. C. Harrison: An analysis of four uncertainty calculi. IEEE Transactions on Systems, Man, and Cybernetics 18 (1988), no. 5, pp. 700–714.

[21] R. Kennes: Evidential reasoning in a categorial perspective: Conjunction and disjunction of belief functions. In: Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence (B. D. d'Ambrosio, P. Smets, P. P. Bonissone, Eds.). Morgan Kaufman, San Matteo, California, 1991, pp. 174–181.

[22] G. J. Klir, T. A. Folger: Fuzzy Sets, Uncertainty, and Information. Prentice–Hall, Englewood Cliffs,, New Jersey, 1988.

[23] J. Kohlas: Mathematical foundations of evidence theory. In: Mathematical Models for Handling Partial Knowledge in Artificial Intelligence (G. Coletti et al, Eds.). Plenum Press, New York, 1994, pp. 31–64.

[24] J. Kohlas, P. A. Monney: A Mathematical Theory of Hints – An Approach to the Dempster–Shafer Theory of Evidence. Lecture Notes in Economics and Mathematical Systems, Springer Verlag, New York, 1995.

[25] A. N. Kolmogorov: Osnovnyje poujatija teoriji vĕrojatnostĕj (Basic Notions of Probability Theory – in Russian). Nauka, Moscow, 1974.

[26] I. Kramosil: Possible–worlds semantics for rule–based expert systems with set–valued weights. In: Proceeding of the Workshop on Uncertainty Processing in Expert Systems, Alšovice, June 20 – 23, 1988. Kybernetika 25 (1989) – supplement, pp. 65–71.

[27] I. Kramosil: Expert systems with boolean-valued uncertainties. Computers and Artificial Intelligence 10 (1991), pp. 43–55.

[28] I. Kramosil: Strong law of large numbers for set-valued random variables. In: WUPES 94 – Proceedings of the 3rd Workshop on Uncertainty Processing in Expert Systems, University of Economics, Prague, 1994, pp. 122–142.

[29] I. Kramosil: Believeability functions induced by partial generalized compatibility relations. In: Fuzzy Sets 94 – Proceedings of the Second International Conference on Fuzzy Sets Theory and its Applications (R. Mesiar and V. Novák, Eds.). Tatra Mountains Mathematical Publications vol. 6 (1995), pp. 103–116.

[30] I. Kramosil: Nonstandard approach to possibility measures. Int. J. of Uncertainty, Fuzziness and Knowledge–Based Systems 4 (1996), no. 3, pp. 275–301.

[31] I. Kramosil: Some remarks on philosophical aspects of the laws of large numbers. In: LOGICA 95 – Proceedings of the Symposium (T. Childers, P. Kolář and V. Svoboda, Eds.), Filosofia, Prague, 1996, pp. 179–198.

[32] I. Kramosil: Belief functions with nonstandard values. In: ECSQARU 97 – Proceedings of the Conference (D. Gabbay, R. Kruse, A. Nonnengart, H. J. Ohlbach, Eds.). Lecture Notes in Artificial Intelligence 1244, Springer Verlag, Berlin – Heidelberg – New York, 1997, pp. 380–391.

[33] I. Kramosil: Probabilistic analysis of Dempster combination rule. In: The Logica Yearbook 1997 (T. Childers, Ed.), Filosofia, Prague, 1998, pp. 175–187.

[34] H. E. Kyburg: Probability and Inductive Logic. Collier–Macmillan, London, 1970.

[35] E. L. Lehman: Testing Statistical Hypotheses. John Wiley, New York, 1947.

[36] M. Loève: Probability Theory. D. van Nonstrand, New York, 1960.

[37] V. Novák: Fuzzy Sets and their Applications. Adam Hilger, Bristol–Philadelphia, 1989.

[38] J. Pearl: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Matteo, California, 1988.

[39] A. Rényi: Probability Theory. North Holland, Amsterdam, 1970.

[40] G. Shafer: A Mathematical Theory of Evidence. Princeton Univ. Press, Princeton, 1976.

[41] G. Shaffer: Allocations of probability. The Annals of Probability 7 (1979), no. 5, pp. 827–839.

[42] R. Sikorski: Boolean Algebras. Springer Verlag, Berlin – Göttingen – Heidelberg, 1960.

[43] Ph. Smets: The nature of the unnormalized beliefs encountered in the transferable belief model. In: Uncertainty in Artificial Intelligence 92 (D. Dubois, M. P. Wellman, B. d'Ambrosio, Ph. Smets, Eds.). Morgan Kaufmann, San Matteo, California, 1992, pp. 292–297.

[44] Ph. Smets: The representation of quantified belief by the transferable belief model. Technical Report no TR/IRIDIA/94–19.1, Institut de Récherches Interdisciplinaires et de Dévelopments en Intelligence Artificielle, Université Libre de Bruxelles, 1994, 49 pp.

[45] Ph. Smets: The canonical decomposition of a weighted belief. In: IJCAI 95 – Proceedings of the 14th International Joint Conference on Artificial Intelligence, Montréal, August 20–25, 1995, Vol. 2, pp. 1896–1901.

[46] J. Štěpán: Teorie pravděpodobnosti (Probability Theory – in Czech). Academia, Prague, 1987.

[47] J. Vejnarová: A few remarks on measures of uncertainty in Dempster–Shafer theory. In: Proceedings of Workshop on Uncertainty in Expert Systems. Czechoslovak Academy of Sciences, Prague, 1991, pages not numbered.

[48] P. Vopěnka, P. Hájek: The Theory of Semisets. Academia, Prague, 1972.

[49] P. Vopěnka: Rozpravy s geometrií I (Discussions with Geometry – in Czech). Panorama, Prague, 1989.

[50] J. Štěpán: Teorie pravděpodobnosti (Probability Theory – in Czech). Academia, Prague, 1987.

[51] R. A. Yager: Entropy and specificity in a mathematical theory of evidence. International Journal of General Systems 9 (1983), no. 4, pp. 249–260.