KYBERNETIKA — VOLUME 38 (2002), NUMBER 1, PAGES 1–12

A NONSTANDARD MODIFICATION OF DEMPSTER COMBINATION RULE[1]

IVAN KRAMOSIL

It is a well-known fact that the Dempster combination rule for combination of uncertainty degrees coming from two or more sources is legitimate only if the combined empirical data, charged with uncertainty and taken as random variables, are statistically (stochastically) independent. We shall prove, however, that for a particular but large enough class of probability measures, an analogy of Dempster combination rule, preserving its extensional character but using some nonstandard and boolean-like structures over the unit interval of real numbers, can be obtained without the assumption of statistical independence of input empirical data charged with uncertainty.

1. BELIEF FUNCTIONS AND DEMPSTER COMBINATION RULE

Let us limit ourselves to a purely theoretical approach focused just to the mathematical apparatus used in order to formalize the basic notions and results of the Dempster–Shafer model of uncertainty quantification and processing.

Let S be a finite nonempty set, let $\mathcal{P}(S)$ denote the system of all subsets of S. *Basic probability assignment* (b.p.a.) over S is a probability distribution m over $\mathcal{P}(S)$, i.e., a mapping $m : \mathcal{P}(S) \rightarrow [0,1]$ (the unit interval of real numbers) such that $\sum_{A \subset S} m(A) = 1$. *Non-normalized belief function* defined (or: induced) by the b.p.a. m over S is the mapping $bel^*_m : \mathcal{P}(S) \rightarrow [0,1]$ such that, for each $A \subset S$,

$$bel^*_m(A) = \sum\nolimits_{\emptyset \neq B \subset A} m(B), \tag{1.1}$$

for the empty subset $\emptyset$ of S we adopt the convention that summing over the empty set of items, the result equals zero. *Normalized belief function* defined (or: induced) by the b.p.a. m over S is the mapping $bel_m : \mathcal{P}(S) \rightarrow [0,1]$ such that, for each $A \subset S$

$$bel_m(A) = (1 - m(\emptyset))^{-1} \sum\nolimits_{\emptyset \neq B \subset A} m(B), \tag{1.2}$$

supposing that $m(\emptyset) < 1$ holds, $bel_m(\emptyset) = 0$ according to the same convention as above. If $m(\emptyset) = 1$, bel_m is not defined.

[1]This work has been supported by Grant A 1030803 of the Grant Agency of the Academy of Sciences of the Czech Republic.

At the same combinatoric level the Dempster combination rule can be also easily defined. Let m_1, m_2 be two b.p.a.'s over the same (nonempty finite) set S. Set, for each $A \subset S$

$$m_3(A) = \sum\nolimits_{B,C\subset S,\, B\cap C=A} m_1(B)\, m_2(C). \tag{1.3}$$

An easy calculation yields that

$$\begin{aligned} & \sum\nolimits_{A\subset S} m_3(A) \qquad (1.4)\\ = \ & \sum\nolimits_{A\subset S} \left(\sum\nolimits_{\langle B,C\rangle\in\mathcal{P}(S)\times\mathcal{P}(S),\, B\cap C=A} m_1(B)\, m_2(C)\right)\\ = \ & \sum\nolimits_{\langle B,C\rangle\in\mathcal{P}(S)\times\mathcal{P}(S)} m_1(B)\, m_2(C) = \sum\nolimits_{B\subset S} m_1(B) \left(\sum\nolimits_{C\subset S} m_2(C)\right)\\ = \ & \sum\nolimits_{B\subset S} m_1(B) = 1, \end{aligned}$$

as the set of pairs $\langle B, C\rangle$ such that $B \cap C = A$ are disjoint for different A's and each $\langle B, C\rangle$, $B \subset S$, $C \subset S$, belongs to just one set of pairs, namely to that with $A = B \cap C$. Hence, the mapping $m_3 : \mathcal{P}(S) \to [0,1]$ is also a probability distribution over $\mathcal{P}(S)$, i.e., a b.p.a. over S, it is called the *Dempster product* of the b.p.a.'s m_1, m_2, and denoted by $m_1 \oplus m_2$. The binary operation $\oplus$, taking pairs of b.p.a.'s over S into the space of b.p.a.'s over the same S, is called the *Dempster combination rule* for b.p.a.'s.

Dempster combination rule for belief functions is defined in such a way that Dempster product $bel^*_{m_1} \oplus bel^*_{m_2}$ ($bel_{m_1} \oplus bel_{m_2}$, resp.) of two belief functions is the belief function induced by the Dempster product of the b.p.a.'s defining the particular belief functions under combination. In symbols,

$$bel^*_{m_1} \oplus bel^*_{m_2} = bel^*_{m_1\oplus m_2}, \tag{1.5}$$

$$bel_{m_1} \oplus bel^*_{m_2} = bel_{m_1\oplus m_2}, \tag{1.6}$$

in the case of (1.6), of course, only when $bel_{m_1\oplus m_2}$ is defined, i.e., when $(m_1 \oplus m_2)(\emptyset) < 1$ holds. Equality symbol in (1.5) and (1.6) denotes the identity relation, i.e., the equality of the corresponding values for each $A \subset S$. As can be easily proved, Dempster combination rule is commutative and associative for b.p.a.'s as well as for belief functions.

Dempster combination rule is usually presented as an appropriate tool how to combine numerical degrees of uncertainties concerning the same field of events but of different provenience or coming from different sources, e.g., from two subjects or experts with their particular pieces of knowledge being charged by a portion of subjectivity. However, when taking both the particular degrees of uncertainty as a priori probability measures not conditioned by each other, then every rule combining these two probability measures into a uniquely determined one necessary introduces into the model in question a hidden assumption of a fixed kind and degree of statistical (stochastical) (in) dependence between the two sources of uncertainty. We will show, in the rest of this chapter and very briefly, referring to [4] or [5] for a more detailed case analysis, that this is, in fact, the case.

Let S be interpreted as a nonempty finite set of possible internal states of a system, just one $s_0 \in S$ being the actual one. The subject's aim is either to identify the actual internal state s_0, or at least to decide whether $s_0 \in T$ holds or does not hold for some (proper, as a rule) subset T of S. The subject is not able to answer this question by the direct observation of s_0, so that she/he has to guess the correct answer on the ground of some observation(s) concerning the system in question and its environment. Let us denote by x this empirical value and by E the space (perhaps a vector one) of all possible empirical values. In order the subject's reasonings were based on some rational grounds, she/he must know at least some relations holding between the actual internal state of the system and the observed empirical values. Namely, the subject has at her/his disposal a relation $\rho \subset S \times E$, also taken as a mapping $\rho : S \times E \to \{0,1\}$, such that, given $s \in S$ and $x \in E$, $\rho(s,x) = 0$ iff the subject knows (or is able to deduce within the scope of her/his deductive abilities) that the actual internal state of the system cannot be s supposing that the value x was observed. If this is not the case, i. e., if $\rho(s,x) = 1$, then the subject cannot avoid the possibility just described, hence, the state s and the empirical value x are *compatible*. Consequently, ρ is called the *compatibility relation* and it will play the role of the keystone in our further considerations and constructions.

In order to describe the supposed random nature of the observed empirical values, we shall suppose that the observed empirical value $x \in E$ is the realization of a random variable X. This random variable is defined as a measurable mapping which takes an abstract and, in what follows, fixed probability space $\langle \Omega, \mathcal{A}, P \rangle$ into the measurable space $\langle E, \mathcal{E} \rangle$ generated over E when choosing and fixing a nonempty σ-field $\mathcal{E} \subset \mathcal{P}(E)$ of subsets of the space E.

Given a compatibility relation $\rho : S \times E \to \{0,1\}$, we denote by $U_\rho(x) = \{s \in S : \rho(s,x) = 1\}$ the set of all states from S which are compatible with the empirical value $x \in E$. Combining this notation with the mapping $X : \langle \Omega, \mathcal{A}, P \rangle \to \langle E, \mathcal{E} \rangle$, we obtain the composed mapping $U_\rho(X(\cdot)) : \Omega \to \mathcal{P}(S)$, hence, for each $\omega \in \Omega$,

$$U_\rho(X(\omega)) = \{s \in S : \rho(s, X(\omega)) = 1\}. \tag{1.7}$$

We will suppose that $U_\rho(X(\cdot))$ is measurable in the sense that

$$\{\{\omega \in \Omega : U_\rho(X(\omega)) = A\} : A \subset S\} \subset \mathcal{A} \tag{1.8}$$

holds. Let us denote, for each $A \subset S$, by $m(A)$ the value

$$m(A) = P\left(\{\omega \in \Omega : U_\rho(X(\omega)) = A\}\right). \tag{1.9}$$

As can be easily proved, the mapping $m : \mathcal{P}(S) \to [0,1]$ is a b.p.a. over S, and for every b.p.a. m^0 over S there exists a mapping $U^0 : \Omega \to \mathcal{P}(S)$ defining m^0 in the sense of (1.9). Moreover, it follows easily that

$$bel^*_m(A) = P\left(\{\omega \in \Omega : \emptyset \neq U_\rho(X(\omega)) \subset A\}\right) \tag{1.10}$$

and

$$bel_m(A) = P\left(\{\omega \in \Omega : U_\rho(X(\omega)) \subset A\} \,/\, \{\omega \in \Omega : U_\rho(X(\omega)) \neq \emptyset\}\right) \tag{1.11}$$

hold for each $A \subset S$ supposing that the conditional probability in (1.11) is defined.

Consider the case when *two* subjects solve the same problem to identify, or at least to specify partially, the actual internal state of the system under consideration. Their observations can be, however, of different kind and nature, so that there are two (possibly different) spaces E_1, E_2 of empirical values, each of them equipped by its own σ-field $\mathcal{E}_1$, $\mathcal{E}_2$ of subsets, and there are two random variables X_1, X_2 both defined on the same probability $\langle \Omega, \mathcal{A}, P \rangle$, but taking their values in $\langle E_1, \mathcal{E}_1 \rangle$ for X_1 and in $\langle E_2, \mathcal{E}_2 \rangle$ for X_2. The apriori knowledge of each of them is defined by compatibility relations $\rho_1 : S \times E_1 \to \{0,1\}$ and $\rho_2 : S \times E_2 \to \{0,1\}$ ($\rho_1 \subset S \times E_1$, $\rho_2 \subset S \times E_2$, under the set-theoretic notation).

Let the two subjects (or some third "meta-subject") decide to combine their a priori knowledge and empirical data in the following way. Let $E_{12} = E_1 \times E_2$ be the Cartesian product of both the empirical spaces, let $\mathcal{E}_{12}$ be the σ-field of subsets of $E_1 \times E_2$ generated by $\mathcal{E}_1$ and $\mathcal{E}_2$. Let $X_{12} = \langle X_1, X_2 \rangle : \Omega \to E_{12}$ be the mapping defined by $X_{12}(\omega) = \langle X_1(\omega), X_2(\omega) \rangle$ for each $\omega \in \Omega$; an elementary result of measure theory (cf. [3], e. g.) then reads that X_{12} is a measurable mapping. Finally, let the compatibility relation $\rho_{12} : S \times E_{12} \to \{0,1\}$ be defined by

$$\rho_{12}(s, \langle x_1, x_2 \rangle) = \min\{\rho_1(s, x_1),\ \rho_2(s, x_2)\} \tag{1.12}$$

for each $s \in S$, $x_1 \in E_1$, and $x_2 \in E_2$. An easy calculation yields that, for each $x = \langle x_1, x_2 \rangle \in E_{12}$,

$$U_{\rho_{12}}(x) = U_{\rho_1}(x_1) \cap U_{\rho_2}(x_2), \tag{1.13}$$

hence, for each $\omega \in \Omega$,

$$U_{\rho_{12}}(X_{12}(\omega)) = U_{\rho_1}(X_1(\omega)) \cap U_{\rho_2}(X_2(\omega)). \tag{1.14}$$

As in the case of ρ_1 and ρ_2, we can define b.p.a. m_{12} and belief functions $bel^*_{m_{12}}$ and $bel_{m_{12}}$ generated by ρ_{12} and X_{12}. We obtain, for each $A \subset S$, that

$$\begin{aligned} m_{12}(A) &= P\left(\{\omega \in \Omega : U_{\rho_{12}}(X_{12}(\omega)) = A\}\right), \\ bel^*_{m_{12}}(A) &= P\left(\{\omega \in \Omega : \emptyset \neq U_{\rho_{12}}(X_{12}(\omega)) \subset A\}\right), \\ bel_{m_{12}}(A) &= P\left(\{\omega \in \Omega : U_{\rho_{12}}(X_{12}(\omega)) \subset A\} \,/\, \{\omega \in \Omega : U_{\rho_{12}}(X_{12}(\omega)) \neq \emptyset\}\right), \end{aligned} \tag{1.15}$$

supposing that the last conditional probability is defined, i. e., supposing that $m_{12}(\emptyset) = P(\{\omega \in \Omega : U_{\rho_{12}}(X_{12}(\omega)) = \emptyset\}) < 1$ holds.

The space S is assumed to be finite, so that $\mathcal{P}(S)$ is finite as well and for each $A_1,\ A_2 \subset S,\ A_1 \neq A_2$,

$$\{\langle B, C \rangle : B,\ C \subset S,\ B \cap C = A_1\} \cap \{\langle B, C \rangle : B,\ C \subset S,\ B \cap C = A_2\} = \emptyset. \tag{1.16}$$

Consequently, (1.15) can be rewritten as

$$m_{12}(A) = \sum\nolimits_{\langle B,C \rangle, B, C \subset S, B \cap C = A} P(\{\omega \in \Omega : U_{\rho_1}(X_1(\omega)) = B, U_{\rho_2}(X_2(\omega)) = C\}), \tag{1.17}$$

$$bel^*_{m_{12}}(A) = \sum\nolimits_{\langle B,C \rangle, B, C \subset S, \emptyset \neq B \cap C \subset A} P(\{\omega \in \Omega : U_{\rho_1}(X_1(\omega)) = B, U_{\rho_2}(X_2(\omega)) = C\}),$$

$$bel_{m_{12}}(A) = \frac{\sum_{\langle B,C \rangle, B, C \subset S, \emptyset \neq B \cap C \subset A} P\left(\{\omega \in \Omega : U_{\rho_1}(X_1(\omega)) = B, U_{\rho_2}(X_2(\omega)) = C\}\right)}{\sum_{\langle B,C \rangle, B, C \subset S, \emptyset \neq B \cap C} P\left(\{\omega \in \Omega : U_{\rho_1}(X_1(\omega)) = B, U_{\rho_2}(X_2(\omega)) = C\}\right)}.$$

Let us state explicitly, now, that we assume that for all $A \subset S$ and each $\rho(\rho_1, \rho_2, \rho_{12}$, resp.)

$$\{\omega \in \Omega : U_\rho(X(\omega)) = A\} = (U_\rho(X))^{-1}(A) \in \mathcal{A} \tag{1.18}$$

holds, so that the values $m(A)$, $m_1(A)$, $m_2(A)$ and $m_{12}(A)$ are defined. This assumption will be accepted also below, moreover, we shall suppose that the random variables X_1 and X_2 are *stochastically (statistically) independent.*

$$\begin{aligned} & P(\{\omega \in \Omega : X_1(\omega) \in F_1,\, X_2(\omega) \in F_2\}) \\ = \; & P(\{\omega \in \Omega : X_1(\omega) \in F_1\})\, P(\{\omega \in \Omega : X_2(\omega) \in F_2\}) \end{aligned} \tag{1.19}$$

holds. Given $A_1, A_2 \subset S$ we obtain that also the set-valued random variables $U_{\rho_1}(X_1(\cdot))$ and $U_{\rho_2}(X_2(\cdot))$ are statistically independent, so that the equality

$$\begin{aligned} & P(\{\omega \in \Omega : U_{\rho_1}(X_1(\omega)) = A_1,\, U_{\rho_2}(X_2(\omega)) = A_2\}) \\ = \; & P(\{\omega \in \Omega : U_{\rho_1}(X_1(\omega)) = A_1\})\, P(\{\omega \in \Omega : U_{\rho_2}(X_2(\omega)) = A_2\}) \end{aligned} \tag{1.20}$$

holds as well. Consequently, (1.17) can be rewritten as

$$\begin{aligned} m_{12}(A) &= \sum\nolimits_{\langle B,C\rangle, B,C \subset S, B\cap C = A} P(\{\omega \in \Omega : U_{\rho_1}(X_1(\omega)) = B\}) \cdot \\ &\qquad \cdot P(\{\omega \in \Omega : U_{\rho_2}(X_2(\omega)) = C\}) \\ &= \sum\nolimits_{\langle B,C\rangle, B,C \subset S, B\cap C = A} m_1(B)\, m_2(C), \\ bel^*_{m_{12}}(A) &= \sum\nolimits_{\langle B,C\rangle, B,C \subset S, \emptyset \neq B\cap C \subset A} m_1(B)\, m_2(C), \\ bel_{m_{12}}(A) &= \frac{\sum_{\langle B,C\rangle, \emptyset \neq B\cap C \subset A} m_1(B)\, m_2(C)}{\sum_{\langle B,C\rangle, \emptyset \neq B\cap C} m_1(B)\, m_2(C)}. \end{aligned} \tag{1.21}$$

Hence, $m_{12} = m_1 \oplus m_2$, $bel^*_{m_{12}} = bel^*_{m_1} \oplus bel^*_{m_2}$, and $bel_{m_{12}} = bel_{m_1} \oplus bel_{m_2}$. In other words said, application of Dempster combination rule when combining two (or more) degrees of uncertainty defined by particular belief functions is sound and justifiable only supposing that the particular random empirical data are statistically independent and that the pieces of knowledge described by particular compatibility relations are combined in the "optimistic" sense. Perhaps strange and interesting enough, we will see, in the rest of this paper, that these assumptions are not necessary when processing the degrees of belief in a nonstandard and, in a sense, boolean-like way. In the next chapter we shall introduce some necessary and very elementary technical preliminaries.

2. ARITHMETICAL AND PROBABILISTIC STRUCTURES OVER BOOLEAN-LIKE PROCESSED REAL NUMBERS

The reader is supposed to be familiar with the notion of Boolean algebra and with the most elementary properties of these structures, cf., e. g. [2] or [8]. Following

[8], we shall define Boolean algebra $\mathcal{B}$ as a quadruple $\langle B, \vee, \wedge, \neg \rangle$, where B is a nonempty set (called the *support* of $\mathcal{B}$), $\vee$ and $\wedge$ are total binary operations taking the Cartesian product $B \times B$ into B, and $\neg$ is a total unary operation taking B into B, such that for all x, y, $z \in B$ the following identities hold

$$\begin{aligned}
&(\text{A1}) \quad x \vee y = y \vee x, \quad x \wedge y = y \wedge x, && (2.1)\\
&(\text{A2}) \quad x \vee (y \vee z) = (x \vee y) \vee z, \quad x \wedge (y \wedge z) = (x \wedge y) \wedge z, \\
&(\text{A3}) \quad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z), \quad x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \\
&(\text{A4}) \quad x \wedge (x \vee y) = x, \quad x \vee (x \wedge y) = x, \\
&(\text{A5}) \quad (x \wedge (\neg x)) \vee y = y, \quad (x \vee (\neg x)) \wedge y = y.
\end{aligned}$$

The zero element $\mathbf{0}_{\mathcal{B}}$ of the Boolean algebra $\mathcal{B}$ is defined by $\mathbf{0}_{\mathcal{B}} = x \wedge (\neg x)$, the unit element $\mathbf{1}_{\mathcal{B}}$ of $\mathcal{B}$ by $\mathbf{1}_{\mathcal{B}} = x \vee (\neg x)$. The binary relation $\prec_{\mathcal{B}}$ defined on B by $x \prec_{\mathcal{B}} y$ iff $x \wedge y = x$ defines obviously a partial ordering in B, moreover, $\vee$ and $\wedge$ are the supremum and the infimum operation with respect to $\prec_{\mathcal{B}}$. Obviously, $\mathbf{0}_{\mathcal{B}} \prec_{\mathcal{B}} x \prec_{\mathcal{B}} \mathbf{1}_{\mathcal{B}}$ holds for each $x \in B$. For every finite subset $C \subset B$ the supremum $\bigvee_{x \in C} x$ ($\bigvee C$, abbreviately) and the infimum $\bigwedge_{x \in C} x$ ($\bigwedge C$, abbreviately) with respect to $\prec_{\mathcal{B}}$ are uniquely defined by recursion. If $\bigvee C$ and $\bigwedge C$ are defined for all $C \subset B$, the Boolean algebra $\mathcal{B}$ is called *complete*.

In what follows, we shall focus our attention to the three following mutually isomorphic Boolean algebras.

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 $-$ be the set-theoretic operations of union, intersection and complement. Then the quadruple $\mathcal{B}_0 = \langle \mathcal{P}(\mathcal{N}^+), \cup, \cap, - \rangle$ is obviously a complete Boolean algebra.

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 $1^\infty = \langle 1, 1, 1, \ldots \rangle$ 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 $\boldsymbol{x} \wedge_1 \boldsymbol{y} = \langle \inf\{x_i, y_i\} \rangle_{i=1}^\infty$ for each $\boldsymbol{x}$, $\boldsymbol{y} \in B_1$; here sup and inf are the usual supremum and infimum operations in $\{0,1\}$. 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 $\mathcal{B}_1 = \langle \{0,1\} \rangle_{i=1}^\infty, \vee_1, \wedge_1, 1^\infty - \cdot \rangle$ is a complete Boolean algebra with the zero element $\mathbf{0}_{\mathcal{B}_1} = 0^\infty$ and the unit element $\mathbf{1}_{\mathcal{B}_1} = 1^\infty$. The Boolean algebras $\mathcal{B}_0$ and $\mathcal{B}_1$ are evidently 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 $[0,1]$ of (standard) real numbers. Let $\mathcal{C}$ be the well-known Cantor subset of $[0,1]$. 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. Hence, the mapping $\varphi_0 : \{0,1\}^\infty \to \mathcal{C}$ 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 $\varphi : \mathcal{P}(\mathcal{N}^+) \to \mathcal{C}$ defined by

$$\varphi(A) = \varphi_0(\chi(A)) = \sum_{i=1}^{\infty} 2\chi(A)_i 3^{-i} \tag{2.2}$$

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

Set, for each α, $\beta \in \mathcal{C}$,

$$\begin{aligned} \alpha \vee_2 \beta &= \varphi\left(\varphi^{-1}(\alpha) \cup \varphi^{-1}(\beta)\right), \\ \alpha \wedge_2 \beta &= \varphi\left(\varphi^{-1}(\alpha) \cap \varphi^{-1}(\beta)\right), \\ 1 \dot{-} \alpha &= \varphi\left(\mathcal{N}^+ - \varphi^{-1}(\alpha)\right), \end{aligned} \tag{2.3}$$

in the last row "$-$" denotes the set theoretic operation of complement. An easy calculation yields that $1 \dot{-} \alpha = 1 - \alpha$ holds for each $\alpha \in \mathcal{C}$. The quadruple $\mathcal{B}_2 = \langle \mathcal{C}, \vee_2, \wedge_2, 1 - \cdot \rangle$ is a complete Boolean algebra, $\mathbf{0}_{\mathcal{B}_2} = 0$ 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 defined above.

The following *partial* operation $\sum^* : \mathcal{C}^\infty \to \mathcal{C}$ 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 proved, 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 $[0,1]$. 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 $\pi : \mathcal{N}^+ \to \mathcal{N}^+$.

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 2.1.

(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 $P : \mathcal{A} \to [0,1]$ ascribing to each $A \in \mathcal{A}$ a real number $P(A)$ from the unit interval of reals is called *(standard) probability measure* (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 $\mathcal{A}$ of subsets of Ω) 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 *nonstandard (Cantor-valued) probability measure* (ns.p.m., abbreviately) on $\langle \Omega, \mathcal{A}\rangle$, if (a) $\mu(\Omega) = 1$ and (b) for each sequence$\langle A_1, A_2, \dots\rangle$ 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.* As a matter of fact, the demands imposed on the nonstandard probability measures are rather restrictive, e. g., the possibility that a set possesses the same value of probability measure as its complement is excluded. Nevertheless, some properties of these nonstandard probability measures, namely their extensionality, seem to be interesting enough to justify a more detailed investigation of such measures.

3. BASIC NONSTANDARD PROBABILITY ASSIGNMENTS AND THEIR PROCESSING

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

$$\mu(\Omega - A) = 1 - \mu(A), \quad \mu(A \cup B) = \mu(A) \vee_2 \mu(B), \quad \mu(A \cap B) = \mu(A) \wedge_2 \mu(B), \quad (3.1)$$

where $\vee_2$ and $\wedge_2$ are the binary operations taking $\mathcal{C} \times \mathcal{C}$ into $\mathcal{C}$ defined by (2.3).

It is perhaps worth being mentioned explicitly, that the set A in (3.3) is defined uniquely. It is caused by the fact that the sets $\varphi^{-1}(m^*(A))$ are disjoint for different A's, this property holding true as the value $\sum^* m^*(A)$ is defined.

Proof. The following relation between the operations $\vee_2$ and $\sum_{i=1}^{\infty^*}$ is 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

$$\sum_{i=1}^{\infty^*} \alpha_i = \alpha_1 \vee_2 \alpha_2 \vee_2 \cdots \vee_2 \alpha_n \tag{3.2}$$

holds and we will use the notation $\sum_{i=1}^{n^*} \alpha_i$ to abbreviate the right-hand side expression in (3.2).

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 $\bigcup_{i=1}^{\infty} E_i = A \cup B$. Hence,

$$\begin{aligned} & \mu(A \cup B) = \sum_{i=1}^{\infty^*} \mu(E_i) = \mu(E_1) \vee_2 \mu(E_2) \vee_2 \mu(E_3) && (3.3)\\ = \; & \mu(A - B) \vee_2 \mu(A \cap B) \vee_2 \mu(B - A) \\ = \; & \varphi(\varphi^{-1}(\mu(A - B)) \cup \varphi^{-1}(\mu(A \cap B))) \vee_2 \mu(B - A) \\ = \; & \varphi\left[\varphi^{-1}(\varphi(\varphi^{-1}(\mu(A - B)) \cup \varphi^{-1}(\mu(A \cap B)))) \cup \varphi^{-1}(\mu(B - A))\right] \\ = \; & \varphi\left[(\varphi^{-1}(\mu(A - B)) \cup \varphi^{-1}(\mu(A \cap B))) \cup (\varphi^{-1}(\mu(A \cap B)) \cup \varphi^{-1}(\mu(B - A)))\right] \\ = \; & \varphi\left[\varphi^{-1}(\mu(A)) \cup \varphi^{-1}(\mu(B))\right] = \mu(A) \vee_2 \mu(B). \end{aligned}$$

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

$$\begin{aligned} & \mu(A \cap B) = \mu(\Omega - ((\Omega - A) \cup (\Omega - B))) && (3.4)\\ = \; & \varphi(\mathcal{N}^+ - \varphi^{-1}(\mu(\Omega - A) \vee_2 \mu(\Omega - B))) \\ = \; & \varphi(\mathcal{N}^+ - \varphi^{-1}[\varphi[\varphi^{-1}(\mu(\Omega - A)) \cup \varphi^{-1}(\mu(\Omega - B))]]) \\ = \; & \varphi(\mathcal{N}^+ - [\mathcal{N}^+ - (\varphi^{-1}(\mu(A)) \cap \varphi^{-1}(\mu(B)))]) \\ = \; & \varphi(\varphi^{-1}(A) \cap \varphi^{-1}(B)) = \mu(A) \wedge_2 \mu(B). \end{aligned}$$

□

Definition 3.1. 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 $i \in \mathcal{N}^+$, $i \geq s$. If this is the case, i.e., if $\sum_{A \subset S}{}^* m^*(A)$ is defined, then obviously $\sum_{i=1}^{\infty^*} \alpha_i = \sum_{i=1}^{s}{}^* m^*(A_i) = m^*(A_1) \vee_2 m^*(A_2) \vee_2 \cdots \vee_2 m^*(A_s)$. 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 chosen ordering $\langle A_1, A_2, \ldots, A_s \rangle$ of all subsets of S.

Theorem 3.2. 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 $A \subset S$,

$$m^*(A) = \mu\left(\{\omega \in \Omega : U_{m^*}(\omega) = A\}\right). \tag{3.5}$$

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 $\mu(\{i\}) = 2 \cdot 3n^{-i}$ for all $i \in \mathcal{N}^+$. Consequently, $\langle \Omega, \mathcal{A}, \mu \rangle = \langle \mathcal{N}^+, \mathcal{P}(\mathcal{N}^+), \mu \rangle$ is a ns. probability space. Indeed, let $A_1, A_2, \ldots \subset \mathcal{N}^+$ be an infinite sequence of mutually disjoint subsets of $\mathcal{N}^+$. Then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{j \in \bigcup_{i=1}^{\infty} A_i} 2 \cdot 3^{-j} = \sum_{i=1}^{\infty} \sum_{j \in A_i} 2 \cdot 3^{-j} \tag{3.6}$$
$$= \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{*\,\infty} \mu(A_i).$$

Let us recall the one-to-one mapping $\varphi : \mathcal{P}(\mathcal{N}^+) \to \mathcal{C}$ (the Cantor set) defined by (2.2) above. For each $\boldsymbol{x} \in \mathcal{C}$, $\boldsymbol{x} \leftrightarrow \langle x_1, x_2, \ldots \rangle \in \{0, 2\}^{\infty}$,

$$\varphi^{-1}(\boldsymbol{x}) = \{i \in \mathcal{N}^+; x_i = 2\}, \tag{3.7}$$

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

$$U_{m^*}(\omega) = A \subset \mathcal{N}^+ \quad \text{iff} \quad \omega \in \varphi^{-1}(m^*(A)). \tag{3.8}$$

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

$$\mu\left(\{\omega \in \Omega : U_{m^*}(\omega) = A\}\right) = \mu\left(\{i \in \mathcal{N}^+ : i \in \varphi^{-1}(m^*(A))\}\right) \tag{3.9}$$
$$= \mu\left(\varphi^{-1}(m^*(A))\right) = \sum_{i \in \varphi^{-1}(m^*(A))} 2 \cdot 3^{-i} = m^*(A)$$

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. Interesting and perhaps important enough, the obtained combination rule conserves the extensional nature of the classical Dempster combination rule, but no assumption concerning the statistical independence (or a special kind and/or degree of dependence) of the random variables in question is needed.

Theorem 3.3. 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 $X_i : \langle \Omega, \mathcal{A}, \mu \rangle \to \langle E_i, \mathcal{E}_i \rangle$, $i = 1, 2$, be measurable mappings (generalized random variables), let $\rho_i : S \times E_i \to \{0, 1\}$, $i = 1, 2$, be compatibility relations over the corresponding spaces. Let the mappings $U_i : \Omega \to \mathcal{P}(S)$ defined, for each $\omega \in \Omega$ and for both $i = 1, 2$, by

$$U_i(\omega) = \{s \in S : \rho_i(s, X_i(\omega)) = 1\} \tag{3.10}$$

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

$$\rho_{12}(s, \langle x_1, x_2 \rangle) = \min\{\rho_1(s, x_1), \rho_2(s, x_2)\} \tag{3.11}$$

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

$$U_{12}(\omega) = \{s \in S : \rho_{12}(s, \langle X_1(\omega), X_2(\omega)\rangle) = 1\} \tag{3.12}$$

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

$$m_i^*(A) = \mu\left(\{\omega \in \Omega : U_i(\omega) = A\}\right). \tag{3.13}$$

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

$$m_{12}^*(A) = \sum\nolimits^*_{B,C\subset S,\, B\cap C = A} m_1^*(B) \wedge_2 m_2^*(C) \tag{3.14}$$

holds for each $A \subset S$, where $\wedge_2$ is the nonstandard infimum operation in $\mathcal{C}$ defined by (2.3).

Proof. For $i = 1,2$, $\{\omega \in \Omega : U_i(\omega) = A\} \in \mathcal{A}$ holds for each $A \subset S$ and both $i = 1,2$. Consequently, $\mu(\{\omega \in \Omega : U_i(\omega) = A\})$ is defined. If $A_1, A_2 \subset S$, $A_1 \neq A_2$, then

$$\{\omega \in \Omega : U_i(\omega) = A_1\} \cap \{\omega \in \Omega : U_i(\omega) = A_2\} = \emptyset \tag{3.15}$$

holds 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 (3.13) are ns. b.p.a.'s over S.

As in the usual case, (3.11) and (3.12) yield that

$$\begin{aligned} & U_{12}(\omega) = \{s \in S : \min\{\rho_1(s, X_1(\omega)),\ \rho_2(s, X_2(\omega))\} = 1\} \\ = \; & U_1(\omega) \cap U_2(\omega). \end{aligned} \tag{3.16}$$

For each $A \subset S$

$$\begin{aligned} & \{\omega \in \Omega : U_{12}(\omega) = A\} = \{\omega \in \Omega : U_1(\omega) \cap U_2(\omega) = A\} \\ = \; & \bigcup\nolimits_{B,C\subset S,\, B\cap C = A} \left(\{\omega \in \Omega : U_1(\omega) = B\} \cap \{\omega \in \Omega : U_2(\omega) = C\}\right). \end{aligned} \tag{3.17}$$

The relation

$$\sum\nolimits^*_{A\subset S} \mu\left(\{\omega \in \Omega : U_{12}(\omega) = A\}\right) = \sum\nolimits^*_{A\subset S} m_{12}^*(A) = 1 \tag{3.18}$$

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,

$$\begin{aligned} & m_{12}^*(A) = \mu(\{\omega \in \Omega : U_{12}(\omega) = A\}) \\ = \; & \mu\Big(\bigcup\nolimits_{B,C\subset S,\, B\cap C = A} (\{\omega \in \Omega : U_1(\omega) = B\} \cap \{\omega \in \Omega : U_2(\omega) = C\})\Big) \\ = \; & \sum\nolimits^*_{B,C\subset S,\, B\cap C = A} \mu\left(\{\omega \in \Omega : U_1(\omega) = B\} \cap \{\omega \in \Omega : U_2(\omega) = C\}\right) \\ = \; & \sum\nolimits^*_{B,C\subset S,\, B\cap C = A} \mu\left(\{\omega \in \Omega : U_1(\omega) = B\}\right) \wedge_2 \mu(\{\omega \in \Omega : U_2(\omega) = C\}) \\ = \; & \sum\nolimits^*_{B,C\subset S,\, B\cap C = A} m_1^*(B) \wedge_2 m_2^*(C) \end{aligned} \tag{3.19}$$

due to Theorem 3.1 and due to the definition of $m_1^*(B)$, $m_2^*(C)$ by (3.13). The assertion is proved. □

In the list of references below, [1] and [7] are already classical sources offering an introduction into the field of the Dempster–Shafer theory. The monograph [6] describes and analyses Dempster–Shafer theory from the probabilistic point of view in more detail than the already published papers [4] and [5].

(Received July 10, 2001.)

REFERENCES

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

[2] R. Faure and E. Heurgon: Structures Ordonnées et Algébres de Boole. Gauthier–Villars, Paris 1971.

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

[4] I. Kramosil: Believeability and plausibility functions over infinite sets. Internat. J. Gen. Systems *23* (1994), 2, 173–198.

[5] I. Kramosil: A probabilistic analysis of Dempster combination rule. In: The Logica Yearbook 1997, Prague 1997, pp. 175–187.

[6] I. Kramosil: Probabilistic Analysis of Belief Functions. Kluwer Academic / Plenum Publishers, New York – Boston – Dordrecht – London – Moscow 2001.

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

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

RNDr. Ivan Kramosil, DrSc., Institute of Computer Science – Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 182 07 Praha 8. Czech Republic.
e-mail: kramosil@cs.cas.cz