Denis de Brucq[1]

[1]Academie des Sciences, Belles-Lettres et Arts de Rouen Rouen, France.

General Fusion Operators from Cox's Postulates

Published in:
Florentin Smarandache & Jean Dezert (Editors)
Advances and Applications of DSmT for Information Fusion
(Collected works), Vol. I
American Research Press, Rehoboth, 2004
ISBN: 1-931233-82-9
Chapter XI, pp. 243 - 261

Abstract: *This chapter presents new important links between the most important theories developed in literature for managing uncertainties (i.e. probability, fuzzy sets and evidence theories). The Information fusion introduces special operators* $\circ$ *in the probability theory, in the fuzzy set theory and in the theory of evidence. The mathematical theory of evidence and the fuzzy set theory often replace probabilities in medicine, economy and automatics. The choice between these three quite distinct theories depends on the intrinsic nature of the data to combine. This chapter shows that same four postulates support actually these apparently distinct theories. We unify these three theories from the four following postulates: non-contradiction, continuity, universality, context dependence and prove that a same functional equation is supported by probability theory, evidence theory and fuzzy set theories. In other words, the same postulates applied on confidences, under different conditions, either in the dependence or independence situation, imply the same foundation for the various modern theories of information fusion in the framework of uncertainty by using deductions that we have unified. The independence between elementary confidences have not to be understood in the sense of probabilistic meaning.*

11.1 About uncertainty

In medical fields as in economics and control, one notes the limitation of the additive probabilities due to the too strong constraints imposed. The modification of basic axioms to overcome these limitations leads to different numerical theories and one finds approaches such as fuzzy set theory. By considering the notion of lower probabilities and upper probabilities, one obtains the credibility and the plausibiliy functions of Dempster-Shafer's theory of evidence [6]. The 60's has seen the development of theories that are not directly linked to probabilities. For instance, Zadeh invented fuzzy set theory in 1965 [15]; he then created the possibility theory in 1978 [16].

With the four postulates, which are the basis of the machines on confidences without adding the additivity postulate that leads to probabilities and by considering the independence of the achievement of these confidences, we obtain the fuzzy set theory.

In fact, we have observed that both basic equalities of information fusion are two continuous, commutative and associative operations on confidences. Let Θ be a discrete body of evidence called frame of discernment. Thus, both combinations can be written in terms of probabilities:

$$\forall A, B \subset \Theta, \qquad P(A \cap B) \triangleq P(A)\ P(B/A) \triangleq P(B)\ P(A/B)$$

and in term of membership functions:

$$\forall A, B \subset \Theta \longrightarrow \mu_{A\cap B}(x) \triangleq \mu_A(x) \wedge \mu_B(x)$$

These two operations had to verify the same basic postulates required to model data fusion.

When analyzing imprecise and uncertain data, all the usual techniques must be changed. It is a fact that logic is only an abstract construction for reasoning and physical laws are only models of material system evolutions. Nothing proves that logic can describe correctly all fusions. Moreover, imprecise and uncertain analyses as in this chapter show that an infinity of fusions are possible. From the principles of this chapter, it is possible to introduce a fusion denoted by the operator $\circ$ with any increasing function from $[0, 1]$ onto $[0, 1]$. More precisely, with two beliefs x, y instead of the product $x * y$ to describe the fusion we write $x \circ y$. For example instead of the probability $P(A \cap B) = P(A)P(B)$ of the intersection $A \cap B$ of two independent sets A, B, we write the belief $[A \textit{ and } B/e] = [A/e] \circ [B/e]$, the fusion $\circ$ of the two beliefs $[A/e]$ and $[B/e]$. Any equation of this book may be changed with this transformation.

Moreover, the hypothesis that the sum of masses of disjoint sets is equal to 1 is a global hypothesis and seems to be hazardous.

We demonstrate that the fusion operation $\circ$ is mainly described by a simple product after transformation. This previous transformation of confidence $c(A) = [A/e]$ on A in the environment e is made by using a continuous and strictly monotone function w. This result is easily understood by comparing the transformation w with the Fourier transformation. The latter transforms the composition product of two functions into the product of their Fourier transform. We observe that convolution is commutative and associative. Similarly, Demspster-Shafer fusion is also commutative and associative. Communality of a fusion is the simple product of the communalities of the sources. Without commutativity or associativity other developments are necesary.

11.1.1 Probabilistic modelling

The probability theory has taken a leap during the 17^{th} century with the study of games for luck calculus. The ultimate objective of probability theory is the study of laws governing the random phenomena, that is the presence of uncertainty. For many years, probabilistic methods have generated many debates, in particular among defenders of the frequentist approach, the objective approach and the subjective approaches. Historically, the formulation of the axiomatic basis and the mathematical foundation of the theory are due to Andreï Kolmogorov in 1933.

Let an uncertain experiment be described by the sample space Ω whose elements, denoted ω are the possible results of that experiment. Let $A \in \mathcal{P}(\Omega)$ be subset of Ω. The subset A is a random *event for this theory* and the event is said to occur when the result ω of the experiment belongs to A. The collection of all the subsets of Ω, $\mathcal{P}(\Omega)$, cannot always be associated to the set $\mathcal{A}$ of possible random events in Ω. For logical coherence purposes, one restricts $\mathcal{A}$ to a σ-algebra, a subset of $\mathcal{P}(\Omega)$ which is closed under countable union and under complement. Thus, the pair $(\Omega, \mathcal{A})$ is a measurable space and a probability measure P over $(\Omega, \mathcal{A})$ is then a positive real-valued function of sets with values in $[0, 1]$ and defined over $\mathcal{A}$.

Definition 1. *A probability measure P over $(\Omega, \mathcal{A})$ is an application of $\mathcal{A}$ with values in $[0, 1]$ satisfying the following axioms (Kolmogorov's axioms): i) For all $A \in \mathcal{A}$*

$$0 \leq P(A) \leq 1 \text{ and } P(\Omega) = 1 \tag{11.1}$$

ii) (additivity) For any finite family $\{A_i, i \in I\}$ of mutually exclusive events, we have:

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i) \tag{11.2}$$

iii) sequential monotonic continuity in $\emptyset$ For any sequence $\{A_n, n \geq 1\}$ of events decreasing to the empty set $\emptyset$ that is $A_1 \supset A_2 \supset A_3 \supset ...$ and $\cap A_n = \emptyset$, we have

$$\lim_n P(A_n) = 0 \tag{11.3}$$

$P(A)$ characterizes the *probability* that the event A occurs. If P is a probability measure on $(\Omega, \mathcal{A})$, the triple $(\Omega, \mathcal{A}, P)$ is a *probability space*. From the previous axioms, one easily deduces the following properties:

$$A_1 \subseteq A_2 \Longrightarrow P(A_1) \leq P(A_2), \tag{11.4}$$

$$P(\emptyset) = 0, \tag{11.5}$$

$$P(A) = 1 - P\left(A\right), \tag{11.6}$$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2). \tag{11.7}$$

The *conditional probability* is one of the most useful notions in probability theory. In practice, it is introduced to allow reasoning on events of a referential. For instance, in the case of an exhaustive draw, it is concerned with the probability of an event A, under the condition that an event E occurs. The random event E represents the environment that is usually expressed as $E = e$. There is no reason for having symmetry between event A and the environment e.

Definition 2. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, the conditional probability $P(A/E)$ of an event A given E such that $P(E) > 0$ is defined as:*

$$P(A/E) = \frac{P(A \cap E)}{P(E)}. \tag{11.8}$$

If $P(E) = 0$, this definition has no sense. If $A \subset E$ then $P(A/E) = \frac{P(A)}{P(E)}$, and one has $P(E/E) = 1$.

Obviously, the conditional probability $P(A/E)$ will be seen as the probability of A when E becomes the certain event following additional information asserting that E satisfies to $(P(E) = 1)$.

The equation (11.8) is generalized by using the well known *Bayes' theorem*. If one considers an event E of which we can estimate, *a priori*, the probability $(P(E) \neq 0)$ and a finite partition $\{H_1, ..., H_n\}$ of Ω (set of mutually exclusive hypotheses describing n modalities of the realization of E). The Bayes' formula then yields:

$$P(H_i/E) = \frac{P(E/H_i)\, P(H_i)}{\sum\limits_{j=1}^{n} P(E/H_j)\, P(H_j)}. \tag{11.9}$$

The conditional probabilities (11.9) allow the modification of the *a priori* probability of event H_i, according to the new knowledge on the realization $E = e$.

Definition 3. *Let $(\Omega, \mathcal{A}, P)$ be a probability space and let A and E be two events of $\mathcal{A}$. The events A and E are two independent events if and only if*

$$P(A \cap E) = P(A)\, P(E). \tag{11.10}$$

Property 1. *Let $(\Omega, \mathcal{A}, P)$ be a probability space and let A and E, two events of $\mathcal{A}$. If $P(E) > 0$, then A and E are two independent events if and only if*

$$P(A/E) = P(A). \tag{11.11}$$

Thus, if A and E are two independent events and if E is not impossible then the probability of A is not modified if one receives information on E being realized.

11.1.2 The mathematical theory of evidence

The evidence theory or Dempster-Shafer's theory (DST) of belief functions was born during a lecture on inference statistics given by Arthur Dempster at Harvard University during the 60's. Dempster's main idea has been reinterpreted by Glenn Shafer in his book entitled "*A Mathematical Theory of Evidence*" [12].

Let us consider two spaces Ω and Θ, and a multivalued relation Γ associating the subset $\Gamma(\omega) \subset \Theta$ to each element $\omega \in \Omega$. Let assume that P is a probability measure defined on $(\Omega, \mathcal{A})$ made of the σ-algebra $\mathcal{A}$ of the subsets of Ω. Considering that P represents the probability of occurrence of an uncertain event $\omega \in \Omega$, and if it is established that this event ω is in correspondence with the events $\theta \in \Gamma(\omega)$, what probability judgment can we make about the occurrence of uncertain events $\theta \in \Theta$?

Dempster's view is that the above consideration leads to the concept of *compatible probability* measures. He then refers to the envelope delimited by the lower probability and upper probability of this probability family.

The probability space $(\Omega, \mathcal{A}, P)$ is the information source which allows the quantification of the (imperfect) state of knowledge over the new referential Θ by means of Γ.

In this study, $(\Omega, P, \Gamma, \Theta)$ is called *belief structure*. By using these mathematical tools, Shafer has proposed another interpretation to Dempster's work. This new interpretation identifies the lower and upper probabilities of the family of compatible measures of probability as authentic confidence measures.

Definition 4. *Let Θ be a finite space and 2^Θ $(= \mathcal{P}(\Theta))$ the power set of Θ. A **credibility** function*[1] *Cr is an application of 2^Θ with values in $[0, 1]$ which satisfies the following conditions :*

(i) $Cr(\emptyset) = 0$,

(ii) $Cr(\Theta) = 1$,

[1]The belief function Cr is denoted Bel in [12]

(iii) For all integer n and all family of subsets $A_1, ..., A_n$ of Θ

$$Cr\left(A_1 \cup \ldots \cup A_n\right) \geq \sum_{\substack{I \subset \{1,\ldots,n\} \\ I \neq \emptyset}} (-1)^{|I|+1} Cr(\cap_{i \in I} A_i) \tag{11.12}$$

The condition (iii) is called the general suradditivity condition. When $n = 2$, (12) becomes,

$$Cr\left(A_1 \cup A_2\right) \geq Cr\left(A_1\right) + Cr\left(A_2\right) - Cr\left(A_1 \cap A_2\right). \tag{11.13}$$

The credibility function allows to quantify the partial information in Θ. In theory, other functions are associated to Cr, which are equivalent to it:

- The plausibility function, dual to the credibilities.
- The elementary probability mass function (also called basic belief assignment or mass function) which is obtained from the credibility function by means of the Möbius transform.

Definition 5. *The basic belief assignment is the function $m : 2^{\Theta} \longrightarrow [0,1]$, that satisfies the following property*

$$\sum_{A \in 2^{\Theta}} m\left(A\right) = 1 \tag{11.14}$$

with

$$m\left(\emptyset\right) = 0. \tag{11.15}$$

The evidence theory is often described as a generalization of probabilistic methods to the treatment of uncertainty as it can handle events which are not necessarily exclusive.

Hence the advantage of being able to represent explicitly the uncertainty from imprecise knowledge. The human being easily handled imprecise knowledge. For example, it does not indicate his age to the day near, or his height to the inch near, even if it has access to sufficient information. A mathematical formulation of the imprecisions has come from Lofti Zadeh through the fuzzy set theory [15]. The modelling of uncertainties due to the imprecisions of knowledge gives rise to possibility theory that constitutes with the fuzzy set theory the general framework of the fuzzy logic.

11.1.3 Fuzzy logic

The fuzzy logic appeared in 1965 with Lofti Zadeh's work. The development of the fuzzy logic was mainly motivated by the need for a conceptual framework that can address the issue of uncertainty and lexical imprecision. From this work, it is necessary to keep the need of formalizing the representation and the processing of imprecise or approximate knowledge with the intention to treat systems with a strong complexity, in which human factors are often present. Thus, fuzzy logic intervenes to deal with imperfect knowledge.

The fuzzy logic is based on two main subject matters [9]: fuzzy set theory and modelling of approximate reasoning in the framework of possibility theory.

The definition of a fuzzy subset answers the need to represent imprecise knowledge. The concept was introduced to avoid abrupt changes of a class to another(black to the white, for example) and to authorize elements so that they cannot belong completely either to one of the classes or to another (to be gray in the example). In a reference set Θ, a fuzzy subset of Θ is characterized by a membership function μ w.r.t. A, defined as:

$$\mu_A : \Theta \longrightarrow [0,1]$$

which is the extension of the classical membership function χ, indicator function of the set A that is:

$$\chi_A : \Theta \longrightarrow \{0,1\} .$$

To emphasize the difference with the ordinary sets of Θ, we use lower case letters for the fuzzy sets of Θ.

Definition 6. *Let a be a fuzzy set of Θ and let α be a real value in $[0,1]$. The $\alpha-cut$ a_α is the subset of Θ defined by:*

$$a_\alpha \triangleq \{\theta \in \Theta;\ \mu_a(\theta) \geq \alpha\} . \tag{11.16}$$

Then $\forall \alpha, \beta \in [0,1]$,

$$\alpha \leq \beta \Longrightarrow a_\beta \subseteq a_\alpha$$

and $\forall \theta \in \Theta$,

$$\mu_a(\theta) = \sup\{\alpha \in [0,1];\ \theta \in a_\alpha\} . \tag{11.17}$$

This allows the passage from the fuzzy sets to ordinary sets and gives immediately the fuzzy versions of the usual operations used for ordinary sets.

Property 2. *Let a and b be two fuzzy sets of Θ defined by their membership functions μ_a and μ_b, one has:*

- *equality*: $a = b \Longleftrightarrow \forall \theta \in \Theta, \quad \mu_a(\theta) = \mu_b(\theta)$
- *inclusion*: $A \subseteq b \Longleftrightarrow \forall \theta \in \Theta, \quad \mu_a(\theta) \leq \mu_b(\theta)$
- *union*: $a \cup b \longleftrightarrow \forall \theta \in \Theta, \quad \mu_{a\cup b}(\theta) = \max(\mu_a(\theta), \mu_b(\theta))$
- *intersection*: $a \cap b \longleftrightarrow \forall \theta \in \Theta, \quad \mu_{a\cap b}(\theta) = \min(\mu_a(\theta), \mu_b(\theta))$
- *complement*: $a \longleftrightarrow \forall \theta \in \Theta, \quad \mu_a(\theta) = (1 - \mu_a(\theta))$

The uncertainties about the truth of a statement are not verified in the case of the fuzzy set theory.

The possibility theory was introduced in 1978 by Lofti Zadeh in order to manipulate non-probabilistic uncertainties for which the probability theory does not give any satisfactory solution. The possibility theory provides a framework in which imprecise knowledge and uncertain knowledge can coexist and can be treated jointly.

Possibility theory provides a method to formalize subjective uncertainties on events. It informs us in which measure the realization of an event is possible and in which measure we are sure without having any evaluation of probabilities at our disposal. One presents the possibility theory in a general form that introduces the concepts of possibility measure and necessity measure.

Consider either the frame Ω (experiment space) or Θ (space of hypotheses). Set $\mathcal{A}$, a family of subsets of Ω or subsets of Θ. When Ω or Θ are finite then $\mathcal{A}$ is the set of all subsets.

Definition 7. *A possibility measure Pos is an application of $\mathcal{A} \subset \mathcal{P}(\Theta)$ in $[0,1]$ such that:*
i) $Pos\,(\emptyset) = 0,\ Pos\,(\Theta) = 1.$
ii) for any finite family $\{A_i, i \in I\}$ of events, one has:

$$Pos\left(\bigcup_i A_i\right) = \sup_i \{Pos\,(A_i)\}\,. \tag{11.18}$$

According to Zadeh, this is the most pessimistic notion or the most prudent notion for a belief. One has in particular:

$$\max\left(Pos\,(A)\,, Pos\,\left(A\right)\right) = 1 \tag{11.19}$$

and then:

$$Pos\,(A) + Pos\,\left(A\right) \geq 1. \tag{11.20}$$

11.1.4 Confidence measures

Definition : A confidence measure c is an application of $\mathcal{P}\,(\Theta)$, parts of Θ, in $[0,1]$ which verifies the following properties:

i) $c\,(\emptyset) = 0$ and $c\,(\Theta) = 1$

ii) (monotony) $\forall A, B \in \mathcal{P}\,(\Theta)\,, \qquad A \subset B \Longrightarrow c\,(A) \leq c\,(B)$

iii) (continuity) For all increasing or decreasing sequences $(A_n)_{\mathbb{N}}$ of elements of $\mathcal{P}\,(\Theta)$, one has :

$$\lim c\,(A_n) = c\,(\lim A_n)\,.$$

Consequently, one has: $\forall\, A, B \in \mathcal{P}(\Theta)$,

$$c(A \cap B) \le \min(c(A), c(B)) \quad \text{and} \quad \max(c(A), c(B)) \le c(A \cup B).$$

The probabilities, the fuzzy sets, the possibility measures are special cases of the general notion of confidence measures.

11.2 Fusions

As with Physics, the information fusion modelling aims at giving the best possible description of the experimental reality. Let us give the postulates [14] that information fusions need to satisfy.

11.2.1 Postulates

1. **Coherence or noncontradiction**
2. **Continuity of the method**
3. **Universality or completeness**
4. **No information refusal**

A first consequence is that postulates 2 and 3 leads to use real numbers to represent and compare degrees of confidence. However postulate 4 leads to hypothetical conditioning: the confidence degree is only known conditionally upon the environment, the context.

The confidence granted to event $A \in \mathcal{P}(\Theta)$ in the environment e is noted $[A/e]$.

From Edwin Thompson Jaynes [10]: *Obviously, the operation of real human brains is so complicated that we can make no pretense of explaining its mysteries; and in any event we are not trying to explain, much less reproduce, all the aberrations and inconsistencies of human brains. To emphasize this, instead of asking, "How can we build a mathematical model of human common sense?" let us ask, "How could we build a machine which would carry out useful plausible reasoning, following clearly defined principles expressing an idealized common sense?"*

11.2.2 Machine on confidence

We develop the approach essentially based on Cox's work [5] later detailed by Tribus [14] while criticized.

$$i = impossible = 0 \le [A/e] \le c = certain = 1$$

The various possible relations are listed by setting $u \triangleq [A \wedge B/e]$ that expresses the confidence provided by the fusion of A and B within the environment e. Let's define:

$$x \triangleq [A/e] \qquad v \triangleq [A/Be] \qquad y \triangleq [B/e] \qquad w \triangleq [B/Ae]$$

Eleven functional relations are possible: $u = F_1(x,v)$, $u = F_2(x,y)$, $u = F_3(x,w)$, $u = F_4(v,y)$, $u = F_5(v,w)$, $u = F_6(y,w)$, $u = F_7(x,v,y)$, $u = F_8(x,v,w)$, $u = F_9(x,y,w)$, $u = F_{10}(v,y,w)$ and $u = F_{11}(x,v,y,w)$

Because of the postulates, the functions F_5, F_8, F_{10} and F_{11} have to be discarded. The symmetries induce simplifications. The functional relations capable to meet the aspirations, are:

$$\begin{aligned} u &= F_2(x,y) = F_2(y,x) \\ u &= F_3(x,w) = F_4(v,y) \\ u &= F_7(x,v,y) = F_9(x,y,w) \end{aligned}$$

The associativity condition on the fusion confidence

$$[A \wedge B \wedge C/e] = [A \wedge (B \wedge C)/e] = [(A \wedge B) \wedge C/e]$$

discards F_7.

On the other hand, F_3 et F_2 verify the same *associativity* equation. By calling $\circ$ the common operation describing all the possible fusions between the confidences, *this unique equation processes two different situations*:

- First case: $u = F_2(x,y) = F_2(y,x)$

  $[A \wedge B/e] = [A/e] \circ [B/Ae] = [B/e] \circ [A/Be]$

- Second case: $u = F_3(x,w) = F_4(v,y)$

  $[A \wedge B/e] = [A/e] \circ [B/e]$.

This second case was not considered by Cox, the consequences of which constitutes the first results of this paper.

11.2.3 Operator

- **First case**:

$$[B/Ae] < [B'/Ae] \Longrightarrow [A \wedge B/e] < [A \wedge B'/e].$$

The first case implies strict inequalities on the second variable. The mathematician Aczél [1] has given the proof based on the strict monotony of one of both variables. The general solution for the functional equation being such that:

$$w\left([A \wedge B/e]\right) = w\left([A/e]\right)\; w\left([B/Ae]\right) = w\left([B/e]\right)\; w\left([A/Be]\right) \tag{11.21}$$

where w is a continuous strictly-monotone function of $[0, 1]$ onto $[0, 1]$. Thus,

$$[A \wedge B/e] = w^{-1}(w\left([A/e]\right) w\left([A/Be]\right)) = [A/e] \circ [B/Ae]$$

The fusion operation $\circ$ is described by a simple product of real numbers after transformation. This previous transformation of confidence $c(A) = [A/e]$ on A in the environment e is made by using a continuous and strictly monotone function w. This result is easily understood by comparing the transformation w with the Fourier transformation. The latter transforms the composition product of two functions into the product of their Fourier transform.

The first case with additional properties gives the probability theory. The problem is to know if there is a similar property in the second case.

- **Second case**: The strict monotony is not obvious.

  If $[A/e] \leq [A/e]$ and $[B/e] \leq [B'/e]$ then $[A \wedge B/e] \leq [A' \wedge B'/e]$. On the other hand, one has the commutativity property and $\circ$ has all the characteristics of a triangular norm, common notion in data processing [9]. In this second case, the confidence fusions are associated to the t-norms. The second case implies the fuzzy theory.

11.3 T-norm

Definition: A triangular norm - called t-norm - is a function $\circ : [0, 1] \times [0, 1] \longrightarrow [0, 1]$ that verifies the following conditions for all x, y, z, t in $[0, 1]$

i) (commutativity) $x \circ y = y \circ x$

ii) (associativity) $(x \circ y) \circ z = x \circ (y \circ z)$

iii) (isotony) if $x \leq z$ and $y \leq t$, $(x \circ y) \leq (z \circ t)$

iv) (neutral element 1) $(x \circ 1) = x$

Example 1. *The operator* $\circ = \min$ *is a t-norm; this is the upper t-norm. For all* x, y *in* $[0, 1]$

$$(x \circ y) \leq \min\left(x, y\right)$$

Lemma 1. *If the associated t-norm is strictly increasing, the operator on the confidences is written as follows:* $w\,[A \wedge B/e] = w\,([A/e])\ \ w\,[B/e]$ *where* w *is a continuous and strictly increasing bijection of* $[0,1]$ *onto* $[0,1]$.

According to the additional hypothesis, we retrieve: $[A \wedge B/e] = w^{-1}(w\,([A/e])\,w\,([B/e]))$.

Theorem 1. *The fuzzy operator* $[A \wedge B/e] = [A/e] \wedge [B/e] = \inf\{[A/e]\,,[B/e]\}$ *is the limit of a sequence of strictly monotone operators* $\circ_n$.

Proof: Let $(T_n)_{n>0}$ be the family of strictly monotone t-norms such that:

$$\forall n \geq 1, \quad T_n(x,y) = \frac{1}{1+\sqrt[n]{\left(\frac{1-x}{x}\right)^n + \left(\frac{1-y}{y}\right)^n}} = w_n^{-1}\left(w_n(x)w_n(y)\right) \quad \text{with} \quad w_n = \exp - \left(\frac{1-x}{x}\right)^n.$$

For all $n \geq 1$, w_n is a continuous and strictly increasing bijection of $[0,1]$ onto $[0,1]$. We have for all x, y :

$$\lim_{n\to\infty} T\,(x,y) = \frac{1}{1+\max\left(\left(\frac{1-x}{x}\right),\left(\frac{1-y}{y}\right)\right)}.$$

In fact, if $0 \leq a \leq b$

$$\lim_{n\to\infty} \sqrt[n]{a^n+b^n} = \lim_{n\to\infty} b\left(1+\left(\frac{a}{b}\right)^n\right)^{\frac{1}{n}} = b$$

therefore

$$\lim_{n\to\infty} T\,(x,y) = f^{-1}\left(\max\left(f\,(x)\,,f\,(y)\right)\right)$$

where $f(x) = \frac{1-x}{x}$

$$\max\left(f\,(x)\,,f\,(y)\right) = f\left(\min(x,y)\right)$$

Since f is strictly decreasing on $[0,1]$, it follows that

$$\lim_{n\to\infty} T\,(x,y) = \min(x,y) \qquad \blacksquare.$$

Here are the results obtained for several fusion operators. On x-axis, x increases by 0.1 jumps and equally on y-axis, y increases by 0.1 jumps.

- **Result obtained with the product operator**: $x \circ y \triangleq x * y$

0.0100	0.0200	0.0300	0.0400	0.0500	0.0600	0.0700	0.0800	0.0900	0.1000
0.0200	0.0400	0.0600	0.0800	0.1000	0.1200	0.1400	0.1600	0.1800	0.2000
0.0300	0.0600	0.0900	0.1200	0.1500	0.1800	0.2100	0.2400	0.2700	0.3000
0.0400	0.0800	0.1200	0.1600	0.2000	0.2400	0.2800	0.3200	0.3600	0.4000
0.0500	0.1000	0.1500	0.2000	0.2500	0.3000	0.3500	0.4000	0.4500	0.5000
0.0600	0.1200	0.1800	0.2400	0.3000	0.3600	0.4200	0.4800	0.5400	0.6000
0.0700	0.1400	0.2100	0.2800	0.3500	0.4200	0.4900	0.5600	0.6300	0.7000
0.0800	0.1600	0.2400	0.3200	0.4000	0.4800	0.5600	0.6400	0.7200	0.8000
0.0900	0.1800	0.2700	0.3600	0.4500	0.5400	0.6300	0.7200	0.8100	0.9000
0.1000	0.2000	0.3000	0.4000	0.5000	0.6000	0.7000	0.8000	0.9000	1.0000

- **Result obtained with the operator**: $x \circ_n y = \frac{1}{1+\sqrt[n]{\left(\frac{1-x}{x}\right)^n+\left(\frac{1-y}{y}\right)^n}}$ for $n = 3$.

0.0810	0.0975	0.0995	0.0999	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
0.0975	0.1656	0.1905	0.1973	0.1992	0.1998	0.1999	0.2000	0.2000	0.2000
0.0995	0.1905	0.2538	0.2838	0.2947	0.2984	0.2996	0.2999	0.3000	0.3000
0.0999	0.1973	0.2838	0.3460	0.3794	0.3933	0.3982	0.3996	0.4000	0.4000
0.1000	0.1992	0.2947	0.3794	0.4425	0.4784	0.4937	0.4987	0.4999	0.5000
0.1000	0.1998	0.2984	0.3933	0.4784	0.5435	0.5810	0.5959	0.5996	0.6000
0.1000	0.1999	0.2996	0.3982	0.4937	0.5810	0.6494	0.6872	0.6988	0.7000
0.1000	0.2000	0.2999	0.3996	0.4987	0.5959	0.6872	0.7605	0.7955	0.8000
0.1000	0.2000	0.3000	0.4000	0.4999	0.5996	0.6988	0.7955	0.8772	0.9000
0.1000	0.2000	0.3000	0.4000	0.5000	0.6000	0.7000	0.8000	0.9000	1.0000

As soon as $n = 3$ we observe how near this operator approximates $x \circ y \triangleq \min(x, y)$.

- **Result obtained with the fusion operator**: $x \circ y \triangleq \min(x, y)$

0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
0.1000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000
0.1000	0.2000	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000
0.1000	0.2000	0.3000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000
0.1000	0.2000	0.3000	0.4000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
0.1000	0.2000	0.3000	0.4000	0.5000	0.6000	0.6000	0.6000	0.6000	0.6000
0.1000	0.2000	0.3000	0.4000	0.5000	0.6000	0.7000	0.7000	0.7000	0.7000
0.1000	0.2000	0.3000	0.4000	0.5000	0.6000	0.7000	0.8000	0.8000	0.8000
0.1000	0.2000	0.3000	0.4000	0.5000	0.6000	0.7000	0.8000	0.9000	0.9000
0.1000	0.2000	0.3000	0.4000	0.5000	0.6000	0.7000	0.8000	0.9000	1.0000

It was not obvious to obtain the functions w_n. The fuzzy operator $\circ = \min$ comes from a limit of fusions $\circ_n$, each admitting after confidence transformation $w_n[A/e]$ and $w_n[B/e]$, a decomposition in a conventional product of real numbers.

11.3.1 Independence-interdependence

The second functional relation

$$w([A \wedge B/e]) = w([A/e])\ w([B/e])$$

is discarded if we consider there is a link between the knowledge of two facts in a given environment. This constraint, admitted by Cox then by Tribus, is however not valid for all uncertainty models. Let us give two examples for which the argument given by Tribus is insufficient. In the probability theory, randomly taking of balls with or without replacement leads to two different models. The testimony of different persons is another example. The testimonies can be obtained separately or in a meeting.

Thus, because of the acquisition conditions of the knowledge, the postulates lead to two distinct theories: the probability theory and the fuzzy logic.

In addition, from the four basic postulates explained above and valid for the three theories (probability theory, evidence theory and fuzzy logic), and while adding the hypothesis of interdependence and admitting a postulate of precision leading to the additive rule, one would obtain the probabilities as well as the transition probabilities and therefore the credibilities.

11.3.2 T-norm description

We have also obtained a result characterizing the t-norms by correcting and extending a previous demonstration [11]. This is our third result.

Theorem 2. *Let $\circ$ be a continuous t-norm of $[0,1] \times [0,1] \to [0,1]$. Then, the interval $[0,1]$ is the union*

1. *of closed intervals $[b,c]$ over which the equality $s \circ s = s$ is satisfied and*

2. *of open intervals (a,b) for which $a \circ a = a$ and $b \circ b = b$ and for which the inequality $s \circ s \neq s$ is satisfied.*

   For the intervals $[b,c]$ of first kind : $\forall x \in [b,c]$, $\quad \forall y \in [x,1]$, $\quad x \circ y = x \wedge y$.

   For each second kind interval (a,b) there exists a function w strictly increasing from $[a,b]$ into $[0,1]$ such that $w(b) = 1$

   If $\forall s \in (a,b) \quad s \circ s \neq a \quad$ then $w(a) = 0$ and $\forall x, y \in [a,b] \quad x \circ y = w^{-1}(w(x)\,w(y))$

   If $\exists s \in (a,b) \quad s \circ s = a$ then $w(a) > 0$ and $\forall x, y \in [a,b] \quad x \circ y = w^{-1}(w(x)\,w(y)) \vee a$

On each like-interval (a,b), the operation $\circ$ can be constant when x varies from a. However, the interval within the function is really constant depending upon the value of the second variable y. The separation curve $\{(x,y) \in [a,b] \times [a,b]\,; x \circ y = a\}$ in the space $[a,b] \times [a,b]$ is given by the equality

$$w(x \circ y) = w(a) = w(x)w(y).$$

This theorem results from the lemmas hereafter.

Lemma 2. *The set* $\{x \in [0,1]\,; T(x,x) = x\}$ *is a union of closed intervals of the interval* $[0,1]$.

Any adherence point s of a sequence $(s_n; n \in \mathbb{N}, T(s_n, s_n) = s_n)$ satisfies $T(s,s) = s$ with respect to the continuity of T, and therefore s belongs to the closed interval. Thus, for example, the set $\left\{[0]\,, \left[\frac{1}{3^n}\,, \frac{2}{3^n}\right]; n \in \mathbb{N}\right\}$ constitutes an infinite family of closed intervals. On each of the open intervals of the countable infinity of the complementary set, it is sufficient to define a t-norm by means of a continuous and increasing function w. Each of these functions w depends on the open interval under consideration.

Lemma 3. *If* α *exists in the open interval* $(0,1)$ *such that* $T(\alpha,\alpha) \neq \alpha$ *then there are two real values* a, b *satisfying the inequalities* $0 \leq a < \alpha < b \leq 1$ *as well as the equalities* $T(a,a) = a$ *and* $T(b,b) = b$. *Furthermore, for all real values in the open interval* (a,b), *the inequality* $T(s,s) \neq s$ *is satisfied.*

Lemma 4. *Let* T *be a continuous t-norm. For all pair* (x,y) *of* $[0,1]$ *such that there exists* a, $x \leq a \leq y$ *with* $T(a,a) = a$, *we have:*

$$T(x,y) = x = min(x,y).$$

Any continuous t-norm T coincides over $[0,1] \times [0,1]$ with the min function, except for the points (x,y), $x \leq y$ for which one cannot find a real α such that:

$$x \leq \alpha \leq y \ et \ T(\alpha,\alpha) = \alpha.$$

One has to study the behavior of T in the regions $[a,b] \times [a,b]$ of the intervals $[a,b]$ of the second kind.

Lemma 5. *Consider the associative and commutative operation* $\circ$ *of* $[a,b] \times [a,b] \to [a,b]$ *which is continuous and decreasing with respect to both variables and such that* $a \circ a = a$ *and* $b \circ b = b$ *but such that for all* s *in the open interval* (a,b), *one has the inequality* $s \circ s \neq s$. *Let* u *be in the closed interval* $[a,b]$, *upper bound of* v *such that* $v \circ v = a$, *that is such that* $u \triangleq \sup\{v \in [a,b]\,; v \circ v = a\}$. *The operation* $\circ$ *is strictly increasing for each of both variables wherever* $x \circ y \neq a$, *and if* $u = a$ *then* $\circ$ *is strictly increasing over* $[a,b] \times [a,b]$.

Lemma 6. *Under valid conditions of application of lemma 5, if* $u = a$, *then for all* α *in* (a,b) *and for all nonzero positive rational number* q, *the real power* $\alpha^{\circ q}$ *is defined and is a real number in the* (a,b).

Remark 1. *It can easily be verified that:*

$$\alpha^{\circ\frac{n}{m}} = \alpha^{\circ\frac{rn}{rm}}$$

and thus:

$$\alpha^{\circ\frac{n}{m}} \circ \alpha^{\circ\frac{r}{s}} = \alpha^{\circ\frac{sn}{sm}} \circ \alpha^{\circ\frac{rm}{sm}} = \alpha^{\circ\frac{rm+sn}{sm}} = \alpha^{\circ(\frac{n}{m}+\frac{r}{s})}$$

Lemma 7. *Under valid conditions of application of lemma 5, if $u = a$ the application $q \in \mathbb{Q}_+^\star \to \alpha^{\circ q} \in (a, b)$ is strictly decreasing and satisfies to $\lim_{q\to 0} \alpha^{\circ q} = b$ and $\lim_{q\to\infty} \alpha^{\circ q} = a$.*

Lemma 8. *The application $r \in [0, \infty) \to \alpha^{\circ r} \triangleq \sup\{\alpha^{\circ q}; r < q\}$ is continuous and decreasing and $\alpha^{\circ r} \triangleq \inf\{\alpha^{\circ q}; q < r\}$.*

Lemma 9. *Under valid conditions of application of lemma 5, if $u > a$, one defines the application $r \in [0, \infty) \to u^{\circ r}$ in $[a, b]$ as previously. With $u^{\circ r}$ strictly decreasing over $[0, 2]$ such that $u^{\circ 0} = b$, $u^{\circ 2} = a$, and for all $r \geq 2 \quad u^{\circ r} = a$.*

Lemma 10. *Under valid conditions of application of lemma 5, if $u > a$, one defines for all $\alpha \in (a, b)$, the application $r \in [0, \infty[\to \alpha^{\circ r}$ in $[a, b]$. In this case, there is a positive real number r_0 such that $\alpha^{\circ r}$ is strictly decreasing over $[0, r_0]$, and $\alpha^{\circ 0} = b$, $\alpha^{\circ r_0} = a$, and for all $r \geq r_0 \quad \alpha^{\circ r} = a$.*

Lemma 11. *Consider the associative and commutative operation $\circ$ of $[a, b] \times [a, b] \to [a, b]$ continuous and strictly increasing with respect to both variables such that $a \circ a = a$ and $b \circ b = b$ but one has the inequality $s \circ s \neq s$ for all s in the open interval (a, b). Therefore, there is a continuous and strictly increasing function w such that:*

$$x \circ y = w^{-1}(w(x)w(y)) \vee a = \max(a, w^{-1}(w(x)w(y))) \tag{11.22}$$

The results of the lemmas finish the justification of the theorem 2.

11.4 Conclusions

Finally, the same postulates applied on confidences, in different environments (either in dependence or independence situation), imply the same foundation for the various modern theories of information fusion in the framework of uncertainty by using deductions that we have unified. The independence between elementary confidences does not need to be understood in the probabilistic sense. The formula $P(A/e) = P(A)$ of the probability of A in the environment e has no sense. One has to find another conceptualization of the notion of independence moving away from the probabilistic concept.

We must make new models when fusion analysis is to be applied in all situations. We take the simple example of logical implication

$$P \quad and \quad Q \Rightarrow R$$

Every logical proposition P, Q, R takes only one of the two numerical values $0, 1$. Yet with Probability these propositions are able to take any numerical value in the interval $[0, 1]$ to represent the statistical limit of existence when the experiment is repeated as often as possible. Nowadays, the numbers $[P/e]$, $[Q/e]$ and $[R/e]$ only give the intuitive beliefs when the conditions e on the surroundings are well defined.

To be more explicit, let take a plausible medical situation. Many patients present chaotic neurologic disorders. Does the deterministic chaos P with the drug Q result in the end of illness R?

We have no reason in such a medical situation to introduce the limitation of logical implication. Moreover, we have the fusion "*and*" about the two beliefs $[P/e]$ on the disorder P and $[Q/e]$ on the efficiency of drug Q and we expect this fusion to give precisely the belief $[R/e]$ of the recovery R from the two beliefs $[P/e]$ and $[Q/e]$.

In addition, let us take the discussion of Zadeh's example, discussed in Chapter 5, in order to make a new analysis with our fusion principles. One has the values

$$m(M) = 0 \qquad m(C) = 0 \qquad m(T) = 1$$

(M standing for Meningitis, C for contusion and T for tumor) for the masses from Dempster-Shafer renormalization where the normalization coefficient is

$$1 - m(\emptyset) = 0.0001$$

From our principles, it is possible to give a belief for the global model. Without renormalization the two doctors give the beliefs

$$[T/e]_1 = 0.01 \qquad [T/e]_2 = 0.01$$

With the principles of this chapter, the numerical value for any fusion arising from these two beliefs is equal to or less than $0.01 = \min([T/e]_1, [T/e]_2)$. So the Dempster-Shafer normalization is not a fusion! The normalization is in contradiction with the arguments of this chapter. Note that the hybrid DSm rule of combination proposed in Chapter 1 provides in this example explained in details in Chapter 5 (Section 5.3.1) $Cr(T) = m(T) = 0.0001 \leq \min([T/e]_1, [T/e]_2)$ which is coherent with a confidence measure.

The probable explanation is that the Dempster-Shafer normalization is the only mistake of the model. One supposes global cohesion between initial mass values coming from Demspster-Shafer rules. In mathematics, we know it is often impossible to adjust analytical functions in the whole complex plan $\mathbb{C}$; global cohesion is impossible! For example the logarithmic function is defined in any neighbourhood but it is not defined in the whole complex plan. The global cohesion is probably the mistake. The DSmT framework seems to provide a better model to satisfy confidence measures and fusion postulates. Some theoretical investigations are currently done to fully analyze DSmT in the context of this work.

Another way to explain losses of mass in Dempster-Shafer theory is to introduce new sets. In any probability diffusion, we observe occasionally probability masses loading infinity with an evolution. Let us take the mass 1 in position $\{n\}$ and increase n to infinity we have no more mass on the real line $\mathbb{R}$. Similarly, let us take the masses 0.5 on $\{-n\}$ and 0.5 on $\{n\}$; this time we load $\{-\infty\}$ and $\{\infty\}$, n increasing to infinity. In Dempster-Shafer model, one sometimes loads the empty set $\{\emptyset\}$ and (or) an extra set, only to explain vanishing masses.

Probably Dempster-Shafer renormalization is the only mistake of the model because false global property of masses is supposed. It is important to know the necessary axioms given renormalization truth.

Surroundings are so different that fusion described only by product is certainly a construction that is too restrictive.

The processing in concrete application of the results presented here suppose additional hypotheses, since any information fusion introduces monotone functions strictly increasing whose existence is proven in this paper. These functions (not only one!) remain to be identified for each application. Theoretical considerations should allow to keep certain typical families of functions. Experimental results would next identify some unknown parameters if some parameterized family of such functions.

Applications of such a methodology on the information fusion such as air pollution measures given by sensors will be processed.

Moreover, during its time evolution, the information data fusion can thus be described by successive t-norms amongst which probability should be introduced.

11.5 References

[1] Aczel J., *Lectures on functional equations and their applications*, Academic press New York and London 1966.

[2] Bloch I., *Incertitude, imprécision et additivité en fusion de données : point de vue historique*, Traitement du Signal, 13(4) : pp. 267-288, 1996.

[3] Bouchon-Meunier B., *Théorie des possibilités. Application à la présentation des connaissances en informatique*, Masson, Paris, 2nd édition, 1988.

[4] De Brucq D., Colot O., Sombo A., *Identical Foundation of Probability theory and Fuzzy Set Theory*, IF 2002, 5th International Conference on InformationFusion, Annapolis, Maryland, July 7-11, pp. 1442-1449, 2002.

[5] Cox R.T., *Pobability frequency and reasonable expectation* , American Journal of Physics, 14(1) pp. 1-14,1946.

[6] Dempster A.P., *Upper and lower probabilities induced by a multivalued mapping*, Annals of mathematical statistics, 1967.

[7] Denœux T., *Decision analysis in evidence theoritical pattern classification*,Pattern recognition, 1997.

[8] Dubois D., Prade H., *Théorie des possibilités. Application à la présentation des connaissances en informatique*, Masson, Paris, 2nd édition, 1988.

[9] Gacogne L., *Eléments de logique floue*, Hermes, Paris, 1997.

[10] Jaynes E.T., *Probability theory: The logic of science*, http://omega.albany.edu:8008/JaynesBook.html, 1994.

[11] Paris J.B., *The uncertain reasoner's companion. A mathematical perspective*, Cambridge University Press, 1994.

[12] Shafer G., *A mathematical theory of evidence*, Princeton Ujiversity Press, 1976.

[13] Smets P., *Probability, possibility, belief : Which and where?*, Handbook of defeasible Reasonning and Uncertainty Management Systems, pp. 1-24. 1998.

[14] Tribus M., *Décisions rationnelles dans l'incertain*, Masson et Cie, Paris, 1972.

[15] Zadeh L.A., *Fuzzy sets*, Volume 8 Information and Control, Academic Press, New York, 1965.

[16] Zadeh L.A., *Fuzzy sets as a basis for a theory of possibility*, Fuzzy Sets and Systems, 1: pp. 3-28,1978.