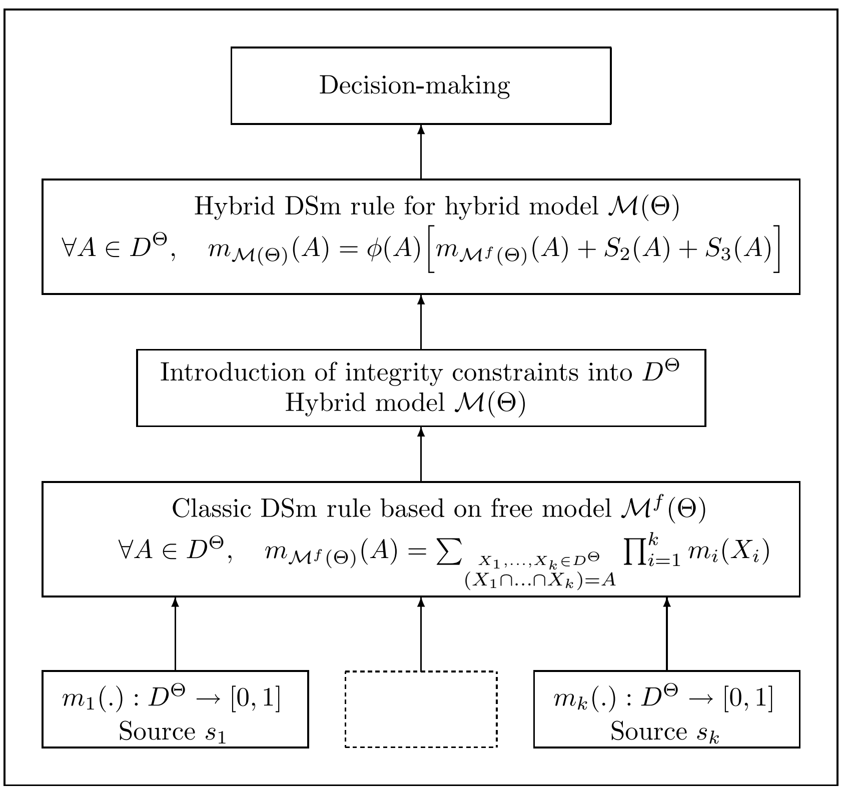
Florentin Smarandache & Jean Dezert

(Editors)

Advances and Applications of DSmT for Information Fusion

(Collected works)



American Research Press

Rehoboth

2004

This book can be ordered in a paper bound reprint from:

Books on Demand

ProQuest Information & Learning

(University of Microfilm International)

300 N. Zeeb Road

P.O. Box 1346, Ann Arbor

MI 48106-1346, U.S.A.

Phone.: 1-800-521-0600 (Customer Service)

`http://wwwlib.umi.com/bod/`

and also be ordered through `Amazon.com` **or downloaded from the web site:**

`http://www.gallup.unm.edu/~smarandache/DSmT-book1.pdf`.

Download books from the Digital Library

`http://www.gallup.unm.edu/~smarandache/eBooks-otherformats.htm`

This book has been peer reviewed and recommended for publication by:

1 - Professor Krassimir Atanassov

Centre of Biomedical Engineering (CLBME)

Bulgarian Academy of Sciences

Bl. 105 Acad. G. Bontchev Str., 1113 Sofia, Bulgaria

2 - Professor Bassel Solaiman

École Nationale Supérieure des Télécommunications de Bretagne

Image and Information Processing Department

Technopôle Brest-Iroise - CS 83818 - 29238 Brest Cedex 3, France

3 - Professor Pierre Valin

Physics Department & Centre de Recherches Mathématiques

University of Montréal

C.P. 6128, succ. Centre-ville, Montréal, H3C 3J7, Canada

ISBN: 1-931233-82-9

Standard Address Number 297-5092

Printed in the United States of America

Contents

Preamble

This book is devoted to an emerging branch of Information Fusion based on new approach for modelling the fusion problematic when the information provided by the sources is both uncertain and (highly) conflicting. This approach, known in literature as DSmT (standing for Dezert-Smarandache Theory), proposes new useful rules of combinations. We gathered in this volume a presentation of DSmT from the beginning to the latest development. Part 1 of this book presents the current state-of-the-art on theoretical investigations while Part 2 presents several applications of this new theory. We hope that this first book on DSmT will stir up some interests to researchers and engineers working in data fusion and in artificial intelligence. Many simple but didactic examples are proposed throughout the book. As a young emerging theory, DSmT is probably not exempt from improvements and its development will continue to evolve over the years. We just want through this book to propose a new look at the Information Fusion problematic and open a new track to attack the combination of information.

We want to thank Prof. Laurence Hubert-Moy, Dr. Anne-Laure Jousselme, Dr. Shubha Kadambe, Dr. Pavlina Konstantinova, Dr. Albena Tchamova, Dr. Hongyan Sun, Samuel Corgne, Dr. Frédéric Dambreville, Dr. Milan Daniel, Prof. Denis de Brucq, Prof. Mohamad Farooq, Dr. Mohammad Khoshnevisan, Patrick Maupin, Dr. Grégoire Mercier and Prof. Tzvetan Semerdjiev for their contributions to this first volume and their interests and support of these new concepts. We encourage all researchers interested in Information Fusion and by DSmT to contribute with papers to a second volume which will be published a few years later. This field of research is promising and currently very dynamic. Any comments, criticisms, notes and articles are welcome.

We are grateful to our colleagues for encouraging us to edit this book and for sharing with us many ideas and questions on DSmT over the last three years. We thank specially Dr. Albena Tchamova and Dr. Milan Daniel for reviewing carefully chapters of this book and also Dr. Frédéric Dambreville, Dr. Anne-Laure Jousselme, Dr. Branko Ristic and Professor Philippe Smets for their valuable recommendations and suggestions for the improvement of Chapter 12 and also Dr. Roy Streit for inciting us to work on the TP2 problem. We want also to thank Professors Krassimir Atanassov, Bassel Solaiman and Pierre Valin for accepting to serve as peer-reviewers for this book.

We are also grateful to International Society of Information Fusion (ISIF) for authorizing us to publish in Chapters 2, 3, 7, 13 and 15 some parts of articles presented during the recent ISIF conferences on Information Fusion.

Jean Dezert is grateful to Department of Information Modelling and Processing (DTIM/IED) at the Office National d'Études et de Recherches Aérospatiales (ONERA), Châtillon, France for encouraging him to carry on this research and for its financial support. He thanks his colleagues and friends Dr. Christian Musso (ONERA/DTIM/IED) for his valuable discussions on DSmT, and also Dr. Frédéric Cassaing (ONERA/DOTA/CC), Dr. Christophe Peyret (ONERA/DSNA) and Dr. Jean-Philippe Ovarlez (ONERA/DEMR/TSI) for their help for typing this book under the LaTeX typesetting system. He wants also to thank Dr. Patrick Vannoorenberghe from PSI Lab., University of Rouen, France, Prof. Tzvetan Semerdjiev from CLPP Bulgarian Academy of Sciences, Sofia, Professors Jean-Pierre Aubin and Patrick Saint-Pierre from University Paris-Dauphine, Paris, France, Prof. Laurence Hubert-Moy from COSTEL Lab./CNRS at University of Rennes 2, France, Dr. Frédéric Dambreville from CTA (Délégation Générale pour l'Armement), Arcueil, France and Prof. Subhash Challa formerly with Melbourne University and currently with University of Technology Sydney, Australia for inviting him to give seminars on DSmT during the last years.

Florentin Smarandache is grateful to The University of New Mexico that many times sponsored him to attend international conferences on Data Fusion or Fuzzy and Neutrosophic Logics in Australia or at the University of Berkeley in California where he met Professor Lofti Zadeh who became interested in DSmT, or in Sweden, and UNM for encouraging him to carry on this research.

We want to thank everyone.

The Editors

Prefaces

Advances in science and technology often result from paradigm shifts. In the 1910's, Einstein tried to reconcile the notion of absolute space and time of Cartesian dynamics, with Maxwell's electrodynamic equations, which introduced an absolute speed for light in vacuum. Addressing this dilemma inevitably lead him to put space and time on an equal footing, for any observer in an inertial frame, and special relativity was born. When he then tried to include gravitation in the picture, space and time became warped by mass (or energy) and general relativity emerged by connecting locally inertial frames. In each case, a new theory arose from relaxing assumptions, which formerly were thought to be immutable. We all know now the ideal regions of applicability of Cartesian dynamics (slow-moving objects) compared to those of special relativity (fast moving objects) and general relativity (cosmology and strong gravitational fields). However general relativity can reduce to special relativity, which itself can become Cartesian dynamics in everyday life. The price to pay in going from Cartesian dynamics to the more general formulations of relativity is increasing complexity of the calculations.

In his classic 1976 book, Shafer stated the paradigm shift, which led him to formulate an alternative to the existing Bayesian formalism for automated reasoning, thus leading to what is commonly known as Dempster-Shafer (DS) evidential reasoning. The basic concept was that an expert's complete ignorance about a statement need not translate into giving 1/2 a probability to the statement and the other 1/2 to its complement, as was assumed in Bayesian reasoning. Furthermore, when there are several possible single mutually exclusive alternatives (singletons) and the expert can only state positively the probabilities of a few of these, the remaining probabilities had to be distributed in some a priori fashion amongst all the other alternatives in Bayesian reasoning. The complete set of all the N alternatives (the frame of discernment) had to be known from the outset, as well as their natural relative frequency of occurrence. By allowing as an alternative that the ignorance could be assigned to the set of all remaining alternatives without any further dichotomy, a new theory was thus born that reasoned over sets of alternatives, DS theory.

Clearly the problem became more complex, as one had to reason over 2^N alternatives, the set of all subsets of the N singletons (under the union operator). When Dempster's orthogonal sum rule is used for combining (fusing) information from experts who might disagree with each other, one obtains the usual Dempster-Shafer (DS) theory. The degree of disagreement, or conflict, enters prominently in the renormalization process of the orthogonal sum rule and signals also when DS theory should be used with extreme caution: the conflict must not be too large. Indeed several paradoxes arise for highly conflicting experts (sources), and these have to be resolved in some way. Going back to relativity for a moment, the twin paradox occurs when one tries to explain it with special relativity, when actually it is a problem that has to be handled by general relativity. A paradigm shift was necessary and one will be needed here to solve the paradoxes (referred to in this book as counter-examples) of DS theory: the relaxation of an a priori completely known frame of discernment made of mutually exclusive singletons, and this is what Dezert-Smarandache (DSm) theory is basically all about.

In the first part of this book, DSm theory is motivated by expanding the frame of discernment to allow for presumed singletons in DS (or Bayesian) theory to actually have a well-defined intersection, which immediately states when this theory should be used: whenever it is impossible to estimate at the outset the granularity required to solve the problem at hand, either by construction (fuzzy concepts which cannot be refined further), or when the problem evolves in time to eventually reveal a finer granularity than originally assumed. It would then be important to continue being able to reason, rather than to go back and expand the frame of discernment and start the reasoning process over again.

However, clearly the problem again becomes more complex than DS theory, as one has to reason now over more alternatives (following Dedekind's sequence of numbers as N increases), consisting of the set of all subsets of the N original singletons (but under the union and the intersection operators). This is still less than would be required for a refined DS theory (if possible), which would consist of 2 to the power $2^N - 1$ alternatives. The classic DSm rule of combination ensures the desired commutativity and associativity properties, which made DS theory viable when the original orthogonal sum rule is used. This classic DSm rule is particularly simple and corresponds to the Free DSm model. Because the classic DSm rule does not involve a renormalization depending on the conflict, it will not exhibit the problems of DS theory under highly conflicting conditions. However since one of the applications of DSm theory involves dealing with problems with dynamic constraints (elements can be known not to occur at all at a certain time), a hybrid rule of combination is also proposed which deals with exclusivity constraints as well (some singletons are known to be truly exclusive). One can think of many examples where such available knowledge fluctuates with time. In this first part, the authors make a special effort to present instructive examples, which highlight both the free DSm model and the hybrid DSm model with exclusivity and/or non-existential constraints. The classic counter-examples to DS theory are presented, together with their solution in DSm theory.

In the second part of the book, data/information fusion applications of DSm theory are presented, including the Tweety Penguin triangle, estimation of target behavior tendencies, generalized data association for multi-target tracking in clutter, Blackman's data association problem, neutrosophic frameworks for situation analysis, land cover change detection from imagery, amongst others. This second part of the book is much more of an applied nature than the theoretical first part. This dual nature of the book makes it interesting reading for all open-minded scientists/engineers. Finally, I would like to thank the authors for having given me the opportunity to peer-review this fascinating book.

Pierre Valin, Prof., Ph.D.
Dept. de Physique
Université de Montréal
Montréal, Québec, Canada
May, 2004

This book presents the foundations, advances and some applications of a new theory of paradoxical and plausible reasoning developed by Jean Dezert and Florentin Smarandache, known as DSmT. This theory proposes a general method for combining uncertain, highly conflicting and imprecise data, provided by independent sources of information. It can be considered as a generalization of classical Dempster-Shafer mathematical theory of evidence, overcoming its inherent constraints, closely related with the acceptance of the law of the excluded middle. Refuting that principle, DSmT proposes a formalism to describe, analyze and combine all the available information, allowing the possibility for paradoxes between the elements of the frame of discernment. It is adapted to deal with each model of fusion occurring, taking into account all possible integrity constraints of the problem under consideration, due to the true nature and granularity of the concepts involved. This theory shows through the considered applications that conclusions drawn from it provides coherent results, which agree with the human reasoning and improves performances with respect to Dempster-Shafer Theory.

Krassimir Atanassov, Prof., Ph.D.
Centre of Biomedical Engineering
Bulgarian Academy of Sciences
Sofia, Bulgaria
May, 2004

Sciences advancement has always been through achievements, ideas and experiences accumulation. New ideas and approaches sometimes suffer misunderstanding and sometimes from a kind of "rejection" because they disturb existing approaches and, humans do not easily accept the changes. Simply, this is the human being history.

Information processing domain is not an exception. While preparing this preface, I remembered what happened when the fuzzy sets theory was developed. In the 1970's, some said "Fuzzy logic is the opium of sciences"! Amazing to see how things have changed since that time and how fuzzy sets theory is now well accepted and so well applied.

The scientific area of Information Fusion is beautifully "disturbing" our ways of thinking. In fact, this area imposes important questions: What is information? What is really informative in information? How to make information fusion? etc. From my own point of view, this area is pushing the scientific community towards promising approaches. One of these approaches is raised by Florentin Smarandache & Jean Dezert in their book: Advances and Applications of DSmT for Information Fusion. This approach aims to formalize the fusion approach in the very particular context of uncertain and highly conflicting information. The Dezert-Smarandache Theory (DSmT) should be considered as an extension of the Dempster-Shafer (DS) as well as the Bayesian theories. From a technical point of view, the fundamental question concerning the granularity of the singletons forming the frame of discernment is clearly raised. The book is not only limited to theoretical developments but also presents a set of very interesting applications, making thus, its reading a real pleasure.

I would like to thank the authors for their original contribution and to encourage the development of this very promising approach.

Bassel Solaiman, Prof., Ph.D.
ENST Bretagne
Brest - France
May, 2004

Part I

Advances on DSmT

Chapter 1

Presentation of DSmT

Jean Dezert	Florentin Smarandache
ONERA	Department of Mathematics
29 Av. de la Division Leclerc	University of New Mexico
92320 Châtillon	Gallup, NM 8730
France	U.S.A.

Abstract: *This chapter presents a general overview and foundations of the DSmT, i.e. the recent theory of plausible and paradoxical reasoning developed by the authors, specially for the static or dynamic fusion of information arising from several independent but potentially highly conflicting, uncertain and imprecise sources of evidence. We introduce and justify here the basis of the DSmT framework with respect to the Dempster-Shafer Theory (DST), a mathematical theory of evidence developed in 1976 by Glenn Shafer. We present the DSm combination rules and provide some simple illustrative examples and comparisons with other main rules of combination available in the literature for the combination of information for simple fusion problems. Detailed presentations on recent advances and applications of DSmT are presented in the next chapters of this book.*

1.1 Introduction

The Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning proposed by the authors in recent years [9, 10, 36] can be considered as an extension of the classical Dempster-Shafer theory (DST) [33] but includes fundamental differences with the DST. DSmT allows to formally combine any types of independent sources of information represented in term of belief functions, but is mainly focused on the fusion of uncertain, highly conflicting and imprecise sources of evidence. DSmT is able to solve complex static or dynamic fusion problems beyond the limits of the DST framework, specially when conflicts between sources become large and when the refinement of the frame of the problem under consideration, denoted Θ, becomes inaccessible because of the vague, relative and imprecise nature of elements of Θ [10].

The foundation of DSmT is based on the definition of the Dedekind's lattice D^Θ also called *hyper-power set* of the frame Θ in the sequel. In the DSmT framework, Θ is first considered as only a set $\{\theta_1, \ldots, \theta_n\}$ of n exhaustive elements (closed world assumption) without introducing other constraint (exclusivity or non-existential constraints). This corresponds to the *free DSm model* on which is based the *classic DSm rule of combination*. The exhaustivity (closed world) assumption is not fundamental actually, because one can always close any open world theoretically, say Θ_{Open} by including into it an extra element/hypothesis θ_0 (although not precisely identified) corresponding to all missing hypotheses of Θ_{Open} to work with the new closed frame $\Theta = \{\theta_0\} \cup \Theta_{\text{Open}} = \{\theta_0, \theta_1, \ldots, \theta_n\}$. This idea has been already proposed and defended by Yager, Dubois & Prade and Testemale in [45, 13, 30] and differs from the Transferable Belief Model (TBM) of Smets [42]. The proper use of the free DSm model for the fusion depends on the intrinsic nature of elements/concepts θ_i involved in the problem under consideration and becomes naturally justified when dealing with vague/continuous elements which cannot be precisely defined and separated (e.g. the relative concepts of smallness/tallness, pleasure/pain, hot/cold, colors (because of the continuous spectrum of the light), etc) so that no refinement of Θ in a new larger set Θ_{ref} of exclusive refined hypotheses is possible. In such case, we just call Θ the *frame* of the problem.

When a complete refinement (or maybe sometimes an only partial refinement) of Θ is possible and thus allows us to work on Θ_{ref}, then we call Θ_{ref} the *frame of discernment* (resp. *frame of partial discernment*) of the problem because some elements of Θ_{ref} are truly exclusive and thus they become (resp. partially) discernable. The refined frame of discernment assuming exclusivity of all elements $\theta_i \in \Theta$ corresponds to the *Shafer's model* on which is based the DST and can be obtained from the free DSm model by introducing into it all exclusivity constraints. All fusion problems dealing with truly exclusive concepts must obviously be based on such model since it describes adequately the real and intrinsic nature of hypotheses. Actually, any constrained model (including Shafer's model) corresponds to what we called an *hybrid DSm model*. DSmT provides a generalized hybrid DSm rule of combination for working with any kind of hybrid models including exclusivity and non-existential constraints as well and it is not only limited to the most constrained one, i.e. Shafer's model (see chapter 4 for a detailed presentation and examples on the hybrid DSm rule). Before going further into this DSmT presentation it is necessary to briefly present the foundations of the DST [33] for pointing out the important differences between these two theories for managing the combination of evidence.

1.2 Short introduction to the DST

In this section, we present a short introduction to the Dempster-Shafer theory. A complete presentation of the Mathematical Theory of Evidence proposed by Glenn Shafer can be found in his milestone book in [33]. Advances on DST can be found in [34, 48] and [49].

1.2.1 Shafer's model and belief functions

Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ be the *frame of discernment* of the fusion problem under consideration having n *exhaustive* and *exclusive elementary* hypotheses θ_i. This corresponds to Shafer's model of the problem. Such a model assumes that an ultimate refinement of the problem is possible (exists and is achievable) so that θ_i are well precisely defined/identified in such a way that we are sure that they are exclusive and exhaustive (closed-world assumption).

The set of all subsets of Θ is called the *power set* of Θ and is denoted 2^Θ. Its cardinality is $2^{|\Theta|}$. Since 2^Θ is closed under unions, intersections, and complements, it defines a *Boolean algebra*.

By example, if $\Theta = \{\theta_1, \theta_2, \theta_3\}$ then $2^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_3, \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cup \theta_2 \cup \theta_3\}$.

In Shafer's model, a *basic belief assignment* (bba) $m(.) : 2^\Theta \rightarrow [0, 1]$ associated to a given body of evidence $\mathcal{B}$ (also called corpus of evidence) is defined by [33]

$$m(\emptyset) = 0 \qquad \text{and} \qquad \sum_{A \in 2^\Theta} m(A) = 1 \tag{1.1}$$

Glenn Shafer defines the belief (credibility) and plausibility functions of $A \subseteq \Theta$ as

$$\text{Bel}(A) = \sum_{B \in 2^\Theta, B \subseteq A} m(B) \tag{1.2}$$

$$\text{Pl}(A) = \sum_{B \in 2^\Theta, B \cap A \neq \emptyset} m(B) = 1 - \text{Bel}(\bar{A}) \tag{1.3}$$

where $\bar{A}$ denotes the complement of the proposition A in Θ.

The belief functions $m(.)$, Bel(.) and Pl(.) are in one-to-one correspondence [33]. The set of elements $A \in 2^\Theta$ having a positive basic belief assignment is called the *core/kernel* of the source of evidence under consideration and is denoted $\mathcal{K}(m)$.

1.2.2 Dempster's rule of combination

Let $\text{Bel}_1(.)$ and $\text{Bel}_2(.)$ be two belief functions provided by two independent (and a priori equally reliable) sources/bodies of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$ over the same frame of discernment Θ and their corresponding bba $m_1(.)$ and $m_2(.)$. Then the combined global belief function denoted $\text{Bel}(.) = \text{Bel}_1(.) \oplus \text{Bel}_2(.)$ is obtained by combining the bba $m_1(.)$ and $m_2(.)$ through the following Dempster rule of combination [33] $m(.) = [m_1 \oplus m_2](.)$ where

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \dfrac{\displaystyle\sum_{\substack{X,Y \in 2^\Theta \\ X \cap Y = A}} m_1(X)m_2(Y)}{1 - \displaystyle\sum_{\substack{X,Y \in 2^\Theta \\ X \cap Y = \emptyset}} m_1(X)m_2(Y)} \qquad \forall (A \neq \emptyset) \in 2^\Theta \end{cases} \tag{1.4}$$

$m(.)$ is a proper basic belief assignment if and only if the denominator in equation (1.4) is non-zero. The *degree of conflict* between the sources $\mathcal{B}_1$ and $\mathcal{B}_2$ is defined by

$$k_{12} \triangleq \sum_{\substack{X,Y \in 2^\Theta \\ X \cap Y = \emptyset}} m_1(X)m_2(Y) \tag{1.5}$$

The effect of the normalizing factor $1 - k_{12}$ in (1.4) consists in eliminating the conflicting pieces of information between the two sources to combine, consistently with the intersection operator. When $k_{12} = 1$, the combined bba $m(.)$ does not exist and the bodies of evidences $\mathcal{B}_1$ and $\mathcal{B}_2$ are said to be in *full contradiction*. Such a case arises when there exists $A \subset \Theta$ such that $\text{Bel}_1(A) = 1$ and $\text{Bel}_2(\bar{A}) = 1$. The core of the bba $m(.)$ equals the intersection of the cores of m_1 and m_2, i.e $\mathcal{K}(m) = \mathcal{K}(m_1) \cap \mathcal{K}(m_2)$. Up to the normalization factor $1 - k_{12}$, Dempster's rule is formally nothing but a random set intersection under stochastic assumption and it corresponds to the conjunctive consensus [13]. Dempster's rule of combination can be directly extended for the combination of N independent and equally reliable sources of evidence and its major interest comes essentially from its commutativity and associativity properties [33]. A recent discussion on Dempster's and Bayesian rules of combination can be found in [5].

1.2.3 Alternatives to Dempster's rule of combination

The DST is attractive for the Information Fusion community because it gives a nice mathematical model for the representation of uncertainty and it includes Bayesian theory as a special case [33] (p. 4). Although very appealing, the DST presents some weaknesses and limitations [27] already reported by Zadeh [50, 51, 52, 53] and Dubois & Prade in the eighties [12] and reinforced by Voorbraak in [43] because of the lack of complete theoretical justification of Dempster's rule of combination, but mainly because of our low confidence to trust the result of Dempster's rule of combination when the conflict becomes important between sources (i.e. $k_{12} \nearrow 1$). Indeed, there exists an infinite class of cases where Dempster's rule of combination can assign certainty to a minority opinion (other infinite classes of counter-examples are discussed in chapter 5) or where the "ignorance" interval disappears forever whenever a single piece of evidence commits all its belief to a proposition and its negation [29]. Moreover, elements of sets with larger cardinality can gain a disproportionate share of belief [43]. These drawbacks have fed intensive debates and research works for the last twenty years:

- either to interpret (and justify as best as possible) the use of Dempster's rule by several approaches and to circumvent numerical problems with it when conflict becomes high. These approaches are mainly based on the extension of the domain of the probability functions from the propositional logic domain to the modal propositional logic domain [31, 32, 28] or on the hint model [22] and probabilistic argumentation systems [14, 15, 1, 2, 16, 17, 18, 19, 20]. Discussions on these interpretations of DST can be found in [38, 40, 42], and also in chapter 12 of this book which analyzes and compares Bayesian reasoning, Dempster-Shafer's reasoning and DSm reasoning on a very simple but interesting example drawn from [28].

- or to propose new alternative rules. DSmT fits in this category since it extends the foundations of DST and also provides a new combination rules as it will be shown in next sections.

Several interesting and valuable alternative rules have thus been proposed in literature to circumvent the limitations of Dempster's rule of combination. The major common alternatives are listed in this section and most of the current available combination rules have been recently unified in a nice general framework by Lefèvre, Colot and Vanoorenberghe in [25]. Their important contribution, although strongly criticized by Haenni in [19] but properly justified by Lefevre et al. in [26], shows clearly that an infinite number of possible rules of combinations can be built from Shafer's model depending on the choice for transfer of the conflicting mass (i.e. k_{12}). A justification of Dempster's rule of combination has been proposed afterwards in the nineties by the axiomatic of Philippe Smets [37, 24, 41, 42] based on his Transferable Belief Model (TBM) related to anterior works of Cheng and Kashyap in [6], a non-probabilistic interpretation of Dempster-Shafer theory (see [3, 4] for discussion).

Here is the list of the most common rules of combination[1] for two independent sources of evidence proposed in the literature in the DST framework as possible alternatives to Dempster's rule of combination to overcome its limitations. Unless explicitly specified, the sources are assumed to be equally reliable.

- **The disjunctive rule of combination** [11, 13, 39]: This commutative and associative rule proposed by Dubois & Prade in 1986 and denoted here by the index $\cup$ is examined in details in chapter 9. $m_{\cup}(.)$ is defined $\forall A \in 2^{\Theta}$ by

$$\begin{cases} m_{\cup}(\emptyset) = 0 \\ m_{\cup}(A) = \displaystyle\sum_{\substack{X,Y \in 2^{\Theta} \\ X \cup Y = A}} m_1(X) m_2(Y) \qquad \forall (A \neq \emptyset) \in 2^{\Theta} \end{cases} \tag{1.6}$$

[1]The MinC rule of combination is not included here since it is covered in details in chapter 10.

The core of the belief function given by $m_\cup$ equals the union of the cores of Bel_1 and Bel_2. This rule reflects the disjunctive consensus and is usually preferred when one knows that one of the source $\mathcal{B}_1$ or $\mathcal{B}_2$ is mistaken but without knowing which one among $\mathcal{B}_1$ and $\mathcal{B}_2$.

- **Murphy's rule of combination** [27]: This commutative (but not associative) trade-off rule, denoted here with index M, drawn from [46, 13] is a special case of convex combination of bba m_1 and m_2 and consists actually in a simple arithmetic average of belief functions associated with m_1 and m_2. $\mathrm{Bel}_M(.)$ is then given $\forall A \in 2^\Theta$ by:

$$\mathrm{Bel}_M(A) = \frac{1}{2}[\mathrm{Bel}_1(A) + \mathrm{Bel}_2(A)] \tag{1.7}$$

- **Smets' rule of combination** [41, 42]: This commutative and associative rule corresponds actually to the non-normalized version of Dempster's rule of combination. It allows positive mass on the null/empty set $\emptyset$. This eliminates the division by $1 - k_{12}$ involved in Dempster's rule (1.4). Smets' rule of combination of two independent (equally reliable) sources of evidence (denoted here by index S) is given by:

$$\begin{cases} m_S(\emptyset) \equiv k_{12} = \displaystyle\sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y=\emptyset}} m_1(X)m_2(Y) \\ m_S(A) = \displaystyle\sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y=A}} m_1(X)m_2(Y) \qquad \forall(A\neq\emptyset)\in 2^\Theta \end{cases} \tag{1.8}$$

- **Yager's rule of combination** [45, 46, 47]: Yager admits that in case of conflict the result is not reliable, so that k_{12} plays the role of an absolute discounting term added to the weight of ignorance. The commutative (but not associative) Yager rule, denoted here by index Y is given[2] by:

$$\begin{cases} m_Y(\emptyset) = 0 \\ m_Y(A) = \displaystyle\sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y=A}} m_1(X)m_2(Y) \qquad \forall A\in 2^\Theta, A\neq\emptyset, A\neq\Theta \\ m_Y(\Theta) = m_1(\Theta)m_2(\Theta) + \displaystyle\sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y=\emptyset}} m_1(X)m_2(Y) \quad \text{when} \quad A=\Theta \end{cases} \tag{1.9}$$

- **Dubois & Prade's rule of combination** [13]: We admit that the two sources are reliable when they are not in conflict, but one of them is right when a conflict occurs. Then if one observes a value in set X while the other observes this value in a set Y, the truth lies in $X \cap Y$ as long $X \cap Y \neq \emptyset$. If $X \cap Y = \emptyset$, then the truth lies in $X \cup Y$ [13]. According to this principle, the commutative (but

[2] Θ represents here the full ignorance $\theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$ on the frame of discernment according the notation used in [33].

not associative) Dubois & Prade hybrid rule of combination, denoted here by index DP, which is a reasonable trade-off between precision and reliability, is defined[3] by:

$$\begin{cases} m_{DP}(\emptyset) = 0 \\ m_{DP}(A) = \displaystyle\sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y = A \\ X\cap Y\neq\emptyset}} m_1(X)m_2(Y) + \sum_{\substack{X,Y\in 2^\Theta \\ X\cup Y = A \\ X\cap Y=\emptyset}} m_1(X)m_2(Y) \qquad \forall A\in 2^\Theta, A\neq\emptyset \end{cases} \tag{1.10}$$

1.2.3.1 The unified formulation for rules of combinations involving conjunctive consensus

We present here the unified framework recently proposed by Lefèvre, Colot and Vanoorenberghe in [25] to embed all the existing (and potentially forthcoming) combination rules involving conjunctive consensus in the same general mechanism of construction. Here is the principle of their general formulation based on two steps.

- **Step 1**: Computation of the total conflicting mass based on the conjunctive consensus

$$k_{12} \triangleq \sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y=\emptyset}} m_1(X)m_2(Y) \tag{1.11}$$

- **Step 2**: This step consists in the reallocation (convex combination) of the conflicting masses on $(A\neq\emptyset)\subseteq\Theta$ with some given coefficients $w_m(A)\in[0,1]$ such that $\sum_{A\subseteq\Theta} w_m(A) = 1$ according to

$$\begin{cases} m(\emptyset) = w_m(\emptyset)k_{12} \\ m(A) = [\displaystyle\sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y=A}} m_1(X)m_2(Y)] + w_m(A)k_{12} \qquad \forall(A\neq\emptyset)\in 2^\Theta \end{cases} \tag{1.12}$$

The particular choice of the set of coefficients $w_m(.)$ provides a particular rule of combination. Actually this nice and important general formulation shows there exists an infinite number of possible rules of combination. Some rules are then justified or criticized with respect to the other ones mainly on their ability to, or not to, preserve the associativity and commutativity properties of the combination. It can be easily shown in [25] that such general procedure provides all existing rules involving conjunctive consensus developed in the literature based on Shafer's model. As examples:

- **Dempster's rule of combination** (1.4) can be obtained from (1.12) by choosing $\forall A\neq\emptyset$

$$w_m(\emptyset) = 0 \qquad \text{and} \qquad w_m(A) = \frac{1}{1-k_{12}} \sum_{\substack{X,Y\in 2^\Theta \\ X\cap Y=A}} m_1(X)m_2(Y) \tag{1.13}$$

[3] taking into account the the correction of the typo error in formula (56) given in [13], page 257.

- **Yager's rule of combination** (1.9) is obtained by choosing

$$w_m(\Theta) = 1 \qquad \text{and} \qquad w_m(A \neq \Theta) = 0 \tag{1.14}$$

- **Smets' rule of combination** (1.8) is obtained by choosing

$$w_m(\emptyset) = 1 \qquad \text{and} \qquad w_m(A \neq \emptyset) = 0 \tag{1.15}$$

- **Dubois and Prade's rule of combination** (1.10) is obtained by choosing

$$\forall A \subseteq \mathcal{P}, \qquad w_m(A) = \frac{1}{1-k_{12}} \sum_{\substack{A_1,A_2 | A_1\cup A_2 = A \\ A_1\cap A_2=\emptyset}} m^\star \tag{1.16}$$

where $m^\star \triangleq m_1(A_1)m_2(A_2)$ corresponds to the partial conflicting mass which is assigned to $A_1 \cup A_2$. $\mathcal{P}$ is the set of all subsets of 2^Θ on which the conflicting mass is distributed. $\mathcal{P}$ is defined by [25]

$$\mathcal{P} \triangleq \{A \in 2^\Theta \mid \exists A_1 \in \mathcal{K}(m_1), \exists A_2 \in \mathcal{K}(m_2), A_1 \cup A_2 = A \text{ and } A_1 \cap A_2 = \emptyset\} \tag{1.17}$$

The computation of the weighting factors $w_m(A)$ of Dubois and Prade's rule of combination does not depend only on propositions they are associated with, but also on belief mass functions which have cause the partial conflicts. Thus the belief mass functions leading to the conflict allow to compute that part of conflicting mass which must be assigned to the subsets of $\mathcal{P}$ [25]. Yager's rule coincides with the Dubois and Prade's rule of combination when $\mathcal{P} = \{\Theta\}$.

1.2.4 The discounting of sources of evidence

Most of the rules of combination proposed in the literature are based on the assumption of the same reliability of sources of evidence. When the sources are known not being equally reliable and the reliability of each source is perfectly known (or at least has been properly estimated when it's possible [42, 25]), then is it natural and reasonable to discount each unreliable source proportionally to its corresponding reliability factor according to method proposed by Shafer in [33], chapter 11. Two methods are usually used for discounting the sources:

- **Classical discounting method** [33, 13, 42, 25]:

  Assume that the reliability/confidence[4] factor $\alpha \in [0, 1]$ of a source is known, then the discounting of the bba $m(.)$ provided by the unreliable source is done to obtain a new (discounted) bba $m'(.)$ as follows:

$$\begin{cases} m'(A) = \alpha \cdot m(A), & \forall A \in 2^\Theta, A \neq \Theta \\ m'(\Theta) = (1-\alpha) + \alpha \cdot m(\Theta) \end{cases} \tag{1.18}$$

[4] We prefer to use here the terminology *confidence* rather than *reliability* since the notion of reliability is closely related to the repetition of experiments with random outputs which may not be always possible in the context of some information fusion applications (see example 1.6 given by Shafer on the life on Sirius in [33], p.23)

$\alpha = 1$ means the total confidence in the source while $\alpha = 0$ means a complete calling in question of the reliability of the source.

- **Discounting by convex combination of sources** [13]: This method of discounting is based on the convex combination of sources by their relative reliabilities, assumed to be known. Let consider two independent unreliable sources of evidence with reliability factors α_1 and α_2 with $\alpha_1, \alpha_2 \in [0, 1]$, then the result of the combination of the discounted sources will be given $\forall A \in 2^\Theta$ by

$$\mathrm{Bel}(A) = \frac{\alpha_1}{\alpha_1 + \alpha_2}\mathrm{Bel}_1(A) + \frac{\alpha_2}{\alpha_1 + \alpha_2}\mathrm{Bel}_2(A) \tag{1.19}$$

  When the sources are highly conflicting and they have been sufficiently discounted, Shafer has shown in [33], p. 253, that the combination of a large number n of equally reliable sources using Dempster's rule on equally discounted belief functions, becomes similar to the convex combination of the n sources with equal reliability factors $\alpha_i = 1/n$. A detailed presentation of discounting methods can be found in [13].

It is important to note that such discounting methods must not be chosen as an ad-hoc tool to adjust the result of the fusion (once obtained) in case of troubles if a counter-intuitive or bad result arises, but only *beforehand* when one has prior information on the quality of sources. In the sequel of the book we will assume that sources under consideration are a priori equally reliable/trustable, unless specified explicitly. Although being very important for practical issues, the case of the fusion of known unreliable sources of information is not considered in this book because it depends on the own choice of the discounting method adopted by the system designer (this is also highly related with the application under consideration and the types of the sources to be combined). Fundamentally the problem of combination of unreliable sources of evidence is the same as working with new sets of basic belief assignments and thus has little interest in the framework of this book.

1.3 Foundations of the DSmT

1.3.1 Notion of free and hybrid DSm models

The development of the DSmT arises from the necessity to overcome the inherent limitations of the DST which are closely related with the acceptance of Shafer's model (the frame of *discernment* Θ defined as a finite set of *exhaustive* and *exclusive* hypotheses θ_i, $i = 1, \dots, n$), the third middle excluded principle (i.e. the existence of the complement for any elements/propositions belonging to the power set of Θ), and the acceptance of Dempter's rule of combination (involving normalization) as the framework for the combination of independent sources of evidence. We argue that these three fundamental conditions of the DST can be removed and another new mathematical approach for combination of evidence is possible.

The basis of the DSmT is the refutation of the principle of the third excluded middle and Shafer's model, since for a wide class of fusion problems the intrinsic nature of hypotheses can be only vague and imprecise in such a way that precise refinement is just impossible to obtain in reality so that the exclusive elements θ_i cannot be properly identified and precisely separated. Many problems involving fuzzy continuous and relative concepts described in natural language and having no absolute interpretation like tallness/smallness, pleasure/pain, cold/hot, Sorites paradoxes, etc, enter in this category. DSmT starts with the notion of *free DSm model*, denoted $\mathcal{M}^f(\Theta)$, and considers Θ only as a frame of exhaustive elements θ_i, $i = 1, \ldots, n$ which can potentially overlap. This model is *free* because no other assumption is done on the hypotheses, but the weak exhaustivity constraint which can always been satisfied according the closure principle explained in the introduction of this chapter. No other constraint is involved in the free DSm model. When the free DSm model holds, the classic commutative and associative DSm rule of combination (corresponding to the conjunctive consensus defined on the free Dedekind's lattice - see next subsection) is performed.

Depending on the intrinsic nature of the elements of the fusion problem under consideration, it can however happen that the free model does not fit the reality because some subsets of Θ can contain elements known to be truly exclusive but also truly non existing at all at a given time (specially when working on dynamic fusion problem where the frame Θ varies with time with the revision of the knowledge available). These integrity constraints are then explicitly and formally introduced into the free DSm model $\mathcal{M}^f(\Theta)$ in order to adapt it properly to fit as close as possible with the reality and permit to construct a *hybrid DSm model* $\mathcal{M}(\Theta)$ on which the combination will be efficiently performed. Shafer's model, denoted $\mathcal{M}^0(\Theta)$, corresponds to a very specific hybrid DSm model including all possible exclusivity constraints. The DST has been developed for working only with $\mathcal{M}^0(\Theta)$ while the DSmT has been developed for working with any kind of hybrid model (including Shafer's model and the free DSm model), to manage as efficiently and precisely as possible imprecise, uncertain and potentially high conflicting sources of evidence while keeping in mind the possible dynamicity of the information fusion problematic. The foundations of the DSmT are therefore totally different from those of all existing approaches managing uncertainties, imprecisions and conflicts. DSmT provides a new interesting way to attack the information fusion problematic with a general framework in order to cover a wide variety of problems. A detailed presentation of hybrid DSm models and hybrid DSm rule of combination is given in chapter 4.

DSmT refutes also the idea that sources of evidence provide their beliefs with the same absolute interpretation of elements of the same frame Θ and the conflict between sources arises not only because of the possible unreliabilty of sources, but also because of possible different and relative interpretation of Θ, e.g. what is considered as good for somebody can be considered as bad for somebody else. There is some unavoidable subjectivity in the belief assignments provided by the sources of evidence, otherwise it would mean that all bodies of evidence have a same objective and universal interpretation (or measure) of the phenomena under consideration, which unfortunately rarely occurs in reality, but when bba are based on some *objective probabilities* transformations. But in this last case, probability theory can handle properly and efficiently the information, and the DST, as well as the DSmT, becomes useless. If we now get out of the probabilistic background argumentation for the construction of bba, we claim that in most of cases, the sources of evidence provide their beliefs about elements of the frame of the fusion problem only based on their own limited knowledge and experience without reference to the (inaccessible) absolute truth of the space of possibilities.

The DSmT includes the possibility to deal with evidences arising from different sources of information which do not have access to the absolute and same interpretation of the elements of Θ under consideration. The DSmT, although not based on probabilistic argumentation can be interpreted as an extension of Bayesian theory and Dempster-Shafer theory in the following sense. Let $\Theta = \{\theta_1, \theta_2\}$ be the simplest frame made of only two hypotheses, then

- the probability theory deals, under the assumptions on exclusivity and exhaustivity of hypotheses, with basic probability assignments (bpa) $m(.) \in [0, 1]$ such that

$$m(\theta_1) + m(\theta_2) = 1$$

- the DST deals, under the assumptions on exclusivity and exhaustivity of hypotheses, with bba $m(.) \in [0, 1]$ such that

$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) = 1$$

- the DSmT theory deals, under only assumption on exhaustivity of hypotheses (i.e. the free DSm model), with the generalized bba $m(.) \in [0, 1]$ such that

$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) + m(\theta_1 \cap \theta_2) = 1$$

1.3.2 Notion of hyper-power set D^Θ

One of the cornerstones of the DSmT is the notion of hyper-power set (see chapters 2 and 3 for examples and a detailed presentation). Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ be a finite set (called frame) of n exhaustive elements[5]. The Dedekind's lattice, also called in the DSmT framework *hyper-power set* D^Θ is defined as the set of all composite propositions built from elements of Θ with $\cup$ and $\cap$ operators[6] such that:

[5] We do not assume here that elements θ_i are necessary exclusive. There is no restriction on θ_i but the exhaustivity.
[6] Θ generates D^Θ under operators $\cup$ and $\cap$

1. $\emptyset, \theta_1, \ldots, \theta_n \in D^\Theta$.

2. If $A, B \in D^\Theta$, then $A \cap B \in D^\Theta$ and $A \cup B \in D^\Theta$.

3. No other elements belong to D^Θ, except those obtained by using rules 1 or 2.

The dual (obtained by switching $\cup$ and $\cap$ in expressions) of D^Θ is itself. There are elements in D^Θ which are self-dual (dual to themselves), for example α_8 for the case when $n = 3$ in the example below. The cardinality of D^Θ is majored by 2^{2^n} when the cardinality of Θ equals n, i.e. $|\Theta| = n$. The generation of hyper-power set D^Θ is closely related with the famous Dedekind problem [8, 7] on enumerating the set of isotone Boolean functions. The generation of the hyper-power set is presented in chapter 2. Since for any given finite set Θ, $|D^\Theta| \geq |2^\Theta|$ we call D^Θ the *hyper-power set* of Θ.

Example of the first hyper-power sets D^Θ

- For the degenerate case ($n = 0$) where $\Theta = \{\}$, one has $D^\Theta = \{\alpha_0 \triangleq \emptyset\}$ and $|D^\Theta| = 1$.

- When $\Theta = \{\theta_1\}$, one has $D^\Theta = \{\alpha_0 \triangleq \emptyset, \alpha_1 \triangleq \theta_1\}$ and $|D^\Theta| = 2$.

- When $\Theta = \{\theta_1, \theta_2\}$, one has $D^\Theta = \{\alpha_0, \alpha_1, \ldots, \alpha_4\}$ and $|D^\Theta| = 5$ with $\alpha_0 \triangleq \emptyset$, $\alpha_1 \triangleq \theta_1 \cap \theta_2$, $\alpha_2 \triangleq \theta_1$, $\alpha_3 \triangleq \theta_2$ and $\alpha_4 \triangleq \theta_1 \cup \theta_2$.

- When $\Theta = \{\theta_1, \theta_2, \theta_3\}$, one has $D^\Theta = \{\alpha_0, \alpha_1, \ldots, \alpha_{18}\}$ and $|D^\Theta| = 19$ with

$$
\begin{array}{ll}
\alpha_0 \triangleq \emptyset & \\
\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 & \alpha_{10} \triangleq \theta_2 \\
\alpha_2 \triangleq \theta_1 \cap \theta_2 & \alpha_{11} \triangleq \theta_3 \\
\alpha_3 \triangleq \theta_1 \cap \theta_3 & \alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \\
\alpha_4 \triangleq \theta_2 \cap \theta_3 & \alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \\
\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 & \alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \\
\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 & \alpha_{15} \triangleq \theta_1 \cup \theta_2 \\
\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 & \alpha_{16} \triangleq \theta_1 \cup \theta_3 \\
\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) & \alpha_{17} \triangleq \theta_2 \cup \theta_3 \\
\alpha_9 \triangleq \theta_1 & \alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3
\end{array}
$$

Note that the complement $\bar{A}$ of any proposition A (except for $\emptyset$ and for the total ignorance $I_t \triangleq \theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$), is not involved within DSmT because of the refutation of the third excluded middle. In other words, $\forall A \in D^\Theta$ with $A \neq \emptyset$ or $A \neq I_t$, $\bar{A} \notin D^\Theta$. Thus $(D^\Theta, \cap, \cup)$ does not define a Boolean algebra. The cardinality of hyper-power set D^Θ for $n \geq 1$ follows the sequence of Dedekind's numbers [35], i.e. 1,2,5,19,167,7580,7828353,... (see next chapter for details).

Elements θ_i, $i = 1, \ldots, n$ of Θ constitute the finite set of hypotheses/concepts characterizing the fusion problem under consideration. D^Θ constitutes what we call the *free DSm model* $\mathcal{M}^f(\Theta)$ and allows to work with fuzzy concepts which depict a continuous and relative intrinsic nature. Such kinds of concepts cannot be precisely refined in an absolute interpretation because of the unapproachable universal truth.

However for some particular fusion problems involving discrete concepts, elements θ_i are truly exclusive. In such case, all the exclusivity constraints on θ_i, $i = 1, \ldots, n$ have to be included in the previous model to characterize properly the true nature of the fusion problem and to fit it with the reality. By doing this, the hyper-power set D^Θ reduces naturally to the classical power set 2^Θ and this constitutes the most restricted hybrid DSm model, denoted $\mathcal{M}^0(\Theta)$, coinciding with Shafer's model. As an exemple, let's consider the 2D problem where $\Theta = \{\theta_1, \theta_2\}$ with $D^\Theta = \{\emptyset, \theta_1 \cap \theta_2, \theta_1, \theta_2, \theta_1 \cup \theta_2\}$ and assume now that θ_1 and θ_2 are truly exclusive (i.e. Shafer's model $\mathcal{M}^0$ holds), then because $\theta_1 \cap \theta_2 \overset{\mathcal{M}^0}{=} \emptyset$, one gets $D^\Theta = \{\emptyset, \theta_1 \cap \theta_2 \overset{\mathcal{M}^0}{=} \emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2\} = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2\} \equiv 2^\Theta$.

Between the class of fusion problems corresponding to the free DSm model $\mathcal{M}^f(\Theta)$ and the class of fusion problems corresponding to Shafer's model $\mathcal{M}^0(\Theta)$, there exists another wide class of hybrid fusion problems involving in Θ both fuzzy continuous concepts and discrete hypotheses. In such (hybrid) class, some exclusivity constraints and possibly some non-existential constraints (especially when working on dynamic[7] fusion) have to be taken into account. Each hybrid fusion problem of this class will then be characterized by a proper hybrid DSm model $\mathcal{M}(\Theta)$ with $\mathcal{M}(\Theta) \neq \mathcal{M}^f(\Theta)$ and $\mathcal{M}(\Theta) \neq \mathcal{M}^0(\Theta)$, see examples presented in chapter 4.

1.3.3 Generalized belief functions

From a general frame Θ, we define a map $m(.) : D^\Theta \to [0, 1]$ associated to a given body of evidence $\mathcal{B}$ as

$$m(\emptyset) = 0 \qquad \text{and} \qquad \sum_{A \in D^\Theta} m(A) = 1 \tag{1.20}$$

The quantity $m(A)$ is called the *generalized basic belief assignment/mass* (gbba) of A.

The *generalized belief and plausibility functions* are defined in almost the same manner as within the DST, i.e.

$$\text{Bel}(A) = \sum_{\substack{B \subseteq A \\ B \in D^\Theta}} m(B) \tag{1.21}$$

$$\text{Pl}(A) = \sum_{\substack{B \cap A \neq \emptyset \\ B \in D^\Theta}} m(B) \tag{1.22}$$

[7]i.e. when the frame Θ is changing with time.

These definitions are compatible with the definitions of classical belief functions in the DST framework when D^Θ reduces to 2^Θ for fusion problems where Shafer's model $\mathcal{M}^0(\Theta)$ holds. We still have $\forall A \in D^\Theta$, $\mathrm{Bel}(A) \leq \mathrm{Pl}(A)$. Note that when working with the free DSm model $\mathcal{M}^f(\Theta)$, one has always $\mathrm{Pl}(A) = 1$ $\forall A \neq \emptyset \in D^\Theta$ which is normal.

1.3.4 The classic DSm rule of combination

When the free DSm model $\mathcal{M}^f(\Theta)$ holds for the fusion problem under consideration, the classic DSm rule of combination $m_{\mathcal{M}^f(\Theta)} \equiv m(.) \triangleq [m_1 \oplus m_2](.)$ of two independent sources of evidences $\mathcal{B}_1$ and $\mathcal{B}_2$ over the same frame Θ with belief functions $\mathrm{Bel}_1(.)$ and $\mathrm{Bel}_2(.)$ associated with gbba $m_1(.)$ and $m_2(.)$ corresponds to the conjunctive consensus of the sources. It is given by [9, 10]:

$$\forall C \in D^\Theta, \qquad m_{\mathcal{M}^f(\Theta)}(C) \equiv m(C) = \sum_{\substack{A,B \in D^\Theta \\ A \cap B = C}} m_1(A) m_2(B) \tag{1.23}$$

Since D^Θ is closed under $\cup$ and $\cap$ set operators, this new rule of combination guarantees that $m(.)$ is a proper generalized belief assignment, i.e. $m(.) : D^\Theta \rightarrow [0, 1]$. This rule of combination is commutative and associative and can always be used for the fusion of sources involving fuzzy concepts. This rule can be directly and easily extended for the combination of $k > 2$ independent sources of evidence (see the expression for $S_1(.)$ in the next section and chapter 4 for details).

This classic DSm rule of combination becomes very expensive in terms of computations and memory size due to the huge number of elements in D^Θ when the cardinality of Θ increases. This remark is however valid only if the cores (the set of focal elements of gbba) $\mathcal{K}_1(m_1)$ and $\mathcal{K}_2(m_2)$ coincide with D^Θ, i.e. when $m_1(A) > 0$ and $m_2(A) > 0$ for all $A \neq \emptyset \in D^\Theta$. Fortunately, it is important to note here that in most of the practical applications the sizes of $\mathcal{K}_1(m_1)$ and $\mathcal{K}_2(m_2)$ are much smaller than $|D^\Theta|$ because bodies of evidence generally allocate their basic belief assignments only over a subset of the hyper-power set. This makes things easier for the implementation of the classic DSm rule (1.23).

The DSm rule is actually very easy to implement. It suffices for each focal element of $\mathcal{K}_1(m_1)$ to multiply it with the focal elements of $\mathcal{K}_2(m_2)$ and then to pool all combinations which are equivalent under the algebra of sets according to figure 1.1.

The figure 1.1 represents the *DSm network architecture* of the DSm rule of combination. The first layer of the network consists in all gbba of focal elements $A_i, i = 1, \ldots, n$ of $m_1(.)$. The second layer of the network consists in all gbba of focal elements $B_j, j = 1, \ldots, k$ of $m_2(.)$. Each node of layer 2 is connected with each node of layer 1. The output layer (on the right) consists in the combined basic belief assignments of all possible intersections $A_i \cap B_j$, $i = 1, \ldots, n$ and $j = 1, \ldots, k$. The last step of the classic DSm rule (not included on the figure) consists in the compression of the output layer by regrouping (summing up) all the combined belief assignments corresponding to the same focal elements (by example if $X = A_2 \cap B_3 = A_4 \cap B_5$, then $m(X) = m(A_2 \cap B_3) + m(A_4 \cap B_5)$). If a third body of evidence provides a new gbba $m_3(.)$, the one combines it by connecting the output layer with the layer associated to $m_3(.)$, and so on. Because of commutativity and associativity properties of the classic DSm rule, the DSm network can be designed with any order of the layers.

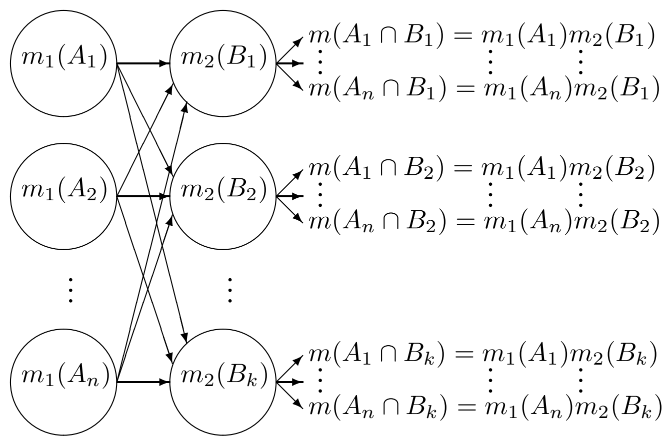


Figure 1.1: Representation of the classic DSm rule on $\mathcal{M}^f(\Theta)$

1.3.5 The hybrid DSm rule of combination

When the free DSm model $\mathcal{M}^f(\Theta)$ does not hold due to the true nature of the fusion problem under consideration which requires to take into account some known integrity constraints, one has to work with a proper hybrid DSm model $\mathcal{M}(\Theta) \neq \mathcal{M}^f(\Theta)$. In such case, the hybrid DSm rule of combination based on the chosen hybrid DSm model $\mathcal{M}(\Theta)$ for $k \geq 2$ independent sources of information is defined for all $A \in D^\Theta$ as (see chapter 4 for details):

$$m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A)\Big[S_1(A) + S_2(A) + S_3(A)\Big] \tag{1.24}$$

where $\phi(A)$ is the *characteristic non-emptiness function* of a set A, i.e. $\phi(A) = 1$ if $A \notin \boldsymbol{\emptyset}$ and $\phi(A) = 0$ otherwise, where $\boldsymbol{\emptyset} \triangleq \{\boldsymbol{\emptyset}_{\mathcal{M}}, \emptyset\}$. $\boldsymbol{\emptyset}_{\mathcal{M}}$ is the set of all elements of D^Θ which have been forced to be empty through the constraints of the model $\mathcal{M}$ and $\emptyset$ is the classical/universal empty set. $S_1(A) \equiv m_{\mathcal{M}^f(\theta)}(A)$, $S_2(A)$, $S_3(A)$ are defined by

$$S_1(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1 \cap X_2 \cap \ldots \cap X_k) = A}} \prod_{i=1}^{k} m_i(X_i) \tag{1.25}$$

$$S_2(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in\boldsymbol{\emptyset}\\ [\mathcal{U}=A]\vee[(\mathcal{U}\in\boldsymbol{\emptyset})\wedge(A=I_t)]}} \prod_{i=1}^{k} m_i(X_i) \tag{1.26}$$

$$S_3(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in D^\Theta\\ (X_1\cup X_2\cup\ldots\cup X_k)=A\\ (X_1\cap X_2\cap\ldots\cap X_k)\in\boldsymbol{\emptyset}}} \prod_{i=1}^{k} m_i(X_i) \tag{1.27}$$

with $\mathcal{U} \triangleq u(X_1) \cup u(X_2) \cup \ldots \cup u(X_k)$ where $u(X)$ is the union of all singletons θ_i that compose X and $I_t \triangleq \theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$ is the total ignorance. $S_1(A)$ corresponds to the classic DSm rule of combination for k independent sources based on the free DSm model $\mathcal{M}^f(\Theta)$; $S_2(A)$ represents the mass of all relatively and absolutely empty sets which is transferred to the total or relative ignorances; $S_3(A)$ transfers the sum of relatively empty sets to the non-empty sets.

The hybrid DSm rule of combination generalizes the classic DSm rule of combination and is not equivalent to Dempter's rule. It works for any models (the free DSm model, Shafer's model or any other hybrid models) when manipulating *precise* generalized (or eventually classical) basic belief functions. An extension of this rule for the combination of *imprecise* generalized (or eventually classical) basic belief functions is presented in chapter 6 and is not reported in this presentation of DSmT.

1.3.6 On the refinement of the frames

Let's bring here a clarification on the notion of refinement and its consequences with respect to DSmT and DST. The refinement of a set of overlapping hypotheses $\Theta = \{\theta_i, i = 1, \ldots, n\}$ consists in getting a new finer set of hypotheses θ'_i, $i = 1, \ldots, n'$, $n' > n\}$ such that we are sure that θ'_i are truly exclusive and $\cup_{i=1}^{n}\theta_i \equiv \cup_{i=1}^{n'}\theta'_i$, i.e. $\Theta = \{\theta'_i, i = 1, \ldots, n' > n\}$. The DST starts with the notion of frame of discernment (finite set of exhaustive and exclusive hypotheses). The DST assumes therefore that a refinement exists to describe the fusion problem and is achievable while DSmT does not make such assumption at its starting. The assumption of existence of a refinement process appears to us as a very strong assumption which reduces drastically the domain of applicability of the DST because the frames for most of problems described in terms of natural language manipulating vague/continuous/relative concepts cannot be formally refined at all. Such an assumption is not fundamental and is relaxed in DSmT.

As a very simple but illustrative example, let's consider Θ defined as $\Theta = \{\theta_1 = \text{Small}, \theta_2 = \text{Tall}\}$. The notions of smallness (θ_1) and tallness (θ_2) cannot be interpreted in an absolute manner actually since these notions are only defined with respect to some reference points chosen arbitrarily. Two independent sources of evidence (human "experts" here) can provide a different interpretation of θ_1 and θ_2 just because they usually do not share the same reference point. θ_1 and θ_2 represent actually fuzzy concepts carrying only a relative meaning. Moreover, these concepts are linked together by a continuous path.

Let's examine now a numerical example. Consider again the frame $\Theta = \{\theta_1 \triangleq \text{Small}, \theta_2 \triangleq \text{Tall}\}$ on the size of person with two independent witnesses providing belief masses

$$m_1(\theta_1) = 0.4 \qquad m_1(\theta_2) = 0.5 \qquad m_1(\theta_1 \cup \theta_2) = 0.1$$

$$m_2(\theta_1) = 0.6 \qquad m_2(\theta_2) = 0.2 \qquad m_2(\theta_1 \cup \theta_2) = 0.2$$

If we admit that θ_1 and θ_2 cannot be precisely refined according to the previous justification, then the result of the classic DSm rule (denoted by index $DSmc$ here) of combination yields:

$$m_{DSmc}(\emptyset) = 0 \quad m_{DSmc}(\theta_1) = 0.38 \quad m_{DSmc}(\theta_2) = 0.22 \quad m_{DSmc}(\theta_1 \cup \theta_2) = 0.02 \quad m_{DSmc}(\theta_1 \cap \theta_2) = 0.38$$

Starting now with the same information, i.e. $m_1(.)$ and $m_2(.)$, we volontary assume that a refinement is possible (even if it does not make sense actually here) in order to compare the previous result with the result one would obtain with Dempster's rule of combination. So, let's assume the existence of an hypothetical refined frame of discernment $\Theta_{ref} \triangleq \{\theta'_1 = \text{Small'}, \theta'_2 \triangleq \text{Medium}, \theta'_3 = \text{Tall'}\}$ where θ'_1, θ'_2 and θ'_3 correspond to some virtual exclusive hypotheses such that $\theta_1 = \theta'_1 \cup \theta'_2$, $\theta_2 = \theta'_2 \cup \theta'_3$ and $\theta_1 \cap \theta_2 = \theta'_2$ and where Small' and Tall' correspond respectively to a *finer notion* of smallness and tallness than in original frame Θ. Because, we don't change the information we have available (that's all we have), the initial bba $m_1(.)$ and $m_2(.)$ expressed now on the virtual refined power set $2^{\Theta_{ref}}$ are given by

$$m'_1(\theta'_1 \cup \theta'_2) = 0.4 \qquad m'_1(\theta'_2 \cup \theta'_3) = 0.5 \qquad m'_1(\theta'_1 \cup \theta'_2 \cup \theta'_3) = 0.1$$

$$m'_2(\theta'_1 \cup \theta'_2) = 0.6 \qquad m'_2(\theta'_2 \cup \theta'_3) = 0.2 \qquad m'_2(\theta'_1 \cup \theta'_2 \cup \theta'_3) = 0.2$$

Because Θ_{ref} is a refined frame, DST works and Dempster's rule applies. Because there is no positive masses for conflicting terms $\theta'_1 \cap \theta'_2$, $\theta'_1 \cap \theta'_3$, $\theta'_2 \cap \theta'_3$ or $\theta'_1 \cap \theta'_2 \cap \theta'_3$, the degree of conflict reduces to $k_{12} = 0$ and the normalization factor involved in Dempster's rule is 1 in this refined example. One gets formally, where index DS denotes here Dempster's rule, the following result:

$$\begin{aligned}
m_{DS}(\emptyset) &= 0 \\
m_{DS}(\theta'_2) &= m'_1(\theta'_1 \cup \theta'_2) m'_2(\theta'_2 \cup \theta'_3) + m'_2(\theta'_1 \cup \theta'_2) m'_1(\theta'_2 \cup \theta'_3) = 0.2 \cdot 0.4 + 0.5 \cdot 0.6 = 0.38 \\
m_{DS}(\theta'_1 \cup \theta'_2) &= m'_1(\theta'_1 \cup \theta'_2) m'_2(\theta'_1 \cup \theta'_2) + m'_1(\theta'_1 \cup \theta'_2 \cup \theta'_3) m'_2(\theta'_1 \cup \theta'_2) + m'_2(\theta'_1 \cup \theta'_2 \cup \theta'_3) m'_1(\theta'_1 \cup \theta'_2) \\
&= 0.4 \cdot 0.6 + 0.1 \cdot 0.6 + 0.2 \cdot 0.4 = 0.38 \\
m_{DS}(\theta'_2 \cup \theta'_3) &= m'_1(\theta'_2 \cup \theta'_3) m'_2(\theta'_2 \cup \theta'_3) + m'_1(\theta'_1 \cup \theta'_2 \cup \theta'_3) m'_2(\theta'_2 \cup \theta'_3) + m'_2(\theta'_1 \cup \theta'_2 \cup \theta'_3) m'_1(\theta'_2 \cup \theta'_3) \\
&= 0.2 \cdot 0.5 + 0.1 \cdot 0.2 + 0.2 \cdot 0.5 = 0.22 \\
m_{DS}(\theta'_1 \cup \theta'_2 \cup \theta'_3) &= m'_1(\theta'_1 \cup \theta'_2 \cup \theta'_3) m'_2(\theta'_1 \cup \theta'_2 \cup \theta'_3) = 0.1 \cdot 0.2 = 0.02
\end{aligned}$$

But since $\theta'_2 = \theta_1 \cap \theta_2$, $\theta'_1 \cup \theta'_2 = \theta_1$, $\theta'_2 \cup \theta'_3 = \theta_2$ and $\theta'_1 \cup \theta'_2 \cup \theta'_3 = \theta_1 \cup \theta_2$, one sees that Dempster's rule reduces to the classic DSm rule of combination, which means that the refinement of the frame Θ does not help to get a more specific (better) result from the DST when the inputs of the problem remain the same. Actually, working on Θ_{ref} with DST does not bring a difference with DSmT, but just brings an useless complexity in derivations. Note that the hybrid DSm rule of combination can also be applied on Shafer's model associated with Θ_{ref}, but it naturally provides the same result as with the classic DSm rule in this case.

If the inputs of the problem are now *changed* by re-asking (assuming that such process is possible) the sources to provide their revised belief assignents directly on Θ_{ref}, with $m'_i(\theta'_1) > 0$, $m'_i(\theta'_2) > 0$ and $m'_i(\theta'_3) > 0$ $(i = 1, 2)$ rather than on Θ, then the hybrid DSm rule of combination will be applied instead of Dempster's rule when adopting the DSmT. The fusion results will then differ, which is normal since the hybrid DSm rule is not equivalent to Dempster's rule, except when the conflict is zero.

1.3.7 On the combination of sources over different frames

In some fusion problems, it can happen that sources provide their basic belief assignment over distinct frames (which can moreover sometimes partially overlap). As simple example, let's consider two equally reliable sources of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$ providing their belief assignments repectively on distinct frames Θ_1 and Θ_2 defined as follows

$$\Theta_1 = \{P \triangleq \text{Plane}, H \triangleq \text{Helicopter}, M \triangleq \text{Missile}\}$$

$$\Theta_2 = \{S \triangleq \text{Slow motion}, F \triangleq \text{Fast motion}\}$$

In other words, $m_1(.)$ associated with $\mathcal{B}_1$ is defined either on D_1^Θ or 2_1^Θ (if Shafer's model is assumed to hold) while $m_2(.)$ associated with $\mathcal{B}_2$ is defined either on D_2^Θ or 2_2^Θ. The problem relates here to the combination of $m_1(.)$ with $m_2(.)$.

The basic solution of this problem consists in working on the global frame[8] $\Theta = \{\Theta_1, \Theta_2\}$ and in following the deconditionning method proposed by Smets in [39] based on the principle on the minimum of specificity to revise the basic belief assignments $m_1(.)$ and $m_2(.)$ on Θ. When additional information on compatibility links between elements of Θ_1 and Θ_2 is known, then the refined method proposed by Janez in [21] is preferred. Once the proper model $\mathcal{M}(\Theta)$ for Θ has been chosen to fit with the true nature of hypotheses and the revised bba $m_1^{rev}(.)$ and $m_2^{rev}(.)$ defined on D^Θ are obtained, the fusion of belief assignments is performed with the hybrid DSm rule of combination.

[8]with suppression of possible redundant elements when Θ_1 and Θ_2 overlap partially.

1.4 Comparison of different rules of combinations

1.4.1 First example

In this section, we compare the results provided by the most common rules of combinations on the following very simple numerical example where only 2 independent sources (a priori assumed equally reliable) are involved and providing their belief initially on the 3D frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$. It is assumed in this example that Shafer's model holds and thus the belief assignments $m_1(.)$ and $m_2(.)$ do not commit belief to internal conflicting information. $m_1(.)$ and $m_2(.)$ are chosen as follows:

$$m_1(\theta_1) = 0.1 \qquad m_1(\theta_2) = 0.4 \qquad m_1(\theta_3) = 0.2 \qquad m_1(\theta_1 \cup \theta_2) = 0.1$$

$$m_2(\theta_1) = 0.5 \qquad m_2(\theta_2) = 0.1 \qquad m_2(\theta_3) = 0.3 \qquad m_2(\theta_1 \cup \theta_2) = 0.1$$

These belief masses are usually represented in the form of a belief mass matrix $\mathbf{M}$ given by

$$\mathbf{M} = \begin{bmatrix} 0.1 & 0.4 & 0.2 & 0.3 \\ 0.5 & 0.1 & 0.3 & 0.1 \end{bmatrix} \tag{1.28}$$

where index i for the rows corresponds to the index of the source no. i and the indexes j for columns of $\mathbf{M}$ correspond to a given choice for enumerating the focal elements of all sources. In this particular example, index $j = 1$ corresponds to θ_1, $j = 2$ corresponds to θ_2, $j = 3$ corresponds to θ_3 and $j = 4$ corresponds to $\theta_1 \cup \theta_2$.

Now let's imagine that one finds out that θ_3 is actually truly empty because some extra and certain knowledge on θ_3 is received by the fusion center. As example, θ_1, θ_2 and θ_3 may correspond to three suspects (potential murders) in a police investigation, $m_1(.)$ and $m_2(.)$ corresponds to two reports of independent witnesses, but it turns out that finally θ_3 has provided a strong alibi to the criminal police investigator once arrested by the policemen. This situation corresponds to set up a hybrid model $\mathcal{M}$ with the constraint $\theta_3 \overset{\mathcal{M}}{\equiv} \emptyset$ (see chapter 4 for a detailed presentation on hybrid models).

Let's examine the result of the fusion in such situation obtained by the Smets', Yager's, Dubois & Prade's and hybrid DSm rules of combinations. First note that, based on the free DSm model, one would get by applying the classic DSm rule (denoted here by index $DSmc$) the following fusion result

$$m_{DSmc}(\theta_1) = 0.21 \qquad m_{DSmc}(\theta_2) = 0.11 \qquad m_{DSmc}(\theta_3) = 0.06 \qquad m_{DSmc}(\theta_1 \cup \theta_2) = 0.03$$

$$m_{DSmc}(\theta_1 \cap \theta_2) = 0.21 \qquad m_{DSmc}(\theta_1 \cap \theta_3) = 0.13 \qquad m_{DSmc}(\theta_2 \cap \theta_3) = 0.14$$

$$m_{DSmc}(\theta_3 \cap (\theta_1 \cup \theta_2)) = 0.11$$

But because of the exclusivity constraints (imposed here by the use of Shafer's model and by the non-existential constraint $\theta_3 \overset{\mathcal{M}}{\equiv} \emptyset$), the total conflicting mass is actually given by

$$k_{12} = 0.06 + 0.21 + 0.13 + 0.14 + 0.11 = 0.65 \qquad \text{(conflicting mass)}$$

- If one applies the **Disjunctive rule** (1.6), one gets:

$$\begin{aligned}
m_{\cup}(\emptyset) &= 0 \\
m_{\cup}(\theta_1) &= m_1(\theta_1)m_2(\theta_1) = 0.1 \cdot 0.5 = 0.05 \\
m_{\cup}(\theta_2) &= m_1(\theta_2)m_2(\theta_2) = 0.4 \cdot 0.1 = 0.04 \\
m_{\cup}(\theta_3) &= m_1(\theta_3)m_2(\theta_3) = 0.2 \cdot 0.3 = 0.06 \\
m_{\cup}(\theta_1 \cup \theta_2) &= [m_1(\theta_1 \cup \theta_2)m_2(\theta_1 \cup \theta_2)] + [m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2)] \\
&\quad + [m_1(\theta_1)m_2(\theta_1 \cup \theta_2) + m_2(\theta_1)m_1(\theta_1 \cup \theta_2)] \\
&\quad + [m_1(\theta_2)m_2(\theta_1 \cup \theta_2) + m_2(\theta_2)m_1(\theta_1 \cup \theta_2)] \\
&= [0.3 \cdot 0.1] + [0.01 + 0.20] + [0.01 + 015] + [0.04 + 0.03] \\
&= 0.03 + 0.21 + 0.16 + 0.007 = 0.47 \\
m_{\cup}(\theta_1 \cup \theta_3) &= m_1(\theta_1)m_2(\theta_3) + m_2(\theta_1)m_1(\theta_3) = 0.03 + 0.10 = 0.13 \\
m_{\cup}(\theta_2 \cup \theta_3) &= m_1(\theta_2)m_2(\theta_3) + m_2(\theta_2)m_1(\theta_3) = 0.12 + 0.02 = 0.14 \\
m_{\cup}(\theta_1 \cup \theta_2 \cup \theta_2) &= m_1(\theta_3)m_2(\theta_1 \cup \theta_2) = 0.02 + 0.09 = 0.11
\end{aligned}$$

- If one applies the **hybrid DSm rule** (1.24) (denoted here by index $DSmh$) for 2 sources ($k = 2$), one gets:

$$\begin{aligned}
m_{DSmh}(\emptyset) &= 0 \\
m_{DSmh}(\theta_1) &= 0.21 + 0.13 = 0.34 \\
m_{DSmh}(\theta_2) &= 0.11 + 0.14 = 0.25 \\
m_{DSmh}(\theta_1 \cup \theta_2) &= 0.03 + [0.2 \cdot 0.1 + 0.3 \cdot 0.3] + [0.1 \cdot 0.1 + 0.5 \cdot 0.4] + [0.2 \cdot 0.3] = 0.41
\end{aligned}$$

- If one applies **Smets' rule** (1.8), one gets:

$$\begin{aligned}
m_S(\emptyset) &= m(\emptyset) = 0.65 \qquad \text{(conflicting mass)} \\
m_S(\theta_1) &= 0.21 \\
m_S(\theta_2) &= 0.11 \\
m_S(\theta_1 \cup \theta_2) &= 0.03
\end{aligned}$$

- If one applies **Yager's rule** (1.9), one gets:

$$m_Y(\emptyset) = 0$$

$$m_Y(\theta_1) = 0.21$$

$$m_Y(\theta_2) = 0.11$$

$$m_Y(\theta_1 \cup \theta_2) = 0.03 + k_{12} = 0.03 + 0.65 = 0.68$$

- If one applies **Dempster's rule** (1.4) (denoted here by index DS), one gets:

$$\begin{aligned}
m_{DS}(\emptyset) &= 0 \\
m_{DS}(\theta_1) &= 0.21/[1-k_{12}] = 0.21/[1-0.65] = 0.21/0.35 = 0.600000 \\
m_{DS}(\theta_2) &= 0.11/[1-k_{12}] = 0.11/[1-0.65] = 0.11/0.35 = 0.314286 \\
m_{DS}(\theta_1 \cup \theta_2) &= 0.03/[1-k_{12}] = 0.03/[1-0.65] = 0.03/0.35 = 0.085714
\end{aligned}$$

- If one applies **Murphy's rule** (1.7), i.e average of masses, one gets:

$$\begin{aligned}
m_M(\emptyset) &= (0+0)/2 = 0 \\
m_M(\theta_1) &= (0.1+0.5)/2 = 0.30 \\
m_M(\theta_2) &= (0.4+0.1)/2 = 0.25 \\
m_M(\theta_3) &= (0.2+0.3)/2 = 0.25 \\
m_M(\theta_1 \cup \theta_2) &= (0.3+0.1)/2 = 0.20
\end{aligned}$$

But if one finds out with certainty that $\theta_3 = \emptyset$, where does $m_M(\theta_3) = 0.25$ go to? Either one accepts here that $m_M(\theta_3)$ goes to $m_M(\theta_1 \cup \theta_2)$ as in Yager's rule, or $m_M(\theta_3)$ goes to $m_M(\emptyset)$ as in Smets' rule. Catherine Murphy does not provide a solution for such a case in her paper [27].

- If one applies **Dubois & Prade's rule** (1.10), one gets because $\theta_3 \overset{\mathcal{M}}{=} \emptyset$:

$$\begin{aligned}
m_{DP}(\emptyset) &= 0 \qquad \text{(by definition of Dubois \& Prade's rule)} \\
m_{DP}(\theta_1) &= [m_1(\theta_1)m_2(\theta_1) + m_1(\theta_1)m_2(\theta_1 \cup \theta_2) + m_2(\theta_1)m_1(\theta_1 \cup \theta_2)] \\
&\quad + [m_1(\theta_1)m_2(\theta_3) + m_2(\theta_1)m_1(\theta_3)] \\
&= [0.1 \cdot 0.5 + 0.1 \cdot 0.1 + 0.5 \cdot 0.3] + [0.1 \cdot 0.3 + 0.5 \cdot 0.2] = 0.21 + 0.13 = 0.34 \\
m_{DP}(\theta_2) &= [0.4 \cdot 0.1 + 0.4 \cdot 0.1 + 0.1 \cdot 0.3] + [0.4 \cdot 0.3 + 0.1 \cdot 0.2] = 0.11 + 0.14 = 0.25 \\
m_{DP}(\theta_1 \cup \theta_2) &= [m_1(\theta_1 \cup \theta_2)m_2(\theta_1 \cup \theta_2)] + [m_1(\theta_1 \cup \theta_2)m_2(\theta_3) + m_2(\theta_1 \cup \theta_2)m_1(\theta_3)] \\
&\quad + [m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2)] \\
&= [0.30.1] + [0.3 \cdot 0.3 + 0.1 \cdot 0.2] + [0.1 \cdot 0.1 + 0.5 \cdot 0.4] = [0.03] + [0.09 + 0.02] + [0.01 + 0.20] \\
&= 0.03 + 0.11 + 0.21 = 0.35
\end{aligned}$$

Now if one adds up the masses, one gets $0+0.34+0.25+0.35 = 0.94$ which is less than 1. Therefore Dubois & Prade's rule of combination does not work when a singleton, or an union of singletons, becomes empty (in a dynamic fusion problem). The products of such empty-element columns of the mass matrix $\mathbf{M}$ are lost; this problem is fixed in DSmT by the sum $S_2(.)$ in (1.24) which transfers these products to the total or partial ignorances.

In this particular example, using the hybrid DSm rule, one transfers the product of the empty-element θ_3 column, $m_1(\theta_3)m_2(\theta_3) = 0.2 \cdot 0.3 = 0.06$, to $m_{DSmh}(\theta_1 \cup \theta_2)$, which becomes equal to $0.35+0.06 = 0.41$.

In conclusion, DSmT is a natural extension of DST and Yager's, Smets' and Dubois & Prade's approaches. When there is no singleton nor union of singletons empty, DSmT is consistent with Dubois & Prade's approach, getting the same results (because the sum $S_2(.)$ is not used in this case in the hybrid DSm rule of combination). Otherwise, Dubois & Prade's rule of combination does not work (giving a sum of fusionned masses less than 1) for dynamic fusion problems involving non existential constraints. Murphy's rule does not work either in this case because the masses of empty sets are not transferred. If the conflict is k_{12} is total (i;e. $k_{12} = 1$, DST does not work at all (one gets 0/0 in Dempster's rule of combination), while Smets' rule gives $m_S(\emptyset) = 1$ which is upon to us for the reasons explained in this introduction and in chapter 5 not necessary justified. When the conflict is total, the DSm rule is consistent with Yager's and Dubois & Prade's rules.

The general hybrid DSm rule of combination works on any models for solving static and dynmaic fusion problems and is designed for all kinds of conflict: $0 \leq m(\text{conflict}) \leq 1$. When the conflict is converging towards zero, all rules (Dempster's, Yager's, Smets', Murphy's, Dubois & Prade's, DSmT) are converging towards the same result. This fact is important because it shows the connection among all of them. But if the conflict is converging towards 1, the results among these rules diverge more and more, getting the point when some rules do not work at all (Dempster's rule). Murphy's rule is the only one which is idempotent (being the average of masses). Dubois & Prade's rule does not work in the Smets' case (when $m(\emptyset) > 0$). For models with all intersections empty (Shafer's model) and conflict 1, Dempster's rule is not defined. See below example on $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ with all θ_i, $i = 1, 2, 3, 4$ exclusive:

$$m_1(\theta_1) = 0.1 \qquad m_1(\theta_2) = 0 \qquad m_1(\theta_3) = 0.7 \qquad m_1(\theta_4) = 0$$

$$m_2(\theta_1) = 0 \qquad m_2(\theta_2) = 0.6 \qquad m_2(\theta_3) = 0 \qquad m_2(\theta_4) = 0.4$$

Using Dempster's rule, one gets 0/0, undefined. Conflicting mass is 1.

Yager's rule provides in this case $m_Y(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = 1$ which does not bring specific information, while Smets' rule gives $m(\emptyset) = 1$ which is also not very useful. Murphy's rule gives $m_M(\theta_1) = 0.15$, $m_M(\theta_2) = 0.30$, $m_M(\theta_3) = 0.35$ and $m_M(\theta_4) = 0.20$ which is very specific while the hybrid DSm rule provides $m_{DSmh}(\theta_1 \cup \theta_2) = 0.18$, $m_{DSmh}(\theta_1 \cup \theta_4) = 0.12$, $m_{DSmh}(\theta_2 \cup \theta_3) = 0.42$ and $m_{DSmh}(\theta_3 \cup \theta_4) = 0.28$ which is less specific than Murphy's result but characterizes adequately the internal conflict between sources after the combination and partial ignorances.

The disjunctive rule gives in this last example $m_\cup(\theta_1 \cup \theta_2) = m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2) = 0.18$. Similarly, one gets $m_\cup(\theta_1 \cup \theta_4) = 0.12$, $m_\cup(\theta_2 \cup \theta_3) = 0.42$ and $m_\cup(\theta_3 \cup \theta_4) = 0.28$. This coincides with the hybrid DSm rule when all intersections are empty.

1.4.2 Second example

This example is an extension of Zadeh's example discussed in chapter 5. Let's consider two independent sources of evidences over the frame $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and assume that Shafer's model holds. The basic belief assignments are chosen as follows:

$$m_1(\theta_1) = 0.998 \qquad m_1(\theta_2) = 0 \qquad m_1(\theta_3) = 0.001 \qquad m_1(\theta_4) = 0.001$$

$$m_2(\theta_1) = 0 \qquad m_2(\theta_2) = 0.998 \qquad m_2(\theta_3) = 0 \qquad m_2(\theta_4) = 0.02$$

In this simple numerical example, Dempster's rule of combination gives the counter-intuitive result

$$m_{DS}(\theta_4) = \frac{0.001 \cdot 0.002}{0.998 \cdot 0.998 + 0.998 \cdot 0.002 + 0.998 \cdot 0.001 + 0.998 \cdot 0.001 + 0.001 \cdot 0.002} = \frac{0.000002}{0.000002} = 1$$

Yager's rule gives $m_Y(\theta_4) = 0.000002$ and $m_Y(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = 0.999998$.

Smets' rule gives $m_S(\theta_4) = 0.000002$ and $m_S(\emptyset) = 0.999998$.

Murphy's rule gives $m_M(\theta_1) = 0.499$, $m_M(\theta_2) = 0.499$, $m_M(\theta_3) = 0.0005$ and $m_M(\theta_4) = 0.0015$.

Dubois & Prade's rule gives $m_{DP}(\theta_4) = 0.000002$, $m_{DP}(\theta_1 \cup \theta_2) = 0.996004$, $m_{DP}(\theta_1 \cup \theta_4) = 0.001996$, $m_{DP}(\theta_2 \cup \theta_3) = 0.000998$, $m_{DP}(\theta_2 \cup \theta_4) = 0.000998$ and $m_{DP}(\theta_3 \cup \theta_4) = 0.000002$. Dubois & Prade's rule works only in Shafer's model $\mathcal{M}^0(\Theta)$, i.e. when all intersections are empty. For other hybrid models, Dubois & Prade's rule of combination fails to provide a reliable and reasonable solution to the combination of sources (see next example).

The classic DSm rule of combination provides $m_{DSmc}(\theta_4) = 0.000002$, $m_{DSmc}(\theta_1 \cap \theta_2) = 0.996004$, $m_{DSmc}(\theta_1 \cap \theta_4) = 0.001996$, $m_{DSmc}(\theta_2 \cap \theta_3) = 0.000998$, $m_{DSmc}(\theta_2 \cap \theta_4) = 0.000998$ and $m_{DSmc}(\theta_3 \cap \theta_4) =$

0.000002. If one now applies the hybrid DSm rule since one assumes here that Shafer's model holds, one gets the same result as Dubois & Prade's. The disjunctive rule coincides with Dubois & Prade's rule and the hybrid DSm rule when all intersections are empty.

1.4.3 Third example

Here is an exemple for the Smets' case (i.e. TBM) when $m(\emptyset) > 0$ where Dubois & Prade's rule of combination does not work. Let's consider the following extended[9] belief assignments

$$m_1(\emptyset) = 0.2 \qquad m_1(\theta_1) = 0.4 \qquad m_1(\theta_2) = 0.4$$

$$m_2(\emptyset) = 0.3 \qquad m_2(\theta_1) = 0.6 \qquad m_2(\theta_2) = 0.1$$

In this specific case, the Dubois & Prade's rule of combination gives (assuming all intersections empty)

$$\begin{aligned}
m_{DP}(\emptyset) &= 0 \qquad \text{(by definition)} \\
m_{DP}(\theta_1) &= m_1(\theta_1)m_2(\theta_1) + [m_1(\emptyset)m_2(\theta_1) + m_2(\emptyset)m_1(\theta_1)] = 0.24 + [0.12 + 0.12] = 0.48 \\
m_{DP}(\theta_2) &= m_1(\theta_2)m_2(\theta_2) + [m_1(\emptyset)m_2(\theta_2) + m_2(\emptyset)m_1(\theta_2)] = 0.04 + [0.02 + 0.12] = 0.18 \\
m_{DP}(\theta_1 \cup \theta_2) &= m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2) = 0.04 + 0.24 = 0.28
\end{aligned}$$

The sum of masses is $0.48 + 0.18 + 0.28 = 0.94 < 1$. Where goes the mass $m_1(\emptyset)m_2(\emptyset) = 0.2 \cdot 0.3 = 0.06$? When using the hybrid DSm rule of combination, one gets $m_{DSmh}(\emptyset) = 0$, $m_{DSmh}(\theta_1) = 0.48$, $m_{DSmh}(\theta_2) = 0.18$ and

$$m_{DSmh}(\theta_1 \cup \theta_2) = [m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2)] + [m_1(\emptyset)m_2(\emptyset)] = [0.28] + [0.2 \cdot 0.3] = 0.34$$

and the masses add up to 1.

The disjunctive rule gives in this example

$$\begin{aligned}
m_{\cup}(\theta_1) &= m_1(\theta_1)m_2(\theta_1) + [m_1(\emptyset)m_2(\theta_1) + m_2(\emptyset)m_1(\theta_1)] = 0.24 + [0.12 + 0.12] = 0.48 \\
m_{\cup}(\theta_2) &= m_1(\theta_2)m_2(\theta_2) + [m_1(\emptyset)m_2(\theta_2) + m_2(\emptyset)m_1(\theta_2)] = 0.04 + [0.02 + 0.12] = 0.18 \\
m_{\cup}(\theta_1 \cup \theta_2) &= m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2) = 0.04 + 0.24 = 0.28 \\
m_{\cup}(\emptyset) &= m_1(\emptyset)m_2(\emptyset) = 0.06 > 0
\end{aligned}$$

One gets the same results for $m_{\cup}(\theta_1)$, $m_{\cup}(\theta_2)$ as with Dubois & Prade's rule and as with the hybrid DSm rule. The distinction is in the reallocation of the empty mass $m_1(\emptyset)m_2(\emptyset) = 0.06$ to $\theta_1 \cup \theta_2$ in the hybrid DSm rule, while in Dubois & Prade's and disjunctive rules it is not.

[9] We mean here non-normalized masses allowing weight of evidence on the empty set as in the TBM of Smets.

A major difference among the hybrid DSm rule and all other combination rules is that DSmT uses from the beginning a hyper-power set, which includes intersections, while other combination rules need to do a refinement in order to get intersections.

1.4.4 Fourth example

Here is another example where Dempster's rule does not work properly (this is different from Zadeh's example). Let's consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and assume that Shafer's model holds. The basic belief assignments are now chosen as follows:

$$m_1(\theta_1) = 0.99 \qquad m_1(\theta_2) = 0 \qquad m_1(\theta_3 \cup \theta_4) = 0.01$$

$$m_2(\theta_1) = 0 \qquad m_2(\theta_2) = 0.98 \qquad m_2(\theta_3 \cup \theta_4) = 0.02$$

Applying Dempster's rule, one gets $m_{DS}(\theta_1) = m_{DS}(\theta_2) = 0$ and

$$m_{DS}(\theta_3 \cup \theta_4) = \frac{0.01 \cdot 0.02}{1 - [0.99 \cdot 0.98 + 0.99 \cdot 0.02 + 0.98 \cdot 0.01]} = \frac{0.0002}{1 - 0.9998} = \frac{0.0002}{0.0002} = 1$$

which is abnormal.

The hybrid DSm rule gives $m_{DSmh}(\theta_1 \cup \theta_2) = 0.99 \cdot 0.98 = 0.9702$, $m_{DSmh}(\theta_1 \cup \theta_3 \cup \theta_4) = 0.0198$, $m_{DSmh}(\theta_2 \cup \theta_3 \cup \theta_4) = 0.0098$ and $m_{DSmh}(\theta_3 \cup \theta_4) = 0.0002$. In this case, Dubois & Prade's rule gives the same results as the hybrid DSm rule. The disjunctive rule provides a combined belief assignment $m_{\cup}(.)$ which is same as $m_{DSmh}(.)$ and $m_{DP}(.)$.

Yager's rule gives $m_Y(\theta_3 \cup \theta_4) = 0.0002$, $m_Y(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = 0.9998$ and Smets' rule gives $m_S(\theta_3 \cup \theta_4) = 0.0002$, $m_S(\emptyset) = 0.9998$. Both Yager's and Smets' results are less specific than the result obtained with the hybrid DSm rule. There is a loss of information somehow when using Yager's or Smets' rules.

1.4.5 Fifth example

Suppose one extends Dubois & Prade's rule from the power set 2^Θ to the hyper-power set D^Θ. It can be shown that Dubois & Prade's rule does not work when (because $S_2(.)$ term is missing):

a) at least one singleton is empty and the element of its column are all non zero

b) at least an union of singletons is empty and elements of its column are all non zero

c) or at least an intersection is empty and the elements of its column are non zero

Here is an example with intersection (Dubois & Prade's rule extended to the hyper-power set). Let's consider two independent sources on $\Theta = \{\theta_1, \theta_2\}$ with

$$m_1(\theta_1) = 0.5 \qquad m_1(\theta_2) = 0.1 \qquad m_1(\theta_1 \cap \theta_2) = 0.4$$

$$m_2(\theta_1) = 0.1 \qquad m_2(\theta_2) = 0.6 \qquad m_2(\theta_1 \cap \theta_2) = 0.3$$

Then the extended Dubois & Prade rule on the hyper-power set gives $m_{DP}(\emptyset) = 0$, $m_{DP}(\theta_1) = 0.05$, $m_{DP}(\theta_2) = 0.06$, $m_{DP}(\theta_1 \cap \theta_2) = 0.04 \cdot 0.3 + 0.5 \cdot 0.6 + 0.5 \cdot 0.3 + 0.1 \cdot 0.4 + 0.1 \cdot 0.3 + 0.6 \cdot 0.4 = 0.89$.

Now suppose one finds out that $\theta_1 \cap \theta_2 = \emptyset$, then the revised masses become

$$\begin{aligned}
m'_{DP}(\emptyset) &= 0 \qquad \text{(by definition)}\\
m'_{DP}(\theta_1) &= 0.05 + [m_1(\theta_1)m_2(\theta_1 \cap \theta_2) + m_2(\theta_1)m_1(\theta_1 \cap \theta_2)] = 0.05 + [0.5 \cdot 0.3 + 0.1 \cdot 0.4] = 0.24\\
m'_{DP}(\theta_2) &= 0.06 + [m_1(\theta_2)m_2(\theta_1 \cap \theta_2) + m_2(\theta_2)m_1(\theta_1 \cap \theta_2)] = 0.06 + [0.1 \cdot 0.3 + 0.6 \cdot 0.4] = 0.33\\
m'_{DP}(\theta_1 \cup \theta_2) &= m_1(\theta_2)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2) = 0.5 \cdot 0.6 + 0.1 \cdot 0.1 = 0.31
\end{aligned}$$

The sum of the masses is $0.24 + 0.33 + 0.31 = 0.88 < 1$. The mass product $m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_2) = 0.4 \cdot 0.3 = 0.12$ has been lost.

When applying the classic DSm rule in this case, one gets exactly the same results as Dubois & Prade, i.e. $m_{DSmc}(\emptyset) = 0$, $m_{DSmc}(\theta_1) = 0.05$, $m_{DSmc}(\theta_2) = 0.06$, $m_{DSmc}(\theta_1 \cap \theta_2) = 0.89$. Now if one takes into account the integrity constraint $\theta_1 \cap \theta_2 = \emptyset$ and using the hybrid DSm rule of combination, one gets

$$\begin{aligned}
m_{DSmh}(\emptyset) &= 0 \qquad \text{(by definition)}\\
m_{DSmh}(\theta_1) &= 0.05 + [m_1(\theta_1)m_2(\theta_1 \cap \theta_2) + m_2(\theta_1)m_1(\theta_1 \cap \theta_2)] = 0.05 + [0.5 \cdot 0.3 + 0.1 \cdot 0.4] = 0.24\\
m_{DSmh}(\theta_2) &= 0.06 + [m_1(\theta_2)m_2(\theta_1 \cap \theta_2) + m_2(\theta_2)m_1(\theta_1 \cap \theta_2)] = 0.06 + [0.1 \cdot 0.3 + 0.6 \cdot 0.4] = 0.33\\
m_{DSmh}(\theta_1 \cup \theta_2) &= [m_1(\theta_2)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2)] + [\underbrace{m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_2)}_{S_2 \text{ in hybrid DSm rule eq.}}] = [0.31] + [0.12] = 0.43
\end{aligned}$$

Thus the sum of the masses obtained by the hybrid DSm rule of combination is $0.24 + 0.33 + 0.43 = 1$.

The disjunctive rule extended on the hyper-power set gives for this example

$$\begin{aligned}
m_\cup(\emptyset) &= 0\\
m_\cup(\theta_1) &= [m_1(\theta_1)m_2(\theta_1)] + [m_1(\theta_1)m_2(\theta_1 \cap \theta_2) + m_2(\theta_1)m_1(\theta_1 \cap \theta_2)] = 0.05 + [0.15 + 0.04] = 0.24\\
m_\cup(\theta_2) &= [m_1(\theta_2)m_2(\theta_2)] + [m_1(\theta_2)m_2(\theta_1 \cap \theta_2) + m_2(\theta_2)m_1(\theta_1 \cap \theta_2)] = 0.06 + [0.15 + 0.04] = 0.33\\
m_\cup(\theta_1 \cup \theta_2) &= [m_1(\theta_2)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2)] = 0.31\\
m_\cup(\theta_1 \cap \theta_2) &= m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_2) = 0.4 \cdot 0.3 = 0.12
\end{aligned}$$

If now one finds out that $\theta_1 \cap \theta_2 = \emptyset$, then the revised masses $m'_{\cup}(.)$ become $m'_{\cup}(\theta_1) = m_{\cup}(\theta_1)$, $m'_{\cup}(\theta_2) = m_{\cup}(\theta_2)$, $m'_{\cup}(\theta_1 \cup \theta_2) = m_{\cup}(\theta_1 \cup \theta_2)$ but $m'_{\cup}(\emptyset) \equiv m_{\cup}(\theta_1 \cap \theta_2) = 0.12 > 0$.

1.5 Summary

DSmT has to be viewed as a general flexible *Bottom-Up* approach for managing uncertainty and conflicts for a wide class of static or dynamic fusion problems where the information to combine is modelled as a finite set of belief functions provided by different independent sources of evidence. The development of DSmT emerged from the fact that the conflict between the sources of evidence arises not only from the unreliability of sources themselves (which can be handled by classical discounting methods), but also from a different interpretation of the frame itself by the sources of evidence due to their limited knowlege and own (local) experience; not to mention the fact that elements of the frame cannot be truly refined at all in many problems involving only fuzzy and continuous concepts. Based on this matter of fact, DSmT proposes, according to the general block-scheme in Figure 1.2, a new appealing mathematical framework.

Here are the major steps for managing uncertain and conflicting information arising from independent sources of evidence in the DSmT framework, once expressed in terms of basic belief functions:

1. **Bottom Level**: The ground level of DSmT is to start from the free DSm model $\mathcal{M}^f(\Theta)$ associated with the frame Θ and the notion of hyper-power set (free Dedekind's lattice) D^Θ. At this level, DSmT provides a general commutative and associative rule of combination of evidences (the conjunctive consensus) to work on $\mathcal{M}^f(\Theta)$.

2. **Higher Level** (only used when necessary): Depending on the absolute true intrinsic nature (assumed to be known by the fusion center) of the elements of the frame Θ of the fusion problem under consideration (which defines a set of integrity constraints on $\mathcal{M}^f(\Theta)$ leading to a particular hybrid DSm model $\mathcal{M}(\Theta)$), DSmT automatically adapts the combination process to work on any hybrid DSm model with the general hybrid DSm rule of combination explaine in details in chapter 4. The taking into account of an integrity constraint consists just in forcing some elements of the Dedekind's lattice D^Θ to be empty, when they truly are, given the problem under consideration.

3. **Decision-Making**: Once the combination is obtained after step 1 (or step 2 when necessary), the Decision-making step follows. Although no real general consensus has emerged in literature over last 30 years to give a well-accepted solution for the decision-making problem in the DST framework, we follow here Smets' idea and his justifications to work at the pignistic level [42] rather than at the credal level when a final decision has to be taken from any combined belief mass $m(.)$. A generalized pignistic transformation is then proposed in chapter 7 based on DSmT.

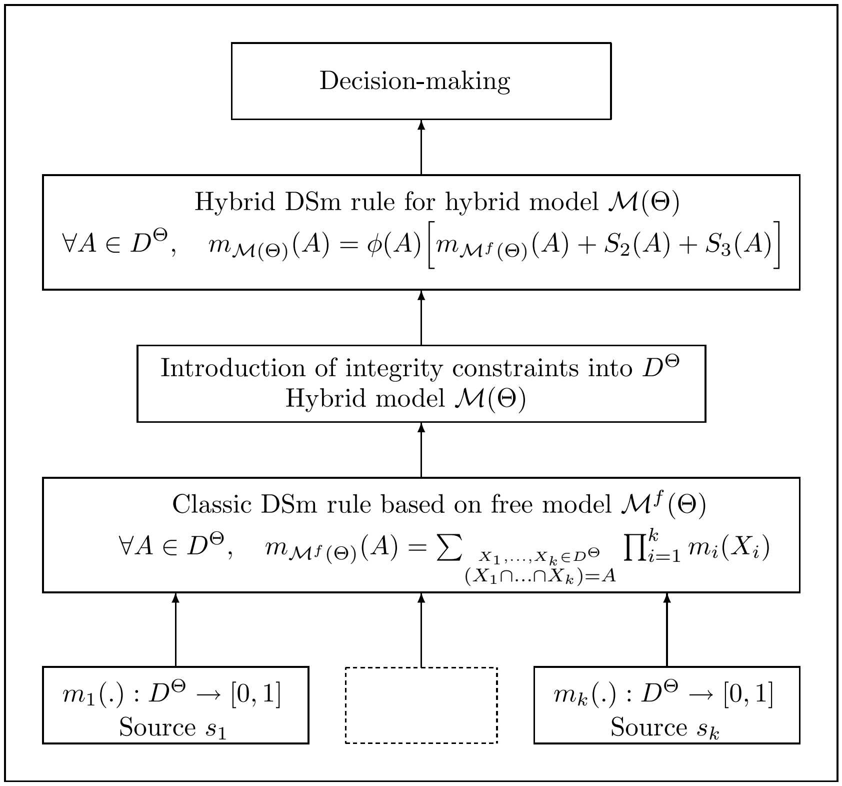


Figure 1.2: Block Scheme of the principle for the DSm fusion

The introduction of a specific integrity constraint in step 2 is like *pushing an elevator button* for going a bit up in the complexity of the processing for managing uncertainty and conflict through the hybrid DSm rule of combination. If one needs to go to a higher level, then one can take into account several integrity constraints as well in the framework of DSmT. If we finally want to take into account *all possible exclusivity constraints only* (when we really know that all elements of the frame of the given problem are truly exclusive), then we go directly to the Top Level (i.e. Shafer's model which serves as foundation for Shafer's mathematical theory of evidence), but we still apply the hybrid DSm rule instead of Dempster's rule of combination. The DSmT approach for modelling the frame and combining information is more general than previous approaches which have been mainly[10] based on the Shafer model (which is a very specific and constrained DSm hybrid model) and works for static fusion problems.

The DSmT framework can easily handle *not only* exclusivity constraints, *but also* non existential constraints or mixed constraints as well which is very useful in some dynamic fusion problems as it is shown in chapter 4. Depending on the nature of the problem, we claim that it is unnecessary to try working at the Top Level (as DST does), when working directly at a lower level is sufficient to manage properly the information to combine using the hybrid DSm rule of combination.

[10]except the Transferable Belief Model of Smets [41] and the trade-off/averaging combination rules.

It is also important to reemphasize here that the general hybrid DSm rule of combination is definitely not equivalent to Dempster's rule of combination (and to all its alternatives involving conjunctive consensus based on the Top level and especially when working with dynamic problems) because DSmT allows to work at any level of modelling for managing uncertainty and conflicts, depending on the intrinsic nature of the problem. The hybrid DSm rule and Dempster's rule do not provide the same results even when working on Shafer's model as it has been shown in examples of the previous section and explained in details in forthcoming chapters 4 and 5.

DSmT differs from DST because it is based on the free Dedekind lattice. It works for any model (free DSm model and hybrid models - including Shafer's model as a special case) which fits adequately with the true nature of the fusion problem under consideration. This ability of DSmT allows to deal formally with any fusion problems expressed in terms of belief functions which can mix discrete concepts with vague/continuous/relative concepts. The DSmT deals with static and dynamic fusion problematics in the same theoretical way taking into account the integrity constraints into the model which are considered either as static or eventually changing with time when necessary. The general hybrid DSm rule of combination of independent sources of evidence works for all possible static or dynamic models and does not require a normalization step. It differs from Dempster's rule of combination and from all its concurrent alternatives. The hybrid DSm rule of combination has been moreover extended to work for the combination of imprecise admissible belief assignments as well. The approach proposed by the DSmT to attack the fusion problematic throughout this book is therefore totally new both by its foundations, its applicability and the solution provided.

1.6 References

[1] Berzati D., Haenni R., Kohlas J.. *Probabilistic Argumentation Systems and Abduction*, Tech. Report 00-2. Institute of Informatics, University of Fribourg, 2000.

[2] Berzati D., Haenni R., Kohlas J.. *Probabilistic Argumentation Systems and Abduction*, C. Baral and M. Truszczynski (eds.), NMR2000, 8th International Workshop on Non-Monotonic Reasoning, Breckenridge, Colorado. 2000.

[3] Bloch I., *Fondements des probabilités et des croyances: une discussion des travaux de Cox et Smets*, 15ième Colloque GRETSI, Juan-Les-Pins, France, Sept. 18-21, 1995, pp. 909-912.

[4] Bloch I., *Incertitude, imprécision et additivité en fusion de données: point de vue historique*, Traitement du Signal, Vol. 13, no. 4, pp. 267-288, 1996.

[5] Challa S., Koks D., *Bayesian and Dempster-Shafer fusion*, Sādhanā, Vol. 29, Part 2, pp. 145-174, April 2004 (`http://www.ias.ac.in/sadhana/`).

[6] Cheng Y., Kashyap R.L., *Study of the different methods for combining evidence*, SPIE Vol. 635 Applications of Artificial Intelligence III, pp. 384-393, 1986.

[7] Comtet L., *Sperner Systems*, sec.7.2 in Advanced Combinatorics: The Art of Finite and Infinite Expansions, D. Reidel Publ. Co., pp. 271-273, 1974.

[8] Dedekind R. *Über Zerlegungen von Zahlen durch ihre grössten gemeinsammen Teiler*, In Gesammelte Werke, Bd. 1. pp. 103-148, 1897.

[9] Dezert J., *Foundations for a new theory of plausible and paradoxical reasoning*, Inform. & Secur. J., Semerdjiev Ed., Bulg. Acad. of Sci., Vol. 9, 2002.

[10] Dezert J., *Fondations pour une nouvelle théorie du raisonnement plausible et paradoxal*, ONERA Tech. Rep. RT 1/06769/DTIM, Jan. 2003.

[11] Dubois D., Prade H., *A Set-Theoretic View of Belief Functions*, International Journal of General Systems, pp.193-226, Vol.12, 1986.

[12] Dubois D., Prade H., *On the unicity of Dempster rule of combination*, International Journal of Intelligent Systems, Vol. 1, pp 133-142, 1986.

[13] Dubois D., Prade H., *Representation and combination of uncertainty with belief functions and possibility measures*, Computational Intelligence, 4, pp. 244-264, 1988.

[14] Haenni R., *Propositional Argumentation Systems and Symbolic Evidence Theory*, Ph.D. Thesis. Institut für Informatik, Universität Freiburg. 1996.

[15] Haenni R., Kohlas J., Lehmann N., *Probabilistic Argumentation Systems*, Tech. Report 99-9. Institute of Informatics, University of Fribourg. 1999.

[16] Haenni R., Lehmann N., *Probabilistic Argumentation Systems: a New Perspective on Dempster-Shafer Theory*, Tech. Report 01-30. Department of Informatics, University of Fribourg. 2001.

[17] Haenni R., Kohlas J., Lehmann N., *Probabilistic Argumentation Systems*, Handbook of Defeasible Reasoning and Uncertainty Management Systems, Volume 5: Algorithms for Uncertainty and Defeasible Reasoning, pp. 221-288, J. Kohlas, S. Moral (eds). Kluwer, Dordrecht. 2001.

[18] Haenni R., Anrig B., Kohlas J., N. Lehmann N., *A Survey on Probabilistic Argumentation*, ECSQARU'01, 6th European Conference on Symbolic and Quantitative Approaches to Reasoning under Uncertainty, Workshop: Adventures in Argumentation, pp. 19-25, Toulouse, France, 2001.

[19] Haenni R., *Are alternatives to Dempster's rule of combination real alternatives ? Comments on " About the belief function combination and the conflict management problem"*, Information Fusion Journal, Elsevier Publisher, Vol. 3, No. 3, pp. 237-239, 2002.

[20] Haenni R., Lehmann N.. *Probabilistic Argumentation Systems: a New Perspective on Dempster-Shafer Theory*, International Journal of Intelligent Systems (Special Issue: the Dempster-Shafer Theory of Evidence). 18 (1), 93-106. 2003.

[21] Janez F., *Fusion de sources d'information définies sur des référentiels non exhaustifs différents*, Ph.D. Thesis, Nov. 13, 1996, Angers Univ., France (also available as Onera Technical Note 1997-2).

[22] Kholas J., Monney P.A., *A mathematical theory of hints: An approach to Dempster-Shafer theory of evidence*, Lectures Notes in Economics and Mathematical Systems, No. 425, Springer-Verlag, 1995.

[23] Kholas J., *Mathematical foundations of evidence theory*, in G. Coletti, D. Dubois, and R. Scozzafava, editors, Mathematical models for handling partial knowledge in artificial intelligence, pp. 31-64, Plenum Press, 1995.

[24] Klawonn F., Smets Ph., *The dynamic of belief in the Transferable Belief Model and specialization-generalization matrices*, In D. Dubois, M.P. Wellman, B. D'Ambrosio, & P. Smets (Eds.), Uncertainty in Artificial Intelligence 92, pp. 130-137, Morgan Kaufman, San Mateo, CA, 1992.

[25] Lefevre E., Colot O., Vannoorenberghe P. *Belief functions combination and conflict management*, Information Fusion Journal, Elsevier Publisher, Vol. 3, No. 2, pp. 149-162, 2002.

[26] Lefevre E., Colot O., Vannoorenberghe P., *Reply to the Comments of R. Haenni on the paper "Belief functions combination and conflict management"*, Information Fusion Journal, Elsevier Publisher, Vol. 4, pp. 63-65, 2003.

[27] Murphy C.K.,*Combining belief functions when evidence conflicts*, Decision Support Systems, Elsevier Publisher, Vol. 29, pp. 1-9, 2000.

[28] Pearl J., *Probabilistic reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, San Mateo, CA, 1988.

[29] Pearl J., *Reasoning with belief functions: an analysis of compatibility*, International Journal of Approximate Reasoning, Vol. 4, pp. 363-389, 1990.

[30] Prade H., Testemale C., *Generalized database relational algebra for the treatment of incomplete/uncertain information and vague queries*, Information Sciences, Vol. 34, pp. 115-143, 1984.

[31] Ruspini E.H., *The logical foundations of evidential reasoning*, Tech. note 408, SRI International, Menlo Park, CA, USA, 1986.

[32] Ruspini E.H., *Epistemic logics, probability and the calculus of evidence*, in IJCAI-97 (Ed.), International Joint Conference on Artificial Intelligence, pp. 924-931, Morgan Kaufman, San Mateo, CA, 1987..

[33] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[34] Shafer G., *Perspectives on the Theory and Practice of Belief Functions*, International Journal of Approximate Reasoning, Vol.4, pp.323-362,1990.

[35] Sloane N.J.A., *The On-line Encyclopedia of Integer Sequences 2003*, (Sequence No. A014466), `http://www.research.att.com/~njas/sequences/`.

[36] Smarandache F. (Editor), *Proceedings of the First International Conference on Neutrosophics*, Univ. of New Mexico, Gallup Campus, NM, USA, 1-3 Dec. 2001, Xiquan, Phoenix, 2002.

[37] Smets Ph., *The combination of evidence in the Transferable Belief Model*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 12, No. 5, pp. 447-458, 1990.

[38] Smets Ph., *Probability of provability and belief functions*, Logique et analyse, 133, pp177-195, 1991.

[39] Smets Ph.,*Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem*, International Journal of Approximate reasoning, Vol. 9, pp. 1-35, 1993.

[40] Smets Ph.,*Probability of deductibility and belief functions*, in M. Clarke, R? Kruse, and S. Moral (Eds.), Symbolic and quantitative approaches to uncertainty, pp. 332-341, Springer-Verlag, Berlin,1993.

[41] Smets Ph., Kennes R., *The transferable belief model*, Artif. Intel., 66(2), pp. 191-234, 1994.

[42] Smets Ph., *Data Fusion in the Transferable Belief Model*, Proceedings of the 3rd International Conference on Information Fusion, Fusion 2000, Paris, July 10-13, 2000, pp PS21-PS33.

[43] Voorbraak F., *On the justification of Dempster's rule of combination*, Artificial Intelligence, 48, pp. 171-197, 1991. (see `http://turing.wins.uva.nl/~fransv/#pub`).

[44] Yager R.R., *Hedging in the combination of evidence*, Journal of Information & Optimization Sciences, Analytic Publishing Co., Vol. 4, No. 1, pp. 73-81, 1983.

[45] Yager R. R., *Hedging in the combination of evidence*, Journal of Information and Optimization Science, Vol. 4, No. 1, pp. 73-81, 1983.

[46] Yager R. R., *On the relationships of methods of aggregation of evidence in expert systems*, Cybernetics and Systems, Vol. 16, pp. 1-21, 1985.

[47] Yager R.R., *On the Dempster-Shafer framework and new combination rules*, Information Sciences, Vol. 41, pp. 93–138, 1987.

[48] Yager R. R., Fedrizzi M., Kacprzyk J. (Editors), *Advances in the Dempster-Shafer theory of evidence*, John Wiley & Sons: New York, 1994, 608 pages (see `http://www.panix.com/~yager/HP/pubs.html`).

[49] Wierzchoń S.T., Kłopotek M.A., *Evidential reasoning. An interpretative investigation*,Wydawnictwo Akademii Podlaskiej Publisher, PL ISSN 0860-2719, 304 pages, Poland, January 2002 (see `http://www.ipipan.waw.pl/staff/m.klopotek/mak/book2a.htm`).

[50] Zadeh L., *On the validity of Dempster's rule of combination*, Memo M 79/24, Univ. of California, Berkeley, 1979.

[51] Zadeh L., *Review of Mathematical theory of evidence, by Glenn Shafer*, AI Magazine, Vol. 5, No. 3, pp. 81-83, 1984.

[52] Zadeh L., *A simple view of the Dempster-Shafer theory of evidence and its implications for the rule of combination*, Berkeley Cognitive Science Report No. 33, University of California, Berkeley, CA, 1985.

[53] Zadeh L., *A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination*, AI Magazine 7, No.2, pp. 85-90, 1986.

Chapter 2

The generation of hyper-power sets

Jean Dezert	Florentin Smarandache
ONERA	Department of Mathematics
29 Av. de la Division Leclerc	University of New Mexico
92320 Châtillon	Gallup, NM 8730
France	U.S.A.

Abstract: *The development of DSmT is based on the notion of Dedekind's lattice, called also hyper-power set in the DSmT framework, on which is defined the general basic belief assignments to be combined. In this chapter, we explain the structure of the hyper-power set, give some examples of hyper-power sets and show how they can be generated from isotone Boolean functions. We also show the interest to work with the hyper-power set rather than the power set of the refined frame of discernment in terms of complexity.*

2.1 Introduction

One of the cornerstones of the DSmT is the notion of Dedekind's lattice, coined as hyper-power set by the authors in literature, which will be defined in next section. The starting point is to consider $\Theta = \{\theta_1, \ldots, \theta_n\}$ as a set of n elements which cannot be precisely defined and separated so that no refinement of Θ in a new larger set Θ_{ref} of disjoint elementary hypotheses is possible. This corresponds to the free DSm model. This model is justified by the fact that in some fusion problems (mainly those manipulating vague or continuous concepts), the refinement of the frame is just impossible to obtain; nevertheless the fusion still applies when working on Dedekind's lattice and based on the DSm rule of

This chapter is based on a paper [6] presented during the International Conference on Information Fusion, Fusion 2003, Cairns, Australia, in July 2003 and is reproduced here with permission of the International Society of Information Fusion.

combination. With the DSmT approach, the refinement of the frame is not prerequisite for managing properly the combination of evidences and one can abandon Shafer's model in general. Even if Shafer's model is justified and adopted in some cases, the hybrid DSm rule of combination appears to be a new interesting and preferred alternative for managing high conflicting sources of evidence. Our approach actually follows the footprints of our predecessors like Yager [23] and Dubois and Prade [7] to circumvent the problem of the applicability of Dempster's rule face to high conflicting sources of evidence but with a new mathematical framework. The major reason for attacking the problem directly from the bottom level, i.e. the free DSm model comes from the fact that in some real-world applications observations/concepts are not unambiguous. The ambiguity of observations is explained by Goodman, Mahler and Nguyen in [9] pp. 43-44. Moreover, the ambiguity can also come from the granularity of knowledge, known as Pawlak's indiscernability or roughness [15].

2.2 Definition of hyper-power set D^Θ

The *hyper-power set* D^Θ is defined as the set of all composite propositions built from elements of Θ with $\cup$ and $\cap$ (Θ generates D^Θ under operators $\cup$ and $\cap$) operators such that

1. $\emptyset, \theta_1, \ldots, \theta_n \in D^\Theta$.
2. If $A, B \in D^\Theta$, then $A \cap B \in D^\Theta$ and $A \cup B \in D^\Theta$.
3. No other elements belong to D^Θ, except those obtained by using rules 1 or 2.

The dual (obtained by switching $\cup$ and $\cap$ in expressions) of D^Θ is itself. There are elements in D^Θ which are self-dual (dual to themselves), for example α_8 for the case when $n = 3$ in the example given in the next section. The cardinality of D^Θ is majored by 2^{2^n} when $\text{Card}(\Theta) = |\Theta| = n$. The generation of hyper-power set D^Θ is closely related with the famous Dedekind problem [4, 3] on enumerating the set of monotone Boolean functions as it will be presented in the sequel with the generation of the elements of D^Θ.

2.3 Example of the first hyper-power sets

- In the degenerate case ($n = 0$) where $\Theta = \{\}$, one has $D^\Theta = \{\alpha_0 \triangleq \emptyset\}$ and $|D^\Theta| = 1$.
- When $\Theta = \{\theta_1\}$, one has $D^\Theta = \{\alpha_0 \triangleq \emptyset, \alpha_1 \triangleq \theta_1\}$ and $|D^\Theta| = 2$.
- When $\Theta = \{\theta_1, \theta_2\}$, one has $D^\Theta = \{\alpha_0, \alpha_1, \ldots, \alpha_4\}$ and $|D^\Theta| = 5$ with $\alpha_0 \triangleq \emptyset$, $\alpha_1 \triangleq \theta_1 \cap \theta_2$, $\alpha_2 \triangleq \theta_1$, $\alpha_3 \triangleq \theta_2$ and $\alpha_4 \triangleq \theta_1 \cup \theta_2$.
- When $\Theta = \{\theta_1, \theta_2, \theta_3\}$, the elements of $D^\Theta = \{\alpha_0, \alpha_1, \ldots, \alpha_{18}\}$ and $|D^\Theta| = 19$ (see [5] for details) are now given by (following the informational strength indexation explained in the next chapter):

Elements of $D^{\Theta=\{\theta_1,\theta_2,\theta_3\}}$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3$	$\alpha_{10} \triangleq \theta_2$
$\alpha_2 \triangleq \theta_1 \cap \theta_2$	$\alpha_{11} \triangleq \theta_3$
$\alpha_3 \triangleq \theta_1 \cap \theta_3$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3$
$\alpha_4 \triangleq \theta_2 \cap \theta_3$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3$
$\alpha_9 \triangleq \theta_1$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3$

Note that the *classical* complementary $\bar{A}$ of any proposition A (except for $\emptyset$ and Θ), is not involved within the free DSm model because of the refutation of the third excluded middle; it can however be introduced if necessary when dealing with hybrid models as it will be shown in chapter 4 if we introduce explicitly some exclusivity constraints into the free DSm model when one has no doubt on the exclusivity between given elements of Θ depending on the nature of the fusion problem. $|D^\Theta|$ for $n \geq 1$ follows the sequence of Dedekind's numbers[1] 1, 2, 5, 19, 167, 7580, 7828353, 56130437228687557907787... [17]. Note also that this huge number of elements of hyper-power set is comparatively far less than the total number of elements of the power set of the refined frame Θ_{ref} if one would to work on $2^{\Theta_{ref}}$ and if we admit the possibility that such refinement exists as it will be seen in section 2.4.1.

2.4 The generation of D^Θ

2.4.1 Memory size requirements and complexity

Before going further on the generation of D^Θ, it is important to estimate the memory size for storing the elements of D^Θ for $|\Theta| = n$. Since each element of D^Θ can be stored as a $2^n - 1$-binary string, the memory size for D^Θ is given by the right column of the following table (we do not count the size for $\emptyset$ which is 0 and the minimum length is considered here as the byte (8 bits)):

[1] Actually this sequence corresponds to the sequence of Dedekind minus one since we don't count the last degenerate isotone function $f_{2^{2^n}-1}(.)$ as element of D^Θ (see section 2.4).

$\|\Theta\| = n$	size/elem.	# of elem.	Size of D^Θ
2	1 byte	4	4 bytes
3	1 byte	18	18 bytes
4	2 bytes	166	0.32 Kb
5	4 bytes	7579	30 Kb
6	8 bytes	7828352	59 Mb
7	16 bytes	$\approx 2.4 \cdot 10^{12}$	$3.6 \cdot 10^4$ Gb
8	32 bytes	$\approx 5.6 \cdot 10^{22}$	$1.7 \cdot 10^{15}$ Gb

This table shows the extreme difficulties for our computers to store all the elements of D^Θ when $|\Theta| > 6$. This complexity remains however smaller than the number of all Boolean functions built from the ultimate refinement (if accessible) $2^{\Theta_{ref}}$ of same initial frame Θ for applying DST. The comparison of $|D^\Theta|$ with respect to $|2^{\Theta_{ref}}|$ is given in the following table

$\|\Theta\| = n$	$\|D^\Theta\|$	$\|2^{\Theta_{ref}}\| = 2^{2^n - 1}$
2	5	$2^3 = 8$
3	19	$2^7 = 128$
4	167	$2^{15} = 32768$
5	7580	$2^{31} = 2147483648$

Fortunately, in most fusion applications only a small subset of elements of D^Θ have a non null basic belief mass because all the commitments are just usually impossible to assess precisely when the dimension of the problem increases. Thus, it is not necessary to generate and keep in memory all elements of D^Θ or $2^{\Theta_{ref}}$ but only those which have a positive belief mass. However there is a real technical challenge on how to manage efficiently all elements of the hyper-power set. This problem is obviously more difficult when working on $2^{\Theta_{ref}}$. Further investigations and research have to be carried out to develop implementable engineering solutions for managing high dimensional problems when the basic belief functions are not degenerated (i.e. all $m(A) > 0$, $A \in D^\Theta$ or $A \in 2^{\Theta_{ref}}$).

2.4.2 Monotone Boolean functions

A simple *Boolean function* $f(.)$ maps n-binary inputs $(x_1, \ldots, x_n) \in \{0,1\}^n \triangleq \{0,1\} \times \ldots \times \{0,1\}$ to a single binary output $y = f(x_1, \ldots, x_n) \in \{0,1\}$. Since there are 2^n possible input states which can map to either 0 or 1 at the output y, the number of possible Boolean functions is 2^{2^n}. Each of these functions can be realized by the logic operations $\wedge$ (and), $\vee$ (or) and $\neg$ (not) [3, 21]. As a simple example, let's consider only a 2-binary input variable $(x_1, x_2) \in \{0,1\} \times \{0,1\}$ then all the $2^{2^2} = 16$ possible Boolean functions $f_i(x_1, x_2)$ built from (x_1, x_2) are summarized in the following tables:

(x_1,x_2)	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7
$(0,0)$	0	0	0	0	0	0	0	0
$(0,1)$	0	0	0	0	1	1	1	1
$(1,0)$	0	0	1	1	0	0	1	1
$(1,1)$	0	1	0	1	0	1	0	1
Notation	False	$x_1 \wedge x_2$	$x_1 \wedge \bar{x}_2$	x_1	$\bar{x}_1 \wedge x_2$	x_2	$x_1 \veebar x_2$	$x_1 \vee x_2$

(x_1,x_2)	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}
$(0,0)$	1	1	1	1	1	1	1	1
$(0,1)$	0	0	0	0	1	1	1	1
$(1,0)$	0	0	1	1	0	0	1	1
$(1,1)$	0	1	0	1	0	1	0	1
Notation	$x_1 \bar{\vee} x_2$	$x_1 \triangle x_2$	$\bar{x}_2$	$x_1 \vee \bar{x}_2$	$\bar{x}_1$	$\bar{x}_1 \vee x_2$	$x_1 \bar{\wedge} x_2$	True

with the notation $\bar{x} \triangleq \neg x$, $x_1 \veebar x_2 \triangleq (x_1 \vee x_2) \wedge (\bar{x}_1 \vee \bar{x}_2)$ (xor), $x_1 \bar{\vee} x_2 \triangleq \neg(x_1 \vee x_2)$ (nor), $x_1 \triangle x_2 \triangleq (x_1 \wedge x_2) \vee (\bar{x}_1 \wedge \bar{x}_2)$ (xnor) and $x_1 \bar{\wedge} x_2 \triangleq \neg(x_1 \wedge x_2)$ (nand).

We denote by $\mathcal{F}_n(\wedge,\vee,\neg) = \{f_0(x_1,\ldots,x_n),\ldots,f_{2^{2^n}-1}(x_1,\ldots,x_n)\}$ the set of all possible Boolean functions built from n-binary inputs. Let $\mathbf{x} \triangleq (x_1,\ldots,x_n)$ and $\mathbf{x}' \triangleq (x'_1,\ldots,x'_n)$ be two vectors in $\{0,1\}^n$. Then $\mathbf{x}$ precedes $\mathbf{x}'$ and we denote $\mathbf{x} \preceq \mathbf{x}'$ if and only if $x_i \leq x'_i$ for $1 \leq i \leq n$ ($\leq$ is applied componentwise). If $x_i < x'_i$ for $1 \leq i \leq n$ then $\mathbf{x}$ strictly precedes $\mathbf{x}'$ which will be denoted as $\mathbf{x} \prec \mathbf{x}'$.

A Boolean function f is said to be a *non-decreasing monotone (or isotone) Boolean function* (or just *monotone Boolean function* for short) if and only if $\forall \mathbf{x},\mathbf{x}' \in \{0,1\}^n$ such that $\mathbf{x} \preceq \mathbf{x}'$, then $f(\mathbf{x}) \preceq f(\mathbf{x}')$ [19]. Since any isotone Boolean function involves only $\wedge$ and $\vee$ operators (no $\neg$ operations) [21] and there exists a parallel between $(\vee,\wedge)$ operators in logics with $(+,\cdot)$ in algebra of numbers and $(\cup,\cap)$ in algebra of sets, the generation of all elements of D^Θ built from Θ with $\cup$ and $\cap$ operator is equivalent to the problem of generating isotone Boolean functions over the vertices of the unit n-cube. We denote by $\mathcal{M}_n(\wedge,\vee)$ the set of all possible monotone Boolean functions built from n-binary inputs. $\mathcal{M}_n(\wedge,\vee)$ is a subset of $\mathcal{F}_n(\wedge,\vee,\neg)$. In the previous example, $f_1(x_1,x_2)$, $f_3(x_1,x_2)$, $f_5(x_1,x_2)$, $f_7(x_1,x_2)$ are isotone Boolean functions but special functions $f_0(x_1,x_2)$ and $f_{2^{2^n}-1}(x_1,\ldots,x_n)$ must also be considered as monotone functions too. All the other functions belonging to $\mathcal{F}_2(\wedge,\vee,\neg)$ do not belong to $\mathcal{M}_2(\wedge,\vee)$ because they require the $\neg$ operator in their expressions and we can check easily that the monotonicity property $\mathbf{x} \preceq \mathbf{x}' \Rightarrow f(\mathbf{x}) \preceq f(\mathbf{x}')$ does not hold for these functions.

The Dedekind's problem [4] is to determine the number $d(n)$ of *distinct* monotone Boolean functions of n-binary variables. Dedekind [4] computed $d(0) = 2$, $d(1) = 3$, $d(2) = 6$, $d(3) = 20$ and $d(4) = 168$. Church [1] computed $d(5) = 7581$ in 1940. Ward [20] computed $d(6) = 7828354$ in 1946. Church [2] then computed $d(7) = 2414682040998$ in 1965. Between sixties and eighties, important advances have been done to obtain upper and lower bounds for $d(n)$ [10, 12, 14]. In 1991, Wiedemann [22] computed $d(8) = 56130437228687557907788$ (200 hours of computing time with a Cray-2 processor) which has recently been validated by Fidytek and al. in [8]. Until now the computation of $d(n)$ for $n > 8$ is still a challenge for mathematicians even if the following direct exact explicit formula for $d(n)$ has been obtained by Kisielewicz and Tombak (see [11, 18] for proof) :

$$d(n) = \sum_{k=1}^{2^{2^n}} \prod_{j=1}^{2^n-1} \prod_{i=0}^{j-1} (1 - b_i^k(1 - b_j^k) \prod_{m=0}^{l(i)} (1 - b_m^i(1 - b_m^j))) \tag{2.1}$$

where $l(0) = 0$ and $l(i) = [\log_2 i]$ for $i > 0$, $b_i^k \triangleq [k/2^i] - 2[k/2^{i+1}]$ and $[x]$ denotes the floor function (i.e. the nearest integer less or equal to x). The difficulty arises from the huge number of terms involved in the formula, the memory size and the high speed computation requirements. The last advances and state of art in counting algorithms of Dedekind's numbers can be found in [18, 8, 19].

2.4.3 Generation of MBF

Before describing the general algorithm for generating the monotone Boolean functions (MBF), let examine deeper the example of section 2.4.2. From the previous tables, one can easily find the set of (restricted) MBF $\mathcal{M}_2^\star(\wedge, \vee) = \{f_0(x_1, x_2) = \text{False}, f_1(x_1, x_2) = x_1 \wedge x_2, f_5(x_1, x_2) = x_2, f_7(x_1, x_2) = x_1 \vee x_2\}$ which is equivalent, using algebra of sets, to hyper-power set $D^X = \{\emptyset, x_1 \cap x_2, x_1, x_2, x_1 \cup x_2\}$ associated with frame of discernment $X = \{x_1, x_2\}$. Since the tautology $f_{15}(x_1, x_2)$ is not involved within DSmT, we do not include it as a proper element of D^X and we consider only $\mathcal{M}_2^\star(\wedge, \vee) \triangleq \mathcal{M}_2(\wedge, \vee) \setminus \{f_{15}\}$ rather than $\mathcal{M}_2(\wedge, \vee)$ itself.

Let's now introduce Smarandache's codification for the enumeration of distinct parts of a Venn diagram X with n partially overlapping elements x_i,$i = 1, 2, \ldots, n$. Such a diagram has $2^n - 1$ disjoint parts. One denotes with only one digit (or symbol) those parts which belong to only one of the elements x_i (one denotes by $< i >$ the part which belongs to x_i only, for $1 \leq i \leq n$), with only two digits (or symbols) those parts which belong to exactly two elements (one denotes by $< ij >$, with $i < j$, the part which belongs to x_i and x_j only, for $1 \leq i < j \leq n$), then with only three digits (or symbols) those parts which belong to exactly three elements (one denotes by $< ijk >$ concatenated numbers, with $i < j < k$, the part which belongs to x_i, x_j, and x_k only, for $1 \leq i < j < k \leq n$), and so on up to $< 12 \ldots n >$ which represents the last part that belongs to all elements x_i. For $1 \leq n \leq 9$, Smarandache's encoding works normally as in base 10. But, for $n \geq 10$, because there occur two (or more) digits/symbols in notation of the elements starting from 10 on, one considers this codification in base $n+1$, i.e. using one symbol to represent two (or more) digits, for example: $A = 10$, $B = 11$, $C = 12$, etc.

- For $n = 1$ one has only one part, coded $< 1 >$.
- For $n = 2$ one has three parts, coded $< 1 >$, $< 2 >$, $< 12 >$. Generally, $< ijk >$ does not represent $x_i \cap x_j \cap x_k$ but only a part of it, the only exception is for $< 12 \ldots n >$.
- For $n = 3$ one has $2^3 - 1 = 7$ disjoint parts, coded $< 1 >$, $< 2 >$, $< 3 >$, $< 12 >$, $< 13 >$, $< 23 >$, $< 123 >$. $< 23 >$ means the part which belongs to x_2 and x_3 only, but $< 23 > \neq x_2 \cap x_3$ because $x_2 \cap x_3 = \{< 23 >, < 123 >\}$ in the Venn diagram of 3 elements x_1, x_2, and x_3 (see next chapter).
- The generalization for $n > 3$ is straightforward. Smarandache's codification can be organized in a numerical increasing order, in lexicographic order or any other orders.

A useful order for organizing Smarandache's codification for the generation of D^X is the *DSm-order* $\mathbf{u}_n = [u_1, \ldots, u_{2^n-1}]'$ based on a recursive construction starting with $\mathbf{u}_1 \triangleq [< 1 >]$. Having constructed $\mathbf{u}_{n-1}$, then we can construct $\mathbf{u}_n$ for $n > 1$ recursively as follows:

- include all elements of $\mathbf{u}_{n-1}$ into $\mathbf{u}_n$;
- afterwards, include element $< n >$ as well in $\mathbf{u}_n$;
- then at the end of each element of $\mathbf{u}_{n-1}$ concatenate the element $< n >$ and get a new set $\mathbf{u}'_{n-1}$ which then is also included in $\mathbf{u}_n$.

This is $\mathbf{u}_n$, which has $(2^{n-1} - 1) + 1 + (2^{n-1} - 1) = 2^n - 1$ components.

For $n = 3$, as example, one gets $\mathbf{u}_3 \triangleq [< 1 > < 2 > < 12 > < 3 > < 13 > < 23 > < 123 >]'$. Because all elements in $\mathbf{u}_n$ are disjoint, we are able to write each element d_i of D^X in a unique way as a linear combination of $\mathbf{u}_n$ elements, i.e.

$$\mathbf{d}_n = [d_1, \ldots, d_{2^n-1}]' = \mathbf{D}_n \cdot \mathbf{u}_n \tag{2.2}$$

Thus $\mathbf{u}_n$ constitutes a basis for generating the elements of D^X. Each row in the matrix $\mathbf{D}_n$ represents the coefficients of an element of D^X with respect to the basis $\mathbf{u}_n$. The rows of $\mathbf{D}_n$ may also be regarded as binary numbers in an increasing order.

Example: For $n = 2$, one has:

$$\underbrace{\begin{bmatrix} d_1 = x_1 \cap x_2 \\ d_2 = x_2 \\ d_3 = x_1 \\ d_4 = x_1 \cup x_2 \end{bmatrix}}_{\mathbf{d}_2} = \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{\mathbf{D}_2} \cdot \underbrace{\begin{bmatrix} <1> \\ <2> \\ <12> \end{bmatrix}}_{\mathbf{u}_2} \tag{2.3}$$

where the "matrix product" is done after identifying $(+, \cdot)$ with $(\cup, \cap)$, $0\cdot <x>$ with $\emptyset$ and $1\cdot <x>$ with $<x>$.

The generation of D^X is then strictly equivalent to generate $\mathbf{u}_n$ and matrix $\mathbf{D}_n$ which can be easily obtained by the following recursive procedure:

- start with $\mathbf{D}_0^c = [0\,1]'$ corresponding to all Boolean functions with no input variable ($n = 0$).
- build the $\mathbf{D}_1^c$ matrix from each row $\mathbf{r}_i$ of $\mathbf{D}_0^c$ by adjoining it to any other row $\mathbf{r}_j$ of $\mathbf{D}_0^c$ such that $\mathbf{r}_i \cup \mathbf{r}_j = \mathbf{r}_j$. This is equivalent here to add either 0 or 1 in front (i.e. left side) of $\mathbf{r}_1 \equiv 0$ but only 1 in front of $\mathbf{r}_2 \equiv 1$. Since the tautology is not involved in the hyper-power set, then one has to remove the first column and the last line of

$$\mathbf{D}_1^c = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{to obtain finally} \quad \mathbf{D}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- build $\mathbf{D}_2^c$ from $\mathbf{D}_1^c$ by adjoining to each row $\mathbf{r}_i$ of $\mathbf{D}_1^c$, any row $\mathbf{r}_j$ of $\mathbf{D}_1^c$ such that $\mathbf{r}_i \cup \mathbf{r}_j = \mathbf{r}_j$ and then remove the first column and the last line of $\mathbf{D}_2^c$ to get $\mathbf{D}_2$ as in (2.3).
- build $\mathbf{D}_3^c$ from $\mathbf{D}_2^c$ by adjoining to each row $\mathbf{r}_i$ of $\mathbf{D}_2^c$ any row $\mathbf{r}_j$ of $\mathbf{D}_2^c$ such that $\mathbf{r}_i \cup \mathbf{r}_j = \mathbf{r}_j$ and then remove the first column and the last line of $\mathbf{D}_3^c$ to get $\mathbf{D}_3$ given by (where $\mathbf{D}'$ denotes here the transposed of the matrix $\mathbf{D}$)

$$\mathbf{D}_3' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

- Likewise, $\mathbf{D}_n^c$ is built from $\mathbf{D}_{n-1}^c$ by adjoining to each row $\mathbf{r}_i$ of $\mathbf{D}_{n-1}^c$ any row $\mathbf{r}_j$ of $\mathbf{D}_{n-1}^c$ such that $\mathbf{r}_i \cup \mathbf{r}_j = \mathbf{r}_j$. Then $\mathbf{D}_n$ is obtained by removing the first column and the last line of $\mathbf{D}_n^c$.

Example for $\Theta = \{\theta_1, \theta_2, \theta_3\}$: Note that the new indexation of elements of D^Θ now follows the MBF generation algorithm.

$$
\underbrace{\begin{bmatrix}
\alpha_0 \triangleq \emptyset \\
\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \\
\alpha_2 \triangleq \theta_2 \cap \theta_3 \\
\alpha_3 \triangleq \theta_1 \cap \theta_3 \\
\alpha_4 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \\
\alpha_5 \triangleq \theta_3 \\
\alpha_6 \triangleq \theta_1 \cap \theta_2 \\
\alpha_7 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \\
\alpha_8 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \\
\alpha_9 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \\
\alpha_{10} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \\
\alpha_{11} \triangleq \theta_2 \\
\alpha_{12} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \\
\alpha_{13} \triangleq (\theta_2 \cup \theta_3) \\
\alpha_{14} \triangleq \theta_1 \\
\alpha_{15} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \\
\alpha_{16} \triangleq (\theta_1 \cup \theta_3) \\
\alpha_{17} \triangleq (\theta_1 \cup \theta_2) \\
\alpha_{18} \triangleq (\theta_1 \cup \theta_2 \cup \theta_3)
\end{bmatrix}}_{\mathbf{d}_3}
=
\underbrace{\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}}_{\mathbf{D}_3}
\cdot
\underbrace{\begin{bmatrix}
<1> \\
<2> \\
<12> \\
<3> \\
<13> \\
<23> \\
<123>
\end{bmatrix}}_{\mathbf{u}_3}
$$

For convenience, we provide in appendix the source code in Matlab[2] language to generate D^Θ. This code includes the identification of elements of D^Θ corresponding to each monotone Boolean function according to Smarandache's codification.

2.5 Conclusion

In this chapter, one has introduced the notion of Dedekind's lattice D^Θ (hyper-power set) on which are defined basic belief functions in the framework of DSmT and the acceptance of the free DSm model. The justification of the free DSm model as a starting point (ground level) for the development of our new theory of plausible and paradoxical reasoning for information fusion has been also given and arises from the necessity to deal with possibly ambiguous concepts which can appear in real-world applications. The lower complexity of the hyper-power set with respect to the complexity of the classical refined power set

[2]Matlab is a trademark of The MathWorks, Inc.

$2^{\Theta_{ref}}$ has been clearly demonstrated here. We have proven the theoretical link between the generation of hyper-power set with Dedekind's problem on counting isotone Boolean functions. A theoretical solution for generating lattices D^Θ has been presented and a MatLab source code has been provided for users convenience.

2.6 References

[1] Church R., *Numerical analysis of certain free distributive structures*, Duke Math. J., Vol. 6, no. 3, pp. 732-734, 1940.

[2] Church R., *Enumeration by rank of the elements of the free distributive lattice with seven generators*, Not. Amer. Math. Soc., Vol. 12, p. 724, 1965.

[3] Comtet L., *Sperner Systems*, sec.7.2 in Advanced Combinatorics: The Art of Finite and Infinite Expansions, D. Reidel Publ. Co., pp. 271-273, 1974.

[4] Dedekind R. *Über Zerlegungen von Zahlen durch ihre grössten gemeinsammen Teiler*, In Gesammelte Werke, Bd. 1. pp. 103-148, 1897.

[5] Dezert J., *Fondations pour une nouvelle théorie du raisonnement plausible et paradoxal*, ONERA Tech. Rep. RT 1/06769/DTIM, Jan. 2003.

[6] Dezert J., Smarandache F., *On the generation of hyper-power sets*, Proc. of Fusion 2003, Cairns, Australia, July 8-11, 2003.

[7] Dubois D., Prade H., *Representation and combination of uncertainy with belief functions and possibility measures*, Computational Intelligence, 4, pp. 244-264, 1988.

[8] Fidytek R., Mostowski A.W., Somla R., Szepietowski A., *Algorithms counting monotone Boolean functions*, Inform. Proc. Letters, Vol. 79, pp. 203-209, 2001.

[9] Goodman I.R., Mahler R.P.S., Nguyen H.T., *Mathematics of Data Fusion*, Kluwer Academic Press, Boston, 1997.

[10] Hansel G., *Sur le nombre des fonctions booléennes monotones de n variables*, C.R. Acad. Sci. Paris, Serie A-B, p. 262, 1966.

[11] Kisielewicz A., *A solution of Dedekind's problem on the number of isotone Boolean functions*, J. reine angew. math., Vol. 386, pp. 139-144, 1988.

[12] Kleitman D., *On Dedekind's problem: The number of Boolean functions*, Proc. of the Amer. Math Society, Vol. 21, pp. 677-682, 1969.

[13] Kleitman D., Markowsky G., *On Dedekind's problem: The number of isotone Boolean functions. II*, Trans. of the Amer. Math. Soc., Vol. 213, 1976.

[14] Korshunov A., *On the number of Boolean functions*, Prob. kibernetiki, Vol. 38, pp. 5-108, 1981.

[15] Pawlak Z., *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dortrecht, 1991.

[16] Shapiro H.N. *On the counting problem for monotone Boolean functions*, Comm. on pure and applied math., Vol. XXIII, pp. 299-312, 1970.

[17] Sloane N.J.A., *The On-line Encyclopedia of Integer Sequences 2003*, (Sequence No. A014466), `http://www.research.att.com/~njas/sequences/`.

[18] Tombak M., Isotamm A., Tamme T., *On logical method for counting Dedekind numbers*, Lect. Notes on Comp.Sci., 2138, p. 424-427, Springer-Verlag, 2001. `www.cs.ut.ee/people/m_tombak/publ.html`.

[19] Triantaphyllou, E., Torvik V.I., *Inference of Monotone Boolean Functions*, Encyclopedia of Optimization, (P.M. Pardalos and C. Floudas, Eds.), Kluwer Acad. Publi., Boston, Vol. 2, pp. 472-480., 2001.

[20] Ward M., *Note on the order of free distributive lattices* , Bull. Amer. Math. Soc., Vol. 52, no. 5, p. 423, 1946.

[21] Weisstein E. W., *CRC Concise Encyclopedia of Mathematics* , CRC Press; 2nd edition, 3252 pages, Dec. 2002. `http://mathworld.wolfram.com/`.

[22] Wiedemann D., *A computation of the eighth Dedekind number*, Order, no. 8, pp. 5-6, 1991.

[23] Yager R.R., *On the Dempster-Shafer Framework and New Combination Rules*, Information Sciences, Vol. 41, pp.93-137, 1987.

Appendix: MatLab code for generating hyper-power sets

```
%**********************************************
% Copyright (c)  2003 J.Dezert and F.Smarandache
%
% Purpose: Generation of D^Theta for the DSmT for
% Theta={theta_1,..,Theta_n}. Due to the huge
% # of elements of D^Theta. only cases up to n<7
% are usually tractable on computers.
%**********************************************
n=input('Enter cardinality for Theta (0<n<6) ? ');
% Generation of the Smarandache codification
% Note: this should be implemented using
% character strings for n>9
u_n=[1];
for nn=2:n
u_n=[u_n nn (u_n*10+nn*ones(1,size(u_n*10,2)))];
end
% Generation of D_n (isotone boolean functions)
D_n1=[0;1];
for nn=1:n, D_n=[];
    for i=1:size(D_n1,1),Li=D_n1(i,:);
        for j=i:size(D_n1,1)
         Lj=D_n1(j,:);Li_inter_Lj=and(Li,Lj);
         Li_union_Lj=or(Li,Lj);
         if((Li_inter_Lj==Li)&(Li_union_Lj==Lj))
                D_n=[D_n; Li Lj];
         end
        end
    end
    D_n1=D_n;
end
DD=D_n;DD(:,1)=[];DD(size(DD,1),:)=[]; D_n=DD;
% Result display
disp(['|Theta|=n=',num2str(n)])
disp(['|D^Theta|=',num2str(size(D_n,1))])
disp('Elem. of D^Theta are obtained by D_n*u_n')
disp(['with u_n=[',num2str(u_n),'] '' and'])
D_n=D_n
```

Matlab source code for generating D^{Θ}

Chapter 3

Partial ordering on hyper-power sets

Jean Dezert	Florentin Smarandache
ONERA	Department of Mathematics
29 Av. de la Division Leclerc	University of New Mexico
92320 Châtillon	Gallup, NM 8730
France	U.S.A.

Abstract: *In this chapter, we examine several issues for ordering or partially ordering elements of hyper-power sets involved in the DSmT. We will show the benefit of some of these issues to obtain a nice and interesting structure of matrix representation of belief functions.*

3.1 Introduction to matrix calculus for belief functions

As rightly emphasized recently by Smets in [9], the mathematic of belief functions is often cumbersome because of the many summations symbols and all its subscripts involved in equations. This renders equations very difficult to read and understand at first sight and might discourage potential readers for their complexity. Actually, this is just an appearance because most of the operations encountered in DST with belief functions and basic belief assignments $m(.)$ are just simple linear operations and can be easily represented using matrix notation and be handled by elementary matrix calculus. We just focus here our presentation on the matrix representation of the relationship between a basic belief assignment $m(.)$ and its associated belief function Bel(.). A nice and more complete presentation of matrix calculus for belief functions can be found in [6, 7, 9]. One important aspect for the simplification of matrix representation and calculus in DST, concerns the choice of the order of the elements of the

This chapter is based on a paper [4] presented during the International Conference on Information Fusion, Fusion 2003, Cairns, Australia, in July 2003 and is reproduced here with permission of the International Society of Information Fusion.

power set 2^Θ. The order of elements of 2^Θ can be chosen arbitrarily actually, and it can be easily seen by denoting $\mathbf{m}$ the bba vector of size $2^n \times 1$ and $\mathbf{Bel}$ its corresponding belief vector of same size, that the set of equations $\mathrm{Bel}(A) = \sum_{B\subseteq A} m(B)$ holding for all $A \subseteq \Theta$ is strictly equivalent to the following general matrix equation

$$\mathbf{Bel} = \mathbf{BM} \cdot \mathbf{m} \quad \Leftrightarrow \quad \mathbf{m} = \mathbf{BM}^{-1} \cdot \mathbf{Bel} \tag{3.1}$$

where the internal structure of $\mathbf{BM}$ depends on the choice of the order for enumerating the elements of 2^Θ. But it turns out that the simplest ordering based on the enumeration of integers from 0 to $2^n - 1$ expressed as n-binary strings with the lower bit on the right (LBR) (where $n = |\Theta|$) to characterize all the elements of power set, is the most efficient solution and best encoding method for matrix calculus and for developing efficient algorithms in MatLab[1] or similar programming languages [9]. By choosing the basic increasing binary enumeration (called *bibe system*), one obtains a very nice recursive algorithm on the dimension n of Θ for computing the matrix $\mathbf{BM}$. The computation of $\mathbf{BM}$ for $|\Theta| = n$ is just obtained from the iterations up to $i + 1 = n$ of the recursive relation [9] starting with $\mathbf{BM}_0 \triangleq [1]$ and where $\mathbf{0}_{i+1}$ denotes the zero-matrix of size $(i + 1) \times (i + 1)$,

$$\mathbf{BM}_{i+1} = \begin{bmatrix} \mathbf{BM}_i & \mathbf{0}_{i+1} \\ \mathbf{BM}_i & \mathbf{BM}_i \end{bmatrix} \tag{3.2}$$

$\mathbf{BM}$ is a binary unimodular matrix ($\det(\mathbf{BM}) = \pm 1$). $\mathbf{BM}$ is moreover triangular inferior and symmetrical with respect to its antidiagonal.

Example for $\Theta = \{\theta_1, \theta_2, \theta_3\}$

The *bibe system* gives us the following order for elements of $2^\Theta = \{\alpha_0, \ldots, \alpha_7\}$:

$$\begin{array}{llll}
\alpha_0 \equiv 000 \equiv \emptyset & \alpha_1 \equiv 001 \equiv \theta_1 & \alpha_2 \equiv 010 \equiv \theta_2 & \alpha_3 \equiv 011 \equiv \theta_1 \cup \theta_2 \\
\alpha_4 \equiv 100 \equiv \theta_3 & \alpha_5 \equiv 101 \equiv \theta_1 \cup \theta_3 & \alpha_6 \equiv 110 \equiv \theta_2 \cup \theta_3 & \alpha_7 \equiv 111 \equiv \theta_1 \cup \theta_2 \cup \theta_3 \equiv \Theta
\end{array}$$

Each element α_i of 2^Θ is a 3-bits string. With this bibe system, one has $\mathbf{m} = [m(\alpha_0), \ldots, m(\alpha_7)]'$ and $\mathbf{Bel} = [\mathrm{Bel}(\alpha_0), \ldots, \mathrm{Bel}(\alpha_7)]'$. The expressions of the matrix $\mathbf{BM}_3$ and its inverse $\mathbf{BM}_3{}^{-1}$ are given by

$$\mathbf{BM}_3 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}$$

[1] Matlab is a trademark of The MathWorks, Inc.

$$\mathbf{BM_3}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix}$$

3.2 Ordering elements of hyper-power set for matrix calculus

As within the DST framework, the order of the elements of D^Θ can be arbitrarily chosen. We denote the Dedekind number or order n as $d(n) \triangleq |D^\Theta|$ for $n = |\Theta|$. We denote also $\mathbf{m}$ the gbba vector of size $d(n) \times 1$ and **Bel** its corresponding belief vector of the same size. The set of equations $\text{Bel}(A) = \sum_{B \in D^\Theta, B \subseteq A} m(B)$ holding for all $A \in D^\Theta$ is then strictly equivalent to the following general matrix equation

$$\mathbf{Bel} = \mathbf{BM} \cdot \mathbf{m} \quad \Leftrightarrow \quad \mathbf{m} = \mathbf{BM}^{-1} \cdot \mathbf{Bel} \tag{3.3}$$

Note the similarity between these relations with the previous ones (3.1). The only difference resides in the size of vectors **Bel** and **m** and the size of matrix **BM** and their components. We explore in the following sections the possible choices for ordering (or partially ordering) the elements of hyper-power set D^Θ, to obtain an interesting matrix structure of **BM** matrix. Only three issues are examined and briefly presented in the sequel. The first method is based on the direct enumeration of elements of D^Θ according to their recursive generation via the algorithm for generating all isotone Boolean functions presented in the previous chapter and in [3]. The second (partial) ordering method is based on the notion of DSm cardinality which will be introduced in section 3.2.2. The last and most interesting solution proposed for partial ordering over D^Θ is obtained by introducing the notion of intrinsic informational strength $s(.)$ associated to each element of hyper-power set.

3.2.1 Order based on the enumeration of isotone Boolean functions

We have presented in chapter 2 a recursive algorithm based on isotone Boolean functions for generating D^Θ with didactic examples. Here is briefly the principle of the method. Let's consider $\Theta = \{\theta_1, \ldots, \theta_n\}$ satisfying the DSm model and the DSm order $\mathbf{u}_n$ of Smarandache's codification of parts of Venn diagram Θ with n partially overlapping elements θ_i, $i = 1, \ldots, n$. All the elements α_i of D^Θ can then be obtained by the very simple linear equation $\mathbf{d}_n = \mathbf{D}_n \cdot \mathbf{u}_n$ where $\mathbf{d}_n \equiv [\alpha_0 \equiv \emptyset, \alpha_1, \ldots, \alpha_{d(n)-1}]'$ is the vector of elements of D^Θ, $\mathbf{u}_n$ is the proper codification vector and D_n a particular binary matrix. The final result $\mathbf{d}_n$ is obtained from the previous *matrix product* after identifying $(+, \cdot)$ with $(\cup, \cap)$ operators, $0 \cdot x$ with $\emptyset$

and $1 \cdot x$ with x. D_n is actually a binary matrix corresponding to isotone (i.e. non-decreasing) Boolean functions obtained by applying recursively the steps (starting with $\mathbf{D}_0^c = [0\,1]'$)

- $\mathbf{D}_n^c$ is built from $\mathbf{D}_{n-1}^c$ by adjoining to each row $\mathbf{r}_i$ of $\mathbf{D}_{n-1}^c$ any row $\mathbf{r}_j$ of $\mathbf{D}_{n-1}^c$ such that $\mathbf{r}_i \cup \mathbf{r}_j = \mathbf{r}_j$. Then $\mathbf{D}_n$ is obtained by removing the first column and the last line of $\mathbf{D}_n^c$.

We denote $r^{iso}(\alpha_i)$ the position of α_i into the column vector $\mathbf{d}_n$ obtained from the previous enumeration/generation system. Such a system provides a total order over D^Θ defined $\forall \alpha_i, \alpha_j \in D^\Theta$ as $\alpha_i \prec \alpha_j$ (α_i precedes α_j) if and only if $r^{iso}(\alpha_i) < r^{iso}(\alpha_j)$. Based on this order, the **BM** matrix involved in (3.3) presents unfortunately no particular interesting structure. We have thus to look for better solutions for ordering the elements of hyper-power sets.

3.2.2 Ordering based on the DSm cardinality

A second possibility for ordering the elements of D^Θ is to (partially) order them by their increasing *DSm cardinality*.

Definition of the DSm cardinality

The *DSm cardinality* of any element $A \in D^\Theta$, denoted $\mathcal{C}_\mathcal{M}(A)$, corresponds to the number of parts of A in the Venn diagram of the problem (model $\mathcal{M}$) taking into account the set of integrity constraints (if any), i.e. all the possible intersections due to the nature of the elements θ_i. This *intrinsic cardinality* depends on the model $\mathcal{M}$. $\mathcal{M}$ is the model that contains A which depends on the dimension of Venn diagram, (i.e. the number of sets $n = |\Theta|$ under consideration), and on the number of non-empty intersections in this diagram. $\mathcal{C}_\mathcal{M}(A)$ must not be confused with the classical cardinality $|A|$ of a given set A (i.e. the number of its distinct elements) - that's why a new notation is necessary here.

Some properties of the DSm cardinality

First note that one has always $1 \leq \mathcal{C}_\mathcal{M}(A) \leq 2^n - 1$. In the (general) case of the free-model $\mathcal{M}^f$ (i.e. the DSm model) where all conjunctions are non-empty, one has for intersections:

$\mathcal{C}_{\mathcal{M}^f}(\theta_1) = \ldots = \mathcal{C}_{\mathcal{M}^f}(\theta_n) = 2^{n-1}$

$\mathcal{C}_{\mathcal{M}^f}(\theta_i \cap \theta_j) = 2^{n-2}$ for $n \geq 2$

$\mathcal{C}_{\mathcal{M}^f}(\theta_i \cap \theta_j \cap \theta_k) = 2^{n-3}$ for $n \geq 3$

It can be proven by induction that for $1 \leq m \leq n$, one has $\mathcal{C}_{\mathcal{M}^f}(\theta_{i_1} \cap \theta_{i_2} \cap \ldots \cap \theta_{i_m}) = 2^{n-m}$. For the cases $n = 1, 2, 3, 4$, this formula can be checked on the corresponding Venn diagrams. Let's consider this formula true for n sets, and prove it for $n+1$ sets (when all intersections/conjunctions are considered non-empty). From the Venn diagram of n sets, we can get a Venn diagram with $n+1$ sets if one draws a closed curve that cuts each of the $2^n - 1$ parts of the previous diagram (and, as a consequence, divides each part into two disjoint subparts). Therefore, the number of parts of each intersection is doubling when passing from a diagram of dimension n to a diagram of dimension $n+1$. Q.e.d.

In the case of the free-model $\mathcal{M}^f$, one has for unions:

$\mathcal{C}_{\mathcal{M}^f}(\theta_i \cup \theta_j) = 3(2^{n-2})$ for $n \geq 2$

$\mathcal{C}_{\mathcal{M}^f}(\theta_i \cup \theta_j \cup \theta_k) = 7(2^{n-3})$ for $n \geq 3$

It can be proven also by induction that for $1 \leq m \leq n$, one has $\mathcal{C}_{\mathcal{M}^f}(\theta_{i_1} \cup \theta_{i_2} \cup \ldots \cup \theta_{i_m}) = (2^m - 1)(2^{n-m})$. The proof is similar to the previous one, and keeping in mind that passing from a Venn diagram of dimension n to a dimension $n+1$, each part that forms the union $\theta_i \cap \theta_j \cap \theta_k$ will be split into two disjoint parts, hence the number of parts is doubling.

For other elements A in D^Θ, formed by unions and intersections, the closed form for $\mathcal{C}_{\mathcal{M}^f}(A)$ seems more complicated to obtain. But from the generation algorithm of D^Θ, DSm cardinal of a set A from D^Θ is exactly equal to the sum of its coefficients in the $\mathbf{u}_n$ basis, i.e. the sum of its row elements in the $\mathbf{D}_n$ matrix, which is actually very easy to compute by programming. The *DSm cardinality* plays in important role in the definition of the Generalized Pignistic Transformation (GPT) for the construction of subjective/pignistic probabilities of elements of D^Θ for decision-making at the pignistic level as explained in chapter 7 and in [5]. If one imposes a constraint that a set B from D^Θ is empty, then one suppresses the columns corresponding to the parts which compose B in the $\mathbf{D}_n$ matrix and the row of B and the rows of all elements of D^Θ which are subsets of B, getting a new matrix $\mathbf{D'}_n$ which represents a new model $\mathcal{M}'$. In the $\mathbf{u}_n$ basis, one similarly suppresses the parts that form B, and now this basis has the dimension $2^n - 1 - \mathcal{C}_{\mathcal{M}}(B)$.

Example of DSm cardinals on $\mathcal{M}^f$

Consider the 3D case $\Theta = \{\theta_1, \theta_2, \theta_3\}$ with the free-model $\mathcal{M}^f$ corresponding to the following Venn diagram (where $< i >$ denotes the part which belongs to θ_i only, $< ij >$ denotes the part which belongs to θ_i and θ_j only, etc; this is Smarandache's codification (see the previous chapter).

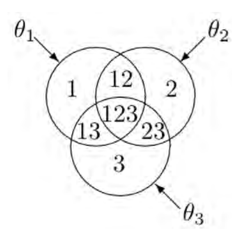


Figure 3.1: Venn Diagram for $\mathcal{M}^f$

The corresponding *partial ordering* for elements of D^Θ is then summarized in the following table:

$A \in D^\Theta$	$\mathcal{C}_{\mathcal{M}^f}(A)$
$\alpha_0 \triangleq \emptyset$	0
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3$	1
$\alpha_2 \triangleq \theta_1 \cap \theta_2$	2
$\alpha_3 \triangleq \theta_1 \cap \theta_3$	2
$\alpha_4 \triangleq \theta_2 \cap \theta_3$	2
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3$	3
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2$	3
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1$	3
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	4
$\alpha_9 \triangleq \theta_1$	4
$\alpha_{10} \triangleq \theta_2$	4
$\alpha_{11} \triangleq \theta_3$	4
$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3$	5
$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2$	5
$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1$	5
$\alpha_{15} \triangleq \theta_1 \cup \theta_2$	6
$\alpha_{16} \triangleq \theta_1 \cup \theta_3$	6
$\alpha_{17} \triangleq \theta_2 \cup \theta_3$	6
$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3$	7

Table 3.1: $\mathcal{C}_{\mathcal{M}^f}(A)$ for free DSm model $\mathcal{M}^f$

Note that this partial ordering doesn't properly catch the intrinsic informational structure/strength of elements since by example $(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$ and θ_1 have the same DSm cardinal although they don't look similar because the part $< 1 >$ in θ_1 belongs only to θ_1 but none of the parts of $(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$ belongs to only one part of some θ_i. A better ordering function is then necessary to catch the intrinsic informational structure of elements of D^Θ. This is the purpose of the next section.

Example of DSm cardinals on an hybrid DSm model $\mathcal{M}$

Consider now the same 3D case with the hybrid DSm model $\mathcal{M} \neq \mathcal{M}^f$ in which we force all possible conjunctions to be empty, but $\theta_1 \cap \theta_2$ according to the following Venn diagram.

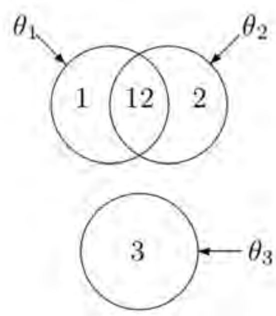


Figure 3.2: Venn Diagram for $\mathcal{M}$

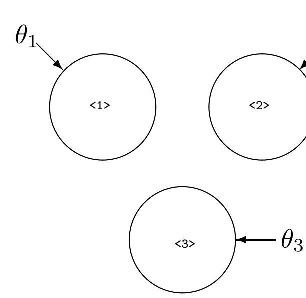
The corresponding *partial ordering* for elements of D^Θ is then summarized in the following table:

$A \in D^\Theta$	$\mathcal{C}_\mathcal{M}(A)$
$\alpha_0 \triangleq \emptyset$	0
$\alpha_1 \triangleq \theta_1 \cap \theta_2$	1
$\alpha_2 \triangleq \theta_3$	1
$\alpha_3 \triangleq \theta_1$	2
$\alpha_4 \triangleq \theta_2$	2
$\alpha_5 \triangleq \theta_1 \cup \theta_2$	3
$\alpha_6 \triangleq \theta_1 \cup \theta_3$	3
$\alpha_7 \triangleq \theta_2 \cup \theta_3$	3
$\alpha_8 \triangleq \theta_1 \cup \theta_2 \cup \theta_3$	4

Table 3.2: $\mathcal{C}_\mathcal{M}(A)$ for hybrid DSm model $\mathcal{M}$

Another example based on Shafer's model

Consider now the same 3D case but including all exclusivity constraints on θ_i, $i = 1, 2, 3$. This corresponds to the 3D Shafer's model $\mathcal{M}^0$ presented in the following Venn diagram.



Then, one gets the following list of elements (with their DSm cardinal) for the restricted D^Θ, which coincides naturally with the classical power set 2^Θ:

$A \in (D^\Theta \equiv 2^\Theta)$	$\mathcal{C}_{\mathcal{M}^0}(A)$
$\alpha_0 \triangleq \emptyset$	0
$\alpha_1 \triangleq \theta_1$	1
$\alpha_2 \triangleq \theta_2$	1
$\alpha_3 \triangleq \theta_3$	1
$\alpha_4 \triangleq \theta_1 \cup \theta_2$	2
$\alpha_5 \triangleq \theta_1 \cup \theta_3$	2
$\alpha_6 \triangleq \theta_2 \cup \theta_3$	2
$\alpha_7 \triangleq \theta_1 \cup \theta_2 \cup \theta_3$	3

Table 3.3: $\mathcal{C}_{\mathcal{M}^0}(A)$ for Shafer's model $\mathcal{M}^0$

The partial ordering of D^Θ based on DSm cardinality does not provide in general an efficient solution to get an interesting structure for the **BM** matrix involved in (3.3), contrary to the structure obtained by Smets in the DST framework as in section 3.1. The partial ordering presented in the sequel will however allow us to get such a nice structure for the matrix calculus of belief functions.

3.2.3 Ordering based on the intrinsic informational content

As already pointed out, the DSm cardinality is insufficient to catch the intrinsic informational content of each element d_i of D^Θ. A better approach to obtain this, is based on the following new function $s(.)$, which describes the *intrinsic information strength* of any $d_i \in D^\Theta$. A previous, but cumbersome, definition of $s(.)$ had been proposed in our previous works [1, 2] but it was difficult to handle and questionable with respect to the formal equivalent (dual) representation of elements belonging to D^Θ.

Definition of the $s(.)$ function

We propose here a better choice for $s(.)$, based on a very simple and natural geometrical interpretation of the relationships between the parts of the Venn diagram belonging to each $d_i \in D^\Theta$. All the values of the $s(.)$ function (stored into a vector $\mathbf{s}$) over D^Θ are defined by the following equation:

$$\mathbf{s} = \mathbf{D}_n \cdot \mathbf{w}_n \tag{3.4}$$

with $\mathbf{s} \triangleq [s(d_0) \ \ldots \ s(d_p)]'$ where p is the cardinal of D^Θ for the model $\mathcal{M}$ under consideration. p is equal to Dedekind's number $d(n) - 1$ if the free-model $\mathcal{M}^f$ is chosen for $\Theta = \{\theta_1, \ldots, \theta_n\}$. $\mathbf{D}_n$ is the *hyper-power set generating matrix*. The components w_i of vector $\mathbf{w}_n$ are obtained from the components of the *DSm encoding basis* vector $\mathbf{u}_n$ as follows (see previous chapter for details about $\mathbf{D}_n$ and $\mathbf{u}_n$) :

$$w_i \triangleq 1/l(u_i) \tag{3.5}$$

where $l(u_i)$ is the length of Smarandache's codification u_i of the part of the Venn diagram of the model $\mathcal{M}$, i.e the number of symbols involved in the codification.

For example, if $u_i =< 123 >$, then $l(u_i) = 3$ just because only three symbols 1, 2, and 3 enter in the codification u_i, thus $w_i = 1/3$.

From this new DSm ordering function s(.) we can partially order all the elements $d_i \in D^\Theta$ by the increasing values of $s(.)$.

Example of ordering on $D^{\Theta=\{\theta_1,\theta_2\}}$ with $\mathcal{M}^f$

In this simple case, the DSm ordering of D^Θ is given by

$\alpha_i \in D^\Theta$	$s(\alpha_i)$
$\alpha_0 = \emptyset$	$s(\alpha_0) = 0$
$\alpha_1 = \theta_1 \cap \theta_2$	$s(\alpha_1) = 1/2$
$\alpha_2 = \theta_1$	$s(\alpha_2) = 1 + 1/2$
$\alpha_3 = \theta_2$	$s(\alpha_3) = 1 + 1/2$
$\alpha_4 = \theta_1 \cup \theta_2$	$s(\alpha_4) = 1 + 1 + 1/2$

Based on this ordering, it can be easily verified that the matrix calculus of the beliefs **Bel** from **m** by equation (3.3), is equivalent to

$$\underbrace{\begin{bmatrix} \text{Bel}(\emptyset) \\ \text{Bel}(\theta_1 \cap \theta_2) \\ \text{Bel}(\theta_1) \\ \text{Bel}(\theta_2) \\ \text{Bel}(\theta_1 \cup \theta_2) \end{bmatrix}}_{\mathbf{Bel}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}}_{\mathbf{BM}_2} \underbrace{\begin{bmatrix} m(\emptyset) \\ m(\theta_1 \cap \theta_2) \\ m(\theta_1) \\ m(\theta_2) \\ m(\theta_1 \cup \theta_2) \end{bmatrix}}_{\mathbf{m}}$$

where the $\mathbf{BM}_2$ matrix has a interesting structure (triangular inferior and unimodular properties, $\det(\mathbf{BM}_2) = \det(\mathbf{BM}_2^{-1}) = 1$). Conversely, the calculus of the generalized basic belief assignment **m** from beliefs **Bel** will be obtained by the inversion of the previous linear system of equations

$$\underbrace{\begin{bmatrix} m(\emptyset) \\ m(\theta_1 \cap \theta_2) \\ m(\theta_1) \\ m(\theta_2) \\ m(\theta_1 \cup \theta_2) \end{bmatrix}}_{\mathbf{m}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 1 \end{bmatrix}}_{\mathbf{MB}_2 = \mathbf{BM}_2^{-1}} \underbrace{\begin{bmatrix} \text{Bel}(\emptyset) \\ \text{Bel}(\theta_1 \cap \theta_2) \\ \text{Bel}(\theta_1) \\ \text{Bel}(\theta_2) \\ \text{Bel}(\theta_1 \cup \theta_2) \end{bmatrix}}_{\mathbf{Bel}}$$

Example of ordering on $D^{\Theta=\{\theta_1,\theta_2,\theta_3\}}$ with $\mathcal{M}^f$

In this more complicated case, the DSm ordering of D^Θ is now given by

$\alpha_i \in D^\Theta, i = 0, ..., 18$	$s(\alpha_i)$
$\emptyset$	0
$\theta_1 \cap \theta_2 \cap \theta_3$	$1/3$
$\theta_1 \cap \theta_2$	$1/3 + 1/2$
$\theta_1 \cap \theta_3$	$1/3 + 1/2$
$\theta_2 \cap \theta_3$	$1/3 + 1/2$
$(\theta_1 \cup \theta_2) \cap \theta_3$	$1/3 + 1/2 + 1/2$
$(\theta_1 \cup \theta_3) \cap \theta_2$	$1/3 + 1/2 + 1/2$
$(\theta_2 \cup \theta_3) \cap \theta_1$	$1/3 + 1/2 + 1/2$
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	$1/3 + 1/2 + 1/2 + 1/2$
θ_1	$1/3 + 1/2 + 1/2 + 1$
θ_2	$1/3 + 1/2 + 1/2 + 1$
θ_3	$1/3 + 1/2 + 1/2 + 1$
$(\theta_1 \cap \theta_2) \cup \theta_3$	$1/3 + 1/2 + 1/2 + 1 + 1/2$
$(\theta_1 \cap \theta_3) \cup \theta_2$	$1/3 + 1/2 + 1/2 + 1 + 1/2$
$(\theta_2 \cap \theta_3) \cup \theta_1$	$1/3 + 1/2 + 1/2 + 1 + 1/2$
$\theta_1 \cup \theta_2$	$1/3 + 1/2 + 1/2 + 1 + 1/2 + 1$
$\theta_1 \cup \theta_3$	$1/3 + 1/2 + 1/2 + 1 + 1/2 + 1$
$\theta_2 \cup \theta_3$	$1/3 + 1/2 + 1/2 + 1 + 1/2 + 1$
$\theta_1 \cup \theta_2 \cup \theta_3$	$1/3 + 1/2 + 1/2 + 1 + 1/2 + 1 + 1$

The order for elements generating the same value of $s(.)$ can be chosen arbitrarily and doesn't change the structure of the matrix $\mathbf{BM}_3$ given right after. That's why only a partial order is possible from $s(.)$. It can be verified that $\mathbf{BM}_3$ holds also the same previous interesting matrix structure properties and that $\det(\mathbf{BM}_3) = \det(\mathbf{BM}_3^{-1}) = 1$. Similar structure can be shown for problems of higher dimensions ($n > 3$).

Although a nice structure for matrix calculus of belief functions has been obtained in this work, and conversely to the recursive construction of $\mathbf{BM}_n$ in DST framework, a recursive algorithm (on dimension n) for the construction of $\mathbf{BM}_n$ from $\mathbf{BM}_{n-1}$ has not yet be found (if such recursive algorithm exists ...) and is still an open difficult problem for further research.

$$
\mathbf{BM}_3 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}
$$

3.3 Conclusion

In this chapter, one has analyzed several issues to obtain an interesting matrix representation of the belief functions defined in the DSmT. For ordering the elements of hyper-power set D^Θ we propose three such orderings: first, using the direct enumeration of isotone Boolean functions, second, based on the DSm cardinality, and third, and maybe the most interesting, by introducing the intrinsic informational strength function $s(.)$ constructed from the DSm encoding basis. The third order permits to get a nice internal structure of the transition matrix $\mathbf{BM}$ in order to compute directly and easily by programming the belief vector $\mathbf{Bel}$ from the basic belief mass vector $\mathbf{m}$ and conversely by inversion of matrix $\mathbf{BM}$.

3.4 References

[1] Dezert J., *Foundations for a New Theory of Plausible and Paradoxical Reasoning*, Information and Security Journal, Edited by Tz. Semerdjiev, ProCon Ltd., Sofia, Bulgaria, Vol. 9, pp. 13-57, 2002.

[2] Dezert J., *Fondations pour une nouvelle théorie du raisonnement plausible et paradoxal*, ONERA Tech. Rep. RT 1/06769/DTIM, Jan. 2003.

[3] Dezert J., Smarandache F., *On the generation of hyper-power sets for the DSmT*, Proc. of the 6th Int. Conf. on inf. fusion (Fusion 2003), Cairns, Australia, july 8-11, 2003.

[4] Dezert J., Smarandache F., *Partial ordering of hyper-power sets and matrix representation of belief functions within DSmT*, Proc. of the 6th Int. Conf. on inf. fusion (Fusion 2003), Cairns, Australia, july 8-11, 2003.

[5] Dezert J., Smarandache F., Daniel M.,*The Generalized Pignistic Transformation*, Proc. of 7th Int.. Conf. on Inf. Fusion, Fusion 2004, Stockholm, Sweden, July 2004.

[6] Kennes R., Smets Ph., *Fast algorithms for Dempster-Shafer theory*, in Uncertainty in Knowledge Bases, B. Bouchon-Meunier, R.R. Yager, L.A. Zadeh (Editors), Lecture Notes in Computer Science 521, Springer-Verlag, Berlin, pp. 14-23, 1991.

[7] Kennes R., *Computational Aspects of the Möbius Transformation of Graphs*, IEEE Trans. on SMC. 22, pp. 201-223, 1992.

[8] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[9] Smets Ph., *Matrix Calculus for Belief Functions*, `http://iridia.ulb.ac.be/~psmets/MatrixRepresentation.pdf`.

Chapter 4

Combination of beliefs on hybrid DSm models

Jean Dezert	Florentin Smarandache
ONERA	Department of Mathematics
29 Av. de la Division Leclerc	University of New Mexico
92320 Châtillon	Gallup, NM 8730
France	U.S.A.

Abstract: *This chapter presents a general method for combining uncertain and paradoxical (i.e. highly conflicting) sources of evidence for a wide class of fusion problems. From the foundations of the DSmT we show how the DSm rule of combination can be extended to take into account all possible integrity constraints (if any) of the problem under consideration due to the true nature of elements/concepts involved into it. We show how Shafer's model can be considered as a specific hybrid DSm model and can be easily handled by the DSmT and one presents here a new efficient alternative to Dempster's rule of combination, following steps of previous researchers towards this quest. Several simple didactic examples are also provided to show the efficiency and the generality of the approach proposed in this work.*

4.1 Introduction

According to each model occurring in real-world fusion problems, we present a general hybrid DSm rule which combines two or more masses of independent sources of information and takes care of constraints, i.e. of sets which might become empty at time t_l or new sets/elements that might arise in the frame at time t_{l+1}. The hybrid DSm rule is applied in a real time when the hyper-power set D^{Θ} changes (i.e. the set of all propositions built from elements of frame Θ with $\cup$ and $\cap$ operators - see [3] for details), either increasing or decreasing its focal elements, or when even Θ decreases or increases influencing the D^Θ as well, thus the dynamicity of our DSmT.

This chapter introduces the reader to the independence of sources of evidences, which needs to be studied deeper in the future, then one defines the models and the hybrid DSm rule, which is different from other rules of combination such as Dempster's, Yager's, Smets', Dubois-Prade's and gives seven numerical examples of applying the hybrid DSm rule in various models and several examples of dynamicity of DSmT, then the Bayesian hybrid DSm models mixture.

4.2 On the independence of the sources of evidences

The notion of independence of the sources of evidence plays a major role in the development of efficient information fusion algorithms but is very difficult to formally establish when manipulating uncertain and paradoxical (i.e. highly conflicting) sources of information. Some attempts to define the independence of uncertain sources of evidences have been proposed by P. Smets and al. in Dempster-Shafer Theory (DST) and Transferable Belief Model in [12, 13, 14] and by other authors in possibility theory [1, 2, 5, 8, 10]. In the following, we consider that n sources of evidences are independent if the internal mechanism by which each source provides its own basic belief assignment doesn't depend on the mechanisms of other sources (i.e. there is no internal relationship between all mechanisms) or if the sources don't share (even partially) same knowledge/experience to establish their own basic belief assignment. This definition doesn't exclude the possibility for independent sources to provide the same (numerical) basic belief assignments. The fusion of dependent uncertain and paradoxical sources is much more complicated because, one has first to identify precisely the piece of redundant information between sources in order to remove it before applying the fusion rules. The problem of combination of dependent sources is under investigation.

4.3 DSm rule of combination for free-DSm models

4.3.1 Definition of the free-DSm model $\mathcal{M}^f(\Theta)$

Let's consider a finite frame $\Theta = \{\theta_1, \dots \theta_n\}$ of the fusion problem under consideration. We abandon Shafer's model by assuming here that the fuzzy/vague/relative nature of elements θ_i $i = 1, \dots, n$ of Θ can be non-exclusive. We assume also that no refinement of Θ into a new finer *exclusive frame of discernment* Θ^{ref} is possible. This is the free-DSm model $\mathcal{M}^f(\Theta)$ which can be viewed as the opposite (if we don't introduce non-existential constraints - see next section) of Shafer's model, denoted $\mathcal{M}^0(\Theta)$ where all θ_i are forced to be exclusive and therefore fully discernable.

4.3.2 Example of a free-DSm model

Let's consider the frame of the problem $\Theta = \{\theta_1, \theta_2, \theta_3\}$. The free Dedekind lattice $D^\Theta = \{\alpha_0, \ldots, \alpha_{18}\}$ over Θ owns the following 19 elements (see chapter 2)

Elements of D^Θ for $\mathcal{M}^f(\Theta)$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \neq \emptyset$	$\alpha_{10} \triangleq \theta_2 \neq \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \neq \emptyset$	$\alpha_{11} \triangleq \theta_3 \neq \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \neq \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \neq \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \neq \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \neq \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \neq \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \neq \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \neq \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \neq \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \neq \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \neq \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \neq \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \neq \emptyset$

The free-DSm model $\mathcal{M}^f(\Theta)$ assumes that all elements α_i, $i > 0$, are non-empty. This corresponds to the following Venn diagram where in Smarandache's codification "i" denotes the part of the diagram which belongs to θ_i only, "ij" denotes the part of the diagram which belongs to θ_i and θ_j only, "ijk" denotes the part of the diagram which belongs to θ_i and θ_j and θ_k only, etc [3]. On such Venn diagram representation of the model, we emphasize the fact that all boundaries of intersections must be seen/interpreted as only vague boundaries just because the nature of elements θ_i can be, in general, only vague, relative and even imprecise (see chapter 6).

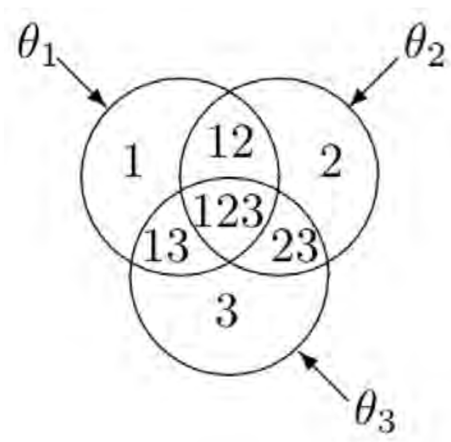


Figure 4.1: Venn Diagram for $\mathcal{M}^f(\Theta)$

For the chapter to be self-contained, we recall here the *classical* DSm rule of combination based on $\mathcal{M}^f(\Theta)$ over the free Dedekind's lattice built from elements of Θ with $\cap$ and $\cup$ operators, i.e. D^Θ.

4.3.3 Classical DSm rule for 2 sources for free-DSm models

For two *independent* uncertain and paradoxical (i.e. highly conflicting) sources of information (experts/-bodies of evidence) providing generalized basic belief assignment $m_1(.)$ and $m_2(.)$ over D^Θ (or over any subset of D^Θ), the classical DSm conjunctive rule of combination $m_{\mathcal{M}_f(\Theta)}(.) \triangleq [m_1 \oplus m_2](.)$ is given by

$$\forall A \neq \emptyset \in D^\Theta, \quad m_{\mathcal{M}^f(\Theta)}(A) \triangleq [m_1 \oplus m_2](A) = \sum_{\substack{X_1, X_2 \in D^\Theta \\ (X_1 \cap X_2) = A}} m_1(X_1) m_2(X_2) \tag{4.1}$$

$m_{\mathcal{M}^f(\Theta)}(\emptyset) = 0$ by definition, unless otherwise specified in special cases when some source assigns a non-zero value to it (like in the Smets TBM approach [9]). This DSm rule of combination is commutative and associative. This rule, dealing with both uncertain and paradoxical/conflicting information, requires no normalization process and can always been applied.

4.3.4 Classical DSm rule for $k \geq 2$ sources for free-DSm models

The above formula can be easily generalized for the free-DSm model $\mathcal{M}^f(\Theta)$ with $k \geq 2$ independent sources in the following way:

$$\forall A \neq \emptyset \in D^\Theta, \quad m_{\mathcal{M}^f(\Theta)}(A) \triangleq [m_1 \oplus \ldots m_k](A) = \sum_{\substack{X_1, \ldots, X_k \in D^\Theta \\ (X_1 \cap \ldots \cap X_k) = A}} \prod_{i=1}^{k} m_i(X_i) \tag{4.2}$$

$m_{\mathcal{M}^f(\Theta)}(\emptyset) = 0$ by definition, unless otherwise specified in special cases when some source assigns a non-zero value to it. This DSm rule of combination is still commutative and associative.

4.4 Presentation of hybrid DSm models

4.4.1 Definition

Let Θ be the general frame of the fusion problem under consideration with n elements θ_1, θ_2, ..., θ_n. A hybrid DSm model $\mathcal{M}(\Theta)$ is defined from the free-DSm model $\mathcal{M}^f(\Theta)$ by introducing some *integrity constraints* on some elements A of D^Θ if one knows with certainty the exact nature of the model corresponding to the problem under consideration. An integrity constraint on A consists in forcing A to be empty (vacuous element), and we will denote such constraint as $A \overset{\mathcal{M}}{\equiv} \emptyset$ which means that A has been forced to $\emptyset$ through the model $\mathcal{M}(\Theta)$. This can be justified by the knowledge of the true nature of each element θ_i of Θ. Indeed, in some fusion problems, some elements θ_i and θ_j of Θ can be fully discernable because they are truly exclusive while other elements cannot be refined into finer exclusive elements. Moreover, it is also possible that for some reason with some new knowledge on the problem, an element or several elements θ_i have to be forced to the empty set (especially if dynamical fusion problems are considered, i.e when Θ varies with space and time). For example, if we consider a list of three potential suspects into a police investigation, it can occur that, during the investigation, one of the suspects can be withdrawn of the initial frame of the problem if his innocence is proven with an ascertainable alibi. The initial basic belief masses provided by sources of information one had on the three suspects, must then be modified by taking into account this new knowledge on the model of the problem.

There exists several possible kinds of integrity constraints which can be introduced in any free-DSm model $\mathcal{M}^f(\Theta)$ actually. The first kind of integrity constraint concerns *exclusivity constraints* by taking into account that some conjunctions of elements $\theta_i,\ldots,\theta_k$ are truly impossible (i.e. $\theta_i\cap\ldots\cap\theta_k \overset{\mathcal{M}}{\equiv} \emptyset$). The second kind of integrity constraint concerns the *non-existential constraints* by taking into account that some disjunctions of elements $\theta_i,\ldots,\theta_k$ are also truly impossible (i.e. $\theta_i\cup\ldots\cup\theta_k \overset{\mathcal{M}}{\equiv} \emptyset$). We exclude from our presentation the completely degenerate case corresponding to the constraint $\theta_1\cup\ldots\cup\theta_n \overset{\mathcal{M}}{\equiv} \emptyset$ (total ignorance) because there is no way and no interest to treat such a vacuous problem. In such a degenerate case, we can just set $m(\emptyset)\triangleq 1$ which is useless because the problem remains vacuous and D^Θ reduces to $\emptyset$. The last kind of possible integrity constraint is a mixture of the two previous ones, like for example $(\theta_i\cap\theta_j)\cup\theta_k$ or any other hybrid proposition/element of D^Θ involving both $\cap$ and $\cup$ operators such that at least one element θ_k is a subset of the constrained proposition. From any $\mathcal{M}^f(\Theta)$, we can thus build several hybrid DSm models depending on the number of integrity constraints one needs to fully characterize the nature of the problem. The introduction of a given integrity constraint $A \overset{\mathcal{M}}{\equiv} \emptyset \in D^\Theta$ implies necessarily the set of inner constraints $B \overset{\mathcal{M}}{\equiv} \emptyset$ for all $B\subset A$. Moreover the introduction of two integrity constraints, say on A and B in D^Θ implies also necessarily the constraint on the emptiness of the disjunction $A\cup B$ which belongs also to D^Θ (because D^Θ is closed under $\cap$ and $\cup$ operators). This implies the emptiness of all $C\in D^\Theta$ such that $C\subset(A\cup B)$. The same remark has to be extended for the case of the introduction of n integrity constraints as well. Shafer's model is the unique and most constrained hybrid DSm model including all possible exclusivity constraints *without non-existential constraint* since all $\theta_i\neq\emptyset\in\Theta$ are forced to be mutually exclusive. Shafer's model is denoted $\mathcal{M}^0(\Theta)$ in the sequel. We denote by $\boldsymbol{\emptyset}_{\mathcal{M}}$ the set of elements of D^Θ which have been forced to be empty in the hybrid DSm model $\mathcal{M}$.

4.4.2 Example 1 : hybrid DSm model with an exclusivity constraint

Let $\Theta=\{\theta_1,\theta_2,\theta_3\}$ be the general frame of the problem under consideration and let's consider the following hybrid DSm model $\mathcal{M}_1(\Theta)$ built by introducing the following exclusivity constraint $\alpha_1\triangleq\theta_1\cap\theta_2\cap\theta_3 \overset{\mathcal{M}_1}{\equiv} \emptyset$. This exclusivity constraint implies however no other constraint because α_1 doesn't contain other elements of D^Θ but itself. Therefore, one has now the following set of elements for D^Θ

Elements of D^Θ for $\mathcal{M}_1(\Theta)$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_1}{\equiv} \emptyset$	$\alpha_{10} \triangleq \theta_2 \neq \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \neq \emptyset$	$\alpha_{11} \triangleq \theta_3 \neq \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \neq \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \neq \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \neq \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \neq \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \neq \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \neq \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \neq \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \neq \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \neq \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \neq \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \neq \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \neq \emptyset$

Hence the initial basic belief mass over D^Θ has to be transferred over the new constrained hyper-power set $D^\Theta(\mathcal{M}_1(\Theta))$ with the 18 elements defined just above (including actually 17 non-empty elements). The mechanism for the transfer of basic belief masses from D^Θ onto $D^\Theta(\mathcal{M}_1(\Theta))$ will be obtained by the hybrid DSm rule of combination presented in the sequel.

4.4.3 Example 2 : hybrid DSm model with another exclusivity constraint

As the second example for a hybrid DSm model $\mathcal{M}_2(\Theta)$, let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and the following exclusivity constraint $\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_2}{\equiv} \emptyset$. This constraint implies also $\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_2}{\equiv} \emptyset$ since $\alpha_1 \subset \alpha_2$. Therefore, one has now the following set of elements for $D^\Theta(\mathcal{M}_2(\Theta))$

Elements of D^Θ for $\mathcal{M}_2(\Theta)$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_2}{\equiv} \emptyset$	$\alpha_{10} \triangleq \theta_2 \neq \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_2}{\equiv} \emptyset$	$\alpha_{11} \triangleq \theta_3 \neq \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \neq \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_2}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \neq \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \neq \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \neq \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \neq \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_2}{\equiv} \alpha_4 \neq \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \neq \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_2}{\equiv} \alpha_3 \neq \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_2}{\equiv} \alpha_5 \neq \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \neq \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \neq \emptyset$

Note that in this case several non-empty elements of $D^\Theta(\mathcal{M}_2(\Theta))$ coincide because of the constraint ($\alpha_6 \overset{\mathcal{M}_2}{\equiv} \alpha_4$, $\alpha_7 \overset{\mathcal{M}_2}{\equiv} \alpha_3$, $\alpha_8 \overset{\mathcal{M}_2}{\equiv} \alpha_5$, $\alpha_{12} \overset{\mathcal{M}_2}{\equiv} \alpha_{11}$). $D^\Theta(\mathcal{M}_2(\Theta))$ has now only 13 different elements. Note that the introduction of both constraints $\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_2}{\equiv} \emptyset$ and $\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_2}{\equiv} \emptyset$ doesn't change the construction of $D^\Theta(\mathcal{M}_2(\Theta))$ because $\alpha_1 \subset \alpha_2$.

4.4.4 Example 3 : hybrid DSm model with another exclusivity constraint

As the third example for a hybrid DSm model $\mathcal{M}_3(\Theta)$, let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and the following exclusivity constraint $\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$. This constraint implies now $\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$ since $\alpha_1 \subset \alpha_6$, but also $\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$ because $\alpha_2 \subset \alpha_6$ and $\alpha_4 \triangleq \theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$ because $\alpha_4 \subset \alpha_6$. Therefore, one has now the following set of elements for $D^\Theta(\mathcal{M}_3(\Theta))$

Elements of D^Θ for $\mathcal{M}_3(\Theta)$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$	$\alpha_{10} \triangleq \theta_2 \neq \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$	$\alpha_{11} \triangleq \theta_3 \neq \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \neq \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_3}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \neq \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \alpha_3 \neq \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_3}{\equiv} \alpha_9 \neq \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \neq \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_3}{\equiv} \alpha_3 \neq \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_3}{\equiv} \alpha_5 \neq \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \neq \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \neq \emptyset$

$D^\Theta(\mathcal{M}_3(\Theta))$ has now only 10 different elements.

4.4.5 Example 4 : Shafer's model

As the fourth particular example for a hybrid DSm model $\mathcal{M}_4(\Theta)$, let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and the following exclusivity constraint $\alpha_8 \triangleq \{(\theta_1 \cap \theta_2) \cup \theta_3\} \cap (\theta_1 \cup \theta_2) \overset{\mathcal{M}_4}{\equiv} \emptyset$. Therefore, one has now the following set of elements for $D^\Theta(\mathcal{M}_4(\Theta))$

Elements of D^Θ for $\mathcal{M}_4(\Theta)$ (Shafer's model)	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{10} \triangleq \theta_2 \neq \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{11} \triangleq \theta_3 \neq \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_4}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_4}{\equiv} \alpha_{10} \neq \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_4}{\equiv} \alpha_9 \neq \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \neq \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_4}{\equiv} \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \neq \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \neq \emptyset$

This model corresponds actually to Shafer's model $\mathcal{M}^0(\Theta)$ because this constraint includes all possible exclusivity constraints between elements θ_i, $i = 1, 2, 3$ since $\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \subset \alpha_8$, $\alpha_2 \triangleq \theta_1 \cap \theta_2 \subset \alpha_8$, $\alpha_3 \triangleq \theta_1 \cap \theta_3 \subset \alpha_8$ and $\alpha_4 \triangleq \theta_2 \cap \theta_3 \subset \alpha_8$. $D^\Theta(\mathcal{M}_4(\Theta))$ has now $2^{|\Theta|} = 8$ different elements and coincides obviously with the classical power set 2^Θ. This corresponds to Shafer's model and serves as the foundation for Dempster-Shafer Theory.

4.4.6 Example 5 : hybrid DSm model with a non-existential constraint

As the fifth example for a hybrid DSm model $\mathcal{M}_5(\Theta)$, let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and the following *non-existential* constraint $\alpha_9 \triangleq \theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$. In other words, we remove θ_1 from the initial frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$. This non-existential constraint implies $\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$, $\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_5}{\equiv} \emptyset$, $\alpha_3 \triangleq \theta_1 \cap \theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$ and $\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$. Therefore, one has now the following set of elements for $D^\Theta(\mathcal{M}_5(\Theta))$

Elements of D^Θ for $\mathcal{M}_5(\Theta)$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$	$\alpha_{10} \triangleq \theta_2 \neq \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_5}{\equiv} \emptyset$	$\alpha_{11} \triangleq \theta_3 \neq \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_5}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \neq \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_5}{\equiv} \alpha_{10} \neq \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_5}{\equiv} \alpha_4 \neq \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_5}{\equiv} \alpha_4 \neq \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_5}{\equiv} \alpha_4 \neq \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \overset{\mathcal{M}_5}{\equiv} \alpha_{10} \neq \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \overset{\mathcal{M}_5}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_5}{\equiv} \alpha_4 \neq \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \overset{\mathcal{M}_5}{\equiv} \alpha_{17} \neq \emptyset$

$D^\Theta(\mathcal{M}_5(\Theta))$ has now 5 different elements and coincides obviously with the hyper-power set $D^{\Theta \setminus \theta_1}$.

4.4.7 Example 6 : hybrid DSm model with two non-existential constraints

As the sixth example for a hybrid DSm model $\mathcal{M}_6(\Theta)$, let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and the following two *non-existential* constraints $\alpha_9 \triangleq \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$ and $\alpha_{10} \triangleq \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$. Actually, these two constraints are equivalent to choose only the following constraint $\alpha_{15} \triangleq \theta_1 \cup \theta_2 \overset{\mathcal{M}_5}{\equiv} \emptyset$. In other words, we remove now both θ_1 and θ_2 from the initial frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$. These non-existential constraints implies now $\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_3 \triangleq \theta_1 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_4 \triangleq \theta_2 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_8 \triangleq \{(\theta_1 \cap \theta_2) \cup \theta_3\} \cap (\theta_1 \cup \theta_2) \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$, $\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$. Therefore, one has now the following set of elements for $D^\Theta(\mathcal{M}_6(\Theta))$:

Elements of D^Θ for $\mathcal{M}_6(\Theta)$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{10} \triangleq \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{11} \triangleq \theta_3 \neq \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \alpha_{11} \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \alpha_{11} \neq \emptyset$

$D^\Theta(\mathcal{M}_6(\Theta))$ reduces now to only two different elements $\emptyset$ and θ_3. $D^\Theta(\mathcal{M}_6(\Theta))$ coincides obviously with the hyper-power set $D^{\Theta\setminus\{\theta_1,\theta_2\}}$. Because there exists only one possible non empty element in $D^\Theta(\mathcal{M}_6(\Theta))$, such kind of a problem is called a *trivial problem*. If one now introduces all non-existential constraints in the free-DSm model, then the initial problem reduces to a *vacuous problem* also called the *impossible problem* corresponding to $m(\emptyset) \equiv 1$ (such kind of a "problem" is not related to reality). Such kinds of trivial or vacuous *problems* are not considered anymore in the sequel since they present no real interest for engineering information fusion problems.

4.4.8 Example 7 : hybrid DSm model with a mixed constraint

As the seventh example for a hybrid DSm model $\mathcal{M}_7(\Theta)$, let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and the following *mixed exclusivity and non-existential* constraint $\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$. This mixed constraint implies $\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\alpha_3 \triangleq \theta_1 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\alpha_4 \triangleq \theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\alpha_8 \triangleq \{(\theta_1 \cap \theta_2) \cup \theta_3\} \cap (\theta_1 \cup \theta_2) \overset{\mathcal{M}_7}{\equiv} \emptyset$ and $\alpha_{11} \triangleq \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$. Therefore, one has now the following set of elements for $D^\Theta(\mathcal{M}_7(\Theta))$

Elements of D^Θ for $\mathcal{M}_7(\Theta)$	
$\alpha_0 \triangleq \emptyset$	
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{10} \triangleq \theta_2 \neq \emptyset$
$\alpha_2 \triangleq \theta_1 \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{11} \triangleq \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$
$\alpha_3 \triangleq \theta_1 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$
$\alpha_4 \triangleq \theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_7}{\equiv} \alpha_{10} \neq \emptyset$
$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_7}{\equiv} \alpha_9 \neq \emptyset$
$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2 \neq \emptyset$
$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \alpha_9 \neq \emptyset$
$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_7}{\equiv} \emptyset$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \alpha_{10} \neq \emptyset$
$\alpha_9 \triangleq \theta_1 \neq \emptyset$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \alpha_{15} \neq \emptyset$

$D^\Theta(\mathcal{M}_7(\Theta))$ reduces now to only four different elements $\emptyset$, θ_1, θ_2, and $\theta_1 \cup \theta_2$.

4.5 DSm rule of combination for hybrid DSm models

In this section, we present a general DSm-hybrid rule of combination able to deal with any hybrid DSm models (including Shafer's model). We will show how this new general rule of combination works with all hybrid DSm models presented in the previous section and we list interesting properties of this new useful and powerful rule of combination.

4.5.1 Notations

Let $\Theta = \{\theta_1, \dots \theta_n\}$ be a frame of *partial discernment* (i.e. a frame Θ for which at least one conjunctive element of $D^\Theta \setminus \{\emptyset\}$ is known to be truly empty) of the constrained fusion problem, and D^Θ the free distributive lattice (hyper-power set) generated by Θ and the empty set $\emptyset$ under $\cap$ and $\cup$ operators. We need to distinguish between the empty set $\emptyset$, which belongs to D^Θ, and by $\emptyset$ we understand a set which is empty all the time (we call it *absolute emptiness* or *absolutely empty*) independent of time, space and model, and all other sets from D^Θ. For example $\theta_1 \cap \theta_2$ or $\theta_1 \cup \theta_2$ or only θ_i itself, $1 \leq i \leq n$, etc, which could be or become empty at a certain time (if we consider a fusion dynamicity) or in a particular model $\mathcal{M}$ (but could not be empty in other model and/or time) (we call a such element *relative emptiness* or *relatively empty*). We'll denote by $\boldsymbol{\emptyset}_{\mathcal{M}}$ the set of relatively empty such elements of D^Θ (i.e. which become empty in a particular model $\mathcal{M}$ or at a specific time). $\boldsymbol{\emptyset}_{\mathcal{M}}$ is the set of integrity constraints which depends on the DSm model $\mathcal{M}$ under consideration, and the model $\mathcal{M}$ depends on the structure of its corresponding *fuzzy Venn Diagram* (number of elements in Θ, number of non-empty intersections, and time in case of dynamic fusion). Through our convention $\emptyset \notin \boldsymbol{\emptyset}_{\mathcal{M}}$. Let's note by $\boldsymbol{\emptyset} \triangleq \{\emptyset, \boldsymbol{\emptyset}_{\mathcal{M}}\}$ the set of all relatively and absolutely empty elements.

For any $A \in D^{\Theta}$, let $\phi(A)$ be the *characteristic non emptiness function* of the set A, i.e. $\phi(A) = 1$ if $A \notin \boldsymbol{\emptyset}$ and $\phi(A) = 0$ otherwise. This function assigns the value zero to all relatively or absolutely empty elements of D^{Θ} through the choice of hybrid DSm model $\mathcal{M}$. Let's define the total ignorance on $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ as $I_t \triangleq \theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$ and the set of relative ignorances as $I_r \triangleq \{\theta_{i_1} \cup \ldots \cup \theta_{i_k}$, where $i_1, ..., i_k \in \{1, 2, ..., n\}$ and $2 \leq k \leq n-1\}$, then the set of all kind of ignorances as $I = I_t \cup I_r$. For any element A in D^{Θ}, one considers $u(A)$ as the union of all singletons θ_i that compose A. For example, if A is a singleton then $u(A) = A$; if $A = \theta_1 \cap \theta_2$ or $A = \theta_1 \cup \theta_2$ then $u(A) = \theta_1 \cup \theta_2$; if $A = (\theta_1 \cap \theta_2) \cup \theta_3$ then $u(A) = \theta_1 \cup \theta_2 \cup \theta_3$. ; by convention $u(\emptyset) \triangleq \emptyset$. The second summation of the hybrid DSm rule (see eq. (4.3) and (4.5) and denoted S_2 in the sequel) transfers the mass of $\emptyset$ [if any; sometimes, in rare cases, $m(\emptyset) > 0$ (for example in Smets' work); we want to catch this particular case as well] to the total ignorance $I_t = \theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$. The other part of the mass of relatively empty elements, θ_i and θ_j together for example, $i \neq j$, goes to the partial ignorance/uncertainty $m(\theta_i \cup \theta_j)$. S_2 multiplies, naturally following the DSm classic network architecture, only the elements of columns of absolutely and relatively empty sets, and then S_2 transfers the mass $m_1(X_1)m_2(X_2)\ldots m_k(X_k)$ either to the element $A \in D^{\theta}$ in the case when $A = u(X_1) \cup u(X_2) \cup \ldots \cup u(X_k)$ is not empty, or if $u(X_1) \cup u(X_2) \cup \ldots \cup u(X_k)$ is empty then the mass $m_1(X_1)m_2(X_2)\ldots m_k(X_k)$ is transferred to the total ignorance. We include all degenerate problems/models in this new DSmT hybrid framework, but the degenerate/vacuous DSm-hybrid model $\mathcal{M}_{\emptyset}$ defined by the constraint $I_t = \theta_1 \cup \theta_2 \cup \ldots \cup \theta_n \overset{\mathcal{M}_{\emptyset}}{\equiv} \emptyset$ which is meaningless and useless.

4.5.2 Programming of the $u(X)$ function

We provide here the issue for programming the calculation of $u(X)$ from the binary representation of any proposition $X \in D^{\Theta}$ expressed in the Dezert-Smarandache order (see chapters 2 and 3). Let's consider the Smarandache codification of elements $\theta_1, \ldots, \theta_n$. One defines the anti-absorbing relationship as follows: element i anti-absorbs element ij (with $i < j$), and let's use the notation $i << ij$, and also $j << ij$; similarly $ij << ijk$ (with $i < j < k$), also $jk << ijk$ and $ik << ijk$. This relationship is transitive, therefore $i << ij$ and $ij << ijk$ involve $i << ijk$; one can also write $i << ij << ijk$ as a chain; similarly one gets $j << ijk$ and $k << ijk$. The anti-absorbing relationship can be generalized for parts with any number of digits, i.e. when one uses the Smarandache codification for the corresponding Venn diagram on $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, with $n \geq 1$. Between elements ij and ik, or between ij and jk there is no anti-absorbing relationship, therefore the anti-absorbing relationship makes a partial order on the parts of the Venn diagram for the free DSm model. If a proposition X is formed by a part only, say $i_1 i_2 \ldots i_r$, in the Smarandache codification, then $u(X) = \theta_{i_1} \cup \theta_{i_2} \cup \ldots \cup \theta_{i_r}$. If X is formed by two or more parts, the first step is to eliminate all anti-absorbed parts, ie. if $A << B$ then $u(A, B) = u(A)$; generally speaking, a part B is anti-absorbed by part A if all digits of A belong to B; for an anti-absorbing chain $A_1 << A_2 << ... << A_s$ one takes A_1 only and the others are eliminated; afterwards, when X is anti-absorbingly irreducible, $u(X)$ will be the unions of all singletons whose indices occur in the remaining parts of X - if one digit occurs many times it is taken only once. For convenience, one provides below the MatLab[1] source code for computing $u(X), X \in D^\Theta$. The input variable u_n of this routine corresponds to the DSm base encoding and can be obtained by the method proposed in chapter 2.

```
%**********************************************
function [UX]=GetUX(u_n,X);
%**********************************************
% GetUX computes the function u(X) involved
% in the DSm hybrid rule of combination.
% Inputs : u_n => Dezert-Smarandache base encoding
%          X   => Element of D^Theta in base u_n
% Example for n=3: if Theta={theta1,theta2,theta3}
%                  then u_3=[1 2 12 3 13 23 123]
% Output : Ux => u(X) expressed in base u_n
% Copyrights (c) 2003 - J. Dezert & F. Smarandache
%**********************************************
UX=zeros(1,size(u_n,2));XP=u_n(find(X==1))';
AF=zeros(size(XP,1),1); XC=[];
for jj=1:size(XP,1)
if(AF(jj)==0),ujj=num2str(XP(jj));
for kk=1:size(XP,1)
if(AF(kk)==0)
ukk=num2str(XP(kk));w=intersect(ujj,ukk);
if(isempty(w)==0),
if((isequal(w,ujj)+isequal(w,ukk))>0)
XC=[XC;str2num(w)];
if(size(ujj,2)<size(ukk,2)),AF(kk)=1;end
if(size(ukk,2)<size(ujj,2)),AF(jj)=1;end
end; end; end; end; end; end
XC=unique(XC);XCS=unique(num2str(XC'));
for ii=1:size(XCS,2), if(XCS(ii)~=' ')
for jj=1:size(u_n,2)
if(isempty(intersect(XCS(ii),num2str(u_n(jj))))==0)
UX(jj)=1;end;end;end;end
```

Matlab source code for computing $u(X), X \in D^\Theta$

Here are some examples for the case $n = 3$: $12 << 123$, i.e. 12 anti-absorbs 123. Between 12 and 23 there is no anti-absorbing relationship.

- If $X = 123$ then $u(X) = \theta_1 \cup \theta_2 \cup \theta_3$.
- If $X = \{23, 123\}$, then $23 << 123$, thus $u(\{23, 123\}) = u(23)$, because 123 has been eliminated, hence $u(X) = u(23) = \theta_2 \cup \theta_3$.
- If $X = \{13, 123\}$, then $13 << 123$, thus $u(\{13, 123\}) = u(13) = \theta_1 \cup \theta_3$.

[1] Matlab is a trademark of The MathWorks, Inc.

- If $X = \{13, 23, 123\}$, then $13 << 123$, thus $u(\{13, 23, 123\}) = u(\{13, 23\}) = \theta_1 \cup \theta_2 \cup \theta_3$ (one takes as theta indices each digit in the $\{13, 23\}$) - if one digit is repeated it is taken only once; between 13 and 23 there is no relation of anti-absorbing.
- If $X = \{3, 13, 23, 123\}$, then $u(X) = u(\{3, 13, 23\})$ because $23 << 123$, then $u(\{3, 13, 23\}) = u(\{3, 13\})$ because $3 << 23$, then $u(\{3, 13\}) = u(3) = \theta_3$ because $3 << 13$.
- If $X = \{1, 12, 13, 23, 123\}$, then one has the anti-absorbing chain: $1 << 12 << 123$, thus $u(X) = u(\{1, 13, 23\}) = u(\{1, 23\})$ because $1 << 13$, and finally $u(X) = \theta_1 \cup \theta_2 \cup \theta_3$.
- If $X = \{1, 2, 12, 13, 23, 123\}$, then $1 << 12 << 123$ and $2 << 23$ thus $u(X) = u(\{1, 2, 13\}) = u(\{1, 2\})$ because $1 << 13$, and finally $u(X) = \theta_1 \cup \theta_2$.
- If $X = \{2, 12, 3, 13, 23, 123\}$, then $2 << 23 << 123$ and $3 << 13$ thus $u(X) = u(\{2, 12, 3\})$, but $2 << 12$ hence $u(X) = u(\{2, 3\}) = \theta_2 \cup \theta_3$.

4.5.3 The hybrid DSm rule of combination for 2 sources

To eliminate the *degenerate vacuous fusion problem* from the presentation, we assume from now on that the given hybrid DSm model $\mathcal{M}$ under consideration is always different from the vacuous model $\mathcal{M}_\emptyset$ (i.e. $I_t \neq \emptyset$). The hybrid DSm rule of combination, associated to a given hybrid DSm model $\mathcal{M} \neq \mathcal{M}_\emptyset$, for two sources is defined for all $A \in D^\Theta$ as:

$$\begin{aligned} m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A)\Big[& \sum_{\substack{X_1, X_2 \in D^\Theta \\ (X_1 \cap X_2) = A}} m_1(X_1) m_2(X_2) \\ & + \sum_{\substack{X_1, X_2 \in \boldsymbol{\emptyset} \\ [(u(X_1) \cup u(X_2)) = A] \vee [(u(X_1) \cup u(X_2) \in \boldsymbol{\emptyset}) \wedge (A = I_t)]}} m_1(X_1) m_2(X_2) \\ & + \sum_{\substack{X_1, X_2 \in D^\Theta \\ (X_1 \cup X_2) = A \\ X_1 \cap X_2 \in \boldsymbol{\emptyset}}} m_1(X_1) m_2(X_2) \Big] \end{aligned} \tag{4.3}$$

The first sum entering in the previous formula corresponds to mass $m_{\mathcal{M}^f(\Theta)}(A)$ obtained by the classic DSm rule of combination (4.1) based on the free-DSm model $\mathcal{M}^f$ (i.e. on the free lattice D^Θ), i.e.

$$m_{\mathcal{M}^f(\Theta)}(A) \triangleq \sum_{\substack{X_1, X_2 \in D^\Theta \\ (X_1 \cap X_2) = A}} m_1(X_1) m_2(X_2) \tag{4.4}$$

The second sum entering in the formula of the DSm-hybrid rule of combination (4.3) represents the mass of all relatively and absolutely empty sets which is transferred to the total or relative ignorances. The third sum entering in the formula of the DSm-hybrid rule of combination (4.3) transfers the sum of relatively empty sets to the non-empty sets in a similar way as it was calculated following the DSm classic rule.

4.5.4 The hybrid DSm rule of combination for $k \geq 2$ sources

The previous formula of hybrid DSm rule of combination can be generalized in the following way for all $A \in D^\Theta$:

$$m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A)\Big[\sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1\cap X_2\cap\ldots\cap X_k)=A}} \prod_{i=1}^{k} m_i(X_i) + \sum_{\substack{X_1,X_2,\ldots,X_k \in \boldsymbol{\emptyset} \\ [(u(X_1)\cup u(X_2)\cup\ldots\cup u(X_k))=A]\vee[(u(X_1)\cup u(X_2)\cup\ldots\cup u(X_k)\in\boldsymbol{\emptyset})\wedge(A=I_t)]}} \prod_{i=1}^{k} m_i(X_i) + \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1\cup X_2\cup\ldots\cup X_k)=A \\ X_1\cap X_2\cap\ldots\cap X_k \in \boldsymbol{\emptyset}}} \prod_{i=1}^{k} m_i(X_i)\Big] \tag{4.5}$$

The first sum entering in the previous formula corresponds to mass $m_{\mathcal{M}^f(\Theta)}(A)$ obtained by the classic DSm rule of combination (4.2) for k sources of information based on the free-DSm model $\mathcal{M}^f$ (i.e. on the free lattice D^Θ), i.e.

$$m_{\mathcal{M}^f(\Theta)}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1\cap X_2\cap\ldots\cap X_k)=A}} \prod_{i=1}^{k} m_i(X_i) \tag{4.6}$$

4.5.5 On the associativity of the hybrid DSm rule

From (4.5) and (4.6), the previous general formula can be rewritten as

$$m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A)\Big[S_1(A) + S_2(A) + S_3(A)\Big] \tag{4.7}$$

where

$$S_1(A) \equiv m_{\mathcal{M}^f(\Theta)}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1\cap X_2\cap\ldots\cap X_k)=A}} \prod_{i=1}^{k} m_i(X_i) \tag{4.8}$$

$$S_2(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in \boldsymbol{\emptyset} \\ [(u(X_1)\cup u(X_2)\cup\ldots\cup u(X_k))=A]\vee[(u(X_1)\cup u(X_2)\cup\ldots\cup u(X_k)\in\boldsymbol{\emptyset})\wedge(A=I_t)]}} \prod_{i=1}^{k} m_i(X_i) \tag{4.9}$$

$$S_3(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1\cup X_2\cup\ldots\cup X_k)=A \\ X_1\cap X_2\cap\ldots\cap X_k \in \boldsymbol{\emptyset}}} \prod_{i=1}^{k} m_i(X_i) \tag{4.10}$$

This rule of combination can be viewed actually as a two-step procedure as follows:

- **Step 1:** Evaluate the combination of the sources over the free lattice D^Θ by the classical DSm rule of combination to get for all $A \in D^\Theta$, $S_1(A) = m_{\mathcal{M}^f(\Theta)}(A)$ using (4.6). This step preserves the commutativity and associativity properties of the combination. When there is no constraint (when using the free DSm model), the hybrid DSm rule reduces to the classic DSm rule because $\boldsymbol{\emptyset} = \{\emptyset\}$ and $m_i(\emptyset) = 0$, $i = 1, \ldots k$ and therefore $\Phi(A) = 1$ and $S_2(A) = S_3(A) = 0$ $\forall A \neq \emptyset \in D^\Theta$. For $A = \emptyset$, $\Phi(A) = 0$ and thus $m_{\mathcal{M}^f}(\emptyset) = 0$.

- **Step 2:** Transfer the masses of the integrity constraints of the hybrid DSm model $\mathcal{M}$ according to formula (4.7). Note that this step is necessary only if one has reliable information about the real integrity constraints involved in the fusion problem under consideration. More precisely, when some constraints are introduced to deal with a given hybrid DSm model $\mathcal{M}(\Theta)$, there exists some propositions $A \overset{\mathcal{M}}{=} \emptyset$ for which $\Phi(A) = 0$. For these propositions, it is actually not necessary to compute $S_1(A)$, $S_2(A)$ and $S_3(A)$ since the product $\Phi(A)[S_1(A) + S_2(A) + S_3(A)]$ equals zero because $\Phi(A) = 0$. This reduces the cost of computations. For propositions $A \overset{\mathcal{M}}{\neq} \emptyset$ characterized by $\Phi(A) = 1$, the derivation of $S_1(A)$, $S_2(A)$ and $S_3(A)$ is necessary to get $m_{\mathcal{M}(\Theta)}(A)$. The last part of the hybrid DSm combination mechanism (called compression step) consists in gathering (summing) all masses corresponding to same proposition because of the constraints of the model. As example, if one considers the 3D frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$ with the constraint $\theta_2 \cap \theta_3 \overset{\mathcal{M}}{=} \emptyset$, then the mass resulting from the hybrid DSm fusion rule (4.7) $m_{\mathcal{M}(\Theta)}(\theta_1 \cup (\theta_2 \cap \theta_3))$ will have to be added to $m_{\mathcal{M}(\Theta)}(\theta_1)$ because $\theta_1 \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}}{=} \theta_1$ due to the constraint $\theta_2 \cap \theta_3 \overset{\mathcal{M}}{=} \emptyset$.

The second step does not preserve the full associativity of the rule (same remark applies also with Yager's or Dubois & Prade's rules), but this is not a fundamental requirement because this problem can be easily circumvented by keeping in parallel the two previous steps 1 and 2. The fusion has to start always on the free-DSm model. The second step is applied only when some integrity constraints are introduced and before the decision-making. In other words, if one has only 2 independent sources of information giving $m_1(.)$ and $m_2(.)$ and some integrity constraints on the frame Θ, one applies step 1 to get[2] $m^{1,2}_{\mathcal{M}^f(\Theta)}(.) = [m_1 \oplus m_2](.)$ defined on the free-DSm model and then one applies step 2 to get the final result $m^{1,2}_{\mathcal{M}(\Theta)}(.)$ on the hybrid-model. If a third source of information is introduced, say $m_3(.)$, one combines it with the two previous ones by step 1 again to get $m^{1,2,3}_{\mathcal{M}^f(\Theta)}(.) = [m_3 \oplus m^{1,2}_{\mathcal{M}^f(\Theta)}](.)$ and then one applies step 2 to get the final result $m^{1,2,3}_{\mathcal{M}(\Theta)}(.)$ on the hybrid-model $\mathcal{M}(\Theta)$.

There is no technical difficulty to process the fusion in this way and that's why the full associativity of the fusion rule is not so fundamental despite of all criticisms against the alternatives to Dempster's rules emerging in litterature over the years. The full/direct associativity property is realized only through Demspter's rule of combination when working on Shafer's model. This is one of reasons for which Dempster's rule is usually preferred to the other fusion rules, but in turn this associativity property (through the normalization factor $1 - m(\emptyset)$) is also one of the main sources of the criticisms for more than twenty years because one knows that Dempster's rule fails to provide coherent results when conflicts become high (see chapters 5 and 12 for examples) and something else must be carried out anyway to prevent problems. This matter of fact is quite paradoxical.

[2] We introduce here the notation $m^{1,2}(.)$ to explicitly express that the resulting mass is related to the combination of sources 1 and 2 only.

To avoid the loss of information in the fusion, one has first to combine all sources using DSm rule on free-DSm model and then to adapt the belief masses according to the integrity constraints of the model $\mathcal{M}$. If one first adapts the local masses $m_1(.), ...m_k(.)$ to the hybrid-model $\mathcal{M}$ and afterwards one applies the combination rule, the fusion becomes only suboptimal because some information is lost forever during the transfer of masses of integrity constraints. The same remark holds if the transfer of masses of integrity constraints is done at some intermediate steps after the fusion of m sources with $m < k$.

Let's note also that this formula of transfer is more general (because we include the possibilities to introduce both exclusivity constraints and non-existential constraints as well) and more precise (because we explicitly consider all different relative emptiness of elements into the general transfer formula (4.7)) than the generic transfer formulas used in the DST framework proposed as alternative rules to Dempster's rule of combination [6] and discussed in section 4.5.10.

4.5.6 Property of the hybrid DSm Rule

The following equality holds:

$$\sum_{A\in D^\Theta} m_{\mathcal{M}(\Theta)}(A) = \sum_{A\in D^\Theta} \phi(A)\Big[S_1(A) + S_2(A) + S_3(A)\Big] = 1 \tag{4.11}$$

Proof: Let's first prove that $\sum_{A\in D^\Theta} m(A) = 1$ where all masses $m(A)$ are obtained by the DSm classic rule. Let's consider each mass $m_i(.)$ provided by the ith source of information, for $1 \leq i \leq k$, as a vector of $d = \mid D^\Theta \mid$ dimension, whose sum of components is equal to one, i.e. $m_i(D^\Theta) = [m_{i1}, m_{i2}, \ldots, m_{id}]$, and $\sum_{j=1,d} m_{ij} = 1$. Thus, for $k \geq 2$ sources of information, the mass matrix becomes

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1d} \\ \ldots & \ldots & \ldots & \ldots \\ m_{k1} & m_{k2} & \ldots & m_{kd} \end{bmatrix}$$

If one denotes the sets in D^Θ by A_1, A_2, ..., A_d (it doesn't matter in what order one lists them) then the column (j) in the matrix represents the masses assigned to A_j by each source of information s_1, s_2, …, s_k; for example $s_i(A_j) = m_{ij}$, where $1 \leq i \leq k$. According to the DSm network architecture [3], all the products in this network will have the form $m_{1j_1} m_{2j_2} \ldots m_{kj_k}$, i.e. one element only from each matrix row, and no restriction about the number of elements from each matrix column, $1 \leq j_1, j_2, \ldots, j_k \leq d$. Each such product will enter in the fusion mass of one set only from D^Θ. Hence the sum of all components of the fusion mass is equal to the sum of all these products, which is equal to

$$\prod_{i=1}^{k} \sum_{j=1}^{d} m_{ij} = \prod_{i=1}^{k} 1 = 1 \tag{4.12}$$

The hybrid DSm rule has three sums S_1, S_2, and S_3. Let's separate the mass matrix $\mathbf{M}$ into two disjoint sub-matrices $\mathbf{M}_\emptyset$ formed by the columns of all absolutely and relatively empty sets, and $\mathbf{M}_N$ formed by the columns of all non-empty sets. According to the DSm network architecture (for $k \geq 2$ rows):

- S_1 is the sum of all products resulted from the multiplications of the columns of $\mathbf{M}_N$ following the DSm network architecture such that the intersection of their corresponding sets is non-empty, i.e. the sum of masses of all non-empty sets before any mass of absolutely or relatively empty sets could be transferred to them;

- S_2 is the sum of all products resulted from the multiplications of $\mathbf{M}_\emptyset$ following the DSm network architecture, i.e. a partial sum of masses of absolutely and relatively empty sets transferred to the ignorances in $I \triangleq I_t \cup I_r$ or to singletons of Θ.

- S_3 is the sum of all the products resulted from the multiplications of the columns of $\mathbf{M}_N$ and $\mathbf{M}_\emptyset$ together, following the DSm network architecture, but such that at least a column is from each of them, and also the sum of all products of columns of $\mathbf{M}_N$ such that the intersection of their corresponding sets is empty (what did not enter into the previous sum S_1), i.e. the remaining sum of masses of absolutely or relatively empty sets transferred to the non-empty sets of the hybrid DSm model $\mathcal{M}$.

If one now considers all the terms (each such term is a product of the form $m_{1j_1}m_{2j_2}\ldots m_{kj_k}$) of these three sums, we get exactly the same terms as in the DSm network architecture for the DSm classic rule, thus the sum of all terms occurring in S_1, S_2, and S_3 is 1 (see formula (4.12)) which completes the proof. The hybrid DSm rule naturally derives from the DSm classic rule. Entire masses of relatively and absolutely empty sets in a given hybrid DSm model $\mathcal{M}$ are transferred to non-empty sets according to the formula (4.7) and thus

$$\forall A \in \boldsymbol{\emptyset} \subset D^\Theta, \quad m_{\mathcal{M}(\Theta)}(A) = 0 \tag{4.13}$$

The entire mass of a relatively empty set (from D^Θ) which has in its expression θ_{j_1}, θ_{j_2}, ..., θ_{j_r}, with $1 \leq r \leq n$ will generally be distributed among the θ_{j_1}, θ_{j_2}, ..., θ_{j_r} or their unions or intersections, and the distribution follows the way of multiplication from the DSm classic rule, explained by the DSm network architecture [3]. Thus, because nothing is lost, nothing is gained, the sum of all $m_{\mathcal{M}(\Theta)}(A)$ is equal to 1 as just proven previously, and fortunately no normalization constant is needed which could bring a loss of information in the fusion rule. The three summations $S_1(.)$, $S_3(.)$ and $S_3(.)$ are disjoint because:

- $S_1(.)$ multiplies the columns corresponding to non-empty sets only - but such that the intersections of the sets corresponding to these columns are non-empty [from the definition of DSm classic rule];

- $S_2(.)$ multiplies the columns corresponding to absolutely and relatively empty sets only;

- $S_3(.)$ multiplies:

a) either the columns corresponding to absolutely or relatively empty sets with the columns corresponding to non-empty sets such that at least a column corresponds to an absolutely or relatively emptyset and at least a column corresponds to a non-emptyset,

b) or the columns corresponding to non-empty sets - but such that the intersections of the sets corresponding to these columns are empty.

The multiplications are following the DSm network architecture, i.e. any product has the above general form: $m_{1j_1}m_{2j_2}\ldots m_{kj_k}$, i.e. any product contains as factor one element only from each row of the mass matrix $\mathbf{M}$ and the total number of factors in a product is equal to k. The function $\phi(A)$ automatically assigns the value zero to the mass of any empty set, and allows the calculation of masses of all non-empty sets.

4.5.7 On the programming of the hybrid DSm rule

We briefly give here an issue for a fast programming of the DSm rule of combination. Let's consider $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, the sources $\mathcal{B}_1$, $\mathcal{B}_2$,..., $\mathcal{B}_k$, and $p = \min\{n, k\}$. One needs to check only the focal sets, i.e. sets (i.e. propositions) whose masses assigned to them by these sources are not all zero. Thus, if $\mathbf{M}$ is the mass matrix, and we consider a set A_j in D^Θ, then the column (j) corresponding to A_j, i.e. $(m_{1j}\ m_{2j}\ \ldots\ m_{kj})$ transposed has not to be identical to the null-vector of k-dimension $(0\ 0\ \ldots\ 0)$ transposed. Let $D^\Theta(\text{step}_1)$ be formed by all focal sets at the beginning (after sources $\mathcal{B}_1$, $\mathcal{B}_2$,..., $\mathcal{B}_k$ have assigned masses to the sets in D^Θ). Applying the DSm classic rule, besides the sets in $D^\Theta(\text{step}_1)$ one adds r-intersections of sets in $D^\Theta(\text{step}_1)$, thus:

$$D^\Theta(\text{step}_2) = D^\Theta(\text{step}_1) \vee \{A_{i_1} \wedge A_{i_2} \wedge \ldots \wedge A_{i_r}\}$$

where A_{i_1}, A_{i_2}, ..., A_{i_r} belong to $D^\Theta(\text{step}_1)$ and $2 \leq r \leq p$.

Applying the hybrid DSm rule, due to its S_2 and S_3 summations, besides the sets in $D^\Theta(\text{step}_2)$ one adds r-unions of sets and the total ignorance in $D^\Theta(\text{step}_2)$, thus:

$$D^\Theta(\text{step}_3) = D^\Theta(\text{step}_2) \vee I_t \vee \{A_{i_1} \vee A_{i_2} \vee \ldots \vee A_{i_r}\}$$

where A_{i_1}, A_{i_2}, ..., A_{i_r} belong to $D^\Theta(\text{step}_2)$ and $2 \leq r \leq p$.

This means that instead of computing the masses of all sets in D^Θ, one needs to first compute the masses of all focal sets (step 1), second the masses of their r-intersections (step 2), and third the masses of r-unions of all previous sets and the mass of total ignorance (step 3).

4.5.8 Application of the hybrid DSm rule on previous examples

We present in this section some numerical results of the hybrid DSm rule of combination for 2 independent sources of information. We examine the seven previous examples in order to help the reader to check by himself (or herself) the validity of our new general formula. We will not go in details in the derivations, but we just present the main intermediary results $S_1(A)$, $S_2(A)$ and $S_3(A)$ (defined in (4.8), (4.9), (4.10)) involved into the general formula (4.3) with setting the number of sources to combine to $k = 2$. Now let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and two independent bodies of evidence with the generalized basic belief assignments[3] $m_1(.)$ and $m_2(.)$ given in the following table[4].

Element A of D^Θ	$m_1(A)$	$m_2(A)$	$m_{\mathcal{M}^f(\Theta)}(A)$
$\emptyset$	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3$	0	0	0.16
$\theta_2 \cap \theta_3$	0	0.20	0.19
$\theta_1 \cap \theta_3$	0.10	0	0.12
$(\theta_1 \cup \theta_2) \cap \theta_3$	0	0	0.01
θ_3	0.30	0.10	0.10
$\theta_1 \cap \theta_2$	0.10	0.20	0.22
$(\theta_1 \cup \theta_3) \cap \theta_2$	0	0	0.05
$(\theta_2 \cup \theta_3) \cap \theta_1$	0	0	0
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	0	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3$	0	0	0
θ_2	0.20	0.10	0.03
$(\theta_1 \cap \theta_3) \cup \theta_2$	0	0	0
$\theta_2 \cup \theta_3$	0	0	0
θ_1	0.10	0.20	0.08
$(\theta_2 \cap \theta_3) \cup \theta_1$	0	0	0.02
$\theta_1 \cup \theta_3$	0.10	0.20	0.02
$\theta_1 \cup \theta_2$	0.10	0	0
$\theta_1 \cup \theta_2 \cup \theta_3$	0	0	0

The right column of the table gives the result obtained by the DSm rule of combination based on the free-DSm model. The following sections give the results obtained by the hybrid DSm rule on the seven previous examples of section 4.3. The tables show the values of $\phi(A)$, $S_1(A)$, $S_2(A)$ and $S_3(A)$ to help the reader to check the validity of these results. It is important to note that the values of $S_1(A)$, $S_2(A)$ and $S_3(A)$ when $\phi(A) = 0$ do not need to be computed in practice but are provided here only for verification.

[3] A general example with $m_1(A) > 0$ and $m_2(A) > 0$ for all $A \neq \emptyset \in D^\Theta$ will be briefly presented in next section.

[4] The order of elements of D^Θ is the order obtained from the generation of isotone Boolean functions - see chapter 2.

4.5.8.1 Application of the hybrid DSm rule on example 1

Here is the numerical result corresponding to example 1 with the hybrid-model $\mathcal{M}_1$ (i.e with the exclusivity constraint $\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_1}{\equiv} \emptyset$). The right column of the table provides the result obtained using the hybrid DSm rule, ie. $\forall A \in D^\Theta$, $m_{\mathcal{M}_1(\Theta)}(A) = \phi(A)\big[S_1(A) + S_2(A) + S_3(A)\big]$

Element A of D^Θ	$\phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}_1(\Theta)}(A)$
$\emptyset$	0	0	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_1}{\equiv} \emptyset$	0	0.16	0	0	0
$\theta_2 \cap \theta_3$	1	0.19	0	0	0.19
$\theta_1 \cap \theta_3$	1	0.12	0	0	0.12
$(\theta_1 \cup \theta_2) \cap \theta_3$	1	0.01	0	0.02	0.03
θ_3	1	0.10	0	0	0.10
$\theta_1 \cap \theta_2$	1	0.22	0	0	0.22
$(\theta_1 \cup \theta_3) \cap \theta_2$	1	0.05	0	0.02	0.07
$(\theta_2 \cup \theta_3) \cap \theta_1$	1	0	0	0.02	0.02
$\{(\theta_1 \cap \theta_2) \cup \theta_3\} \cap (\theta_1 \cup \theta_2)$	1	0	0	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3$	1	0	0	0.07	0.07
θ_2	1	0.03	0	0	0.03
$(\theta_1 \cap \theta_3) \cup \theta_2$	1	0	0	0.01	0.01
$\theta_2 \cup \theta_3$	1	0	0	0	0
θ_1	1	0.08	0	0	0.08
$(\theta_2 \cap \theta_3) \cup \theta_1$	1	0.02	0	0.02	0.04
$\theta_1 \cup \theta_3$	1	0.02	0	0	0.02
$\theta_1 \cup \theta_2$	1	0	0	0	0
$\theta_1 \cup \theta_2 \cup \theta_3$	1	0	0	0	0

$$\mathbf{D}_{\mathcal{M}_1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

From the previous table of this first numerical example, we see in column corresponding to $S_3(A)$ how the initial combined mass $m_{\mathcal{M}^f(\Theta)}(\theta_1 \cap \theta_2 \cap \theta_3) \equiv S_1(\theta_1 \cap \theta_2 \cap \theta_3) = 0.16$ is transferred (due to the constraint of $\mathcal{M}_1$) only onto the elements $(\theta_1 \cup \theta_2) \cap \theta_3$, $(\theta_1 \cup \theta_3) \cap \theta_2$, $(\theta_2 \cup \theta_3) \cap \theta_1$, $(\theta_1 \cap \theta_2) \cup \theta_3$, $(\theta_1 \cap \theta_3) \cup \theta_2$, and $(\theta_2 \cap \theta_3) \cup \theta_1$ of D^Θ. We can easily check that the sum of the elements of the column for $S_3(A)$ is equal to $m_{\mathcal{M}^f(\Theta)}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.16$ (i.e. to the sum of $S_1(A)$ for which $\phi(A) = 0$) and that the sum of $S_2(A)$ for which $\phi(A) = 1$ is equal to the sum of $S_3(A)$ for which $\phi(A) = 0$ (in this example the sum is zero). Thus after introducing the constraint, the initial hyper-power set D^Θ reduces to 18 elements as follows

$$D^\Theta_{\mathcal{M}_1} = \{\emptyset, \theta_2 \cap \theta_3, \theta_1 \cap \theta_3, (\theta_1 \cup \theta_2) \cap \theta_3, \theta_3, \theta_1 \cap \theta_2, (\theta_1 \cup \theta_3) \cap \theta_2, (\theta_2 \cup \theta_3) \cap \theta_1, \{(\theta_1 \cap \theta_2) \cup \theta_3\} \cap (\theta_1 \cup \theta_2),$$
$$(\theta_1 \cap \theta_2) \cup \theta_3, \theta_2, (\theta_1 \cap \theta_3) \cup \theta_2, \theta_2 \cup \theta_3, \theta_1, (\theta_2 \cap \theta_3) \cup \theta_1, \theta_1 \cup \theta_3, \theta_1 \cup \theta_2, \theta_1 \cup \theta_2 \cup \theta_3\}$$

As detailed in chapter 2, the elements of $D^{\Theta}_{\mathcal{M}_1}$ can be described and encoded by the matrix product $\mathbf{D}_{\mathcal{M}_1} \cdot \mathbf{u}_{\mathcal{M}_1}$ with $\mathbf{D}_{\mathcal{M}_1}$ given above and the basis vector $\mathbf{u}_{\mathcal{M}_1}$ defined[5] as $\mathbf{u}_{\mathcal{M}_1} = [<1><2><12><3><13><23>]'$. Actually $\mathbf{u}_{\mathcal{M}_1}$ is directly obtained from $\mathbf{u}_{\mathcal{M}^f}$ by removing its component $<123>$ corresponding to the constraint introduced by the model $\mathcal{M}_1$. In general, the encoding matrix $\mathbf{D}_{\mathcal{M}}$ for a given hybrid DSm model $\mathcal{M}$ is obtained from $\mathbf{D}_{\mathcal{M}^f}$ by removing all its columns corresponding to the constraints of the chosen model $\mathcal{M}$ and all the rows corresponding to redundant/equivalent propositions. In this particular example with model $\mathcal{M}_1$, we will just have to remove the last column of $\mathbf{D}_{\mathcal{M}^f}$ to get $\mathbf{D}_{\mathcal{M}_1}$ and no row is removed from $\mathbf{D}_{\mathcal{M}^f}$ because there is no redundant/equivalent proposition involved in this example. This suppression of some rows of $\mathbf{D}_{\mathcal{M}^f}$ will however occur in the next examples.

4.5.8.2 Application of the hybrid DSm rule on example 2

Here is the numerical result corresponding to example 2 with the hybrid-model $\mathcal{M}_2$ (i.e with the exclusivity constraint $\theta_1 \cap \theta_2 \overset{\mathcal{M}_2}{\equiv} \emptyset \Rightarrow \theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_2}{\equiv} \emptyset$). One gets now

Element A of D^{Θ}	$\phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}_2(\Theta)}(A)$
$\emptyset$	0	0	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_2}{\equiv} \emptyset$	0	0.16	0	0	0
$\theta_2 \cap \theta_3$	1	0.19	0	0	0.19
$\theta_1 \cap \theta_3$	1	0.12	0	0	0.12
$(\theta_1 \cup \theta_2) \cap \theta_3$	1	0.01	0	0.02	0.03
θ_3	1	0.10	0	0	0.10
$\theta_1 \cap \theta_2 \overset{\mathcal{M}_2}{\equiv} \emptyset$	0	0.22	0	0.02	0
$(\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_2}{\equiv} \theta_2 \cap \theta_3$	1	0.05	0	0.02	0.07
$(\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_2}{\equiv} \theta_1 \cap \theta_3$	1	0	0	0.02	0.02
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_2}{\equiv} (\theta_1 \cup \theta_2) \cap \theta_3$	1	0	0	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_2}{\equiv} \theta_3$	1	0	0	0.07	0.07
θ_2	1	0.03	0	0.05	0.08
$(\theta_1 \cap \theta_3) \cup \theta_2$	1	0	0	0.01	0.01
$\theta_2 \cup \theta_3$	1	0	0	0	0
θ_1	1	0.08	0	0.04	0.12
$(\theta_2 \cap \theta_3) \cup \theta_1$	1	0.02	0	0.02	0.04
$\theta_1 \cup \theta_3$	1	0.02	0	0.04	0.06
$\theta_1 \cup \theta_2$	1	0	0.02	0.07	0.09
$\theta_1 \cup \theta_2 \cup \theta_3$	1	0	0	0	0

[5] $\mathbf{D}_{\mathcal{M}^f}$ was denoted $\mathbf{D}_n$ and $\mathbf{u}_{\mathcal{M}^f}$ as $\mathbf{u}_n$ in chapter 2.

From the previous table of this numerical example, we see in the column corresponding to $S_3(A)$ how the initial combined masses $m_{\mathcal{M}^f(\Theta)}(\theta_1 \cap \theta_2 \cap \theta_3) \equiv S_1(\theta_1 \cap \theta_2 \cap \theta_3) = 0.16$ and $m_{\mathcal{M}^f(\Theta)}(\theta_1 \cap \theta_2) \equiv S_1(\theta_1 \cap \theta_2) = 0.22$ are transferred (due to the constraint of $\mathcal{M}_2$) onto some elements of D^Θ. We can easily check that the sum of the elements of the column for $S_3(A)$ is equal to $0.16 + 0.22 = 0.38$ (i.e. to the sum of $S_1(A)$ for which $\phi(A) = 0$) and that the sum of $S_2(A)$ for which $\phi(A) = 1$ is equal to the sum of $S_3(A)$ for which $\phi(A) = 0$ (this sum is 0.02). Because some elements of D^Θ are now equivalent due to the constraints of $\mathcal{M}_2$, we have to sum all the masses corresponding to same equivalent propositions/elements (by example $\{(\theta_1 \cap \theta_2) \cup \theta_3\} \cap (\theta_1 \cup \theta_2) \overset{\mathcal{M}_2}{\equiv} (\theta_1 \cup \theta_2) \cap \theta_3$). This can be viewed as the final compression step. One then gets the reduced hyper-power set $D^\Theta_{\mathcal{M}_2}$ having now 13 different elements with the combined belief masses presented in the following table.

The basis vector $\mathbf{u}_{\mathcal{M}_2}$ and the encoding matrix $\mathbf{D}_{\mathcal{M}_2}$ for the elements of $D^\Theta_{\mathcal{M}_2}$ are given by $\mathbf{u}_{\mathcal{M}_2} = [< 1 >< 2 >< 3 >< 13 >< 23 >]'$ and below. Actually $\mathbf{u}_{\mathcal{M}_2}$ is directly obtained from $\mathbf{u}_{\mathcal{M}^f}$ by removing its components $< 12 >$ and $< 123 >$ corresponding to the constraints introduced by the model $\mathcal{M}_2$.

Element A of $D^\Theta_{\mathcal{M}_2}$	$m_{\mathcal{M}_2(\Theta)}(A)$
$\emptyset$	0
$\theta_2 \cap \theta_3$	$0.19 + 0.07 = 0.26$
$\theta_1 \cap \theta_3$	$0.12 + 0.02 = 0.14$
$(\theta_1 \cup \theta_2) \cap \theta_3$	$0.03 + 0 = 0.03$
θ_3	$0.10 + 0.07 = 0.17$
θ_2	0.08
$(\theta_1 \cap \theta_3) \cup \theta_2$	0.01
$\theta_2 \cup \theta_3$	0
θ_1	0.12
$(\theta_2 \cap \theta_3) \cup \theta_1$	0.04
$\theta_1 \cup \theta_3$	0.06
$\theta_1 \cup \theta_2$	0.09
$\theta_1 \cup \theta_2 \cup \theta_3$	0

and

$$\mathbf{D}_{\mathcal{M}_2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

4.5.8.3 Application of the hybrid DSm rule on example 3

Here is the numerical result corresponding to example 3 with the hybrid-model $\mathcal{M}_3$ (i.e with the exclusivity constraint $(\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$). This constraint implies directly $\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$, $\theta_1 \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$ and $\theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$. One gets now

Element A of D^Θ	$\phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}_3(\Theta)}(A)$
$\emptyset$	0	0	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$	0	0.16	0	0	0
$\theta_2 \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \emptyset$	0	0.19	0	0	0
$\theta_1 \cap \theta_3$	1	0.12	0	0	0.12
$(\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_3}{\equiv} \theta_1 \cap \theta_3$	1	0.01	0	0.02	0.03
θ_3	1	0.10	0	0.06	0.16
$\theta_1 \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$	0	0.22	0	0.02	0
$(\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_3}{\equiv} \emptyset$	0	0.05	0	0.02	0
$(\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_3}{\equiv} \theta_1 \cap \theta_3$	1	0	0	0.02	0.02
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_3}{\equiv} \theta_1 \cap \theta_3$	1	0	0	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_3}{\equiv} \theta_3$	1	0	0	0.07	0.07
θ_2	1	0.03	0	0.09	0.12
$(\theta_1 \cap \theta_3) \cup \theta_2$	1	0	0	0.01	0.01
$\theta_2 \cup \theta_3$	1	0	0	0.05	0.05
θ_1	1	0.08	0	0.04	0.12
$(\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_3}{\equiv} \theta_1$	1	0.02	0	0.02	0.04
$\theta_1 \cup \theta_3$	1	0.02	0	0.06	0.08
$\theta_1 \cup \theta_2$	1	0	0.02	0.09	0.11
$\theta_1 \cup \theta_2 \cup \theta_3$	1	0	0.02	0.05	0.07

We see in the column corresponding to $S_3(A)$ how the initial combined masses $m_{\mathcal{M}^f(\Theta)}((\theta_1 \cup \theta_3) \cap \theta_2) \equiv S_1((\theta_1 \cup \theta_3) \cap \theta_2) = 0.05$, $m_{\mathcal{M}^f(\Theta)}(\theta_1 \cap \theta_2 \cap \theta_3) \equiv S_1(\theta_1 \cap \theta_2 \cap \theta_3) = 0.16$, $m_{\mathcal{M}^f(\Theta)}(\theta_2 \cap \theta_3) \equiv S_1(\theta_2 \cap \theta_3) = 0.19$ and $m_{\mathcal{M}^f(\Theta)}(\theta_1 \cap \theta_2) \equiv S_1(\theta_1 \cap \theta_2) = 0.22$ are transferred (due to the constraint of $\mathcal{M}_3$) onto some elements of D^Θ. We can easily check that the sum of the elements of the column for $S_3(A)$ is equal to $0.05 + 0.16 + 0.19 + 0.22 = 0.62$ (i.e. to the sum of $S_1(A)$ for which $\phi(A) = 0$) and that the sum of $S_2(A)$ for which $\phi(A) = 1$ is equal to $0.02+0.02 = 0.04$ (i.e. to the sum of $S_3(A)$ for which $\phi(A) = 0$). Due to the model $\mathcal{M}_3$, one has to sum all the masses corresponding to same equivalent propositions. Thus after the final compression step, one gets the reduced hyper-power set $D^\Theta_{\mathcal{M}_3}$ having only 10 different elements with the following combined belief masses. The basis vector $\mathbf{u}_{\mathcal{M}_3}$ is given by $\mathbf{u}_{\mathcal{M}_3} = [<1><2><3><13>]'$ and the encoding matrix $\mathbf{D}_{\mathcal{M}_3}$ is shown just right after.

Element A of $D^{\Theta}_{\mathcal{M}_3}$	$m_{\mathcal{M}_3(\Theta)}(A)$
$\emptyset$	0
$\theta_1 \cap \theta_3$	$0.12 + 0.03 + 0.02 + 0 = 0.17$
θ_3	$0.16 + 0.07 = 0.23$
θ_2	0.12
$(\theta_1 \cap \theta_3) \cup \theta_2$	0.01
$\theta_2 \cup \theta_3$	0.05
θ_1	$0.12 + 0.04 = 0.16$
$\theta_1 \cup \theta_3$	0.08
$\theta_1 \cup \theta_2$	0.11
$\theta_1 \cup \theta_2 \cup \theta_3$	0.07

and

$$\mathbf{D}_{\mathcal{M}_3} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

4.5.8.4 Application of the hybrid DSm rule on example 4 (Shafer's model)

Here is the result obtained with the hybrid-model $\mathcal{M}_4$, i.e. Shafer's model.

Element A of D^{Θ}	$\phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}_4(\Theta)}(A)$
$\emptyset$	0	0	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0.16	0	0	0
$\theta_2 \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0.19	0	0	0
$\theta_1 \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0.12	0	0	0
$(\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0.01	0	0.02	0
θ_3	1	0.10	0	0.07	0.17
$\theta_1 \cap \theta_2 \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0.22	0	0.02	0
$(\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0.05	0	0.02	0
$(\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0	0	0.02	0
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_4}{\equiv} \emptyset$	0	0	0	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_4}{\equiv} \theta_3$	1	0	0	0.07	0.07
θ_2	1	0.03	0	0.09	0.12
$(\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_4}{\equiv} \theta_2$	1	0	0	0.01	0.01
$\theta_2 \cup \theta_3$	1	0	0	0.05	0.05
θ_1	1	0.08	0	0.06	0.14
$(\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_4}{\equiv} \theta_1$	1	0.02	0	0.02	0.04
$\theta_1 \cup \theta_3$	1	0.02	0	0.15	0.17
$\theta_1 \cup \theta_2$	1	0	0.02	0.09	0.11
$\theta_1 \cup \theta_2 \cup \theta_3$	1	0	0.06	0.06	0.12

From the previous table of this numerical example, we see in column corresponding to $S_3(A)$ how the initial combined masses of the eight elements forced to the empty set by the constraints of the model $\mathcal{M}_4$ are transferred onto some elements of D^Θ. We can easily check that the sum of the elements of the column for $S_3(A)$ is equal to $0.16 + 0.19 + 0.12 + 0.01 + 0.22 + 0.05 = 0.75$ (i.e. to the sum of $S_1(A)$ for which $\phi(A) = 0$) and that the sum of $S_2(A)$ for which $\phi(A) = 1$ is equal to the sum of $S_3(A)$ for which $\phi(A) = 0$ (this sum is $0.02 + 0.06 = 0.08 = 0.02 + 0.02 + 0.02 + 0.02$).

After the final compression step (i.e. the clustering of all equivalent propositions), one gets the reduced hyper-power set $D^\Theta_{\mathcal{M}_4}$ having only $2^3 = 8$ (corresponding to the classical power set 2^Θ) with the following combined belief masses:

Element A of $D^\Theta_{\mathcal{M}_4}$	$m_{\mathcal{M}_4(\Theta)}(A)$
$\emptyset$	0
θ_3	$0.17 + 0.07 = 0.24$
θ_2	$0.12 + 0.01 = 0.13$
$\theta_2 \cup \theta_3$	0.05
θ_1	$0.14 + 0.04 = 0.18$
$\theta_1 \cup \theta_3$	0.17
$\theta_1 \cup \theta_2$	0.11
$\theta_1 \cup \theta_2 \cup \theta_3$	0.12

and $\mathbf{D}_{\mathcal{M}_4} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

The basis vector $\mathbf{u}_{\mathcal{M}_4}$ is given by $\mathbf{u}_{\mathcal{M}_4} = [<1><2><3>]'$ and the encoding matrix $\mathbf{D}_{\mathcal{M}_4}$ is shown just above.

4.5.8.5 Application of the hybrid DSm rule on example 5

The following table presents the numerical result corresponding to example 5 with the hybrid-model $\mathcal{M}_5$ including the non-existential constraint $\theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$. This non-existential constraint implies $\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$, $\theta_1 \cap \theta_2 \overset{\mathcal{M}_5}{\equiv} \emptyset$, $\theta_1 \cap \theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$ and $(\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$.

From the table, we see in the column corresponding to $S_3(A)$ how the initial combined masses of the 5 elements forced to the empty set by the constraints of the model $\mathcal{M}_5$ are transferred onto some elements of D^Θ. We can easily check that the sum of the elements of the column for $S_3(A)$ is equal to $0 + 0.16 + 0.12 + 0.22 + 0 + 0.08 = 0.58$ (i.e. to the sum of $S_1(A)$ for which $\phi(A) = 0$) and that the sum of $S_2(A)$ for which $\phi(A) = 1$ is equal to the sum of $S_3(A)$ for which $\phi(A) = 0$ (this sum is $0.02 + 0.06 + 0.04 = 0.12 = 0.02 + 0.02 + 0.08$).

Element A of D^Θ	$\phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}_5(\Theta)}(A)$
$\emptyset$	0	0	0	0	0
$\theta_1\cap\theta_2\cap\theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$	0	0.16	0	0	0
$\theta_2\cap\theta_3$	1	0.19	0	0	0.19
$\theta_1\cap\theta_3 \overset{\mathcal{M}_5}{\equiv} \emptyset$	0	0.12	0	0	0
$(\theta_1\cup\theta_2)\cap\theta_3 \overset{\mathcal{M}_5}{\equiv} \theta_2\cap\theta_3$	1	0.01	0	0.02	0.03
θ_3	1	0.10	0	0.01	0.11
$\theta_1\cap\theta_2 \overset{\mathcal{M}_5}{\equiv} \emptyset$	0	0.22	0	0.02	0
$(\theta_1\cup\theta_3)\cap\theta_2 \overset{\mathcal{M}_5}{\equiv} \theta_2\cap\theta_3$	1	0.05	0	0.02	0.07
$(\theta_2\cup\theta_3)\cap\theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$	0	0	0	0.02	0
$(\theta_1\cap\theta_2)\cup(\theta_1\cap\theta_3)\cup(\theta_2\cap\theta_3) \overset{\mathcal{M}_5}{\equiv} \theta_2\cap\theta_3$	1	0	0	0	0
$(\theta_1\cap\theta_2)\cup\theta_3 \overset{\mathcal{M}_5}{\equiv} \theta_3$	1	0	0	0.07	0.07
θ_2	1	0.03	0	0.05	0.08
$(\theta_1\cap\theta_3)\cup\theta_2 \overset{\mathcal{M}_5}{\equiv} \theta_2$	1	0	0	0.01	0.01
$\theta_2\cup\theta_3$	1	0	0	0	0
$\theta_1 \overset{\mathcal{M}_5}{\equiv} \emptyset$	0	0.08	0.02	0.08	0
$(\theta_2\cap\theta_3)\cup\theta_1 \overset{\mathcal{M}_5}{\equiv} \theta_2\cap\theta_3$	1	0.02	0	0.02	0.04
$\theta_1\cup\theta_3 \overset{\mathcal{M}_5}{\equiv} \theta_3$	1	0.02	0.02	0.17	0.21
$\theta_1\cup\theta_2 \overset{\mathcal{M}_5}{\equiv} \theta_2$	1	0	0.06	0.09	0.15
$\theta_1\cup\theta_2\cup\theta_3 \overset{\mathcal{M}_5}{\equiv} \theta_2\cup\theta_3$	1	0	0.04	0	0.04

After the final compression step (i.e. the clustering of all equivalent propositions), one gets the reduced hyper-power set $D^\Theta_{\mathcal{M}_5}$ having only 5 different elements according to:

Element A of $D^\Theta_{\mathcal{M}_5}$	$m_{\mathcal{M}_5(\Theta)}(A)$
$\emptyset$	0
$\theta_2\cap\theta_3$	$0.19+0.03+0.07+0+0.04=0.33$
θ_3	$0.11+0.07+0.21=0.39$
θ_2	$0.08+0.01+0.15=0.24$
$\theta_2\cup\theta_3$	$0+0.04=0.04$

and $$\mathbf{D}_{\mathcal{M}_5} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

The basis vector $\mathbf{u}_{\mathcal{M}_5}$ is given by $\mathbf{u}_{\mathcal{M}_5} = [< 2 >< 3 >< 23 >]'$. and the encoding matrix $\mathbf{D}_{\mathcal{M}_5}$ is shown just above.

4.5.8.6 Application of the hybrid DSm rule on example 6

Here is the numerical result corresponding to example 6 with the hybrid-model $\mathcal{M}_6$ including the two non-existential constraint $\theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$ and $\theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$. This is a degenerate example actually, since no uncertainty arises in such trivial model. We just want to show here that the hybrid DSm rule still works in this example and provide a legitimate result. By applying the hybrid DSm rule of combination, one now gets:

Element A of D^Θ	$\phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}_6(\Theta)}(A)$
$\emptyset$	0	0	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.16	0	0	0
$\theta_2 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.19	0	0	0
$\theta_1 \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.12	0	0	0
$(\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.01	0	0.02	0
θ_3	1	0.10	0	0.07	0.17
$\theta_1 \cap \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.22	0	0.02	0
$(\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.05	0	0.02	0
$(\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0	0	0.02	0
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0	0	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \theta_3$	1	0	0	0.07	0.07
$\theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.03	0.02	0.11	0
$(\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0	0	0.01	0
$\theta_2 \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \theta_3$	1	0	0.04	0.05	0.09
$\theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.08	0	0.08	0
$(\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0.02	0	0.02	0
$\theta_1 \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \theta_3$	1	0.02	0.02	0.19	0.23
$\theta_1 \cup \theta_2 \overset{\mathcal{M}_6}{\equiv} \emptyset$	0	0	0.21	0.12	0
$\theta_1 \cup \theta_2 \cup \theta_3 \overset{\mathcal{M}_6}{\equiv} \theta_3$	1	0	0.36	0.08	0.44

We can still verify that the sum of $S_3(A)$ (i.e. 0.88) equals the sum of $S_1(A)$ for which $\phi(A) = 0$ and that the sum of $S_2(A)$ for which $\phi(A) = 1$ (i.e. 0.42) equals the sum of $S_3(A)$ for which $\phi(A) = 0$. After the clustering of all equivalent propositions, one gets the reduced hyper-power set $D^\Theta_{\mathcal{M}_6}$ having only 2 different elements according to:

Element A of $D^\Theta_{\mathcal{M}_6}$	$m_{\mathcal{M}_6(\Theta)}(A)$
$\emptyset$	0
θ_3	$0.17 + 0.07 + 0.09 + 0.23 + 0.44 = 1$

The encoding matrix $\mathbf{D}_{\mathcal{M}_6}$ and the basis vector $\mathbf{u}_{\mathcal{M}_6}$ for the elements of $D^\Theta_{\mathcal{M}_6}$ reduce to $\mathbf{D}_{\mathcal{M}_6} = [01]'$ and $\mathbf{u}_{\mathcal{M}_6} = [< 3 >]$.

4.5.8.7 Application of the hybrid DSm rule on example 7

Here is the numerical result corresponding to example 7 with the hybrid-model $\mathcal{M}_7$ including the *mixed exclusivity and non-existential* constraint $(\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$. This mixed constraint implies $\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\theta_1 \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\theta_1 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $(\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $(\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $(\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_7}{\equiv} \emptyset$, $\{(\theta_1 \cap \theta_2) \cup \theta_3\} \cap (\theta_1 \cup \theta_2) \overset{\mathcal{M}_7}{\equiv} \emptyset$ and $\theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$. By applying the hybrid DSm rule of combination, one gets:

Element A of D^Θ	$\phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}_7(\Theta)}(A)$
$\emptyset$	0	0	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0.16	0	0	0
$\theta_2 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0.19	0	0	0
$\theta_1 \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0.12	0	0	0
$(\theta_1 \cup \theta_2) \cap \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0.01	0	0.02	0
$\theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0.10	0.03	0.10	0
$\theta_1 \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0.22	0	0.02	0
$(\theta_1 \cup \theta_3) \cap \theta_2 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0.05	0	0.02	0
$(\theta_2 \cup \theta_3) \cap \theta_1 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0	0	0.02	0
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3) \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0	0	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \emptyset$	0	0	0	0.07	0
θ_2	1	0.03	0	0.09	0.12
$(\theta_1 \cap \theta_3) \cup \theta_2 \overset{\mathcal{M}_7}{\equiv} \theta_2$	1	0	0	0.01	0.01
$\theta_2 \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \theta_2$	1	0	0.06	0.05	0.11
θ_1	1	0.08	0	0.06	0.14
$(\theta_2 \cap \theta_3) \cup \theta_1 \overset{\mathcal{M}_7}{\equiv} \theta_1$	1	0.02	0	0.02	0.04
$\theta_1 \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \theta_1$	1	0.02	0.01	0.22	0.25
$\theta_1 \cup \theta_2$	1	0	0.02	0.09	0.11
$\theta_1 \cup \theta_2 \cup \theta_3 \overset{\mathcal{M}_7}{\equiv} \theta_1 \cup \theta_2$	1	0	0.16	0.06	0.22

After the clustering of all equivalent propositions, one gets the reduced hyper-power set $D^\Theta_{\mathcal{M}_7}$ having only 4 different elements according to:

Element A of $D^\Theta_{\mathcal{M}_7}$	$m_{\mathcal{M}_7(\Theta)}(A)$
$\emptyset$	0
θ_2	$0.12 + 0.01 + 0.11 = 0.24$
θ_1	$0.14 + 0.04 + 0.25 = 0.43$
$\theta_1 \cup \theta_2$	$0.11 + 0.22 = 0.33$

The basis vector $\mathbf{u}_{\mathcal{M}_7}$ and the encoding matrix $\mathbf{D}_{\mathcal{M}_7}$ for the elements of $D^{\Theta}_{\mathcal{M}_7}$ are given by

$$\mathbf{u}_{\mathcal{M}_7} = [<1><2>]' \qquad \text{and} \qquad \mathbf{D}_{\mathcal{M}_7} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

We can still verify that the sum of $S_3(A)$ (i.e. 0.85) equals the sum of $S_1(A)$ for which $\phi(A) = 0$ and that the sum of $S_2(A)$ for which $\phi(A) = 1$ (i.e. 0.25) equals the sum of $S_3(A)$ for which $\phi(A) = 0$.

4.5.9 Example with more general basic belief assignments $m_1(.)$ and $m_2(.)$

We present in this section the numerical results of the hybrid DSm rule of combination applied upon the seven previous models $\mathcal{M}_i$, $i = 1, ..., 7$ with two general basic belief assignments $m_1(.)$ and $m_2(.)$ such that $m_1(A) > 0$ and $m_2(A) > 0$ for all $A \neq \emptyset \in D^{\Theta=\{\theta_1,\theta_2,\theta_3\}}$. We just provide here the results. The verification is left to the reader. The following table presents the numerical values chosen for $m_1(.)$ and $m_2(.)$ and the result of the fusion obtained by the classical DSm rule of combination

Element A of D^{Θ}	$m_1(A)$	$m_2(A)$	$m_{\mathcal{M}^f}(A)$
$\emptyset$	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3$	0.01	0.40	0.4389
$\theta_2 \cap \theta_3$	0.04	0.03	0.0410
$\theta_1 \cap \theta_3$	0.03	0.04	0.0497
$(\theta_1 \cup \theta_2) \cap \theta_3$	0.01	0.02	0.0257
θ_3	0.03	0.04	0.0311
$\theta_1 \cap \theta_2$	0.02	0.20	0.1846
$(\theta_1 \cup \theta_3) \cap \theta_2$	0.02	0.01	0.0156
$(\theta_2 \cup \theta_3) \cap \theta_1$	0.03	0.04	0.0459
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	0.04	0.03	0.0384
$(\theta_1 \cap \theta_2) \cup \theta_3$	0.04	0.03	0.0296
θ_2	0.02	0.01	0.0084
$(\theta_1 \cap \theta_3) \cup \theta_2$	0.01	0.02	0.0221
$\theta_2 \cup \theta_3$	0.20	0.02	0.0140
θ_1	0.01	0.02	0.0109
$(\theta_2 \cap \theta_3) \cup \theta_1$	0.02	0.01	0.0090
$\theta_1 \cup \theta_3$	0.04	0.03	0.0136
$\theta_1 \cup \theta_2$	0.03	0.04	0.0175
$\theta_1 \cup \theta_2 \cup \theta_3$	0.40	0.01	0.0040

The following table shows the results obtained by the hybrid DSm rule before the final compression step of all redundant propositions for the hybrid DSm models presented in the previous examples.

Element A of D^Θ	$m_{\mathcal{M}_1}(A)$	$m_{\mathcal{M}_2}(A)$	$m_{\mathcal{M}_3}(A)$	$m_{\mathcal{M}_4}(A)$	$m_{\mathcal{M}_5}(A)$	$m_{\mathcal{M}_6}(A)$	$m_{\mathcal{M}_7}(A)$
$\emptyset$	0	0	0	0	0	0	0
$\theta_1 \cap \theta_2 \cap \theta_3$	0	0	0	0	0	0	0
$\theta_2 \cap \theta_3$	0.0573	0.0573	0	0	0.0573	0	0
$\theta_1 \cap \theta_3$	0.0621	0.0621	0.0621	0	0	0	0
$(\theta_1 \cup \theta_2) \cap \theta_3$	0.0324	0.0324	0.0335	0	0.0334	0	0
θ_3	0.0435	0.0435	0.0460	0.0494	0.0459	0.0494	0
$\theta_1 \cap \theta_2$	0.1946	0	0	0	0	0	0
$(\theta_1 \cup \theta_3) \cap \theta_2$	0.0323	0.0365	0	0	0.0365	0	0
$(\theta_2 \cup \theta_3) \cap \theta_1$	0.0651	0.0719	0.0719	0	0	0	0
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	0.0607	0.0704	0.0743	0	0.0764	0	0
$(\theta_1 \cap \theta_2) \cup \theta_3$	0.0527	0.0613	0.0658	0.0792	0.0687	0.0792	0
θ_2	0.0165	0.0207	0.0221	0.0221	0.0207	0	0.0221
$(\theta_1 \cap \theta_3) \cup \theta_2$	0.0274	0.0309	0.0340	0.0375	0.0329	0	0.0375
$\theta_2 \cup \theta_3$	0.0942	0.1346	0.1471	0.1774	0.1518	0.1850	0.1953
θ_1	0.0151	0.0175	0.0175	0.0195	0	0	0.0195
$(\theta_2 \cap \theta_3) \cup \theta_1$	0.0182	0.0229	0.0243	0.0295	0.0271	0	0.0295
$\theta_1 \cup \theta_3$	0.0299	0.0385	0.0419	0.0558	0.0489	0.0589	0.0631
$\theta_1 \cup \theta_2$	0.0299	0.0412	0.0452	0.0544	0.0498	0	0.0544
$\theta_1 \cup \theta_2 \cup \theta_3$	0.1681	0.2583	0.3143	0.4752	0.3506	0.6275	0.5786

The next tables present the final results of the hybrid DSm rule of combination after the compression step (the merging of all equivalent redundant propositions) presented in previous examples.

Element A of $D^\Theta_{\mathcal{M}_7}$	$m_{\mathcal{M}_7(\Theta)}(A)$
$\emptyset$	0
θ_2	0.2549
θ_1	0.1121
$\theta_1 \cup \theta_2$	0.6330

On example no. 7

Element A of $D^\Theta_{\mathcal{M}_6}$	$m_{\mathcal{M}_6(\Theta)}(A)$
$\emptyset$	0
θ_3	1

On example no. 6

Element A of $D^\Theta_{\mathcal{M}_5}$	$m_{\mathcal{M}_5(\Theta)}(A)$
$\emptyset$	0
$\theta_2 \cap \theta_3$	0.2307
θ_3	0.1635
θ_2	0.1034
$\theta_2 \cup \theta_3$	0.5024

On example no. 5

Element A of $D^{\Theta}_{\mathcal{M}_4}$	$m_{\mathcal{M}_4(\Theta)}(A)$
$\emptyset$	0
θ_3	0.1286
θ_2	0.0596
$\theta_2 \cup \theta_3$	0.1774
θ_1	0.0490
$\theta_1 \cup \theta_3$	0.0558
$\theta_1 \cup \theta_2$	0.0544
$\theta_1 \cup \theta_2 \cup \theta_3$	0.4752

On example no 4

Element A of $D^{\Theta}_{\mathcal{M}_3}$	$m_{\mathcal{M}_3(\Theta)}(A)$
$\emptyset$	0
$\theta_1 \cap \theta_3$	0.2418
θ_3	0.1118
θ_2	0.0221
$(\theta_1 \cap \theta_3) \cup \theta_2$	0.0340
$\theta_2 \cup \theta_3$	0.1471
θ_1	0.0418
$\theta_1 \cup \theta_3$	0.0419
$\theta_1 \cup \theta_2$	0.0452
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3143

On example no 3

Element A of $D^{\Theta}_{\mathcal{M}_2}$	$m_{\mathcal{M}_2(\Theta)}(A)$
$\emptyset$	0
$\theta_2 \cap \theta_3$	0.0938
$\theta_1 \cap \theta_3$	0.1340
$(\theta_1 \cup \theta_2) \cap \theta_3$	0.1028
θ_3	0.1048
θ_2	0.0207
$(\theta_1 \cap \theta_3) \cup \theta_2$	0.0309
$\theta_2 \cup \theta_3$	0.1346
θ_1	0.0175
$(\theta_2 \cap \theta_3) \cup \theta_1$	0.0229
$\theta_1 \cup \theta_3$	0.0385
$\theta_1 \cup \theta_2$	0.0412
$\theta_1 \cup \theta_2 \cup \theta_3$	0.2583

On example no 2

Element A of $D^{\Theta}_{\mathcal{M}_1}$	$m_{\mathcal{M}_1(\Theta)}(A)$
$\emptyset$	0
$\theta_2 \cap \theta_3$	0.0573
$\theta_1 \cap \theta_3$	0.0621
$(\theta_1 \cup \theta_2) \cap \theta_3$	0.0324
θ_3	0.0435
$\theta_1 \cap \theta_2$	0.1946
$(\theta_1 \cup \theta_3) \cap \theta_2$	0.0323
$(\theta_2 \cup \theta_3) \cap \theta_1$	0.0651
$(\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	0.0607
$(\theta_1 \cap \theta_2) \cup \theta_3$	0.0527
θ_2	0.0165
$(\theta_1 \cap \theta_3) \cup \theta_2$	0.0274
$\theta_2 \cup \theta_3$	0.0942
θ_1	0.0151
$(\theta_2 \cap \theta_3) \cup \theta_1$	0.0182
$\theta_1 \cup \theta_3$	0.0299
$\theta_1 \cup \theta_2$	0.0299
$\theta_1 \cup \theta_2 \cup \theta_3$	0.1681

On example no 1

4.5.10 The hybrid DSm rule versus Dempster's rule of combination

In its essence, the hybrid DSm rule of combination is close to Dubois and Prade's rule of combination (see chapter 1 and [4]) but more general and precise because it works on $D^\Theta \supset 2^\Theta$ and allows us to include all possible exclusivity and non-existential constraints for the model one has to work with. The advantage of using the hybrid DSm rule is that it does not require the calculation of weighting factors, nor a normalization. The hybrid DSm rule of combination is definitely not equivalent to Dempster's rule of combination as one can easily prove in the following very simple example:

Let's consider $\Theta = \{\theta_1, \theta_2\}$ and the two sources in full contradiction providing the following basic belief assignments

$$m_1(\theta_1) = 1 \qquad m_1(\theta_2) = 0$$

$$m_2(\theta_1) = 0 \qquad m_2(\theta_2) = 1$$

Using the classic DSm rule of combination working with the free DSm model $\mathcal{M}^f$, one gets

$$m_{\mathcal{M}^f}(\theta_1) = 0 \qquad m_{\mathcal{M}^f}(\theta_2) = 0 \qquad m_{\mathcal{M}^f}(\theta_1 \cap \theta_2) = 1 \qquad m_{\mathcal{M}^f}(\theta_1 \cup \theta_2) = 0$$

If one forces θ_1 and θ_2 to be exclusive to work with Shafer's model $\mathcal{M}^0$, then the Dempster's rule of combination can not be applied in this limit case because of the full contradiction of the two sources of information. One gets the undefined operation $0/0$. But the hybrid DSm rule can be applied in such limit case because it transfers the mass of this empty set ($\theta_1 \cap \theta_2 \equiv \emptyset$ because of the choice of the model $\mathcal{M}^0$) to non-empty set(s), and one gets:

$$m_{\mathcal{M}^0}(\theta_1) = 0 \qquad m_{\mathcal{M}^0}(\theta_2) = 0 \qquad m_{\mathcal{M}^0}(\theta_1 \cap \theta_2) = 0 \qquad m_{\mathcal{M}^0}(\theta_1 \cup \theta_2) = 1$$

This result is coherent in this very simple case with Yager's and Dubois-Prade's rule of combination [11, 4].

Now let examine the behavior of the numerical result when introducing a small variation $\epsilon > 0$ on initial basic belief assignments $m_1(.)$ and $m_2(.)$ as follows:

$$m_1(\theta_1) = 1 - \epsilon \qquad m_1(\theta_2) = \epsilon \qquad \text{and} \qquad m_2(\theta_1) = \epsilon \qquad m_2(\theta_2) = 1 - \epsilon$$

As shown in figure 4.2, $\lim_{\epsilon \to 0} m_{DS}(.)$, where $m_{DS}(.)$ is the result obtained from the Dempster's rule of combination, is given by

$$m_{DS}(\theta_1) = 0.5 \qquad m_{DS}(\theta_2) = 0.5 \qquad m_{DS}(\theta_1 \cap \theta_2) = 0 \qquad m_{DS}(\theta_1 \cup \theta_2) = 0$$

This result is very questionable because it assigns same belief on θ_1 and θ_2 which is more informational than to assign all the belief to the total ignorance. The assignment of the belief to the total ignorance appears to be more justified from our point of view because it properly reflects the almost total contradiction between the two sources and in such cases, it seems legitimate that the information can be drawn from the fusion. When we apply the hybrid DSm rule of combination (using Shafer's model $\mathcal{M}^0$), one gets the expected belief assignment on the total ignorance, i.e. $m_{\mathcal{M}^0}(\theta_1 \cup \theta_2) = 1$. The figure below shows the evolution of belief assignments on θ_1, θ_2 and $\theta_1 \cup \theta_2$ with ϵ obtained with the classical Dempster rule and the hybrid DSm rule based on Shafer's model $\mathcal{M}^0$ (i.e. $\theta_1 \cap \theta_2 \overset{\mathcal{M}_0}{\equiv} \emptyset$).

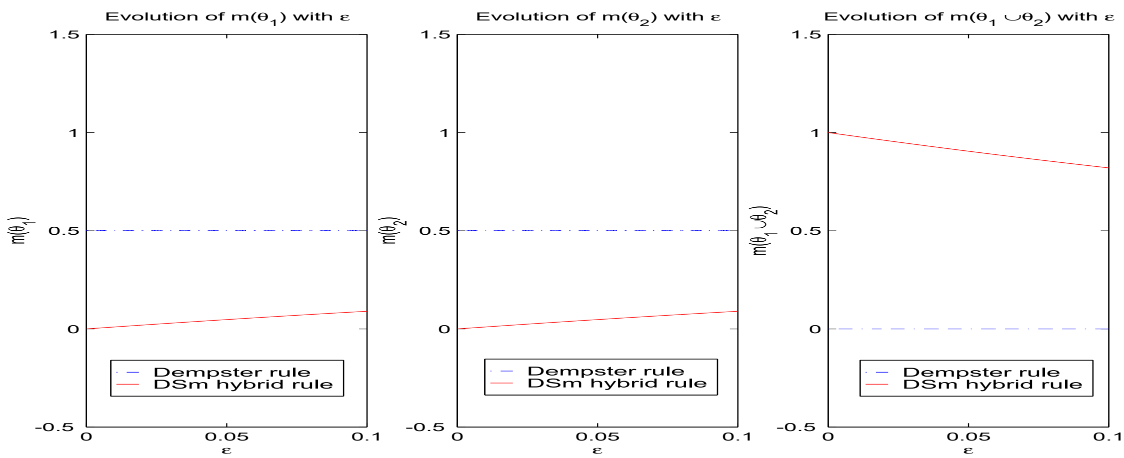


Figure 4.2: Comparison of Dempster's rule with the hybrid DSm rule on $\Theta = \{\theta_1, \theta_2\}$

4.6 Dynamic fusion

The hybrid DSm rule of combination presented in this paper has been developed for *static* problems, but is also directly applicable for easily handling *dynamic fusion* problems in real time as well, since at each temporal change of the models, one can still apply such a hybrid rule. If D^Θ changes, due to the dynamicity of the frame Θ, from time t_l to time t_{l+1}, i.e. some of its elements which at time t_l were not empty become (or are proven) empty at time t_{l+1}, or vice versa: if new elements, empty at time t_l, arise non-empty at time t_{l+1}, this hybrid DSm rule can be applied again at each change. If Θ stays the same but its set non-empty elements of D^Θ increases, then again apply the hybrid DSm rule.

4.6.1 Example 1

Let's consider the testimony fusion problem[6] with the frame

$$\Theta(t_l) \triangleq \{\theta_1 \equiv \text{young}, \theta_2 \equiv \text{old}, \theta_3 \equiv \text{white hairs}\}$$

with the following two basic belief assignments

$$m_1(\theta_1) = 0.5 \qquad m_1(\theta_3) = 0.5 \qquad \text{and} \qquad m_2(\theta_2) = 0.5 \qquad m_2(\theta_3) = 0.5$$

[6]This problem has been proposed to the authors in a private communication by L. Cholvy in 2002.

By applying the classical DSm fusion rule, one then gets

$$m_{\mathcal{M}^f(\Theta(t_l))}(\theta_1 \cap \theta_2) = 0.25 \qquad m_{\mathcal{M}^f(\Theta(t_l))}(\theta_1 \cap \theta_3) = 0.25$$

$$m_{\mathcal{M}^f(\Theta(t_l))}(\theta_2 \cap \theta_3) = 0.25 \qquad m_{\mathcal{M}^f(\Theta(t_l))}(\theta_3) = 0.25$$

Suppose now that at time t_{l+1}, one knows that young people don't have white hairs (i.e $\theta_1 \cap \theta_3 \equiv \emptyset$). How can we update the previous fusion result with this new information on the model of the problem? We solve it with the hybrid DSm rule, which transfers the mass of the empty sets (imposed by the constraints on the new model $\mathcal{M}$ available at time t_{l+1}) to the non-empty sets of D^Θ, going on the track of the DSm classic rule. Using the hybrid DSm rule with the constraint $\theta_1 \cap \theta_3 \equiv \emptyset$, one then gets:

$$m_{\mathcal{M}}(\theta_1 \cap \theta_2) = 0.25 \qquad m_{\mathcal{M}}(\theta_2 \cap \theta_3) = 0.25 \qquad m_{\mathcal{M}}(\theta_3) = 0.25$$

and the mass $m_{\mathcal{M}}(\theta_1 \cap \theta_3) = 0$, because $\theta_1 \cap \theta_3 = \{\text{young}\} \cap \{\text{white hairs}\} \overset{\mathcal{M}}{\equiv} \emptyset$ and its previous mass $m_{\mathcal{M}^f(\Theta(t_l))}(\theta_1 \cap \theta_3) = 0.25$ is transferred to $m_{\mathcal{M}}(\theta_1 \cup \theta_3) = 0.25$ by the hybrid DSm rule.

4.6.2 Example 2

Let $\Theta(t_l) = \{\theta_1, \theta_2, \dots, \theta_n\}$ be a list of suspects and let's consider two observers who eyewitness the scene of plunder at a museum in Bagdad and who testify to the radio and TV the identities of thieves using the basic beliefs assignments $m_1(.)$ and $m_2(.)$ defined on $D^{\Theta(t_l)}$, where t_l represents the time of the observation. Afterwards, at time t_{l+1}, one finds out that one suspect, among this list $\Theta(t_l)$, say θ_i, could not be a suspect because he was on duty in another place, evidence which was certainly confirmed. Therefore he has to be taken off the suspect list $\Theta(t_l)$, and a new frame of discernment results in $\Theta(t_{l+1})$. If this one changes again, one applies again the hybrid DSm of combining of evidences, and so on. This is a typically dynamical example where models change with time and where one needs to adapt fusion results with the current model over time. In the meantime, one can also take into account new observations/testimonies in the hybrid DSm fusion rule as soon as they become available to the fusion system.

If Θ (and therefore D^Θ) diminish (i.e. some of their elements are proven to be empty sets) from time t_l to time t_{l+1}, then one applies the hybrid DSm rule in order to transfer the masses of empty sets to the non-empty sets (in the way of the DSm classic rule) getting an updated basic belief assignment $m_{t_{l+1}|t_l}(.)$. Contrarily, if Θ and D^Θ increase (i.e. new elements arise in Θ, and/or new elements in D^Θ are proven different from the empty set and as a consequence a basic belief assignment for them is required), then new masses (from the same or from the other sources of information) are needed to describe these new elements, and again one combines them using the hybrid DSm rule.

4.6.3 Example 3

Let's consider a fusion problem at time t_l characterized by the frame $\Theta(t_l) \triangleq \{\theta_1, \theta_2\}$ and two independent sources of information providing the basic belief assignments $m_1(.)$ and $m_2(.)$ over $D^{\Theta(t_l)}$ and assume that at time t_{l+1} a new hypothesis θ_3 is introduced into the previous frame $\Theta(t_l)$ and a third source of evidence available at time t_{l+1} provides its own basic belief assignment $m_3(.)$ over $D^{\Theta(t_{l+1})}$ where

$$\Theta(t_{l+1}) \triangleq \{\Theta(t_l), \theta_3\} \equiv \{\theta_1, \theta_2, \theta_3\}$$

To solve such kind of dynamical fusion problems, we just use the classical DSm fusion rule as follows:

- combine $m_1(.)$ and $m_2(.)$ at time t_l using classical DSm fusion rule to get $m_{12}(.) = [m_1 \oplus m_2](.)$ over $D^{\Theta(t_l)}$
- because $D^{\Theta(t_l)} \subset D^{\Theta(t_{l+1})}$, $m_{12}(.)$ assigns the combined basic belief on a subset of $D^{\Theta(t_{l+1})}$, it is still directly possible to combine $m_{12}(.)$ with $m_3(.)$ at time t_{l+1} by the classical DSm fusion rule to get the final result $m_{123}(.)$ over $D^{\Theta(t_{l+1})}$ given by

$$m_{t_{l+1}}(.) \triangleq m_{123}(.) = [m_{12} \oplus m_3](.) = [(m_1 \oplus m_2) \oplus m_3](.) \equiv [m_1 \oplus m_2 \oplus m_3](.)$$

- eventually apply hybrid DSm rule if some integrity constraints have to be taken into account in the model $\mathcal{M}$ of the problem

This method can be directly generalized to any number of sources of evidences and, in theory, to any structure/dimension of the frames $\Theta(t_l)$, $\Theta(t_{l+1})$, ... In practice however, due to the huge number of elements of hyper-power sets, the dimension of the frames $\Theta(t_l)$, $\Theta(t_{l+1})$, $\ldots$ must be not too large. This practical limitation depends on the computer resources available for the real-time processing. Specific suboptimal implementations of DSm rule will have to be developed to deal with fusion problems of large dimension.

It is also important to point out here that DSmT can easily deal, not only with dynamical fusion problems but with decentralized fusion problems as well working on non exhaustive frames. For example, let's consider a set of two independent sources of information providing the basic belief assignments $m_1(.)$ and $m_2(.)$ over $D^{\Theta_{12}(t_l)=\{\theta_1,\theta_2\}}$ and another group of three independent sources of information providing the basic belief assignments $m_3(.)$, $m_4(.)$ and $m_5(.)$ over $D^{\Theta_{345}(t_l)=\{\theta_3,\theta_4,\theta_5,\theta_6\}}$, then it is still possible to combine all information in a decentralized manner as follows:

- combine $m_1(.)$ and $m_2(.)$ at time t_l using classical DSm fusion rule to get $m_{12}(.) = [m_1 \oplus m_2](.)$ over $D^{\Theta_{12}(t_l)}$.
- combine $m_3(.)$, $m_4(.)$ and $m_5(.)$ at time t_l using classical DSm fusion rule to get $m_{345}(.) = [m_3 \oplus m_4 \oplus m_5](.)$ over $D^{\Theta_{345}(t_l)}$.

- consider now the global frame $\Theta(t_l) \triangleq \{\Theta_{12}(t_l), \Theta_{345}(t_l)\}$.
- eventually apply hybrid DSm rule if some integrity constraints have to be taken into account in the model $\mathcal{M}$ of the problem.

Note that this *static* decentralized fusion can also be extended to decentralized dynamical fusion also by mixing the two previous approaches.

One can even combine all five masses together by extending the vectors $m_i(.)$, $1 \leq i \leq 5$, with null components for the new elements arisen from enlarging Θ to $\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ and correspondingly enlarging D^Θ, and using the hybrid DSm rule for $k = 5$. And more general combine the masses of any $k \geq 2$ sources.

We give now several simple numerical examples for such dynamical fusion problems involving non exclusive frames.

4.6.3.1 Example 3.1

Let's consider $\Theta(t_l) \triangleq \{\theta_1, \theta_2\}$ and the two following basic belief assignments available at time t_l:

$$m_1(\theta_1) = 0.1 \qquad m_1(\theta_2) = 0.2 \qquad m_1(\theta_1 \cup \theta_2) = 0.3 \qquad m_1(\theta_1 \cap \theta_2) = 0.4$$

$$m_2(\theta_1) = 0.5 \qquad m_2(\theta_2) = 0.3 \qquad m_2(\theta_1 \cup \theta_2) = 0.1 \qquad m_2(\theta_1 \cap \theta_2) = 0.1$$

The classical DSm rule of combination gives

$$m_{12}(\theta_1) = 0.21 \qquad m_{12}(\theta_2) = 0.17 \qquad m_{12}(\theta_1 \cup \theta_2) = 0.03 \qquad m_{12}(\theta_1 \cap \theta_2) = 0.59$$

Now let's consider at time t_{l+1} the frame $\Theta(t_{l+1}) \triangleq \{\theta_1, \theta_2, \theta_3\}$ and a third source of evidence with the following basic belief assignment

$$m_3(\theta_3) = 0.4 \qquad m_3(\theta_1 \cap \theta_3) = 0.3 \qquad m_3(\theta_2 \cup \theta_3) = 0.3$$

Then the final result of the fusion is obtained by combining $m_3(.)$ with $m_{12}(.)$ by the classical DSm rule of combination. One thus obtains:

$$m_{123}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.464 \quad m_{123}(\theta_2 \cap \theta_3) = 0.068 \quad m_{123}(\theta_1 \cap \theta_3) = 0.156 \quad m_{123}((\theta_1 \cup \theta_2) \cap \theta_3) = 0.012$$

$$m_{123}(\theta_1 \cap \theta_2) = 0.177 \quad m_{123}(\theta_1 \cap (\theta_2 \cup \theta_3)) = 0.063 \quad m_{123}(\theta_2) = 0.051 \quad m_{123}((\theta_1 \cap \theta_3) \cup \theta_2) = 0.009$$

4.6.3.2 Example 3.2

Let's consider $\Theta(t_l) \triangleq \{\theta_1, \theta_2\}$ and the two previous following basic belief assignments $m_1(.)$ and $m_2(.)$ available at time t_l. The classical DSm fusion rule gives gives as before

$$m_{12}(\theta_1) = 0.21 \qquad m_{12}(\theta_2) = 0.17 \qquad m_{12}(\theta_1 \cup \theta_2) = 0.03 \qquad m_{12}(\theta_1 \cap \theta_2) = 0.59$$

Now let's consider at time t_{l+1} the frame $\Theta(t_{l+1}) \triangleq \{\theta_1, \theta_2, \theta_3\}$ and the third source of evidence as in previous example with the basic belief assignment

$$m_3(\theta_3) = 0.4 \qquad m_3(\theta_1 \cap \theta_3) = 0.3 \qquad m_3(\theta_2 \cup \theta_3) = 0.3$$

The final result of the fusion obtained by the classical DSm rule of combination corresponds to the result of the previous example, but suppose now one finds out that the integrity constraint $\theta_3 = \emptyset$ holds, which implies also constraints $\theta_1 \cap \theta_2 \cap \theta_3 = \emptyset$, $\theta_1 \cap \theta_3 = \emptyset$, $\theta_2 \cap \theta_3 = \emptyset$ and $(\theta_1 \cup \theta_2) \cap \theta_3 = \emptyset$. This is the hybrid DSm model $\mathcal{M}$ under consideration here. We then have to readjust the mass $m_{123}(.)$ of the previous example by the hybrid DSm rule and one finally gets

$$m_{\mathcal{M}}(\theta_1) = 0.147 \qquad m_{\mathcal{M}}(\theta_2) = 0.060 + 0.119 = 0.179$$

$$m_{\mathcal{M}}(\theta_1 \cup \theta_2) = 0 + 0 + 0.021 = 0.021 \qquad m_{\mathcal{M}}(\theta_1 \cap \theta_2) = 0.240 + 0.413 = 0.653$$

Therefore, when we restrain back $\theta_3 = \emptyset$ and apply the hybrid DSm rule, we don't get back the same result (i.e. $m_{\mathcal{M}}(.) \neq m_{12}(.)$) because still remains some information from $m_3(.)$ on θ_1, θ_2, $\theta_1 \cup \theta_2$, or $\theta_1 \cap \theta_2$, i.e. $m_3(\theta_2) = 0.3 > 0$.

4.6.3.3 Example 3.3

Let's consider $\Theta(t_l) \triangleq \{\theta_1, \theta_2\}$ and two previous following basic belief assignments $m_1(.)$ and $m_2(.)$ available at time t_l. The classical DSm fusion rule gives as before

$$m_{12}(\theta_1) = 0.21 \qquad m_{12}(\theta_2) = 0.17 \qquad m_{12}(\theta_1 \cup \theta_2) = 0.03 \qquad m_{12}(\theta_1 \cap \theta_2) = 0.59$$

Now let's consider at time t_{l+1} the frame $\Theta(t_{l+1}) \triangleq \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and another third source of evidence with the following basic belief assignment

$$m_3(\theta_3) = 0.5 \qquad m_3(\theta_4) = 0.3 \qquad m_3(\theta_3 \cap \theta_4) = 0.1 \qquad m_3(\theta_3 \cup \theta_4) = 0.1$$

Then, the DSm rule applied at time t_{l+1} provides the following combined belief assignment

$$m_{123}(\theta_1 \cap \theta_3) = 0.105 \quad m_{123}(\theta_1 \cap \theta_4) = 0.063 \quad m_{123}(\theta_1 \cap (\theta_3 \cup \theta_4)) = 0.021 \quad m_{123}(\theta_1 \cap \theta_3 \cap \theta_4) = 0.021$$

$$m_{123}(\theta_2 \cap \theta_3) = 0.085 \quad m_{123}(\theta_2 \cap \theta_4) = 0.051 \quad m_{123}(\theta_2 \cap (\theta_3 \cup \theta_4)) = 0.017 \qquad m_{123}(\theta_2 \cap \theta_3 \cap \theta_4) = 0.017$$

$$m_{123}(\theta_3 \cap (\theta_1 \cup \theta_2)) = 0.015 \quad m_{123}(\theta_4 \cap (\theta_1 \cup \theta_2)) = 0.009 \quad m_{123}((\theta_1 \cup \theta_2) \cap (\theta_3 \cup \theta_4)) = 0.003$$

$$m_{123}((\theta_1 \cup \theta_2) \cap (\theta_3 \cap \theta_4)) = 0.003 \quad m_{123}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.295 \quad m_{123}(\theta_1 \cap \theta_2 \cap \theta_4) = 0.177$$

$$m_{123}((\theta_1 \cap \theta_2) \cap (\theta_3 \cup \theta_4)) = 0.059 \quad m_{123}(\theta_1 \cap \theta_2 \cap \theta_3 \cap \theta_4) = 0.059$$

Now suppose at time t_{l+2} one finds out that $\theta_3 = \theta_4 = \emptyset$, then one applies the hybrid DSm rule after re-adjusting the combined belief mass $m_{123}(.)$ by cumulating the masses of all empty sets. Using the hybrid DSm rule, one finally gets:

$$\begin{aligned} m_{t_{l+2}}(\theta_1) &= m_{123}(\theta_1) + \{m_{12}(\theta_1)m_3(\theta_3) + m_{12}(\theta_1)m_3(\theta_4) + m_{12}(\theta_1)m_3(\theta_3 \cup \theta_4) + m_{12}(\theta_1)m_3(\theta_3 \cap \theta_4)\} \\ &= 0 + \{(0.21 \times 0.5) + (0.21 \times 0.3) + (0.21 \times 0.1) + (0.21 \times 0.1)\} = 0.21 \end{aligned}$$

$$\begin{aligned} m_{t_{l+2}}(\theta_2) &= m_{123}(\theta_2) + \{m_{12}(\theta_2)m_3(\theta_3) + m_{12}(\theta_2)m_3(\theta_4) + m_{12}(\theta_2)m_3(\theta_3 \cup \theta_4) + m_{12}(\theta_2)m_3(\theta_3 \cap \theta_4)\} \\ &= 0 + \{(0.17 \times 0.5) + (0.17 \times 0.3) + (0.17 \times 0.1) + (0.17 \times 0.1)\} = 0.17 \end{aligned}$$

$$\begin{aligned} m_{t_{l+2}}(\theta_1 \cup \theta_2) &= m_{123}(\theta_1 \cup \theta_2) + \{m_{12}(\theta_1 \cup \theta_2)m_3(\theta_3) + m_{12}(\theta_1 \cup \theta_2)m_3(\theta_4) \\ &\quad + m_{12}(\theta_1 \cup \theta_2)m_3(\theta_3 \cup \theta_4) + m_{12}(\theta_1 \cup \theta_2)m_3(\theta_3 \cap \theta_4)\} \\ &\quad + \sum_{X_1, X_2 \in \{\theta_3, \theta_4, \theta_3 \cup \theta_4, \theta_3 \cap \theta_4\}} m_{12}(X_1)m_3(X_2) \\ &= 0 + \{(0.03 \times 0.5) + (0.03 \times 0.3) + (0.03 \times 0.1) + (0.03 \times 0.1)\} + \{0\} = 0.03 \end{aligned}$$

$$\begin{aligned} m_{t_{l+2}}(\theta_1 \cap \theta_2) &= m_{123}(\theta_1 \cap \theta_2) + \{m_{12}(\theta_1 \cap \theta_2)m_3(\theta_3) + m_{12}(\theta_1 \cap \theta_2)m_3(\theta_4) \\ &\quad + m_{12}(\theta_1 \cap \theta_2)m_3(\theta_3 \cup \theta_4) + m_{12}(\theta_1 \cap \theta_2)m_3(\theta_3 \cap \theta_4)\} \\ &= 0 + \{(0.59 \times 0.5) + (0.59 \times 0.3) + (0.59 \times 0.1) + (0.59 \times 0.1)\} = 0.59 \end{aligned}$$

Thus we get the same result as for $m_{12}(.)$ at time t_l, which is normal.

Remark: note that if the third source of information doesn't assign non-null masses to θ_1, or θ_2 (or to their combinations using $\cup$ or $\cap$ operators), then one obtains the same result at time t_{l+2} as at time t_l as in this example 3.3, i.e. $m_{l+2}(.) = m_l(.)$, when imposing back $\theta_3 = \theta_4 = \emptyset$. But, if the third source of information assigns non-null masses to either θ_1, or θ_2, or to some of their combinations $\theta_1 \cup \theta_2$ or $\theta_1 \cap \theta_2$, then when one returns from 4 singletons to 2 singletons for Θ, replacing $\theta_3 = \theta_4 = \emptyset$ and using the hybrid DSm rule, the fusion results at time t_{l+2} is different from that at time t_l, and this is normal because some information/mass is left from the third source and is now fusioned with that of the previous sources (as in example 3.2 or in the next example 3.4).

In general, let's suppose that the fusion of $k \geq 2$ masses provided by the sources $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_k$ has been done at time t_l on $\Theta(t_l) = \{\theta_1, \theta_2, ..., \theta_n\}$. At time t_{l+1} new non-empty elements θ_{n+1}, θ_{n+2}, $\ldots$, θ_{n+m} appear, $m \geq 1$, thus $\Theta(t_{l+1}) = \{\theta_1, \theta_2, ..., \theta_n, \theta_{n+1}, \theta_{n+2}, \ldots, \theta_{n+m}\}$ and of course one or more sources (i.e. bodies of evidences) $\mathcal{B}_{k+1}$, $\ldots$, $\mathcal{B}_{k+l}$, where $l \geq 1$, appear to assign masses to these new elements.

a) If all these new sources $\mathcal{B}_{k+1}, \ldots, \mathcal{B}_{k+l}$ assign null masses to all elements from $D^{\Theta(t_{l+1})}$ which contain in their structure/composition at least one of the singletons $\theta_1, \theta_2, \ldots, \theta_n$, then at time t_{l+2} if one sets back the constraints that $\theta_{n+1} = \theta_{n+2} = \ldots = \theta_{n+m} = \emptyset$, then using the hybrid DSm rule, one obtains the same result as at time t_l, i.e. $m_{l+2}(.) = m_l(.)$.

b) Otherwise, the fusion at time t_{l+2} will be different from the fusion at time t_l because there still remains some information/mass from sources $\mathcal{B}_{k+1}, \ldots, \mathcal{B}_{k+l}$ on singletons $\theta_1, \theta_2, \ldots, \theta_n$ or on some elements from $D^{\Theta(t_l)}$ which contain at least one such singleton, information/mass which fusions with the previous sources.

4.6.3.4 Example 3.4

Let's consider $\Theta(t_l) \triangleq \{\theta_1, \theta_2\}$ and the two following basic belief assignments available at time t_l:

$$m_1(\theta_1) = 0.6 \qquad m_1(\theta_2) = 0.4 \qquad \text{and} \qquad m_2(\theta_1) = 0.7 \qquad m_2(\theta_2) = 0.3$$

The classical DSm rule of combination gives $m_{12}(\theta_1) = 0.42$, $m_{12}(\theta_2) = 0.12$ and $m_{12}(\theta_1 \cap \theta_2) = 0.46$. Now let's consider at time t_{l+1} the frame $\Theta(t_{l+1}) \triangleq \{\theta_1, \theta_2, \theta_3\}$ and a third source of evidence with the following basic belief assignment $m_3(\theta_1) = 0.5$, $m_3(\theta_2) = 0.2$ and $m_3(\theta_3) = 0.3$. Then the final result obtained from the classical DSm rule of combination is still as before

$$m_{123}(\theta_1) = 0.210 \quad m_{123}(\theta_2) = 0.024 \quad m_{123}(\theta_1 \cap \theta_2) = 0.466 \quad m_{123}(\theta_1 \cap \theta_3) = 0.126$$
$$m_{123}(\theta_2 \cap \theta_3) = 0.036 \quad m_{123}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.138$$

Suppose now one finds out that the integrity constraint $\theta_1 \cap \theta_3 = \emptyset$ which also implies $\theta_1 \cap \theta_2 \cap \theta_3 = \emptyset$. This is the hybrid DSm model $\mathcal{M}$ under consideration. By applying the hybrid DSm fusion rule, one forces $m_{\mathcal{M}}(\theta_1 \cap \theta_3) = 0$ and $m_{\mathcal{M}}(\theta_1 \cap \theta_2 \cap \theta_3) = 0$ and we transfer $m_{123}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.138$ towards $m_{\mathcal{M}}((\theta_1 \cap \theta_2) \cup \theta_3)$ and the mass $m_{123}(\theta_1 \cap \theta_3) = 0.126$ has to be transferred towards $m_{\mathcal{M}}(\theta_1 \cup \theta_3)$. One then gets finally

$$m_{\mathcal{M}}(\theta_1) = 0.210 \quad m_{\mathcal{M}}(\theta_2) = 0.024 \quad m_{\mathcal{M}}(\theta_1 \cap \theta_2) = 0.466 \quad m_{\mathcal{M}}(\theta_2 \cap \theta_3) = 0.036$$
$$m_{\mathcal{M}}((\theta_1 \cap \theta_2) \cup \theta_3) = 0.138 \quad m_{\mathcal{M}}(\theta_1 \cup \theta_3) = 0.126$$

4.6.3.5 Example 3.5

Let's consider $\Theta(t_l) \triangleq \{\theta_1, \theta_2\}$ and the two previous basic belief assignments available at time t_l as in previous example, i.e.

$$m_1(\theta_1) = 0.6 \qquad m_1(\theta_2) = 0.4 \qquad \text{and} \qquad m_2(\theta_1) = 0.7 \qquad m_2(\theta_2) = 0.3$$

The classical DSm rule of combination gives

$$m_{12}(\theta_1) = 0.42 \qquad m_{12}(\theta_2) = 0.12 \quad m_{12}(\theta_1 \cap \theta_2) = 0.46$$

Now let's consider at time t_{l+1} the frame $\Theta(t_{l+1}) \triangleq \{\theta_1, \theta_2, \theta_3\}$ and a third source of evidence with the following basic belief assignment

$$m_3(\theta_1) = 0.5 \qquad m_3(\theta_2) = 0.2 \qquad m_3(\theta_3) = 0.3$$

Then the final result of the fusion is obtained by combining $m_3(.)$ with $m_{12}(.)$ by the classical DSm rule of combination. One thus obtains now

$$m_{123}(\theta_1) = 0.210 \quad m_{123}(\theta_2) = 0.024 \quad m_{123}(\theta_1 \cap \theta_2) = 0.466 \quad m_{123}(\theta_1 \cap \theta_3) = 0.126$$
$$m_{123}(\theta_2 \cap \theta_3) = 0.036 \quad m_{123}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.138$$

But suppose one finds out that the integrity constraint is now $\theta_3 = \emptyset$ which implies necessarily also $\theta_1 \cap \theta_3 = \theta_2 \cap \theta_3 = \theta_1 \cap \theta_2 \cap \theta_3 \equiv \emptyset$ and $(\theta_1 \cup \theta_2) \cap \theta_3 = \emptyset$ (this is our new hybrid DSm model $\mathcal{M}$ under consideration in this example). By applying the hybrid DSm fusion rule, one gets finally the non-null masses

$$m_{\mathcal{M}}(\theta_1) = 0.336 \qquad m_{\mathcal{M}}(\theta_2) = 0.060 \qquad m_{\mathcal{M}}(\theta_1 \cap \theta_2) = 0.604$$

4.6.3.6 Example 3.6

Let's consider $\Theta(t_l) \triangleq \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and the following basic belief assignments available at time t_l :

$$\begin{cases} m_1(\theta_1) = 0.5 \qquad m_1(\theta_2) = 0.4 \qquad m_1(\theta_1 \cap \theta_2) = 0.1 \\ m_2(\theta_1) = 0.3 \qquad m_2(\theta_2) = 0.2 \qquad m_2(\theta_1 \cap \theta_3) = 0.1 \qquad m_2(\theta_4) = 0.4 \end{cases}$$

The classical DSm rule of combination gives

$$m_{12}(\theta_1) = 0.15 \qquad m_{12}(\theta_2) = 0.08 \qquad m_{12}(\theta_1 \cap \theta_2) = 0.27 \qquad m_{12}(\theta_1 \cap \theta_3) = 0.05 \qquad m_{12}(\theta_1 \cap \theta_4) = 0.20$$
$$m_{12}(\theta_2 \cap \theta_4) = 0.16 \qquad m_{12}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.05 \qquad m_{12}(\theta_1 \cap \theta_2 \cap \theta_4) = 0.04$$

Now assume that at time t_{l+1} one finds out that $\theta_1 \cap \theta_2 \overset{\mathcal{M}}{\equiv} \theta_1 \cap \theta_3 \overset{\mathcal{M}}{\equiv} \emptyset$. Using the hybrid DSm rule, one gets:

$$\begin{cases}
m_{\mathcal{M}}(\theta_1 \cap \theta_2) = m_{\mathcal{M}}(\theta_1 \cap \theta_3) = m_{\mathcal{M}}(\theta_1 \cap \theta_2 \cap \theta_3) = m_{\mathcal{M}}(\theta_1 \cap \theta_2 \cap \theta_4) = 0 \\
m_{\mathcal{M}}(\theta_1) = m_{12}(\theta_1) + m_2(\theta_1)m_1(\theta_1 \cap \theta_2) + m_1(\theta_1)m_2(\theta_1 \cap \theta_3) = 0.15 + 0.03 + 0.05 = 0.23 \\
m_{\mathcal{M}}(\theta_2) = m_{12}(\theta_2) + m_2(\theta_2)m_1(\theta_1 \cap \theta_2) + m_1(\theta_2)m_2(\theta_1 \cap \theta_3) = 0.08 + 0.02 + 0.04 = 0.14 \\
m_{\mathcal{M}}(\theta_4) = m_{12}(\theta_4) + m_1(\theta_1 \cap \theta_2)m_2(\theta_4) = 0 + 0.04 = 0.04 \\
m_{\mathcal{M}}(\theta_1 \cap \theta_4) = m_{12}(\theta_1 \cap \theta_4) = 0.20 \\
m_{\mathcal{M}}(\theta_2 \cap \theta_4) = m_{12}(\theta_2 \cap \theta_4) = 0.16 \\
m_{\mathcal{M}}(\theta_1 \cup \theta_2) = m_{12}(\theta_1 \cup \theta_2) + m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2) + m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_2) = 0.22 \\
m_{\mathcal{M}}(\theta_1 \cup \theta_2 \cup \theta_3) = m_{12}(\theta_1 \cup \theta_2 \cup \theta_3) + m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_3) + m_2(\theta_1 \cap \theta_2)m_1(\theta_1 \cap \theta_3) \\
\qquad\qquad +m_1(\theta_1 \cap \theta_2 \cap \theta_3)m_2(\theta_1 \cap \theta_2 \cap \theta_3) = 0.01
\end{cases}$$

4.6.3.7 Example 3.7

Let's consider $\Theta(t_l) \triangleq \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and the following basic belief assignments available at time t_l :

$$\begin{cases} m_1(\theta_1) = 0.2 \quad m_1(\theta_2) = 0.4 \quad m_1(\theta_1 \cap \theta_2) = 0.1 \quad m_1(\theta_1 \cap \theta_3) = 0.2 \quad m_1(\theta_4) = 0.1 \\ m_2(\theta_1) = 0.1 \quad m_2(\theta_2) = 0.3 \quad m_2(\theta_1 \cap \theta_2) = 0.2 \quad m_2(\theta_1 \cap \theta_3) = 0.1 \quad m_2(\theta_4) = 0.3 \end{cases}$$

The classical DSm rule of combination gives

$$m_{12}(\theta_1) = 0.02 \quad m_{12}(\theta_2) = 0.12 \quad m_{12}(\theta_1 \cap \theta_2) = 0.28 \quad m_{12}(\theta_1 \cap \theta_3) = 0.06 \quad m_{12}(\theta_4) = 0.03$$

$$m_{12}(\theta_1 \cap \theta_4) = 0.07 \quad m_{12}(\theta_2 \cap \theta_4) = 0.15 \quad m_{12}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.15$$

$$m_{12}(\theta_1 \cap \theta_2 \cap \theta_4) = 0.05 \quad m_{12}(\theta_1 \cap \theta_3 \cap \theta_4) = 0.07$$

Now assume that at time t_{l+1} one finds out that $\theta_1 \cap \theta_2 \overset{\mathcal{M}}{\equiv} \theta_1 \cap \theta_3 \overset{\mathcal{M}}{\equiv} \emptyset$. Using the hybrid DSm rule, one gets:

$$\begin{cases} m_{\mathcal{M}}(\theta_1 \cap \theta_2) = m_{\mathcal{M}}(\theta_1 \cap \theta_3) = m_{\mathcal{M}}(\theta_1 \cap \theta_2 \cap \theta_3) = m_{\mathcal{M}}(\theta_1 \cap \theta_2 \cap \theta_4) = 0 \\ m_{\mathcal{M}}(\theta_1) = m_{12}(\theta_1) + m_1(\theta_1)m_2(\theta_1 \cap \theta_2) + m_2(\theta_1)m_1(\theta_1 \cap \theta_2) + m_1(\theta_1)m_2(\theta_1 \cap \theta_3) \\ \qquad\qquad + m_2(\theta_1)m_1(\theta_1 \cap \theta_3) = 0.11 \\ m_{\mathcal{M}}(\theta_2) = m_{12}(\theta_2) + m_1(\theta_2)m_2(\theta_1 \cap \theta_2) + m_2(\theta_2)m_1(\theta_1 \cap \theta_2) + m_1(\theta_2)m_2(\theta_1 \cap \theta_3) \\ \qquad\qquad + m_2(\theta_2)m_1(\theta_1 \cap \theta_3) = 0.33 \\ m_{\mathcal{M}}(\theta_4) = m_{12}(\theta_4) + m_1(\theta_4)m_2(\theta_1 \cap \theta_2) + m_2(\theta_4)m_1(\theta_1 \cap \theta_2) + m_1(\theta_4)m_2(\theta_1 \cap \theta_3) \\ \qquad\qquad + m_2(\theta_4)m_1(\theta_1 \cap \theta_3) = 0.15 \\ m_{\mathcal{M}}(\theta_1 \cap \theta_4) = m_{12}(\theta_1 \cap \theta_4) = 0.07 \\ m_{\mathcal{M}}(\theta_2 \cap \theta_4) = m_{12}(\theta_2 \cap \theta_4) = 0.15 \\ m_{\mathcal{M}}(\theta_1 \cup \theta_2) = m_{12}(\theta_1 \cup \theta_2) + m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_2) + m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2) = 0.12 \\ m_{\mathcal{M}}(\theta_1 \cup \theta_3) = m_{12}(\theta_1 \cup \theta_3) + m_1(\theta_1 \cap \theta_3)m_2(\theta_1 \cap \theta_3) = 0.02 \\ m_{\mathcal{M}}(\theta_1 \cup \theta_2 \cup \theta_3) = m_{12}(\theta_1 \cup \theta_2 \cup \theta_3) + m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_3) + m_2(\theta_1 \cap \theta_2)m_1(\theta_1 \cap \theta_3) = 0.05 \end{cases}$$

4.7 Bayesian mixture of hybrid DSm models

In the preceding, one has first shown how to combine generalized basic belief assignments provided by $k \geq 2$ independent and equally reliable sources of information with the general hybrid DSm rule of combination for dealing with all possible kinds of integrity constraints involved in a model. This approach implicitly assumes that one knows/trusts with certainty that the model $\mathcal{M}$ (usually a hybrid DSm model) of the problem is valid and corresponds to the true model. In some complex fusion problems however (static or dynamic ones), one may have some doubts about the validity of the model $\mathcal{M}$ on which is based the fusion because of the nature and evolution of elements of the frame Θ. In such situations, we propose to consider a set of exclusive and exhaustive models $\{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_K\}$ with some probabilities $\{P\{\mathcal{M}_1\}, P\{\mathcal{M}_2\}, \ldots, P\{\mathcal{M}_K\}\}$. We don't go here deeper on the justification/acquisition of such probabilities because this is highly dependent on the nature of the fusion problem under consideration. We just assume here that such probabilities are available at any given time t_l when the fusion has to be done. We propose then to use the Bayesian mixture of combined masses $m_{\mathcal{M}_i(\Theta)}(.)$ $i = 1, \ldots, K$ to obtain the final result :

$$\forall A \in D^{\Theta}, \qquad m_{\mathcal{M}_1,\ldots,\mathcal{M}_K}(A) = \sum_{i=1}^{K} P\{\mathcal{M}_i\} m_{\mathcal{M}_i(\Theta)}(A) \tag{4.14}$$

4.8 Conclusion

In this chapter we have extended the DSmT and the classical DSm rule of combination to the case of any kind of hybrid model for the frame Θ involved in many complex fusion problems. The free-DSm model (which assumes that none of the elements of the frame is refinable) can be interpreted as the opposite of Shafer's model (which assumes that all elements of the frame are truly exclusive) on which is based the mathematical theory of evidence (Dempster-Shafer Theory - DST). Between these two extreme models, there exists actually many possible hybrid models for the frames Θ depending on the real intrinsic nature of elements of the fusion problem under consideration. For real problems, some elements of Θ can appear to be truly exclusive whereas some others cannot be considered as fully discernable or refinable. This present research work proposes a new hybrid DSm rule of combination for hybrid models based on the DSmT. The hybrid DSm rule works in any model and is involved in calculation of mass fusion of any number of sources of information, no matter how big is the conflict/paradoxism of sources, and on any frame (exhaustive or non-exhaustive, with elements which may be exclusive or non-exclusive or both). This is an important rule since does not require the calculation of weighting factors, neither normalization as other rules do, and the transfer of masses of empty-sets to the masses of non-empty sets is naturally done following the DSm network architecture which is derived from the DSm classic rule. DSmT together with hybrid DSm rule is a new solid alternative to classical approaches and to existing combination rules. This new result is appealing for the development of future complex (uncertain/incomplete/paradoxical/dynamical) information fusion systems.

4.9 References

[1] Dawid A.P., *Conditional Independence*, 14th Conf. on Uncertainty and Artificial Intelligence, USA, 1998.

[2] Dawid A.P., *Conditional Independence*, In Encyclopedia of Statistical Science (Update) Volume 3, Wiley, New York, 1999.

[3] Dezert J., Smarandache F., *On the generation of hyper-power sets for the DSmT*, Proceedings of the 6th International Conference on Information Fusion, Cairns, Australia, July 8-11, 2003.

[4] Dubois D., Prade H., *Representation and combination of uncertainty with belief functions and possibility measures*, Computational Intelligence, Vol. 4, pp. 244-264, 1988.

[5] Fonck P., *Conditional Independence in Possibility Theory*, Uncertainty and Artificial Intelligence, pp. 221-226, 1994.

[6] Lefevre E., Colot O., Vannoorenberghe P. *Belief functions combination and conflict management*, Information Fusion Journal, Elsevier, 2002.

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

[8] Shenoy P., *Conditional Independence in Valuation-Based Systems*, International Journal of Approximate reasoning, VoL. 10, pp. 203-234, 1994.

[9] Smets Ph., Kennes R., *The transferable belief model*, Artificial Intelligence, 66(2), pp. 191-234, 1994.

[10] Studeny M., *Formal properties of Conditional Independence in Different Calculi of AI*, Proc. of ECSQARU'93, (Clarke K., Kruse R. and Moral S., Eds.), Springer-Verlag, 1993.

[11] Yager R.R., *On the Dempster-Shafer framework and new combination rules*, Information Sciences, Vol. 41, pp. 93-138, 1987..

[12] Yaghlane B.B., Smets Ph., Mellouli K., *Independence and Non-Interactivity in the Transferable Belief Model*, Workshop on Conditional Independence Structure and graphical Models, Eds. F. Matus and M. Studeny, Toronto, CA, 1999.

[13] Yaghlane B.B., Smets Ph., Mellouli K., *Belief Function Independence: I The marginal case*, International Journal of Approximate Reasoning, Vol. 29, pp. 47-70, 2002.

[14] Yaghlane B.B., Smets Ph., Mellouli K., *Belief Function Independence: II conditional case*, International Journal of Approximate Reasoning, Vol. 31, pp. 31-75, 2002.

Chapter 5

Counter-examples to Dempster's rule of combination

Jean Dezert	Florentin Smarandache	Mohammad Khoshnevisan
ONERA	Department of Mathematics	Griffith Business School
29 Av. de la Division Leclerc	University of New Mexico	Griffith University
92320 Châtillon	Gallup, NM 8730	Gold Coast Campus
France	U.S.A.	Australia

Abstract: *This chapter presents several classes of fusion problems which cannot be directly approached by the classical mathematical theory of evidence, also known as Dempster-Shafer Theory (DST), either because Shafer's model for the frame of discernment is impossible to obtain, or just because Dempster's rule of combination fails to provide coherent results (or no result at all). We present and discuss the potentiality of the DSmT combined with its classical (or hybrid) rule of combination to attack these infinite classes of fusion problems.*

5.1 Introduction

In this chapter we focus our attention on the limits of the validity of Dempster's rule of combination in Dempster-Shafer theory (DST) [5]. We provide several infinite classes of fusion problems where Dempster rule of combination fails to provide coherent results and we show how these problems can be attacked directly by the DSmT presented in previous chapters. DST and DSmT are based on a different approach for modelling the frame Θ of the problem (Shafer's model versus free-DSm, or hybrid-DSm model), on the choice of the space (classical power set 2^{Θ} versus hyper-power set D^{Θ}) on which will be defined the basic belief assignment functions $m_i(.)$ to be combined, and on the fusion rules to apply (Dempster rule versus DSm rule or hybrid DSm rule of combination).

5.2 First infinite class of counter examples

The first infinite class of counter examples for Dempster's rule of combination consists trivially in all cases for which Dempster's rule becomes mathematically not defined, i.e. one has 0/0, because of full conflicting sources. The first sub-class presented in subsection 5.2.1 corresponds to Bayesian belief functions. The subsection 5.2.2 will present counter-examples for more general conflicting sources of evidence.

5.2.1 Counter-examples for Bayesian sources

The following examples are devoted only to Bayesian sources, i.e. sources for which the focal elements of belief functions coincide only with some singletons θ_i of Θ.

5.2.1.1 Example with $\Theta = \{\theta_1, \theta_2\}$

Let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2\}$, two independent experts, and the basic belief masses:

$$m_1(\theta_1) = 1 \qquad m_1(\theta_2) = 0$$

$$m_2(\theta_1) = 0 \qquad m_2(\theta_2) = 1$$

We represent these belief assignments by the mass matrix

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- Dempster's rule can not be applied because one formally gets $m(\theta_1) = 0/0$ and $m(\theta_2) = 0/0$ as well, i.e. undefined.

- The DSm rule works here because one obtains $m(\theta_1) = m(\theta_2) = 0$ and $m(\theta_1 \cap \theta_2) = 1$ (the total paradox, which it really is! if one accepts the free-DSm model). If one adopts Shafer's model and applies the hybrid DSm rule, then one gets $m_h(\theta_1 \cup \theta_2) = 1$ which makes sense in this case. The index h denotes here the mass obtained with the hybrid DSm rule to avoid confusion with result obtained with the DSm classic rule.

5.2.1.2 Example with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

Let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, two independent experts, and the mass matrix

$$\begin{bmatrix} 0.6 & 0 & 0.4 & 0 \\ 0 & 0.2 & 0 & 0.8 \end{bmatrix}$$

- Again, Dempster's rule can not be applied because: $\forall 1 \leq j \leq 4$, one gets $m(\theta_j) = 0/0$ (undefined!).
- But the DSm rule works because one obtains: $m(\theta_1) = m(\theta_2) = m(\theta_3) = m(\theta_4) = 0$, and $m(\theta_1 \cap \theta_2) = 0.12$, $m(\theta_1 \cap \theta_4) = 0.48$, $m(\theta_2 \cap \theta_3) = 0.08$, $m(\theta_3 \cap \theta_4) = 0.32$ (partial paradoxes/conflicts).
- Suppose now one finds out that all intersections are empty (Shafer's model), then one applies the hybrid DSm rule and one gets (index h stands here for *hybrid* rule): $m_h(\theta_1 \cup \theta_2) = 0.12$, $m_h(\theta_1 \cup \theta_4) = 0.48$, $m_h(\theta_2 \cup \theta_3) = 0.08$ and $m_h(\theta_3 \cup \theta_4) = 0.32$.

5.2.1.3 Another example with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

Let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, three independent experts, and the mass matrix

$$\begin{bmatrix} 0.6 & 0 & 0.4 & 0 \\ 0 & 0.2 & 0 & 0.8 \\ 0 & 0.3 & 0 & 0.7 \end{bmatrix}$$

- Again, Dempster's rule can not be applied because: $\forall 1 \leq j \leq 4$, one gets $m(\theta_j) = 0/0$ (undefined!).
- But the DSm rule works because one obtains: $m(\theta_1) = m(\theta_2) = m(\theta_3) = m(\theta_4) = 0$, and

$$m(\theta_1 \cap \theta_2) = 0.6 \cdot 0.2 \cdot 0.3 = 0.036$$

$$m(\theta_1 \cap \theta_4) = 0.6 \cdot 0.8 \cdot 0.7 = 0.336$$

$$m(\theta_2 \cap \theta_3) = 0.4 \cdot 0.2 \cdot 0.3 = 0.024$$

$$m(\theta_3 \cap \theta_4) = 0.4 \cdot 0.8 \cdot 0.7 = 0.224$$

$$m(\theta_1 \cap \theta_2 \cap \theta_4) = 0.6 \cdot 0.2 \cdot 0.7 + 0.6 \cdot 0.3 \cdot 0.8 = 0.228$$

$$m(\theta_2 \cap \theta_3 \cap \theta_4) = 0.2 \cdot 0.4 \cdot 0.7 + 0.3 \cdot 0.4 \cdot 0.8 = 0.152$$

  (partial paradoxes/conflicts) and the others equal zero. If we add all these masses, we get the sum equals to 1.

- Suppose now one finds out that all intersections are empty (Shafer's model), then one applies the hybrid DSm rule and one gets: $m_h(\theta_1 \cup \theta_2) = 0.036$, $m_h(\theta_1 \cup \theta_4) = 0.336$, $m_h(\theta_2 \cup \theta_3) = 0.024$, $m_h(\theta_3 \cup \theta_4) = 0.224$, $m_h(\theta_1 \cup \theta_2 \cup \theta_4) = 0.228$, $m_h(\theta_2 \cup \theta_3 \cup \theta_4) = 0.152$.

5.2.1.4 More general

Let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, with $n \geq 2$, and k experts, for $k \geq 2$. Let $\mathbf{M} = [a_{ij}]$, $1 \leq i \leq k$, $1 \leq j \leq n$, be the mass matrix with k rows and n columns. If each column of the mass matrix contains at least a zero, then Dempster's rule can not be applied because one obtains for all $1 \leq j \leq n$, $m(\theta_j) = 0/0$ which is undefined! The degree of conflict is 1. However, one can use the classical DSm rule and one obtains: for all $1 \leq j \leq n$, $m(\theta_j) = 0$, and also partial paradoxes/conflicts: $\forall 1 \leq v_s \leq n$, $1 \leq s \leq w$, and $2 \leq w \leq k$, $m(\theta_{v_1} \cap \theta_{v_2} \cap \ldots \cap \theta_{v_w}) = \sum (a_{1t_1}) \cdot (a_{2t_2}) \cdot \ldots \cdot (a_{kt_k})$, where the set $T = \{t_1, t_2, \ldots, t_k\}$ is equal to the set $V = \{v_1, v_2, \ldots, v_w\}$ but the order may be different and the elements in the set T could be repeated; we mean from set V one obtains set T if one repeats some elements of V; therefore: summation $\sum$ is done upon all possible combinations of elements from columns $v_1, v_2, \ldots, v_w$ such that at least one element one takes from each of these columns $v_1, v_2, \ldots, v_w$ and also such that from each row one takes one element only; the product $(a_{1t_1}) \cdot (a_{2t_2}) \cdot \ldots \cdot (a_{kt_k})$ contains one element only from each row 1, 2, ..., k respectively, and one or more elements from each of the columns $v_1, v_2, \ldots, v_w$ respectively.

5.2.2 Counter-examples for more general sources

We present in this section two numerical examples involving general (i.e. non Bayesian) sources where Dempster's rule cannot be applied.

5.2.2.1 Example with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

Let's consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, two independent experts, and the mass matrix:

	θ_1	θ_2	θ_3	θ_4	$\theta_1 \cup \theta_2$
$m_1(.)$	0.4	0.5	0	0	0.1
$m_2(.)$	0	0	0.3	0.7	0

Dempster's rule cannot apply here because one gets 0/0 for all $m(\theta_i)$, $1 \leq i \leq 4$, but the DSm rules (classical or hybrid) work.

Using the DSm classical rule: $m(\theta_1 \cap \theta_3) = 0.12$, $m(\theta_1 \cap \theta_4) = 0.28$, $m(\theta_2 \cap \theta_3) = 0.15$, $m(\theta_2 \cap \theta_4) = 0.35$, $m(\theta_3 \cap (\theta_1 \cup \theta_2)) = 0.03$, $m(\theta_4 \cap (\theta_1 \cup \theta_2)) = 0.07$.

Suppose now one finds out that one has a Shafer model; then one uses the hybrid DSm rule (denoted here with index h): $m_h(\theta_1 \cup \theta_3) = 0.12$, $m_h(\theta_1 \cup \theta_4) = 0.28$, $m_h(\theta_2 \cup \theta_3) = 0.15$, $m_h(\theta_2 \cup \theta_4) = 0.35$, $m_h(\theta_3 \cup \theta_1 \cup \theta_2) = 0.03$, $m_h(\theta_4 \cup \theta_1 \cup \theta_2) = 0.07$.

5.2.2.2 Another example with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

Let's consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, three independent experts, and the mass matrix:

	θ_1	θ_2	θ_3	θ_4	$\theta_1 \cup \theta_2$	$\theta_3 \cup \theta_4$
$m_1(.)$	0.4	0.5	0	0	0.1	0
$m_2(.)$	0	0	0.3	0.6	0	0.1
$m_3(.)$	0.8	0	0	0	0.2	0

Dempster's rule cannot apply here because one gets 0/0 for all $m(\theta_i)$, $1 \leq i \leq 4$, but the DSm rules (classical or hybrid) work.

Using the DSm classical rule, one gets:

$$m(\theta_1) = m(\theta_2) = m(\theta_3) = m(\theta_4) = 0$$
$$m(\theta_1 \cap \theta_3) = 0.096$$
$$m(\theta_1 \cap \theta_4) = 0.192$$
$$m(\theta_1 \cap (\theta_3 \cup \theta_4)) = 0.032$$
$$m(\theta_2 \cap \theta_3 \cap \theta_1) = 0.120$$
$$m(\theta_2 \cap \theta_4 \cap \theta_1) = 0.240$$
$$m(\theta_2 \cap (\theta_3 \cup \theta_4) \cap \theta_1) = m((\theta_1 \cap \theta_2) \cap (\theta_3 \cup \theta_4)) = 0.040$$
$$m((\theta_1 \cup \theta_2) \cap \theta_3 \cap \theta_1) = m(\theta_1 \cap \theta_3) = 0.024$$
$$m((\theta_1 \cup \theta_2) \cap \theta_4 \cap \theta_1) = m(\theta_1 \cap \theta_4) = 0.048$$
$$m((\theta_1 \cup \theta_2) \cap (\theta_3 \cup \theta_4) \cap \theta_1) = m(\theta_1 \cap (\theta_3 \cup \theta_4)) = 0.008$$
$$m(\theta_1 \cup \theta_2) = m(\theta_3 \cup \theta_4) = 0$$
$$m(\theta_1 \cap \theta_3 \cap (\theta_1 \cup \theta_2)) = m(\theta_1 \cap \theta_3) = 0.024$$
$$m(\theta_1 \cap \theta_4 \cap (\theta_1 \cup \theta_2)) = m(\theta_1 \cap \theta_4) = 0.048$$
$$m(\theta_1 \cap (\theta_3 \cup \theta_4) \cap (\theta_1 \cup \theta_2)) = m(\theta_1 \cap (\theta_3 \cup \theta_4)) = 0.008$$
$$m(\theta_2 \cap \theta_3 \cap (\theta_1 \cup \theta_2)) = m(\theta_2 \cap \theta_3) = 0.030$$
$$m(\theta_2 \cap \theta_4 \cap (\theta_1 \cup \theta_2)) = m(\theta_2 \cap \theta_4) = 0.060$$
$$m(\theta_2 \cap (\theta_3 \cup \theta_4) \cap (\theta_1 \cup \theta_2)) = m(\theta_2 \cap (\theta_3 \cup \theta_4)) = 0.010$$
$$m((\theta_1 \cup \theta_2) \cap \theta_3) = 0.006$$
$$m((\theta_1 \cup \theta_2) \cap \theta_4) = 0.012$$
$$m((\theta_1 \cup \theta_2) \cap (\theta_3 \cup \theta_4)) = 0.002$$

After cumulating, one finally gets with DSm classic rule:

$$m(\theta_1 \cap \theta_3) = 0.096 + 0.024 + 0.024 = 0.144$$
$$m(\theta_1 \cap \theta_4) = 0.192 + 0.048 + 0.048 = 0.288$$
$$m(\theta_2 \cap \theta_3) = 0.030$$
$$m(\theta_2 \cap \theta_4) = 0.060$$
$$m(\theta_1 \cap \theta_2 \cap \theta_3) = 0.120$$
$$m(\theta_1 \cap \theta_2 \cap \theta_4) = 0.240$$
$$m((\theta_1 \cup \theta_2) \cap \theta_3) = 0.006$$
$$m((\theta_1 \cup \theta_2) \cap \theta_4) = 0.012$$
$$m(\theta_1 \cap (\theta_3 \cup \theta_4)) = 0.032 + 0.008 + 0.008 = 0.048$$
$$m(\theta_1 \cap \theta_2 \cap (\theta_3 \cup \theta_4)) = 0.040$$
$$m(\theta_2 \cap (\theta_3 \cup \theta_4)) = 0.010$$
$$m((\theta_1 \cup \theta_2) \cap (\theta_3 \cup \theta_4)) = 0.002$$

Suppose now, one finds out that all intersections are empty. Using the hybrid DSm rule one gets:

$$m_h(\theta_1 \cup \theta_3) = 0.144$$
$$m_h(\theta_1 \cup \theta_4) = 0.288$$
$$m_h(\theta_2 \cup \theta_3) = 0.030$$
$$m_h(\theta_2 \cup \theta_4) = 0.060$$
$$m_h(\theta_1 \cup \theta_2 \cup \theta_3) = 0.120 + 0.006 = 0.126$$
$$m_h(\theta_1 \cup \theta_2 \cup \theta_4) = 0.240 + 0.012 = 0.252$$
$$m_h(\theta_1 \cup \theta_3 \cup \theta_4) = 0.048$$
$$m_h(\theta_2 \cup \theta_3 \cup \theta_4) = 0.010$$
$$m_h(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = 0.040 + 0.002 = 0.042$$

5.2.2.3 More general

Let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, with $n \geq 2$, and k experts, for $k \geq 2$, and the mass matrix $\mathbf{M}$ with k rows and $n+u$ columns, where $u \geq 1$, corresponding to θ_1, θ_2, $\ldots$, θ_n, and u uncertainties $\theta_{i_1} \cup \ldots \cup \theta_{i_s}$, $\ldots$, $\theta_{j_1} \cup \ldots \cup \theta_{j_t}$ respectively.

If the following conditions occur:

- each column contains at least one zero;
- all uncertainties are different from the total ignorance $\theta_1 \cup \ldots \cup \theta_n$ (i.e., they are partial ignorances);
- the partial uncertainties are disjoint two by two;
- for each non-null uncertainty column c_j, $n+1 \leq j \leq n+u$, of the form say $\theta_{p_1} \cup \ldots \cup \theta_{p_w}$, there exists a row such that all its elements on columns p_1, $\ldots$, p_w, and c_j are zero.

then Dempster's rule of combination cannot apply for such infinite class of fusion problems because one gets $0/0$ for all $m(\theta_i)$, $1 \leq i \leq n$. The DSm rules (classical or hybrid) work for such infinite class of examples.

5.3 Second infinite class of counter examples

This second class of counter-examples generalizes the famous Zadeh example given in [7, 8].

5.3.1 Zadeh's example

Two doctors examine a patient and agree that it suffers from either meningitis (M), contusion (C) or brain tumor (T). Thus $\Theta = \{M, C, T\}$. Assume that the doctors agree in their low expectation of a tumor, but disagree in likely cause and provide the following diagnosis

$$m_1(M) = 0.99 \qquad m_1(T) = 0.01 \qquad \text{and} \qquad m_2(C) = 0.99 \qquad m_2(T) = 0.01$$

If we combine the two basic belief functions using Dempster's rule of combination, one gets the unexpected final conclusion

$$m(T) = \frac{0.0001}{1 - 0.0099 - 0.0099 - 0.9801} = 1$$

which means that the patient suffers with certainty from brain tumor !!!. This unexpected result arises from the fact that the two bodies of evidence (doctors) agree that the patient most likely does not suffer from tumor but are in almost full contradiction for the other causes of the disease. This very simple but interesting example shows the limitations of the practical use of the DST for automated reasoning.

This example has been examined in literature by several authors to explain the anomaly of the result of Dempster's rule of combination in such case. Due to the high degree of conflict arising in such extreme case, willingly pointed out by Zadeh to show the weakness of this rule, it is often argued that in such case the result of Dempster's rule must not be taken directly without checking the level of the conflict between sources of evidence. This is trivially true but there is no theoretical way to decide *beforehand* if one can trust or not the result of such rule of combination, especially in complex systems involving many sources and many hypotheses. This is one of its major drawback. The issue consists generally in choosing rather somewhat arbitrarily or heuristically some threshold value on the degree of conflict between sources to accept or reject the result of the fusion [9]. Such approach can't be solidly justified from theoretical analysis. Assuming such threshold is set to a given value, say 0.70 for instance, is it acceptable to reject the fusion result if the conflict appears to be 0.7001 and accept it when the conflict becomes 0.6999? What to do when the decision about the fusion result is rejected and one has no assessment on the reliability of the sources or when the sources have the same reliability/confidence but an important decision has to be taken anyway? There is no theoretical solid justification which can reasonably support such kind of approaches commonly used in practice up to now.

The two major explanations of this problem found in literature are mainly based, either on the fact that problem arises from the closed-world assumption of Shafer's model Θ and it is suggested to work rather with an open-world model, and/or the fact that sources of evidence are not reliable. These explanations although being admissible are not necessarily the only correct (sufficient) explanations. Note that the open-world assumption can always be easily relaxed advantageously by introducing a new hypothesis, say θ_0 in the initial frame $\Theta = \{\theta_1, \ldots, \theta_n\}$ in order to close it. θ_0 will then represent all possible alternatives (although remaining unknown) of initial hypotheses $\theta_1, \ldots \theta_n$. This idea has been already proposed by Yager in [6] through his hedging solution. Upon our analysis, it is not necessary to adopt/follow the open-world model neither to admit the assumption about the reliability of the sources to find a justification in this counter-intuitive result. Actually, both sources can have the same reliability and Shafer's model can be accepted for the combination of the two reports by using another rule of combination. This is exactly the purpose of the hybrid DSm rule of combination. Of course when one has some prior information on the reliability of sources, one has to take them into account properly by some discounting methods. The discounting techniques can also apply in the DSmT framework and there is no incompatibility to mix both (i.e. discounting techniques with DSm rules of combinations) when necessary (when there is strong reason to justify doing it, i.e. when one has prior reliable information on reliability of the sources). The discounting techniques must never been used as an artificial ad-hoc mechanism to update Dempster's result once problem has arisen. We strongly disagree with the idea that all problems with Dempster's rule can be solved *beforehand* by discounting techniques. This can help obviously to improve the assessment of belief function to be combined when used properly and fairly, but this does not fundamentally solve the inherent problem of Dempster's rule itself when conflict remains high.

The problem comes from the fact that both sources provide essentially their belief with respect only to their own limited knowledge and experience. It is also possible in some cases, that sources of information even don't have the same interpretation of concepts included in the frame of the problem. Such kind of situation frequently appears for example in debates on TV, on radio or in most of the meetings where important decision/approval have to be drawn and when the sources don't share the same opinion. This is what happens daily in real life and one has to deal with such conflicting situations anyway. In other words, the sources do not speak about the same events or even they do, they there is a possibility that they do not share the same interpretation of the events. This has already been pointed out by Dubois and Prade in [3] (p. 256). In Zadeh's controversy example, it is possible that the first doctor is expert mainly in meningitis and in brain tumor while the second doctor is expert mainly in cerebral contusion and in brain tumor. Because of their limited knowledges and experiences, both doctors can also have also the same reliability. If they have been asked to give their reports only on $\Theta = \{M, C, T\}$ (but not on an extended frame), their reports have to be taken with same weight and the combination has to be done anyway when one has no solid reason to reject one report with respect to the other one; the result of the Demsper's rule still remains very questionable. No rational brain surgeon would take the decision for a brain intervention (i.e. a risky tumor ablation) based on Dempster's rule result, neither the family of the patient. Therefore upon our analysis, the two previous explanations given in literature (although being possible and admissible in some cases) are not necessary and sufficient to explain the source of the anomaly. Several alternatives to Dempster's rule to circumvent this anomaly have been proposed in literature mainly through the works of R. Yager [6], D. Dubois and H. Prade [2] already reported in chapter 1 or by Daniel in [1]. The DSmT offers just a new issue for solving also such controversy example as it will be shown. In summary, some extreme caution on the degree of conflict of the sources must always be taken before taking a final decision based on Dempster's rule of combination, especially when vital wagers are involved.

If we now adopt the free-DSm model, i.e. we replace the initial Shafer model by accepting the possibility of non null intersections between hypotheses M, C and T and by working directly on hyper-power set D^Θ then one gets directly and easily the following result with the classical DSm rule of combination:

$$m(M \cap C) = 0.9801 \qquad m(M \cap T) = 0.0099 \qquad m(C \cap T) = 0.0099 \qquad m(T) = 0.0001$$

which makes sense when working with such a new model. Obviously same result can be obtained (the proof is left here to the reader) when working with Dempster's rule based on the following refined frame Θ_{ref} defined with basic belief functions on power set $2^{\Theta_{ref}}$:

$$\Theta_{ref} = \{\theta_1 = M \cap C \cap T, \theta_2 = M \cap C \cap \bar{T}, \theta_3 = M \cap \bar{C} \cap T, \theta_4 = \bar{M} \cap C \cap T,$$
$$\theta_5 = M \cap \bar{C} \cap \bar{T}, \theta_6 = \bar{M} \cap C \cap \bar{T}, \theta_7 = \bar{M} \cap \bar{C} \cap T\}$$

where $\bar{T}$,$\bar{C}$ and $\bar{M}$ denote respectively the complement of T, C and M.

The equality of both results (i.e. by the classical DSm rule based on the free-DSm model and by Dempster's rule based on the refined frame) is just normal since the normalization factor $1-k$ of Dempster's rule in this case reduces to 1 because of the new choice of the new model. Based on this remark, one could then try to argue that DSmT (together with its DSm classical rule for free-DSm model) is superfluous. Such claim is obviously wrong for the two following reasons: it is unecessary to work with a bigger space (keeping in mind that $|D^\Theta| < |2^{\Theta_{ref}}|$) to get the result (the DSm rule offers just a direct and more convenient issue to get the result), but also because in some fusion problems involving vague/-continuous concepts, the refinement is just impossible to obtain and we are unfortunately forced to deal with ambiguous concepts/hypotheses (see [4] for details and justification).

If one has no doubt on the reliability of both Doctors (or no way to assess it) and if one is absolutely sure that the true origin of the suffering of the patient lies only in the frame $\Theta = \{M, C, T\}$ and we consider these origins as truly exclusive, then one has to work with the initial frame of discernment Θ satisfying Shafer's model. As previously shown, Dempster's rule fails to provide a reasonable and acceptable conclusion in such high conflicting case. However, this case can be easily handled by the hybrid DSm rule of combination. The hybrid DSm rule applies now because Shafer's model is nothing but a particular hybrid model including all exclusivity constraints between hypotheses of the frame Θ (see chapter 4 for details). One then gets with the hybrid DSm rule for this simple case (more general and complex examples have been already presented in chapter 4), after the proper mass transfer of all sources of the conflicts:

$$m(M \cup C) = 0.9801 \qquad m(M \cup T) = 0.0099 \qquad m(C \cup T) = 0.0099 \qquad m(T) = 0.0001$$

This result is not surprising and makes perfectly sense with common intuition actually since it provides a coherent and reasonable solution to the problem. It shows clearly that a brain intervention for ablation of an hypothetical tumor is not recommended, but preferentially a better examination of the patient focused on Meningitis or Contusion as possible source of the suffering. The consequence of the results of Dempster's rule and the hybrid DSm rule is therefore totally different.

5.3.2 Generalization with $\Theta = \{\theta_1, \theta_2, \theta_3\}$

Let's consider $0 < \epsilon_1, \epsilon_2 < 1$ be two very tiny positive numbers (close to zero), the frame of discernment be $\Theta = \{\theta_1, \theta_2, \theta_3\}$, have two experts (independent sources of evidence s_1 and s_2) giving the belief masses

$$m_1(\theta_1) = 1 - \epsilon_1 \quad m_1(\theta_2) = 0 \quad m_1(\theta_3) = \epsilon_1$$

$$m_2(\theta_1) = 0 \quad m_2(\theta_2) = 1 - \epsilon_2 \quad m_2(\theta_3) = \epsilon_2$$

From now on, we prefer to use matrices to describe the masses, i.e.

$$\begin{bmatrix} 1-\epsilon_1 & 0 & \epsilon_1 \\ 0 & 1-\epsilon_2 & \epsilon_2 \end{bmatrix}$$

- Using Dempster's rule of combination, one gets

  $$m(\theta_3) = \frac{(\epsilon_1\epsilon_2)}{(1-\epsilon_1)\cdot 0 + 0\cdot(1-\epsilon_2) + \epsilon_1\epsilon_2} = 1$$

  which is absurd (or at least counter-intuitive). Note that whatever positive values for ϵ_1, ϵ_2 are, Dempster's rule of combination provides always the same result (one) which is abnormal. The only acceptable and correct result obtained by Dempster's rule is really obtained only in the trivial case when $\epsilon_1 = \epsilon_2 = 1$, i.e. when both sources agree in θ_3 with certainty which is obvious.

- Using the DSm rule of combination based on free-DSm model, one gets $m(\theta_3) = \epsilon_1\epsilon_2$, $m(\theta_1 \cap \theta_2) = (1-\epsilon_1)(1-\epsilon_2)$, $m(\theta_1 \cap \theta_3) = (1-\epsilon_1)\epsilon_2$, $m(\theta_2 \cap \theta_3) = (1-\epsilon_2)\epsilon_1$ and the others are zero which appears more reliable/trustable.

- Going back to Shafer's model and using the hybrid DSm rule of combination, one gets $m(\theta_3) = \epsilon_1\epsilon_2$, $m(\theta_1 \cup \theta_2) = (1-\epsilon_1)(1-\epsilon_2)$, $m(\theta_1 \cup \theta_3) = (1-\epsilon_1)\epsilon_2$, $m(\theta_2 \cup \theta_3) = (1-\epsilon_2)\epsilon_1$ and the others are zero.

Note that in the special case when $\epsilon_1 = \epsilon_2 = 1/2$, one has

$$m_1(\theta_1) = 1/2 \quad m_1(\theta_2) = 0 \quad m_1(\theta_3) = 1/2 \qquad \text{and} \qquad m_2(\theta_1) = 0 \quad m_2(\theta_2) = 1/2 \quad m_2(\theta_3) = 1/2$$

Dempster's rule of combinations still yields $m(\theta_3) = 1$ while the hybrid DSm rule based on the same Shafer's model yields now $m(\theta_3) = 1/4$, $m(\theta_1 \cup \theta_2) = 1/4$, $m(\theta_1 \cup \theta_3) = 1/4$, $m(\theta_2 \cup \theta_3) = 1/4$ which is normal.

5.3.3 Generalization with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

Let's consider $0 < \epsilon_1, \epsilon_2, \epsilon_3 < 1$ be three very tiny positive numbers, the frame of discernment be $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, have two experts giving the mass matrix

$$\begin{bmatrix} 1-\epsilon_1-\epsilon_2 & 0 & \epsilon_1 & \epsilon_2 \\ 0 & 1-\epsilon_3 & 0 & \epsilon_3 \end{bmatrix}$$

Again using Dempster's rule of combination, one gets $m(\theta_4) = 1$ which is absurd while using the DSm rule of combination based on free-DSm model, one gets $m(\theta_4) = \epsilon_2\epsilon_3$ which is reliable. Using the DSm classical rule: $m(\theta_1 \cap \theta_2) = (1-\epsilon_1-\epsilon_2)(1-\epsilon_3)$, $m(\theta_1 \cap \theta_4) = (1-\epsilon_1-\epsilon_3)\epsilon_3$, $m(\theta_3 \cap \theta_2) = \epsilon_1(1-\epsilon_3)$, $m(\theta_3 \cap \theta_4) = \epsilon_1\epsilon_3$, $m(\theta_4) = \epsilon_2\epsilon_3$. Suppose one finds out that all intersections are empty, then one applies the hybrid DSm rule: $m_h(\theta_1 \cup \theta_2) = (1-\epsilon_1-\epsilon_2)(1-\epsilon_3)$, $m_h(\theta_1 \cup \theta_4) = (1-\epsilon_1-\epsilon_3)\epsilon_3$, $m_h(\theta_3 \cup \theta_2) = \epsilon_1(1-\epsilon_3)$, $m_h(\theta_3 \cup \theta_4) = \epsilon_1\epsilon_3$, $m_h(\theta_4) = \epsilon_2\epsilon_3$.

5.3.4 More general

Let's consider $0 < \epsilon_1, \ldots, \epsilon_n < 1$ be very tiny positive numbers, the frame of discernment be $\Theta = \{\theta_1, \ldots, \theta_n, \theta_{n+1}\}$, have two experts giving the mass matrix

$$\begin{bmatrix} 1-S_1^p & 0 & \epsilon_1 & 0 & \epsilon_2 & \ldots & 0 & \epsilon_p \\ 0 & 1-S_{p+1}^n & 0 & \epsilon_{p+1} & 0 & \ldots & \epsilon_{n-1} & \epsilon_n \end{bmatrix}$$

where $1 \leq p \leq n$ and $S_1^p \triangleq \sum_{i=1}^p \epsilon_i$ and $S_{p+1}^n \triangleq \sum_{i=p+1}^n \epsilon_i$. Again using Dempster's rule of combination, one gets $m(\theta_{n+1}) = 1$ which is absurd while using the DSm rule of combination based on free-DSm model, one gets $m(\theta_{n+1}) = \epsilon_p\epsilon_n$ which is reliable. This example is similar to the previous one, but generalized.

5.3.5 Even more general

Let's consider $0 < \epsilon_1, \ldots, \epsilon_n < 1$ be very tiny positive numbers (close to zero), the frame of discernment be $\Theta = \{\theta_1, \ldots, \theta_n, \theta_{n+1}\}$, have $k \geq 2$ experts giving the mass matrix of k rows and $n+1$ columns such that:

- one column, say column j, is $(\epsilon_{j_1}, \epsilon_{j_2}, \ldots, \epsilon_{j_k})'$ (transposed vector), where $1 \leq j \leq n+1$ where $\{\epsilon_{j_1}, \epsilon_{j_2}, \ldots, \epsilon_{j_k}\}$ is included in $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_n\}$;
- and each column (except column j) contains at least one element equals to zero.

Then Dempster's rule of combination gives $m(\theta_j) = 1$ which is absurd, while the classical DSm rule gives $m(\theta_j) = \epsilon_{j_1} \cdot \epsilon_{j_2} \cdot \ldots \cdot \epsilon_{j_k} \neq 0$ which is reliable.

Actually, we need to set restrictions only for ϵ_{j_1}, ϵ_{j_2}, ..., and ϵ_{j_k} to be very tiny positive numbers, not for all ϵ_1, ϵ_2, ..., ϵ_n (the others can be anything in the interval $[0, 1)$ such that the sum of elements on each row be equal 1).

5.4 Third infinite class of counter examples

This third class of counter-examples deals with belief functions committing a non null mass to some uncertainties.

5.4.1 Example with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

Let's consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, two independent experts, and the mass matrix:

	θ_1	θ_2	θ_3	θ_4	$\theta_3 \cup \theta_4$
$m_1(.)$	0.99	0	0	0	0.01
$m_2(.)$	0	0.98	0	0	0.02

If one applies Dempster's rule, one gets

$$m(\theta_3 \cup \theta_4) = \frac{(0.01 \cdot 0.02)}{(0 + 0 + 0 + 0 + 0.01 \cdot 0.02)} = 1$$

(total ignorance), which doesn't bring any information to the fusion. This example looks similar to Zadeh's example, but is different because it is referring to uncertainty (not to contradictory) result. Using the DSm classical rule: $m(\theta_1 \cap \theta_2) = 0.9702$, $m(\theta_1 \cap (\theta_3 \cup \theta_4)) = 0.0198$, $m(\theta_2 \cap (\theta_3 \cup \theta_4)) = 0.0098$, $m(\theta_3 \cup \theta_4) = 0.0002$. Suppose now one finds out that all intersections are empty (i.e. one adopts Shafer's model). Using the hybrid DSm rule one gets: $m_h(\theta_1 \cup \theta_2) = 0.9702$, $m_h(\theta_1 \cup \theta_3 \cup \theta_4) = 0.0198$, $m_h(\theta_2 \cup \theta_3 \cup \theta_4) = 0.0098$, $m_h(\theta_3 \cup \theta_4) = 0.0002$.

5.4.2 Example with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$

Let's consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, , \theta_5\}$, three independent experts, and the mass matrix:

	θ_1	θ_2	θ_3	θ_4	θ_5	$\theta_4 \cup \theta_5$
$m_1(.)$	0.99	0	0	0	0	0.01
$m_2(.)$	0	0.98	0.01	0	0	0.01
$m_3(.)$	0.01	0.01	0.97	0	0	0.01

- If one applies Dempster's rule, one gets

$$m(\theta_4 \cup \theta_5) = \frac{(0.01 \cdot 0.01 \cdot 0.01)}{(0 + 0 + 0 + 0 + 0.01 \cdot 0.01 \cdot 0.01)} = 1$$

(total ignorance), which doesn't bring any information to the fusion.

- Using the DSm classical rule one gets:

$$m(\theta_1 \cap \theta_2) = 0.99 \cdot 0.98 \cdot 0.01 + 0.99 \cdot 0.98 \cdot 0.01 = 0.019404$$

$$m(\theta_1 \cap \theta_3) = 0.99 \cdot 0.01 \cdot 0.01 + 0.99 \cdot 0.01 \cdot 0.97 = 0.009702$$

$$m(\theta_1 \cap \theta_2 \cap \theta_3) = 0.99 \cdot 0.98 \cdot 0.97 + 0.99 \cdot 0.01 \cdot 0.01 = 0.941193$$

$$m(\theta_1 \cap \theta_3 \cap (\theta_4 \cup \theta_5)) = 0.99 \cdot 0.01 \cdot 0.01 + 0.99 \cdot 0.01 \cdot 0.97 + 0.01 \cdot 0.01 \cdot 0.01 = 0.009703$$

$$m(\theta_1 \cap (\theta_4 \cup \theta_5)) = 0.99 \cdot 0.01 \cdot 0.01 + 0.99 \cdot 0.01 \cdot 0.01 + 0.01 \cdot 0.01 \cdot 0.01 = 0.000199$$

$$m((\theta_4 \cup \theta_5) \cap \theta_2 \cap \theta_1) = 0.01 \cdot 0.98 \cdot 0.01 + 0.99 \cdot 0.01 \cdot 0.01 + 0.99 \cdot 0.98 \cdot 0.01 = 0.009899$$

$$m((\theta_4 \cup \theta_5) \cap \theta_2) = 0.01 \cdot 0.98 \cdot 0.01 + 0.01 \cdot 0.98 \cdot 0.01 + 0.01 \cdot 0.01 \cdot 0.01 = 0.000197$$

$$m((\theta_4 \cup \theta_5) \cap \theta_2 \cap \theta_3) = 0.01 \cdot 0.98 \cdot 0.97 + 0.01 \cdot 0.01 \cdot 0.01 = 0.009507$$

$$m((\theta_4 \cup \theta_5) \cap \theta_3) = 0.01 \cdot 0.01 \cdot 0.97 + 0.01 \cdot 0.01 \cdot 0.01 + 0.01 \cdot 0.01 \cdot 0.97 = 0.000195$$

$$m(\theta_4 \cup \theta_5) = 0.01 \cdot 0.01 \cdot 0.01 = 0.000001$$

The sum of all masses is 1.

- Suppose now one finds out that all intersections are empty (Shafer's model), then one uses the hybrid DSm rule and one gets:

$$\begin{array}{ll} m_h(\theta_1 \cup \theta_2) = 0.019404 & m_h(\theta_1 \cup \theta_3) = 0.009702 \\ m_h(\theta_1 \cup \theta_2 \cup \theta_3) = 0.941193 & m_h(\theta_1 \cup \theta_3 \cup \theta_4 \cup \theta_5) = 0.009703 \\ m_h(\theta_1 \cup \theta_4 \cup \theta_5) = 0.000199 & m_h(\theta_4 \cup \theta_5 \cup \theta_2 \cup \theta_1) = 0.009899 \\ m_h(\theta_4 \cup \theta_5 \cup \theta_2) = 0.000197 & m_h(\theta_4 \cup \theta_5 \cup \theta_2 \cup \theta_3) = 0.009507 \\ m_h(\theta_4 \cup \theta_5 \cup \theta_3) = 0.000195 & m_h(\theta_4 \cup \theta_5) = 0.000001 \end{array}$$

The sum of all masses is 1.

5.4.3 More general

Let $\Theta = \{\theta_1, \ldots, \theta_n\}$, where $n \geq 2$, k independent experts, $k \geq 2$, and the mass matrix $\mathbf{M}$ of k rows and $n+1$ columns, corresponding to θ_1, θ_2, $\ldots$, θ_n, and one uncertainty (different from the total uncertainty $\theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$) say $\theta_{i_1} \cup \ldots \cup \theta_{i_s}$ respectively. If the following conditions occur:

- each column contains at least one zero, except the last column (of uncertainties) which has only non-null elements, $0 < \epsilon_1, \epsilon_2, \ldots, \epsilon_k < 1$, very tiny numbers (close to zero);
- the columns corresponding to the elements $\theta_{i_1}, \ldots, \theta_{i_s}$ are null (all their elements are equal to zero).

If one applies Dempster's rule, one gets $m(\theta_{i_1} \cup \ldots \cup \theta_{i_s}) = 1$ (total ignorance), which doesn't bring any information to the fusion.

5.4.4 Even more general

One can extend the previous case even more, considering to u uncertainty columns, $u \geq 1$ as follows.

Let $\Theta = \{\theta_1, \ldots, \theta_n\}$, where $n \geq 2$, k independent experts, $k \geq 2$, and the mass matrix $\mathbf{M}$ of k rows and $n+u$ columns, corresponding to θ_1, θ_2, $\ldots$, θ_n, and u uncertainty columns (different from the total uncertainty $\theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$) respectively. If the following conditions occur:

- each column contains at least one zero, except one column among the last u uncertainty ones which has only non-null elements $0 < \epsilon_1, \epsilon_2, \ldots, \epsilon_k < 1$, very tiny numbers (close to zero);

- the columns corresponding to all elements $\theta_{i_1},\ldots,\ \theta_{i_s},\ldots,\ \theta_{r_1},\ldots,\ \theta_{r_s}$ (of course, these elements should not be all $\theta_1,\ \theta_2,\ldots,\ \theta_n$, but only a part of them) that occur in all uncertainties are null (i.e., all their elements are equal to zero).

If one applies Dempster's rule, one gets $m(\theta_{i_1}\cup\ldots\cup\theta_{i_s})=1$ (total ignorance), which doesn't bring any information to the fusion.

5.5 Fourth infinite class of counter examples

This infinite class of counter-examples concerns Dempster's rule of conditioning defined as [5] :

$$\forall B\in 2^\Theta,\quad m(B|A)=\frac{\sum_{X,Y\in 2^\Theta,(X\cap Y)=B} m(X)m_A(Y)}{1-\sum_{X,Y\in 2^\Theta,(X\cap Y)=\emptyset} m(X)m_A(Y)}$$

where $m(.)$ is any proper basic belief function defined over 2^Θ and $m_A(.)$ is a particular belief function defined by choosing $m_A(A)=1$ for any $A\in 2^\Theta$ with $A\neq\emptyset$.

5.5.1 Example with $\Theta=\{\theta_1,\ldots,\theta_6\}$

Let's consider $\Theta=\{\theta_1,\ldots,\theta_6\}$, one expert and a certain body of evidence over θ_2, with the mass matrix:

	θ_1	θ_2	θ_3	$\theta_4\cup\theta_5$	$\theta_5\cup\theta_6$
$m_1(.)$	0.3	0	0.4	0.2	0.1
$m_{\theta_2}(.)$	0	1	0	0	0

- Using Dempster's rule of conditioning, one gets: $m(.|\theta_2)=0/0$ for all the masses.

- Using the DSm classical rule, one gets:

$$m(\theta_1\cap\theta_2|\theta_2)=0.3\qquad m(\theta_2\cap\theta_3|\theta_2)=0.4\qquad m(\theta_2\cap(\theta_4\cup\theta_5)|\theta_2)=0.2\qquad m(\theta_2\cap(\theta_5\cup\theta_6)|\theta_2)=0.1$$

- If now, one finds out that all intersections are empty (we adopt Shafer's model), then using the hybrid DSm rule, one gets:

$$m_h(\theta_1\cup\theta_2|\theta_2)=0.3\qquad m_h(\theta_2\cup\theta_3|\theta_2)=0.4\qquad m_h(\theta_2\cup\theta_4\cup\theta_5|\theta_2)=0.2\qquad m_h(\theta_2\cup\theta_5\cup\theta_6|\theta_2)=0.1$$

5.5.2 Another example with $\Theta=\{\theta_1,\ldots,\theta_6\}$

Let's change the previous counter-example and use now the following mass matrix:

	θ_1	θ_2	θ_3	$\theta_4\cup\theta_5$	$\theta_5\cup\theta_6$
$m_1(.)$	1	0	0	0	0
$m_{\theta_2}(.)$	0	1	0	0	0

- Using Dempster's rule of conditioning, one gets: $m(.|\theta_2) = 0/0$ for all the masses.

- Using the DSm classical rule, one gets: $m(\theta_1 \cap \theta_2|\theta_2) = 1$, and others 0.

- If now, one finds out that all intersections are empty (we adopt Shafer's model), then using the hybrid DSm rule, one gets: $m_h(\theta_1 \cup \theta_2|\theta_2) = 1$, and others 0.

5.5.3 Generalization

Let $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, where $n \geq 2$, and two basic belief functions/masses $m_1(.)$ and $m_2(.)$ such that there exist $1 \leq (i \neq j) \leq n$, where $m_1(\theta_i) = m_2(\theta_j) = 1$, and 0 otherwise. Then Dempster's rule of conditioning can not be applied because one gets division by zero.

5.5.4 Example with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and ignorance

Let's consider $\Theta = \{\theta_1, \theta_2, \theta_1, \theta_2\}$, one expert and a certain ignorant body of evidence over $\theta_3 \cup \theta_4$, with the mass matrix:

	θ_1	θ_2	$\theta_3 \cup \theta_4$
$m_1(.)$	0.3	0.7	0
$m_{\theta_3 \cup \theta_4}(.)$	0	0	1

- Using Dempster's rule of conditioning, one gets 0/0 for all masses $m(.|\theta_3 \cup \theta_4)$.

- Using the classical DSm rule, one gets: $m(\theta_1 \cap (\theta_3 \cup \theta_4)|\theta_3 \cup \theta_4) = 0.3$, $m(\theta_2 \cap (\theta_3 \cup \theta_4)|\theta_3 \cup \theta_4) = 0.7$ and others 0.

- If now one finds out that all intersections are empty (Shafer's model), using the hybrid DSm rule, one gets $m(\theta_1 \cup \theta_3 \cup \theta_4|\theta_3 \cup \theta_4) = 0.3$, $m(\theta_2 \cup \theta_3 \cup \theta_4|\theta_3 \cup \theta_4) = 0.7$ and others 0.

5.5.5 Generalization

Let $\Theta = \{\theta_1, \theta_2, \dots, \theta_n, \theta_{n+1}, \dots, \theta_{n+m}\}$, for $n \geq 2$ and $m \geq 2$. Let's consider the mass $m_1(.)$, which is a row of its values assigned for $\theta_1, \theta_2, \dots, \theta_n$, and some unions among the elements $\theta_{n+1}, \dots, \theta_{n+m}$ such that all unions are disjoint with each other. If the second mass $m_A(.)$ is a conditional mass, where A belongs to $\{\theta_1, \theta_2, \dots, \theta_n\}$ or unions among $\theta_{n+1}, \dots, \theta_{n+m}$, such that $m_1(A) = 0$, then Dempster's rule of conditioning can not be applied because one gets division by zero, which is undefined. [We did not consider any intersection of θ_i because Dempster's rule of conditioning doesn't accept paradoxes]. But the DSm rule of conditioning does work here as well.

5.5.6 Example with a paradoxical source

A counter-example with a paradox (intersection) over a non-refinable frame, where Dempster's rule of conditioning can not be applied because Dempster-Shafer theory does not accept paradoxist/conflicting information between elementary elements θ_i of the frame Θ:

Let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2\}$, one expert and a certain body of evidence over θ_2, with the mass matrix:

	θ_1	θ_2	$\theta_1 \cap \theta_2$	$\theta_1 \cup \theta_2$
$m_1(.)$	0.2	0.1	0.4	0.3
$m_{\theta_2}(.)$	0	1	0	0

Using the DSm rule of conditioning, one gets

$$m(\theta_1|\theta_2) = 0 \qquad m(\theta_2|\theta_2) = 0.1 + 0.3 = 0.4 \qquad m(\theta_1 \cap \theta_2|\theta_2) = 0.2 + 0.4 = 0.6 \qquad m(\theta_1 \cup \theta_2|\theta_2) = 0$$

and the sum of fusion results is equal to 1.

Suppose now one finds out that all intersections are empty. Using the hybrid DSm rule when $\theta_1 \cap \theta_2 = \emptyset$, one has:

$$\begin{aligned}
m_h(\theta_1 \cap \theta_2|\theta_2) &= 0 \\
m_h(\theta_1|\theta_2) &= m(\theta_1|\theta_2) + [m_1(\theta_1)m_2(\theta_1 \cap \theta_2) + m_2(\theta_1)m_1(\theta_1 \cap \theta_2)] = 0 \\
m_h(\theta_2|\theta_2) &= m(\theta_2|\theta_2) + [m_1(\theta_2)m_2(\theta_1 \cap \theta_2) + m_2(\theta_2)m_1(\theta_1 \cap \theta_2)] = 0.4 + 0.1(0) + 1(0.4) = 0.8 \\
m_h(\theta_1 \cup \theta_2|\theta_2) &= m(\theta_1 \cup \theta_2|\theta_2) + [m_1(\theta_1)m_2(\theta_2) + m_2(\theta_1)m_1(\theta_2)] \\
&\quad + [m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cup \theta_2) + m_2(\theta_1 \cap \theta_2)m_1(\theta_1 \cup \theta_2)] + [m_1(\theta_1 \cap \theta_2)m_2(\theta_1 \cap \theta_2)] \\
&= 0 + [0.2(1) + 0(0.1)] + [0.4(0) + 0(0.3)] + [0.4(0)] \\
&= 0.2 + [0] + [0] + [0] = 0.2
\end{aligned}$$

5.6 Conclusion

Several infinite classes of counter-examples to Dempster's rule of combination have been presented in this chapter for didactic purposes to show the limitations of this rule in the DST framework. These infinite classes of fusion problems bring the necessity of generalizing the DST to a more flexible theory which permits the combination of any kind of sources of information with any degree of conflict and working on any frame with exclusive or non-exclusive elements. The DSmT with the hybrid DSm rule of combination proposes a new issue to satisfy these requirements based on a new mathematical framework.

5.7 References

[1] Daniel M., *Distribution of contradictive belief masses in combination of belief functions*, in B. Bouchon-Meunier, R.R. Yager, and L. Zadeh, editors, Information, Uncertainty and Fusion, Kluwer Academic Publishers, pp. 431-446, 2000.

[2] Dubois D., Prade H., *On the unicity of the Demspter rule of combination*, International Journal of Intelligent System, 1, pp. 133-142, 1986.

[3] Dubois D., Prade H., *Representation and combination of uncertainy with belief functions and possibility measures*, Computational Intelligence, 4, pp. 244-264, 1988.

[4] Goodman I.R., Mahler R.P.S., Nguyen H.T., *Mathematics of Data Fusion*, Kluwer Academic Press, pp. 43-44, 1997.

[5] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[6] Yager R.R., *Hedging in the combination of evidence*, Journal of Information & Optimization Sciences, Vol. 4, No. 1, pp. 73-81, 1983.

[7] Zadeh L., *On the validity of Dempster's rule of combination*, Memo M 79/24, Univ. of California, Berkeley, 1979.

[8] Zadeh L., *A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination*, AI Magazine 7, pp. 85-90, 1986.

[9] Zeigler B.P., *Some properties of modified Dempster-Shafer operators in rule based inference systems*, International Journal of General Systems, Vol. 14, no. 4, pp. 345-356, 1988.

Chapter 6

Fusion of imprecise beliefs

Jean Dezert	Florentin Smarandache
ONERA	Department of Mathematics
29 Av. de la Division Leclerc	University of New Mexico
92320 Châtillon	Gallup, NM 8730
France	U.S.A.

Abstract: *In this chapter one studies, within the DSmT framework, the case when the sources of information provide imprecise belief functions/masses, and we generalize the DSm rules of combination (classic or hybrid rules) from scalar fusion to sub-unitary interval fusion and, more generally, to any set of sub-unitary interval fusion. This work generalizes previous works available in literature which appear limited to IBS (Interval-valued Belief Structures) in the Transferable Belief Model framework. Numerical didactic examples of these new DSm fusion rules for dealing with imprecise information are also presented.*

6.1 Introduction

In the previous chapters, we had focused our efforts on the fusion of *precise* uncertain and conflicting/-paradoxical generalized basic belief assignments (gbba). We mean here by precise gbba, basic belief functions/masses $m(.)$ defined precisely on the hyper-power set D^Θ where each mass $m(X)$, where X belongs to D^Θ, is represented by only one real number belonging to $[0, 1]$ such that $\sum_{X \in D^\Theta} m(X) = 1$. In this chapter, we extend the DSm fusion rules for dealing with *admissible imprecise generalized basic belief assignments* $m^I(.)$ defined as real subunitary intervals of $[0, 1]$, or even more general as real subunitary sets [i.e. sets, not necessarily intervals]. An imprecise belief assignment $m^I(.)$ over D^Θ is said admissible if and only if there exists for every $X \in D^\Theta$ at least one real number $m(X) \in m^I(X)$ such that

$\sum_{X \in D^\Theta} m(X) = 1$. The idea to work with imprecise belief structures represented by real subset intervals of $[0, 1]$ is not new and we strongly encourage the reader to examine the previous works of Lamata & Moral and also Denœux for instance on this topic in [5, 1, 2] and references therein. The proposed works available in the literature, upon our knowledge were limited only to sub-unitary interval combination in the framework of Transferable Belief Model (TBM) developed by Smets [12, 13]. We extend the approach of Lamata & Moral and Denœux based on subunitary interval-valued masses to subunitary set-valued masses; therefore the closed intervals used by Denœux to denote imprecise masses are generalized to any sets included in [0,1], i.e. in our case these sets can be unions of (closed, open, or half-open/half-closed) intervals and/or scalars all in $[0, 1]$. In this work, the proposed extension is done in the context of the DSmT framework, although it can also apply directly to fusion of IBS within TBM as well if the user prefers to adopt TBM rather than DSmT.

In many fusion problems, it seems very difficult (if not impossible) to have precise sources of evidence generating precise basic belief assignments (especially when belief functions are provided by human experts), and a more flexible plausible and paradoxical theory supporting imprecise information becomes necessary. This chapter proposes a new way to deal with the fusion of imprecise, uncertain and conflicting source of information. The section 6.2 presents briefly the DSm rule of combination for precise belief functions. In section 6.3, we present the operations on sets for the chapter to be self-contained and necessary to deal with imprecise nature of information in our framework. In section 6.4, we propose a method to combine simple imprecise belief assignment corresponding only to sub-unitary intervals also known as IBS (Interval-valued belief structures) in [1]. In section 6.5, we present the generalization of our new fusion rules to combine any type of imprecise belief assignment which may be represented by the union of several sub-unitary (half-) open intervals, (half-)closed intervals and/or sets of points belonging to [0,1]. Several numerical examples are also given. In the sequel, one uses the notation (a, b) for an open interval, $[a, b]$ for a closed interval, and $(a, b]$ or $[a, b)$ for a half open and half closed interval.

6.2 Combination of precise beliefs

6.2.1 General DSm rule of combination

Let's consider a frame of discernment of a fusion problem $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, its hyper-power set D^Θ (i.e. the set of all propositions built from elements θ_i of Θ with $\cap$ and $\cup$ operators (see chapter 2), and k independent (precise) sources of information $\mathcal{B}_1$, $\mathcal{B}_2$, $\ldots$, $\mathcal{B}_k$ with their associated generalized basic belief assignments (gbba) $m_1(.)$, $m_2(.)$, $\ldots$, $m_k(.)$ defined over D^Θ. Let $\mathbf{M}$ be the mass matrix

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1d} \\ m_{21} & m_{22} & \ldots & m_{2d} \\ \ldots & \ldots & \ldots & \ldots \\ m_{k1} & m_{k2} & \ldots & m_{kd} \end{bmatrix}$$

where $d = \mid D^\Theta \mid$ is the dimension of the hyper-power set, and $m_{ij} \in [0,1]$ for all $1 \leq i \leq k$ and $1 \leq j \leq d$, is the mass assigned by source $\mathcal{B}_i$ to the element $A_j \in D^\Theta$. We use the DSm ordering procedure presented in chapter 3 for enumerating the elements A_1, A_2, …, A_d of the hyper-power set D^Θ. The matrix $\mathbf{M}$ characterizes all information available which has to be combined to solve the fusion problem under consideration. Since $m_1(.)$, $m_2(.)$, …, $m_k(.)$ are gbba, the summation on each row of the matrix must be one. For any (possibly hybrid) model $\mathcal{M}(\Theta)$, we apply the DSm general rule of combination (also called hybrid DSm rule) for $k \geq 2$ sources to fuse the masses (see chapter 4) defined for all $A \in D^\Theta$ as:

$$m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A)\Big[S_1(A) + S_2(A) + S_3(A)\Big] \tag{6.1}$$

$\phi(A)$ is the *characteristic non emptiness function* of the set A, i.e. $\phi(A) = 1$ if $A \notin \boldsymbol{\emptyset}$ and $\phi(A) = 0$ otherwise. $\boldsymbol{\emptyset} \triangleq \{\emptyset, \boldsymbol{\emptyset}_\mathcal{M}\}$ represents the set absolutely empty and of all relatively empty elements belonging to D^Θ (elements/propositions which have been forced to empty set in the chosen hybrid model $\mathcal{M}(\Theta)$). If no constraint is introduced in the model, $\boldsymbol{\emptyset}$ reduces to $\{\emptyset\}$ and this corresponds to the free DSm model (see chapter 4). If all constraints of exclusivity between elements $\theta_i \in \Theta$ are introduced, the hybrid model $\mathcal{M}(\Theta)$ corresponds to Shafer's model on which is based Dempster-Shafer Theory (DST) [9]. $S_1(A)$, $S_2(A)$ and $S_3(A)$ are defined by

$$S_1(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1 \cap X_2 \cap \ldots \cap X_k) = A}} \prod_{i=1}^{k} m_i(X_i) \tag{6.2}$$

$$S_2(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in \boldsymbol{\emptyset} \\ [\mathcal{U}=A] \vee [(\mathcal{U} \in \boldsymbol{\emptyset}) \wedge (A = I_t)]}} \prod_{i=1}^{k} m_i(X_i) \tag{6.3}$$

$$S_3(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1 \cup X_2 \cup \ldots \cup X_k) = A \\ (X_1 \cap X_2 \cap \ldots \cap X_k) \in \boldsymbol{\emptyset}}} \prod_{i=1}^{k} m_i(X_i) \tag{6.4}$$

where $I_t \triangleq \theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$ and $\mathcal{U} \triangleq u(X_1) \cup u(X_2) \cup \ldots \cup u(X_k)$. $u(X)$ is the union of all singletons θ_i that compose X. For example, if X is a singleton then $u(X) = X$; if $X = \theta_1 \cap \theta_2$ or $X = \theta_1 \cup \theta_2$ then $u(X) = \theta_1 \cup \theta_2$; if $X = (\theta_1 \cap \theta_2) \cup \theta_3$ then $u(X) = \theta_1 \cup \theta_2 \cup \theta_3$, etc; by convention $u(\emptyset) \triangleq \emptyset$.

6.2.2 Examples

Let's consider at time t the frame of discernment $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and two independent bodies of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$ with the generalized basic belief assignments $m_1(.)$ and $m_2(.)$ given by:

$A \in D^\Theta$	$m_1(A)$	$m_2(A)$
θ_1	0.1	0.5
θ_2	0.2	0.3
θ_3	0.3	0.1
$\theta_1 \cap \theta_2$	0.4	0.1

Table 6.1: Inputs of the fusion with precise bba

Based on the free DSm model and the classical DSm rule (6.2), the combination denoted by the symbol $\oplus$ (i.e. $m(.) = [m_1 \oplus m_2](.)$) of these two precise sources of evidence is

$A \in D^\Theta$	$m(A) = [m_1 \oplus m_2](A)$
θ_1	0.05
θ_2	0.06
θ_3	0.03
$\theta_1 \cap \theta_2$	0.52
$\theta_1 \cap \theta_3$	0.16
$\theta_2 \cap \theta_3$	0.11
$\theta_1 \cap \theta_2 \cap \theta_3$	0.07

Table 6.2: Fusion with DSm classic rule

Then, assume at time $t+1$ one finds out for some reason that the free DSm model has to be changed by introducing the constraint $\theta_1 \cap \theta_2 = \emptyset$ which involves also $\theta_1 \cap \theta_2 \cap \theta_3 = \emptyset$. This characterizes the hybrid-model $\mathcal{M}$ we have to work with. Then one uses the general hybrid DSm rule of combination for scalars (i.e. for precise masses $m_1(.)$ and $m_2(.)$ to get the new result of the fusion at time $t+1$. According to (6.1), one obtains $m(\theta_1 \cap \theta_2 \overset{\mathcal{M}}{\equiv} \emptyset) = 0$, $m(\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}}{\equiv} \emptyset) = 0$ and

$A \in D^\Theta$	$m(A)$
θ_1	$0.05 + [0.1(0.1) + 0.5(0.4)] = 0.26$
θ_2	$0.06 + [0.2(0.1) + 0.3(0.4)] = 0.20$
θ_3	$0.03 + [0.3(0.1) + 0.1(0.4)] = 0.10$
$\theta_1 \cap \theta_3$	0.16
$\theta_2 \cap \theta_3$	0.11
$\theta_1 \cup \theta_2$	$0 + [0.13] + [0.04] = 0.17$

Table 6.3: Fusion with hybrid DSm rule for model $\mathcal{M}$

6.3 Operations on sets

To manipulate imprecise information and for the chapter to be self-contained, we need to introduce operations on sets as follows (detailed presentations on Interval Analysis and Methods can be found in [3, 4, 6, 7, 8]). The interval operations defined here about imprecision are similar to the rational interval extension through the interval arithmetics [10], but they are different from Modal Interval Analysis which doesn't serve our fusion needs. We are not interested in a dual of an interval $[a, b]$, used in the Modal Interval Analysis, because we always consider $a \leq b$, while its dual, $\mathrm{Du}([a, b]) = [b, a]$, doesn't occur. Yet, we generalize the interval operations to any set operations. Of course, for the fusion we only need real sub-unitary sets, but these defined set operations can be used for any kind of sets.

Let S_1 and S_2 be two (unidimensional) real standard subsets of the unit interval $[0, 1]$, and a number $k \in [0, 1]$, then one defines [11] :

- **Addition of sets**

$$S_1 \boxplus S_2 = S_2 \boxplus S_1 \triangleq \{x \mid x = s_1 + s_2, s_1 \in S_1, s_2 \in S_2\} \quad \text{with} \quad \begin{cases} \inf(S_1 \boxplus S_2) = \inf(S_1) + \inf(S_2) \\ \sup(S_1 \boxplus S_2) = \sup(S_1) + \sup(S_2) \end{cases}$$

and, as a particular case, we have

$$\{k\} \boxplus S_2 = S_2 \boxplus \{k\} = \{x \mid x = k + s_2, s_2 \in S_2\} \quad \text{with} \quad \begin{cases} \inf(\{k\} \boxplus S_2) = k + \inf(S_2) \\ \sup(\{k\} \boxplus S_2) = k + \sup(S_2) \end{cases}$$

Examples:

$[0.1, 0.3] \boxplus [0.2, 0.5] = [0.3, 0.8]$ because $0.1 + 0.2 = 0.3$ and $0.3 + 0.5 = 0.8$;

$(0.1, 0.3] \boxplus [0.2, 0.5] = (0.3, 0.8]$;

$[0.1, 0.3] \boxplus (0.2, 0.5] = (0.3, 0.8]$;

$[0.1, 0.3) \boxplus [0.2, 0.5] = [0.3, 0.8)$;

$[0.1, 0.3] \boxplus [0.2, 0.5) = [0.3, 0.8)$;

$(0.1, 0.3] \boxplus (0.2, 0.5) = (0.3, 0.8)$;

$[0.7, 0.8] \boxplus [0.5, 0.9] = [1.2, 1.7]$;

$\{0.4\} \boxplus [0.2, 0.5] = [0.2, 0.5] \boxplus \{0.4\} = [0.6, 0.9]$ because $0.4 + 0.2 = 0.6$ and $0.4 + 0.5 = 0.9$;

$\{0.4\} \boxplus (0.2, 0.5] = (0.6, 0.9]$;

$\{0.4\} \boxplus [0.2, 0.5) = [0.6, 0.9)$;

$\{0.4\} \boxplus (0.2, 0.5) = (0.6, 0.9)$.

- **Subtraction of sets**

$$S_1 \boxminus S_2 \triangleq \{x \mid x = s_1 - s_2, s_1 \in S_1, s_2 \in S_2\} \quad \text{with} \quad \begin{cases} \inf(S_1 \boxminus S_2) = \inf(S_1) - \sup(S_2) \\ \sup(S_1 \boxminus S_2) = \sup(S_1) - \inf(S_2) \end{cases}$$

and, as a particular case, we have

$$\{k\} \boxminus S_2 = \{x \mid x = k - s_2, s_2 \in S_2\} \quad \text{with} \quad \begin{cases} \inf(\{k\} \boxminus S_2) = k - \sup(S_2) \\ \sup(\{k\} \boxminus S_2) = k - \inf(S_2) \end{cases}$$

and similarly for $S_2 \boxminus \{k\}$ with $\begin{cases} \inf(S_2 \boxminus \{k\}) = \inf(S_2) - k \\ \sup(S_2 \boxminus \{k\}) = \sup(S_2) - k \end{cases}$

Examples:

$[0.3, 0.7] \boxminus [0.2, 0.3] = [0.0, 0.5]$ because $0.3 - 0.3 = 0.0$ and $0.7 - 0.2 = 0.5$;

$[0.3, 0.7] \boxminus \{0.1\} = [0.2, 0.6]$;

$\{0.8\} \boxminus [0.3, 0.7] = [0.1, 0.5]$ because $0.8 - 0.7 = 0.1$ and $0.8 - 0.3 = 0.5$;

$[0.1, 0.8] \boxminus [0.5, 0.6] = [-0.5, 0.3]$;

$[0.1, 0.8] \boxminus [0.2, 0.9] = [-0.8, 0.6]$;

$[0.2, 0.5] \boxminus [0.1, 0.6] = [-0.4, 0.4]$.

- **Multiplication of sets**

$$S_1 \boxdot S_2 \triangleq \{x \mid x = s_1 \cdot s_2, s_1 \in S_1, s_2 \in S_2\} \quad \text{with} \quad \begin{cases} \inf(S_1 \boxdot S_2) = \inf(S_1) \cdot \inf(S_2) \\ \sup(S_1 \boxdot S_2) = \sup(S_1) \cdot \sup(S_2) \end{cases}$$

and, as a particular case, we have

$$\{k\} \boxdot S_2 = S_2 \boxdot \{k\} = \{x \mid x = k \cdot s_2, s_2 \in S_2\} \quad \text{with} \quad \begin{cases} \inf(\{k\} \boxdot S_2) = k \cdot \inf(S_2) \\ \sup(\{k\} \boxdot S_2) = k \cdot \sup(S_2) \end{cases}$$

Examples:

$[0.1, 0.6] \boxdot [0.8, 0.9] = [0.08, 0.54]$ because $0.1 \cdot 0.8 = 0.08$ and $0.6 \cdot 0.9 = 0.54$;

$[0.1, 0.6] \boxdot \{0.3\} = \{0.3\} \boxdot [0.1, 0.6] = [0.03, 0.18]$ because $0.3 \cdot 0.1 = 0.03$ and $0.3 \cdot 0.6 = 0.18$.

- **Division of sets**

In our fusion context, the division of sets is not necessary since the DSm rules of combination (classic or hybrid ones) do not require a normalization procedure and thus a division operation. Actually, the DSm rules require only addition and multiplication operations. We however give here the definition of division of sets only for the reader's interest and curiosity. The division of sets is defined as follows:

If $0 \notin S_2$, then $S_1 \boxslash S_2 \triangleq \{x \mid x = s_1/s_2, s_1 \in S_1, s_2 \in S_2\}$ with $\begin{cases} \inf(S_1 \boxslash S_2) = \inf(S_1)/\sup(S_2) \\ \sup(S_1 \boxslash S_2) = \sup(S_1)/\inf(S_2) \text{ if } 0 \notin S_2 \\ \sup(S_1 \boxslash S_2) = +\infty \text{ if } 0 \in S_2 \end{cases}$

If $0 \in S_2$, then $S_1 \boxslash S_2 = [\inf(S_1)/\sup(S_2), +\infty)$

and as some particular cases, we have for $k \neq 0$,

$$\{k\} \boxslash S_2 = \{x \mid x = k/s_2, \text{where } s_2 \in S_2 \setminus \{0\}\} \quad \text{with} \quad \begin{cases} \inf(\{k\} \boxslash S_2) = k/\sup(S_2) \\ \sup(\{k\} \boxslash S_2) = k/\inf(S_2) \end{cases}$$

and if $0 \in S_2$ then $\sup(\{k\} \boxslash S_2) = +\infty$

One has also as some particular case for $k \neq 0$,

$$S_2 \boxslash \{k\} = \{x \mid x = s_2/k, \text{where } s_2 \in S_2\} \quad \text{with} \quad \begin{cases} \inf(S_2 \boxslash \{k\}) = \inf(S_2)/k \\ \sup(S_2 \boxslash \{k\}) = \sup(S_2)/k \end{cases}$$

Examples:

$[0.4, 0.6] \boxslash [0.1, 0.2] = [2, 6]$ because $0.4/0.2 = 2$ and $0.6/0.1 = 6$;

$[0.4, 0.6] \boxslash \{0.4\} = [1, 1.5]$ because $0.4/0.4 = 1$ and $0.6/0.4 = 1.5$;

$\{0.8\} \boxslash [0.2, 0.5] = [1.6, 4]$ because $0.8/0.2 = 4$ and $0.8/0.5 = 1.6$;

$[0, 0.5] \boxslash [0.1, 0.2] = [0, 5]$: $[0, 0.5] \boxslash \{0.4\} = [0, 1.25]$ because $0/0.4 = 0$ and $0.5/0.4 = 1.25$;

$[0.3, 0.9] \boxslash [0, 0.2] = [1.5, +\infty)$ because $0.3/0.2 = 1.5$ and since $0 \in (S_2 = [0, 0.2])$, $\sup([0.3, 0.9] \boxslash [0, 0.2]) = +\infty$;

$[0, 0.9] \boxslash [0, 0.2] = [0, +\infty)$:

$\{0.7\} \boxslash [0, 0.2] = [3.5, +\infty)$ because $0.7/0.2 = 3.5$ and $0 \in (S_2 = [0, 0.2])$, $\sup(\{0.7\} \boxslash [0, 0.2]) = +\infty$;

$\{0\} \boxslash [0, 0.2] = [0, +\infty)$: $[0.3, 0.9] \boxslash \{0\} = +\infty$:

$[0, 0.9] \boxslash \{0\} = +\infty$:

$[0.2, 0.7] \boxslash [0, 0.8] = [0.25, +\infty)$.

These operations can be directly extended for any types of sets (not necessarily sub-unitary subsets as it will be shown in our general examples of section 6), but for simplicity, we will start the presentation in the following section only for sub-unitary subsets.

Due to the fact that the fusion of imprecise information must also be included in the unit interval $[0, 1]$ as it happens with the fusion of precise information, if the masses computed are less than 0 one replaces them by 0, and similarly if they are greater than 1 one replaces them by 1. For example (specifically in our fusion context): $[0.2, 0.4] \boxplus [0.5, 0.8] = [0.7, 1.2]$ will be forced to $[0.7, 1]$.

6.4 Fusion of beliefs defined on single sub-unitary intervals

6.4.1 DSm rules of combination

Let's now consider some given sources of information which are not able to provide us a specific/precise mass $m_{ij} \in [0, 1]$, but only an interval centered[1] in m_{ij}, i.e. $I_{ij} = [m_{ij} - \epsilon_{ij}, m_{ij} + \epsilon_{ij}]$ where $0 \leq \epsilon_{ij} \leq 1$ and $I_{ij} \subseteq [0, 1]$ for all $1 \leq i \leq k$ and $1 \leq j \leq d$. The cases when I_{ij} are half-closed or open are similarly treated.

Lemma 1: if $A, B \subseteq [0, 1]$ and $\alpha \in [0, 1]$ then:

$$\begin{cases} \inf(A \boxdot B) = \inf(A) \cdot \inf(B) \\ \sup(A \boxdot B) = \sup(A) \cdot \sup(B) \end{cases} \qquad \begin{cases} \inf(A \oplus B) = \inf(A) + \inf(B) \\ \sup(A \oplus B) = \sup(A) + \sup(B) \end{cases}$$

$$\begin{cases} \inf(\alpha \cdot A) = \alpha \cdot \inf(A) \\ \sup(\alpha \cdot A) = \alpha \cdot \sup(A) \end{cases} \qquad \begin{cases} \inf(\alpha + A) = \alpha + \inf(A) \\ \sup(\alpha + A) = \alpha + \sup(A) \end{cases}$$

We can regard a scalar α as a particular interval $[\alpha, \alpha]$, thus all operations of the previous lemma are reduced to multiplications and additions of sub-unitary intervals. Therefore, the DSm general rule (6.1), which operates (multiplies and adds) sub-unitary scalars, can be extended to operate sub-unitary intervals. The formula (6.1) remains the same, but $m_i(X_i)$, $1 \leq i \leq k$, are sub-unitary intervals I_{ij}. The

[1]This interval centered assumption is not important actually but has been adopted here only for notational convenience.

mass matrix $\mathbf{M}$ is extended to:

$$\inf(\mathbf{M}) = \begin{bmatrix} m_{11}-\epsilon_{11} & m_{12}-\epsilon_{12} & \ldots & m_{1d}-\epsilon_{1d} \\ m_{21}-\epsilon_{21} & m_{22}-\epsilon_{22} & \ldots & m_{2d}-\epsilon_{2d} \\ \ldots & \ldots & \ldots & \ldots \\ m_{k1}-\epsilon_{k1} & m_{k2}-\epsilon_{k2} & \ldots & m_{kd}-\epsilon_{kd} \end{bmatrix}$$

$$\sup(\mathbf{M}) = \begin{bmatrix} m_{11}+\epsilon_{11} & m_{12}+\epsilon_{12} & \ldots & m_{1d}+\epsilon_{1d} \\ m_{21}+\epsilon_{21} & m_{22}+\epsilon_{22} & \ldots & m_{2d}+\epsilon_{2d} \\ \ldots & \ldots & \ldots & \ldots \\ m_{k1}+\epsilon_{k1} & m_{k2}+\epsilon_{k2} & \ldots & m_{kd}+\epsilon_{kd} \end{bmatrix}$$

Notations: Let's distinguish between DSm general rule for scalars, noted as usual $m_{\mathcal{M}(\Theta)}(A)$, or $m_i(X_i)$, etc., and the DSm general rule for intervals noted as $m^I_{\mathcal{M}(\Theta)}(A)$, or $m^I_i(X_i)$, etc. Hence, the DSm general rule for interval-valued masses is:

$$\inf(m^I_{\mathcal{M}(\Theta)}(A)) \triangleq \phi(A)\Big[S_1^{\inf}(A) + S_2^{\inf}(A) + S_3^{\inf}(A)\Big] \tag{6.5}$$

with

$$S_1^{\inf}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in D^\Theta \\ (X_1\cap X_2\cap\ldots\cap X_k)=A}} \prod_{i=1}^{k} \inf(m_i^I(X_i))$$

$$S_2^{\inf}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in\boldsymbol{\emptyset} \\ [\mathcal{U}=A]\vee[(\mathcal{U}\in\boldsymbol{\emptyset})\wedge(A=I_t)]}} \prod_{i=1}^{k} \inf(m_i^I(X_i))$$

$$S_3^{\inf}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in D^\Theta \\ (X_1\cup X_2\cup\ldots\cup X_k)=A \\ (X_1\cap X_2\cap\ldots\cap X_k)\in\boldsymbol{\emptyset}}} \prod_{i=1}^{k} \inf(m_i^I(X_i))$$

and

$$\sup(m^I_{\mathcal{M}(\Theta)}(A)) \triangleq \phi(A)\Big[S_1^{\sup}(A) + S_2^{\sup}(A) + S_3^{\sup}(A)\Big] \tag{6.6}$$

with

$$S_1^{\sup}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in D^\Theta \\ (X_1\cap X_2\cap\ldots\cap X_k)=A}} \prod_{i=1}^{k} \sup(m_i^I(X_i))$$

$$S_2^{\sup}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in\boldsymbol{\emptyset} \\ [\mathcal{U}=A]\vee[(\mathcal{U}\in\boldsymbol{\emptyset})\wedge(A=I_t)]}} \prod_{i=1}^{k} \sup(m_i^I(X_i))$$

$$S_3^{\sup}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k\in D^\Theta \\ (X_1\cup X_2\cup\ldots\cup X_k)=A \\ (X_1\cap X_2\cap\ldots\cap X_k)\in\boldsymbol{\emptyset}}} \prod_{i=1}^{k} \sup(m_i^I(X_i))$$

Actually formula (6.5) results from applying the hybrid DSm rule for scalars to the matrix $\inf(\mathbf{M})$, while formula (6.6) results from applying the hybrid DSm rule for scalars to the matrix $\sup(\mathbf{M})$. The bounds of the DSm classic rule for the free DSm model are given for all $A \in D^\Theta$ by $S_1^{\inf}(A)$ and $S_1^{\sup}(A)$. Combining (6.5) and (6.6), one gets directly:

$$m^I_{\mathcal{M}(\Theta)}(A) = [\inf m^I_{\mathcal{M}(\Theta)}(A), \sup m^I_{\mathcal{M}(\Theta)}(A)] \tag{6.7}$$

Of course, the closeness of this interval to the left and/or to the right depends on the closeness of the combined intervals I_{ij}. If all of them are closed to the left, then $m^I_{\mathcal{M}(\Theta)}(A)$ is also closed to the left. But, if at least one is open to the left, then $m^I_{\mathcal{M}(\Theta)}(A)$ is open to the left. Similarly for the closeness to the right. Because one has $\forall i = 1, \ldots, k$ and $\forall j = 1, \ldots, d$:

$$\lim_{\epsilon_{ij}\to 0}(\inf(\mathbf{M})) = \lim_{\epsilon_{ij}\to 0}(\sup(\mathbf{M})) = \mathbf{M} \tag{6.8}$$

It results the following theorem.

Theorem 1: $\forall A \in D^\Theta$, $\forall i = 1, \ldots, k$ and $\forall j = 1, \ldots, d$, one has:

$$\lim_{\epsilon_{ij}\to 0} m^I_{\mathcal{M}(\Theta)}(A) = [\lim_{\inf_{ij}}(A), \lim_{\sup_{ij}}(A)] \quad \text{with} \quad \begin{cases} \lim_{\inf_{ij}}(A) \triangleq \lim_{\epsilon_{ij}\to 0}(\inf(m^I_{\mathcal{M}(\Theta)}(A))) \\ \lim_{\sup_{ij}}(A) \triangleq \lim_{\epsilon_{ij}\to 0}(\sup(m^I_{\mathcal{M}(\Theta)}(A))) \end{cases} \tag{6.9}$$

In other words, if all centered sub-unitary intervals converge to their corresponding mid points (the imprecision becomes zero), then the DSm rule for intervals converges towards the DSm rule for scalars.

Normally we must apply the DSm classical or hybrid rules directly to the interval-valued masses, but this is equivalent to applying the DSm rules to the inferior and superior bounds of each mass. If, after fusion, the sum of inferior masses is < 1 (which occurs all the time because combining incomplete masses one gets incomplete results) and the sum of superior masses is ≥ 1 (which occurs all the time because combining paraconsistent masses one gets paraconsistent results), then there exist points in each resulted interval-valued mass such that their sum is 1 (according to a continuity theorem - see section 6.5.2).

6.4.2 Example with the DSm classic rule

Let's take back the previous example (see section 6.2.2), but let's now suppose the sources of information give at time t imprecise generalized basic belief assignments, i.e. interval-valued masses centered in the scalars given in section 6.2.2, of various radii according to table 6.4.

Based on the free DSm model and the classical DSm rule applied to imprecise basic belief assignments following the method proposed in previous section, one has:

$$m^I(\theta_1) = [0.05, 0.15] \boxdot [0.4, 0.6] = [0.020, 0.090]$$

$$m^I(\theta_2) = [0.1, 0.3] \boxdot [0.1, 0.5] = [0.010, 0.150]$$

$$m^I(\theta_3) = [0.15, 0.45] \boxdot [0, 0.2] = [0, 0.090]$$

$A \in D^\Theta$	$m_1^I(A)$	$m_2^I(A)$
θ_1	$[0.05, 0.15]$	$[0.4, 0.6]$
θ_2	$[0.1, 0.3]$	$[0.1, 0.5]$
θ_3	$[0.15, 0.45]$	$[0, 0.2]$
$\theta_1 \cap \theta_2$	$[0.2, 0.6]$	$[0.05, 0.15]$

Table 6.4: Inputs of the fusion with imprecise bba

$$m^I(\theta_1 \cap \theta_3) = [[0.05, 0.15] \boxdot [0, 0.2]] \boxplus [[0.4, 0.6] \boxdot [0.15, 0.45]] = [0, 0.030] \boxplus [0.060, 0.270] = [0.060, 0.300]$$

$$m^I(\theta_2 \cap \theta_3) = [[0.1, 0.3] \boxdot [0, 0.2]] \boxplus [[0.1, 0.5] \boxdot [0.15, 0.45]] = [0, 0.06] \boxplus [0.015, 0.225] = [0.015, 0.285]$$

$$\begin{aligned} m^I(\theta_1 \cap \theta_2 \cap \theta_3) &= [[0.15, 0.45] \boxdot [0.05, 0.15]] \boxplus [[0, 0.2] \boxdot [0.2, 0.6]] \\ &= [0.0075, 0.0675] \boxplus [0, 0.12] \\ &= [0.0075, 0.1875] \end{aligned}$$

$$\begin{aligned} m^I(\theta_1 \cap \theta_2) = & [[0.2, 0.6] \boxdot [0.05, 0.15]] \boxplus [[0.05, 0.15] \boxdot [0.05, 0.15]] \boxplus [[0.4, 0.6] \boxdot [0.2, 0.6]] \boxplus \\ & [[0.1, 0.3] \boxdot [0.05, 0.15]] \boxplus [[0.1, 0.5] \boxdot [0.2, 0.6]] \boxplus \\ & [[0.05, 0.15] \boxdot [0.1, 0.5]] \boxplus [[0.4, 0.6] \boxdot [0.1, 0.3]] \\ = & [0.010, 0.90] \boxplus [0.0025, 0.0225] \boxplus [0.08, 0.36] \boxplus [0.005, 0.045] \boxplus \\ & [0.02, 0.30] \boxplus [0.005, 0.075] \boxplus [0.04, 0.18] = [0.1625, 1.0725] \equiv [0.1625, 1] \end{aligned}$$

The last *equality* comes from the absorption of $[0.1625, 1.0725]$ into $[0.1625, 1]$ according to operations on sets defined in this fusion context. Thus, the final result of combination $m^I(.) = [m_1^I \oplus m_2^I](.)$ of these two imprecise sources of evidence is given in table 6.5.

$A \in D^\Theta$	$m^I(A) = [m_1^I \oplus m_2^I](A)$
θ_1	$[0.020, 0.090]$
θ_2	$[0.010, 0.150]$
θ_3	$[0, 0.090]$
$\theta_1 \cap \theta_2$	$[0.1625, 1.0725 \to 1]$
$\theta_1 \cap \theta_3$	$[0.060, 0.300]$
$\theta_2 \cap \theta_3$	$[0.015, 0.285]$
$\theta_1 \cap \theta_2 \cap \theta_3$	$[0.0075, 0.1875]$

Table 6.5: Fusion with DSm classic rule for free DSm model

There exist some points, for example 0.03, 0.10. 0.07, 0.4, 0.1, 0.2, 0.1 from the intervals $[0.020, 0.090], \ldots,$ $[0.0075, 0.1875]$ respectively such that their sum is 1 and therefore the admissibility of the fusion result holds. Note that this fusion process is equivalent to using the DSm classic rule for scalars for inferior limit and incomplete information (see table 6.6), and the same rule for superior limit and paraconsistent information (see table 6.7).

$A \in D^\Theta$	$m_1^{\inf}(A)$	$m_2^{\inf}(A)$	$m^{\inf}(A)$
θ_1	0.05	0.4	0.020
θ_2	0.1	0.1	0.010
θ_3	0.15	0	0
$\theta_1 \cap \theta_2$	0.2	0.05	0.1625
$\theta_1 \cap \theta_3$	0	0	0.060
$\theta_2 \cap \theta_3$	0	0	0.015
$\theta_1 \cap \theta_2 \cap \theta_3$	0	0	0.0075

Table 6.6: Fusion with DSm classic rule on lower bounds

$A \in D^\Theta$	$m_1^{\sup}(A)$	$m_2^{\sup}(A)$	$m^{\sup}(A)$
θ_1	0.15	0.6	0.090
θ_2	0.3	0.5	0.150
θ_3	0.45	0.2	0.090
$\theta_1 \cap \theta_2$	0.6	0.15	$1.0725 \to 1$
$\theta_1 \cap \theta_3$	0	0	0.300
$\theta_2 \cap \theta_3$	0	0	0.285
$\theta_1 \cap \theta_2 \cap \theta_3$	0	0	0.1875

Table 6.7: Fusion with DSm classic rule on upper bounds

6.4.3 Example with the hybrid DSm rule

Then, assume at time $t+1$, that one finds out for some reason that the free DSm model has to be changed by introducing the constraint $\theta_1 \cap \theta_2 = \emptyset$ which involves also $\theta_1 \cap \theta_2 \cap \theta_3 = \emptyset$. One directly applies the hybrid DSm rule for set to get the new belief masses:

$$\begin{aligned} m^I(\theta_1) &= [0.020, 0.090] \boxplus [[0.05, 0.15] \boxdot [0.05, 0.15]] \boxplus [[0.4, 0.6] \boxdot [0.2, 0.6]] \\ &= [0.020, 0.090] \boxplus [0.0025, 0.0225] \boxplus [0.08, 0.36] = [0.1025, 0.4725] \end{aligned}$$

$$\begin{aligned} m^I(\theta_2) &= [0.010, 0.150] \boxplus [[0.1, 0.3] \boxdot [0.05, 0.15]] \boxplus [[0.1, 0.5] \boxdot [0.2, 0.6]] \\ &= [0.010, 0.150] \boxplus [0.005, 0.045] \boxplus [0.02, 0.30] = [0.035, 0.495] \end{aligned}$$

$$\begin{aligned} m^I(\theta_3) &= [0, 0.090] \boxplus [[0.15, 0.45] \boxdot [0.05, 0.15]] \boxplus [[0, 0.2] \boxdot [0.2, 0.6]] \\ &= [0, 0.090] \boxplus [0.0075, 0.0675] \boxplus [0, 0.12] = [0.0075, 0.2775] \end{aligned}$$

$$\begin{aligned} m^I(\theta_1 \cup \theta_2) &= [[02, 0.6] \boxdot [0.05, 0.15]] \boxplus [[0.05, 0.15] \boxdot [0.1, 0.5]] \boxplus [[0.4, 0.6] \boxdot [0.1, 0.3]] \\ &= [0.010, 0.090] \boxplus [0.005, 0.075] \boxplus [0.04, 0.18] = [0.055, 0.345] \end{aligned}$$

$m^I(\theta_1 \cap \theta_2) = m^I(\theta_1 \cap \theta_2 \cap \theta_3) = 0$ by definition of empty masses (due to the choice of the hybrid model $\mathcal{M}$). $m^I(\theta_1 \cap \theta_3) = [0.060, 0.300]$ and $m^I(\theta_2 \cap \theta_3) = [0.015, 0.285]$ remain the same. Finally, the result of the fusion of imprecise belief assignments for the chosen hybrid model $\mathcal{M}$, is summarized in table 6.8.

$A \in D^\Theta$	$m^I(A) = [m^{\inf}(A), m^{\sup}(A)]$
θ_1	$[0.1025, 0.4725]$
θ_2	$[0.035, 0.495]$
θ_3	$[0.0075, 0.2775]$
$\theta_1 \cap \theta_2 \overset{\mathcal{M}}{\equiv} \emptyset$	$[0, 0] = 0$
$\theta_1 \cap \theta_3$	$[0, 060, 0.300]$
$\theta_2 \cap \theta_3$	$[0.015, 0.285]$
$\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}}{\equiv} \emptyset$	$[0, 0] = 0$
$\theta_1 \cup \theta_2$	$[0.055, 0.345]$

Table 6.8: Fusion with hybrid DSm rule for model $\mathcal{M}$

The admissibility of the fusion result still holds since there exist some points, for example 0.1, 0.3, 0.1, 0, 0.2, 0.1, 0, 0.2 from the intervals $[0.1025, 0.4725]$, …, $[0.055, 0.345]$ respectively such that their sum is 1. Actually in each of these examples there are infinitely many such groups of points in each respective interval whose sum is 1. This can be generalized for any examples.

6.5 Generalization of DSm rules for sets

In this section, we extend the previous results on the fusion of admissible imprecise information defined only on single sub-unitary intervals to the general case where the imprecision is defined on sets. In other words, in the previous section we dealt with admissible imprecise masses having the form $m^I(A) = [a, b] \subseteq [0, 1]$, and now we deals with admissible imprecise masses having the form $m^I(A) = [a_1, b_1] \cup \ldots \cup [a_m, b_m] \cup (c_1, d_1) \cup \ldots \cup (c_n, d_n) \cup (e_1, f_1] \cup \ldots \cup (e_p, f_p] \cup [g_1, h_1) \cup \ldots \cup [g_q, h_q) \cup \{A_1, \ldots, A_r\}$ where all the bounds or elements involved into $m^I(A)$ belong to $[0, 1]$.

6.5.1 General DSm rules for imprecise beliefs

From our previous results, one can generalize the DSm classic rule from scalars to sets in the following way: $\forall A \neq \emptyset \in D^\Theta$,

$$m^I(A) = \boxed{\sum}_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1 \cap X_2 \cap \ldots \cap X_k) = A}} \boxed{\prod}_{i=1,\ldots,k} m_i^I(X_i) \tag{6.10}$$

where $\boxed{\sum}$ and $\boxed{\prod}$ represent the summation, and respectively product, of sets.

Similarly, one can generalize the hybrid DSm rule from scalars to sets in the following way:

$$m^I_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A) \boxdot \left[S_1^I(A) \boxplus S_2^I(A) \boxplus S_3^I(A) \right] \tag{6.11}$$

$\phi(A)$ is the *characteristic non emptiness function* of the set A and $S_1^I(A)$, $S_2^I(A)$ and $S_3^I(A)$ are defined by

$$S_1^I(A) \triangleq \boxed{\sum}_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1 \cap X_2 \cap \ldots \cap X_k) = A}} \boxed{\prod}_{i=1,\ldots,k} m_i^I(X_i) \tag{6.12}$$

$$S_2^I(A) \triangleq \boxed{\sum}_{\substack{X_1,X_2,\ldots,X_k \in \boldsymbol{\emptyset} \\ [\mathcal{U}=A] \vee [(\mathcal{U} \in \boldsymbol{\emptyset}) \wedge (A = I_t)]}} \boxed{\prod}_{i=1,\ldots,k} m_i^I(X_i) \tag{6.13}$$

$$S_3^I(A) \triangleq \boxed{\sum}_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1 \cup X_2 \cup \ldots \cup X_k) = A \\ (X_1 \cap X_2 \cap \ldots \cap X_k) \in \boldsymbol{\emptyset}}} \boxed{\prod}_{i=1,\ldots,k} m_i^I(X_i) \tag{6.14}$$

In the case when all sets are reduced to points (numbers), the set operations become normal operations with numbers; the sets operations are generalizations of numerical operations.

6.5.2 Some lemmas and a theorem

Lemma 2: Let the scalars $a, b \geq 0$ and the intervals $I_1, I_2 \subseteq [0, 1]$, with $a \in I_1$ and $b \in I_2$. Then obviously $(a + b) \in I_1 \boxplus I_2$ and $(a \cdot b) \in I_1 \boxdot I_2$.

Because in DSm rules of combining imprecise information, one uses only additions and subtractions of sets, according to this lemma if one takes at random a point of each mass set and one combines them using the DSm rules for scalars, the resulting point will belong to the resulting set from the fusion of mass sets using the DSm rules for sets.

Lemma 3: Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ and $K \geq 2$ independent sources of information, and $d = \dim(D^\Theta)$. By combination of incomplete information in DSmT, one gets incomplete information.

Proof: Suppose the masses of the sources of information on D^Θ are for all $1 \leq j \leq K$, represented by the mass-vector $\mathbf{m}_j = [m_{j_1}, m_{j_2}, \ldots, m_{j_d}]$ with $0 \leq \sum_{r=1}^{d} m_{j_r} < 1$. According to the DSm network architecture, no matter what DSm rule of combination is applied (classic or hybrid), the sum of all resulted masses has the form:

$$\prod_{j=1}^{K}(m_{j_1} + m_{j_2} + \ldots + m_{j_d}) < (\underbrace{1 \times 1 \times \ldots \times 1}_{K \text{ times}}) = 1 \tag{6.15}$$

Lemma 4: By combination of paraconsistent information, one gets paraconsistent information.

Proof: Using the same notations and similar reasoning, one has for all $1 \leq j \leq K$, $\mathbf{m}_j = [m_{j_1}, m_{j_2}, \ldots, m_{j_d}]$, with $\sum_{r=1}^{d} m_{j_r} > 1$. Then

$$\prod_{j=1}^{K}(m_{j_1} + m_{j_2} + \ldots + m_{j_d}) > (\underbrace{1 \times 1 \times \ldots \times 1}_{K \text{ times}}) = 1$$

Lemma 5: Combining incomplete (sum of masses < 1) with complete (sum of masses $= 1$) information, one gets incomplete information.

Lemma 6: Combining complete information, one gets complete information.

Remark: Combining incomplete with paraconsistent (sum of masses > 1) information can give any result. For example:

- If the sum of masses of the first source is 0.99 (incomplete) and the sum of masses of the second source is 1.01 (paraconsistent), then the sum of resulted masses is $0.99 \times 1.01 = 0.9999$ (i.e. incomplete)

- But if the first is 0.9 (incomplete) and the second is 1.2 (paraconsistent), then the resulted sum of masses is $0.9 \times 1.2 = 1.08$ (i.e. paraconsistent).

We can also have: incomplete information fusionned with paraconsistent information and get complete information. For example: $0.8 \times 1.25 = 1$.

Admissibility condition:

An *imprecise mass* on D^Θ is considered *admissible* if there exist at least a point belonging to $[0,1]$ in each mass set such that the sum of these points is equal to 1 (i.e. complete information for at least a group of selected points).

Remark: A complete scalar information is admissible. Of course, for the incomplete scalar information and paraconsistent scalar information there can not be an admissibility condition, because by definitions the masses of these two types of informations do not add up to 1 (i.e. to the complete information).

Theorem of Admissibility:

Let a frame $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, with $n \geq 2$, its hyper-power set D^Θ with $\dim(D^\Theta) = d$, and $K \geq 2$ sources of information providing imprecise admissible masses on D^Θ. Then, the resulted mass, after fusion of the imprecise masses of these sources of information with the DSm rules of combination, is also admissible.

Proof: Let s_j, $1 \leq j \leq K$, be an imprecise source of information, and its imprecise admissible mass $\mathbf{m}_j^I = [m_{j_1}^I, m_{j_2}^I, \dots, m_{j_d}^I]$. We underline that all $m_{j_r}^I$, for $1 \leq r \leq d$, are sets (not scalars); if there is a scalar α, we treat it as a set $[\alpha, \alpha]$. Because $\mathbf{m}_j^I$ is admissible, there exist the points (scalars in $[0,1]$) $m_{j_1}^s \in m_{j_1}^I$, $m_{j_2}^s \in m_{j_2}^I$,$\dots$,$m_{j_d}^s \in m_{j_d}^I$ such that $\sum_{r=1}^d m_{j_1}^s = 1$. This property occurs for all sources of information, thus there exist such points $m_{j_r}^s$ for any $1 \leq j \leq K$ and any $1 \leq r \leq d$. Now, if we fusion, as a particular case, the masses of only these points, using DSm classic or hybrid rules, and according to lemmas, based on DSm network architecture, one gets complete information (i.e. sum of masses equals to 1). See also Lemma 2.

6.5.3 An example with multiple-interval masses

We present here a more general example with multiple-interval masses. For simplicity, this example is a particular case when the theorem of admissibility is verified by a few points, which happen to be just on the bounders. More general and complex examples (not reported here due to space limitations), can be given and verified as well. It is however an extreme example, because we tried to comprise all kinds of possibilities which may occur in the imprecise or very imprecise fusion. So, let's consider a fusion problem over $\Theta = \{\theta_1, \theta_2\}$, two independent sources of information with the following imprecise admissible belief assignments

$A \in D^\Theta$	$m_1^I(A)$	$m_2^I(A)$
θ_1	$[0.1, 0.2] \cup \{0.3\}$	$[0.4, 0.5]$
θ_2	$(0.4, 0.6) \cup [0.7, 0.8]$	$[0, 0.4] \cup \{0.5, 0.6\}$

Table 6.9: Inputs of the fusion with imprecise bba

Using the DSm classic rule for sets, one gets

$$\begin{aligned}
m^I(\theta_1) &= ([0.1, 0.2] \cup \{0.3\}) \boxdot [0.4, 0.5] \\
&= ([0.1, 0.2] \boxdot [0.4, 0.5]) \cup (\{0.3\} \boxdot [0.4, 0.5]) \\
&= [0.04, 0.10] \cup [0.12, 0.15]
\end{aligned}$$

$$\begin{aligned}
m^I(\theta_2) &= ((0.4, 0.6) \cup [0.7, 0.8]) \boxdot ([0, 0.4] \cup \{0.5, 0.6\}) \\
&= ((0.4, 0.6) \boxdot [0, 0.4]) \cup ((0.4, 0.6) \boxdot \{0.5, 0.6\}) \cup ([0.7, 0.8] \boxdot [0, 0.4]) \cup ([0.7, 0.8] \boxdot \{0.5, 0.6\}) \\
&= (0, 0.24) \cup (0.20, 0.30) \cup (0.24, 0.36) \cup [0, 0.32] \cup [0.35, 0.40] \cup [0.42, 0.48] \\
&= [0, 0.40] \cup [0.42, 0.48]
\end{aligned}$$

$$\begin{aligned}
m^I(\theta_1 \cap \theta_2) &= [([0.1, 0.2] \cup \{0.3\}) \boxdot ([0, 0.4] \cup \{0.5, 0.6\})] \boxplus [[0.4, 0.5] \boxdot ((0.4, 0.6) \cup [0.7, 0.8])] \\
&= [([0.1, 0.2] \boxdot [0, 0.4]) \cup ([0.1, 0.2] \boxdot \{0.5, 0.6\}) \cup (\{0.3\} \boxdot [0, 0.4]) \cup (\{0.3\} \boxdot \{0.5, 0.6\})] \\
&\quad \boxplus [([0.4, 0.5] \boxdot (0.4, 0.6)) \cup ([0.4, 0.5] \boxdot [0.7, 0.8])] \\
&= [[0, 0.08] \cup [0.05, 0.10] \cup [0.06, 0.12] \cup [0, 0.12] \cup \{0.15, 0.18\}] \boxplus [(0.16, 0.30) \cup [0.28, 0.40]] \\
&= [[0, 0.12] \cup \{0.15, 0.18\}] \boxplus (0.16, 0.40] \\
&= (0.16, 0.52] \cup (0.31, 0.55] \cup (0.34, 0.58] \\
&= (0.16, 0.58]
\end{aligned}$$

Hence finally the fusion admissible result is given by:

$A \in D^\Theta$	$m^I(A) = [m_1^I \oplus m_2^I](A)$
θ_1	$[0.04, 0.10] \cup [0.12, 0.15]$
θ_2	$[0, 0.40] \cup [0.42, 0.48]$
$\theta_1 \cap \theta_2$	$(0.16, 0.58]$
$\theta_1 \cup \theta_2$	0

Table 6.10: Fusion result with the DSm classic rule

If one finds out that $\theta_1 \cap \theta_2 \overset{\mathcal{M}}{\equiv} \emptyset$ (this is our hybrid model $\mathcal{M}$ one wants to deal with), then one uses the hybrid DSm rule for sets (6.11): $m^I_{\mathcal{M}}(\theta_1 \cap \theta_2) = 0$ and $m^I_{\mathcal{M}}(\theta_1 \cup \theta_2) = (0.16, 0.58]$, the others imprecise masses are not changed. In other words, one gets now with hybrid DSm rule applied to imprecise beliefs:

$A \in D^\Theta$	$m^I_{\mathcal{M}}(A) = [m^I_1 \oplus m^I_2](A)$
θ_1	$[0.04, 0.10] \cup [0.12, 0.15]$
θ_2	$[0, 0.40] \cup [0.42, 0.48]$
$\theta_1 \cap \theta_2 \overset{\mathcal{M}}{\equiv} \emptyset$	0
$\theta_1 \cup \theta_2$	$(0.16, 0.58]$

Table 6.11: Fusion result with the hybrid DSm rule for $\mathcal{M}$

Let's check now the admissibility conditions and theorem. For the source 1, there exist the precise masses $(m_1(\theta_1) = 0.3) \in ([0.1, 0.2] \cup \{0.3\})$ and $(m_1(\theta_2) = 0.7) \in ((0.4, 0.6) \cup [0.7, 0.8])$ such that $0.3 + 0.7 = 1$. For the source 2, there exist the precise masses $(m_1(\theta_1) = 0.4) \in ([0.4, 0.5])$ and $(m_2(\theta_2) = 0.6) \in ([0, 0.4] \cup \{0.5, 0.6\})$ such that $0.4 + 0.6 = 1$. Therefore both sources associated with $m^I_1(.)$ and $m^I_2(.)$ are admissible imprecise sources of information.

It can be easily checked that the DSm classic fusion of $m_1(.)$ and $m_2(.)$ yields the paradoxical basic belief assignment $m(\theta_1) = [m_1 \oplus m_2](\theta_1) = 0.12$, $m(\theta_2) = [m_1 \oplus m_2](\theta_2) = 0.42$ and $m(\theta_1 \cap \theta_2) = [m_1 \oplus m_2](\theta_1 \cap \theta_2) = 0.46$. One sees that the admissibility theorem is satisfied since $(m(\theta_1) = 0.12) \in (m^I(\theta_1) = [0.04, 0.10] \cup [0.12, 0.15])$, $(m(\theta_2) = 0.42) \in (m^I(\theta_2) = [0, 0.40] \cup [0.42, 0.48])$ and $(m(\theta_1 \cap \theta_2) = 0.46) \in (m^I(\theta_1 \cap \theta_2) = (0.16, 0.58])$ such that $0.12 + 0.42 + 0.46 = 1$. Similarly if one finds out that $\theta_1 \cap \theta_2 = \emptyset$, then one uses the hybrid DSm rule and one gets: $m(\theta_1 \cap \theta_2) = 0$ and $m(\theta_1 \cup \theta_2) = 0.46$; the others remain unchanged. The admissibility theorem still holds.

6.6 Conclusion

In this chapter, we proposed from the DSmT framework, a new general approach to combine, imprecise, uncertain and possibly paradoxical sources of information to cover a wider class of fusion problems. This work was motivated by the fact that in most of practical and real fusion problems, the information is rarely known with infinite precision and the admissible belief assignment masses, for each element of the hyper-power set of the problem, have to be taken/chosen more reasonably as sub-unitary (or as a set of sub-unitary) intervals rather than a pure and simple scalar values. This is a generalization of previous available works proposed in literature (mainly IBS restricted to TBM framework). One showed that it is possible to fusion directly interval-valued masses using the DSm rules (classic or hybrid ones) and the operations on sets defined in this work. Several illustrative and didactic examples have been given throughout this chapter to show the application of this new approach. The method developed here can also combine incomplete and paraconsistent imprecise, uncertain and paradoxical sources of information as well. This approach (although focused here only on the derivation of imprecise basic belief assignments) can be extended without difficulty to the derivation of imprecise belief and plausibility functions as well as to imprecise pignistic probabilities according to the generalized pignistic transformation presented in chapter 7. This work allows the DSmT to cover a wider class of fusion problems.

6.7 References

[1] Denœux T, *Reasoning with imprecise belief structures*, Technical Report Heudiasys 97/44, available at `http://www.hds.utc.fr/~tdenoeux/`.

[2] Denœux T, *Reasoning with imprecise belief structures*, International Journal of Approximate Reasoning, 20, pp. 79-111, 1999.

[3] Hayes B., *A lucid interval*, American Scientist, Vol. 91, No. 6, Nov.-Dec. 2003, pp. 484-488.

[4] Jaulin E., Kieffer M., Didrit O., Walter E., *Applied Interval Analysis*, Springer Verlag, 384 pages, Dec. 2001.

[5] Lamata M., Moral S., *Calculus with linguistic probabilities and beliefs*, In R. R. Yager, M. Fedrizzi, and J. Kacprzyk, editors, Advances in Dempster-Shafer Theory of Evidence, pp. 133-152, Wiley.

[6] Moore R. E., *Interval Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1966.

[7] Moore R. E., *Methods and Applications of Interval Analysis*, Philadelphia: Society for Industrial and Applied Mathematics, 1979.

[8] Neumaier A., *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990.

[9] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[10] SIGLA/X Group, *Modal Interval Analysis: An Introduction*, `http://ima.udg.es/SIGLA/X/mod_interval/`.

[11] Smarandache F. (Editor), *Proceedings of the First International Conference on Neutrosophics*, Univ. of New Mexico, Gallup Campus, NM, USA, 1-3 Dec. 2001, Xiquan, Phoenix, 2002.

[12] Smets Ph., Kennes R., *The Transferable Belief Model*, Artificial Intelligence, 66, pp. 191-243, 1994.

[13] Smets Ph., *Data Fusion in the Transferable Belief Model*, Proc. of Third Int. Conf. on Information Fusion, Fusion 2000, Paris, July 2000.

Chapter 7

A Generalized Pignistic Transformation

Jean Dezert
ONERA
29 Av. de la Division Leclerc
92320 Châtillon
France

Florentin Smarandache
Department of Mathematics
University of New Mexico
Gallup, NM 8730
U.S.A.

Milan Daniel[1]
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, CZ - 182 07 Prague 8
Czech Republic

Abstract: *This chapter introduces a generalized pignistic transformation (GPT) developed in the DSmT framework as a tool for decision-making at the pignistic level. The GPT allows to construct quite easily a subjective probability measure from any generalized basic belief assignment provided by any corpus of evidence. We focus our presentation on the 3D case and we provide the full result obtained by the proposed GPT and its validation drawn from the probability theory.*

This chapter is based on a paper [3] and is reproduced here with permission of the International Society of Information Fusion. 1. Milan Daniel thanks the COST action 274 TARSKI for supporting this work.

7.1 A short introduction to the DSm cardinality

One important notion involved in the definition of the Generalized Pignistic Transformation (GPT) is the *DSm cardinality* introduced in chapter 3 (section 3.2.2) and in [1]. The *DSm cardinality* of any element A of hyper-power set D^Θ, denoted $\mathcal{C}_\mathcal{M}(A)$, corresponds to the number of parts of A in the corresponding fuzzy/vague Venn diagram of the problem (model $\mathcal{M}$) taking into account the set of integrity constraints (if any), i.e. all the possible intersections due to the nature of the elements θ_i. This *intrinsic cardinality* depends on the model $\mathcal{M}$ (free, hybrid or Shafer's model). $\mathcal{M}$ is the model that contains A, which depends both on the dimension $n = |\Theta|$ and on the number of non-empty intersections present in its associated Venn diagram. The DSm cardinality depends on the cardinal of $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ and on the model of D^Θ (i.e., the number of intersections and between what elements of Θ - in a word the structure) at the same time; it is not necessarily that every singleton, say θ_i, has the same DSm cardinal, because each singleton has a different structure; if its structure is the simplest (no intersection of this elements with other elements) then $\mathcal{C}_\mathcal{M}(\theta_i) = 1$, if the structure is more complicated (many intersections) then $\mathcal{C}_\mathcal{M}(\theta_i) > 1$; let's consider a singleton θ_i: if it has 1 intersection only then $\mathcal{C}_\mathcal{M}(\theta_i) = 2$, for 2 intersections only $\mathcal{C}_\mathcal{M}(\theta_i)$ is 3 or 4 depending on the model $\mathcal{M}$, for m intersections it is between $m+1$ and 2^m depending on the model; the maximum DSm cardinality is 2^{n-1} and occurs for $\theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$ in the free model $\mathcal{M}^f$; similarly for any set from D^Θ: the more complicated structure it has, the bigger is the DSm cardinal; thus the DSm cardinality measures the complexity of en element from D^Θ, which is a nice characterization in our opinion; we may say that for the singleton θ_i not even $|\Theta|$ counts, but only its structure (= how many other singletons intersect θ_i). Simple illustrative examples have already been presented in chapter 3. One has $1 \leq \mathcal{C}_\mathcal{M}(A) \leq 2^n - 1$. $\mathcal{C}_\mathcal{M}(A)$ must not be confused with the classical cardinality $|A|$ of a given set A (i.e. the number of its distinct elements) - that's why a new notation is necessary here.

It has been shown in [1], that $\mathcal{C}_\mathcal{M}(A)$, is exactly equal to the sum of the elements of the row of $\mathbf{D}_n$ corresponding to proposition A in the $\mathbf{u}_n$ basis (see chapter 2). Actually $\mathcal{C}_\mathcal{M}(A)$ is very easy to compute by programming from the algorithm of generation of D^Θ given in chapter 2 and in [2].

If one imposes a constraint that a set B from D^Θ is empty (i.e. we choose a hybrid DSm model), then one suppresses the columns corresponding to the parts which compose B in the matrix $\mathbf{D}_n$ and the row of B and the rows of all elements of D^Θ which are subsets of B, getting a new matrix $\mathbf{D'}_n$ which represents a new hybrid DSm model $\mathcal{M}'$. In the $\mathbf{u}_n$ basis, one similarly suppresses the parts that form B, and now this basis has the dimension $2^n - 1 - \mathcal{C}_\mathcal{M}(B)$.

7.2 The Classical Pignistic Transformation (CPT)

We follow here Smets' point of view [8] about the assumption that beliefs manifest themselves at two mental levels: the *credal* level where beliefs are entertained and the *pignistic* level where belief functions are used to make decisions. Pignistic terminology has been coined by Philippe Smets and comes from *pignus*, a bet in Latin. The probability functions, usually used to quantify the beliefs at both levels, are actually used here only to quantify the uncertainty when a decision is really necessary, otherwise we argue as Philippe Smets does, that beliefs are represented by belief functions. To take a rational decision, we propose to transform generalized beliefs into pignistic probability functions through the Generalized Pignistic Transformation (the GPT) which will be presented in the following. We first recall the Classical Pignistic Transformation (the CPT) based on Dempster-Shafer Theory (DST) and then we generalize it within the Dezert-Smarandache Theory (DSmT) framework.

When a decision must be taken, we use the expected utility theory which requires to construct a probability function $P\{.\}$ from basic belief function $m(.)$ [8]. This is achieved by the so-called classical Pignistic Transformation. In the Transferable Belief Model (the TBM) context [7] with open-world assumption, Philippe Smets derives the pignistic probabilities from any non normalized basic belief assignment $m(.)$ (i.e. for which $m(\emptyset) \geq 0$) by the following formula [8]:

$$P\{A\} = \sum_{X \subseteq \Theta} \frac{|X \cap A|}{|X|} \frac{m(X)}{1 - m(\emptyset)} \tag{7.1}$$

where $|A|$ denotes the number of worlds in the set A (with convention $|\emptyset|/|\emptyset| = 1$, to define $P\{\emptyset\}$). $P\{A\}$ corresponds to $BetP(A)$ in Smets' notation [8]. Decisions are achieved by computing the expected utilities of the acts using the subjective/pignistic $P\{.\}$ as the probability function needed to compute expectations. Usually, one uses the maximum of the pignistic probability as decision criterion. The max. of $P\{.\}$ is often considered as a prudent betting decision criterion between the two other alternatives (max of plausibility or max. of credibility which appears to be respectively too optimistic or too pessimistic). It is easy to show that $P\{.\}$ is indeed a probability function (see [7]).

It is important to note that if the belief mass $m(.)$ results from the combination of two independent sources of evidence (i.e. $m(.) = [m_1 \oplus m_2](.)$) then, at the pignistic level, the classical pignistic probability measure $P(.)$ remains the same when using Dempster's rule or when using Smets' rule in his TBM open-world approach working with $m(\emptyset) > 0$. Thus the problem arising with the combination of highly conflicting sources when using Dempster's rule (see chapter 5), and apparently circumvented with the TBM at the credal level, still fundamentally remains at the pignistic level. The problem is only transferred from credal level to pignistic level when using TBM. TBM does not help to improve the reliability of the decision-making with respect to Dempster's rule of combination because the pignistic probabilities are strictly and mathematically equivalent. In other words, if the result of the combination is wrong or at least very questionable or counter-intuitive when the degree of the conflict $m(\emptyset)$ becomes high, then the decision based on pignistic probabilities will become inevitably wrong or very questionable too.

Taking into account the previous remark, we rather prefer to adopt from now on the classical Shafer's definition for basic belief assignment $m(.) : 2^\Theta \to [0,1]$ which imposes to take $m(\emptyset) = 0$ and $\sum_{X\in 2^\Theta} m(X) = 1$. We adopt therefore the following definition for the Classical Pignistic Transformation (CPT):

$$P\{A\} = \sum_{X\in 2^\Theta} \frac{|X\cap A|}{|X|} m(X) \tag{7.2}$$

7.3 A Generalized Pignistic Transformation (GPT)

7.3.1 Definition

To take a rational decision within the DSmT framework, it is necessary to generalize the Classical Pignistic Transformation in order to construct a pignistic probability function from any generalized basic belief assignment $m(.)$ drawn from the DSm rules of combination (the classic or the hybrid ones - see chapter 1). We propose here the simplest and direct extension of the CPT to define a Generalized Pignistic Transformation as follows:

$$\forall A \in D^\Theta, \qquad P\{A\} = \sum_{X\in D^\Theta} \frac{\mathcal{C}_\mathcal{M}(X\cap A)}{\mathcal{C}_\mathcal{M}(X)} m(X) \tag{7.3}$$

where $\mathcal{C}_\mathcal{M}(X)$ denotes the DSm cardinal of proposition X for the DSm model $\mathcal{M}$ of the problem under consideration.

The decision about the solution of the problem is usually taken by the maximum of pignistic probability function $P\{.\}$. Let's remark the close ressemblance of the two pignistic transformations (7.2) and (7.3). It can be shown that (7.3) reduces to (7.2) when the hyper-power set D^Θ reduces to classical power set 2^Θ if we adopt Shafer's model. But (7.3) is a generalization of (7.2) since it can be used for computing pignistic probabilities for any models (including Shafer's model).

7.3.2 $P\{.\}$ is a probability measure

It is important to prove that $P\{.\}$ built from GPT is indeed a (subjective/pignistic) probability measure satisfying the following axioms of probability theory [4, 5]:

- **Axiom 1** (nonnegativity): The (generalized pignistic) probability of any event A is bounded by 0 and 1

$$0 \leq P\{A\} \leq 1$$

- **Axiom 2** (unity): Any sure event (the sample space) has unity (generalized pignistic) probability

$$P\{\mathcal{S}\} = 1$$

- **Axiom 3** (additivity over mutually exclusive events): If A, B are disjoint (i.e. $A \cap B = \emptyset$) then $P(A \cup B) = P(A) + P(B)$

The axiom 1 is satisfied because, by the definition of the generalized basic belief assignment $m(.)$, one has $\forall \alpha_i \in D^\Theta$, $0 \leq m(\alpha_i) \leq 1$ with $\sum_{\alpha_i \in D^\Theta} m(\alpha_i) = 1$ and since all coefficients involved within GPT are bounded by 0 and 1, it follows directly that pignistic probabilities are also bounded by 0 and 1.

The axiom 2 is satisfied because all the coefficients involved in the sure event $\mathcal{S} \triangleq \theta_1 \cup \theta_2 \cup ... \cup \theta_n$ are equal to one because $\mathcal{C}_\mathcal{M}(X \cap \mathcal{S})/\mathcal{C}_\mathcal{M}(X) = \mathcal{C}_\mathcal{M}(X)/\mathcal{C}_\mathcal{M}(X) = 1$, so that $P\{\mathcal{S}\} \equiv \sum_{\alpha_i \in D^\Theta} m(\alpha_i) = 1$.

The axiom 3 is satisfied. Indeed, from the definition of GPT, one has

$$P\{A \cup B\} = \sum_{X \in D^\Theta} \frac{\mathcal{C}_\mathcal{M}(X \cap (A \cup B))}{\mathcal{C}_\mathcal{M}(X)} m(X) \tag{7.4}$$

But if we consider A and B exclusive (i.e. $A \cap B = \emptyset$), then it follows:

$$\mathcal{C}_\mathcal{M}(X \cap (A \cup B)) = \mathcal{C}_\mathcal{M}((X \cap A) \cup (X \cap B)) = \mathcal{C}_\mathcal{M}(X \cap A) + \mathcal{C}_\mathcal{M}(X \cap B)$$

By substituting $\mathcal{C}_\mathcal{M}(X \cap (A \cup B))$ by $\mathcal{C}_\mathcal{M}(X \cap A) + \mathcal{C}_\mathcal{M}(X \cap B)$ into (7.4), it comes:

$$\begin{aligned}
P\{A \cup B\} &= \sum_{X \in D^\Theta} \frac{\mathcal{C}_\mathcal{M}(X \cap A) + \mathcal{C}_\mathcal{M}(X \cap B)}{\mathcal{C}_\mathcal{M}(X)} m(X) \\
&= \sum_{X \in D^\Theta} \frac{\mathcal{C}_\mathcal{M}(X \cap A)}{\mathcal{C}_\mathcal{M}(X)} m(X) + \sum_{X \in D^\Theta} \frac{\mathcal{C}_\mathcal{M}(X \cap B)}{\mathcal{C}_\mathcal{M}(X)} m(X) \\
&= P\{A\} + P\{B\}
\end{aligned}$$

which completes the proof. From the coefficients $\frac{\mathcal{C}_\mathcal{M}(X \cap A)}{\mathcal{C}_\mathcal{M}(X)}$ involved in (7.3), it can also be easily checked that $A \subset B \Rightarrow P\{A\} \leq P\{B\}$. One can also easily prove the Poincaré' equality: $P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}$ because $\mathcal{C}_\mathcal{M}(X \cap (A \cup B) = \mathcal{C}_\mathcal{M}((X \cap A) \cup (X \cap B)) = \mathcal{C}_\mathcal{M}(X \cap A) + \mathcal{C}_\mathcal{M}(X \cap B) - \mathcal{C}_\mathcal{M}(X \cap (A \cap B))$ (one has substracted $\mathcal{C}_\mathcal{M}(X \cap (A \cap B))$, i.e. the number of parts of $X \cap (A \cap B)$ in the Venn diagram, due to the fact that these parts were added twice: once in $\mathcal{C}_\mathcal{M}(X \cap A)$ and second time in $\mathcal{C}_\mathcal{M}(X \cap B)$.

7.4 Some examples for the GPT

7.4.1 Example for the 2D case

- **With the free DSm model**:

Let's consider $\Theta = \{\theta_1, \theta_2\}$ and the generalized basic belief function $m(.)$ over the hyper-power set $D^\Theta = \{\emptyset, \theta_1 \cap \theta_2, \theta_1, \theta_2, \theta_1 \cup \theta_2\}$. It is easy to construct the pignistic probability $P\{.\}$. According to the definition of the GPT given in (7.3), one gets:

$$\begin{aligned}
P\{\emptyset\} &= 0\\
P\{\theta_1\} &= m(\theta_1) + \frac{1}{2}m(\theta_2) + m(\theta_1 \cap \theta_2) + \frac{2}{3}m(\theta_1 \cup \theta_2)\\
P\{\theta_2\} &= m(\theta_2) + \frac{1}{2}m(\theta_1) + m(\theta_1 \cap \theta_2) + \frac{2}{3}m(\theta_1 \cup \theta_2)\\
P\{\theta_1 \cap \theta_2\} &= \frac{1}{2}m(\theta_2) + \frac{1}{2}m(\theta_1) + m(\theta_1 \cap \theta_2) + \frac{1}{3}m(\theta_1 \cup \theta_2)\\
P\{\theta_1 \cup \theta_2\} &= P\{\Theta\} = m(\theta_1) + m(\theta_2) + m(\theta_1 \cap \theta_2) + m(\theta_1 \cup \theta_2) = 1
\end{aligned}$$

It is easy to prove that $0 \leq P\{.\} \leq 1$ and $P\{\theta_1 \cup \theta_2\} = P\{\theta_1\} + P\{\theta_2\} - P\{\theta_1 \cap \theta_2\}$

- **With Shafer's model**:

If one adopts Shafer's model (we assume $\theta_1 \cap \theta_2 \overset{\mathcal{M}^0}{\equiv} \emptyset$), then after applying the hybrid DSm rule of combination, one gets a basic belief function with non null masses only on θ_1, θ_2 and $\theta_1 \cup \theta_2$. By applying the GPT, one gets:

$$P\{\emptyset\} = 0$$

$$P\{\theta_1 \cap \theta_2\} = 0$$

$$P\{\theta_1\} = m(\theta_1) + \frac{1}{2}m(\theta_1 \cup \theta_2)$$

$$P\{\theta_2\} = m(\theta_2) + \frac{1}{2}m(\theta_1 \cup \theta_2)$$

$$P\{\theta_1 \cup \theta_2\} = m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) = 1$$

which naturally corresponds in this case to the pignistic probability built with the classical pignistic transformation (7.2).

7.4.2 Example for the 3D case

- **With the free DSm model**:

X	$\frac{\mathcal{C}_{\mathcal{M}}(X\cap\alpha_6)}{\mathcal{C}_{\mathcal{M}}(X)}$	$\leq$	$\frac{\mathcal{C}_{\mathcal{M}}(X\cap\alpha_{10})}{\mathcal{C}_{\mathcal{M}}(X)}$	X	$\frac{\mathcal{C}_{\mathcal{M}}(X\cap\alpha_6)}{\mathcal{C}_{\mathcal{M}}(X)}$	$\leq$	$\frac{\mathcal{C}_{\mathcal{M}}(X\cap\alpha_{10})}{\mathcal{C}_{\mathcal{M}}(X)}$
α_1	1	$\leq$	1	α_{10}	(3/4)	$\leq$	1
α_2	1	$\leq$	1	α_{11}	(2/4)	$\leq$	(2/4)
α_3	(1/2)	$\leq$	(1/2)	α_{12}	(3/5)	$\leq$	(3/5)
α_4	1	$\leq$	1	α_{13}	(3/5)	$\leq$	(4/5)
α_5	(2/3)	$\leq$	(2/3)	α_{14}	(3/5)	$\leq$	(3/5)
α_6	1	$\leq$	1	α_{15}	(3/6)	$\leq$	(4/6)
α_7	(2/3)	$\leq$	(2/3)	α_{16}	(3/6)	$\leq$	(3/6)
α_8	(3/4)	$\leq$	(3/4)	α_{17}	(3/6)	$\leq$	(4/6)
α_9	(2/4)	$\leq$	(2/4)	α_{18}	(3/7)	$\leq$	(4/7)

Table 7.1: Coefficients $\frac{\mathcal{C}_{\mathcal{M}}(X\cap\alpha_6)}{\mathcal{C}_{\mathcal{M}}(X)}$ and $\frac{\mathcal{C}_{\mathcal{M}}(X\cap\alpha_{10})}{\mathcal{C}_{\mathcal{M}}(X)}$

Let's consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$, its hyper-power set $D^\Theta = \{\alpha_0, \dots, \alpha_{18}\}$ (with α_i, $i = 0, \dots, 18$ corresponding to propositions shown in table 3.1 of chapter 3, and the generalized basic belief assignment $m(.)$ over the hyper-power set D^Θ. The six tables presented in the appendix show the full derivations of all generalized pignistic probabilities $P\{\alpha_i\}$ for $i = 1, \dots, 18$ ($P\{\emptyset\} = 0$ by definition) according to the GPT formula (7.3).

Note that $P\{\alpha_{18}\} = 1$ because $(\theta_1 \cup \theta_2 \cup \theta_3)$ corresponds to the sure event in our subjective probability space and $\sum_{\alpha_i \in D^\Theta} m(\alpha_i) = 1$ by the definition of any generalized basic belief assignment $m(.)$ defined on D^Θ.

It can be verified (as expected) on this example, although being a quite tedious task, that Poincaré' s equality holds:

$$P\{A_1 \cup \ldots \cup A_n\} = \sum_{\substack{I \subset \{1,\ldots,n\} \\ I \neq \emptyset}} (-1)^{|I|+1} P\{\bigcap_{i \in I} A_i\} \tag{7.5}$$

It is also easy to verify that $\forall A \subset B \Rightarrow P\{A\} \leq P\{B\}$ holds. By example, for $(\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2) \subset \alpha_{10} \triangleq \theta_2)$ and from the expressions of $P\{\alpha_6\}$ and $P\{\alpha_{10}\}$ given in appendix, we directly conclude that $P\{\alpha_6\} \leq P\{\alpha_{10}\}$ because

$$\forall X \in D^\Theta, \qquad \frac{\mathcal{C}_{\mathcal{M}}(X \cap \alpha_6)}{\mathcal{C}_{\mathcal{M}}(X)} \leq \frac{\mathcal{C}_{\mathcal{M}}(X \cap \alpha_{10})}{\mathcal{C}_{\mathcal{M}}(X)} \tag{7.6}$$

as shown in the table above.

- **Example with a given hybrid DSm model**:

Consider now the hybrid DSm model $\mathcal{M} \neq \mathcal{M}^f$ in which we force all possible conjunctions to be empty, but $\theta_1 \cap \theta_2$ according to the second Venn diagram presented in Chapter 3 and shown in Figure 3.2. In this case the hyper-power set D^Θ reduces to 9 elements $\{\alpha_0, \ldots, \alpha_8\}$ shown in table 3.2 of Chapter 3. The following tables present the full derivations of the pignistic probabilities $P\{\alpha_i\}$ for $i = 1, \ldots, 8$ from the GPT formula (7.3) applied to this hybrid DSm model.

$P\{\alpha_1\} =$	$P\{\alpha_2\} =$	$P\{\alpha_3\} =$	$P\{\alpha_4\} =$
$(1/1)m(\alpha_1)$	$(0/1)m(\alpha_1)$	$(1/1)m(\alpha_1)$	$(1/1)m(\alpha_1)$
$+(0/1)m(\alpha_2)$	$+(1/1)m(\alpha_2)$	$+(0/2)m(\alpha_2)$	$+(0/1)m(\alpha_2)$
$+(1/2)m(\alpha_3)$	$+(0/2)m(\alpha_3)$	$+(2/2)m(\alpha_3)$	$+(1/2)m(\alpha_3)$
$+(1/2)m(\alpha_4)$	$+(0/2)m(\alpha_4)$	$+(1/2)m(\alpha_4)$	$+(2/2)m(\alpha_4)$
$+(1/3)m(\alpha_5)$	$+(0/3)m(\alpha_5)$	$+(2/3)m(\alpha_5)$	$+(2/3)m(\alpha_5)$
$+(1/3)m(\alpha_6)$	$+(1/3)m(\alpha_6)$	$+(2/3)m(\alpha_6)$	$+(1/3)m(\alpha_6)$
$+(1/3)m(\alpha_7)$	$+(1/3)m(\alpha_7)$	$+(1/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$
$+(1/4)m(\alpha_8)$	$+(1/4)m(\alpha_8)$	$+(2/4)m(\alpha_8)$	$+(2/4)m(\alpha_8)$

Table 7.2: Derivation of $P\{\alpha_1 \triangleq \theta_1 \cap \theta_2\}$, $P\{\alpha_2 \triangleq \theta_3\}$, $P\{\alpha_3 \triangleq \theta_1\}$ and $P\{\alpha_4 \triangleq \theta_2\}$

$P\{\alpha_5\} =$	$P\{\alpha_6\} =$	$P\{\alpha_7\} =$	$P\{\alpha_8\} =$
$(1/1)m(\alpha_1)$	$(1/1)m(\alpha_1)$	$(1/1)m(\alpha_1)$	$(1/1)m(\alpha_1)$
$+(0/1)m(\alpha_2)$	$+(1/1)m(\alpha_2)$	$+(2/2)m(\alpha_2)$	$+(2/2)m(\alpha_2)$
$+(2/2)m(\alpha_3)$	$+(2/2)m(\alpha_3)$	$+(1/2)m(\alpha_3)$	$+(2/2)m(\alpha_3)$
$+(2/2)m(\alpha_4)$	$+(1/2)m(\alpha_4)$	$+(2/2)m(\alpha_4)$	$+(2/2)m(\alpha_4)$
$+(3/3)m(\alpha_5)$	$+(2/3)m(\alpha_5)$	$+(2/3)m(\alpha_5)$	$+(3/3)m(\alpha_5)$
$+(2/3)m(\alpha_6)$	$+(3/3)m(\alpha_6)$	$+(2/3)m(\alpha_6)$	$+(3/3)m(\alpha_6)$
$+(2/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$	$+(3/3)m(\alpha_7)$	$+(3/3)m(\alpha_7)$
$+(3/4)m(\alpha_8)$	$+(3/4)m(\alpha_8)$	$+(3/4)m(\alpha_8)$	$+(4/4)m(\alpha_8)$

Table 7.3: Derivation of $P\{\alpha_5 \triangleq \theta_1 \cup \theta_2\}$, $P\{\alpha_6 \triangleq \theta_1 \cup \theta_3\}$, $P\{\alpha_7 \triangleq \theta_2 \cup \theta_3\}$ and $P\{\alpha_8 \triangleq \theta_1 \cup \theta_2 \cup \theta_3\}$

- **Example with Shafer's model**:

Consider now Shafer's model $\mathcal{M}^0 \neq \mathcal{M}^f$ in which we force all possible conjunctions to be empty according to the third Venn diagram presented in Chapter 3. In this case the hyper-power set

D^Θ reduces to the classical power set 2^Θ with 8 elements $\{\alpha_0, \ldots, \alpha_7\}$ explicated in table 3.3 of Chapter 3. Applying, the GPT formula (7.3), one gets the following pignistic probabilities $P\{\alpha_i\}$ for $i = 1, \ldots, 7$ which naturally coincide, in this particular case, with the values obtained directly by the classical pignistic transformation (7.2):

$P\{\alpha_1\} =$	$P\{\alpha_2\} =$	$P\{\alpha_3\} =$
$(1/1)m(\alpha_1)$	$(0/1)m(\alpha_1)$	$(0/1)m(\alpha_1)$
$+(0/1)m(\alpha_2)$	$+(1/1)m(\alpha_2)$	$+(0/1)m(\alpha_2)$
$+(0/1)m(\alpha_3)$	$+(0/1)m(\alpha_3)$	$+(1/1)m(\alpha_3)$
$+(1/2)m(\alpha_4)$	$+(1/2)m(\alpha_4)$	$+(0/2)m(\alpha_4)$
$+(1/2)m(\alpha_5)$	$+(0/2)m(\alpha_5)$	$+(1/2)m(\alpha_5)$
$+(0/2)m(\alpha_6)$	$+(1/2)m(\alpha_6)$	$+(1/2)m(\alpha_6)$
$+(1/3)m(\alpha_7)$	$+(1/3)m(\alpha_7)$	$+(1/3)m(\alpha_7)$

Table 7.4: Derivation of $P\{\alpha_1 \triangleq \theta_1\}$, $P\{\alpha_2 \triangleq \theta_2\}$ and $P\{\alpha_3 \triangleq \theta_3\}$

$P\{\alpha_4\} =$	$P\{\alpha_5\} =$	$P\{\alpha_6\} =$	$P\{\alpha_7\} =$
$(1/1)m(\alpha_1)$	$(1/1)m(\alpha_1)$	$(0/1)m(\alpha_1)$	$(1/1)m(\alpha_1)$
$+(1/1)m(\alpha_2)$	$+(0/1)m(\alpha_2)$	$+(1/1)m(\alpha_2)$	$+(1/1)m(\alpha_2)$
$+(0/1)m(\alpha_3)$	$+(1/1)m(\alpha_3)$	$+(1/1)m(\alpha_3)$	$+(1/1)m(\alpha_3)$
$+(2/2)m(\alpha_4)$	$+(1/2)m(\alpha_4)$	$+(1/2)m(\alpha_4)$	$+(2/2)m(\alpha_4)$
$+(1/2)m(\alpha_5)$	$+(2/2)m(\alpha_5)$	$+(1/2)m(\alpha_5)$	$+(2/2)m(\alpha_5)$
$+(1/2)m(\alpha_6)$	$+(1/2)m(\alpha_6)$	$+(2/2)m(\alpha_6)$	$+(2/2)m(\alpha_6)$
$+(2/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$	$+(3/3)m(\alpha_7)$

Table 7.5: Derivation of $P\{\alpha_4 \triangleq \theta_1 \cup \theta_2\}$, $P\{\alpha_5 \triangleq \theta_1 \cup \theta_3\}$, $P\{\alpha_6 \triangleq \theta_2 \cup \theta_3\}$ and $P\{\alpha_7 \triangleq \theta_1 \cup \theta_2 \cup \theta_3\} = 1$

7.5 Conclusion

A generalization of the classical pignistic transformation developed originally within the DST framework has been proposed in this chapter. This generalization is based on the new theory of plausible and paradoxical reasoning (DSmT) and provides a new mathematical issue to help the decision-making under uncertainty and paradoxical (i.e. highly conflicting) sources of information. The generalized pignistic transformation (GPT) proposed here allows to build a subjective/pignistic probability measure over the hyper-power set of the frame of the problem under consideration for all kinds of models (free, hybrid or Shafer's model). The GPT coincides naturally with the classical pignistic transformation whenever Shafer's model is adopted. It corresponds with the assumptions of classical pignistic probability generalized to the free DSm model. A relation of GPT on general hybrid DSm models to assumptions of classical PT is still in the process of investigation. Several examples for the 2D and 3D cases for different kinds of models have been presented to illustrate the validity of the GPT.

7.6 References

[1] Dezert J., Smarandache F., *Partial ordering of hyper-power sets and matrix representation of belief functions within DSmT*, Proc. of Fusion 2003 Conf., pp. 1230-1238, Cairns, Australia, July, 2003.

[2] Dezert J., Smarandache F., *On the generation of hyper-power sets for the DSmT*, Proceedings of the 6th Int. Conf. on Information Fusion, Cairns, Australia, pp. 1118-1125, July, 2003.

[3] Dezert J., Smarandache F., Daniel M., *The Generalized Pignistic Transformation*, Proceedings of 7th International Conference on Information Fusion, Fusion 2004, Stockholm, Sweden, June, 2004.

[4] Li X.-R., *Probability, Random Signals and Statistics*, CRC Press, 1999.

[5] Papoulis A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill series in electrical engineering (6th printing), 1989.

[6] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[7] Smets Ph., Kennes R., *The Transferable Belief Model*, Artificial Intelligence, 66, pp. 191-234, 1994.

[8] Smets Ph., *Data Fusion in the transferable Belief Model*, Proc. of 3rd Int. Conf. on Inf. Fusion, Fusion 2000, pp. PS21-PS33, Paris, France, July 10-13, 2000.

[9] Tchamova A., Semerdjiev T., Dezert J., *Estimation of Target Behavior Tendencies using Dezert-Smarandache Theory*, Proc. of Fusion 2003 Conf., pp. 1230-1238, Cairns, Australia, July 8-11, 2003.

Appendix: Derivation of the GPT for the 3D free DSm model

$P\{\alpha_1\} =$	$P\{\alpha_2\} =$	$P\{\alpha_3\} =$
$m(\alpha_1)$	$m(\alpha_1)$	$m(\alpha_1)$
$+(1/2)m(\alpha_2)$	$+m(\alpha_2)$	$+(1/2)m(\alpha_2)$
$+(1/2)m(\alpha_3)$	$+(1/2)m(\alpha_3)$	$+m(\alpha_3)$
$+(1/2)m(\alpha_4)$	$+(1/2)m(\alpha_4)$	$+(1/2)m(\alpha_4)$
$+(1/3)m(\alpha_5)$	$+(1/3)m(\alpha_5)$	$+(2/3)m(\alpha_5)$
$+(1/3)m(\alpha_6)$	$+(2/3)m(\alpha_6)$	$+(1/3)m(\alpha_6)$
$+(1/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$
$+(1/4)m(\alpha_8)$	$+(2/4)m(\alpha_8)$	$+(2/4)m(\alpha_8)$
$+(1/4)m(\alpha_9)$	$+(2/4)m(\alpha_9)$	$+(2/4)m(\alpha_9)$
$+(1/4)m(\alpha_{10})$	$+(2/4)m(\alpha_{10})$	$+(1/4)m(\alpha_{10})$
$+(1/4)m(\alpha_{11})$	$+(1/4)m(\alpha_{11})$	$+(2/4)m(\alpha_{11})$
$+(1/5)m(\alpha_{12})$	$+(2/5)m(\alpha_{12})$	$+(2/5)m(\alpha_{12})$
$+(1/5)m(\alpha_{13})$	$+(2/5)m(\alpha_{13})$	$+(2/5)m(\alpha_{13})$
$+(1/5)m(\alpha_{14})$	$+(2/5)m(\alpha_{14})$	$+(2/5)m(\alpha_{14})$
$+(1/6)m(\alpha_{15})$	$+(2/6)m(\alpha_{15})$	$+(2/6)m(\alpha_{15})$
$+(1/6)m(\alpha_{16})$	$+(2/6)m(\alpha_{16})$	$+(2/6)m(\alpha_{16})$
$+(1/6)m(\alpha_{17})$	$+(2/6)m(\alpha_{17})$	$+(2/6)m(\alpha_{17})$
$+(1/7)m(\alpha_{18})$	$+(2/7)m(\alpha_{18})$	$+(2/7)m(\alpha_{18})$

Derivation of $P\{\alpha_1\}$, $P\{\alpha_2\}$ and $P\{\alpha_3\}$

$P\{\alpha_{10}\} =$	$P\{\alpha_{11}\} =$	$P\{\alpha_{12}\} =$
$m(\alpha_1)$	$m(\alpha_1)$	$m(\alpha_1)$
$+m(\alpha_2)$	$+(1/2)m(\alpha_2)$	$+m(\alpha_2)$
$+(1/2)m(\alpha_3)$	$+m(\alpha_3)$	$+m(\alpha_3)$
$+m(\alpha_4)$	$+m(\alpha_4)$	$+m(\alpha_4)$
$+(2/3)m(\alpha_5)$	$+m(\alpha_5)$	$+m(\alpha_5)$
$+m(\alpha_6)$	$+(2/3)m(\alpha_6)$	$+m(\alpha_6)$
$+(2/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$	$+m(\alpha_7)$
$+(3/4)m(\alpha_8)$	$+(3/4)m(\alpha_8)$	$+m(\alpha_8)$
$+(2/4)m(\alpha_9)$	$+(2/4)m(\alpha_9)$	$+(3/4)m(\alpha_9)$
$+m(\alpha_{10})$	$+(2/4)m(\alpha_{10})$	$+(3/4)m(\alpha_{10})$
$+(2/4)m(\alpha_{11})$	$+m(\alpha_{11})$	$+m(\alpha_{11})$
$+(3/5)m(\alpha_{12})$	$+(4/5)m(\alpha_{12})$	$+m(\alpha_{12})$
$+(4/5)m(\alpha_{13})$	$+(3/5)m(\alpha_{13})$	$+(4/5)m(\alpha_{13})$
$+(3/5)m(\alpha_{14})$	$+(3/5)m(\alpha_{14})$	$+(4/5)m(\alpha_{14})$
$+(4/6)m(\alpha_{15})$	$+(3/6)m(\alpha_{15})$	$+(4/6)m(\alpha_{15})$
$+(3/6)m(\alpha_{16})$	$+(4/6)m(\alpha_{16})$	$+(4/6)m(\alpha_{16})$
$+(4/6)m(\alpha_{17})$	$+(4/6)m(\alpha_{17})$	$+(5/6)m(\alpha_{17})$
$+(4/7)m(\alpha_{18})$	$+(4/7)m(\alpha_{18})$	$+(5/7)m(\alpha_{18})$

Derivation of $P\{\alpha_{10}\}$, $P\{\alpha_{11}\}$ and $P\{\alpha_{12}\}$

$P\{\alpha_4\} =$	$P\{\alpha_5\} =$	$P\{\alpha_6\} =$
$m(\alpha_1)$	$m(\alpha_1)$	$m(\alpha_1)$
$+(1/2)m(\alpha_2)$	$+(1/2)m(\alpha_2)$	$+m(\alpha_2)$
$+(1/2)m(\alpha_3)$	$+m(\alpha_3)$	$+(1/2)m(\alpha_3)$
$+m(\alpha_4)$	$+m(\alpha_4)$	$+m(\alpha_4)$
$+(2/3)m(\alpha_5)$	$+m(\alpha_5)$	$+(2/3)m(\alpha_5)$
$+(2/3)m(\alpha_6)$	$+(2/3)m(\alpha_6)$	$+m(\alpha_6)$
$+(1/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$
$+(2/4)m(\alpha_8)$	$+(3/4)m(\alpha_8)$	$+(3/4)m(\alpha_8)$
$+(1/4)m(\alpha_9)$	$+(2/4)m(\alpha_9)$	$+(2/4)m(\alpha_9)$
$+(2/4)m(\alpha_{10})$	$+(2/4)m(\alpha_{10})$	$+(3/4)m(\alpha_{10})$
$+(2/4)m(\alpha_{11})$	$+(3/4)m(\alpha_{11})$	$+(2/4)m(\alpha_{11})$
$+(2/5)m(\alpha_{12})$	$+(3/5)m(\alpha_{12})$	$+(3/5)m(\alpha_{12})$
$+(2/5)m(\alpha_{13})$	$+(3/5)m(\alpha_{13})$	$+(3/5)m(\alpha_{13})$
$+(2/5)m(\alpha_{14})$	$+(3/5)m(\alpha_{14})$	$+(3/5)m(\alpha_{14})$
$+(2/6)m(\alpha_{15})$	$+(3/6)m(\alpha_{15})$	$+(3/6)m(\alpha_{15})$
$+(2/6)m(\alpha_{16})$	$+(3/6)m(\alpha_{16})$	$+(3/6)m(\alpha_{16})$
$+(2/6)m(\alpha_{17})$	$+(3/6)m(\alpha_{17})$	$+(3/6)m(\alpha_{17})$
$+(2/7)m(\alpha_{18})$	$+(3/7)m(\alpha_{18})$	$+(3/7)m(\alpha_{18})$

Derivation of $P\{\alpha_4\}$, $P\{\alpha_5\}$ and $P\{\alpha_6\}$

$P\{\alpha_{10}\} =$	$P\{\alpha_{11}\} =$	$P\{\alpha_{12}\} =$
$m(\alpha_1)$	$m(\alpha_1)$	$m(\alpha_1)$
$+m(\alpha_2)$	$+(1/2)m(\alpha_2)$	$+m(\alpha_2)$
$+(1/2)m(\alpha_3)$	$+m(\alpha_3)$	$+m(\alpha_3)$
$+m(\alpha_4)$	$+m(\alpha_4)$	$+m(\alpha_4)$
$+(2/3)m(\alpha_5)$	$+m(\alpha_5)$	$+m(\alpha_5)$
$+m(\alpha_6)$	$+(2/3)m(\alpha_6)$	$+m(\alpha_6)$
$+(2/3)m(\alpha_7)$	$+(2/3)m(\alpha_7)$	$+m(\alpha_7)$
$+(3/4)m(\alpha_8)$	$+(3/4)m(\alpha_8)$	$+m(\alpha_8)$
$+(2/4)m(\alpha_9)$	$+(2/4)m(\alpha_9)$	$+(3/4)m(\alpha_9)$
$+m(\alpha_{10})$	$+(2/4)m(\alpha_{10})$	$+(3/4)m(\alpha_{10})$
$+(2/4)m(\alpha_{11})$	$+m(\alpha_{11})$	$+m(\alpha_{11})$
$+(3/5)m(\alpha_{12})$	$+(4/5)m(\alpha_{12})$	$+m(\alpha_{12})$
$+(4/5)m(\alpha_{13})$	$+(3/5)m(\alpha_{13})$	$+(4/5)m(\alpha_{13})$
$+(3/5)m(\alpha_{14})$	$+(3/5)m(\alpha_{14})$	$+(4/5)m(\alpha_{14})$
$+(4/6)m(\alpha_{15})$	$+(3/6)m(\alpha_{15})$	$+(4/6)m(\alpha_{15})$
$+(3/6)m(\alpha_{16})$	$+(4/6)m(\alpha_{16})$	$+(4/6)m(\alpha_{16})$
$+(4/6)m(\alpha_{17})$	$+(4/6)m(\alpha_{17})$	$+(5/6)m(\alpha_{17})$
$+(4/7)m(\alpha_{18})$	$+(4/7)m(\alpha_{18})$	$+(5/7)m(\alpha_{18})$

Derivation of $P\{\alpha_{13}\}$, $P\{\alpha_{14}\}$ and $P\{\alpha_{15}\}$

$P\{\alpha_7\} =$	$P\{\alpha_8\} =$	$P\{\alpha_9\} =$
$m(\alpha_1)$	$m(\alpha_1)$	$m(\alpha_1)$
$+m(\alpha_2)$	$+m(\alpha_2)$	$+m(\alpha_2)$
$+m(\alpha_3)$	$+m(\alpha_3)$	$+m(\alpha_3)$
$+(1/2)m(\alpha_4)$	$+m(\alpha_4)$	$+(1/2)m(\alpha_4)$
$+(2/3)m(\alpha_5)$	$+m(\alpha_5)$	$+(2/3)m(\alpha_5)$
$+(2/3)m(\alpha_6)$	$+m(\alpha_6)$	$+(2/3)m(\alpha_6)$
$+m(\alpha_7)$	$+m(\alpha_7)$	$+m(\alpha_7)$
$+(3/4)m(\alpha_8)$	$+m(\alpha_8)$	$+(3/4)m(\alpha_8)$
$+(3/4)m(\alpha_9)$	$+(3/4)m(\alpha_9)$	$+m(\alpha_9)$
$+(2/4)m(\alpha_{10})$	$+(3/4)m(\alpha_{10})$	$+(2/4)m(\alpha_{10})$
$+(2/4)m(\alpha_{11})$	$+(3/4)m(\alpha_{11})$	$+(2/4)m(\alpha_{11})$
$+(3/5)m(\alpha_{12})$	$+(4/5)m(\alpha_{12})$	$+(3/5)m(\alpha_{12})$
$+(3/5)m(\alpha_{13})$	$+(4/5)m(\alpha_{13})$	$+(3/5)m(\alpha_{13})$
$+(3/5)m(\alpha_{14})$	$+(4/5)m(\alpha_{14})$	$+(4/5)m(\alpha_{14})$
$+(3/6)m(\alpha_{15})$	$+(4/6)m(\alpha_{15})$	$+(4/6)m(\alpha_{15})$
$+(3/6)m(\alpha_{16})$	$+(4/6)m(\alpha_{16})$	$+(4/6)m(\alpha_{16})$
$+(3/6)m(\alpha_{17})$	$+(4/6)m(\alpha_{17})$	$+(3/6)m(\alpha_{17})$
$+(3/7)m(\alpha_{18})$	$+(4/7)m(\alpha_{18})$	$+(4/7)m(\alpha_{18})$

Derivation of $P\{\alpha_7\}$, $P\{\alpha_8\}$ and $P\{\alpha_9\}$

$P\{\alpha_{16}\} =$	$P\{\alpha_{17}\} =$	$P\{\alpha_{18}\} =$
$m(\alpha_1)$	$m(\alpha_1)$	$m(\alpha_1)$
$+m(\alpha_2)$	$+m(\alpha_2)$	$+m(\alpha_2)$
$+m(\alpha_3)$	$+m(\alpha_3)$	$+m(\alpha_3)$
$+m(\alpha_4)$	$+m(\alpha_4)$	$+m(\alpha_4)$
$+m(\alpha_5)$	$+m(\alpha_5)$	$+m(\alpha_5)$
$+m(\alpha_6)$	$+m(\alpha_6)$	$+m(\alpha_6)$
$+m(\alpha_7)$	$+m(\alpha_7)$	$+m(\alpha_7)$
$+m(\alpha_8)$	$+m(\alpha_8)$	$+m(\alpha_8)$
$+m(\alpha_9)$	$+(3/4)m(\alpha_9)$	$+m(\alpha_9)$
$+(3/4)m(\alpha_{10})$	$+m(\alpha_{10})$	$+m(\alpha_{10})$
$+m(\alpha_{11})$	$+m(\alpha_{11})$	$+m(\alpha_{11})$
$+m(\alpha_{12})$	$+m(\alpha_{12})$	$+m(\alpha_{12})$
$+(4/5)m(\alpha_{13})$	$+m(\alpha_{13})$	$+m(\alpha_{13})$
$+m(\alpha_{14})$	$+(4/5)m(\alpha_{14})$	$+m(\alpha_{14})$
$+(5/6)m(\alpha_{15})$	$+(5/6)m(\alpha_{15})$	$+m(\alpha_{15})$
$+m(\alpha_{16})$	$+(5/6)m(\alpha_{16})$	$+m(\alpha_{16})$
$+(5/6)m(\alpha_{17})$	$+m(\alpha_{17})$	$+m(\alpha_{17})$
$+(6/7)m(\alpha_{18})$	$+(6/7)m(\alpha_{18})$	$+m(\alpha_{18})$

Derivation of $P\{\alpha_{16}\}$, $P\{\alpha_{17}\}$ and $P\{\alpha_{18}\}$

Chapter 8

Probabilized logics related to DSmT and Bayes inference

Frédéric Dambreville

Délégation Générale pour l'Armement, DGA/CTA/DT/GIP/PRO

16 Bis, Avenue Prieur de la Côte d'Or

94114, Arcueil Cedex France

Abstract: *This work proposes a logical interpretation of the non hybrid Dezert Smarandache Theory (DSmT). As probability is deeply related to a classical semantic, it appears that DSmT relies on an alternative semantic of decision. This semantic is characterized as a probabilized multi-modal logic. It is noteworthy that this interpretation justifies clearly some hypotheses usually made about the fusion rule (*ie. *the independence between the sensors). At last, a conclusion arises: there could be many possible fusion rules, depending on the chosen semantic of decision; and the choice of a semantic depends on how the actual problem is managed. Illustrating this fact, a logical interpretation of the Bayesian inference is proposed as a conclusion to this chapter.*

8.1 Introduction

When a non deterministic problem appears to be too badly shaped, it becomes difficult to make a coherent use of the probabilistic models. A particular difficulty, often neglected, comes from the interpretation of the raw data. The raw data could have a good probabilistic modelling, but in general such informations are useless: an interpretation is necessary. Determining the model of interpretation, and its probabilistic law, is the true issue. Due to the *forgotten/unknown case* syndrome, it is possible that such model cannot be entirely constructed. In some cases, only a rather weak approximation of the model is possible. Such approximated model of interpretation may produce paradoxical results. This is particularly true in information fusion problems.

Several new theories have been proposed for managing these difficulties. Dempster Shafer Theory of evidence [1, 5] is one of them. In this paper, we are interested in the Dezert Smarandache Theory (DSmT) [3], a closely related theory. These theories, and particularly the DSmT, are able to manipulate the model contradictions. But a difficulty remains: it seems uneasy to link these various theories. In particular, their relation with the theory of probability seems unclear. Such a relation is perhaps not possible, as could claim some authors, but it is necessary: it is sometimes needed to combine methods and algorithms based on different theories. This paper intends to establish such relations. A probabilized multi-modal logic is constructed. This probabilized logic, intended for the information fusion, induces the same conjunctive fusion operator as DSmT (*ie.* operator $\oplus$). By the way, the necessity of independent sources for applying the operator $\oplus$ is clarified and confirmed. Moreover, this logical interpretation induces a possible semantic of the DSmT, and somehow enlightens the intuitions behind this theory. Near the end, the paper keeps going by giving a similar interpretation of the Bayes inference. Although the Bayes inference is not related to the DSmT, this last result suggests that probabilized logics could be a possible common frame for several non deterministic theories.

Section 8.2 is beginning by a general discussion about probability. It is shown that probabilistic modellings are sometimes questionable. Following this preliminary discussion, two versions of the theory of evidence are introduced: the historical Dempster Shafer Theory and the Transferable Belief Model of Smets [8]. Section 8.3 makes a concise presentation of the Dezert Smarandache Theory. The short section 8.4 establishes some definitions about probability (and partial probability) over a set of logical propositions. These general definitions are needed in the following sections. Section 8.5 gives a logical interpretation of the DSmT on a small example. This section does not enter the theory too deeply: the modal logic associated to this interpretation is described with *practical words*, not with formulae! Section 8.6 generalizes the results to any cases. This section is much more theoretic. The modal logic is defined mathematically. Section 8.7 proposes a similar logical interpretation of the Bayesian inference. Section 8.8 concludes.

8.2 Belief Theory Models

8.2.1 Preliminary: about probability

This subsection argues about the difficulty to modelize “everything” with probability. Given a measurable universe of abstract events (or propositions) $\Omega = \{\omega_i, i \in I\}$, a probability P could be defined as a bounded and normalized measure over Ω. In this paper, we are interested in finite models (I is finite).

A probability P could also be defined from the probabilities $\rho(\omega)$ of the elementary events $\omega \in \Omega$. The density of probability ρ should verify (finite case):

$$\rho : \Omega \mapsto \mathbb{R}^+,$$

and:

$$\sum_{\omega\in\Omega} \rho(\omega) = 1\ .$$

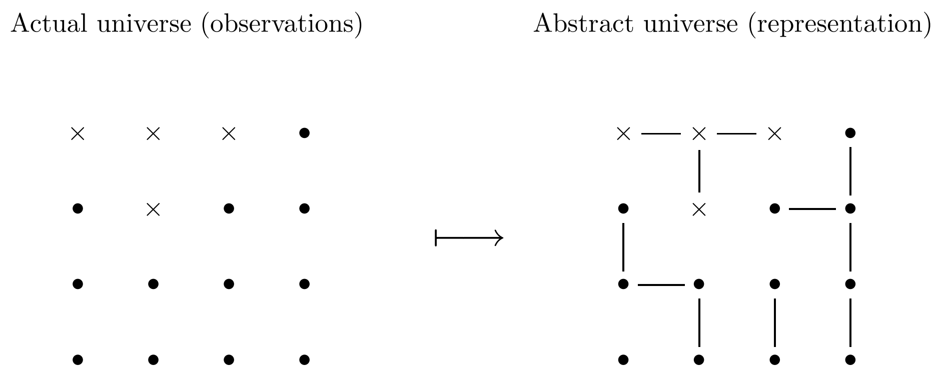
The probability P is recovered by means of the additivity property:

$$\forall A \subset \Omega,\ P(A) = \sum_{\omega\in A} \rho(\omega).$$

It is important to remember how such abstract definitions are related to a concrete notion of “chance” in the actual universe. Behind the formalism, behind the abstract events, there are *actual events*. The formalism introduced by the abstract universe Ω is just a modelling of the actual universe. Such a model is expected to be more suitable to mathematical manipulations and reasoning. But there is no reason that these actual events are compatible with the abstract events. Probability theory assumes this compatibility. More precisely, probability assumes that either the abstract and actual events are the same, either *there is a mapping from the actual events to the abstract events* (figure 8.1). When this mapping hypothesis is made, the density function makes sense then, in regard to the observation. Indeed, a practical construction of ρ becomes possible with a frequentist taste:

1. Set $\rho(\omega) = 0$ for all $\omega \in \Omega$,
2. Make N tossing of an actual event. For each tossed event, a, do:
   (a) Select the $\omega \in \Omega$ such that a maps to ω,
   (b) Set $\rho(\omega) = \rho(\omega) + 1$,
3. Set $\rho(\omega) \simeq \frac{1}{N}\rho(\omega)$ for all $\omega \in \Omega$.

The next paragraph explains why the mapping from the actual events to the abstract events is not always possible and how to overcome this difficulty.



An abstract event is a connected component; in this example,
the ×-ed observations map to the unique ×-ed component

Figure 8.1: Event mapping: probabilist case

8.2.1.1 The impossible quest of the perfect universe

It is always possible to assume that there is a perfect universe, where all problems could be modeled, but we are not able to construct it or to manipulate it practically. However, we are able to think with it. Let $\mathcal{A}$ be the actual universe, let Ω be the abstract universe, and let $\mathcal{Z}$ be this perfect universe.

The structure of Ω is well known; it describes our modelling of the actual world. *This is how we interpret the observations.* Practically, such interpretation is almost always necessary, while the raw observation may be useless. But Ω is only an hypothesis: our knowledge about the observation is generally insufficient for a *true interpretation*.

The universe $\mathcal{A}$ is observed, but like $\mathcal{Z}$ its structure is not really known: although an observation is possible, it is not necessary possible to know the meaning, the *true interpretation*, of this observation. For example, what is the meaning of an observation for a situation never seen before?

The universe $\mathcal{Z}$ is perfect, which means that it *contains* the two other, and is unknown. The word *contains* has a logical signification here, *ie.* the events/propositions of $\mathcal{A}$ or Ω are macro-events/macro-propositions of $\mathcal{Z}$ (figure 8.2) :

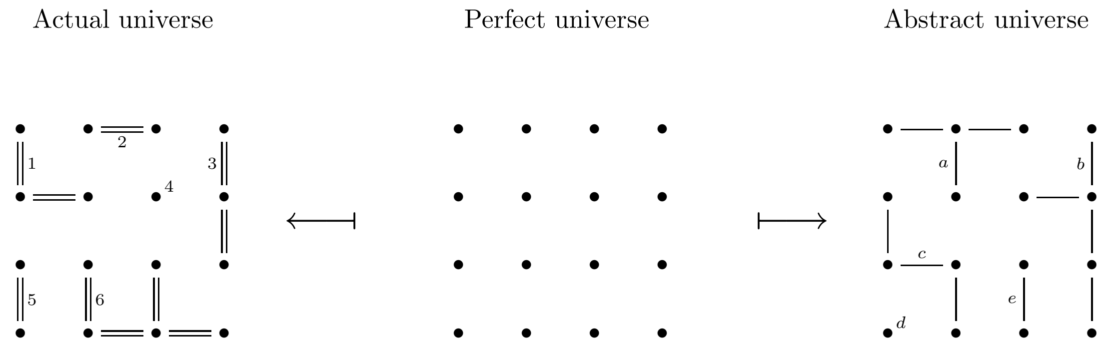
$$\mathcal{A} \subset \mathcal{P}(\mathcal{Z}) \quad \text{and} \quad \Omega \subset \mathcal{P}(\mathcal{Z}) \ ,$$

with the following exhaustiveness (**x**) and coherence (**c**) hypotheses for $\mathcal{A}$ and Ω :

x. $\mathcal{Z} = \bigcup_{a \in \mathcal{A}} a = \bigcup_{\omega \in \Omega} \omega \,,$

c1. $[a_1, a_2 \in \mathcal{A}\,,\, a_1 \neq a_2] \Rightarrow a_1 \cap a_2 = \emptyset \,,$

c2. $[\omega_1, \omega_2 \in \Omega\,,\, \omega_1 \neq \omega_2] \Rightarrow \omega_1 \cap \omega_2 = \emptyset \,.$



An abstract event (*ie.* a, b, c, d, e) is a $-$ connected component

An actual event (*ie.* $1, 2, 3, 4, 5, 6$) is a $=$ connected component

Figure 8.2: Event mapping: general case

The exhaustiveness and coherence hypotheses are questionable; it will be seen that these hypotheses induce contradictions when fusing informations.

Of course, the abstract universe Ω is a coherent interpretation of the observations, when any actual event $a \in \mathcal{A}$ is a subevent of an abstract event $\omega \in \Omega$. But since the interpretation of $\mathcal{A}$ is necessarily partial and subjective, this property does not hold in general. The figure 8.2 gives an example of erroneous interpretation of the observations: the actual event 5 intersects both the abstract event d and the abstract event c. More precisely, if an actual event $a \in \mathcal{A}$ happens, there is a perfect event $z \in a$ which has happened. Since $\mathcal{Z}$ contains (*ie.* maps to) Ω, there is an unique abstract event, $\omega \in \Omega$, which checks z, *ie.* $z \in \omega$. As a conclusion, when a given actual event a happens, any abstract event $\omega \in \Omega$ such that $\omega \cap a \neq \emptyset$ is likely to happen. Practically, such situation is easy to decide, since it just happens when a doubt appears in a measure classification. The table 8.1, refering to the example of figure 8.2, gives the possible abstract events related to each tossed observation.

Finally, it does not seem possible to define a density of probability for unique abstract events from partially decidable observations. But it is possible to define a density function for multiple events. Again, a construction of such function, still denoted ρ, is possible in a frequentist manner:

1. Set $\rho(\phi) = 0$ for all $\phi \subset \Omega$,
2. Make N tossing of an actual event. For each tossed event, a, do:
   (a) Define the set $\phi(a) = \{\omega \in \Omega / \omega \cap a \neq \emptyset\}$,
   (b) Set $\rho(\phi(a)) = \rho(\phi(a)) + 1$,
3. Set $\rho(\phi) \simeq \frac{1}{N}\rho(\phi)$ for all $\phi \subset \Omega$.

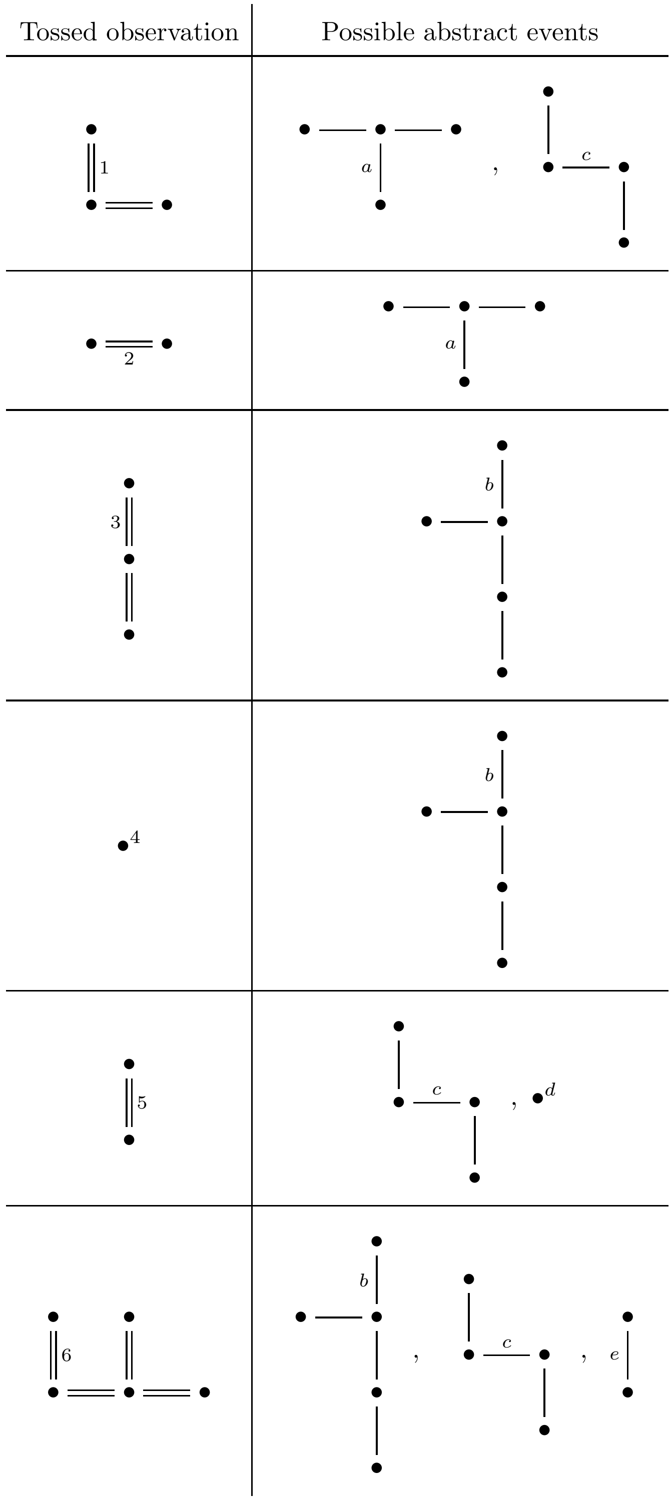


Table 8.1: Event *multi-mapping* for figure 8.2

In particular, $\rho(\emptyset) = 0$.

In the particular case of table 8.1, this density is related to the probability of observation by:

$$\rho\{a, c\} = p(1)\,,\ \rho\{a\} = p(2)\,,\ \rho\{b\} = p(3) + p(4)\,,\ \rho\{c, d\} = p(5)\,,\ \rho\{b, c, e\} = p(6)\,.$$

The previous discussion has shown that the definition of a density of probability for the abstract events does not make sense, when the interpretations of the observations are approximative. However, it is possible to construct a density for multiple abstract events. Such a density looks quite similarly to the *Basic Belief Assignment* of DST, defined in the next section.

8.2.2 Dempster Shafer Theory

8.2.2.1 Definition

A Dempster Shafer model [1, 2, 5] is characterized by a pair (Ω, m), where Ω is a set of abstract events and the basic belief assignment (bba) m is a non negatively valued function defined over $\mathcal{P}(\Omega)$, the set of subsets of Ω, such that:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{\phi \subset \Omega} m(\phi) = 1 \, .$$

A DSm (Ω, m) could be seen as a non deterministic interpretation of the actuality. Typically, it is a tool providing informations from a sensor.

8.2.2.2 Belief of a proposition

Let $\phi \subset \Omega$ be a proposition. Assume a basic belief assignment m. The *degree of belief of* ϕ, $\mathrm{Bel}(\phi)$, and the *plausibility* of ϕ, $\mathrm{Pl}(\phi)$, are defined by:

$$\mathrm{Bel}(\phi) = \sum_{\psi \subset \phi} m(\psi) \quad \text{and} \quad \mathrm{Pl}(\phi) = \sum_{\psi \cap \phi \neq \emptyset} m(\psi) \, .$$

Bel and Pl do not satisfy the additivity property of probability. $\mathrm{Bel}(\phi)$ and $\mathrm{Pl}(\phi)$ are the lower and upper measures of the "credibility" of the proposition ϕ. These measures are sometimes considered as the lower bound and the upper bound of the probability of ϕ:

$$\mathrm{Bel}(\phi) \leq P(\phi) \leq \mathrm{Pl}(\phi) \, .$$

This interpretation is dangerous, since it is generally admitted that probability and DST are quite different theories.

8.2.2.3 Fusion rule

Assume two bba m_1 and m_2, defined on the same universe Ω, obtained from two different sources. It is generally admitted that the sources are independent. Then, the bba $m_1 \oplus m_2$ is defined by:

$$\begin{cases} m_1 \oplus m_2(\emptyset) = 0 \, , \\ m_1 \oplus m_2(\phi) = \dfrac{1}{Z} \displaystyle\sum_{\psi_1 \cap \psi_2 = \phi} m_1(\psi_1) m_2(\psi_2) \, , \quad \text{where } Z = 1 - \displaystyle\sum_{\psi_1 \cap \psi_2 = \emptyset} m_1(\psi_1) m_2(\psi_2) \, . \end{cases}$$

The operator $\oplus$ describes the (conjunctive) information fusion between two bba.

The normalizer Z is needed since the bba is zeroed for the empty set $\emptyset$. Except some specific cases, it is indeed possible that:

$$m_1(\psi_1)m_2(\psi_2) > 0 \ , \tag{8.1}$$

$$\psi_1 \cap \psi_2 = \emptyset \ . \tag{8.2}$$

In particular, the property (8.2) is related to an implied coherence hypothesis; more precisely, since the universe Ω is defined as a *set* of events, the intersection of distinct singletons is empty:

$$\forall \{\omega_1\}, \{\omega_2\} \subset \Omega \, , \ \{\omega_1\} \neq \{\omega_2\} \Rightarrow \{\omega_1\} \cap \{\omega_2\} = \emptyset \ .$$

Notice that this hypothesis is quite similar to the hypothesis **c2.** of section 8.2.1 . *The coherence hypothesis seems to be the source of the contradictions in the abstract model, when fusing informations.* Finally, $Z < 1$ means that our abstract universe Ω has been incorrectly defined and is thus unable to fit the both sensors. Z measures the error in our model of interpretation. This ability of the rule $\oplus$ is really new in comparison with probabilistic rules.

8.2.3 Transferable Belief Model

Smets has made an extensive explanation of TBM [8]. This section focuses on a minimal and somewhat simplified description of the model.

8.2.3.1 Definition

A Transferable Belief Model is characterized by a pair (Ω, m), where Ω is a set of abstract events and the basic belief assignment m is a non negatively valued function defined over $\mathcal{P}(\Omega)$ such that:

$$\sum_{\phi \subset \Omega} m(\phi) = 1 \ .$$

In this definition, *the hypothesis* $m(\emptyset) = 0$ *does not hold anymore.*

8.2.3.2 Fusion rule

Smets' rule looks like a refinement of Dempster and Shafer's rule:

$$m_1 \oplus m_2(\phi) = \sum_{\psi_1 \cap \psi_2 = \phi} m_1(\psi_1)m_2(\psi_2) \ .$$

Notice that the normalizer does not exist anymore. The measure of contradiction has been moved into $m(\emptyset)$. *This theory has been justified from an axiomatization of the fusion rule.*

8.2.3.3 TBM generalizes DST

First, notice that any bba for DST is a valid bba for TBM, but the converse is false because of $\emptyset$. Now, for any bba m_T of TBM such that $m_T(\emptyset) < 1$, construct the bba $\Delta(m_T)$ of DST defined by:

$$\Delta(m_T)(\emptyset) = 0 \text{ and } \forall\phi \subset \Omega : \phi \neq \emptyset\,,\ \Delta(m_T)(\phi) = \frac{m_T(\phi)}{1 - m_T(\emptyset)}\,.$$

Δ is an onto mapping. Any bba m_D of DST is a bba of TBM, and $\Delta(m_D) = m_D$.

Δ is a morphism for $\oplus$. *IE.* $\Delta(m_{T,1} \oplus m_{T,2}) = \Delta(m_{T,1}) \oplus \Delta(m_{T,2})$.

Proof. By definition, it is clear that:

$$\Delta(m_{T,1}) \oplus \Delta(m_{T,2})(\emptyset) = 0 = \Delta(m_{T,1} \oplus m_{T,2})(\emptyset)\,.$$

Now, for any $\phi \subset \Omega$, such that $\phi \neq \emptyset$:

$$\Delta(m_{T,1}) \oplus \Delta(m_{T,2})(\phi) = \frac{\displaystyle\sum_{\psi_1\cap\psi_2=\phi} \Delta(m_{T,1})(\psi_1)\Delta(m_{T,2})(\psi_2)}{\displaystyle\sum_{\phi\neq\emptyset}\sum_{\psi_1\cap\psi_2=\phi} \Delta(m_{T,1})(\psi_1)\Delta(m_{T,2})(\psi_2)}$$

$$= \frac{\displaystyle\sum_{\psi_1\cap\psi_2=\phi} \frac{m_{T,1}(\psi_1)}{1 - m_{T,1}(\emptyset)} \times \frac{m_{T,2}(\psi_2)}{1 - m_{T,2}(\emptyset)}}{\displaystyle\sum_{\phi\neq\emptyset}\sum_{\psi_1\cap\psi_2=\phi} \frac{m_{T,1}(\psi_1)}{1 - m_{T,1}(\emptyset)} \times \frac{m_{T,2}(\psi_2)}{1 - m_{T,2}(\emptyset)}} = \frac{\displaystyle\sum_{\psi_1\cap\psi_2=\phi} m_{T,1}(\psi_1)m_{T,2}(\psi_2)}{\displaystyle\sum_{\phi\neq\emptyset}\sum_{\psi_1\cap\psi_2=\phi} m_{T,1}(\psi_1)m_{T,2}(\psi_2)}$$

$$= \frac{m_{T,1} \oplus m_{T,2}(\phi)}{\displaystyle\sum_{\phi\neq\emptyset} m_{T,1} \oplus m_{T,2}(\phi)} = \frac{m_{T,1} \oplus m_{T,2}(\phi)}{1 - m_{T,1} \oplus m_{T,2}(\emptyset)} = \Delta(m_{T,1} \oplus m_{T,2})(\phi)\,.$$

□□□

Since Δ is an onto morphism, TBM is a generalization of DST. More precisely, a bba of TBM contains more information than a bba of DST, *ie.* the measure of contradiction $m(\emptyset)$, but this complementary information remains compatible with the fusion rule of DST.

The Dezert Smarandache Theory is introduced in the next section. This theory shares many common points with TBM. But there is a main and fundamental contribution of this theory. It does not make the coherence hypothesis anymore and the contradictions are managed differently: the abstract model is more flexible to the interpretation and it is not needed to rebuild the model in case of contradicting sensors.

8.3 Dezert Smarandache Theory (DSmT)

Both in DST and in TBM, the difficulty in the model definition appears when dealing with the contradictions between the informations. But contradictions are unavoidable, when dealing with imprecise informations. This assertion is illustrated by the following example.

B&W example. Assume a sensor s_1 which tells us if an object is white (W) or not (NW), and gives no answer (NA_1) in litigious cases. The actual universe for this sensor is $\mathcal{A}_1 = \{\mathrm{W}, \mathrm{NW}, \mathrm{NA}_1\}$. Assume a sensor s_2 which tells us if an object is black (B) or not (NB), and gives no answer (NA_2) in litigious cases. The actual universe for this sensor is $\mathcal{A}_2 = \{\mathrm{B}, \mathrm{NB}, \mathrm{NA}_2\}$. These characteristics are not known, but the sensors have been tested with black or white objects. For this reason, it is natural to model our world by $\Omega = \{\text{black}, \text{white}\}$. When a litigious case happens, its interpretation will just be the pair $\{\text{black}, \text{white}\}$. Otherwise the good answer is expected. The following properties are then verified:

$$\mathrm{B}, \mathrm{NW} \subset \text{black} \quad \text{and} \quad \mathrm{W}, \mathrm{NB} \subset \text{white}\,.$$

The coherence hypothesis is assumed, that is $\text{black} \cap \text{white} = \emptyset$. The event $\text{black} \cap \text{white}$ is impossible. This model works well, as long as the sensors work separately or the objects are still black or white. Now, in a true universe there are many objects which are neither white and neither black, and this without any litigation. For example: gray objects. Assume that the two sensors are activated. Then, the fused sensors will answer $\mathrm{NW} \cap \mathrm{NB}$, which will be interpreted by $\text{black} \cap \text{white}$. This contradicts the coherence hypothesis.

Conclusion. This example is a sketch of what generally happens, when constructing a system of decision. Several sources of information are available (two sensors here). These sources have different discrimination abilities. In fact, these discrimination abilities are not really known, but by running these sources on several test samples (black and white objects here), a model of theses abilities is obtained (here it is learned within Ω that our sensors distinguish between black and white objects). Of course, it is never sure that this model is complete. It is still possible actually that some new unknown cases could be discriminated by the information sources. In the example, the combination of two sensors made it possible to discriminate a new class of objects: the neither black, neither white objects. But when fusing these sensors, the new cases will become contradictions regarding the coherence hypothesis. Not only the coherence hypothesis makes our model contradictory, but it also prevents us from discovering new cases. The coherence hypothesis should be removed! Dezert and Smarandache proposed a model without the coherence hypothesis.

8.3.1 Dezert Smarandache model

In DST and TBM, the coherence hypothesis was implied by the use of a set, Ω, to represent the abstract universe. Moreover, the set operators $\cap$, $\cup$ and c (*ie.* set complement) were used to explain the interactions between the propositions $\phi \subset \Omega$. In fact, the notion of propositions is related to the notion of Boolean Algebra. Sets together with set operators are particular models of Boolean Algebra. Since DSmT does not make the coherence hypothesis, DSmT cannot rely on the set formalism. However, some boolean relations are needed to explain the relations between propositions. Another fundamental Boolean

Algebra is the propositional logic. This model should be used for the representation of the propositions of DSmT. Nevertheless, the negation operator will be removed from our logic, since it implies itself some coherence hypotheses, *eg.* $\phi \wedge \neg\phi \equiv \bot$! By identifying the equivalent propositions of the resulting logic, an *hyper-power set* of propositions is obtained. *Hyper-power sets are used as models of universe for the DSmT.*

8.3.1.1 Hyper-power set

Let $\Phi = \{\phi_i / i \in I\}$ be a set of propositions. The hyper-power set $< \Phi >$ is the free boolean pre-algebra generated by Φ and the boolean operators $\wedge$ and $\vee$:

$$\Phi, < \Phi > \wedge < \Phi >, < \Phi > \vee < \Phi > \subset < \Phi >$$

and $\wedge, \vee$ verify the properties:

Commutative. $\phi \wedge \psi \equiv \psi \wedge \phi$ and $\phi \vee \psi \equiv \psi \vee \phi$,

Associative. $\phi \wedge (\psi \wedge \eta) \equiv (\phi \wedge \psi) \wedge \eta$ and $\phi \vee (\psi \vee \eta) \equiv (\phi \vee \psi) \vee \eta$,

Distributive. $\phi \wedge (\psi \vee \eta) \equiv (\phi \wedge \psi) \vee (\phi \wedge \eta)$ and $\phi \vee (\psi \wedge \eta) \equiv (\phi \vee \psi) \wedge (\phi \vee \eta)$,

Idempotent. $\phi \wedge \phi \equiv \phi$ and $\phi \vee \phi \equiv \phi$,

Neutral sup/sub-elements. $\phi \wedge (\phi \vee \psi) \equiv \phi$ and $\phi \vee (\phi \wedge \psi) \equiv \phi$,

for any $\phi, \psi, \eta \in < \Phi >$.

Unless more specifications about the free pre-algebra are made, this definition forbids the propositions to be exclusive (no coherence assumption) or to be exhaustive. In particular, the negation operator, $\neg$, and the *never happen/always happen*, $\bot/\top$, are excluded from the formalism. Indeed, the negation is related to the coherence hypothesis, since $\top$ is related to the exhaustiveness hypothesis.

Property. It is easily proved from the definition that:

$$\forall \phi, \psi \in < \Phi >,\ \phi \wedge \psi \equiv \phi \iff \phi \vee \psi \equiv \psi\ .$$

The order $\leq$ is a meta-operator defined over $< \Phi >$ by:

$$\phi \leq \psi \iff \phi \wedge \psi \equiv \phi \iff \phi \vee \psi \equiv \psi\ .$$

The order $<$ is a meta-operator defined over $< \Phi >$ by:

$$\phi < \psi \iff \left[\phi \leq \psi \text{ and } \phi \not\equiv \psi \right]\ .$$

The hyper-power set order $\leq$ is the analogue of the set order $\subset$.

8.3.1.2 Dezert Smarandache Model

A Dezert Smarandache model (DSmm) is a pair (Φ, m), where the (abstract) universe Φ is a set of propositions and the basic belief assignment m is a non negatively valued function defined over $< \Phi >$ such that:

$$\sum_{\phi \in <\Phi>} m(\phi) = 1 \, .$$

8.3.1.3 Belief Function

The belief function Bel is defined by:

$$\forall \phi \in < \Phi >, \, \mathrm{Bel}(\phi) = \sum_{\psi \in <\Phi> : \psi \leq \phi} m(\psi) \, . \tag{8.3}$$

Since propositions are never exclusive within $< \Phi >$, the (classical) plausibility function is just equal to 1. The equation (8.3) is invertible:

$$\forall \phi \in < \Phi >, \, m(\phi) = \mathrm{Bel}(\phi) - \sum_{\psi \in <\Phi> : \psi < \phi} m(\psi) \, .$$

8.3.2 Fusion rule

For a given universe Φ , and two basic belief assignments m_1 and m_2, associated to different sensors, the fused basic belief assignment is $m_1 \oplus m_2$, defined by:

$$m_1 \oplus m_2(\phi) = \sum_{\psi_1 \wedge \psi_2 \equiv \phi} m_1(\psi_1) m_2(\psi_2) \, . \tag{8.4}$$

8.3.2.1 Dezert & Smarandache's example

Assume a thief (45 years old) witnessed by a granddad and a grandson. The witnesses answer the question: *is the thief young or old?* The universe is then $\Phi = \{\mathrm{young}, \mathrm{old}\}$. The granddad answers that the thief is rather young. Its testimony is described by the bba:

$$m_1(\mathrm{young}) = 0.9 \quad \text{and} \quad m_1(\mathrm{young} \vee \mathrm{old}) = 0.1 \text{ (slight unknown)} \, .$$

Of course, the grandson thinks he is rather old:

$$m_2(\mathrm{old}) = 0.9 \quad \text{and} \quad m_2(\mathrm{young} \vee \mathrm{old}) = 0.1 \text{ (slight unknown)} \, .$$

How to interpret the testimonies? The fusion rule says:

$$\begin{cases} m_1 \oplus m_2(\mathrm{young} \wedge \mathrm{old}) = 0.9801 \quad \text{(highly contradicts} \rightarrow \text{third case)} \\ m_1 \oplus m_2(\mathrm{young}) = m_1 \oplus m_2(\mathrm{old}) = 0.0099 \\ m_1 \oplus m_2(\mathrm{young} \vee \mathrm{old}) = 0.0001 \end{cases}$$

Our hypotheses contradict. There were a third case: the thief is middle aged.

8.3.2.2 Comments

In DSmT, there is not a clear distinction between the notion of conjunction, $\wedge$, the notion of *third case* and the notion of contradiction. The model does not decide for that and leaves this distinction to our last interpretation. It is our interpretation of the model which will make the distinction. Thus, the DSm model avoids any *over-abstraction* of the actual universe. Consequently, it never fails although we could fail in the last instance by interpreting it. Another good consequence is that *DSmT specifies any contradiction/third case*: the contradiction $\phi \wedge \psi$ is not just a contradiction, it is the contradiction between ϕ and ψ.

8.4 Probability over logical propositions

Probabilities are classically defined over measurable sets. However, this is only a manner to modelize the notion of probability, which is essentially a measure of the belief of logical propositions. Probability could be defined without reference to the measure theory, at least when the number of propositions is finite. In this section, the notion of probability is explained within a strict logical formalism. This formalism is of constant use in the sequel.

Intuitively, a probability over a set of logical propositions is a measure of belief which is additive (disjoint propositions are adding their chances) and increasing with the proposition (weak propositions are more probable). This measure should be zeroed for the impossible propositions and full for the ever-true propositions. Moreover, *a probability is a multiplicative measure for independent propositions.* The independence of propositions is a meta-relation between propositions, which generally depends on the problem setting.

These intuitions are formalized now. It is assumed that the reader is used with some logical notions.

8.4.1 Definition

Let L be at least an extension of the classical logic of propositions, that is L contains the operators $\wedge$, $\vee$, $\neg$ (and, or, negation) and the propositions $\bot$, $\top$ (always false, always true). Assume moreover that some propositions pairs of L are *recognized* as *independent propositions* (this is a meta-relation not necessarily related to the logic itself). A probability p over L is a $\mathbb{R}^+$ valued function such that for any proposition ϕ and ψ of L:

Additivity. $p(\phi \wedge \psi) + p(\phi \vee \psi) = p(\phi) + p(\psi)$,

Coherence. $p(\bot) = 0$,

Finiteness. $p(\top) = 1$,

Multiplicativity. When ϕ and ψ are *independent* propositions, then $p(\phi \wedge \psi) = p(\phi)p(\psi)$.

8.4.2 Property

The coherence and additivity implies the increaseness of p:

Increaseness. $p(\phi \wedge \psi) \leq p(\phi)$.

Proof. Since $\phi \equiv (\phi \wedge \psi) \vee (\phi \wedge \neg\psi)$ and $(\phi \wedge \psi) \wedge (\phi \wedge \neg\psi) \equiv \bot$, it follows from the additivity:

$$p(\phi) + p(\bot) = p(\phi \wedge \psi) + p(\phi \wedge \neg\psi) \ .$$

From the coherence $p(\bot) = 0$, it is deduced $p(\phi) = p(\phi \wedge \psi) + p(\phi \wedge \neg\psi)$. Since p is non negatively valued, $p(\phi) \geq p(\phi \wedge \psi)$.

□□□

8.4.3 Partially defined probability

In the sequel, knowledges are alternately described by partially known probabilities over a logical system. Typically, the probability p will be known only for a subset of propositions $\ell \subset L$.

Partial probabilities have been investigated by other works [9], for the representation of partial knowledge. In these works, the probabilities are characterized by constraints. It is believed that this area has been insufficiently investigated. And although our presentation is essentially focused on the logical aspect of the knowledge representation, it should be noticed that it is quite related to this notion of partial probability. In particular, the knowledge of the probability for a subset of propositions implies the definition of constraints for the probability over the whole logical system. For example, the knowledge of $\pi = p(\phi \wedge \psi)$ implies a lower bound for $p(\phi)$ and $p(\psi)$: $p(\phi) \geq \pi$ and $p(\psi) \geq \pi$.

The next section introduces, on a small example, a new interpretation of DSmT by means of probabilized logic.

8.5 Logical interpretation of DSmT: an example

A bipropositional DSm model $\Delta = (\{\phi_1, \phi_2\}, m)$ is considered. This section proposes an interpretation of this DSm model by means of probabilized modal propositions.

8.5.1 A possible modal interpretation

Consider the following *modal* propositions:

U **.** Unable to decide between the ϕ_i's ,

α_i **.** Proposition ϕ_i is sure ; No Other Information (NOI) ,

I **.** Contradiction between the ϕ_i's .

It is noticeable that these propositions are exclusive:

$$\forall\, a, b \in \{U, \alpha_1, \alpha_2, I\}\,,\ a \not\equiv b \Rightarrow a \wedge b \equiv \perp\,. \tag{8.5}$$

These propositions are clearly related to the propositions ϕ_i :

$$\begin{cases} I \leq \phi_1 \wedge \phi_2\,,\ \phi_1\,,\ \phi_2\,,\ \phi_1 \vee \phi_2 & [\text{the } \textit{contradiction } I \text{ implies everything}] \\ \alpha_i \leq \phi_i\,,\ \phi_1 \vee \phi_2\,,\ \ \text{for } i = 1, 2 & [\alpha_i \text{ implies } \phi_i \text{ and } \phi_1 \vee \phi_2] \\ U \leq \phi_1 \vee \phi_2 & [U \text{ only implies } \phi_1 \vee \phi_2] \end{cases} \tag{8.6}$$

These propositions are also exhaustive; *ie.* in the universe Φ , either one of the propositions I, α_1, α_2, U should be verified:

$$I \vee \alpha_1 \vee \alpha_2 \vee U \equiv \phi_1 \vee \phi_2\,. \tag{8.7}$$

Since the propositions α_i, U, I are characterizing the knowledge about ϕ_i (with NOI), the doubt or the contradiction, it seems natural to associate to these propositions a belief equivalent to $m(\phi_i)$, $m(\phi_1 \vee \phi_2)$ and $m(\phi_1 \wedge \phi_2)$. These beliefs will be interpreted as probabilities over I, U and α_i :

$$p(I) = m(\phi_1 \wedge \phi_2)\,, \qquad p(U) = m(\phi_1 \vee \phi_2)\,. \qquad p(\alpha_i) = m(\phi_i)\,,\ \ \text{for } i = 1, 2\,. \tag{8.8}$$

Such probabilistic interpretation is natural but questionable: it mixes probabilities together with bba. Since the propositions ϕ_i are not directly manipulated, this interpretation is not forbidden however. In fact, it will be shown next that this interpretation *implies* the fusion rule $\oplus$ and this will be a posterior justification of such hypothesis .

8.5.2 Deriving a fusion rule

In this section, a fusion rule is deduced from the previous *probabilized modal interpretation.* This rule happens to be the (conjunctive) fusion rule of DSmT.
Let $\Delta_j = (\{\phi_1, \phi_2\}, m_j)$ be the DSm models associated to sensors $j = 1, 2$ working beside the same abstract universe $\{\phi_1, \phi_2\}$. Define the set of modal propositions $S_j = \{I_j, \alpha_{j1}, \alpha_{j2}, U_j\}$:

U_j **.** Unable to decide between the ϕ_i's, according to sensor j ,

α_{ji} **.** Proposition ϕ_i is sure and NOI, according to sensor j ,

I_j **.** Contradiction between the ϕ_i's, according to sensor j .

The propositions of S_j verify of course the properties (8.5), (8.6), (8.7) and (8.8), the subscript $_j$ being added when needed. Define:

$$S = S_1 \wedge S_2 = \{a_1 \wedge a_2 / a_1 \in S_1 \text{ and } a_2 \in S_2\} .$$

Consider $a \equiv a_1 \wedge a_2$ and $b \equiv b_1 \wedge b_2$, two distinct elements of S. Then, either $a_1 \not\equiv b_1$ or $a_2 \not\equiv b_2$. Since S_j verifies (8.5), it follows $a_1 \wedge b_1 \equiv \bot$ or $a_2 \wedge b_2 \equiv \bot$, thus yielding:

$$(a_1 \wedge a_2) \wedge (b_1 \wedge b_2) \equiv (a_1 \wedge b_1) \wedge (a_2 \wedge b_2) \equiv \bot .$$

S is made of exclusive elements. It is also known from (8.7) that $\phi_1 \vee \phi_2 \equiv \bigvee_{a_j \in S_j} a_j$; S_j is exhaustive. It follows:

$$\phi_1 \vee \phi_2 \equiv (\phi_1 \vee \phi_2) \wedge (\phi_1 \vee \phi_2) \equiv \bigwedge_{j=1}^{2} \bigvee_{a_j \in S_j} a_j \equiv \bigvee_{a \in S} a .$$

S is exhaustive. In fact, S enumerates all the possible cases of observation. It is thus reasonable to think that the fused knowledge of these sensors could be constructed from S. The question then arising is: what is the signification of a proposition $a_1 \wedge a_2 \in S$? It is remembered that a proposition of S_j just tells *what is known for sure according to sensor j*. But the semantic for combining sure or unsure propositions is quite natural:[1]

- unsure + unsure = unsure
- unsure + sure = sure
- sure + sure = sure OR contradiction
- anything + contradiction = contradiction

In particular contradiction arises, when two informations are sure *and* these informations are known contradictory. This conduces to a general interpretation of S:

$\wedge$	I_2	α_{21}	α_{22}	U_2
I_1	Contradiction	Contradiction	Contradiction	Contradiction
α_{11}	Contradiction	ϕ_1 is sure	Contradiction	ϕ_1 is sure
α_{12}	Contradiction	Contradiction	ϕ_2 is sure	ϕ_2 is sure
U_1	Contradiction	ϕ_1 is sure	ϕ_2 is sure	Unsure

At last, any proposition of S is a sub-event of a proposition I, α_1, α_2 or U, defined by:

U **.** The sensors are unable to decide between the ϕ_i's,

α_i **.** The sensors are sure of the proposition ϕ_i, but do not know anything else,

I **.** The sensors contradict.

[1]In fact, the independence of the sensors is implicitly hypothesized in such combining rule (refer to next section).

Since S is exhaustive, the propositions U, α_i, I are entirely determined by S:

- $I \equiv (I_1 \wedge I_2) \vee (I_1 \wedge \alpha_{21}) \vee (I_1 \wedge \alpha_{22}) \vee (I_1 \wedge U_2) \vee (\alpha_{11} \wedge I_2) \vee$
  $(\alpha_{12} \wedge I_2) \vee (U_1 \wedge I_2) \vee (\alpha_{12} \wedge \alpha_{21}) \vee (\alpha_{11} \wedge \alpha_{22})$,
- $\alpha_i \equiv (\alpha_{1i} \wedge \alpha_{2i}) \vee (U_1 \wedge \alpha_{2i}) \vee (\alpha_{1i} \wedge U_2)$,
- $U \equiv U_1 \wedge U_2$.

The propositions I, α_i, U are thus entirely specified and since S is made of exclusive elements, their probabilities are given by:

- $p(I) = p(I_1 \wedge I_2) + p(I_1 \wedge \alpha_{21}) + p(I_1 \wedge \alpha_{22}) + p(I_1 \wedge U_2) + \cdots + p(\alpha_{11} \wedge \alpha_{22})$,
- $p(\alpha_i) = p(\alpha_{1i} \wedge \alpha_{2i}) + p(U_1 \wedge \alpha_{2i}) + p(\alpha_{1i} \wedge U_2)$,
- $p(U) = p(U_1 \wedge U_2)$.

At this point, *the independence of the sensors is needed.* The hypothesis implies $p(a_1 \wedge a_2) = p(a_1)p(a_2)$. The constraints (8.8) for each sensor j then yield:

- $p(I) = m_1(\phi_1 \wedge \phi_2)m_2(\phi_1 \wedge \phi_2) + m_1(\phi_1 \wedge \phi_2)m_2(\phi_1) + \cdots + m_1(\phi_1)m_2(\phi_2)$,
- $p(\alpha_i) = m_1(\phi_i)m_2(\phi_i) + m_1(\phi_1 \vee \phi_2)m_2(\phi_i) + m_1(\phi_i)m_2(\phi_1 \vee \phi_2)$,
- $p(U) = m_1(\phi_1 \vee \phi_2)m_2(\phi_1 \vee \phi_2)$.

The definition of $m_1 \oplus m_2$ implies finally:

$$p(I) = m_1 \oplus m_2(\phi_1 \wedge \phi_2)\,, \qquad p(\alpha_i) = m_1 \oplus m_2(\phi_i)\,, \qquad \text{and} \qquad p(U) = m_1 \oplus m_2(\phi_1 \vee \phi_2)\,.$$

Our interpretation of DSmT by means of probabilized modal propositions has implied the fusion rule $\oplus$.

This result is investigated rigorously and generally in the next section.

8.6 Multi-modal logic and information fusion

This section generalizes the results of the previous section. The presentation is more formalized. In particular, a multi-modal logic for the information fusion is constructed. This presentation is not fully detailed and it is assumed that the reader is acquainted with some logical notions.

8.6.1 Modal logic

In this introductory section, we are just interested in modal logic, and particularly in the T-system. There is no need to argue about a better system, since we are only interested in manipulating the modalities

$\Box$, $\neg\Box$, $\diamond$ and $\neg\diamond$.

Being given Φ a set of atomic propositions, the set of classical propositions, $C(\Phi)$ more simply denoted C, is defined by:

- $\Phi \subset C$, $\bot \in C$ and $\top \in C$,
- If $\phi, \psi \in C$, then $\neg\phi \in C$, $\phi \wedge \psi \in C$, $\phi \vee \psi \in C$ and $\phi \rightarrow \psi \in C$.

The set of modal propositions, $M(\Phi)$ also denoted M, is constructed as follows:

- $C \subset M$,
- If $\phi \in M$, then $\Box\phi \in M$ and $\diamond\phi \in M$,
- If $\phi, \psi \in M$, then $\neg\phi \in M$, $\phi \wedge \psi \in M$, $\phi \vee \psi \in M$ and $\phi \rightarrow \psi \in M$.

The proposition $\Box\phi$ will mean that *the proposition ϕ is true for sure*. The proposition $\diamond\phi$ will mean that *the proposition ϕ is possibly true*.

In the sequel, the notation $\vdash \phi$ means that ϕ *is proved in* T. A proposition ϕ such that $\vdash \phi$ is also called an axiom. The notation $\phi \equiv \psi$ means both $\vdash \phi \rightarrow \psi$ and $\vdash \psi \rightarrow \phi$.

All axioms are defined recursively by assuming some deduction rules and initial axioms.

Modus Ponens (MP). For any proposition $\phi, \psi \in M$, such that $\vdash \phi$ and $\vdash \phi \rightarrow \psi$, it is deduced $\vdash \psi$.

Classical axioms. For any $\phi, \psi, \eta \in M$, it is assumed the axioms:

1. $\vdash \top$,
2. $\vdash \phi \rightarrow (\psi \rightarrow \phi)$,
3. $\vdash (\eta \rightarrow (\phi \rightarrow \psi)) \rightarrow ((\eta \rightarrow \phi) \rightarrow (\eta \rightarrow \psi))$,
4. $\vdash (\neg\phi \rightarrow \neg\psi) \rightarrow ((\neg\phi \rightarrow \psi) \rightarrow \phi)$,
5. $\bot \equiv \neg\top$,
6. $\phi \rightarrow \psi \equiv \neg\phi \vee \psi$,
7. $\phi \wedge \psi \equiv \neg(\neg\phi \vee \neg\psi)$.

It is deduced from these axioms that:

- The relation $\vdash \phi \rightarrow \psi$ is a pre-order with a minimum $\bot$ and a maximum $\top$: $\bot$ is the strongest proposition, $\top$ is the weakest proposition,
- The relation $\equiv$ is an equivalence relation.

Modal axioms and rule. Let $\phi, \psi \in M$.

i. From $\vdash \phi$ is deduced $\vdash \Box\phi$; axioms are sure. *This does not mean* $\vdash \phi \rightarrow \Box\phi$ *which is false!*

ii. $\vdash \Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$; when the inference is sure and the premise is sure, the conclusion is sure,

iii. $\vdash \Box\phi \to \phi$; sure propositions are true,

iv. $\diamond\phi \equiv \neg\Box\neg\phi$; is unsure what cannot be false for sure.

It is deduced that the proposition $\Box\phi$ is stronger than ϕ which is stronger than $\diamond\phi$.

Notation. In the sequel, $\psi \leq \phi$ means $\vdash \psi \to \phi$, and $\psi < \phi$ means both $\psi \leq \phi$ and $\phi \not\equiv \psi$.

The logical operators are compatible with $\equiv$. Denote $\phi_{/\equiv} = \{\psi \in M/\psi \equiv \phi\}$, the class of equivalence of ϕ. Let $\phi, \psi \in M$, $\hat{\phi} \in \phi_{/\equiv}$ and $\hat{\psi} \in \psi_{/\equiv}$. Then holds:

- $\hat{\phi} \to \hat{\psi} \in (\phi \to \psi)_{/\equiv}$
- $\neg\hat{\phi} \in (\neg\phi)_{/\equiv}$
- $\hat{\phi} \wedge \hat{\psi} \in (\phi \wedge \psi)_{/\equiv}$
- $\hat{\phi} \vee \hat{\psi} \in (\phi \vee \psi)_{/\equiv}$
- $\Box\hat{\phi} \in (\Box\phi)_{/\equiv}$
- $\diamond\hat{\phi} \in (\diamond\phi)_{/\equiv}$

The logical operators over M are thus extended naturally to the classes of M by setting:

- $\phi_{/\equiv} \to \psi_{/\equiv} \triangleq (\phi \to \psi)_{/\equiv}$
- $\neg\phi_{/\equiv} \triangleq (\neg\phi)_{/\equiv}$
- $\phi_{/\equiv} \wedge \psi_{/\equiv} \triangleq (\phi \wedge \psi)_{/\equiv}$
- $\phi_{/\equiv} \vee \psi_{/\equiv} \triangleq (\phi \vee \psi)_{/\equiv}$
- $\Box\phi_{/\equiv} \triangleq (\Box\phi)_{/\equiv}$
- $\diamond\phi_{/\equiv} \triangleq (\diamond\phi)_{/\equiv}$

From now on, the class $\phi_{/\equiv}$ is simply denoted ϕ.

Hyper-power set. Construct the subset of classical propositions $F(\Phi)$ recursively by the properties $\Phi \subset F(\Phi)$ and $\forall\phi, \psi \in F(\Phi)$, $[\,\phi\wedge\psi \in F(\Phi)$ and $\phi\vee\psi \in F(\Phi)\,]$. The hyper-power of Φ, denoted $<\Phi>$, is the set of equivalence classes of $F(\Phi)$ according to the relation $\equiv$:

$$< \Phi >= F(\Phi)_{/\equiv} = \{\phi_{/\equiv} \,/\, \phi \in F(\Phi)\} \,.$$

8.6.1.1 Useful theorems

Let $\phi, \psi \in M$.

1. $\vdash (\Box\phi \wedge \Box\psi) \to \Box(\phi \wedge \psi)$ and $\vdash \Box(\phi \wedge \psi) \to (\Box\phi \wedge \Box\psi)$

2. $\vdash (\diamond\phi \vee \diamond\psi) \to \diamond(\phi \vee \psi)$ and $\vdash \diamond(\phi \vee \psi) \to (\diamond\phi \vee \diamond\psi)$

3. $\vdash (\Box\phi \vee \Box\psi) \to \Box(\phi \vee \psi)$ **but** $\nvdash \Box(\phi \vee \psi) \to (\Box\phi \vee \Box\psi)$

4. $\vdash \diamond(\phi \wedge \psi) \to (\diamond\phi \wedge \diamond\psi)$ **but** $\nvdash (\diamond\phi \wedge \diamond\psi) \to \diamond(\phi \wedge \psi)$

Proof. Theorem 1 and theorem 2 are dual and thus equivalent (rules 7 and iv.). It is exactly the same thing for theorem 3 and theorem 4.

Proof of $\vdash (\Box\phi \wedge \Box\psi) \to \Box(\phi \wedge \psi)$.

Classical rules yield the axiom:

$\vdash \phi \rightarrow (\psi \rightarrow (\phi \wedge \psi))$

Rule i. implies then:

$\vdash \Box(\phi \rightarrow (\psi \rightarrow (\phi \wedge \psi)))$

Applying rule ii. twice, it is deduced:

$\vdash \Box\phi \rightarrow \Box(\psi \rightarrow (\phi \wedge \psi))$

$\vdash \Box\phi \rightarrow (\Box\psi \rightarrow \Box(\phi \wedge \psi))$

The proof is concluded by applying the classical rules.

Proof of $\vdash \Box(\phi \wedge \psi) \rightarrow (\Box\phi \wedge \Box\psi)$.

Classical rules yield the axioms:

$\vdash (\phi \wedge \psi) \rightarrow \phi$ and $\vdash (\phi \wedge \psi) \rightarrow \psi$

Rule i. implies then:

$\vdash \Box((\phi \wedge \psi) \rightarrow \phi)$ and $\vdash \Box((\phi \wedge \psi) \rightarrow \psi)$

Applying rule ii., it is deduced:

$\vdash \Box(\phi \wedge \psi) \rightarrow \Box\phi$ and $\vdash \Box(\phi \wedge \psi) \rightarrow \Box\psi$

The proof is concluded by applying the classical rules.

Proof of $\vdash (\Box\phi \vee \Box\psi) \rightarrow \Box(\phi \vee \psi)$.

Classical rules yield the axioms:

$\vdash \phi \rightarrow (\phi \vee \psi)$ and $\vdash \psi \rightarrow (\phi \vee \psi)$

Rule i. implies then:

$\vdash \Box(\phi \rightarrow (\phi \vee \psi))$ and $\vdash \Box(\psi \rightarrow (\phi \vee \psi))$

Applying rule ii., it is deduced:

$\vdash \Box\phi \rightarrow \Box(\phi \vee \psi)$ and $\vdash \Box\psi \rightarrow \Box(\phi \vee \psi)$

The proof is concluded by applying the classical rules.

Why $\nvdash \Box(\phi \vee \psi) \rightarrow (\Box\phi \vee \Box\psi)$?

To answer this question precisely, the Kripke semantic should be introduced. Such discussion is outside the scope of this paper. However, some practical considerations will clarify this assertion. When $\phi \vee \psi$ *is sure*, does that mean that ϕ *is sure* or ψ *is sure*? Not really since we know that ϕ or ψ is true, but *we do not know which one* is true. Moreover, it may happen that ϕ is true sometimes, while ψ is true the other times. As a conclusion, we are not sure of ϕ and are not sure of ψ.

This example is a counter-example of $\vdash \Box(\phi \vee \psi) \rightarrow (\Box\phi \vee \Box\psi)$.

$\Box\Box\Box$

8.6.2 A multi-modal logic

Assume that several informations are obtained from different sources. Typically, these informations are modalities such as *"according to the source σ, the proposition ϕ is sure"*. Such a modality could be naturally denoted $\Box_\sigma \phi$ (a modality depending on σ). A more readable notation $[\phi|\sigma] \stackrel{\Delta}{\equiv} \Box_\sigma \phi$ is prefered. Take note that $[\phi|\sigma]$ is not related to the Bayes inference $(\phi|\sigma)$! Now, the question arising is *how to combine these modalities?* For example, is it possible to deduce something from $[\phi_1|\sigma_1] \wedge [\phi_2|\sigma_2]$? Without any relations between *heterogeneous* modalities, it is not possible to answer this question. Such a relation, however, is foreseeable. Assume that the source τ involves the source σ, *ie.* $\tau \to \sigma$. Now assume that the proposition ϕ should be known from the source σ, *ie.* $[\phi|\sigma]$. Since τ involves σ, it is natural to state that ϕ should be known from the source τ, *ie.* $[\phi|\tau]$. This natural deduction could be formalized by the rule:

$$\vdash \tau \to \sigma \quad \text{implies} \quad \vdash [\phi|\sigma] \to [\phi|\tau]\ .$$

With this rule, it is now possible to define the logic.

The set of multi-modal propositions, $mM(\Phi)$ also denoted mM, is defined recursively:

- $C \subset mM$,
- If $\phi, \sigma \in mM$, then $[\phi|\sigma] \in mM$,
- If $\phi, \psi \in mM$, then $\neg\phi \in mM$, $\phi \wedge \psi \in mM$, $\phi \vee \psi \in mM$ and $\phi \to \psi \in mM$.

The multi-modal logic obeys to the following rules and axioms:

Modus Ponens.

Classical axioms. Axioms 1 to 7,

Modal axioms and rule. Let $\sigma, \tau, \phi, \psi \in mM$.

m.i. From $\vdash \phi$ is deduced $\vdash [\phi|\sigma]$: axioms are sure, according to any sources,

m.ii. $\vdash [\phi \to \psi|\sigma] \to ([\phi|\sigma] \to [\psi|\sigma])$. If a source of information asserts a proposition and recognizes a weaker proposition, then it asserts this weaker proposition,

m.iii. $\vdash [\phi|\sigma] \to \phi$. *The sources of information always tell the truth.* If a source asserts a proposition, this proposition is actually true,

m.iv. $\vdash \tau \to \sigma$ implies $\vdash [\phi|\sigma] \to [\phi|\tau]$. Knowledge increases with stronger sources of information.

The axiom m.iii. is questionable and may be changed. But the work presented in this paper is restricted to this axiom.

It is also possible to consider some exotic rules like $\phi \equiv [\phi|\bot]$, *ie.* a perfect source of information $\bot$ yields a perfect knowledge of the propositions ϕ. Similarly, the modality $[\phi|\top]$ could be interpreted as the proposition *"ϕ is an absolute truth"* or *"ϕ has a proof"*: one does not need any source of information to assert an absolute truth…

8.6.3 Some multi-modal theorems

8.6.3.1 Modal theorems map into multi-modal logic

Let $\mu \in M$ be a modal proposition. Let $\sigma \in mM$ be a multi-modal proposition. Let $\mu[\sigma] \in mM$ be the multi-modal proposition obtained by replacing $\Box$ by $[\cdot|\sigma]$ and $\diamond$ by $\neg[\neg \cdot |\sigma]$ in the proposition μ. Then $\vdash \mu$ implies $\vdash \mu[\sigma]$.

8.6.3.2 Useful multi-modal theorems

If the source σ asserts ϕ and the source τ asserts ψ, then the fused sources assert $\phi \wedge \psi$:

$$\boxed{\vdash \big([\phi|\sigma] \wedge [\psi|\tau]\big) \rightarrow [\phi \wedge \psi|\sigma \wedge \tau]}$$

Proof. From the axioms $\vdash (\sigma \wedge \tau) \rightarrow \sigma$ and $\vdash (\sigma \wedge \tau) \rightarrow \tau$, it is deduced:

$$\vdash [\phi|\sigma] \rightarrow [\phi|\sigma \wedge \tau] \,,$$

and

$$\vdash [\psi|\tau] \rightarrow [\psi|\sigma \wedge \tau] \,.$$

From the useful theorems proved for modal logic, it is deduced:

$$[\phi|\sigma \wedge \tau] \wedge [\psi|\sigma \wedge \tau] \equiv [\phi \wedge \psi|\sigma \wedge \tau] \,.$$

The proof is concluded by applying the classical rules.

□□□

If one of the sources σ or τ asserts ϕ, then the fused sources assert ϕ :

$$\vdash \big([\phi|\sigma] \vee [\phi|\tau]\big) \rightarrow [\phi|\sigma \wedge \tau] \,.$$

Proof. This results directly from $\vdash [\phi|\sigma] \rightarrow [\phi|\sigma \wedge \tau]$ and $\vdash [\phi|\tau] \rightarrow [\phi|\sigma \wedge \tau]$.

□□□

The converse is not necessarily true:

$$\nvdash [\phi|\sigma \wedge \tau] \rightarrow \big([\phi|\sigma] \vee [\phi|\tau]\big) \,.$$

In fact, *when sensors are not independent and possibly interactive, it is possible that the fused sensor* $\sigma\wedge\tau$ *works better than* σ *and* τ *separately!* On the other hand, this converse property could be considered as a necessary condition for the sensor independence. This discussion leads to the introduction of a new axiom, *the independence axiom* m.indep. :

$$\text{m.indep.}\ \ \vdash [\phi|\sigma\wedge\tau] \rightarrow \big([\phi|\sigma]\vee[\phi|\tau]\big)\ .$$

8.6.4 Sensor fusion

8.6.4.1 The context

Two sensors, σ and τ, are generating informations about a set of atomic propositions Φ . More precisely, the sensors will measure *independently* the probability for each proposition $\phi \in <\Phi>$ to be sure. In this section, it is discussed about fusing these sources of information.

This problem is clearly embedded in the multi-modal logic formalism. In particular, the modality $[\cdot|\sigma]$ characterizes the knowledge of σ about the universe $<\Phi>$. More precisely, the proposition $[\phi|\sigma]$ explains if ϕ is sure according to σ or not. This knowledge is probabilistic: the working data are the probabilities $p([\phi|\sigma])$ and $p([\phi|\tau])$ for $\phi \in <\Phi>$. The problem setting is more formalized in the next section.

Notation. From now on, the notation $p[\phi|\sigma]$ is used instead of $p([\phi|\sigma])$. *Beware* that $p[\phi|\sigma]$ *is not* the conditional probability $p(\phi|\sigma)$!

8.6.4.2 Sensors model and problem setting

The set of multi-modal propositions, $mM(\Theta)$, is constructed from the set $\Theta = \Phi\cup\{\sigma,\tau\}$. The propositions σ and τ are referring to the two independent sensors. The proposition $\sigma\wedge\tau$ is referring to the fused sensor. It is assumed for sure that $\bigvee_{\phi\in\Phi}\phi$ is true:

$$\vdash \left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\top\right] .$$

Consequently:

$$\left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\sigma\wedge\tau\right] \equiv \left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\sigma\right] \equiv \left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\tau\right] \equiv \left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\top\right] \equiv \top\ ,$$

and:

$$p\left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\sigma\wedge\tau\right] = p\left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\sigma\right] = p\left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\tau\right] = 1\ .$$

The sensors σ and τ are giving probabilistic informations about the certainty of the other propositions. More precisely, it is known:

$$p[\phi|\sigma] \quad\text{and}\quad p[\phi|\tau]\ ,\quad \text{for any } \phi\in<\Phi>\ .$$

Since the propositions σ and τ are referring to independent sensors, it is assumed that:

- The axiom m.indep. holds for σ and τ ,
- For any $\phi, \psi \in <\Phi>$, $p([\phi|\sigma] \wedge [\psi|\tau]) = p[\phi|\sigma] p[\psi|\tau]$.

A pertinent fused information is expected:

How to compute $p[\phi|\sigma \wedge \tau]$ for any $\phi \in <\Phi>$?

8.6.4.3 Constructing the fused belief

Defining tools. The useful propositions $\phi^{(\sigma)}$ are defined for any $\phi \in <\Phi>$:

$$\phi^{(\sigma)} \stackrel{\Delta}{\equiv} [\phi|\sigma] \wedge \neg \left(\bigvee_{\psi \in <\Phi> : \psi < \phi} [\psi|\sigma] \right) .$$

The same propositions $\phi^{(\tau)}$ are defined for τ :

$$\phi^{(\tau)} \stackrel{\Delta}{\equiv} [\phi|\tau] \wedge \neg \left(\bigvee_{\psi \in <\Phi> : \psi < \phi} [\psi|\tau] \right) .$$

Properties.

The propositions $\phi^{(\sigma)}$ are exclusive:

$$\phi^{(\sigma)} \wedge \psi^{(\sigma)} \equiv \bot , \quad \text{for any } \phi \not\equiv \psi .$$

Proof. Since $[\phi|\sigma] \wedge [\psi|\sigma] \equiv [\phi \wedge \psi|\sigma]$, it is deduced:

$$\phi^{(\sigma)} \wedge \psi^{(\sigma)} \equiv [\phi \wedge \psi|\sigma] \wedge \neg \left(\bigvee_{\eta : \eta < \phi} [\eta|\sigma] \right) \wedge \neg \left(\bigvee_{\eta : \eta < \psi} [\eta|\sigma] \right) .$$

It follows:

$$\phi^{(\sigma)} \wedge \psi^{(\sigma)} \equiv [\phi \wedge \psi|\sigma] \wedge \left(\bigwedge_{\eta : \eta < \phi} \neg [\eta|\sigma] \right) \wedge \left(\bigwedge_{\eta : \eta < \psi} \neg [\eta|\sigma] \right) .$$

Since $\phi \wedge \psi < \phi$ or $\phi \wedge \psi < \psi$ when $\phi \not\equiv \psi$, the property is deduced.

□□□

Lemma: $\bigvee_{\psi : \psi < \phi} [\psi|\sigma] \leq [\phi|\sigma]$.

Proof. The property $\psi < \phi$ implies successively $\vdash \psi \rightarrow \phi$, $\vdash [\psi \rightarrow \phi|\sigma]$ and $\vdash [\psi|\sigma] \rightarrow [\phi|\sigma]$. The lemma is then deduced.

□□□

The propositions $\phi^{(\sigma)}$ are exhaustive:

$$\bigvee_{\psi : \psi \leq \phi} \psi^{(\sigma)} \equiv [\phi|\sigma] , \quad \text{and in particular:} \quad \bigvee_{\phi \in <\Phi>} \phi^{(\sigma)} \equiv \top .$$

Proof. The proof is recursive. First, the smallest element of $<\Phi>$ is $\mu \equiv \bigwedge_{\phi\in<\Phi>}\phi$ and verifies:

$$\mu^{(\sigma)} \equiv [\mu|\sigma]\,.$$

Secondly:

$$\bigvee_{\psi:\psi\leq\phi}\psi^{(\sigma)} \equiv \phi^{(\sigma)}\vee\left(\bigvee_{\psi:\psi<\phi}\psi^{(\sigma)}\right) \equiv \phi^{(\sigma)}\vee\left(\bigvee_{\psi:\psi<\phi}\;\bigvee_{\eta:\eta\leq\psi}\eta^{(\sigma)}\right) \equiv \phi^{(\sigma)}\vee\left(\bigvee_{\psi:\psi<\phi}[\psi|\sigma]\right).$$

Since $\phi^{(\sigma)} \equiv [\phi|\sigma]\wedge\neg\left(\bigvee_{\psi:\psi<\phi}[\psi|\sigma]\right)$ and $\bigvee_{\psi:\psi<\phi}[\psi|\sigma] \leq [\phi|\sigma]$, it follows:

$$\bigvee_{\psi:\psi\leq\phi}\psi^{(\sigma)} \equiv [\phi|\sigma]\,.$$

The second part of the property results from:

$$\bigvee_{\phi\in<\Phi>}\phi^{(\sigma)} \equiv \bigvee_{\psi:\psi\leq\bigvee_{\phi:\phi\in\Phi}\phi}\psi^{(\sigma)} \quad\text{and}\quad \left[\bigvee_{\phi\in\Phi}\phi\,\middle|\,\sigma\right] \equiv \top\,.$$

□□□

The propositions $\phi^{(\tau)}$ are also exclusive and exhaustive:

$$\phi^{(\tau)}\wedge\psi^{(\tau)} \equiv \bot\,,\quad \text{for any } \phi\not\equiv\psi\,,$$

and:

$$\bigvee_{\phi\in<\Phi>}\phi^{(\tau)} \equiv \top\,.$$

It is then deduced that *the propositions $\phi^{(\sigma)}\wedge\psi^{(\tau)}$ are exclusive and exhaustive:*[2]

$$\forall\phi_1,\psi_1,\phi_2,\psi_2 \in<\Phi>\,,\;\; (\phi_1,\psi_1)\not\equiv(\phi_2,\psi_2) \Longrightarrow \left(\phi_1^{(\sigma)}\wedge\psi_1^{(\tau)}\right)\wedge\left(\phi_2^{(\sigma)}\wedge\psi_2^{(\tau)}\right) \equiv \bot\,, \tag{8.9}$$

and:

$$\bigvee_{\phi,\psi\in<\Phi>}\left(\phi^{(\sigma)}\wedge\psi^{(\tau)}\right) \equiv \top\,. \tag{8.10}$$

From the properties (8.9) and (8.10), it results that the set:

$$\Sigma = \left\{\phi^{(\sigma)}\wedge\psi^{(\tau)}\,/\,\phi,\psi\in<\Phi>\right\}$$

is *a partition* of $\top$. This property is particularly interesting, since it makes possible the computation of the probability of any proposition factorized within Σ:

$$\forall\Lambda\subset\Sigma\,,\; p\left(\bigvee_{\phi\in\Lambda}\phi\right) = \sum_{\phi\in\Lambda}p(\phi)\,. \tag{8.11}$$

[2]The notation $(\phi_1,\psi_1)\equiv(\phi_2,\psi_2)$ means $\phi_1\equiv\phi_2$ and $\psi_1\equiv\psi_2$.

Factorizing $\phi^{(\sigma\wedge\tau)}$. It has been shown that $\vdash ([\phi|\sigma]\wedge[\psi|\tau])\rightarrow[\phi\wedge\psi|\sigma\wedge\tau]$. It follows:

$$\vdash \bigvee_{\phi\wedge\psi\leq\eta}([\phi|\sigma]\wedge[\psi|\tau])\rightarrow[\eta|\sigma\wedge\tau]\;, \tag{8.12}$$

The axiom m.indep. says $\vdash [\eta|\sigma\wedge\tau]\rightarrow\big([\eta|\sigma]\vee[\eta|\tau]\big)$. Since $\left[\bigvee_{\phi\in\Phi}\phi\middle|\sigma\right]\equiv\left[\bigvee_{\phi\in\Phi}\phi\middle|\tau\right]\equiv\top$, it is deduced:

$$\vdash [\eta|\sigma\wedge\tau]\rightarrow\left(\left([\eta|\sigma]\wedge\left[\bigvee_{\phi\in\Phi}\phi\middle|\tau\right]\right)\vee\left(\left[\bigvee_{\phi\in\Phi}\phi\middle|\sigma\right]\wedge[\eta|\tau]\right)\right).$$

At last

$$[\eta|\sigma\wedge\tau]\equiv\bigvee_{\phi\wedge\psi\leq\eta}([\phi|\sigma]\wedge[\psi|\tau])\;. \tag{8.13}$$

It is then deduced:

$$\phi^{(\sigma\wedge\tau)}\equiv[\phi|\sigma\wedge\tau]\wedge\neg\left(\bigvee_{\psi<\phi}[\psi|\sigma\wedge\tau]\right)\equiv\left(\bigvee_{\eta\wedge\zeta\leq\phi}([\eta|\sigma]\wedge[\zeta|\tau])\right)\wedge\neg\left(\bigvee_{\psi<\phi}\bigvee_{\eta\wedge\zeta\leq\psi}([\eta|\sigma]\wedge[\zeta|\tau])\right).$$

Now:

$$\bigvee_{\eta\wedge\zeta\leq\phi}([\eta|\sigma]\wedge[\zeta|\tau])\equiv\bigvee_{\eta\wedge\zeta\leq\phi}\left(\left(\bigvee_{\xi\leq\eta}\xi^{(\sigma)}\right)\wedge\left(\bigvee_{\chi\leq\zeta}\chi^{(\tau)}\right)\right)$$
$$\equiv\bigvee_{\eta\wedge\zeta\leq\phi}\bigvee_{\xi\leq\eta}\bigvee_{\chi\leq\zeta}\big(\xi^{(\sigma)}\wedge\chi^{(\tau)}\big)\equiv\bigvee_{\eta\wedge\zeta\leq\phi}\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\;.$$

At last:

$$\phi^{(\sigma\wedge\tau)}\equiv\left(\bigvee_{\eta\wedge\zeta\leq\phi}\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\right)\wedge\neg\left(\bigvee_{\psi<\phi}\bigvee_{\eta\wedge\zeta\leq\psi}\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\right)$$
$$\equiv\left(\bigvee_{\eta\wedge\zeta\leq\phi}\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\right)\wedge\neg\left(\bigvee_{\eta\wedge\zeta<\phi}\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\right).$$

Since Σ is a partition, it is deduced the final results:

$$\boxed{\phi^{(\sigma\wedge\tau)}\equiv\bigvee_{\eta\wedge\zeta\equiv\phi}\eta^{(\sigma)}\wedge\zeta^{(\tau)}} \tag{8.14}$$

and:

$$p\big(\phi^{(\sigma\wedge\tau)}\big)=\sum_{\eta\wedge\zeta\equiv\phi}p\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\;. \tag{8.15}$$

This last result is sufficient to derive $p[\phi|\sigma\wedge\tau]$, as soon as we are able to compute the probability over Σ. The next paragraph explains this computation.

The probability over Σ.

Computing $p(\phi^{(\sigma)})$ and $p(\phi^{(\tau)})$. These probabilities are computed recursively from $p[\phi|\sigma]$ and $p[\phi|\tau]$. More precisely, it is derived from the definition $\phi^{(\sigma)}\equiv[\phi|\sigma]\wedge\neg\big(\bigvee_{\psi<\phi}[\psi|\sigma]\big)$ and the property $[\phi|\sigma]\equiv\bigvee_{\psi\leq\phi}\psi^{(\sigma)}$ that:

$$\phi^{(\sigma)}\equiv[\phi|\sigma]\wedge\neg\left(\bigvee_{\psi<\phi}\psi^{(\sigma)}\right).$$

Since the propositions $\phi^{(\sigma)}$ are exclusive and $\bigvee_{\psi<\phi}\psi^{(\sigma)}<[\phi|\sigma]$, it follows:

$$p\big(\phi^{(\sigma)}\big)=p[\phi|\sigma]-\sum_{\psi<\phi}p\big(\psi^{(\sigma)}\big)\;. \tag{8.16}$$

This equation, related to the *Moebius transform*, is sufficient for computing $p\big(\phi^{(\sigma)}\big)$ recursively.

Deriving $p\big(\phi^{(\sigma\wedge\tau)}\big)$. First, it is shown recursively that:

$$p\big(\phi^{(\sigma)}\wedge\psi^{(\tau)}\big)=p\big(\phi^{(\sigma)}\big)p\big(\psi^{(\tau)}\big)\;. \tag{8.17}$$

Proof - step 1. For the smallest element $\mu\equiv\bigwedge_{\phi\in<\Phi>}\phi$, it happens that $\mu^{(\sigma)}\equiv[\mu|\sigma]$ and $\mu^{(\tau)}\equiv[\mu|\tau]$. Since $[\mu|\sigma]$ and $[\mu|\tau]$ are independent propositions, it follows then $p(\mu^{(\sigma)}\wedge\mu^{(\tau)})=p\big(\mu^{(\sigma)}\big)p\big(\mu^{(\tau)}\big)$.

Proof - step 2. Being given $\phi,\psi\in<\Phi>$, assume $p\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)=p\big(\eta^{(\sigma)}\big)p\big(\zeta^{(\tau)}\big)$ for any $\eta,\zeta\in<\Phi>$ such that ($\eta\leq\phi$ and $\zeta<\psi$) or ($\eta<\phi$ and $\zeta\leq\psi$). From $[\phi|\sigma]\equiv\bigvee_{\eta\leq\phi}\eta^{(\sigma)}$ and $[\psi|\tau]\equiv\bigvee_{\zeta\leq\psi}\zeta^{(\tau)}$ it is deduced:

$$[\phi|\sigma]\wedge[\psi|\tau]\equiv\left(\bigvee_{\eta\leq\phi}\eta^{(\sigma)}\right)\wedge\left(\bigvee_{\zeta\leq\psi}\zeta^{(\tau)}\right)\equiv\bigvee_{\eta\leq\phi}\;\bigvee_{\zeta\leq\psi}\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\;.$$

It follows:

$$\begin{cases}p\big([\phi|\sigma]\wedge[\psi|\tau]\big)=\displaystyle\sum_{\eta\leq\phi}\sum_{\zeta\leq\psi}p\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)\\ p[\phi|\sigma]=\displaystyle\sum_{\eta\leq\phi}p\big(\eta^{(\sigma)}\big)\\ p[\psi|\tau]=\displaystyle\sum_{\zeta\leq\psi}p\big(\zeta^{(\tau)}\big)\end{cases}$$

Now, $[\phi|\sigma]$ and $[\psi|\tau]$ are independent and $p\big([\phi|\sigma]\wedge[\psi|\tau]\big)=p[\phi|\sigma]p[\psi|\tau]$. Then:

$$\sum_{\eta\leq\phi}\sum_{\zeta\leq\psi}p\big(\eta^{(\sigma)}\wedge\zeta^{(\tau)}\big)=\sum_{\eta\leq\phi}\sum_{\zeta\leq\psi}p\big(\eta^{(\sigma)}\big)p\big(\zeta^{(\tau)}\big)\;.$$

From the recursion assumption, it is deduced $p\big(\phi^{(\sigma)}\wedge\psi^{(\tau)}\big)=p\big(\phi^{(\sigma)}\big)p\big(\psi^{(\tau)}\big)$.

□□□

From the factorization (8.15), it is deduced:

$$\boxed{p\big(\phi^{(\sigma\wedge\tau)}\big)=\sum_{\eta\wedge\zeta\equiv\phi}p\big(\eta^{(\sigma)}\big)p\big(\zeta^{(\tau)}\big)} \tag{8.18}$$

This result relies strongly on the independence hypothesis about the sensors.

Back to $[\phi|\sigma\wedge\tau]$. Reminding that $[\phi|\sigma\wedge\tau]\equiv\bigvee_{\psi\leq\phi}\psi^{(\sigma\wedge\tau)}$, the fused probability $p[\phi|\sigma\wedge\tau]$ is deduced from $p\big(\psi^{(\sigma\wedge\tau)}\big)$ by means of the relation:

$$p[\phi|\sigma\wedge\tau]=\sum_{\psi\leq\phi}p\big(\psi^{(\sigma\wedge\tau)}\big)\;. \tag{8.19}$$

Conclusion. It is possible to derive an exact fused information $p[\phi|\sigma\wedge\tau]$, $\phi\in<\Phi>$ from the informations $p[\phi|\sigma]$, $\phi\in<\Phi>$ and $p[\phi|\tau]$, $\phi\in<\Phi>$ obtained from two *independent* sensors σ and τ . This derivation is done in 3 steps:

- Compute $p\big(\phi^{(\sigma)}\big)$ and $p\big(\phi^{(\tau)}\big)$ by means of (8.16) ,
- Compute $p\big(\phi^{(\sigma\wedge\tau)}\big)$ by means of (8.18) ,
- Compute $p[\phi|\sigma\wedge\tau]$ by means of (8.19) .

8.6.4.4 Link with the DSmT

It is noteworthy that the relation (8.18) looks strangely like the DSmT fusion rule (8.4) , although these two results have been obtained from quite different viewpoints. In fact the similarity is not just related to the fusion rule and the whole construction is identical. More precisely, let us now consider the problem from the DSmT viewpoint.

Let be defined for two sensors σ and τ the respective bba m_σ and m_τ over $<\Phi>$. The belief function associated to these two bba, denoted respectively Bel_σ and Bel_τ , are just verifying:

$$\mathrm{Bel}_\sigma(\phi)=\sum_{\psi\leq\phi}m_\sigma(\psi)\quad\text{and}\quad\mathrm{Bel}_\tau(\phi)=\sum_{\psi\leq\phi}m_\tau(\psi)\;.$$

Conversely, the bba m_σ is recovered by means of the recursion:

$$\forall\phi\in<\Phi>,\, m_\sigma(\phi)=\mathrm{Bel}_\sigma(\phi)-\sum_{\psi<\phi}m_\sigma(\psi)\;.$$

The fused bba $m_\sigma\oplus m_\tau$ is defined by:

$$m_\sigma\oplus m_\tau(\phi)=\sum_{\psi\wedge\eta\equiv\phi}m_\sigma(\psi)m_\tau(\eta)\,.$$

Now make the hypothesis that the probabilities $p[\phi|\sigma]$ and $p[\phi|\tau]$ are initialized for any $\phi\in<\Phi>$ by:

$$p[\phi|\sigma]=\mathrm{Bel}_\sigma(\phi)\quad\text{and}\quad p[\phi|\tau]=\mathrm{Bel}_\tau(\phi)\;.$$

Then, the following results are obviously obtained:

- $p\big(\phi^{(\sigma)}\big)=m_\sigma(\phi)$ and $p\big(\phi^{(\tau)}\big)=m_\tau(\phi)$,
- $p\big(\phi^{(\sigma\wedge\tau)}\big)=m_\sigma\oplus m_\tau(\phi)$,
- $p[\phi|\sigma\wedge\tau]=\mathrm{Bel}_\sigma\oplus\mathrm{Bel}_\tau(\phi)$, where $\mathrm{Bel}_\sigma\oplus\mathrm{Bel}_\tau$ is the belief function associated to $m_\sigma\oplus m_\tau$.

From this discussion, *it seems natural to consider the probabilized multi-modal logic mM as a possible logical interpretation of DSmT.*

Evaluate the consequence of the independence axiom. By using the axiom m.indep. , it is possible to prove (8.13) . Otherwise, it is only possible to prove (8.12) , which means that possibly more belief is put on the smallest propositions, in comparison with the independent sensors case. Such a property expresses a better and more precise knowledge about the world. Then it appears, accordingly to the mM interpretation of DSmT, that *the fusion rule $\oplus$ is an optimal rule only for fusing* <u>*independent*</u> *and (strictly) reliable sensors.*

8.7 Logical interpretation of the Bayes inference

Notation. In the sequel, $\phi \leftrightarrow \psi$ is just an equivalent notation for $(\psi \to \phi) \wedge (\phi \to \psi)$.

General discussion. The Bayes inference explains the probability of a proposition ψ , while is known a proposition ϕ . This probability is expressed as follows by the quantity $p(\psi|\phi)$:

$$p(\phi \wedge \psi) = p(\phi)p(\psi|\phi) \ .$$

From this implicit and probabilistic definition, $(\psi|\phi)$ appears more like a mathematical artifice than an actual "logical" operator. However, $(\psi|\phi)$ has clearly a meta-logical meaning although it is intuitive and just implied: it characterizes the knowledge about ψ , when a prior information ϕ is known. In this section, we are trying to interpret the Bayes operator $(\ |\ )$ as a logical operator. The author admits that this viewpoint seems extremely suspicious: the Bayes inference implies a change of the probabilistic universe, and then a change of the truth values! It makes no sense to put at the same level a conditional probability with an unconditional probability! But in fact, there are logics which handle multiple truths: the modal logics, and more precisely, the multi-modal logics. However, the model we are defining here is quite different from the usual modal models.

From now on, we are assuming a same logic involving the whole operators, *ie.* $\wedge$, $\neg$, $\vee$, $\to$ *and* $(\ |\ )$, and a *same* probability function p defined over the resulting propositions.

When defining a logic, a first step is perhaps to enumerate the intuitive properties the new logic should have, and then derive new language and rules. Since a probability is based on a Boolean algebra, this logic will include the classical logic. A first question arises then: *is the Bayes inference* $(\ |\ )$ *the same inference than in classical logic ?* More precisely, do we have $(\psi|\phi) \equiv \phi \to \psi$? If our logical model is coherent with the probability, this should imply:

$$p(\psi|\phi) = p(\phi \to \psi) = p(\neg\phi \vee \psi) \ .$$

Applying the Bayes rule, it is deduced:

$$p(\phi \wedge \psi) = p(\phi)p(\neg\phi \vee \psi) = \big(p(\phi \wedge \psi) + p(\phi \wedge \neg\psi)\big)\big(1 - p(\phi \wedge \neg\psi)\big) \, .$$

This is clearly false: *eg.* taking $p(\phi \wedge \neg\psi) = \frac{1}{4}$ and $p(\phi \wedge \psi) = \frac{1}{2}$ results in $\frac{1}{2} = \frac{9}{16}$! *The Bayes inference* $(\psi|\phi)$ *is not a classical inference.* Since it is a new kind of inference, we have to explain the meaning of this inference.

The Bayes inference seems to rely on the following principles:

- Any proposition ϕ induces a sub-universe, entirely characterized by the Bayes operator $(\cdot|\phi)$. For this reason, $(\cdot|\phi)$ could be seen as a conditional modality. But this modality possesses a strange quality: the implied sub-universe is essentially *classical*. From now on, $(\cdot|\phi)$ refers both to the modality and its induced sub-universe,
- The sub-universe $(\cdot|\top)$ is just the whole universe. The empty universe $(\cdot|\bot)$ is a singularity which cannot be manipulated,
- The sub-universe $(\cdot|\phi)$ is a projection of the sup-universe (which could be another sub-universe) into ϕ. In particular, the axioms of $(\cdot|\phi)$ result from the propositions which are axioms within the range ϕ in the sup-universe. Moreover, the modus ponens should work in the sub-universes,
- Any sub-proposition $(\psi|\phi)$ implies the infered proposition $\phi \to \psi$ in the sup-universe. This last point in not exactly the converse of the previous point. The previous point concerns axioms, while any possible propositions are considered here. This (modal-like) difference is necessary and makes the distinction between $(\,|\,)$ and $\to$,
- Since sub-universes are classical, the negation has a classical behavior: the double negation vanishes,
- The sub-universe of a sub-universe is the intersected sub-universe. For example, "considering blue animals within a universe of birds" means "considering blue birds".

In association with the Bayes inference is the notion of independence between propositions, described by the meta-operator $\times$, which is not an operator of the logic. More precisely, ψ is independent to ϕ, *ie.* $\psi \times \phi$, when it is equivalent to consider ψ within the sub-universe ϕ or within the sup-universe. Deciding whether this meta-operator is symmetric or not is probably another philosophical issue. In the sequel, this hypothesis is made possible in the axiomatization but is not required. Moreover, it seems reasonable that complementary propositions like ϕ and $\neg\phi$ cannot be independent unless $\phi \equiv \top$. In the following discussion, such a rule is proposed but not required.

8.7.1 Definitions

8.7.1.1 Bayesian modal language

The set of the *Bayesian* propositions bM is constructed recursively:

- $C \subset bM$,

- If $\phi, \psi \in bM$, then $(\psi|\phi) \in bM$,
- If $\phi, \psi \in bM$, then $\neg\phi \in bM$, $\phi \wedge \psi \in bM$, $\phi \vee \psi \in bM$ and $\phi \rightarrow \psi \in bM$.

8.7.1.2 Bayesian logical rules

The logic over bM obeys the following rules and axioms:

- Classical axioms and modus ponens ,

b.i. $(\phi|\top) \equiv \phi$; *the sub-universe of* $\top$ *is of course the whole universe,*

b.ii. It is assumed $\not\vdash \neg\phi$. Then, $\vdash \phi \rightarrow \psi$ implies $\vdash (\psi|\phi)$; *axioms within the range* ϕ *are axioms of the sub-universe* $(\cdot|\phi)$ *,*

b.iii. It is assumed $\not\vdash \neg\phi$. Then, $\vdash (\psi \rightarrow \eta|\phi) \rightarrow \big((\psi|\phi) \rightarrow (\eta|\phi)\big)$; *when both an inference and a premise are recognized true in a sub-universe, the conclusion also holds true in this sub-universe. This property allows the modus ponens within sub-universes,*

b.iv. It is assumed $\not\vdash \neg\phi$. Then, $\vdash (\psi|\phi) \rightarrow (\phi \rightarrow \psi)$; *the modality* $(\cdot|\phi)$ *implies the truth within the range* ϕ*,*

b.v. It is assumed $\not\vdash \neg\phi$. Then, $\neg(\neg\psi|\phi) \equiv (\psi|\phi)$; *there is no doubt within the modality* $(\cdot|\phi)$ *. Sub-universes have a classical negation operator. However, truth may change depending on the proposition of reference* ϕ *,*

b.vi. It is assumed $\not\vdash \neg(\phi \wedge \psi)$.[3] Then, $\big((\eta|\psi)\big|\phi\big) \equiv (\eta|\psi \wedge \phi)$; *the sub-universe* $(\cdot|\psi)$ *of a sub-universe* $(\cdot|\phi)$ *is the intersected sub-universe* $(\cdot|\phi \wedge \psi)$ *,*

b.vii. $\psi \times \phi$ means $\vdash (\psi|\phi) \leftrightarrow \psi$; *a proposition* ψ *is independent to a proposition* ϕ *when it makes no difference to observe it in the sub-universe* $(\cdot|\phi)$ *or not,*

b.viii. (optional) $\psi \times \phi$ implies $\phi \times \psi$; *the independence relation is symmetric,*

b.ix. (optional) Assuming $\phi \times \psi$ and $\vdash \phi \vee \psi$ implies $\vdash \phi$ or $\vdash \psi$; *this uncommon logical rule explains that complementary and non trivial propositions cannot be independent.* EG. *to an extreme degree,* ϕ *and* $\neg\phi$ *are strictly complementary and at the same time are not independent unless* $\phi \equiv \top$ *or* $\phi \equiv \bot$*.*

These axioms leave the modality $(\cdot|\bot)$ undefined, by requiring the condition $\not\vdash \neg\phi$ for any deduction on the sub-universe $(\cdot|\phi)$. In fact, the modality $(\cdot|\bot)$ is a singularity which cannot be defined according to the common axioms and rules. Otherwise, it would be deduced from $\vdash \bot \rightarrow \phi$ that $\vdash (\phi|\bot)$; this last deduction working for any ϕ would contradict the negation rule $\neg(\neg\phi|\bot) \equiv (\phi|\bot)$. Nevertheless, the axioms **b.vii.** and **b.viii.** induces a definition of $\times$ for any pair of propositions, except $(\bot, \bot)$.

[3] It will be proved that the hypothesis $\not\vdash \neg(\phi \wedge \psi)$ implies the hypotheses $\not\vdash \neg\phi$ and $\not\vdash (\neg\psi|\phi)$.

8.7.2 Properties

8.7.2.1 Probability over bM

A probability p over bM is defined according to the definition of section 8.4. In particular, since the meta-operator $\times$ characterizes an independence between propositions, it is naturally hypothesized that:

$$\phi \times \psi \quad \text{implies} \quad p(\phi \wedge \psi) = p(\phi)p(\psi) \ .$$

8.7.2.2 Useful theorems

Sub-universes are classical. It is assumed $\nvdash \neg\phi$. Then:

- $(\neg\psi|\phi) \equiv \neg(\psi|\phi)$,
- $(\psi \wedge \eta|\phi) \equiv (\psi|\phi) \wedge (\eta|\phi)$,
- $(\psi \vee \eta|\phi) \equiv (\psi|\phi) \vee (\eta|\phi)$,
- $(\psi \rightarrow \eta|\phi) \equiv (\psi|\phi) \rightarrow (\eta|\phi)$.

Proof. The first theorem is a consequence of axiom **b.v.**

From axiom **b.iii.**, it is deduced $\vdash (\neg\psi \vee \eta|\phi) \rightarrow \big(\neg(\psi|\phi) \vee (\eta|\phi)\big)$. Applying the first theorem, it is deduced $\vdash (\neg\psi \vee \eta|\phi) \rightarrow \big((\neg\psi|\phi) \vee (\eta|\phi)\big)$. At last:

$$\vdash (\psi \vee \eta|\phi) \rightarrow \big((\psi|\phi) \vee (\eta|\phi)\big) \ . \tag{8.20}$$

It is deduced $\vdash \neg\big((\psi|\phi) \vee (\eta|\phi)\big) \rightarrow \neg(\psi \vee \eta|\phi)$ and, by applying the first theorem,

$$\vdash \big((\neg\psi|\phi) \wedge (\neg\eta|\phi)\big) \rightarrow (\neg\psi \wedge \neg\eta|\phi) \ .$$

At last:

$$\vdash \big((\psi|\phi) \wedge (\eta|\phi)\big) \rightarrow (\psi \wedge \eta|\phi) \ .$$

Now, it is deduced from $\vdash \phi \rightarrow \big((\psi \wedge \eta) \rightarrow \psi\big)$ that:

$$\vdash \big((\psi \wedge \eta) \rightarrow \psi|\phi\big) \ .$$

By applying the axiom **b.iii.**:

$$\vdash (\psi \wedge \eta|\phi) \rightarrow (\psi|\phi) \ .$$

It is similarly proved that $\vdash (\psi \wedge \eta|\phi) \rightarrow (\eta|\phi)$ and finally:

$$\vdash (\psi \wedge \eta|\phi) \rightarrow \big((\psi|\phi) \wedge (\eta|\phi)\big) \ .$$

The second theorem is then proved.

Third theorem is a consequence of the first and second theorem.

Last theorem is a consequence of the first and third theorem.

□□□

Inference property. It is assumed $\nvdash \neg\phi$. Then $(\psi|\phi)\wedge\phi \equiv \phi\wedge\psi$. In particular, the hypothesis $\nvdash \neg(\phi\wedge\psi)$ implies the hypotheses $\nvdash \neg\phi$ and $\nvdash (\neg\psi|\phi)$.

Proof. From **b.iv.** it comes $\vdash (\psi|\phi)\to(\phi\to\psi)$. Then $\vdash \neg(\phi\to\psi)\to\neg(\psi|\phi)$ and $\vdash (\phi\wedge\neg\psi)\to(\neg\psi|\phi)$.

It follows $\vdash (\phi\wedge\psi)\to(\psi|\phi)$ and finally:

$$\vdash (\phi\wedge\psi)\to\big((\psi|\phi)\wedge\phi\big)\ .$$

The converse is more simple. From $\vdash (\psi|\phi)\to(\phi\to\psi)$, it follows:

$$\vdash \big((\psi|\phi)\wedge\phi\big)\to\big((\phi\to\psi)\wedge\phi\big)\ .$$

Since $(\phi\to\psi)\wedge\phi\equiv\phi\wedge\psi$, the converse is proved.

□□□

Intra-independence. It is assumed $\nvdash \neg\phi$. Then $(\eta|\phi)\times(\psi|\phi)$ is equivalently defined by the property $\vdash \big((\eta|\psi)\leftrightarrow\eta\big|\phi\big)$.

Proof.

$$\big((\eta|\psi)\leftrightarrow\eta\big|\phi\big)\equiv\big((\eta|\psi)\big|\phi\big)\leftrightarrow(\eta|\phi)\equiv(\eta|\phi\wedge\psi)\leftrightarrow(\eta|\phi)$$
$$\equiv\big(\eta\big|\phi\wedge(\psi|\phi)\big)\leftrightarrow(\eta|\phi)\equiv\big((\eta|\phi)\big|(\psi|\phi)\big)\leftrightarrow(\eta|\phi)\ .$$

□□□

Independence invariant. $\psi\times\phi$ implies $\neg\psi\times\phi$.

Proof.

$$(\neg\psi|\phi)\leftrightarrow\neg\psi\equiv\neg(\psi|\phi)\leftrightarrow\neg\psi\equiv(\psi|\phi)\leftrightarrow\psi\ .$$

□□□

Inter-independence. It is assumed $\nvdash \neg\phi$. Then $(\psi|\phi)\times\phi$.

Proof. From axiom **b.vi.**:

$$((\psi|\phi)|\phi)\equiv(\psi|\phi\wedge\phi)\equiv(\psi|\phi)\ .$$

It is deduced $(\psi|\phi)\times\phi$.

□□□

Corollary: assuming the rules b.viii. and b.ix., the hypotheses $\nvdash \neg\phi$ and $\nvdash (\neg\psi|\phi)$ imply the hypothesis $\nvdash \neg(\phi\wedge\psi)$.

Proof. Assume $\vdash \neg(\phi\wedge\psi)$. Then $\vdash \neg\big(\phi\wedge(\psi|\phi)\big)$ and $\vdash \neg\phi\vee\neg(\psi|\phi)$. Since $(\neg\psi|\phi)\times\phi$, it follows $\phi\times(\neg\psi|\phi)$ from rule **b.viii.** And then $\neg\phi\times\neg(\psi|\phi)$. Now, applying the rule **b.ix.** to $\vdash \neg\phi\vee\neg(\psi|\phi)$, it is deduced $\vdash \neg\phi$ or $\vdash \neg(\psi|\phi)$.

□□□

A proposition is true in its proper sub-universe. It is assumed $\nvdash \neg\phi$. Then $\vdash (\phi|\phi)$.

Proof. Obvious from $\vdash \phi \to \phi$.

□□□

Narcissist independence. It is assumed $\nvdash \neg\phi$. Then, $\phi\times\phi$ implies $\vdash \phi$ and conversely . In particular, $\phi \times \phi$ implies $\phi \equiv \top$.

Proof.

$$(\phi|\phi) \leftrightarrow \phi \equiv \top \leftrightarrow \phi \equiv \phi \,.$$

□□□

Non transitivity (modus barbara fails). It is assumed $\nvdash \neg\phi$ and $\nvdash \neg\psi$. Then

$$\nvdash (\psi|\phi) \to \big((\eta|\psi) \to (\eta|\phi)\big) \,.$$

Proof. The choice $\psi \equiv \top$, $\eta \equiv \neg\phi$ and $\phi \not\equiv \top$ is a counter example:

$$(\top|\phi) \to \big((\neg\phi|\top) \to (\neg\phi|\phi)\big) \equiv \top \to (\neg\phi \to \bot) \equiv \phi \,.$$

□□□

8.7.2.3 Axioms and rules extend to sub-universes

Assume $\nvdash \neg\phi$. The rules and axioms of bM extend on the sub-universe $(\cdot|\phi)$:

- $\vdash \psi$ implies $\vdash (\psi|\phi)$,
- It is assumed $\nvdash \neg(\phi\wedge\psi)$. Then $\vdash (\psi \to \eta|\phi)$ implies $\vdash \big((\eta|\psi)|\phi\big)$,
- It is assumed $\nvdash \neg(\phi\wedge\psi)$. Then $\vdash \big((\eta \to \zeta|\psi)\big|\phi\big) \to \big((\eta|\psi) \to (\zeta|\psi)\big|\phi\big)$,
- It is assumed $\nvdash \neg(\phi\wedge\psi)$. Then $\vdash \big((\eta|\psi)\big|\phi\big) \to (\psi \to \eta|\phi)$.

Proof. $\vdash \psi$ implies $\vdash \phi \to \psi$ and then $\vdash (\psi|\phi)$. *First point is then proved.*

It is successively implied from $\vdash (\psi \to \eta|\phi)$:

$\vdash (\psi|\phi) \to (\eta|\phi)$,

$\vdash \big((\eta|\phi)\big|(\psi|\phi)\big)$,

$\vdash \big(\eta\big|\phi \wedge (\psi|\phi)\big)$,

$\vdash (\eta|\phi\wedge\psi)$,

$\vdash \big((\eta|\psi)\big|\phi\big)$.

Second point is then proved.

By applying axiom **b.iii.** and first point, it comes:

$$\vdash \Big((\eta \to \zeta|\psi) \to \big((\eta|\psi) \to (\zeta|\psi) \big) \Big| \phi \Big) \;.$$

It follows:

$$\vdash \big((\eta \to \zeta|\psi) \big| \phi \big) \to \big((\eta|\psi) \to (\zeta|\psi) \big| \phi \big) \;.$$

Third point is proved.

By applying axiom **b.iv.** and first point, it comes:

$$\vdash \big((\eta|\psi) \to (\psi \to \eta) \big| \phi \big) \;.$$

It follows:

$$\vdash \big((\eta|\psi) \big| \phi \big) \to (\psi \to \eta|\phi) \;.$$

Fourth point is proved.

□□□

8.7.2.4 Bayes inference

It is assumed $\nvdash \neg\phi$. Define $p(\psi|\phi)$ as an abbreviation for $p\big((\psi|\phi)\big)$. Then:

$$p(\psi|\phi)p(\phi) = p(\phi \wedge \psi) \;.$$

Proof. This result is implied by the theorems $(\psi|\phi) \wedge \phi \equiv \phi \wedge \psi$ and $(\psi|\phi) \times \phi$.

□□□

8.7.2.5 Conclusion

Finally, the Bayes inference has been recovered from our axiomatization of the operator $(\cdot|\cdot)$. Although this result needs more investigation, in particular for the justification of the coherence of bM, it appears that the Bayesian inference could be interpreted logically as a manner to handle the knowledges. A similar result has been obtained for the fusion rule of DSmT. At last, it seems possible to conjecture that logics and probability could be mixed in order to derive many other belief rules or inferences.

8.8 Conclusion

In this contribution, it has been shown that DSmT was interpretable in the paradigm of probabilized multi-modal logic. This logical characterization has made apparent the true necessity of an independence hypothesis about the sensors, when applying the $\oplus$ fusion rule. Moreover, it is expected that our work has given some clarifications about the semantic associated with the conjunctive rule of DSmT.

A similar logical interpretation of the Bayes inference has been constructed, although this preliminary work should be improved. At last, it seems possible to handle probabilized logics as a relatively general framework for manipulating non deterministic informations. This is perhaps a generic method for constructing new customized belief theories. The principle is first to construct a logic well adapted to the problem, second to probabilize this logic, and third to derive the implied new belief theory (and forget then the mother logic!) :

$$\begin{array}{ccc} \text{Classical Logic} & & \text{Probability} \\ \downarrow & \overset{\text{Probabilized}}{\underset{\text{propositions}}{\Longrightarrow}} & \downarrow \\ \text{New Logic} & & \text{New Belief Theory} \end{array}$$

It seems obviously that there could be many theories and rules for manipulating non deterministic informations. This is not a new result and I feel necessary to refer to the works of *Sombo, Lefèvre, De Brucq and al.* [6, 4, 7] , which have investigated such questions.

At last, a common framework for both DSmT and Bayesian inference could be certainly derived by fusing the logics mM and bM .

Acknowledgment

I am grateful to Doctor Dezert and Professor Smarandache for their interest in this work.

8.9 References

[1] Dempster A. P., *Upper and lower probabilities induced by a multiple valued mapping*, Ann. Math. Statistics, no. 38, pp. 325–339, 1967.

[2] Dempster A. P., *A generalization of Bayesian inference*, J. Roy. Statist. Soc., Vol. B, No. 30, pp. 205–247, 1968.

[3] Dezert J., *Foundations for a new theory of plausible and paradoxical reasoning*, Information & Security, An international Journal, edited by Prof. Tzv. Semerdjiev, CLPP, Bulg. Acad. of Sciences, Vol. 9, pp. 13-57, 2002.

[4] Lefèvre E., Colot O., Vannoorenberghe P., *Belief functions combination and conflict management*, Information Fusion Journal, Elsevier, 2002.

[5] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[6] Sombo A., *Thesis Manuscript: Contribution à la modélisation mathématique de l'imprécis et de l'incertain*, Applied Mathematics, University of Rouen, France, 2001.

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

[8] Smets Ph., *The combination of evidences in the transferable belief model*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, no. 5, pp. 447–458, 1990.

[9] Voorbraak F., *Partial probability: theory and applications*, ISIPT A'99 First International Symposium on Imprecise Probabilities, Ghent, Belgium, pp. 655–662, 1999.

Chapter 9

On conjunctive and disjunctive combination rules of evidence

Hongyan Sun and **Mohamad Farooq**

Department of Electrical & Computer Engineering

Royal Military College of Canada

Kingston, ON, Canada, K7K 7B4

Abstract: *In this chapter, the Dempster-Shafer (DS) combination rule is examined based on the multi-valued mapping (MVM) and the product combination rule of multiple independent sources of information. The shortcomings in DS rule are correctly interpreted via the product combination rule of MVM. Based on these results, a new justification of the disjunctive rule is proposed. This combination rule depends on the logical judgment of OR and overcomes the shortcomings of DS rule, especially, in the case of the counter-intuitive situation. The conjunctive, disjunctive and hybrid combination rules of evidence are studied and compared. The properties of each rule are also discussed in details. The role of evidence of each source of information, the comparison of the combination judgment belief and ignorance of each rule, the treatment of conflicting judgments given by sources, and the applications of combination rules are discussed. The new results yield valuable theoretical insight into the rules that can be applied to a given situation. Zadeh's example is also included in this chapter for the evaluation of the performance and the efficiency of each combination rule of evidence in case of conflicting judgments.*

9.1 Introduction

Combination theory of multiple sources of information is always an important area of research in information processing of multiple sources. The initial important contribution in this area is due to Dempster in terms of Dempster's rule [1]. Dempster derived the combination rule for multiple independent sources of information based on the product space of multiple sources of information and multi-valued mappings. In the product space, combination-mapping of multiple multi-valued mappings is defined as the intersection of each multi-valued mapping, that is, an element can be judged by combination sources of information if and only if it can be judged by each source of information simultaneously, irrespective of the magnitude of the basic judgment probability. Shafer extended Dempster's theory to the space with all the subsets of a given set (i.e. the power set) and defined the frame of discernment, degree of belief, and, furthermore, proposed a new combination rule of the multiple independent sources of information in the form of Dempster-Shafer's (DS) combination rule [2]. However, the interpretation, implementation, or computation of the technique are not described in a sufficient detail in [2]. Due to the lack of details in [2], the literature is full of techniques to arrive at DS combination rule. For example, compatibility relations [3, 4], random subsets [5, 6, 7], inner probability [8, 9], joint (conjunction) entropy [10] etc. have been utilized to arrive at the results in [2]. In addition, the technique has been applied in various fields such as engineering, medicine, statistics, psychology, philosophy and accounting [11], and multi-sensor information fusion [12, 13, 14, 15, 16] etc. DS combination rule is more efficient and effective than the Bayesian judgment rule because the former does not require a priori probability and can process ignorance. A number of researchers have documented the drawbacks of DS techniques, such as the counter-intuitive results for some pieces of evidence [17, 18, 19], computational expenses and independent sources of information [20, 21].

One of the problems in DS combination rule of evidence is that the measure of the basic probability assignment of combined empty set is not zero, i.e. $m(\emptyset) \neq 0$, however, it is supposed to be zero, i.e. $m(\emptyset) = 0$. In order to overcome this problem, the remaining measure of the basic probability assignment is reassigned via the orthogonal technique [2]. This has created a serious problem for the combination of the two sharp sources of information, especially, when two sharp sources of information have only one of the same focal elements (i.e. two sources of information are in conflict), thus resulting in a counter-intuitive situation as demonstrated by Zadeh. In addition, DS combination rule cannot be applied to two sharp sources of information that have none of the same focal elements. These problems are not essentially due to the orthogonal factor in DS combination rule (see references [22, 23]).

In general, there are two main techniques to resolve the Shafer problem. One is to suppose $m(\emptyset) \neq 0$ or $m(\emptyset) > 0$ as it is in reality. The Smets transferable belief model (TBM), and Yager, Dubois &

Prade and Dezert-Smarandache (DSm) combination rules are the ones that utilize this fact in references [20, 24, 25, 26, 27, 28]. The other technique is that the empty set in the combined focal elements is not allowed and this idea is employed in the disjunctive combination rule [22, 23, 29, 30, 31]. Moreover, E. Lefèvre et al. propose a general combination formula of evidence in [32] and further conjunctive combination rules of evidence can been derived from it.

In this chapter, we present some of work that we have done in the combination rules of evidence. Based on a multi-valued mapping from a probability space (X, Ω, μ) to space S, a probability measure over a class 2^S of subsets of S is defined. Then, using the product combination rule of multiple information sources, Dempster-Shafer's combination rule is derived. The investigation of the two rules indicates that Dempster's rule and DS combination rule are for different spaces. Some problems of DS combination rule are correctly interpreted via the product combination rule that is used for multiple independent information sources. An error in multi-valued mappings in [11] is pointed out and proven.

Furthermore, a novel justification of the disjunctive combination rule for multiple independent sources of information based on the redefined combination-mapping rule of multiple multi-valued mappings in the product space of multiple independent sources of information is being proposed. The combination rule reveals a type of logical inference in the human judgment, that is, the OR rule. It overcomes the shortcoming of DS combination rule with the AND rule, especially, the one that is counter-intuitive, and provides a more plausible judgment than DS combination rule over different elements that are judged by different sources of information.

Finally, the conjunctive and disjunctive combination rules of evidence, namely, DS combination rule, Yager's combination rule, Dubois and Prade's (DP) combination rule, DSm's combination rule and the disjunctive combination rule, are studied for the two independent sources of information. The properties of each combination rule of evidence are discussed in detail, such as the role of evidence of each source of information in the combination judgment, the comparison of the combination judgment belief and ignorance of each combination rule, the treatment of conflict judgments given by the two sources of information, and the applications of combination rules. The new results yield valuable theoretical insight into the rules that can be applied to a given situation. Zadeh's example is included in the chapter to evaluate the performance as well as efficiency of each combination rule of evidence for the conflict judgments given by the two sources of information.

9.2 Preliminary

9.2.1 Source of information and multi-valued mappings

Consider n sources of information and corresponding multi-valued mappings [1]. They are mathematically defined by n basic probability spaces (X_i, Ω_i, μ_i) and multi-valued mappings Γ_i which assigns a subset $\Gamma_i x_i \subset S$ to every $x_i \in X_i$, $i = 1, 2, \ldots, n$. The space S into which Γ_i maps is the same for each i, namely: n different sources yield information about the same uncertain outcomes in S.

Let n sources be independent. Then based on the definition of the statistical independence, the combined sources (X, Ω, μ) can be defined as

$$X = X_1 \times X_2 \times \ldots \times X_n \tag{9.1}$$

$$\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n \tag{9.2}$$

$$\mu = \mu_1 \times \mu_2 \times \ldots \times \mu_n \tag{9.3}$$

for all $x \in X$, Γ is defined as

$$\Gamma x = \Gamma_1 x \cap \Gamma_2 x \cap \ldots \cap \Gamma_n x \tag{9.4}$$

The definition of Γ implies that $x_i \in X_i$ is consistent with a particular $s \in S$ if and only if $s \in \Gamma_i x_i$, for $i = 1, 2, \ldots, n$, and consequently $x = (x_1, x_2, \ldots, x_n) \in X$ is consistent with s if and only if $s \in \Gamma_i x_i$ for all $i = 1, 2, \ldots, n$ [1].

For finite $S = \{s_1, s_2, \ldots, s_n\}$, suppose $S_{\delta_1 \delta_2 \ldots \delta_n}$ denotes the subset of S which contains s_j if $\delta_j = 1$ and excludes s_j if $\delta_j = 0$, for $j = 1, 2, \ldots, n$. Then the 2^n subsets of S so defined are possible for all $\Gamma_i x_i$ $(i = 1, 2, \ldots, n)$, and partition X_i into

$$X_i = \bigcup_{\delta_1 \delta_2 \ldots \delta_m} X^{(i)}_{\delta_1 \delta_2 \ldots \delta_m} \tag{9.5}$$

where

$$X^{(i)}_{\delta_1 \delta_2 \ldots \delta_n} = \{x_i \in X_i, \Gamma_i x_i = S_{\delta_1 \delta_2 \ldots \delta_n}\} \tag{9.6}$$

and define [1]

$$p^{(i)}_{\delta_1 \delta_2 \ldots \delta_n} = \mu(X^{(i)}_{\delta_1 \delta_2 \ldots \delta_n}) \tag{9.7}$$

9.2.2 Dempster's combination rule of independent information sources

Based on (9.1) - (9.7), the combination of probability judgments of multiple independent information sources is characterized by [1] $p^{(i)}_{\delta_1\delta_2\ldots\delta_n}$, $i = 1, 2, \ldots, n$. That is

$$p_{\delta_1\delta_2\ldots\delta_n} = \sum_{\delta_i=\delta_i^{(1)}\delta_i^{(2)}\ldots\delta_i^{(n)}} p^{(1)}_{\delta_1^{(1)}\delta_2^{(1)}\ldots\delta_n^{(1)}} p^{(2)}_{\delta_1^{(2)}\delta_2^{(2)}\ldots\delta_n^{(2)}} \cdots p^{(n)}_{\delta_1^{(n)}\delta_2^{(n)}\ldots\delta_n^{(n)}} \tag{9.8}$$

Equation (9.8) indicates that the combination probability judgment of n independent information sources for any element $S_{\delta_1\delta_2\ldots\delta_n}$ of S equals the sum of the product of simultaneously doing probability judgment of each independent information source for the element. It emphasizes the common role of each independent information source. That is characterized by the product combination rule.

9.2.3 Degree of belief

Definition 1:

If Θ is a frame of discernment, then function $m : 2^\Theta \to [0,1]$ is called[1] a *basic belief assignment* whenever

$$m(\emptyset) = 0 \tag{9.9}$$

and

$$\sum_{A\subseteq\Theta} m(A) = 1 \tag{9.10}$$

The quantity $m(A)$ is called the belief mass of A (or *basic probability number* in [2]).

Definition 2:

A function $\mathrm{Bel} : 2^\Theta \to [0,1]$ is called a *belief function* over Θ [2] if it is given by

$$\mathrm{Bel}(A) = \sum_{B\subseteq A} m(B) \tag{9.11}$$

for some basic probability assignment $m : 2^\Theta \to [0,1]$.

Definition 3:

A subset A of a frame Θ is called a *focal element* of a belief function Bel over Θ [2] if $m(A) > 0$. The union of all the focal elements of a belief function is called its *core*.

Theorem 1:

If Θ is a frame of discernment, then a function $\mathrm{Bel} : 2^\Theta \to [0,1]$ is a belief function if and only if it satisfies the three following conditions [2]:

[1] also called *basic probability assignment* in [2].

1.
$$\mathrm{Bel}(\emptyset) = 0 \tag{9.12}$$

2.
$$\mathrm{Bel}(\Theta) = 1 \tag{9.13}$$

3. For every positive integer n and every collection $A_1, \ldots, A_n$ of subsets of Θ,

$$\mathrm{Bel}(A_1 \cup \ldots \cup A_n) = \sum_{\substack{I \subset \{1,\ldots,n\} \\ I \neq \emptyset}} (-1)^{|I|+1} \mathrm{Bel}(\cap_{i \in I} A_i) \tag{9.14}$$

Definition 4:

The function $\mathrm{Pl} : 2^\Theta \to [0,1]$ defined by

$$\mathrm{Pl}(A) = 1 - \mathrm{Bel}(\bar{A}) \tag{9.15}$$

is called the *plausibility function* for Bel. $\bar{A}$ denotes the complement of A in 2^Θ.

Definition 5:

If Θ is a frame of discernment, then a function $\mathrm{Bel} : 2^\Theta \to [0,1]$ is called *Bayesian belief* [2] if and only if

1. $\mathrm{Bel}(\emptyset) = 0$ (9.16)
2. $\mathrm{Bel}(\Theta) = 1$ (9.17)
3. If $A, B \subset \Theta$ and $A \cap B = \emptyset$, then $\mathrm{Bel}(A \cup B) = \mathrm{Bel}(A) + \mathrm{Bel}(B)$ (9.18)

Theorem 2:

If $\mathrm{Bel} : 2^\Theta \to [0,1]$ is a belief function over Θ, Pl is a plausibility corresponding to it, then the following conclusions are equal [2]

1. The belief is a Bayesian belief.

2. Each focal element of Bel is a single element set.

3. $\forall A \subset \Theta$, $\mathrm{Bel}(A) + \mathrm{Bel}(\bar{A}) = 1$.

9.2.4 The DS combination rule

Theorem 3:

Suppose Bel_1 and Bel_2 are belief functions over the same frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ with basic belief assignments m_1 and m_2, and focal elements $A_1, A_2, \ldots, A_k$ and $B_1, B_2, \ldots, B_l$, respectively. Suppose

$$\sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i) m_2(B_j) < 1 \tag{9.19}$$

Then the function $m : 2^\Theta \to [0,1]$ defined by $m(\emptyset) = 0$ and

$$m(A) = \frac{\sum\limits_{\substack{i,j\\ A_i \cap B_j = A}} m_1(A_i)m_2(B_j)}{1 - \sum\limits_{\substack{i,j\\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j)} \tag{9.20}$$

for all non-empty $A \subseteq \Theta$ is a basic belief assignment [2]. The core of the belief function given by m is equal to the intersection of the cores of Bel_1 and Bel_2. This defines Dempster-Shafer's rule of combination (denoted as the DS combination rule in the sequel).

9.3 The DS combination rule induced by multi-valued mapping

9.3.1 Definition of probability measure over the mapping space

Given a probability space (X, Ω, μ) and a space S with a multi-valued mapping:

$$\Gamma : X \to S \tag{9.21}$$

$$\forall x \in X, \Gamma x \subset S \tag{9.22}$$

The problem here is that if the uncertain outcome is known to correspond to an uncertain outcome $s \in \Gamma x$, then the probability judgement of the uncertain outcome $s \in \Gamma x$ needs to be determined.

Assume S consists of n elements, i.e. $S = \{s_1, s_2, \ldots, s_n\}$. Let's denote $S_{\delta_1\delta_2\ldots\delta_n}$ the subsets of S, where $\delta_i = 1$ or 0, $i = 1, 2, \ldots, n$, and

$$S_{\delta_1\delta_2\ldots\delta_n} = \bigcup_{i \neq j, \delta_i = 1, \delta_j = 0} s_i \tag{9.23}$$

then from mapping (9.21)-(9.22) it is evident that $S_{\delta_1\delta_2\ldots\delta_n}$ is related to Γx. Therefore, the 2^S subsets such as in equation (9.23) of S yield a partition of X

$$X = \bigcup_{\delta_1\delta_2\ldots\delta_n} X_{\delta_1\delta_2\ldots\delta_n} \tag{9.24}$$

where

$$X_{\delta_1\delta_2\ldots\delta_n} = \{x \in X, \Gamma x = S_{\delta_1\delta_2\ldots\delta_n}\} \tag{9.25}$$

Define a probability measure over $2^S = \{S_{\delta_1\delta_2\ldots\delta_n}\}$ as $M : 2^S = \{S_{\delta_1\delta_2\ldots\delta_n}\} \to [0,1]$ with

$$M(S_{\delta_1\delta_2\ldots\delta_n}) = \begin{cases} 0, S_{\delta_1\delta_2\ldots\delta_n} = \emptyset \\ \frac{\mu(X_{\delta_1\delta_2\ldots\delta_n})}{1-\mu(X_{00\ldots0})}, S_{\delta_1\delta_2\ldots\delta_n} \neq \emptyset \end{cases} \tag{9.26}$$

where M is the probability measure over a class $2^S = \{S_{\delta_1\delta_2\ldots\delta_n}\}$ of subsets of space S which Γ maps X into.

9.3.2 Derivation of the DS combination rule

Given two $n = 2$ independent information sources, then from equation (9.8), we have

$$\mu(X_{\delta_1\delta_2\ldots\delta_n}) = \sum_{\Gamma X_{\delta_1\delta_2\ldots\delta_n} = \Gamma^{(1)} X^{(1)}_{\delta'_1\delta'_2\ldots\delta'_n} \cap \Gamma^{(2)} X^{(2)}_{\delta''_1\delta''_2\ldots\delta''_n}} \mu^{(1)}(X^{(1)}_{\delta'_1\delta'_2\ldots\delta'_n})\mu^{(2)}(X^{(2)}_{\delta''_1\delta''_2\ldots\delta''_n}) \tag{9.27}$$

From equation (9.26), if $S_{\delta_1\delta_2\ldots\delta_n} \neq \emptyset$, we have for $i = 1, 2$

$$\mu^{(i)}(X_{\delta_1\delta_2\ldots\delta_n}) = M^{(i)}(S_{\delta_1\delta_2\ldots\delta_n})(1 - \mu^{(i)}(X_{00\ldots0})) \tag{9.28}$$

and

$$\mu(X_{\delta_1\delta_2\ldots\delta_n}) = M(S_{\delta_1\delta_2\ldots\delta_n})(1 - \mu(X_{00\ldots0})) \tag{9.29}$$

where equations (9.28) and (9.29) correspond to information source i, $(i = 1, 2)$ and their combined information sources, respectively. Substituting equations (9.28)-(9.29) into equation (9.27), we have

$$M(S_{\delta_1\delta_2\ldots\delta_n}) = \frac{\sum\limits_{\delta=\delta'\delta''} M^{(1)}(S_{\delta'_1\delta'_2\ldots\delta'_n})M^{(2)}(S_{\delta''_1\delta''_2\ldots\delta''_n})[1 - \mu^{(1)}(X^{(1)}_{00\ldots0})[1 - \mu^{(2)}(X^{(2)}_{00\ldots0})]}{1 - \mu(X_{00\ldots0})} \tag{9.30}$$

and

$$\begin{aligned}\frac{[1 - \mu^{(1)}(X^{(1)}_{00\ldots0})][1 - \mu^{(2)}(X^{(2)}_{00\ldots0})]}{1 - \mu(X_{00\ldots0})} &= \frac{[1 - \mu^{(1)}(X^{(1)}_{00\ldots0})][1 - \mu^{(2)}(X^{(2)}_{00\ldots0})]}{\sum\limits_{\Gamma_1 X^{(1)}_{\delta'_1\delta'_2\ldots\delta'_n} \cap \Gamma_2 X^{(2)}_{\delta''_1\delta''_2\ldots\delta''_n} \neq \emptyset} \mu^{(1)}(X^{(1)}_{\delta'_1\delta'_2\ldots\delta'_n})\mu^{(2)}(X^{(2)}_{\delta''_1\delta''_2\ldots\delta''_n})} \\ &= \frac{1}{\sum\limits_{S_{\delta'_1\delta'_2\ldots\delta'_n} \cap S_{\delta''_1\delta''_2\ldots\delta''_n} \neq \emptyset} M^{(1)}(S_{\delta'_1\delta'_2\ldots\delta'_n})M^{(2)}(S_{\delta''_1\delta''_2\ldots\delta''_n})}\end{aligned} \tag{9.31}$$

Substitute (9.31) back into (9.30), hence we have

$$M(S_{\delta_1\delta_2\ldots\delta_n}) = \frac{\sum\limits_{S_{\delta'_1\delta'_2\ldots\delta'_n} \cap S_{\delta''_1\delta''_2\ldots\delta''_n} = S_{\delta_1\delta_2\ldots\delta_n}} M^{(1)}(S_{\delta'_1\delta'_2\ldots\delta'_n})M^{(2)}(S_{\delta''_1\delta''_2\ldots\delta''_n})}{1 - \sum\limits_{S_{\delta'_1\delta'_2\ldots\delta'_n} \cap S_{\delta''_1\delta''_2\ldots\delta''_n} = \emptyset} M^{(1)}(S_{\delta'_1\delta'_2\ldots\delta'_n})M^{(2)}(S_{\delta''_1\delta''_2\ldots\delta''_n})} \tag{9.32}$$

when $S_{\delta_1\delta_2\ldots\delta_n} = \emptyset$,

$$M(S_{\delta_1\delta_2\ldots\delta_n}) \triangleq 0 \tag{9.33}$$

Thus, equations (9.32) and (9.33) are DS combination rule. Where space $S = \{s_1, s_2, \ldots, s_n\}$ is the frame of discernment.

The physical meaning of equations (9.8) and (9.32)-(9.33) is different. Equation (9.8) indicates the probability judgement combination in the combination space (X, Ω, μ) of n independent information sources, while equations (9.32)-(9.33) denotes the probability judgement combination in the mapping space $(S, 2^S, M)$ of n independent information sources. The mappings of Γ and Γ_i, $(i = 1, 2, \ldots, n)$ relate equations (9.8) and (9.32)-(9.33). This shows the difference between Dempster's rule and DS combination rule.

9.3.3 New explanations for the problems in DS combination rule

From the above derivation, it can be seen that DS combination rule is mathematically based on the product combination rule of multiple independent information sources as evident from equations (9.1)-(9.8). For each of the elements in the space, the combination probability judgement of independent information sources is the result of the simultaneous probability judgement of each independent information source. That is, if each information source yields simultaneously its probability judgement for the element, then the combination probability judgement for the element can be obtained by DS combination rule, regardless of the magnitude of the judgement probability of each information source. Otherwise, it is the opposite. This gives raise to the following problems:

1. *The counter-intuitive results*

   Suppose a frame of discernment is $S = \{s_1, s_2, s_3\}$, the probability judgments of two independent information sources, (X_i, Ω_i, μ_i), $i = 1, 2$, are m_1 and m_2, respectively. That is:

   $$(X_1, \Omega_1, \mu_1): \quad m_1(s_1) = 0.99, \quad m_1(s_2) = 0.01$$

   and

   $$(X_2, \Omega_2, \mu_2): \quad m_2(s_2) = 0.01, \quad m_2(s_3) = 0.99$$

   Using DS rule to combine the above two independent probability judgements, results in

   $$m(s_2) = 1, m(s_1) = m(s_3) = 0 \tag{9.34}$$

   This is counter-intuitive. The information source (X_1, Ω_1, μ_1) judges s_1 with a very large probability measure, 0.99, and judges s_2 with a very small probability measure, 0.01, while the information source (X_2, Ω_2, μ_2) judges s_3 with a very large probability measure, 0.99, and judges s_2 with a very small probability measure, 0.01. However, the result of DS combination rule is that s_2 occurs with probability measure, 1, and others occur with zero probability measure. The reason for this result is that the two information sources simultaneously give their judgement only for an element s_2 of space $S = \{s_1, s_2, s_3\}$ although the probability measures from the two information sources for the element are very small and equal to 0.01, respectively. The elements s_1 and s_3 are not judged by the two information sources simultaneously. According to the product combination rule, the result in equation (9.34) is as expected.

It should be pointed out that this counter-intuitive result is not completely due to the normalization factor in highly conflicting evidence [17, 18, 19] of DS combination rule. This can be proven by the following example.

Suppose for the above frame of discernment, the probability judgments of another two independent information sources, (X_i, Ω_i, μ_i), $i = 3, 4$, are m_2 and m_4, are chosen. That is:

$$(X_3, \Omega_3, \mu_3): \quad m_3(s_1) = 0.99, \quad m_3(S) = 0.01$$

and

$$(X_4, \Omega_4, \mu_4): \quad m_4(s_3) = 0.99, \quad m_4(S) = 0.01$$

The result of DS combination rule is

$$m'(s_1) = 0.4975, m'(s_3) = 0.4975, m'(S) = 0.0050$$

This result is very different from that in equation (9.34) although the independent probability judgements of the two information sources are also very conflicting for elements s_1 and s_3. That is, the information source, (X_3, Ω_3, μ_3), judges s_1 with a very large probability measure, 0.99, and judges S with a very small probability measure, 0.01, while the information source (X_4, Ω_4, μ_4) judges s_3 with a very large probability measure, 0.99, and judges S with a very small probability measure, 0.01.

This is due to the fact that the same element $S = \{s_1, s_2, s_3\}$ of the two information sources includes elements s_1 and s_3. So, the element s_1 in the information source, (X_3, Ω_3, μ_3), and the element $S = \{s_1, s_2, s_3\}$ in the information source, (X_4, Ω_4, μ_4) have the same information, and the element $S = \{s_1, s_2, s_3\}$ in information source, (X_3, Ω_3, μ_3), and the element s_3 in information source, (X_4, Ω_4, μ_4) have the same information. Thus, the two independent information sources can simultaneously give information for the same probability judgement element $S = \{s_1, s_2, s_3\}$, and also simultaneously yield the information for the conflicting elements s_1 and s_3, respectively. That is required by the product combination rule.

2. *The combination of Bayesian (sensitive) information sources*

If two Bayesian information sources cannot yield the information about any element of the frame of discernment simultaneously, then the two Bayesian information sources cannot be combined by DS combination rule. For example, there are two Bayesian information sources (X_1, Ω_1, μ_1)

and (X_2, Ω_2, μ_2) over the frame of discernment, $S = \{s_1, s_2, s_3, s_4\}$, and the basic probability assignments are, respectively,

$$(X_1, \Omega_1, \mu_1): \quad m_1(s_1) = 0.4, \quad m_1(s_2) = 0.6$$

and

$$(X_2, \Omega_2, \mu_2): \quad m_2(s_3) = 0.8, \quad m_2(s_4) = 0.2$$

then their DS combination rule is

$$m(s_1) = m(s_2) = m(s_3) = m(s_4) = 0$$

This indicates that every element of the frame of discernment occurs with zero basic probability after DS combination rule is applied. This is a conflict. This is because the source (X_1, Ω_1, μ_1) gives probability judgements for elements s_1 and s_2 of the frame of discernment, $S = \{s_1, s_2, s_3, s_4\}$, while the source (X_2, Ω_2, μ_2) gives probability judgements for elements s_3 and s_4 of the frame of discernment, $S = \{s_1, s_2, s_3, s_4\}$. The two sources cannot simultaneously give probability judgements for any element of the frame of discernment, $S = \{s_1, s_2, s_3, s_4\}$. Thus, the product combination rule does not work for this case.

Based on the above analysis, a possible solution to the problem is to relax the conditions used in the product combination rule (equations (9.1)-(9.4)) for practical applications, and establish a new theory for combining information of multiple sources (see sections 9.4 and 9.5).

9.3.4 Remark about "multi-valued mapping" in Shafer's paper

On page 331 of [11] where G. Shafer explains the concept of multi-valued mappings of DS combination rule, the Dempter's rule is considered as belief, $\mathrm{Bel}(T) = P\{x|\Gamma(x) \subseteq T, \forall T \subset S\}$, combination. The following proof shows this is incorrect.

Proof: Given the two independent information sources, equations (9.1)-(9.4) become as the followings:

$$X = X_1 \times X_2 \tag{9.35}$$

$$\Omega = \Omega_1 \times \Omega_2 \tag{9.36}$$

$$\mu = \mu_1 \times \mu_2 \tag{9.37}$$

$$\Gamma x = \Gamma_1 x \cap \Gamma_2 x \tag{9.38}$$

then

$$\mathrm{Bel}(T) \neq \mathrm{Bel}_1(T) \oplus \mathrm{Bel}_2(T)$$

in fact, $\forall T \subset S$,

$$\{\Gamma(x) \subseteq T\} \nRightarrow \{\Gamma(x_1) \subseteq T\} \cap \{\Gamma(x_2) \subseteq T\}$$

hence,

$$\{x \in X | \Gamma(x) \subseteq T\} \neq \{x_1 \in X_1 | \Gamma(x_1) \subseteq T\} \times \{x_2 \in X_2 | \Gamma(x_2) \subseteq T\}$$

i.e. the product combination rule in equations (9.35)-(9.38) is not satisfied by the defined belief $\mathrm{Bel}(T) = P\{x|\Gamma(x) \subseteq T, \forall T \subset S\}$. Therefore, the combination belief cannot be obtained from equations (9.35)-(9.38) with the belief, $\mathrm{Bel}(T) = P\{x|\Gamma(x) \subseteq T, \forall T \subset S\}$. When we examine the product combination rule in equations (9.1)-(9.4), it is known that the combination rule is neither for upper probabilities, nor for lower probabilities (belief), nor for probabilities of the type, $p_{\delta_1\delta_2\ldots\delta_n} = \mu(X_{\delta_1\delta_2\ldots\delta_n})$ [1]. It is simply for probability spaces of multiple independent information sources with multi-valued mappings.

9.4 A new combination rule of probability measures over mapping space

It has been demonstrated in section 9.3 that DS combination rule is mathematically based on the product combination rule of multiple independent information sources. The combination probability judgment of n independent information sources for each element is the result of the simultaneous probability judgment of each independent information source. That is, if each information source yields simultaneously its probability judgment for the element, then the combination probability judgment for the element can be obtained by DS combination rule regardless of the magnitude of the judgment probability of each information source. Otherwise, such results are not plausible. This is the main reason that led to the counter-intuitive results in [17, 18, 19]. We will redefine the combination-mapping rule Γ using n independent mapping Γ_i, $i = 1, 2, \ldots, n$ in order to relax the original definition in equation (9.4) in section 9.2.1. The combination of probabilities of type $p^{(i)}_{\delta_2\delta_1\ldots\delta_n}$ in the product space (X, Ω, μ) will then be realized, and, furthermore, the combination rule of multiple sources of information over mapping space S will also be established.

9.4.1 Derivation of combination rule of probabilities $p^{(i)}_{\delta_1\delta_2\ldots\delta_n}$

Define a new combination-mapping rule for multiple multi-valued mappings as

$$\Gamma x = \Gamma_1 x \cup \Gamma_2 x \cup \ldots \cup \Gamma_n x \tag{9.39}$$

It shows that $x_i \in X$ is consistent with a particular $s \in S$ if and only if $s \in \Gamma_i x_i$, for $i = 1, 2, \ldots, n$, and consequently $x = \{x_1, x_2, \ldots, x_n\} \in X$ is consistent with that s if and only if there exist certain $i \in \{1, 2, \ldots, n\}$, such that $s \in \Gamma_i x_i$.

For any $T \subset S$, we construct sets

$$\bar{T} = \{x \in X, \Gamma x \subset T\} \tag{9.40}$$

$$\bar{T}_i = \{x_i \in X_i, \Gamma_i x_i \subset T\} \tag{9.41}$$

and let

$$\lambda(T) = \mu(\bar{T}) \tag{9.42}$$

$$\lambda^{(i)}(T) = \mu_i(\bar{T}_i) \tag{9.43}$$

Hence,

$$\bar{T} = \bar{T}_1 \times \bar{T}_2 \times \ldots \times \bar{T}_n \tag{9.44}$$

and

$$\lambda(T) = \lambda^{(1)}(T) \times \lambda^{(2)}(T) \times \ldots \times \lambda^{(n)}(T) \tag{9.45}$$

Consider a finite $S = \{s_1, s_2, s_3\}$ and two independent sources of information characterized by $p_{000}^{(i)}$, $p_{100}^{(i)}$, $p_{010}^{(i)}$, $p_{001}^{(i)}$, $p_{110}^{(i)}$, $p_{101}^{(i)}$, $p_{011}^{(i)}$ and $p_{111}^{(i)}$, $i = 1, 2$. Suppose $\lambda^{(i)}(T)$, $(i = 1, 2)$ corresponding to $T = \emptyset$, $\{s_1\}$, $\{s_2\}$, $\{s_3\}$, $\{s_1, s_2\}$, $\{s_2, s_3\}$, $\{s_1, s_3\}$, $\{s_1, s_2, s_3\}$ is expressed as $\lambda_{000}^{(i)}$, $\lambda_{100}^{(i)}$, $\lambda_{010}^{(i)}$, $\lambda_{001}^{(i)}$, $\lambda_{110}^{(i)}$, $\lambda_{101}^{(i)}$, $\lambda_{011}^{(i)}$ and $\lambda_{111}^{(i)}$, $i = 1, 2$. Then for $i = 1, 2$,

$$\lambda_{000}^{(i)} = p_{000}^{(i)} \tag{9.46}$$

$$\lambda_{100}^{(i)} = p_{000}^{(i)} + p_{100}^{(i)} \tag{9.47}$$

$$\lambda_{010}^{(i)} = p_{000}^{(i)} + p_{010}^{(i)} \tag{9.48}$$

$$\lambda_{001}^{(i)} = p_{000}^{(i)} + p_{001}^{(i)} \tag{9.49}$$

$$\lambda_{110}^{(i)} = p_{000}^{(i)} + p_{100}^{(i)} + p_{010}^{(i)} + p_{110}^{(i)} \tag{9.50}$$

$$\lambda_{101}^{(i)} = p_{000}^{(i)} + p_{100}^{(i)} + p_{001}^{(i)} + p_{101}^{(i)} \tag{9.51}$$

$$\lambda_{011}^{(i)} = p_{000}^{(i)} + p_{010}^{(i)} + p_{001}^{(i)} + p_{011}^{(i)} \tag{9.52}$$

$$\lambda_{111}^{(i)} = p_{000}^{(i)} + p_{100}^{(i)} + p_{010}^{(i)} + p_{001}^{(i)} + p_{110}^{(i)} + p_{101}^{(i)} + p_{011}^{(i)} + p_{111}^{(i)} \tag{9.53}$$

If $\lambda_{\delta_1\delta_2\delta_3}$ and $p_{\delta_1\delta_2\delta_3}$ ($\delta_i = 1$ or 0, $i = 1, 2, 3$) are used to express the combined probability measure of two independent sources of information in spaces $S = \{s_1, s_2, s_3\}$ and (X, Ω, μ), respectively, then based on equation (9.45) and through equations (9.46)-(9.53), the following can be obtained

$$p_{000} = p^{(1)}_{000}p^{(2)}_{000} \tag{9.54}$$

$$p_{100} = p^{(1)}_{000}p^{(2)}_{100} + p^{(1)}_{100}p^{(2)}_{000} + p^{(1)}_{100}p^{(2)}_{100} \tag{9.55}$$

$$p_{010} = p^{(1)}_{000}p^{(2)}_{010} + p^{(1)}_{010}p^{(2)}_{000} + p^{(1)}_{010}p^{(2)}_{010} \tag{9.56}$$

$$p_{001} = p^{(1)}_{000}p^{(2)}_{001} + p^{(1)}_{001}p^{(2)}_{000} + p^{(1)}_{001}p^{(2)}_{001} \tag{9.57}$$

$$\begin{aligned} p_{110} = {} & p^{(1)}_{000}p^{(2)}_{110} + p^{(1)}_{100}p^{(2)}_{010} + p^{(1)}_{100}p^{(2)}_{110} + p^{(1)}_{010}p^{(2)}_{100} \\ & + p^{(1)}_{010}p^{(2)}_{110} + p^{(1)}_{110}p^{(2)}_{000} + p^{(1)}_{110}p^{(2)}_{100} + p^{(1)}_{110}p^{(2)}_{010} + p^{(1)}_{110}p^{(2)}_{110} \end{aligned} \tag{9.58}$$

$$\begin{aligned} p_{101} = {} & p^{(1)}_{000}p^{(2)}_{101} + p^{(1)}_{100}p^{(2)}_{001} + p^{(1)}_{100}p^{(2)}_{101} + p^{(1)}_{001}p^{(2)}_{100} \\ & + p^{(1)}_{001}p^{(2)}_{101} + p^{(1)}_{101}p^{(2)}_{000} + p^{(1)}_{101}p^{(2)}_{100} + p^{(1)}_{101}p^{(2)}_{001} + p^{(1)}_{101}p^{(2)}_{101} \end{aligned} \tag{9.59}$$

$$\begin{aligned} p_{011} = {} & p^{(1)}_{000}p^{(2)}_{011} + p^{(1)}_{010}p^{(2)}_{001} + p^{(1)}_{010}p^{(2)}_{011} + p^{(1)}_{001}p^{(2)}_{010} \\ & + p^{(1)}_{001}p^{(2)}_{011} + p^{(1)}_{011}p^{(2)}_{000} + p^{(1)}_{011}p^{(2)}_{010} + p^{(1)}_{011}p^{(2)}_{001} + p^{(1)}_{011}p^{(2)}_{011} \end{aligned} \tag{9.60}$$

$$\begin{aligned} p_{111} = {} & p^{(1)}_{000}p^{(2)}_{111} + p^{(1)}_{100}p^{(2)}_{011} + p^{(1)}_{100}p^{(2)}_{111} + p^{(1)}_{010}p^{(2)}_{101} \\ & + p^{(1)}_{010}p^{(2)}_{111} + p^{(1)}_{001}p^{(2)}_{110} + p^{(1)}_{001}p^{(2)}_{111} + p^{(1)}_{011}p^{(2)}_{100} + p^{(1)}_{011}p^{(2)}_{101} \\ & + p^{(1)}_{011}p^{(2)}_{110} + p^{(1)}_{011}p^{(2)}_{111} + p^{(1)}_{101}p^{(2)}_{010} + p^{(1)}_{101}p^{(2)}_{011} + p^{(1)}_{101}p^{(2)}_{110} \\ & + p^{(1)}_{101}p^{(2)}_{111} + p^{(1)}_{110}p^{(2)}_{001} + p^{(1)}_{110}p^{(2)}_{011} + p^{(1)}_{110}p^{(2)}_{101} + p^{(1)}_{110}p^{(2)}_{111} \\ & + p^{(1)}_{111}p^{(2)}_{000} + p^{(1)}_{111}p^{(2)}_{100} + p^{(1)}_{111}p^{(2)}_{010} + p^{(1)}_{111}p^{(2)}_{001} + p^{(1)}_{111}p^{(2)}_{011} \\ & + p^{(1)}_{111}p^{(2)}_{101} + p^{(1)}_{111}p^{(2)}_{110} + p^{(1)}_{111}p^{(2)}_{111} \end{aligned} \tag{9.61}$$

For the case of $S = \{s_1, s_2, \ldots, s_n\}$, the general combination rule is

$$p_{\delta_1\delta_2\ldots\delta_n} = \sum_{\substack{\delta_i=\delta'_i\cup\delta''_i \\ i=1,2,\ldots,m}} p^{(1)}_{\delta'_1\delta'_2\ldots\delta'_n} p^{(2)}_{\delta''_1\delta''_2\ldots\delta''_n} \tag{9.62}$$

for all $(\delta'_1, \delta'_2, \ldots, \delta'_m, \delta''_1, \delta''_2, \ldots, \delta''_n)$.

9.4.2 Combination rule of probability measures in space S

Define a probability measure over $2^S = \{S_{\delta_1\delta_2\ldots\delta_n}\}$ as $M : 2^S = \{S_{\delta_1\delta_2\ldots\delta_m}\} \to [0, 1]$ with

$$M(S_{\delta_1\delta_2\ldots\delta_n}) = \begin{cases} 0, & S_{\delta_1\delta_2\ldots\delta_n} = S_{00\ldots0} \\ \frac{\mu(X_{\delta_1\delta_2\ldots\delta_n})}{1-\mu(X_{00\ldots0})}, & S_{\delta_1\delta_2\ldots\delta_n} \neq S_{00\ldots\delta_0} \end{cases} \tag{9.63}$$

where M is the probability measure over a class $2^S = \{S_{\delta_1\delta_2\ldots\delta_n}\}$ of subsets of space S and Γ maps X into S.

The combination rule:

Given two independent sources of information (X_i, Ω_i, μ_i), $i = 1, 2$, and the corresponding mapping space, $S = \{s_1, s_2, \ldots, s_n\} = \{S_{\delta_1\delta_2\ldots\delta_n}\}$, where Γ_i maps X_i into S. Based on equation (9.62), we have

$$\mu(X_{\delta_1\delta_2\ldots\delta_n}) = \sum_{\substack{\delta_i=\delta_i'\cup\delta_i''\\ i=1,2,\ldots,n}} \mu^{(1)}(X^{(1)}_{\delta_1'\delta_2'\ldots\delta_n'})\mu^{(2)}(X^{(2)}_{\delta_1''\delta_2''\ldots\delta_n''}) \tag{9.64}$$

From equation (9.63), for any $S_{\delta_1\delta_2\ldots\delta_m} \neq S_{00\ldots0}$, there exists

$$\mu^{(1)}(X^{(1)}_{\delta_1'\delta_2'\ldots\delta_n'}) = M^{(1)}(S_{\delta_1\delta_2\ldots\delta_n})(1 - \mu^{(1)}(X^{(1)}_{00\ldots0})) \tag{9.65}$$

$$\mu^{(2)}(X^{(2)}_{\delta_1''\delta_2''\ldots\delta_n''}) = M^{(2)}(S_{\delta_1\delta_2\ldots\delta_n})(1 - \mu^{(2)}(X^{(2)}_{00\ldots0})) \tag{9.66}$$

and

$$\mu(X_{\delta_1\delta_2\ldots\delta_n}) = M(S_{\delta_1\delta_2\ldots\delta_n})(1 - \mu(X_{00\ldots0})) \tag{9.67}$$

such that equation (9.64) becomes

$$M(S_{\delta_1\delta_2\ldots\delta_n}) = \frac{\sum\limits_{\substack{\delta_i=\delta_i'\cup\delta_i''\\ i=1,2,\ldots,n}} M^{(1)}(S_{\delta_1\delta_2\ldots\delta_n})M^{(2)}(S_{\delta_1\delta_2\ldots\delta_n})[1 - \mu^{(1)}(X^{(1)}_{00\ldots0})[1 - \mu^{(2)}(X^{(2)}_{00\ldots0})]}{1 - \mu(X_{00\ldots0})} \tag{9.68}$$

and

$$\frac{[1 - \mu^{(1)}(X^{(1)}_{00\ldots0})][1 - \mu^{(2)}(X^{(2)}_{00\ldots0})]}{1 - \mu(X_{00\ldots0})} = \frac{1}{\sum\limits_{\substack{\delta_i'\cup\delta_i''\neq 0\\ i=1,2,\ldots,n}} M^{(1)}(S_{\delta_1'\delta_2'\ldots\delta_n'})M^{(2)}(S_{\delta_1''\delta_2''\ldots\delta_n''})} \tag{9.69}$$

Substitute (9.69) into (9.68),

$$\begin{aligned} M(S_{\delta_1\delta_2\ldots\delta_n}) &= \frac{\sum\limits_{\substack{\delta_i=\delta_i'\cup\delta_i''\\ i=1,2,\ldots,n}} M^{(1)}(S_{\delta_1'\delta_2'\ldots\delta_n'})M^{(2)}(S_{\delta_1''\delta_2''\ldots\delta_n''})}{1 - \sum\limits_{\substack{\delta_i'\cup\delta_i''\neq 0\\ i=1,2,\ldots,n}} M^{(1)}(S_{\delta_1'\delta_2'\ldots\delta_n'})M^{(2)}(S_{\delta_1''\delta_2''\ldots\delta_n''})} \\ &= \sum_{\substack{\delta_i=\delta_i'\cup\delta_i''\\ i=1,2,\ldots,n}} M^{(1)}(S_{\delta_1'\delta_2'\ldots\delta_n'})M^{(2)}(S_{\delta_1''\delta_2''\ldots\delta_n''}) \end{aligned} \tag{9.70}$$

If $S_{\delta_1\delta_2\ldots\delta_n} = S_{00\ldots0}$, we define

$$M(S_{\delta_1\delta_2\ldots\delta_n}) \triangleq 0 \tag{9.71}$$

Hence, equations (9.70)- (9.71) express the combination of two sources of information, (X_i, Ω_i, μ_i), $i = 1, 2$, for the mapping space, $S = \{s_1, s_2, \ldots, s_n\} = S_{\delta_1\delta_2\ldots\delta_n}$, where Γ_i maps X_i into S.

9.5 The disjunctive combination rule

Based on the results in section 9.4, the disjunctive combination rule for two independent sources of information is obtained as follows:

Theorem 4:

Suppose $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ is a frame of discernment with n elements. The basic probability assignments of the two sources of information, (X_1, Ω_1, μ_2) and (X_2, Ω_2, μ_2) over the same frame of discernment are m_1 and m_2, and focal elements A_1, A_2, …, A_k and B_1, B_2, …, B_l, respectively. Then the combined basic probability assignment of the two sources of information can be defined as

$$m(C) = \begin{cases} 0, \qquad C = \emptyset \\ \displaystyle\sum_{C = A_i \cup B_j} m_1(A_i) m_2(B_j), \qquad C \neq \emptyset \end{cases} \tag{9.72}$$

Proof: Since $m(\emptyset) = 0$ by definition, m is a basic probability assignment provided only that the $m(C)$ sum to one. In fact,

$$\begin{aligned} \sum_{C \subseteq \Theta} m(C) &= m(\emptyset) + \sum_{\substack{C \subset \Theta \\ C \neq \emptyset}} m(C) \\ &= \sum_{\substack{C \subset \Theta \\ C \neq \emptyset}} \sum_{\substack{C = A_i \cup B_j \\ i \in \{1,2,\ldots,k\}, j \in \{1,2,\ldots,l\}}} m_1(A_i) m_2(B_j) \\ &= \sum_{\substack{A_i \cup B_j \neq \emptyset \\ i \in \{1,2,\ldots,k\}, j \in \{1,2,\ldots,l\}}} m_1(A_i) m_2(B_j) \\ &= \sum_{\substack{A_i \subset \Theta \\ A_i \neq \emptyset}} m_1(A_i) \sum_{\substack{B_j \subset \Theta \\ B_j \neq \emptyset}} m_2(B_j) \end{aligned}$$

Hence, m is a basic probability assignment over the frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$. Its focal elements are

$$C = (\bigcup_{i=1,2,\ldots,k} A_i) \bigcup (\bigcup_{j=1,2,\ldots,l} B_l)$$

Based on theorem 4, theorem 5 can be stated as follows. A similar result can be found in [29, 31].

Theorem 5:

If Bel_1 and Bel_2 are belief functions over the same frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ with basic probability assignments m_1 and m_2, and focal elements A_1, A_2, …, A_k and B_1, B_2, …, B_l, respectively, then the function $m : 2^\Theta \to [0, 1]$ defined as

$$m(C) = \begin{cases} 0, & C = \emptyset \\ \displaystyle\sum_{C = A_i \cup B_j} m_1(A_i) m_2(B_j), & C \neq \emptyset \end{cases} \tag{9.73}$$

yields a basic probability assignment. The core of the belief function given by m is equal to the union of the cores of Bel_1 and Bel_2.

Physical interpretations of the combination rule for two independent sources of information are:

1. The combination rule in theorem 4 indicates a type of logical inference in human judgments, namely: the OR rule. That is, for a given frame of discernment, the elements that are simultaneously judged by each source of information will also be judgment elements of the combined source of information; otherwise, it will result in uncertainty so the combination judgments of the elements will be ignorance.

2. The essential difference between the new combination rule and DS combination rule is that the latter is a type of logical inference with AND or conjunction, while the former is based on OR or disjunction. The new combination rule (or the OR rule) overcomes the shortcomings of DS combination rule with AND, such as in the counter-intuitive situation and in the combination of sharp sources of information.

3. The judgment with OR has the advantage over that with AND in treating elements that are not simultaneously judged by each independent source of information. The OR rule gives more plausible judgments for these elements than the AND rule. The judgment better fits to the logical judgment of human beings.

Example 1

Given the frame of discernment $\Theta = \{\theta_1, \theta_2\}$, the judgments of the basic probability from two sources of information are m_1 and m_2 as follows:

$$m_1(\theta_1) = 0.2, \quad m_1(\theta_2) = 0.4, \quad m_1(\theta_1, \theta_2) = 0.4$$

$$m_2(\theta_1) = 0.4, \quad m_2(\theta_2) = 0.4, \quad m_2(\theta_1, \theta_2) = 0.2$$

Then through theorem 4, the combination judgment is

$$m(\theta_1) = 0.08, \quad m(\theta_2) = 0.16, \quad m(\theta_1, \theta_2) = 0.76$$

Comparing the combined basic probabilities of θ_1 and θ_2, the judgment of θ_2 occurs more often than θ_1, but the whole combination doesn't decrease the uncertainty of the judgments, which is evident from the above results.

Example 2 (the counter-intuitive situation)

Zadeh's example:

The frame of discernment about the patient is $\Theta = \{M, C, T\}$where M denotes meningitis, C represents contusion and T indicates tumor. The judgments of two doctors about the patient are

$$m_1(M) = 0.99, \quad m_1(T) = 0.01$$

$$m_2(C) = 0.99, \quad m_2(T) = 0.01$$

Combining these judgments through theorem 4, results in

$$m(M \cup C) = 0.9801, \quad m(M \cup T) = 0.0099, \quad m(C \cup T) = 0.0099, \quad m(T) = 0.0001$$

From $m(M \cup T) = 0.0099$ and $m(C \cup T) = 0.0099$, it is clear that there are less uncertainties between T and M, as well as T and C; which implies that T can easily be distinguished from M and C. Also, T occurs with the basic probability $m(T) = 0.0001$, i.e. T probably will not occur in the patient. The patient may be infected with M or C. Furthermore, because of $m(M \cup C) = 0.9801$, there is a bigger uncertainty with 0.9801 between M and C, so the two doctors cannot guarantee that the patient has meningitis (M) or contusion (C) except that the patient has no tumor (T). The patient needs to be examined by more doctors to assure the diagnoses.

We see the disjunctive combination rule can be used to this case very well. It fits to the human intuitive judgment.

9.6 Properties of conjunctive and disjunctive combination rules

In the section, the conjunctive and disjunctive combination rules, namely, Dempster-Shafer's combination rule, Yager's combination rule, Dubois and Prade's (DP) combination rule, DSm's combination rule and the disjunctive combination rule, are studied. The properties of each combination rule of evidence are discussed in detail, such as the role of evidence of each source of information in the combination judgment, the comparison of the combination judgment belief and ignorance of each combination rule, the treatment of conflict judgments given by the two sources of information, and the applications of combination rules. Zadeh's example is included in this section to evaluate the performance as well as efficiency of each combination rule of evidence for the conflict judgments given by the two sources of information.

9.6.1 The combination rules of evidence

9.6.1.1 Yager's combination rule of evidence

Suppose Bel_1 and Bel_2 are belief functions over the same frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ with basic probability assignments m_1 and m_2 , and focal elements A_1, A_2, ..., A_k and B_1, B_2, ..., B_l, respectively. Then Yager's combined basic probability assignment of the two sources of information can be defined as [20]

$$m_Y(C) = \begin{cases} \displaystyle\sum_{\substack{i,j \\ C=A_i\cap B_j}} m_1(A_i)m_2(B_j), & C \neq \Theta, \emptyset \\ \displaystyle m_1(\Theta)m_2(\Theta) + \sum_{\substack{i,j \\ A_i\cap B_j=\emptyset}} m_1(A_i)m_2(B_j), & C = \Theta \\ 0, \quad C = \emptyset \end{cases} \tag{9.74}$$

9.6.1.2 Dubois & Prade (DP)'s combination rule of evidence

Given the same conditions as in Yager's combination rule, Dubois and Prade's combined basic probability assignment of the two sources of information can be defined as [26]

$$m_{DP}(C) = \begin{cases} \displaystyle\sum_{\substack{i,j \\ C=A_i\cap B_j}} m_1(A_i)m_2(B_j) + \sum_{\substack{i,j \\ C=A_i\cup B_j \\ A_i\cap B_j=\emptyset}} m_1(A_i)m_2(B_j), & C \neq \emptyset \\ 0, \quad C = \emptyset \end{cases} \tag{9.75}$$

9.6.1.3 DSm combination rules of evidence

These rules are presented in details in chapters 1 and 4 and are just recalled briefly here for convenience for the two independent sources of information.

- *The classical DSm combination rule for free DSm model* [27]

$$\forall C \in D^\Theta, \qquad m(C) = \sum_{\substack{A,B\in D^\Theta \\ A\cap B=C}} m_1(A)m_2(B) \tag{9.76}$$

  where D^Θ denotes the hyper-power set of the frame Θ (see chapters 2 and 3 for details).

- *The general DSm combination rule for hybrid DSm model* $\mathcal{M}$

  We consider here only the two sources combination rule.

$$\forall A \in D^\Theta, \qquad m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A)\Big[S_1(A) + S_2(A) + S_3(A)\Big] \tag{9.77}$$

where $\phi(A)$ is the characteristic non emptiness function of a set A, i.e. $\phi(A) = 1$ if $A \notin \boldsymbol{\emptyset}$ and $\phi(A) = 0$ otherwise, where $\boldsymbol{\emptyset} \triangleq \{\boldsymbol{\emptyset}_{\mathcal{M}}, \emptyset\}$. $\boldsymbol{\emptyset}_{\mathcal{M}}$ is the set of all elements of D^Θ which have been forced to be empty through the constraints of the model $\mathcal{M}$ and $\emptyset$ is the classical/universal empty set. $S_1(A) \equiv m_{\mathcal{M}^f(\Theta)}(A)$, $S_2(A)$, $S_3(A)$ are defined by (see chapter 4)

$$S_1(A) \triangleq \sum_{\substack{X_1,X_2\in D^\Theta \\ X_1\cap X_2=A}} \prod_{i=1}^{2} m_i(X_i) \tag{9.78}$$

$$S_2(A) \triangleq \sum_{\substack{X_1,X_2\in \boldsymbol{\emptyset} \\ [\mathcal{U}=A]\vee[(\mathcal{U}\in\boldsymbol{\emptyset})\wedge(A=I_t)]}} \prod_{i=1}^{2} m_i(X_i) \tag{9.79}$$

$$S_3(A) \triangleq \sum_{\substack{X_1,X_2\in D^\Theta \\ X_1\cup X_2=A \\ X_1\cap X_2\in\boldsymbol{\emptyset}}} \prod_{i=1}^{2} m_i(X_i) \tag{9.80}$$

with $\mathcal{U} \triangleq u(X_1)\cup u(X_2)$ where $u(X)$ is the union of all singletons θ_i that compose X and $I_t \triangleq \theta_1\cup\theta_2$ is the total ignorance. $S_1(A)$ corresponds to the classic DSm rule of combination based on the free DSm model; $S_2(A)$ represents the mass of all relatively and absolutely empty sets which is transferred to the total or relative ignorances; $S_3(A)$ transfers the sum of relatively empty sets to the non-empty sets.

9.6.1.4 The disjunctive combination rule of evidence

This rule has been presented and justified previously in this chapter and can be found also in [22, 23, 29, 30, 31].

Suppose $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ is a frame of discernment with n elements (it is the same as in theorem 3). The basic probability assignments of the two sources of information over the same frame of discernment are m_1 and m_2, and focal elements A_1, A_2, …, A_k and B_1, B_2, …, B_l, respectively. Then the combined basic probability assignment of the two sources of information can be defined as

$$m_{Dis}(C) = \begin{cases} \sum\limits_{\substack{i,j \\ C=A_i\cup B_j}} m_1(A_i)m_2(B_j), & C \neq \emptyset \\ 0, \quad C = \emptyset \end{cases} \tag{9.81}$$

for any $C \subset \Theta$. The core of the belief function given by m is equal to the union of the cores of Bel_1 and Bel_2.

9.6.2 Properties of combination rules of evidence

Given two independent sources of information defined over the frame of discernment $\Theta = \{\theta_1, \theta_2\}$, their basic probability assignments or basic belief masses over Θ are

$$S1: \quad m_1(\theta_1) = 0.4, \quad m_1(\theta_2) = 0.3, \quad m_1(\theta_1 \cup \theta_2) = 0.3$$

$$S2: \quad m_2(\theta_1) = 0.5, \quad m_2(\theta_2) = 0.3, \quad m_2(\theta_1 \cup \theta_2) = 0.2$$

Then the results of each combination rule of evidence for the two independent sources of information are as follows. For the frame of discernment with n elements, similar results can be obtained.

$S2\ (m_2)$ \ $S1\ (m_1)$	$\{\theta_1\}$ (0.4)	$\{\theta_2\}$ (0.3)	$\{\theta_1, \theta_2\}$ (0.3)
$\{\theta_1\}$ (0.5)	$\{\theta_1\}$ (0.2)	$\{\theta_1\} \cap \{\theta_2\} \Rightarrow k$ (0.15)	$\{\theta_1\}$ (0.15)
$\{\theta_2\}$ (0.3)	$\{\theta_1\} \cap \{\theta_2\} \Rightarrow k$ (0.12)	$\{\theta_2\}$ (0.09)	$\{\theta_2\}$ (0.09)
$\{\theta_1, \theta_2\}$ (0.2)	$\{\theta_1\}$ (0.08)	$\{\theta_2\}$ (0.06)	$\{\theta_1, \theta_2\}$ (0.06)

Table 9.1: The conjunctive combination of evidence (DS)

$S2\ (m_2)$ \ $S1\ (m_1)$	$\{\theta_1\}$ (0.4)	$\{\theta_2\}$ (0.3)	$\{\theta_1, \theta_2\}$ (0.3)
$\{\theta_1\}$ (0.5)	$\{\theta_1\}$ (0.2)	$\{\theta_1\} \cap \{\theta_2\} \Rightarrow \Theta$ (0.15)	$\{\theta_1\}$ (0.15)
$\{\theta_2\}$ (0.3)	$\{\theta_1\} \cap \{\theta_2\} \Rightarrow \Theta$ (0.12)	$\{\theta_2\}$ (0.09)	$\{\theta_2\}$ (0.09)
$\{\theta_1, \theta_2\}$ (0.2)	$\{\theta_1\}$ (0.08)	$\{\theta_2\}$ (0.06)	$\{\theta_1, \theta_2\}$ (0.06)

Table 9.2: The conjunctive and disjunctive combination of evidence (Yager)

$S2\ (m_2)$ \ $S1\ (m_1)$	$\{\theta_1\}$ (0.4)	$\{\theta_2\}$ (0.3)	$\{\theta_1, \theta_2\}$ (0.3)
$\{\theta_1\}$ (0.5)	$\{\theta_1\}$ (0.2)	$\{\theta_1\} \cap \{\theta_2\} \Rightarrow \{\theta_1\} \cup \{\theta_2\}$ (0.15)	$\{\theta_1, \theta_2\}$ (0.15)
$\{\theta_2\}$ (0.3)	$\{\theta_1\} \cap \{\theta_2\} \Rightarrow \{\theta_1\} \cup \{\theta_2\}$ (0.12)	$\{\theta_2\}$ (0.09)	$\{\theta_2\}$ (0.09)
$\{\theta_1, \theta_2\}$ (0.2)	$\{\theta_1\}$ (0.08)	$\{\theta_2\}$ (0.06)	$\{\theta_1, \theta_2\}$ (0.06)

Table 9.3: The conjunctive and disjunctive combination of evidence (Dubois-Prade)

Property 1: the role of evidence of each source of information in the combination judgment:

1. With DS combination rule of evidence [2], the combined judgment for element θ_i $(i = 1, 2)$ consists of two parts. One is from the simultaneous support judgment of two sources of information for the element θ_i $(i = 1, 2)$ and the other is that one of two sources of information yields a support judgment, while the second source is ignorant for the element θ_i $(i = 1, 2)$ (i.e. ignorance). The

$A \in D^{\Theta}$	m_1	m_2	$\Phi(A)$	$S_1(A)$	$S_2(A)$	$S_3(A)$	$m_{\mathcal{M}(\Theta)}(A)$
$\emptyset$	0	0	0	0	0	0	0
$\{\theta_1\}$	0.4	0.5	1	0.43	0	0	0.43
$\{\theta_2\}$	0.3	0.3	1	0.24	0	0	0.24
$\{\theta_1 \cap \theta_2\} \overset{\mathcal{M}(\Theta)}{=} \emptyset$	0	0	0	0.27	0	0	0
$\{\theta_1 \cup \theta_2\}$	0.3	0.2	1	0.06	0	0.27	0.33

Table 9.4: The hybrid DSm combination of evidence

$S2\ (m_2) \quad \backslash \quad S1\ (m_1)$	$\{\theta_1\}$ (0.4)	$\{\theta_2\}$ (0.3)	$\{\theta_1, \theta_2\}$ (0.3)
$\{\theta_1\}$ (0.5)	$\{\theta_1\}$ (0.2)	$\{\theta_1\} \cup \{\theta_2\}$ (0.15)	$\{\theta_1, \theta_2\}$ (0.15)
$\{\theta_2\}$ (0.3)	$\{\theta_1\} \cup \{\theta_2\}$ (0.12)	$\{\theta_2\}$ (0.09)	$\{\theta_1, \theta_2\}$ (0.09)
$\{\theta_1, \theta_2\}$ (0.2)	$\{\theta_1, \theta_2\}$ (0.08)	$\{\theta_1, \theta_2\}$ (0.06)	$\{\theta_1, \theta_2\}$ (0.06)

Table 9.5: The disjunctive combination of evidence

combined total ignorance is from the total ignorance of both sources of information. The failure combination judgment for some element is from the conflict judgments given by two sources of information for the element.

2. The difference between Yager's combination rule of evidence [20] and DS combination rule of evidence [2] is that the conflict judgments of combination given by two sources of information for some element is considered to be a part of combined ignorance i.e. it is added into the total ignorance.

3. Dubois and Prade's combination rule of evidence [26] is different from that of Yager's combination rule [20] in that when two sources of information give the conflict judgments for an element in the frame of discernment, one of two judgments is at least thought as a reasonable judgment. The conflict judgments of combination for the two conflict elements are distributed to the judgment corresponding to union of the two conflict elements.

4. The classical DSm combination rule of evidence [27] is different from those of Dubois and Prade's [26], Yager's [20] and DS [2]. The conflict judgments given by two sources of information for an element in the frame of discernment are considered as paradox. These paradoxes finally support the combination judgment of each element θ_i $(i = 1, 2)$. For the hybrid DSm combination rule, see chapter 4, it consists of three parts. The first one is from the classic DSm rule of combination based on the free-DSm model;the second one is the mass of all relatively and absolutely empty sets which are transferred to the total or relative ignorance, while the third one is the mass that transfers the all relatively empty sets to union of the elements that are included in the sets.

5. With the disjunctive combination rule of evidence [22, 23, 29, 30, 31], the combination judgment for each element is only from the simultaneous support judgment of each source of information for the element θ_i ($i = 1, 2$). The combined ignorance consists of the combination of conflict judgments given by two sources of information, the combination of the ignorance given by one source of information and the support judgment for any element given by another source, and the combination of the ignorance from both sources of information simultaneously. There is no failure combination judgment. However, the combined belief is decreased and the ignorance is increased.

6. The combination rules of evidence of DS and the classical DSm are the conjunctive rule, the disjunctive combination rule of evidence is the disjunctive rule, while the combination rule of evidence of Yager, Dubois & Prade, and the hybrid DSm are hybrid of the conjunctive and disjunctive rules.

Property 2: the comparison of combination judgment belief (Bel(.)) and ignorance (Ign(.) = Pl(.) − Bel(.)) of each combination rule is:

$$\mathrm{Bel}_{DS}(\theta_i) > \mathrm{Bel}_{DSm}(\theta_i) > \mathrm{Bel}_{DP}(\theta_i) > \mathrm{Bel}_{Y}(\theta_i) > \mathrm{Bel}_{Dis}(\theta_i), \quad i = 1, 2 \tag{9.82}$$

$$\mathrm{Ign}_{DS}(\theta_i) < \mathrm{Ign}_{DSm}(\theta_i) > \mathrm{Ign}_{DP}(\theta_i) < \mathrm{Ign}_{Y}(\theta_i) < \mathrm{Ign}_{Dis}(\theta_i), \quad i = 1, 2 \tag{9.83}$$

In fact, for the above two sources of information, the results from each combination rule are as the following:

Combination rule	$m(\theta_1)$	$m(\theta_2)$	$m(\Theta)$	$\mathrm{Bel}(\theta_1)$	$\mathrm{Bel}(\theta_2)$	$\mathrm{Bel}(\Theta)$	$\mathrm{Ign}(\theta_1)$	$\mathrm{Ign}(\theta_2)$
DS	0.589	0.329	0.082	0.589	0.329	1	0.082	0.082
Yager	0.43	0.24	0.33	0.43	0.24	1	0.33	0.33
DP	0.43	0.24	0.33	0.43	0.24	1	0.33	0.33
Hybrid DSm	0.43	0.24	0.33	0.43	0.24	1	0.33	0.33
Disjunctive	0.20	0.09	0.71	0.20	0.09	1	0.71	0.71

From the results in the above table, it can be observed that the hybrid DSm's, Yager's and Dubois & Prade's combination judgments are identical for the two independent sources of information. However, for more than two independent sources of information, the results of combination judgments are as in equations (9.82) and (9.83) (i.e. the results are different, the hybrid DSm model is more general than Dubois-Prade's and Yager's, while Dubois-Prade's model has less total ignorance than Yager's).

Property 3: The conflict judgments given by two sources of information for the frame of discernment:

Under DS combination rule, the combined conflict judgments are thought as failures and are deducted from the total basic probability assignment of combination, while under Yager's combination rule, they are thought as the total ignorance; under Dubois & Prade's combination rule; they are distributed to the union of the two conflict elements. That means one of conflict judgments is at least reasonable. Under the classical DSm combination rule, they constitute paradoxes to support the combined judgment belief of each element, and are also thought as a new event that takes part in the subsequent judgment when new evidences occur. While for the hybrid DSm combination rule, the treatment of conflict evidence is similar to Dubois & Prade's approach. For the disjunctive combination rule, the conflict judgments of combination constitute ignorance, and take part in the subsequent judgment when the new evidences occur.

Property 4: using them in applications:

Based on properties 1-3, when the two independent sources of information are not very conflict, the disjunctive combination rule is more conservative combination rule. The combined results are uncertain when conflict judgments of two sources of information occur and hence the final judgment is delayed until more evidence comes into the judgment systems. Also, the combined judgment belief for each element in the frame of discernment is decreased, and ignorance is increased as the new evidences come. Hence, the disjunctive combination rule is not more efficient when we want the ignorance be decreased in the combination of evidence. It is fair to assume that for the case when the two (conflict) judgments are not exactly known which one is more reasonable, however, at least one of them should provide a reasonable judgment. But DS combination rule is contrary to the disjunctive combination rule. It can make the final judgment faster than other rules (see equations (9.82)-(9.83)), but the disjunctive combination rule will make less erroneous judgments than other rules. The cases for the combination rules of the hybrid DSm, Dubois & Prade, and Yager's combination rule fall between the above two. For the other properties, for instance, the two conflict independent sources of information, see the next section and the example that follows.

9.6.3 Example

In this section, we examine the efficiency of each combination rule for conflict judgments via Zadeh's famous example. Let the frame of discernment of a patient be $\Theta = \{M, C, T\}$ where M denotes meningitis, C represents contusion and T indicates tumor. The judgments of two doctors about the patient are

$$m_1(M) = 0.99, m_1(T) = 0.01 \qquad \text{and} \qquad m_2(C) = 0.99, m_2(T) = 0.01$$

The results from each combination rule of evidence are:

Rules	$m(T)$	$m(M \cup C)$	$m(C \cup T)$	$m(M \cup T)$	$m(\Theta)$
DS	1	0	0	0	0
Yager	0.0001	0	0	0	0.9999
DP	0.0001	0.9801	0.0099	0.0099	0
Hybrid DSm	0.0001	0.9801	0.0099	0.0099	0
Disjunctive	0.0001	0.9801	0.0099	0.0099	0

The basic belief masses $m(M \cap C)$, $m(C \cap T)$ and $m(M \cap T)$ equal zero with all five rules of combination and the belief of propositions $M \cap C$, $C \cap T$, $M \cap T$, $M \cup C$, $C \cup T$, $M \cup T$, M, C, T and $M \cup C \cup T$ are given in the next tables:

Rules	Bel$(M \cap C)$	Bel$(C \cap T)$	Bel$(M \cap T)$	Bel$(M \cup C)$	Bel$(C \cup T)$	Bel$(M \cup T)$
DS	0	0	0	0	0	0
Yager	0	0	0	0	0	0
DP	0	0	0	0.9801	0.01	0.01
Hybrid DSm	0	0	0	0.9801	0.01	0.01
Disjunctive	0	0	0	0.9801	0.01	0.01

Rules	Bel(M)	Bel(C)	Bel(T)	Bel$(M \cup C \cup T)$
DS	0	0	1	1
Yager	0	0	0.0001	1
DP	0	0	0.0001	1
Hybrid DSm	0	0	0.0001	1
Disjunctive	0	0	0.0001	1

Comparison and analysis of the fusion results:

1. DS combination judgment belief of each element is:

$$\text{Bel}_{DS}(T) = 1, \qquad \text{Bel}_{DS}(M) = \text{Bel}_{DS}(C) = 0$$

It means that the patient must have disease T with a degree of belief of 1 and must not have diseases M and C, because their degrees of belief are 0, respectively. It is a counter-intuitive situation with $\mathrm{Bel}_{DS,1}(M) = \mathrm{Bel}_{DS,2}(C) = 0.99$, $\mathrm{Bel}_{DS,1}(T) = \mathrm{Bel}_{DS,2}(T) = 0.01$. Moreover, in spite of the basic probability assignment values over diseases T, M and C, the judgment of the two doctors for DS combination rule will always be T with the degree of belief of 1, and each M and C with degree of belief of 0. It shows DS combination rule is not effective in this case. The main reason for this situation has been presented in sections 9.3-9.5.

2. Yager's combination judgment belief of each element is:

$$\mathrm{Bel}_Y(T) = 0.0001, \qquad \mathrm{Bel}_Y(M) = \mathrm{Bel}_Y(C) = 0$$

This degree of belief is too small to make the final judgment. Therefore, Yager's combination rule of evidence will wait for the new evidence to come in order to obtain more accurate judgment. The reason for this result is that the rule transforms all conflict judgments into the total ignorance.

3. For Dubois & Prade's combination rule, there is

$$\mathrm{Bel}_{DP}(T) = 0.0001, \qquad \mathrm{Bel}_{DP}(M \cup C) = 0.9801, \qquad \mathrm{Bel}_{DP}(M \cup T) = \mathrm{Bel}_{DP}(C \cup T) = 0.01$$

This result is the same as that of the disjunctive combination rule and the hybrid DSm combination rule. With a belief of T, $\mathrm{Bel}_{DP}(T) = 0.0001$, we can judge that the patient having disease T is less probable event. Furthermore, $\mathrm{Bel}_{DP}(M \cup T) = \mathrm{Bel}_{DP}(C \cup T) = 0.01$, hence the patient may have disease M or C. Also, $\mathrm{Bel}_{DP}(M \cup C) = 0.9801$, this further substantiates the fact that the patient has either M or C, or both. For the final judgment, one needs the new evidence or diagnosis by the third doctor.

Based on the judgments of two doctors, the different judgment results of each combination rules are clearly demonstrated. For this case, the results from Dubois & Prade's rulr, the hybrid DSm rule and from the disjunctive combination rule are more suitable to human intuitive judgment; the result from Yager's combination rule, can't make the final judgment immediately because of less degree of judgment belief and more ignorance, while the results of DS combination rule is counter-intuitive. These results demonstrate the efficiency of each combination rule for the conflict judgments given by two sources of information for the element in the frame of discernment.

9.7 Conclusion

In this chapter, DS combination rule is examined based on multi-valued mappings of independent information sources and the product combination rule of multiple independent information sources. It is obtained that Dempster's rule is different from DS combination rule and shortcomings in DS combination rule are due to the result of the product combination rule. The drawback in the explanation of multi-valued mappings when applied to Dempster's rule were pointed out and proven. Furthermore, based on these results, a novel justification of the disjunctive combination rule for two independent sources of information based on the redefined combination-mapping rule of multiple multi-valued mappings in the product space of multiple sources of information mappings has been proposed. The combination rule depends on the logical judgment of OR. It overcomes the shortcomings of Dempster-Shafer's combination rule, especially, in resolving the counter-intuitive situation. Finally, the conjunctive and disjunctive combination rules of evidence, namely, Dempster-Shafer's (DS) combination rule, Yager's combination rule, Dubois & Prade's (DP) combination rule, DSm's combination rule and the disjunctive combination rule, are studied for the two independent sources of information. The properties of each combination rule of evidence are discussed in detail, such as the role of evidence of each source of information in the combination judgment, the comparison of the combination judgment belief and ignorance of each combination rule, the treatment of conflict judgments given by the two sources of information, and the applications of combination rules. The new results yield valuable theoretical insight into the rules that can be applied to a given situation. Zadeh's typical example is included in this chapter to evaluate the performance as well as efficiency of each combination rule of evidence for the conflict judgments given by the two sources of information.

9.8 References

[1] Dempster A.P., *Upper and Lower Probabilities Induced by a Multivalued Mapping*, Ann. Math. Statistics, Vol.38, pp.325-339, 1967.

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

[3] Lowrance D.J., *Automated Argument Construction*, J. Stat.Planning Inference,20,369-387,1988.

[4] Shafer G., *Belief Functions and Possibility Measures*, in The Analysis of Fuzzy Information, Vol.1 (J.C.Bezdek, Ed.), CRC Press, Boca Raton, Fla.,1987.

[5] Goodman I.R., Nguyen H.T., *Uncertainty Models for Knowledge-Based Systems*, Horth-Holland, New York, 1985.

[6] Nguyen H.T., *On Random Sets and Belief Functions*, J. Math. Anal. Appl. 65, pp.531-542,1987.

[7] Shafer G., Shenoy P.P., Mellouli K., *Propagating Belief Functions in Qualitative Markov Trees*, Int. J. Approximate Reasoning 1, pp.349-400,1987.

[8] Fagin R., Halpern J.Y., *Uncertainty, Belief, and Probability*, IJCAI-89, pp.1161-1167, 1989.

[9] Ruspini E.H., *Epistemic Logics, Probability, and The Calculus of Evidence*, IJCAI-87, pp.924-931, 1987.

[10] Sun H., He K., Zhang B., Mao S., *The Performance of Fusion Judgment on Dempster-Shafer Rule*, Chinese Journal of Electronics, Vol.8, No.1, pp.86-91. Jan. 1999.

[11] Shafer G., *Perspectives on the Theory and Practice of Belief Functions*, International Journal of Approximate Reasoning, Vol.4, pp.323-362,1990.

[12] Waltz E., Llinas J. , *Multisensor Data Fusion*, Artech Norwood, Massachusetts,1990.

[13] Biswas G., Anand T.S., *Using the Dempster-Shafer Scheme in a Mixed-initiatice Expert System Shell*, in Uncertainty in Artificial Intelligence, Vol.3 (L.N.Kanal, T.S.Levitt, and J.F.Lemmer, Eds.), pp.223-239, North-Holland, New York, 1989.

[14] Bogler Ph.L., *Shafer-Dempster Reasoning with Applications to Multisensor Target Identification Systems*, IEEE Trans. on Systems, Man and Cybernetics, Vol.17, No.6, pp.968-977, 1987.

[15] Buede D.M., *Shafer-Dempster and Bayesian Reasoning: A Response to 'Shafer-Dempster Reasoning with Applications to Multisensor Target Identification Systems'*, IEEE Trans. on Systems, Man and Cybernetics, Vol.18, No.6, pp.1009-10, 1988.

[16] Buede D.M., Girardi P., *A Target Identification Comparison of Bayesian and Dempster-Shafer Multisensor Fusion*, IEEE Trans. on Systems, Man and Cybernetics-Part A: Systems and Humans, Vol.27, No.5, Sept.1997.

[17] Zadeh L.A., *On the Validity of Dempster's Rule of Combination of Evidence*, ERL Memo M 79/24, Univ. of California, Berkeley, 1979.

[18] Zadeh L.A., *Review of Shafer's A Mathematical Theory of Evidence*, AI Magazine 5(3), pp.81-83, 1984.

[19] Zadeh L.A., *A Simple View of the Dempster-Shafer Theory of Evidence and its Implications for the Rule of Combination*, AI Mag. 7(2),85-90, 1986.

[20] Yager R.R., *On the Dempster-Shafer Framework and New Combination Rules*, Information Sciences, Vol. 41, pp.93-137, 1987.

[21] Wu U.G., Yang J.Y., Lin K., Liu L.J., *On the Evidence Inference Theory*, Information Sciences, Vol.89, pp.245-260, 1996.

[22] Sun H., Majdi F., Farooq M., *Dempster-Shafer Combination Rule Induced by Multi-valued Mapping*, Proceedings of SPIE Vol.4729, 1-5 April, 2002, Orlando, Florida USA.

[23] Sun H., Farooq M., *A Novel Justification of the Disjunctive Combination Rule*, Proceedings of the Sixth International Conference on Information Fusion, pp. 1244-1251, July, 2003, Cairns, Queensland, Australia.

[24] Smets Ph., *The Combination of Evidence in the Transferable Belief Model*, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (5), pp.447-458, 1990.

[25] Smets Ph., Kennes R., *The Transferable Belief Model*, Artificial Intelliegence 66 (2), pp. 191-234, 1994.

[26] Dubois D., Prade H., *Representation and Combination of Uncertainty with Belief Functions and Possibility Measures*, Computational Intelligence, 4, pp.244-264, 1988.

[27] Dezert J., *Foundations for a New Theory of Plausible and Paradoxical Reasoning*, Information and Security Journal, Edited by Tz. Semerdjiev, ProCon Ltd., Sofia, Bulgaria, Vol. 9, pp. 13-57, 2002.

[28] Dezert J., Smarandache F., *Fusion of Uncertain and Paradoxical Evidences with the General DSm Hybrid Rule of Combination*, In Advances and Applications of DSmT for Information Fusion (Collected Works) by F. Smarandache & J. Dezert (Editors), American Research Press, Rehoboth, 2004.

[29] Smets Ph., *Belief Functions: The Disjunctive Rule of Combination and the Generalized Bayesian Theorem*, International Journal of Approximate Reasoning, Vol.1, 1-35, 1993.

[30] Smets Ph., *The Alpha-junctions: Combination Operatiors Aapplicable to Belief Function*, Qualitative and Quantitative Practical Reasoning, Springer, Gabbay, D.M. and Kruse R. and Nonnengart A. and Ohlbach H. J. Editors, 1997, 131-153.

[31] Dubois D., Prade H., *A Set-Theoretic View of Belief Functions*, International Journal of General Systems, pp.193-226, Vol.12, 1986.

[32] Lefevre E., Colot O., Vannoorenberghe P., *Belief Function Combination and Conflict Management*, Information Fusion, 3, pp. 149-162, 2002.

[33] Sun H., Mao S., *Multisensor Data Fusion for Target Identification*, Chinese Journal of Electronics, Vol.4,No.3, pp.78-84 , July 1995 and Acta Electronic Sinica, Vol.23, No.10,1995.

[34] Sun H., Chen G., Zhang B., He K., Mao S., *Dempster-Shafer Rule Fusion Identification on Sensitive Infromation Sources*, Chinese Journal of Electronics, Vol.11, No.1, pp.56-61, Jan. 2002.

[35] Sun H., Farooq M., *Conjunctive and Disjunctive Combination Rules of Evidence*, Proceedings of SPIE Vol.5429, April, 2004, Orlando, Florida USA.

[36] Gordon J., Shortliffe E.H., *The Depster-Shafer Theory of Evidence, in Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, (B.G.Buchanan and E.H. Shortliffe, Eds.), Addison-Wesley, Reading, Mass., 1984.

[37] Klopotek M.A., *Interpretation of Belief Function in Dempster-Shafer Theory*, Foundations of Computing and Decision Sciences, Vol.20, No.4, pp.287-306, 1995.

[38] Klopotek M.A., *Identification of Belief Structure in Dempster-Shafer Theory*, Foundations of Computing and Decision Sciences, Vol.21, No.1, pp. 35-54.1996.

Chapter 10

Comparison between DSm and MinC combination rules

Milan Daniel

Institute of Computer Science

Academy of Sciences of the Czech Republic

Pod vodárenskou věží 2, CZ - 182 07 Prague 8

Czech Republic

Abstract: *Both DSm and minC rules of combination endeavor to process conflicts among combined beliefs better. The nature of conflicts as well as their processing during the belief combination is sketched. An presentation of the minC combination, an alternative to Dempster's rule of combination, follows. Working domains, structures and mechanisms of the DSm and minC combination rules are compared in the body of this chapter. Finally, some comparative examples are presented.*

10.1 Introduction

The classical DSm rule of combination, originally presented in [5, 6], has served for combination of two or several beliefs on the free DSm model. Later, a hybrid DSm combination rule has been developed to be applicable also on the classical Shafer (or Dempster-Shafer, DS) and the hybrid DSm model. The present state of the DSm rule is described in Chapter 4, see Equations (4.7)-(4.10).

Partial support by the COST action 274 TARSKI is acknowledged.

MinC combination (minimal conflict/minimal contradiction) rule introduced in [2, 4] is an alternative to the Dempter's rule of combination on the classical DS model. This rule has been developed for better handling of conflicting situations, which is a weak point of the classical Dempster rule. A brief description of the idea of the minC combination is presented in Section 10.3.

Both arguments and results of the DSm rule are beliefs in a DSm model, which admits intersections of elements of the frame of discernment in general. The minC combination serves for combination of classical belief functions (BFs) where all intersections of elements (of the frame of discernment) are empty and their resulting basic belief masses should be 0.

For finer processing of conflicts than the classical normalization in Dempster rule, a system of different types of conflict (or empty set) is introduced. For representation of intermediate results, generalized BFs serve on generalized frames of discernment which contains elements of the classical DS frame of discernment and correspondent types of conflict.

Even if the two developed approaches were originally different (disjoint), as well as the paradigms of both approaches, the intermediate working generalized beliefs of the minC combination are similar to those in the free DSm model, and the way of combination on the generalized level is analogous to that in the free DSm model. This surprising fact is the main reason why we compare these two seemingly incomparable, and originally quite disjoint approaches.

Now, after the development of the DSm combination for any hybrid DSm model, it is, moreover, possible to compare behavior of both approaches on classical BFs, i.e. in the application domain of the minC combination.

10.2 Conflict in belief combination

In the DSm combination, which is specially designed for conflicting situations, there are no problems with conflicts.

The common similar principle for Dempster rule, the minC combination and the DSm combination rule is that the basic belief assignment/mass (bbm) $m_1(X)$, assigned to set X by the first basic belief assignment (bba) m_1, multiplied by bbm $m_2(Y)$, assigned to set Y by the second bba m_2, is assigned to the set $X \cap Y$ by the resulting bba m_{12}, i.e. $m_1(X)m_2(Y)$ is a part of $m_{12}(X \cap Y)$.

This principle works relatively nicely if sets X and Y are not disjoint. There is also no problem for the DSm rule because $X \cap Y$ is always an element of D^Θ and its positive value is accepted even in the case of sets X and Y without any common element of Θ.

In Dempster's rule, disjoint X and Y tend to a conflict situation. All the conflicts are summed up together and reallocated onto 2^Θ by normalization in the classical normalized Dempster's rule, see [9], or stored as $m(\emptyset)$ in the non-normalized Dempster's rule in Transferable Belief Model (TBM) by Smets, see [10, 11]. It is a fact that in Smets' approach the normalization is only postponed from the combination process phase to the decisional one, as the normalization is the first step of computation of the classical pignistic transformation in TBM. The non-normalized Dempster rule commutes with the normalization, hence the pignistic probability is always the same in both the cases of normalized and non-normalized Dempster's rule.

A weak point of Dempster's rule — combination of conflicting beliefs is caused by normalization or by grouping all the conflicts together by the non-normalized version of Dempster's rule. Therefore, different types of conflict were introduced and a minC combination rule has been developed for a better handling of conflicting situations.

10.3 The minC combination

The minC combination (the minimal contradiction/conflict combination) of belief functions was developed [2, 4] with an effort to find a new associative combination which processes conflicts better than Dempster's rule. The classical Shafer model from Dempster-Shafer theory is supposed for both input and resulting belief functions. The minC combination is a generalization[1] of the un-normalized Dempster's rule. $m(\emptyset)$ is not considered as an argument for new unknown elements of the frame of discernment, $m(\emptyset)$ is considered as a conflict[2] arising by conjunctive combination. To handle it, a system of different types of conflicts is considered with respect to sets which produce the conflicts.

10.3.1 A system of different types of conflicts

We distinguish conflicts according to the sets to which the original bbms were assigned by m_i. There is only one type of conflict among the belief functions defined on a binary frame of discernment, hence the minC combination coincides with the non-normalized conjunctive rule there.

[1]Note that, on the other hand, the minC combination approach is a special case of an even more general approach of combination belief functions 'per elements', see [3]

[2]The term "contradiction" is used in [2, 4], while we use "conflict" here in order to have a uniform terminology.

In the case of an n-ary frame of discernment we distinguish different types of conflicts, e.g. $\{\theta_1\}\times\{\theta_2\}$, $\{\theta_1\}\times\{\theta_2,\theta_3\}$, $\{\theta_1\}\times\{\theta_2\}\times\{\theta_3\}$, $\{\theta_i,\theta_j,\theta_k\}\times\{\theta_m,\theta_n,\theta_o\}$ etc. The symbol $\times$ serves here for a denotation of conflicts, it is not used as any new operation on sets. Thus e.g. $\{\theta_1\}\times\{\theta_2,\theta_3\}$ simply denotes the conflict between sets $\{\theta_1\}$ and $\{\theta_2,\theta_3\}$.

We assume that products of the conflicting bbms are temporarily assigned (we all the time keep in mind that Shafer's constraints should be satisfied) to the corresponding conflicts: e.g. $m_1(\{\theta_1\})m_2(\{\theta_2\})$ is assigned to the conflict $\{\theta_1\}\times\{\theta_2\}$. In this way we obtain so called *generalized bbas*, and *generalized BFs* on a *generalized frame of discernment* given by Θ.

When combining 2 BFs defined on 3D frame $\Theta = \{\theta_1,\theta_2,\theta_3\}$ we obtain the following conflicts as intersections of disjoint subsets of Θ: $\{\theta_1\}\times\{\theta_2\}$, $\{\theta_1\}\times\{\theta_3\}$, $\{\theta_2\}\times\{\theta_3\}$, $\{\theta_1,\theta_2\}\times\{\theta_3\}$, $\{\theta_1,\theta_3\}\times\{\theta_2\}$, and $\{\theta_2,\theta_3\}\times\{\theta_1\}$.

Because we need a classical BF as a result of the combination, we have to reallocate bbms assigned to conflicts among subsets of Θ after the combination. These bbms are proportionalized, i.e. proportionally distributed, among subsets of Θ corresponding to the conflicts. A few such *proportionalizations* are presented in [4]. Unfortunately, all these proportionalizations break required associativity of the conjunctive combination. To keep the associativity as long as possible we must be able to combine the generalized belief functions with other BFs and generalized BFs. From this reason other conflicts arise: e.g. $\{\theta_1\}\times\{\theta_2\}\times\{\theta_3\}$, $(\{\theta_1,\theta_2\}\times\{\theta_1,\theta_3\})\times\{\theta_2\}\times\{\theta_3\}$, $(\{\theta_1,\theta_2\}\times\{\theta_3\})\times(\{\theta_2\}\times\{\theta_3\})$, etc.

A very important role for keeping associativity is played by so called *partial* or *potential conflicts* [3], e.g. a partial conflict $\{\theta_1,\theta_2\}\times\{\theta_2,\theta_3\}$ which is not a conflict in the case of combination of two beliefs $\{\theta_1,\theta_2\}\cap\{\theta_2,\theta_3\}=\{\theta_2\}$, but it can cause a conflict in a later combination with another belief, e.g. *pure* or *real conflict* [4] $\{\theta_1,\theta_2\}\times\{\theta_2,\theta_3\}\times\{\theta_1,\theta_3\}$ because there is $\{\theta_1,\theta_2\}\cap\{\theta_2,\theta_3\}\cap\{\theta_1,\theta_3\}=\emptyset$, in Shafer's model.

In order not to have an infinite number of different conflicts, the conflicts are divided into classes of equivalence $\sim$ which are called *types of conflicts*, e.g. $\{\theta_1\}\times\{\theta_2\}\sim\{\theta_2\}\times\{\theta_1\}\sim\{\theta_1\}\times\{\theta_2\}\times\{\theta_2\}\times\{\theta_2\}\times\{\theta_1\}\times\{\theta_1\}\times\{\theta_1\}$, etc. The minC combination works with these classes of equality (types of conflict) instead of the set of all different conflicts. For more details see [4].

[3] *Potential contradictions* in the original terminology of [2, 4]

[4] *A real contradiction* in [2, 4].

The conflicts are considered "per elements" in the following way: conflict $\{\theta_1, \theta_2\} \times \{\theta_3\}$ is considered as a set of *elementary conflicts* $\{\{\theta_1\} \times \{\theta_3\},\ \{\theta_2\} \times \{\theta_3\}\}$, i.e. set of conflicts between/among singletons. Analogically, potential conflict $\{\theta_1, \theta_2\} \times \{\theta_2, \theta_3\}$ is considered as a set of elementary conflicts $\{\{\theta_1\}\times\{\theta_2\},\ \{\theta_1\}\times\{\theta_3\},\ \{\theta_2\},\ \{\theta_2\}\times\{\theta_3\}\}$, where $\{\theta_2\} \sim \{\theta_2\} \times \{\theta_2\}$ is so called *trivial conflict*[5], i.e. no conflict in fact. Note that any partial conflict contains at least one trivial conflict. The set of elementary conflicts is constructed similarly to the Cartesian product of conflicting sets, where $\{\theta_1\} \times \{\theta_2\} \times ... \times \{\theta_k\}$ is used instead on n-tuple $[\theta_1, \theta_2, ..., \theta_k]$. As the above equivalence $\sim$ of elementary conflicts is used, we have elementary conflicts of different n-arity in the same set, thus we do not use n-tuples as it is usual in the Cartesian product. The idea of "conflicts per elements" was generalized also for non-conflicting sets in the "combination per elements", see [3].

For further decreasing of the number of types of conflicts we consider only minimal conflicts in the following sense: $\{\theta_1\} \times \{\theta_2\},\ \{\theta_3\}$, are minimal conflicts of the set $\{\{\theta_1\} \times \{\theta_2\},\ \{\theta_3\},\ \{\theta_1\} \times \{\theta_2\} \times \{\theta_3\}, \{\theta_1\}\times\{\theta_2\}\times\{\theta_4\}\times\{\theta_5\},\ \{\theta_1\}\times\{\theta_3\}\times\{\theta_5\}\}$; i.e. the set of singletons contained in a minimal conflict is minimal from the point of view of inclusion among all sets of singletons corresponding to elementary conflicts. Thus $\{\{\theta_1\}\times\{\theta_2\}, \{\theta_3\}\} \sim \{\{\theta_1\}\times\{\theta_2\}, \{\theta_3\}, \{\theta_1\}\times\{\theta_2\}\times\{\theta_3\}, \{\theta_1\}\times\{\theta_2\}\times\{\theta_4\}\times\{\theta_5\}, \{\theta_1\}\times\{\theta_3\}\times\{\theta_5\}\}$. Our concentration only to minimal conflicts brings us a simplification, which is closer to Shafer's model, and it has no influence on associativity of combination.

In this way we obtain 8 types of conflicts ($\{\theta_1\}\times\{\theta_2\},\ \{\theta_1\}\times\{\theta_3\},\ \{\theta_2\}\times\{\theta_3\},\ \{\theta_1\}\times\{\theta_2\}\times\{\theta_3\}, \{\{\theta_1\}\times\{\theta_2\},\ \{\theta_1\}\times\{\theta_3\}\},\ \{\{\theta_1\}\times\{\theta_2\},\ \{\theta_2\}\times\{\theta_3\}\},\ \{\{\theta_1\}\times\{\theta_3\},\ \{\theta_2\}\times\{\theta_3\}\},\ \{\{\theta_1\}\times\{\theta_2\},\ \{\theta_1\}\times\{\theta_3\},\ \{\theta_2\}\times\{\theta_3\}\}$) and 3 types of potential conflicts ($\{\{\theta_1\},\ \{\theta_2\} \times \{\theta_3\}\},\ \{\{\theta_2\},\ \{\theta_1\} \times \{\theta_3\}\},\ \{\{\theta_3\},\ \{\theta_1\} \times \{\theta_2\}\}$) in a 3D case $\Theta = \{\theta_1, \theta_2, \theta_3\}$. Together with 7 non-conflicting subsets of Θ we have 18 sets of conflicts to which nonnegative bbms can be assigned in the 3D case, or 18 elements of a generalized 3D frame of discernment.

10.3.2 Combination on generalized frames of discernment

As minC combination has a nature of a conjunctive rule of combination, $m_1(X)m_2(Y)$ is assigned to $X \cap Y$, if it is non-empty, or to $X \times Y$ otherwise. More precisely the least representative of the type of conflict of $X \times Y$ is considered instead of $X \times Y$. It is unique but an order of elementary conflicts and an order of elements inside elementary conflicts. A fixation of these orders enables a unique selection of representatives of $\sim$ classes of conflicts. A complete 18x18 table of minC combination for 3D is presented in [2, 4]. We include here only an illustrative part of it, see Table 10.1. The resulting value $m^0(Z)$ of the generalized bba is computed as a sum of all $m_1(X)m_2(Y)$ for which the field of the complete table in the

[5] *A trivial contradiction.*

row corresponding to X and column corresponding to Y contains Z. In other words, generalized $m^0(Z)$ is computed as a sum of all $m_1(X)m_2(Y)$ for which $Z = X \cap Y$ if $(X \subseteq Y) \vee (Y \subseteq X)$ or $Z \sim X \times Y$ otherwise, where $\sim$ is the equivalence of conflicts from the previous subsection (Z and $X \times Y$ are in the same $\sim$ class of conflicts.); i.e.

$$m^0(Z) = \sum_{\substack{Z=X\cap Y \\ X\subseteq Y \vee Y\subseteq X}} m_1(X)m_2(Y) + \sum_{\substack{Z\sim X\times Y \\ X\nsubseteq Y \& Y\nsubseteq X}} m_1(X)m_2(Y). \tag{10.1}$$

In order to decrease the size of the table below, the following abbreviations are used in this table: A stands for $\{A\}$, similarly AB stands for $\{A, B\}$, and ABC stands for $\{A, B, C\}$, $A \times B$ stands for $\{A\} \times \{B\}$, similarly $A \times BC$ stands for $\{A\} \times \{B, C\}$, $\times$ stands for $\{A\} \times \{B\} \times \{C\}$, $\square A$ stands for $\square\{A\}$, and $\square$ stands for $\{A, B\} \times \{A, C\} \times \{B, C\}$, and similarly.

	A	B	AB	ABC	$A\times B$	$A\times BC$	$\times$	$\square$	$\square A$
A	A	$A\times B$	A	A	$A\times B$	$A\times BC$	$\times$	$A\times BC$	A
B	$A\times B$	B	B	B	$A\times B$	$A\times B$	$\times$	$B\times AC$	$B\times AC$
C	$A\times C$	$B\times C$	$C\times AB$	C	$\times$	$A\times C$	$\times$	$C\times AB$	$C\times AB$
BC	$A\times BC$	B	$\square B$	BC	$A\times B$	$A\times BC$	$\times$	$\square$	$\square$
AC	A	$B\times AC$	$\square A$	AC	$A\times B$	$A\times BC$	$\times$	$\square$	$\square A$
AB	A	B	AB	AB	$A\times B$	$A\times BC$	$\times$	$\square$	$\square A$
ABC	A	B	AB	ABC	$A\times B$	$A\times BC$	$\times$	$\square$	$\square A$
$A\times B$	$A\times B$	$A\times B$	$A\times B$	$A\times B$	$A\times B$	$A\times B$	$\times$	$A\times B$	$A\times B$
$A\times C$	$A\times C$	$\times$	$A\times C$	$A\times C$	$\times$	$\times$	$\times$	$A\times C$	$A\times C$
$B\times C$	$\times$	$B\times C$	$B\times C$	$B\times C$	$\times$	$A\times C$	$\times$	$B\times C$	$B\times C$
$A\times BC$	$A\times BC$	$A\times B$	$A\times BC$	$A\times BC$	$A\times B$	$A\times BC$	$\times$	$A\times BC$	$A\times BC$
$B\times AC$	$A\times B$	$B\times AC$	$B\times AC$	$B\times AC$	$A\times B$	$A\times B$	$\times$	$B\times AC$	$B\times AC$
$C\times AB$	$A\times C$	$B\times C$	$C\times AB$	$C\times AB$	$\times$	$A\times C$	$\times$	$C\times AB$	$C\times AB$
$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
$\square$	$A\times BC$	$B\times AC$	$\square$	$\square$	$A\times B$	$A\times BC$	$\times$	$\square$	$\square$
$\square A$	A	$B\times AC$	$\square A$	$\square A$	$A\times B$	$A\times BC$	$\times$	$\square$	$\square A$
$\square B$	$A\times BC$	B	$\square B$	$\square B$	$A\times B$	$A\times BC$	$\times$	$\square$	$\square$
$\square C$	$A\times BC$	$B\times AC$	$\square$	$\square C$	$A\times B$	$A\times BC$	$\times$	$\square$	$\square$

Table 10.1: A partial table of combination of 2 generalized BFs on $\Theta = \{A, B, C\}$.

The minC combination is commutative and associative on generalized BFs. It overcomes some disadvantages of both Dempster's rules (normalized and un-normalized). This theoretically nice combining rule has however a computational complexity rapidly increasing with the size of the frame of discernment.

10.3.3 Reallocation of belief masses of conflicts

Due to the belief masses being assigned also to types of conflicts and partial conflicts, the result of the minC combination is a generalized belief function even if it is applied to classical BFs. To obtain a classical belief function on Shafer's model we have to do the following two steps: we first reassign the bbms of partial conflicts to their non contradictive elements and then we proportionalize bbms of pure (real) conflicts. Because of a different nature of pure and partial conflicts, also these two steps of bbms reallocation are different.

10.3.3.1 Reallocation of gbbms of partial conflicts

Gbbms of partial conflicts (potential contradictions) are simply reassigned to the sets of their trivial conflicts, i.e. to the sets of their non-contradictive elements (e.g. $m^0(\{\theta_i,\theta_j\}\times\{\theta_i,\theta_k\})$ is reallocated to $\{\theta_i\}$). We denote resulting gbba of this step with m^1 to distinguish it from gbba m^0 on the completely generalized level. Thus we obtain $m^1(\{\theta_i,\theta_j\}\times\{\theta_i,\theta_k\}) = 0$ and $m^1(\{\theta_i\})$ is a sum of all $m^0(X)$, where $\{\theta_i\}$ is maximal nonconflicting part of X. Nothing is performed with gbbms of pure conflicts in this step, hence $m^1(Y) = m^0(Y)$ for any pure conflict Y.

10.3.3.2 Proportionalization of gbbms of pure conflicts

Let us present two ways how to accomplish a proportionalization of gbbms which has been assigned by m^0 to pure (real) conflicts . The basic belief mass of a conflict $X \times Y$ between two subsets of Θ can be proportionalized, i.e. reallocated according to the proportions of the corresponding non-conflicting bbms:

a) among X, Y, and $X \cup Y$ as originally designed for so called proportionalized combination rule in [1].

b) among all nonempty subsets of $X \cup Y$. This way combines the original idea of proportionalization with the consideration of conflict "per elements".

For a conflict X of several subsets of a frame of discernment $X_1, X_2, ..., X_k \subset \Theta$, e.g. for $\{\theta_1\}\times\{\theta_2\}\times\{\theta_3\}$ and $\Box \sim \{\{\theta_1\}\times\{\theta_2\}, \{\theta_1\}\times\{\theta_3\}, \{\theta_2\}\times\{\theta_3\}\} \sim \{\theta_1,\theta_2\}\times\{\theta_1,\theta_3\}\times\{\theta_2,\theta_3\}$ in 3D and further conflicts from nD case, we have to generalize the above description of proportionalization in the following way. The bbm of contradiction $X = X_1 \times X_2 \times ... \times X_k$ can be proportionalized:

a) among all unions $\bigcup_{i=1}^{j} X_i$ of $j \leq k$ sets X_i from $\{X_1, X_2, ..., X_k\}$.

b) among all nonempty subsets of $X_1 \cup X_2 \cup ... \cup X_k$.

For an explicit expression, the conflicts of the subsets of 3D $\Theta = \{\theta_1,\theta_2,\theta_3\}$ should be proportionalized among, see Table 10.2. The bbms of conflicts in the first column should be proportionalized by the proportionalization ad a) among sets in the second column and by the proportionalization ad b) among the sets in the third column.

If gbbms $m^1(X_i) = 0$ for all X_i then we divide the proportionalized gbbm $m^1(X_1 \times X_2 \times ... \times X_k)$ by number of the sets among them the gbbm should be proportionalized, i.e. by $2^k - 1$ in the proportionalization a) and by $2^m - 1$, where $m = |X_1 \cup X_2 \cup ... \cup X_k|$ in the case b).

Type of conflict	Proportionalization ad a)	Proportionalization ad b)
$\{\theta_1\}\times\{\theta_2\}$	$\{\theta_1\}, \{\theta_2\}, \{\theta_1, \theta_2\}$	$\{\theta_1\}, \{\theta_2\}, \{\theta_1, \theta_2\}$
$\{\theta_1\}\times\{\theta_2, \theta_3\}$	$\{\theta_1\}, \{\theta_2, \theta_3\}, \{\theta_1, \theta_2, \theta_3\}$	$\mathcal{P}(\{\theta_1, \theta_2, \theta_3\}) - \emptyset$
$\{\theta_1, \theta_2\}\times\{\theta_1, \theta_3\}\times\{\theta_2, \theta_3\}$	$\{\theta_1, \theta_2\}, \{\theta_1, \theta_3\}, \{\theta_2, \theta_3\}, \{\theta_1, \theta_2, \theta_3\}$	$\mathcal{P}(\{\theta_1, \theta_2, \theta_3\}) - \emptyset$
$\{\theta_1\}\times\{\theta_2\}\times\{\theta_3\}$	$\mathcal{P}(\{\theta_1, \theta_2, \theta_3\}) - \emptyset$	$\mathcal{P}(\{\theta_1, \theta_2, \theta_3\}) - \emptyset$

Table 10.2: Proportionalizations on a 3D frame of discernment

A proportionalization of the types of the conflicts from the Table is the same even if $\{\theta_1, \theta_2, \theta_3\} \subsetneq \Theta$. Hence we can see from the Table that the proportionalization is something like 'local normalization' on the power set of $\Theta' \subsetneq \Theta$ in the case b) or on a subset of such power set. E. g. $m^1(\{\theta_1\}\times\{\theta_2,\theta_3\})$ is proportionalized with proportionalization a) among $\{\theta_1\}, \{\theta_2, \theta_3\}, \{\theta_1, \theta_2, \theta_3\}$ so that $\frac{m^1(\{\theta_1\})}{m^1(\{\theta_1\})+m^1(\{\theta_2,\theta_3\})+m^1(\{\theta_1,\theta_2,\theta_3\})}$ $m^1(\{\theta_1\}\times\{\theta_2,\theta_3\})$ is assigned to $\{\theta_1\}$, $\frac{m^1(\{\theta_2,\theta_3\})}{m^1(\{\theta_1\})+m^1(\{\theta_2,\theta_3\})+m^1(\{\theta_1,\theta_2,\theta_3\})}$ $m^1(\{\theta_1\}\times\{\theta_2,\theta_3\})$ is assigned to $\{\theta_2, \theta_3\}$, and $\frac{m^1(\{\theta_1,\theta_2,\theta_3\})}{m^1(\{\theta_1\})+m^1(\{\theta_2,\theta_3\})+m^1(\{\theta_1,\theta_2,\theta_3\})}$ $m^1(\{\theta_1\}\times\{\theta_2,\theta_3\})$ is assigned to $\{\theta_1, \theta_2, \theta_3\}$. Analogically $\frac{m^1(\{\theta_2,\theta_3\})}{m^1(\{\theta_1\})+m^1(\{\theta_2\})+m^1(\{\theta_3\})+m^1(\{\theta_1,\theta_2\})+m^1(\{\theta_1,\theta_3\})+m^1(\{\theta_2,\theta_3\})+m^1(\{\theta_1,\theta_2,\theta_3\})}$ $m^1(\{\theta_1\}\times\{\theta_2,\theta_3\})$ is assigned to $\{\theta_2, \theta_3\}$ with proportionalization b), and similarly for other subsets of $\{\theta_1, \theta_2, \theta_3\}$. For single elementary conflicts both the proportionalizations coincide, see e.g. the 1st and the 4th rows of the Table 10.2. Specially there is the only proportionalization in the 2D case because, there is the only conflict and it is an elementary one. This proportionalization actually coincides with the classical normalization, see examples in Section 10.5.

Let us remember that neither the reallocation of gbbms of partial conflicts nor the proportionalization does not keep associativity of minC combination of the generalized level. Hence we have always to keep in the consideration and to save the generalized version of the result to be prepared for a later combination with another belief.

10.3.4 Summary of the idea of the minC combination

We can summarize the process of the minC combination of $n \geq$ beliefs as follows:

1. we apply $(n - 1)$ times the generalized version of minC, to compute gbba m^0, see formula (10.1);
2. after we once apply a reallocation of gbbms of the partial conflicts to produce gbba m^1 and finally we once apply the proportionalization a) or b) to obtain the final bbm m. If we want to keep as

much as possible of associativity for future combining, we have to remember also the gbbm m^0 and continue further combination (if there is any) from it.

10.4 Comparison

10.4.1 Comparison of generalized frames of discernment

As has been already mentioned in the introduction of this chapter, DSm and minC rules of combination arise from completely different assumptions and ideas. On the other hand, 18 different subsets of a frame of discernment and types of conflicts and potential conflicts (7+8+3) in 3D case or 18 elements of a generalized 3D frame of discernment correspond to 18 non empty elements of hyper-power set D^Θ in the free DSm model. Moreover, if we rewrite subsets of the frame of discernment, e.g. $\{\theta_i, \theta_j, \theta_k\}$, and sets of elementary conflicts as unions of their elements, e.g. $\{\theta_i, \theta_j, \theta_k\} \sim \theta_i \cup \theta_j \cup \theta_k$, and conflicts as intersections, e.g. $\{\theta_i\} \times \{\theta_j\} \sim \theta_i \cap \theta_j$, $\{\theta_i, \theta_j\} \times \{\theta_i, \theta_k\} \sim (\theta_i \cup \theta_j) \cap (\theta_i \cup \theta_k)$, $\{\{\theta_i\} \times \{\theta_j\}, \{\theta_j\} \times \{\theta_k\}, \{\theta_i\} \times \{\theta_j\} \times \{\theta_k\}\} \sim (\theta_i \cap \theta_j) \cup (\theta_j \cap \theta_k) \cup (\theta_i \cap \theta_j \cap \theta_k)$, then we obtain the following:

$\{\theta_1\} \sim \theta_1 = \alpha_9$

$\{\theta_2\} \sim \theta_2 = \alpha_{10}$

$\{\theta_3\} \sim \theta_3 = \alpha_{11}$

$\{\theta_1, \theta_2\} \sim \theta_1 \cup \theta_2 = \alpha_{15}$

$\{\theta_1, \theta_3\} \sim \theta_1 \cup \theta_3 = \alpha_{16}$

$\{\theta_2, \theta_3\} \sim \theta_2 \cup \theta_3 = \alpha_{17}$

$\{\theta_1, \theta_2, \theta_3\} \sim \theta_1 \cup \theta_2 \cup \theta_3 = \alpha_{18}$

$\{\theta_1\} \times \{\theta_2\} \sim \theta_1 \cap \theta_2 = \alpha_2$

$\{\theta_1\} \times \{\theta_3\} \sim \theta_1 \cap \theta_3 = \alpha_3$

$\{\theta_2\} \times \{\theta_3\} \sim \theta_2 \cap \theta_3 = \alpha_4$

$\{\theta_1\} \times \{\theta_2, \theta_3\} = \{\{\theta_1\} \times \{\theta_2\}, \{\theta_1\} \times \{\theta_3\}\} \sim \theta_1 \cap (\theta_2 \cup \theta_3) = \alpha_7$

$\{\theta_2\} \times \{\theta_1, \theta_3\} = \{\{\theta_1\} \times \{\theta_2\}, \{\theta_2\} \times \{\theta_3\}\} \sim \theta_2 \cap (\theta_1 \cup \theta_3) = \alpha_6$

$\{\theta_3\} \times \{\theta_1, \theta_2\} = \{\{\theta_3\} \times \{\theta_1\}, \{\theta_3\} \times \{\theta_2\}\} \sim \theta_3 \cap (\theta_1 \cup \theta_2) = \alpha_5$

$\{\theta_1\} \times \{\theta_2\} \times \{\theta_3\} \sim \theta_1 \cap \theta_2 \cap \theta_3 = \alpha_1$

$\{\{\theta_1\} \times \{\theta_2\}, \{\theta_1\} \times \{\theta_3\}, \{\theta_2\} \times \{\theta_3\}\} \sim (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_1 \cap \theta_3) = \alpha_8$

$\square\theta_1 = \{\{\theta_1\}, \{\theta_2\} \times \{\theta_3\}\} \sim \theta_1 \cup (\theta_2 \cap \theta_3) = \alpha_{14}$

$\square\theta_2 = \{\{\theta_2\}, \{\theta_1\} \times \{\theta_3\}\} \sim \theta_2 \cup (\theta_1 \cap \theta_3) = \alpha_{13}$

$\square\theta_3 = \{\{\theta_3\}, \{\theta_1\} \times \{\theta_2\}\} \sim \theta_3 \cup (\theta_1 \cap \theta_2) = \alpha_{12}$.

Thus a generalized frame of discernment from the minC approach uniquely corresponds to $D^\Theta - \emptyset$. Hence the minC approach is an alternative way how to generate Dedekind's lattice.

10.4.2 Comparison of principles of combination

For bbms of two non-conflicting sets $X, Y \subset \Theta$ both the minC and the DSm rules assign the product of the belief masses to the intersection of the sets[6]. If one of the sets (or both of them) is (are) conflicting, then the minC combination assigns the product of their bbms to the conflict $X \times Y$. Similarly as above, we can consider this conflict as an intersection $X \cap Y$. We should verify whether $X \cap Y$ really corresponds to the corresponding field of the minC combination table.

As first example, let's denote by definition $A_1 \triangleq \{\theta_1, \theta_3\} \times (\{\theta_3\} \times \{\theta_1, \theta_2\})$, then one has

$$\begin{aligned}
A_1 &\sim (\theta_1 \cup \theta_3) \cap (\theta_3 \cap (\theta_1 \cup \theta_2)) = (\theta_1 \cap (\theta_3 \cap (\theta_1 \cup \theta_2))) \cup (\theta_3 \cap (\theta_3 \cap (\theta_1 \cup \theta_2))) \\
&= (\theta_3 \cap (\theta_1 \cap (\theta_1 \cup \theta_2))) \cup (\theta_3 \cap (\theta_1 \cup \theta_2)) = (\theta_3 \cap \theta_1) \cup (\theta_3 \cap (\theta_1 \cup \theta_2)) = (\theta_3 \cap (\theta_1 \cup \theta_2)) \\
&\sim \{\theta_3\} \times \{\theta_1, \theta_2\}
\end{aligned}$$

As second example, let's denote $A_2 \triangleq (\{\theta_1\} \times \{\theta_2\} \times \{\theta_3\}) \times \{\{\theta_1\} \times \{\theta_2\}, \{\theta_1\} \times \{\theta_3\}, \{\theta_2\} \times \{\theta_3\}\}$, then one has

$$\begin{aligned}
A_2 &\sim (\theta_1 \cap \theta_2 \cap \theta_3) \times ((\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)) \\
&\sim (\theta_1 \cap \theta_2 \cap \theta_3) \cap ((\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)) \\
&= \theta_1 \cap \theta_2 \cap \theta_3 \cap ((\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)) = (\theta_1 \cap \theta_2 \cap \theta_3) \cup (\theta_1 \cap \theta_2 \cap \theta_3) \cup (\theta_1 \cap \theta_2 \cap \theta_3) \\
&= \theta_1 \cap \theta_2 \cap \theta_3 \sim \{\theta_1\} \times \{\theta_2\} \times \{\theta_3\}
\end{aligned}$$

As third example, let's denote $A_3 \triangleq \Box\{\theta_1\} \times (\theta_1 \times \{\theta_2, \theta_3\})$, then one has

$$\begin{aligned}
A_3 &= \{\{\theta_1\}, \{\theta_2 \times \theta_3\}\} \times (\theta_1 \times \{\theta_2, \theta_3\}) \\
&\sim (\theta_1 \cup (\theta_2 \cap \theta_3)) \cap (\theta_1 \cap (\theta_2 \cup \theta_3)) = (\theta_1 \cup (\theta_2 \cap \theta_3)) \cap (\theta_1 \cap (\theta_2 \cup \theta_3)) \\
&= (\theta_1 \cap (\theta_1 \cap (\theta_2 \cup \theta_3))) \cup ((\theta_2 \cap \theta_3)) \cap (\theta_1 \cap (\theta_2 \cup \theta_3)) \\
&= (\theta_1 \cap (\theta_2 \cup \theta_3)) \cup ((\theta_2 \cap \theta_3) \cap (\theta_1 \cap (\theta_2 \cup \theta_3))) \\
&= (\theta_1 \cap (\theta_2 \cup \theta_3)) \cup ((\theta_2 \cap \theta_3 \cap \theta_1 \cap \theta_2) \cup (\theta_2 \cap \theta_3 \cap \theta_1 \cap \theta_3)) \\
&= (\theta_1 \cap (\theta_2 \cup \theta_3)) \cup ((\theta_1 \cap \theta_2 \cap \theta_3) \cup (\theta_1 \cap \theta_2 \cap \theta_3)) \\
&= (\theta_1 \cap (\theta_2 \cup \theta_3)) \cup (\theta_1 \cap \theta_2 \cap \theta_3) = (\theta_1 \cap (\theta_2 \cup \theta_3)) \sim (\theta_1 \times \{\theta_2, \theta_3\})
\end{aligned}$$

[6] We have to mention here that the minC combination rule has never been formulated as a k-ary operator for combination of $k \geq 2$ belief sources, analogically to the DSm combination rule, see Equations (4.2) and (4.5). Nevertheless, it is theoretically very easy to explicitly formulate it similarly to the DSm rule for k sources. Moreover, because of its associativity on the generalized level we can obtain the same result by step-wise ($(k-1)$-times) application of the binary form, and continue with reallocation of bbms of conflicts as is usual.

In the case of $\{\theta_1,\theta_3\}\times\{\theta_1,\theta_2\} \sim (\theta_1\cup\theta_3)\times(\theta_1\cup\theta_2) \sim (\theta_1\cup\theta_3)\cap(\theta_1\cup\theta_2) = (\theta_1\cap(\theta_1\cup\theta_2)\cup(\theta_3\cap(\theta_1\cup\theta_2) = (\theta_1\cap\theta_1\cup\theta_1\cap\theta_2)\cup(\theta_3\cap\theta_1\cup\theta_3\cap\theta_2) = (\theta_1)\cup(\theta_3\cap\theta_1)\cup(\theta_3\cap\theta_2) = (\theta_1)\cup(\theta_3\cap\theta_2) \sim \{\{\theta_1\},\{\theta_2\times\theta_3\}\} \sim \Box\{\theta_1\}$ we can show again that minC combination of bbms of sets $\{\theta_1,\theta_3\},\{\theta_1,\theta_2\}$ corresponds to the intersection of the corresponding elements of D^Θ: $(\theta_1\cup\theta_3)$ and $(\theta_1\cup\theta_2)$, i.e. to $\theta_1\cup(\theta_3\cap\theta_2)$. Moreover, this shows a rise and the importance of a partial conflict (or potential contradiction) between two sets with non-empty intersection $\{\theta_1,\theta_3\}\cap\{\theta_1,\theta_2\} = \{\theta_1\}$ in Shafer's model. This intersection $\{\theta_1\}$ which is used in Dempster's rule, is different from the generalized minC and the free DSm intersection $\{\theta_1,\theta_3\}\cap\{\theta_1,\theta_2\} \sim (\theta_1\cup\theta_3)\cap(\theta_1\cup\theta_2) = (\theta_1)\cup(\theta_3\cap\theta_2) \sim \Box\{\theta_1\}$ on the generalized level.

Analogically we can verify that all the fields in the complete minC combination table uniquely correspond to intersections of corresponding sets. For a general nD case it is possible to verify that the similarity relation $\sim$ on conflicts corresponds with properties of the lattice $\{\Theta,\cap,\cup\}$. Thus the minC combination equation (10.1) corresponds with the classical DSm combination equation (4.1).

Hence the minC combination[7] on a generalized level fully corresponds to the DSm combination rule on a free DSm model.

10.4.3 Two steps of combination

Because minC is not designed for the DSm model but for the classical Shafer's model, we have to compare it in the context of the special Shaferian case of the hybrid DSm rule. According to the present development state of the hybrid DSm rule, see Chapter 4, in the first step all the combination is done on the free DSm model — it is fully equivalent to the generalized minC combination — and in the second step constraints are introduced. The second step is analogous to the reallocation in the minC approach. It does not explicitly distinguish anything like partial conflicts and pure conflicts, but analogically to the minC combination, bbms are reallocated in two different ways. An introduction of constraints can joint two or more elements of D^Θ, e.g. see Example 4 in Chapter 4, where the element α_9 is joined with the element α_{14}, and the elements α_{10}, and α_{11} are joined with α_{13} and α_{12} respectively. Gbbms of such elements are actually reallocated within this process. Really, the gbbms $m_{\mathcal{M}^f}(\alpha_9)$, $m_{\mathcal{M}^f}(\alpha_{10})$, and $m_{\mathcal{M}^f}(\alpha_{11})$ are reallocated to $m_{\mathcal{M}^0}(\alpha_{14})$, $m_{\mathcal{M}^0}(\alpha_{13})$ and $m_{\mathcal{M}^0}(\alpha_{12})$ respectively, as an analogy of the reallocation of partial conflicts in the minC approach. We can verify that the elements α_9, α_{10}, α_{11} really correspond to the partial conflicts of the minC approach. The step 2 consists further in grouping of all empty sets together and in the reallocation of their bbms. This action fully corresponds to a proportionalization of pure conflicts in the minC approach.

[7] For a comparison of the minC combination with other approaches for combination of conflicting beliefs, see [8].

Hence, the only principal difference between the minC and the DSm combination rules consists in reallocation of the bbms of conflicting (or empty) sets to non-conflicting (non-empty) ones, i.e. to the subsets of the frame of discernment, because the reallocation performed in the 2nd step of the hybrid DSm combination does not correspond to any of the above proportionalizations used in minC either.

10.4.4 On the associativity of the combination rules

As it was already mentioned both the DSm rule and the minC combination rule are fully associative on the generalized level, i.e. on the free DSm model in DSm terminology. Steps 2 in both the combinations, i.e. the introduction of constraints in DSm combination and the reallocation of conflicts including both the proportionalizations, do not keep associativity. If we use results of combination with all the constraints as an input for another combination, we obtain suboptimal results, see Section 4.5.5 in Chapter 4.

In order to keep as much associativity of the combination on the generalized level as possible, we have to use n-ary version of DSm rule. In the case where k input beliefs have been already combined, we have to save all the k input belief functions. If we want to combine the previous result with the new $(k+1)$th input m_{k+1}, then we have either to repeat all the n-ary combination for $k+1$ inputs this time, or we can use the free DSm result of the previous combination (the result of the last application of the Step 1) and apply the binary Step 1 to combine the new input (we obtain the same result as with an application of n-ary version for $k+1$ inputs). Nevertheless, after it we have to apply n-ary version of the Step 2 for introduction of all constraints at the end.

There is another situation in the case of the minC combination. Because we consider only minimal conflicts, the result of the Step 2 depends only on the generalized result m^0 of the Step 1 and we need not the input belief functions for the reallocation of partial conflicts and for the proportionalization. The non-normalized combination rule including the generalized one, provides the same result either if n-ary version is used for k inputs or if step-wise $k-1$ times the binary version is applied. Hence binary version of the generalized minC combination and unary reallocation satisfy for the optimal results in the sense of Chapter 4. If we already have k inputs combined, it is enough to save and store only the generalized result instead of all inputs. We perform the generalized combination with the input m_{k+1} after. And in the end we perform Step 2 for obtaining classical Shaferian result. Of course it is also possible to store all the inputs and to make a new combination, analogically, to the DSm approach.

10.4.5 The special cases

Specially in the 2D case minC corresponds to Dempster's rule — there is only one type of conflict and both the presented proportionalizations a) and b) coincide with normalization there. While the 2D DSm corresponds to Yager's rule, see [12], where $m_1(X)m_2(Y)$ is assigned to $X \cap Y$ if it is non-empty or to Θ for $X \cap Y = \emptyset$, and it also coincides with Dubois-Prade's rule, see [7], where $m_1(X)m_2(Y)$ is assigned to $X \cap Y$ if it is non-empty or to $X \cup Y$ otherwise. To complete the 2D comparison, it is necessary to add that the classical DSm combination rule for the 2D free DSm model corresponds to the non-normalized Dempster's rule used in TBM. For examples see Table 10.3 in Section10.5.

In an nD case for $n > 2$ neither the minC nor DSm rule correspond to any version of Dempster's or Yager's rules. On the other hand the binary version of the hybrid DSm rule coincides with Dubois-Prade's rule on Shafer's model, for an example see Table 10.6 in Section10.5.

10.4.6 Comparison of expressivity of DSm and minC approaches

As the minC combination is designed for combination of classical belief functions on frames of discernment with exclusive elements, we cannot explicitly express that 2 elements of frame have a non-empty intersection. The only way for it is a generalized result of combination of 2 classical BFs. On the other hand, even if the hyper-power set D^Θ has more elements than the number of parts in the corresponding Venn's diagram, we cannot assign belief mass to θ_1 but not to θ_2 in DSm approach. I. e. we cannot assign bbms in such a way that for generalized pignistic probability, see Chapter 7, the following holds: $P(\theta_1) > 0$ and $P(\theta_2) = 0$. The intersection $\theta_1 \cap \theta_2$ is always a subset both of θ_1 and θ_2. Hence from $m(\theta_1) > 0$ we always obtain $P(\theta_1 \cap \theta_2) > 0$ and $P(\theta_2) > 0$. We cannot assign any gbbm to $\theta_1 - \theta_2$. The only way how to do it is to add an additional constraint $\theta_1 \cap \theta_2 = \emptyset$, but such a constraint should be applied to all beliefs in the model and not only to one or several specific ones. As Shafer's model has already all the exclusivity constraints, the above described property is not related to it. Hence both the DSm approach and the minC combination have the comparable expressivity on Shafer's model. The DSm approach utilizes, in addition to it, its capability to express positive belief masses of the intersections.

10.5 Examples

In this section we present a comparison on examples of combination. The first 2D example simply compares not only the DSm and minC combination rules but also both the normalized and non-normalized Dempster's rule, Yager's rule, and Dubois-Prade's rule of belief combination, see Table 10.3. Because the proportionalizations a) and b) coincide in the 2D case, and subsequently the corresponding bbas $m_{12}^{a)}$ and $m_{12}^{b)}$ also coincide, we use m^{minC} for $m_{12}^{a)} \equiv m_{12}^{b)}$. This example enables us to make a wide comparison, but it does not really discover a nature of the presented approaches to the belief combination. For this reason we present also a more complicated 3D example, see Tables 10.4 and 10.5, which show us how conflicts and partial conflicts arise during combination, how constraints are introduced, and how proportionalizations are performed.

	m_1	m_2	$m_{12}^{\mathcal{M}^f}$	$m_{12}^{\mathcal{M}^0}$	m_{12}^0	m_{12}^{minC}	m_{12}^{TBM}	m_{12}^Y	m_{12}^{DP}	$m_{12}^{\oplus}$
$\theta_1 \sim \{\theta_1\}$	0.6	0.2	0.48	0.48	0.48	0.6000	0.48	0.48	0.48	0.6000
$\theta_2 \sim \{\theta_2\}$	0.1	0.3	0.17	0.17	0.17	0.2125	0.17	0.17	0.17	0.2125
$\theta_1 \cup \theta_2 \sim \{\theta_1, \theta_2\}$	0.3	0.5	0.15	0.35	0.15	0.1875	0.15	0.35	0.35	0.1875
$\theta_1 \cap \theta_2 \sim \{\theta_1\}\times\{\theta_2\} \sim \emptyset$			0.20		0.20		0.20			

Table 10.3: Comparison of combination of 2D belief functions

Table 10.4 provides a comparison of combination of 3D belief functions based on the free DSm model with the classic DSm rule and on Shafer's model with the hybrid DSm rule. The 5th column ($m_{12}^{\mathcal{M}^f}$) gives the result of the combination of the sources 1 and 2 obtained with the classic DSm rule based on the free DSm model. The 7th column ($m_{123}^{\mathcal{M}^f}$) gives the result of the combination of the sources 1, 2 and 3 obtained with the classic DSm rule based also on the free DSm model. Column 6 ($m_{12}^{\mathcal{M}^0}$) presents the result of the hybrid DSm combination of sources 1 and 2 based on Shafer's model $\mathcal{M}^0$. Column 8 ($m_{123}^{\mathcal{M}^0}$) presents the result of the hybrid DSm combination of sources 1, 2 and 3 based on Shafer's model $\mathcal{M}^0$. Column 9 and 10 shows the results obtained when performing suboptimal fusion. $\circledD$ stands for the DSm rule on the free DSm model and blank fields stand for 0.

Table 10.5 presents the results drawn from the minC combination rule. m^0 corresponds to the gbba on the generalized frame of discernment, m^1 to the gbba after reallocation of bbms of partial conflicts, $m^{a)}$ to the bba after proportionalization a) and $m^{b)}$ to the bba after proportionalization b). $m^0_{12^b3}$ denotes $(m_{12}^{b)} \circledm m_3)^0$, and m_{12^b3} denotes $(m_{12}^{b)} \circledm m_3)^{b)}$, where $\circledm$ stands for the generalized minC combination, blank fields stand for 0.

Table 10.6 presents the results of several rules of combination for 3D belief functions for sources 1 and 2 on Shafer's model, i.e. on the hybrid DSm model $\mathcal{M}^0$ (for the source bbas m_1, m_2, and m_3 see Table 10.4). $m^{a)}$ corresponds to the bba of the minC combination (the minC combination of m_1 and m_2 or m_1, m_2 and m_3 respectively) with proportionalization a); $m^{b)}$ corresponds to the bba of the minC combination with proportionalization b); $m^{\mathcal{M}^0}$ corresponds to the bba of the DSm combination. m^{TBM} corresponds to the bba of the combination with the TBM's non-normalized Demspter's rule; m^Y corresponds to the bba of the Yager's combination; m^{DB} corresponds to the bba of Dubois-Prade's combination and $m^{\oplus}$ corresponds to the bba of the normalized Dempster's combination.

	m_1	m_2	m_3	$m_{12}^{\mathcal{M}^f}$	$m_{12}^{\mathcal{M}^0}$	$m_{123}^{\mathcal{M}^f}$	$m_{123}^{\mathcal{M}^0}$	$(m_{12}^{\mathcal{M}^0} \textcircled{D} m_3)^{\mathcal{M}^f}$	$(m_{12}^{\mathcal{M}^0} \textcircled{D} m_3)^{\mathcal{M}^0}$
$\alpha_9 \sim \{\theta_1\}$	0.3	0.1	0.2	0.19	0.20	0.165	0.188	0.216	0.258
$\alpha_{10} \sim \{\theta_2\}$	0.2	0.1	0.1	0.15	0.17	0.090	0.109	0.119	0.145
$\alpha_{11} \sim \{\theta_3\}$	0.1	0.2	0.1	0.14	0.16	0.088	0.110	0.119	0.150
$\alpha_{15} \sim \{\theta_1,\theta_2\}$	0.1	0.0	0.2	0.03	0.08	0.021	0.056	0.058	0.112
$\alpha_{16} \sim \{\theta_1,\theta_3\}$	0.1	0.1	0.2	0.06	0.13	0.030	0.082	0.073	0.125
$\alpha_{17} \sim \{\theta_2,\theta_3\}$	0.0	0.2	0.1	0.04	0.09	0.014	0.039	0.035	0.068
$\alpha_{18} \sim \{\theta_1,\theta_2,\theta_3\}$	0.2	0.3	0.1	0.06	0.17	0.006	0.416	0.017	0.142
$\alpha_2 \sim \{\theta_1\}\times\{\theta_2\}$				0.05		0.106		0.054	
$\alpha_3 \sim \{\theta_1\}\times\{\theta_3\}$				0.07		0.120		0.052	
$\alpha_4 \sim \{\theta_2\}\times\{\theta_3\}$				0.05		0.074		0.033	
$\alpha_7 \sim \{\theta_1\}\times\{\theta_2,\theta_3\}$				0.06		0.083		0.038	
$\alpha_6 \sim \{\theta_2\}\times\{\theta_1,\theta_3\}$				0.03		0.060		0.047	
$\alpha_5 \sim \{\theta_3\}\times\{\theta_1,\theta_2\}$				0.02		0.048		0.040	
$\alpha_1 \sim \times$						0.022			
$\alpha_8 \sim \Box$						0.009			
$\alpha_{14} \sim \Box\theta_1$				0.01		0.023		0.042	
$\alpha_{13} \sim \Box\theta_2$				0.02		0.019		0.026	
$\alpha_{12} \sim \Box\theta_3$				0.02		0.022		0.031	

Table 10.4: Comparison of combination of 3D belief functions based on DSm rules of combination.

We can see that during the combination of 2 belief functions a lot of types of conflict arise, but some of them still remain with 0 bbm ($\alpha_1 \sim \times$ and $\alpha_8 \sim \Box$). We can see how these conflicts arise when the 3rd BF is combined. We can see the difference between the combination of 3 BFs on the generalized level (see m_{123}^0) and the suboptimal combination of the 3rd belief with an intermediate result to which constraints have already been introduced (see $(m_{12}^{DSm} \textcircled{D} m_3)^0$ and $(m_{12}^{b)} \textcircled{m} m_3)^0$). We can see how the gbbms are reallocated among the subsets of Θ during the second step of minC combination and finally how the gbbms of all pure conflicts are reallocated in both ways a) and b).

The final results of DSm and minC combinations are compared in Table 10.6. We can note that the small subsets of Θ (singletons in our 3D example) have greater bbms after the minC combination while the great sets (2-element sets and namely whole $\{\theta_1, \theta_2, \theta_3\}$ in our case) have greater bbms after application of the DSm combination rule. I. e. the DSm combining rule is more cautious than the minC combination within the reallocation of the conflicting gbbms. Thus we see that the minC combination

	m^0_{12}	m^1_{12}	$m^{a)}_{12}$	$m^{b)}_{12}$	m^0_{123}	m^1_{123}	$m^{a)}_{123}$	$m^{b)}_{123}$	$m^0_{12^b3}$	m_{12^b3}
$\alpha_9 \sim \{\theta_1\}$	0.19	0.20	0.2983	0.2889	0.165	0.165	0.4031	0.4068	0.2396	0.4113
$\alpha_{10} \sim \{\theta_2\}$	0.15	0.17	0.2318	0.2402	0.090	0.090	0.2301	0.2306	0.1360	0.2319
$\alpha_{11} \sim \{\theta_3\}$	0.14	0.16	0.2311	0.2327	0.088	0.088	0.2288	0.2363	0.1364	0.2372
$\alpha_{15} \sim \{\theta_1,\theta_2\}$	0.03	0.03	0.0362	0.0383	0.021	0.021	0.0390	0.0377	0.0253	0.0354
$\alpha_{16} \sim \{\theta_1,\theta_3\}$	0.06	0.06	0.0762	0.0792	0.030	0.030	0.0586	0.0549	0.0376	0.0522
$\alpha_{17} \sim \{\theta_2,\theta_3\}$	0.04	0.04	0.0534	0.0515	0.014	0.014	0.0264	0.0249	0.0172	0.0236
$\alpha_{18} \sim \{\theta_1,\theta_2,\theta_3\}$	0.06	0.06	0.0830	0.0692	0.006	0.006	0.0140	0.0088	0.0069	0.0084
$\alpha_2 \sim \{\theta_1\}\times\{\theta_2\}$	0.05	0.05			0.106	0.106			0.0769	
$\alpha_3 \sim \{\theta_1\}\times\{\theta_3\}$	0.07	0.07			0.120	0.120			0.0754	
$\alpha_4 \sim \{\theta_2\}\times\{\theta_3\}$	0.05	0.05			0.074	0.074			0.0473	
$\alpha_7 \sim \{\theta_1\}\times\{\theta_2,\theta_3\}$	0.06	0.06			0.083	0.083			0.0392	
$\alpha_6 \sim \{\theta_2\}\times\{\theta_1,\theta_3\}$	0.03	0.03			0.060	0.060			0.0560	
$\alpha_5 \sim \{\theta_3\}\times\{\theta_1,\theta_2\}$	0.02	0.02			0.048	0.048			0.0504	
$\alpha_1 \sim \times$					0.022	0.022				
$\alpha_8 \sim \square$					0.009	0.009				
$\alpha_{14} \sim \square\theta_1$	0.01				0.023				0.0235	
$\alpha_{13} \sim \square\theta_2$	0.02				0.019				0.0141	
$\alpha_{12} \sim \square\theta_3$	0.02				0.022				0.0182	

Table 10.5: Comparison of combination of 3D belief functions with the minC rule.

rule produces more specified results than the DSm rule does. The last three columns of the table show us that the DSm and the minC with both the proportionalizations produce results different from those of Yager's rule and of both the versions of Dempster's rule (see m^Y, m^{TMB}, and $m^{\oplus}$ respectively). While binary DSm result on Shafer's model ($\mathcal{M}^0$) coincides with the results of Dubois-Prade's rule of combination.

Let us present numeric examples of parts of computation m^0, m^1, $m^{a)}$, and $m^{b)}$ for readers which are interested in detail. We begin with a non-conflicting set $\{\theta_1,\theta_2\}$, i.e. with $\alpha_{15} = \theta_1 \cup \theta_2$ in the DSm notation. It is an intersection with itself or with the whole $\Theta = \{\theta_1,\theta_2,\theta_3\}$ (i.e. $\theta_1\cup\theta_2\cup\theta_3$ in DSm), and it is not $\sim$ equivalent to any other element of D^Θ. Thus $m^0_{12}(\theta_1\cup\theta_2) = m_1(\theta_1\cup\theta_2)m_2(\theta_1\cup\theta_2) + m_1(\theta_1\cup\theta_2)m_2(\theta_1\cup\theta_2\cup\theta_3) + m_1(\theta_1\cup\theta_2\cup\theta_3)m_2(\theta_1\cup\theta_2) = 0.1\cdot 0.0 + 0.1\cdot 0.3 + 0.0\cdot 0.2 = 0.00 + 0.03 + 0.00 = 0.03$. α_{15} is a non-conflicting element of D^Θ, hence it is not further reassigned or proportionalized, i. e. its bbm will not be decreased. α_{15} is not a non-conflicting part of any other element of D^Θ, thus $m^1_{12}(\alpha_{15}) = m^0_{12}(\alpha_{15})$.

	$m_{12}^{\mathcal{M}^0}$	$m_{12}^{a)}$	$m_{12}^{b)}$	$m_{123}^{\mathcal{M}^0}$	$m_{123}^{a)}$	$m_{123}^{b)}$	m_{12}^{TBM}	m_{12}^{Y}	m_{12}^{DP}	$m_{12}^{\oplus}$
$\alpha_9 \sim \{\theta_1\}$	0.20	0.2983	0.2889	0.188	0.4031	0.4068	0.20	0.20	0.20	0.2778
$\alpha_{10} \sim \{\theta_2\}$	0.17	0.2318	0.2402	0.109	0.2301	0.2306	0.17	0.17	0.17	0.2361
$\alpha_{11} \sim \{\theta_3\}$	0.16	0.2311	0.2327	0.110	0.2288	0.2363	0.16	0.16	0.16	0.2222
$\alpha_{15} \sim \{\theta_1,\theta_2\}$	0.08	0.0362	0.0383	0.056	0.0390	0.0377	0.04	0.04	0.08	0.0556
$\alpha_{16} \sim \{\theta_1,\theta_3\}$	0.13	0.0762	0.0792	0.082	0.0586	0.0549	0.06	0.06	0.13	0.0833
$\alpha_{17} \sim \{\theta_2,\theta_3\}$	0.09	0.0534	0.0515	0.039	0.0264	0.0249	0.03	0.03	0.09	0.0417
$\alpha_{18} \sim \{\theta_1,\theta_2,\theta_3\}$	0.17	0.0830	0.6992	0.416	0.0140	0.0088	0.06	0.34	0.17	0.0833
$\emptyset$							0.28			

Table 10.6: Comparison of combinations of sources 1 and 2 on Shafer's model (i.e. on the hybrid DSm model $\mathcal{M}^0$).

$m_{12}^{a)}(\alpha_{15}) > m_{12}^{1}(\alpha_{15})$ because gbbms of some other elements are proportionalized, among others, also to α_{15}. For the same reason it holds also $m_{12}^{b)}(\alpha_{15}) > m_{12}^{1}(\alpha_{15})$.

A potential conflict $\Box\{\theta_1\} \sim (\theta_1\cup\theta_2)\cap(\theta_1\cup\theta_3) = \alpha_{14}$ is equivalent to $\Box\{\theta_1\}\times\Box\{\theta_1\}$, to $\Box\{\theta_1\}\times X$, and to $X\times\Box\{\theta_1\}$, where $\{\theta_1\}\subset X$ in Shafer's model, see Table 10.1; or $\alpha_{14} = (\theta_1\cup\theta_2)\cap(\theta_1\cup\theta_3)$ is an intersection of itself with X, where $\alpha_{14}\subseteq X\subseteq \theta_1\cup\theta_2\cup\theta_3$ in the DSm terminology. I.e. $m_{12}^0(\alpha_{14}) = m^0(\theta_1\cap(\theta_2\cup\theta_3)) = m_1(\alpha_{14})m_2(\alpha_{14}) + m_1(\theta_1\cup\theta_2)m_2(\theta_1\cup\theta_3) + m_1(\theta_1\cup\theta_3)m_2(\theta_1\cup\theta_2) + m_1(\alpha_{14})(m_2(\theta_1\cup\theta_2)+m_2(\theta_1\cup\theta_3)+m_2(\theta_1\cup\theta_2\cup\theta_3))+(m_1(\theta_1\cup\theta_2)+m_1(\theta_1\cup\theta_3)+m_1(\theta_1\cup\theta_2\cup\theta_3))m_2(\alpha_{14}) = 0.0\cdot0.0+0.1\cdot0.1+0.1\cdot0.0+0.0\cdot(0.1+0.1+0.2)+(0.0+0.1+0.3)\cdot0.0 = 0+0.01+0+0+0 = 0.01$. $\alpha_9 = \{\theta_1\}$ is a non-conflicting part of $\theta_1\cap(\theta_2\cup\theta_3)$, thus $m^0(\alpha_{14})$ is reallocated to θ_1. On the other hand $\{\theta_1\}$ is not a non-conflicting part of any other element of D^Θ, hence $m^1(\alpha_9) = m^0(\alpha_9)+m^0(\alpha_{14}) = 0.19+0.01 = 0.20$. After this reallocation, the bbm of α_{14} equals 0, hence $m^1(\alpha_{14}) = m^{a)}(\alpha_{14}) = m^{b)}(\alpha_{14}) = 0$.

A pure conflict $\{\theta_1\}\times\{\theta_2,\theta_3\}\sim\theta_1\cap(\theta_2\cup\theta_3) = \alpha_7$ is contained in 24 fields of the full minC combination table (for its part see Table 10.1), e. g. in the fields corresponding to $\{A\}\times(\{A\}\times\{B,C\})$, $\{A\}\times\{B,C\}$, $\{A,B\}\times(\{A\}\times\{B,C\})$, but only some of them correspond to the Shaferian input beliefs (i. e. only some of them are positive). Thus $m^1(\alpha_7) = m^0(\alpha_7) = m_1(\theta_1)m_2(\theta_2\cup\theta_3) + m_1(\theta_2\cup\theta_3)m_2(\theta_1) = 0.3\cdot0.2+0.1\cdot0.0 = 0.06+0.00 = 0.06$. As α_7 is a pure conflict, thus its bbm is not changing during the reallocation substep, and it is proportionalized among $\{\theta_1\},\{\theta_2,\theta_3\},\{\theta_1,\theta_2,\theta_3\}$ with the proportionalization a), and among all the subsets of $\Theta = \{\theta_1,\theta_2,\theta_3\}$ with the proportionalization b). Thus $m^1(\alpha_7)\cdot\frac{m^1(\theta_1)}{m^1(\theta_1)+m^1(\theta_2\cup\theta_3)+m^1(\theta_1\cup\theta_2\cup\theta_3)} = 0.06\frac{0.20}{0.20+0.04+0.06} = 0.06\frac{0.20}{0.30} = 0.040$ is reassigned to $\theta_1 = \alpha_9$; $m^1(\alpha_7)\cdot\frac{m^1(\theta_2\cup\theta_3)}{m^1(\theta_1)+m^1(\theta_2\cup\theta_3)+m^1(\theta_1\cup\theta_2\cup\theta_3)} = 0.06\frac{0.04}{0.20+0.04+0.06} = 0.06\frac{0.04}{0.30} = 0.008$ is reassigned to $\theta_2 \cup \theta_3 = \alpha_{17}$; and $m^1(\alpha_7) \cdot \frac{m^1(\theta_1\cup\theta_2\cup\theta_3)}{m^1(\theta_1)+m^1(\theta_2\cup\theta_3)+m^1(\theta_1\cup\theta_2\cup\theta_3)} = 0.06\frac{0.06}{0.20+0.04+0.06} = 0.06\frac{0.06}{0.30} = 0.012$ is reassigned to $\theta_1 \cup \theta_2 \cup \theta_3 = \alpha_{18}$ with the proportionalization a). As belief masses $0.05\frac{0.20}{0.20+0.17+0.03} = 0.05 \cdot 0.5 = 0.0250$ and $0.07\frac{0.20}{0.20+0.16+0.06} = 0.07 \cdot 0.4762 = 0.0333$ are analogically proportionalized with the proportionalization a) also to θ_1, so we obtain $m^{a)}_{12}(\theta_1) = m^1(\theta_1) + 0.040 + 0.0250 + 0.0333 = 0.2000 + 0.040 + 0.0250 + 0.0333 = 0.2983$. A value $m^{b)}_{12}(\theta_1)$ is computed analogically; where e.g. $0.06\frac{0.20}{0.20+0.17+0.16+0.03+0.06+0.04+0.06} = 0.06\frac{0.20}{0.72} = 0.06 \cdot 0.2777 = 0.0166$ is proportionalized from $m^1(\alpha_7)$.

10.6 Conclusion

In this chapter we have compared two independently developed approaches to combination of conflicting beliefs. Motivations and the starting points of the approaches are significantly different. The classical frame of discernment with mutually exclusive elements is the starting point for the minC combination, whereas the free DSm model is the starting point for the classical DSm approach. The approaches were originally rather complementary than comparable.

Surprisingly, the internal combining structures and mechanisms of both these combination rules are the same and the results of the classical DSm rule for the free DSm model are the same as the intermediate results of the minC combination on a generalized frame of discernment. Nevertheless, this common step is followed by reallocation of the belief masses temporarily assigned to conflicts to obtain classical belief functions as results in the case of the minC combination.

After the recent development of versions of the DSm rule for Shafer's model and for general hybrid DSm models, which consider 2 steps of combination, the minC combination becomes an alternative to the special case of the DSm combination rule for Shafer's model.

The first step — a combination on a generalized frame — is the same again. Also a reallocation of the generalized basic belief masses of potential conflicts is analogous. The main difference consists in different reallocations of the generalized basic belief masses (gbbm) of pure conflicts: it is a reassigning of the gbbms to the union of the corresponding sets in the DSm rule, whereas a proportionalization in the minC approach.

In spite of this difference, we can also consider the DSm introduction of constraints as an alternative to a reallocation of the belief masses of conflicts in the minC approach.

10.7 References

[1] Daniel M., *Distribution of Contradictive Belief Masses in Combination of Belief Functions*, Information, Uncertainty and Fusion, Eds. Bouchon-Meunier B., Yager R. R., and Zadeh L. A., Kluwer Academic Publishers, 2000, pp. 431-446.

[2] Daniel M., *Associativity in Combination of Belief Functions*, Proceedings of 5th Workshop on Uncertainty Processing (WUPES 2000), Ed. Vejnarová, J., Ediční oddělení VŠE (Editorial Department of University of Economics), 2000, pp. 41–54.

[3] Daniel M., *Associativity and Contradiction in Combination of Belief Functions*, Proceedings Eighth International conference IPMU, Universidad Politecnica de Madrid, 2000, Vol. I., pp. 133–140.

[4] Daniel M., *Associativity in Combination of belief functions; a derivation of minC combination*, Soft Computing, 7(5), pp. 288–296, 2003.

[5] Dezert J., *Combination of Paradoxical Sources of Information within the Neutrosophic Framework*, Proceedings of the 1st Int. Conf. Neutrosophics, Univ. of New Mexico, Gallup Campus, 2001.

[6] Dezert J., *Foundations for a New Theory of Plausible and Paradoxical Reasoning*, Information and Security Journal, Edited by Tz. Semerdjiev, ProCon Ltd., Sofia, Bulgaria, Vol. 9, pp. 13-57, 2002.

[7] Dubois D., Prade H., *Representation and combination of uncertainty with belief functions and possibility measures*, Computational Intelligence, Vol. 4, pp. 244-264, 1988.

[8] Jøsang A., Daniel M., Vannoorenberghe P., *Strategies for Combining Conflicting Dogmatic Beliefs*, Proceedings of Fusion 2003 International Conference, pp. 1133–1140, Cairns, Australia, July 2003.

[9] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[10] Smets Ph., *The combination of evidence in the transferable belief model*, IEEE-Pattern analysis and Machine Intelligence, Vol. 12, pp. 447–458, 1990.

[11] Smets Ph., Kennes R., *The transferable belief model*, Artificial Intelligence, 66(2), pp. 191–234, 1994.

[12] Yager R.R., *On the Dempster-Shafer framework and new combination rules*, Information Sciences, Vol. 41, pp. 93–138, 1987.

Chapter 11

General Fusion Operators from Cox's Postulates

Denis de Brucq

Académie des Sciences, Belles-Lettres et Arts de Rouen

Rouen, France.

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:

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

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 $[A\ and\ B/e] = [A/e] \circ [B/e]$, 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:

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

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

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

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

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:*

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

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*: $\overline{a} \longleftrightarrow \forall \theta \in \Theta, \quad \mu_{\overline{a}}(\theta) = (1 - \mu_a(\theta))$

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:

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

and then:

$$Pos(A) + Pos(\overline{A}) \geq 1. \tag{11.20}$$

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

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

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.

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

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:

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

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

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

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* $[0,1]$

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

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

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

Theorem 1. *The fuzzy operator* $[A \wedge B/e] = [A/e] \wedge [B/e] = \inf\left\{[A/e], [B/e]\right\}$ *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 :

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

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

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

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

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

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

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

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 : $\forall x\in[b,c],\quad \forall y\in[x,1],\qquad x\circ y = x\wedge y$.

   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 $\forall s\in(a,b)\quad s\circ s\neq a$ 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* $u \stackrel{\triangle}{=} \sup\{v \in [a, b]\,; v \circ v = a\}$. *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

$$P \quad \textit{and} \quad Q \Rightarrow R$$

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.

Part II

Applications of DSmT

Chapter 12

On the Tweety Penguin Triangle Problem

Jean Dezert
ONERA
29 Av. de la Division Leclerc
92320 Châtillon
France

Florentin Smarandache
Department of Mathematics
University of New Mexico
Gallup, NM 8730
U.S.A.

Abstract: *In this chapter, one studies the famous well-known and challenging Tweety Penguin Triangle Problem (TPTP or TP2) pointed out by Judea Pearl in one of his books. We first present the solution of the TP2 based on the fallacious Bayesian reasoning and prove that reasoning cannot be used to conclude on the ability of the penguin-bird Tweety to fly or not to fly. Then we present in details the counter-intuitive solution obtained from the Dempster-Shafer Theory (DST). Finally, we show how the solution can be obtained with our new theory of plausible and paradoxical reasoning (DSmT).*

12.1 Introduction

Judea Pearl claimed that DST of evidence fails to provide a reasonable solution for the combination of evidence even for apparently very simple fusion problem [11, 12]. Most criticisms are answered by Philippe Smets in [22, 23]. The Tweety Penguin Triangle Problem (TP2) is one of the typical exciting and challenging problem for all theories managing uncertainty and conflict because it shows the real difficulty to maintain truth for automatic reasoning systems when the classical property of transitivity (which is basic to the material-implication) does not hold. In his book, Judea Pearl presents and discusses in details the semantic clash between Bayes vs. Dempster-Shafer reasoning. We present here our analysis on this problem and provide a new solution of the Tweety Penguin Triangle Problem based on our new theory of plausible and paradoxical reasoning, known as DSmT (Dezert-Smarandache Theory). We show how this problem can be attacked and solved by our new reasoning with help of the (hybrid) DSm rule of combination (see chapter 4). The purpose of this chapter is not to browse all approaches available in literature for attacking the TP2 problem but only to provide a comparison of the DSm reasoning with respect to the Bayesian reasoning and to the plausible reasoning of DST framework. Interesting but complex analysis on this problem based on default reasoning and ϵ-belief functions can be also found by example in [22] and [1]. Other interesting and promising issues for the TP2 problem based on the fuzzy logic of Zadeh [25] jointly with the theory of possibilities [5, 6] are under investigations. Some theoretical research works on new conditional event algebras (CEA) have emerged in literature [7] since last years and could offer a new track for attacking the TP2 problem although unfortunately no clear didactic, simple and convincing examples are provided to show the real efficiency and usefulness of these theoretical investigations.

12.2 The Tweety Penguin Triangle Problem

This very important and challenging problem, as known as the Tweety Penguin Triangle Problem (TP2) in literature, is presented in details by Judea Pearl in [11]. We briefly present here the TP2 and the solutions based first on fallacious Bayesian reasoning and then on the Dempster-Shafer reasoning. We will then focus our analysis of this problem from the DSmT framework and the DSm reasoning.

Let's consider the set $R = \{r_1, r_2, r_3\}$ of given rules (as known as defaults in [1]):

- r_1: "**Penguins normally don't fly**" $\Leftrightarrow (p \rightarrow \neg f)$
- r_2: "**Birds normally fly**" $\Leftrightarrow (b \rightarrow f)$
- r_3: "**Penguins are birds**" $\Leftrightarrow (p \rightarrow b)$

To emphasize our strong conviction in these rules we commit them some high confidence weights w_1, w_2 and w_3 in $[0, 1]$ with $w_1 = 1 - \epsilon_1$, $w_2 = 1 - \epsilon_2$ and $w_3 = 1$ (where ϵ_1 and ϵ_2 are small positive quantities). The conviction in these rules is then represented by the set $W = \{w_1, w_2, w_3\}$ in the sequel.

Another useful and general notation adopted by Judea Pearl in the first pages of his book [11] to characterize these three weighted rules is the following one (where $w_1, w_2, w_3 \in [0, 1]$):

$$r_1 : p \overset{w_1}{\rightarrow} (\neg f) \qquad r_2 : b \overset{w_2}{\rightarrow} f \qquad r_3 : p \overset{w_3}{\rightarrow} b$$

When $w_1, w_2, w_3 \in \{0, 1\}$ the classical logic is the perfect tool to conclude on the truth or on the falsity of a proposition built from these rules based on the standard propositional calculus mainly with its three fundamental rules (Modus Ponens, Modus Tollens and Modus Barbara - i.e. transitivity rule). When $0 < w_1, w_2, w_3 < 1$, the classical logic can't be applied because the Modus Ponens, the Modus Tollens and the Modus Barbara do not longer hold and some other tools must be chosen. This will discussed in detail in section 3.2.

Question: Assume we observe an animal called Tweety (T) that is categorically classified as a bird (b) and a penguin (p), i.e. our observation is $O \triangleq [T = (b \cap p)] = [(T = b) \cap (T = p)]$. The notation $T = (b \cap p)$ stands here for "Entity T holds property $(b \cap p)$". What is the belief (or the probability - if such probability exists) that Tweety can fly given the observation O and all information available in our knowledge base (i.e. our rule-based system R and W) ?

The difficulty of this problem for most of artificial reasoning systems (ARS) comes from the fact that, in this example, the property of transitivity, usually supposed satisfied from material-implication interpretation [11], $(p \to b, b \to f) \Rightarrow (p \to f)$ does not hold here (see section 12.3.2). In this interesting example, the classical property of inheritance is thus broken. Nevertheless a powerful artificial reasoning system must be able to deal with such kind of difficult problem and must provide a reliable conclusion by a general mechanism of reasoning whatever the values of convictions are (not only restricted to values close to either 0 or 1). We examine now three ARS based on the Bayesian reasoning [11] which turns to be fallacious and actually not appropriate for this problem and we explain why, on the Dempster-Shafer Theory (DST) [16] and on the Dezert-Smarandache Theory (DSmT) (see part I of this book).

12.3 The fallacious Bayesian reasoning

We first present the fallacious Bayesian reasoning solution drawn from the J. Pearl's book in [11] (pages 447-449) and then we explain why the solution which seems at the first glance correct with intuition is really fallacious. We then explain why the common rational intuition turns actually to be wrong and show the weakness of Pearl's analysis.

12.3.1 The Pearl's analysis

To preserve mathematical rigor, we introduce explicitly all information available in the derivations. In other words, one wants to evaluate using the Bayesian reasoning, the conditional probability, if it exists, $P(T = f|O, R, W) = P(T = f|T = p, T = b, R, W)$. The Pearl's analysis is based on the assumption that a conviction on a given rule can be interpreted as a conditional probability (see [11] page 4). In other words if one has a given rule $a \xrightarrow{w} b$ with $w \in [0, 1]$ then one can interpret, at least for the calculus, w as $P(b|a)$ and thus the probability theory and Bayesian reasoning can help to answer to the question. We prove in the following section that such model cannot be reasonably adopted. For now, we just assume that such probabilistic model holds effectively as Judea Pearl does. Based on this assumption, since the conditional term/information $(T = p, T = b, R, W)$ is strictly equivalent to $(T = p, R, W)$ because of the knowledge of rule r_3 with certainty (since $w_3 = 1$), one gets easily the fallacious intuitive expected Pearl's result:

$$
\begin{aligned}
P(T = f|O, R, W) &= P(T = f|T = p, T = b, R, W) \\
P(T = f|O, R, W) &\equiv P(T = f|T = p, R, W) \\
P(T = f|O, R, W) &= 1 - P(T = \neg f|T = p, R, W) \\
P(T = f|O, R, W) &= 1 - w_1 = \epsilon_1
\end{aligned}
$$

From this simple analysis, the Tweety's "birdness" does not render her a better flyer than an ordinary penguin as intuitively expected and the probability that Tweety can fly remains very low which looks normal. We reemphasize here the fact, that in his Bayesian reasoning J. Pearl assumes that the weight w_1 for the conviction in rule r_1 can be interpreted in term of a real probability measure $P(\neg f|p)$. This assumption is necessary to provide the rigorous derivation of $P(T = f|O, R, W)$. It turns out however that convictions w_i on logical rules cannot be interpreted in terms of probabilities as we will prove in the next section.

When rule r_3 is not asserted with absolute certainty (i.e. $w_3 = 1$) but is subject to exceptions, i.e. $w_3 = 1 - \epsilon_3 < 1$, the fallacious Bayesian reasoning yields (where notations $T = f$, $T = b$ and $T = p$ are replaced by f, b and p due to space limitations):

$$
\begin{aligned}
P(f|O, R, W) &= P(f|p, b, R, W) \\
P(f|O, R, W) &= \frac{P(f, p, b|R, W)}{P(p, b|R, W)} \\
P(f|O, R, W) &= \frac{P(f, b|p, R, W)P(p|R, W)}{P(b|p, R, W)P(p|R, W)}
\end{aligned}
$$

By assuming $P(p|R, W) > 0$, one gets after simplification by $P(p|R, W)$

$$
\begin{aligned}
P(f|O, R, W) &= \frac{P(f, b|p, R, W)}{P(b|p, R, W)} \\
P(f|O, R, W) &= \frac{P(b|f, p, R, W)P(f|p, R, W)}{P(b|p, R, W)}
\end{aligned}
$$

If one assumes $P(b|p, R, W) = w_3 = 1 - \epsilon_3$ and $P(f|p, R, W) = 1 - P(\neg f|p, R, W) = 1 - w_1 = \epsilon_1$, one gets

$$
P(f|O, R, W) = P(b|f, p, R, W) \times \frac{\epsilon_1}{1 - \epsilon_3}
$$

Because $0 \leq P(b|f, p, R, W) \leq 1$, one finally gets the Pearl's result [11] (p.448)

$$P(f|O, R, W) \leq \frac{\epsilon_1}{1-\epsilon_3} \tag{12.1}$$

which states that the observed animal Tweety (a penguin-bird) has a very small probability of flying as long as ϵ_3 remains small, regardless of how many birds cannot fly (ϵ_2), and has consequently a high probability of not flying because $P(f|O, R, W) + P(\bar{f}|O, R, W) = 1$ since the events f and $\bar{f}$ are mutually exclusive and exhaustive (assuming that the Pearl's probabilistic model holds ...).

12.3.2 The weakness of the Pearl's analysis

We prove now that the previous Bayesian reasoning is really fallacious and the problem is truly undecidable to conclude about the ability of Tweety to fly or not to fly if a deep analysis is done. Actually, the Bayes' inference is not a classical inference (see chapter 8 for justification). Indeed, before applying blindly the Bayesian reasoning as in the previous section, one first has to check that the probabilistic model is well-founded to characterize the convictions of the rules of the rule-based system under analysis. We prove here that such probabilistic model doesn't hold for a suitable and useful representation of the problem and consequently for any problems based on the weighting of logical rules (with positive weighting factors/convictions below than 1).

12.3.2.1 Preliminaries

We just remind here only few important principles of the propositional calculus of the classical Mathematical Logic which will be used in our demonstration. A simple notation, which may appear as unusual for logicians, is adopted here just for convenience. A detailed presentation of the propositional calculus and Mathematical Logic can be easily found in many standard mathematical textbooks like [15, 10, 9]. Here are these important principles:

- **Third middle excluded principle** : A logical variable is either true or false, i.e.

$$a \vee \neg a \tag{12.2}$$

- **Non-contradiction law** : A logical variable can't be both true and false, i.e.

$$\neg(a \wedge \neg a) \tag{12.3}$$

- **Modus Ponens** : This rule of the propositional calculus states that if a logical variable a is true and $a \rightarrow b$ is true, then b is true (syllogism principle), i.e.

$$(a \wedge (a \rightarrow b)) \rightarrow b \tag{12.4}$$

- **Modus Tollens** : This rule of the propositional calculus states that if a logical variable $\neg b$ is true and $a \rightarrow b$ is true, then $\neg a$ is true, i.e.

$$(\neg b \wedge (a \rightarrow b)) \rightarrow \neg a \tag{12.5}$$

- **Modus Barbara** : This rule of the propositional calculus states that if $a \rightarrow b$ is true and $b \rightarrow c$ is true then $a \rightarrow c$ is true (transitivity property), i.e.

$$((a \rightarrow b) \wedge (b \rightarrow c)) \rightarrow (a \rightarrow c) \tag{12.6}$$

From these principles, one can prove easily, based on the truth table method, the following property (more general deducibility theorems in Mathematical Logic can be found in [18, 19]) :

$$((a \rightarrow b) \wedge (c \rightarrow d)) \rightarrow ((a \wedge c) \rightarrow (b \wedge d)) \tag{12.7}$$

12.3.2.2 Analysis of the problem when $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0$

We first examine the TP2 when one has no doubt in the rules of our given rule-based systems, i.e.

$$\begin{cases} r_1 : p \overset{w_1=1-\epsilon_1=1}{\rightarrow} (\neg f) \\ r_2 : b \overset{w_2=1-\epsilon_2=1}{\rightarrow} f \\ r_3 : p \overset{w_3=1-\epsilon_3=1}{\rightarrow} b \end{cases}$$

From rules r_1 and r_2 and because of property (12.7), one concludes that

$$p \wedge b \rightarrow (f \wedge \neg f)$$

and using the non-contradiction law (12.3) with the Modus Tollens (12.5), one finally gets

$$\neg(f \wedge \neg f) \rightarrow \neg(p \wedge b)$$

which proves that $p \wedge b$ is always false whatever the rule r_3 is. Interpreted in terms of the probability theory, the event $T = p \cap b$ corresponds actually and truly to the impossible event $\emptyset$ since $T = f$ and $T = \bar{f}$ are exclusive and exhaustive events. Under such conditions, the analysis proves the non-existence of the penguin-bird Tweety.

If one adopts the notations[1] of the probability theory, trying to derive $P(T = f|T = p \cap b)$ and $P(T = \bar{f}|T = p \cap b)$ with the Bayesian reasoning is just impossible because from one of the axioms of the probability theory, one must have $P(\emptyset) = 0$ and from the conditioning rule, one would get expressly for this problem the indeterminate expressions:

[1] Because probabilities are related to sets, we use here the common set-complement notation $\bar{f}$ instead of the logical negation notation $\neg f$, $\cap$ for $\wedge$ and $\cup$ for $\vee$ if necessary.

$$\begin{aligned}
P(T=f|T=p\cap b) &= P(T=f|T=\emptyset)\\
P(T=f|T=p\cap b) &= \frac{P(T=f\cap\emptyset)}{P(T=\emptyset)}\\
P(T=f|T=p\cap b) &= \frac{P(T=\emptyset)}{P(T=\emptyset)}\\
P(T=f|T=p\cap b) &= \frac{0}{0} \qquad\qquad \text{(indeterminate)}
\end{aligned}$$

and similarly

$$\begin{aligned}
P(T=\bar{f}|T=p\cap b) &= P(T=\bar{f}|T=\emptyset)\\
P(T=\bar{f}|T=p\cap b) &= \frac{P(T=\bar{f}\cap\emptyset)}{P(T=\emptyset)}\\
P(T=\bar{f}|T=p\cap b) &= \frac{P(T=\emptyset)}{P(T=\emptyset)}\\
P(T=\bar{f}|T=p\cap b) &= \frac{0}{0} \qquad\qquad \text{(indeterminate)}
\end{aligned}$$

12.3.2.3 Analysis of the problem when $0 < \epsilon_1, \epsilon_2, \epsilon_3 < 1$

Let's examine now the general case when one allows some little doubt on the rules characterized by taking $\epsilon_1 \gtrsim 0$, $\epsilon_2 \gtrsim 0$ and $\epsilon_3 \gtrsim 0$ and examine the consequences on the probabilistic model on these rules.

First note that, because of the third middle excluded principle and the assumption of the existence of a probabilistic model for a weighted rule, then one should be able to consider simultaneously both "probabilistic/Bayesian" rules

$$\begin{cases} a \overset{P(b|a)=w}{\longrightarrow} b \\ a \overset{P(\bar{b}|a)=1-w}{\longrightarrow} \neg b \end{cases} \tag{12.8}$$

In terms of classical (objective) probability theory, these weighted rules just indicate that in $100\times w$ percent of cases the logical variable b is true if a is true, or equivalently, that in $100\times w$ percent of cases the random event b occurs when the random event a occurs. When we don't refer to classical probability theory, the weighting factors w and $1-w$ indicate just the level of conviction committed to the validity of the rules. Although very appealing at the first glance, this probabilistic model hides actually a strong drawback/weakness especially when dealing with several rules as shown right below.

Let's prove first that from a "probabilized" rule $a \overset{P(b|a)=w}{\longrightarrow} b$ one cannot assess rigorously the convictions onto its Modus Tollens. In other words, from (12.8) what can we conclude on

$$\begin{cases} \neg b \overset{P(\bar{a}|\bar{b})=?}{\longrightarrow} \neg a \\ b \overset{P(\bar{a}|b)=?}{\longrightarrow} \neg a \end{cases} \tag{12.9}$$

From the Bayes' rule of conditioning (which must hold if the probabilitic model holds), one can express $P(\bar{a}|\bar{b})$ and $P(\bar{a}|b)$ as follows

$$\begin{cases} P(\bar{a}|\bar{b}) = 1 - P(a|\bar{b}) = 1 - \frac{P(a\cap\bar{b})}{1-P(b)} = 1 - \frac{P(\bar{b}|a)P(a)}{1-P(b)} \\ P(\bar{a}|b) = 1 - P(a|b) = 1 - \frac{P(a\cap b)}{P(b)} = 1 - \frac{P(b|a)P(a)}{P(b)} \end{cases}$$

or equivalently by replacing $P(b|a)$ and $P(\bar{b}|a)$ by their values w and $1-w$, one gets

$$\begin{cases} P(\bar{a}|\bar{b}) = 1 - (1-w)\frac{P(a)}{1-P(b)} \\ P(\bar{a}|b) = 1 - w\frac{P(a)}{P(b)} \end{cases} \tag{12.10}$$

These relationships show that one cannot fully derive in theory $P(\bar{a}|\bar{b})$ and $P(\bar{a}|b)$ because the prior probabilities $P(a)$ and $P(b)$ are unknown.

A simplistic solution, based on the principle of indifference, is then just to assume without solid justification that $P(a) = P(\bar{a}) = 1/2$ and $P(b) = P(\bar{b}) = 1/2$. With such assumption, then one gets the following estimates $\hat{P}(\bar{a}|\bar{b}) = w$ and $\hat{P}(\bar{a}|b) = 1-w$ for $P(\bar{a}|\bar{b})$ and $P(\bar{a}|b)$ respectively and we can go further in the derivations.

Now let's go back to our Tweety Penguin Triangle Problem. Based on the probabilistic model (assumed to hold), one starts now with both

$$\begin{cases} r_1 : p \overset{P(\bar{f}|p)=1-\epsilon_1}{\longrightarrow} \neg f \\ r_2 : b \overset{P(f|b)=1-\epsilon_2}{\longrightarrow} f \\ r_3 : p \overset{P(b|p)=1-\epsilon_3}{\longrightarrow} b \end{cases} \qquad \begin{cases} p \overset{P(f|p)=\epsilon_1}{\longrightarrow} f \\ b \overset{P(\bar{f}|b)=\epsilon_2}{\longrightarrow} \neg f \\ p \overset{P(\bar{b}|p)=\epsilon_3}{\longrightarrow} \neg b \end{cases} \tag{12.11}$$

Note that taking into account our preliminary analysis and *accepting the principle of indifference*, one has also the two sets of weighted rules either

$$\begin{cases} f \overset{\hat{P}(\bar{p}|f)=1-\epsilon_1}{\longrightarrow} \neg p \\ \neg f \overset{\hat{P}(\bar{b}|\bar{f})=1-\epsilon_2}{\longrightarrow} \neg b \\ \neg b \overset{\hat{P}(\bar{p}|\bar{b})=1-\epsilon_3}{\longrightarrow} \neg p \end{cases} \qquad \begin{cases} \neg f \overset{\hat{P}(\bar{p}|\bar{f})=\epsilon_1}{\longrightarrow} \neg p \\ f \overset{\hat{P}(\bar{b}|f)=\epsilon_2}{\longrightarrow} \neg b \\ b \overset{\hat{P}(\bar{p}|b)=\epsilon_3}{\longrightarrow} \neg p \end{cases} \tag{12.12}$$

One wants to assess the convictions (assumed to correspond to some conditional probabilities) into the following rules

$$p \wedge b \overset{P(f|p\cap b)=?}{\longrightarrow} f \tag{12.13}$$

$$p \wedge b \overset{P(\bar{f}|p\cap b)=?}{\longrightarrow} \neg f \tag{12.14}$$

The question is to derive rigorously $P(f|p \cap b)$ and $P(\bar{f}|p \cap b)$ from all previous available information. It turns out that the derivation is impossible without unjustified extra assumption on conditional independence. Indeed, $P(f|p \cap b)$ and $P(\bar{f}|p \cap b)$ are given by

$$\begin{cases} P(f|p \cap b) = \frac{P(f,p,b)}{P(p,b)} = \frac{P(p,b|f)P(f)}{P(b|p)P(p)} \\ \\ P(\bar{f}|p \cap b) = \frac{P(\bar{f},p,b)}{P(p,b)} = \frac{P(p,b|\bar{f})P(\bar{f})}{P(b|p)P(p)} \end{cases} \tag{12.15}$$

If one assumes as J. Pearl does, that the conditional independence condition also holds, i.e. $P(p,b|f) = P(p|f)P(b|f)$ and $P(p,b|\bar{f}) = P(p|\bar{f})P(b|\bar{f})$, then one gets

$$\begin{cases} P(f|p \cap b) = \frac{P(p|f)P(b|f)P(f)}{P(b|p)P(p)} \\ \\ P(\bar{f}|p \cap b) = \frac{P(p|\bar{f})P(b|\bar{f})P(\bar{f})}{P(b|p)P(p)} \end{cases}$$

By accepting again the principle of indifference, $P(f) = P(\bar{f}) = 1/2$ and $P(p) = P(\bar{p}) = 1/2$, one gets the following expressions

$$\begin{cases} \hat{P}(f|p \cap b) = \frac{P(p|f)P(b|f)}{P(b|p)} \\ \\ \hat{P}(\bar{f}|p \cap b) = \frac{P(p|\bar{f})P(b|\bar{f})}{P(b|p)} \end{cases} \tag{12.16}$$

Replacing probabilities $P(p|f), P(b|f), P(b|p), P(p|\bar{f})$ and $P(b|\bar{f})$ by their values in the formula (12.16), one finally gets

$$\begin{cases} \hat{P}(f|p \cap b) = \frac{\epsilon_1(1-\epsilon_2)}{1-\epsilon_3} \\ \\ \hat{P}(\bar{f}|p \cap b) = \frac{(1-\epsilon_1)\epsilon_2}{1-\epsilon_3} \end{cases} \tag{12.17}$$

Therefore we see that, even if one accepts the principle of indifference together with the conditional independence assumption, the approximated "probabilities" remain both small and do not correspond to a real measure of probability since the conditional probabilities of exclusive elements f and $\bar{f}$ do not add up to one. When ϵ_1, ϵ_2 and ϵ_3 tends towards 0, one has

$$\hat{P}(f|p \cap b) + \hat{P}(\bar{f}|p \cap b) \approx 0$$

Actually our analysis based on the principle of indifference, the conditional independence assumption and the model proposed by Judea Pearl, proves clearly the impossibility of the Bayesian reasoning to be applied rigorously on such kind of weighted rule-based system, because no probabilistic model exists for describing correctly the problem. This conclusion is actually not surprising taking into account the Lewis' theorem [13] explained in details in [7] (chapter 11).

Let's now explain the reason of the error in the fallacious reasoning which was looking coherent with the common intuition. The problem arises directly from the fact that penguin class and bird class are defined in this problem only with respect to the "flying" and "not-flying" properties. If one considers only these properties, then none Tweety animal can be categorically classified as a penguin-bird, because penguin-birdness doesn't not hold in reality based on these exclusive and exhaustive properties (if we consider only the information given within the rules r_1, r_2 and r_3). Actually everybody knows that penguins are effectively classified as bird because "birdness" property is not defined with respect to the "flying" or "not-flying" abilities of the animal but by other zoological characteristics C (birds are vertebral oviparous animals with hot blood, a beak, feather and anterior members are wings) and such information must be properly taken into account in the rule-based systems to avoid to fall in the trap of such fallacious reasoning. The intuition (which seems to justify the fallacious reasoning conclusion) for TP2 is actually biased because one already knows that penguins (which are truly classified as birds by some other criterions) do not fly in real world and thus we commit a low conviction (which is definitely not a probability measure, but rather a belief) to the fact that a penguin-bird can fly. Thus the Pear'ls analysis proposed in [11] appears to the authors to be unfortunately incomplete and somehow fallacious.

12.4 The Dempster-Shafer reasoning

As pointed out by Judea Pearl in [11], the Dempster-Shafer reasoning yields, for this problem, a very counter-intuitive result: birdness seems to endow Tweety with extra flying power ! We present here our analysis of this problem based on the Dempster-Shafer reasoning.

Let's examine in detail the available prior information summarized by the rule r_1: "*Penguins normally don't fly*" $\Leftrightarrow (p \rightarrow \neg f)$ with the conviction $w_1 = 1 - \epsilon_1$ where ϵ_1 is a small positive number close to zero. This information, in the DST framework, has to be correctly represented in term of a conditional belief $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$ rather than directly the mass $m_1(\bar{f} \cap p) = 1 - \epsilon_1$.

Choosing $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$ means that there is a high degree of belief that a penguin-animal is also a nonflying-animal (whatever kind of animal we are observing). This representation reflects perfectly our prior knowledge while the erroneous coarse modeling based on the commitment $m_1(\bar{f} \cap p) = 1 - \epsilon_1$ is unable to distinguish between rule r_1 and another (possibly erroneous) rule like $r_1' : (\neg f \rightarrow p)$ having same conviction value w_1. This correct model allows us to distinguish between r_1 and r_1' (even if they have the same numerical level of conviction) by considering the two different conditional beliefs $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$ and $\text{Bel}_{1'}(p|\bar{f}) = 1 - \epsilon_1$. The coarse/inadequate basic belief assignment modeling (if adopted) in contrary would make no distinction between those two rules r_1 and r_1' since one would have to take $m_1(\bar{f} \cap p) = m_{1'}(p \cap \bar{f})$ and therefore cannot serve as the starting model for the analysis

Similarly, the prior information relative to rules $r_2 : (b \to f)$ and $r_3 : (p \to b)$ with convictions $w_2 = 1 - \epsilon_2$ and $w_3 = 1 - \epsilon_3$ has to be modeled by the conditional beliefs $\mathrm{Bel}_2(f|b) = 1 - \epsilon_2$ and $\mathrm{Bel}_3(b|p) = 1 - \epsilon_3$ respectively.

The first problem we have to face now is the combination of these three prior information characterized by $\mathrm{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$, $\mathrm{Bel}_2(f|b) = 1 - \epsilon_2$ and $\mathrm{Bel}_3(b|p) = 1 - \epsilon_3$. All the available prior information can be viewed actually as three independent bodies of evidence $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$ providing separately the partial knowledges summarized through the values of $\mathrm{Bel}_1(\bar{f}|p)$, $\mathrm{Bel}_2(f|b)$ and $\mathrm{Bel}_3(b|p)$. To achieve the combination, one needs to define complete basic belief assignments $m_1(.)$, $m_2(.)$ and $m_3(.)$ compatible with the partial conditional beliefs $\mathrm{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$, $\mathrm{Bel}_2(f|b) = 1 - \epsilon_2$ and $\mathrm{Bel}_3(b|p) = 1 - \epsilon_3$ without introducing extra knowledge. We don't want to introduce in the derivations some extra-information we don't have in reality. We present in details the justification for the choice of assignment $m_1(.)$. The choice for $m_2(.)$ and $m_3(.)$ will follow similarly.

The body of evidence $\mathcal{B}_1$ provides some information only about $\bar{f}$ and p through the value of $\mathrm{Bel}_1(\bar{f}|p)$ and without reference to b. Therefore the frame of discernment Θ_1 induced by $\mathcal{B}_1$ and satisfying Shafer's model (i.e. a finite set of exhaustive and exclusive elements) corresponds to

$$\Theta_1 = \{\theta_1 \triangleq \bar{f} \cap \bar{p}, \theta_2 \triangleq f \cap \bar{p}, \theta_3 \triangleq \bar{f} \cap p, \theta_4 \triangleq f \cap p\}$$

schematically represented by

$$f = \theta_2 \cup \theta_4 \left\{ \begin{array}{c} \overbrace{\boxed{\theta_4 \triangleq f \cap p} \quad \boxed{\theta_3 \triangleq \bar{f} \cap p}}^{p=\theta_3\cup\theta_4} \\ \underbrace{\boxed{\theta_2 \triangleq f \cap \bar{p}} \quad \boxed{\theta_1 \triangleq \bar{f} \cap \bar{p}}}_{\bar{p}=\theta_1\cup\theta_2} \end{array} \right\} \bar{f} = \theta_1 \cup \theta_3$$

The complete basic assignment $m_1(.)$ we are searching for and defined over the power set 2^{Θ_1} which must be compatible with $\mathrm{Bel}_1(\bar{f}|p)$ is actually the result of the Dempster's combination of an unknown (for now) basic belief assignment $m_1'(.)$ with the particular assignment $m_1''(.)$ defined by $m_1''(p \triangleq \theta_3 \cup \theta_4) = 1$; in other worlds, one has

$$m_1(.) = [m_1' \oplus m_1''](.)$$

From now on, we introduce explicitly the conditioning term in our notation to avoid confusion and thus we use $m_1(.|p) = m_1(.|\theta_3 \cup \theta_4)$ instead $m_1(.)$. From $m_1''(p \triangleq \theta_3 \cup \theta_4) = 1$ and from any generic unknow basic assignment $m_1'(.)$ defined by its components $m_1'(\emptyset) \triangleq 0$, $m_1'(\theta_1)$, $m_1'(\theta_2)$, $m_1'(\theta_3)$, $m_1'(\theta_4)$, $m_1'(\theta_1 \cup \theta_2)$, $m_1'(\theta_1 \cup \theta_3)$, $m_1'(\theta_1 \cup \theta_4)$, $m_1'(\theta_2 \cup \theta_3)$, $m_1'(\theta_2 \cup \theta_4)$, $m_1'(\theta_3 \cup \theta_4)$, $m_1'(\theta_1 \cup \theta_2 \cup \theta_3)$, $m_1'(\theta_1 \cup \theta_2 \cup \theta_4)$,

$m_1'(\theta_1 \cup \theta_3 \cup \theta_4)$, $m_1'(\theta_2 \cup \theta_3 \cup \theta_4)$, $m_1'(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4)$ and applying Dempter's rule, one gets easily the following expressions for $m_1(.|\theta_3 \cup \theta_4)$. All $m_1(.|\theta_3 \cup \theta_4)$ masses are zero except theoretically

$$m_1(\theta_3|\theta_3 \cup \theta_4) = \overbrace{m_1''(\theta_3 \cup \theta_4)}^{1}[m_1'(\theta_3) + m_1'(\theta_1 \cup \theta_3) + m_1'(\theta_2 \cup \theta_3) + m_1'(\theta_1 \cup \theta_2 \cup \theta_3)]/K_1$$

$$m_1(\theta_4|\theta_3 \cup \theta_4) = \overbrace{m_1''(\theta_3 \cup \theta_4)}^{1}[m_1'(\theta_4) + m_1'(\theta_1 \cup \theta_4) + m_1'(\theta_2 \cup \theta_4) + m_1'(\theta_1 \cup \theta_2 \cup \theta_4)]/K_1$$

$$m_1(\theta_3 \cup \theta_4|\theta_3 \cup \theta_4) = \overbrace{m_1''(\theta_3 \cup \theta_4)}^{1}[m_1'(\theta_3 \cup \theta_4) + m_1'(\theta_1 \cup \theta_3 \cup \theta_4) + m_1'(\theta_2 \cup \theta_3 \cup \theta_4) + m_1'(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4)]/K_1$$

with

$$K_1 \triangleq 1 - \overbrace{m_1''(\theta_3 \cup \theta_4)}^{1}[m_1'(\theta_1) + m_1'(\theta_2) + m_1'(\theta_1 \cup \theta_2)]$$

To complete the derivation of $m_1(.|\theta_3 \cup \theta_4)$, one needs to use the fact that one knows that $\mathrm{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$ which, by definition [16], is expressed by

$$\mathrm{Bel}_1(\bar{f}|p) = \mathrm{Bel}_1(\theta_1 \cup \theta_3|\theta_3 \cup \theta_4) = m_1(\theta_1|\theta_3 \cup \theta_4) + m_1(\theta_3|\theta_3 \cup \theta_4) + m_1(\theta_1 \cup \theta_3|\theta_3 \cup \theta_4) = 1 - \epsilon_1$$

But from the generic expression of $m_1(.|\theta_3 \cup \theta_4)$, one knows also that $m_1(\theta_1|\theta_3 \cup \theta_4) = 0$ and $m_1(\theta_1 \cup \theta_3|\theta_3 \cup \theta_4) = 0$. Thus the knowledge of $\mathrm{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$ implies to have

$$m_1(\theta_3|\theta_3 \cup \theta_4) = [m_1'(\theta_3) + m_1'(\theta_1 \cup \theta_3) + m_1'(\theta_2 \cup \theta_3) + m_1'(\theta_1 \cup \theta_2 \cup \theta_3)]/K_1 = 1 - \epsilon_1$$

This is however not sufficient to fully define the values of all components of $m_1(.|\theta_3 \cup \theta_4)$ or equivalently of all components of $m_1'(.)$. To complete the derivation without extra unjustified specific information, one needs to apply the minimal commitment principle (MCP) which states that one should never give more support to the truth of a proposition than justified [8]. According to this principle, we commit a non null value only to the less specific proposition involved into $m_1(\theta_3|\theta_3 \cup \theta_4)$ expression. In other words, the MCP allows us to choose legitimately

$$m_1'(\theta_1) = m_1'(\theta_2) = m_1'(\theta_3) = 0$$
$$m_1'(\theta_1 \cup \theta_2) = m_1'(\theta_1 \cup \theta_3) = m_1'(\theta_2 \cup \theta_3) = 0$$
$$m_1'(\theta_1 \cup \theta_2 \cup \theta_3) \neq 0$$

Thus $K_1 = 1$ and $m_1(\theta_3|\theta_3 \cup \theta_4)$ reduces to

$$m_1(\theta_3|\theta_3 \cup \theta_4) = m_1'(\theta_1 \cup \theta_2 \cup \theta_3) = 1 - \epsilon_1$$

Since the sum of basic belief assignments must be one, one must also have for the remaining (uncommitted for now) masses of $m_1'(.)$ the constraint

$$\begin{aligned} m_1'(\theta_4) + m_1'(\theta_1 \cup \theta_4) + m_1'(\theta_2 \cup \theta_4) + m_1'(\theta_1 \cup \theta_2 \cup \theta_4) \\ + m_1'(\theta_3 \cup \theta_4) + m_1'(\theta_1 \cup \theta_3 \cup \theta_4) + m_1'(\theta_2 \cup \theta_3 \cup \theta_4) \\ + m_1'(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = \epsilon_1 \end{aligned}$$

By applying a second time the MCP, one chooses $m'_1(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = \epsilon_1$.

Finally, the complete and less specific belief assignment $m_1(.|p)$ compatible with the available prior information $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$ provided by the source $\mathcal{B}_1$ reduces to

$$m_1(\theta_3|\theta_3 \cup \theta_4) = m'_1(\theta_1 \cup \theta_2 \cup \theta_3) = 1 - \epsilon_1 \tag{12.18}$$

$$m_1(\theta_3 \cup \theta_4|\theta_3 \cup \theta_4) = m'_1(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = \epsilon_1 \tag{12.19}$$

or equivalently

$$m_1(\bar{f} \cap p|p) = m'_1(\bar{p} \cup \bar{f}) = 1 - \epsilon_1 \tag{12.20}$$

$$m_1(p|p) = m'_1(\bar{p} \cup \bar{f} \cup p \cup f) = \epsilon_1 \tag{12.21}$$

It is easy to check, from the mass $m_1(.|p)$, that one gets effectively $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$. Indeed:

$$\begin{aligned}
\text{Bel}_1(\bar{f}|p) &= \text{Bel}_1(\theta_1 \cup \theta_3|p) \\
\text{Bel}_1(\bar{f}|p) &= \text{Bel}_1((\bar{f} \cap \bar{p}) \cup (\bar{f} \cap p)|p) \\
\text{Bel}_1(\bar{f}|p) &= \underbrace{m_1(\bar{f} \cap \bar{p}|p)}_{0} + m_1(\bar{f} \cap p|p) \\
&\quad + \underbrace{m_1((\bar{f} \cap \bar{p}) \cup (\bar{f} \cap p)|p)}_{0} \\
\text{Bel}_1(\bar{f}|p) &= m_1(\bar{f} \cap p|p) \\
\text{Bel}_1(\bar{f}|p) &= 1 - \epsilon_1
\end{aligned}$$

In a similar way, for the source $\mathcal{B}_2$ with Θ_2 defined as

$$\Theta_2 = \{\theta_1 \triangleq f \cap \bar{b}, \theta_2 \triangleq \bar{b} \cap \bar{f}, \theta_3 \triangleq f \cap b, \theta_4 \triangleq \bar{f} \cap b\}$$

schematically represented by

$$\bar{f} = \theta_2 \cup \theta_4 \left\{ \underbrace{\overbrace{\begin{array}{cc} \boxed{\theta_4 \triangleq \bar{f} \cap b} & \boxed{\theta_3 \triangleq f \cap b} \\ \boxed{\theta_2 \triangleq \bar{f} \cap \bar{b}} & \boxed{\theta_1 \triangleq f \cap \bar{b}} \end{array}}^{b=\theta_3\cup\theta_4}}_{\bar{b}=\theta_1\cup\theta_2} \right\} f = \theta_1 \cup \theta_3$$

one looks for $m_2(.|b) = [m'_2 \oplus m''_2](.)$ with $m''_2(b) = m''_2(\theta_3 \cup \theta_4) = 1$. From the MCP, the condition $\text{Bel}_2(f|b) = 1 - \epsilon_2$ and with simple algebraic manipulations, one finally gets

$$m_2(\theta_3|\theta_3 \cup \theta_4) = m'_2(\theta_1 \cup \theta_2 \cup \theta_3) = 1 - \epsilon_2 \tag{12.22}$$

$$m_2(\theta_3 \cup \theta_4|\theta_3 \cup \theta_4) = m'_2(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = \epsilon_2 \tag{12.23}$$

or equivalently

$$m_2(f \cap b|b) = m_2'(\bar{b} \cup f) = 1 - \epsilon_2 \tag{12.24}$$

$$m_2(b|b) = m_2'(\bar{b} \cup \bar{f} \cup b \cup f) = \epsilon_2 \tag{12.25}$$

In a similar way, for the source $\mathcal{B}_3$ with Θ_3 defined as

$$\Theta_3 = \{\theta_1 \triangleq b \cap \bar{p}, \theta_2 \triangleq \bar{b} \cap \bar{p}, \theta_3 \triangleq p \cap b, \theta_4 \triangleq \bar{b} \cap p\}$$

schematically represented by

$$\bar{b} = \theta_2 \cup \theta_4 \left\{ \overbrace{\underbrace{\begin{array}{cc} \boxed{\theta_4 \triangleq \bar{b} \cap p} & \boxed{\theta_3 \triangleq b \cap p} \\ \boxed{\theta_2 \triangleq \bar{b} \cap \bar{p}} & \boxed{\theta_1 \triangleq b \cap \bar{p}} \end{array}}_{\bar{p}=\theta_1\cup\theta_2}}^{p=\theta_3\cup\theta_4} \right\} b = \theta_1 \cup \theta_3$$

one looks for $m_3(.|p) = [m_3' \oplus m_3''](.)$ with $m_3''(p) = m_3''(\theta_3 \cup \theta_4) = 1$. From the MCP, the condition $\text{Bel}_3(b|p) = 1 - \epsilon_3$ and with simple algebraic manipulations, one finally gets

$$m_3(\theta_3|\theta_3 \cup \theta_4) = m_3'(\theta_1 \cup \theta_2 \cup \theta_3) = 1 - \epsilon_3 \tag{12.26}$$

$$m_3(\theta_3 \cup \theta_4|\theta_3 \cup \theta_4) = m_3'(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = \epsilon_3 \tag{12.27}$$

or equivalently

$$m_3(b \cap p|p) = m_3'(\bar{p} \cup b) = 1 - \epsilon_3 \tag{12.28}$$

$$m_3(p|p) = m_3'(\bar{b} \cup \bar{p} \cup b \cup p) = \epsilon_3 \tag{12.29}$$

Since all the complete prior basic belief assignments are available, one can combine them with the Dempster's rule to summarize all our prior knowledge drawn from our simple rule-based expert system characterized by rules $R = \{r_1, r_2, r_3\}$ and convictions/confidences $W = \{w_1, w_2, w_3\}$ in these rules.

The fusion operation requires to primilarily choose the following frame of discernment Θ (satisfying Shafer's model) given by

$$\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8\}$$

where

$$\begin{array}{ll}
\theta_1 \triangleq f \cap b \cap p & \theta_5 \triangleq \bar{f} \cap b \cap p \\
\theta_2 \triangleq f \cap b \cap \bar{p} & \theta_6 \triangleq \bar{f} \cap b \cap \bar{p} \\
\theta_3 \triangleq f \cap \bar{b} \cap p & \theta_7 \triangleq \bar{f} \cap \bar{b} \cap p \\
\theta_4 \triangleq f \cap \bar{b} \cap \bar{p} & \theta_8 \triangleq \bar{f} \cap \bar{b} \cap \bar{p}
\end{array}$$

The fusion of masses $m_1(.)$ given by eqs. (12.20)-(12.21) with $m_2(.)$ given by eqs. (12.24)-(12.25) using the Demspter's rule of combination [16] yields $m_{12}(.) = [m_1 \oplus m_2](.)$ with the following non null components

$$m_{12}(f \cap b \cap p) = \epsilon_1(1-\epsilon_2)/K_{12}$$

$$m_{12}(\bar{f} \cap b \cap p) = \epsilon_2(1-\epsilon_1)/K_{12}$$

$$m_{12}(b \cap p) = \epsilon_1\epsilon_2/K_{12}$$

with $K_{12} \triangleq 1 - (1-\epsilon_1)(1-\epsilon_2) = \epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2$.

The fusion of all prior knowledge by the Dempster's rule $m_{123}(.) = [m_1 \oplus m_2 \oplus m_3](.) = [m_{12} \oplus m_3](.)$ yields the final result :

$$m_{123}(f \cap b \cap p) = m_{123}(\theta_1) = \epsilon_1(1-\epsilon_2)/K_{123}$$

$$m_{123}(\bar{f} \cap b \cap p) = m_{123}(\theta_5) = \epsilon_2(1-\epsilon_1)/K_{123}$$

$$m_{123}(b \cap p) = m_{123}(\theta_1 \cup \theta_5) = \epsilon_1\epsilon_2/K_{123}$$

with $K_{123} = K_{12} \triangleq 1 - (1-\epsilon_1)(1-\epsilon_2) = \epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2$.

which defines actually and precisely the conditional belief assignment $m_{123}(.|p \cap b)$. It turns out that the fusion with the last basic belief assignment $m_3(.)$ brings no change with respect to previous fusion result $m_{12}(.)$ in this particular problem.

Since we are actually interested to assess the belief that our observed particular penguin-animal named Tweety (denoted as $T = (p \cap b)$) can fly, we need to combine all our prior knowledge $m_{123}(.)$ drawn from our rule-based system with the belief assignment $m_o(T = (p \cap b)) = 1$ characterizing the observation about Tweety. Applying again the Demspter's rule, one finally gets the resulting conditional basic belief function $m_{o123} = [m_o \oplus m_{123}](.)$ defined by

$$m_{o123}(T = (f \cap b \cap p)|T = (p \cap b)) = \epsilon_1(1-\epsilon_2)/K_{12}$$

$$m_{o123}(T = (\bar{f} \cap b \cap p)|T = (p \cap b)) = \epsilon_2(1-\epsilon_1)/K_{12}$$

$$m_{o123}(T = (b \cap p)|T = (p \cap b)) = \epsilon_1\epsilon_2/K_{12}$$

From the Dempster-Shafer reasoning, the belief and plausibity that Tweety can fly are given by [16]

$$\text{Bel}(T = f|T = (p \cap b)) = \sum_{x \in 2^\Theta, x \subseteq f} m_{o123}(T = x|T = (p \cap b))$$

$$\text{Pl}(T = f|T = (p \cap b)) = \sum_{x \in 2^\Theta, x \cap f \neq \emptyset} m_{o123}(T = x|T = (p \cap b))$$

Because $f = [(f \cap b \cap p) \cup (f \cap b \cap \bar{p}) \cup (f \cap \bar{b} \cap p) \cup (f \cap \bar{b} \cap \bar{p})]$ and the specific values of the masses defining $m_{o123}(.)$, one has

$$\text{Bel}(T = f|T = (p \cap b)) = m_{o123}(T = (f \cap b \cap p)|T = (p \cap b))$$

$$\text{Pl}(T = f|T = (p \cap b)) = m_{o123}(T = (f \cap b \cap p)|T = (p \cap b)) + m_{o123}(T = (b \cap p)|T = (p \cap b))$$

and finally

$$\text{Bel}(T = f|T = (p \cap b)) = \frac{\epsilon_1(1 - \epsilon_2)}{K_{12}} \tag{12.30}$$

$$\text{Pl}(T = f|T = (p \cap b)) = \frac{\epsilon_1(1 - \epsilon_2)}{K_{12}} + \frac{\epsilon_1\epsilon_2}{K_{12}} = \frac{\epsilon_1}{K_{12}} \tag{12.31}$$

In a similar way, one will get for the belief and the plausibility that Tweety cannot fly

$$\text{Bel}(T = \bar{f}|T = (p \cap b)) = \frac{\epsilon_2(1 - \epsilon_1)}{K_{12}} \tag{12.32}$$

$$\text{Pl}(T = \bar{f}|T = (p \cap b)) = \frac{\epsilon_2(1 - \epsilon_1)}{K_{12}} + \frac{\epsilon_1\epsilon_2}{K_{12}} = \frac{\epsilon_2}{K_{12}} \tag{12.33}$$

Using the first order approximation when ϵ_1 and ϵ_2 are very small positive numbers, one gets finally

$$\text{Bel}(T = f|T = (p \cap b)) = \text{Pl}(T = f|T = (p \cap b)) \approx \frac{\epsilon_1}{\epsilon_1 + \epsilon_2}$$

In a similar way, one will get for the belief that Tweety cannot fly

$$\text{Bel}(T = \bar{f}|T = (p \cap b)) = \text{Pl}(T = \bar{f}|T = (p \cap b)) \approx \frac{\epsilon_2}{\epsilon_1 + \epsilon_2}$$

This result coincides with the Judea Pearl's result but a different analysis and detailed presentation has been done here. It turns out that this simple and complete analysis corresponds actually to the ballooning extension and the generalized Bayesian theorem proposed by Smets in [21, 24] and discussed by Shafer in [17] although it was carried out independently of Smets' works. As pointed out by Judea Pearl, this result based on DST and the Dempster's rule of combination looks very paradoxical/counter-intuitive since it means that if nonflying birds are very rare, i.e. $\epsilon_2 \approx 0$, then penguin-birds like our observed penguin-bird Tweety, have a very big chance of flying. As stated by Judea Pearl in [11] pages 448-449: *"The clash with intuition revolves not around the exact numerical value of Bel(f) but rather around the unacceptable phenomenon that rule r_3, stating that penguins are a subclass of birds, plays no role in the analysis. Knowing that Tweety is both a penguin and a bird renders Bel($T = f|T = (p \cap b)$) solely a function of $m_1(.)$ and $m_2(.)$, regardless of how penguins and birds are related. This stands contrary to common discourse, where people expect class properties to be overridden by properties of more specific subclasses. While in classical logic the three rules in our example would yield an unforgivable contradiction, the uncertainties attached to these rules, together with Dempster's normalization, now* *render them manageable. However, they are managed in the wrong way whenever we interpret if-then rules as randomized logical formulas of the material-implication type, instead of statements of conditional probabilities".* Keep in mind that this Pearl's statement is however given to show the semantic clash between the Dempster-Shafer reasoning vs. the fallacious Bayesian reasoning to support the Bayesian reasoning approach.

12.5 The Dezert-Smarandache reasoning

We analyze here the Tweety penguin triangle problem with the DSmT (see Part I of this book for a presentation of DSmT). The prior knowledge characterized by the rules $R = \{r_1, r_2, r_3\}$ and convictions $W = \{w_1, w_2, w_3\}$ is modeled as three independent sources of evidence defined on separate minimal and potentially paradoxical (i.e internal conflicting) frames $\Theta_1 \triangleq \{p, \bar{f}\}$, $\Theta_2 \triangleq \{b, f\}$ and $\Theta_3 \triangleq \{p, b\}$ since the rule r_1 doesn't refer to the existence of b, the rule r_2 doesn't refer to the existence of p and the rule r_3 doesn't refer to the existence of f or $\bar{f}$. Let's note that the DSmT doesn't require the refinement of frames as with DST (see previous section). We follow the same analysis as in previous section but now based on our DSm reasoning and the DSm rule of combination.

The first source $\mathcal{B}_1$ relative to r_1 with confidence $w_1 = 1 - \epsilon_1$ provides us the conditional belief $\text{Bel}_1(\bar{f}|p)$ which is now defined from a paradoxical basic belief assignment $m_1(.)$ resulting of the DSm combination of $m''_1(p) = 1$ with $m''_1(.)$ defined on the hyper-power set $D^{\Theta_1} = \{\emptyset, p, \bar{f}, p \cap \bar{f}, p \cup \bar{f}\}$. The choice for $m'_1(.)$ results directly from the derivation of the DSm rule and the application of the MCP. Indeed, the non null components of $m_1(.)$ are given by (we introduce explicitly the conditioning term in notation for convenience):

$$m_1(p|p) = \overbrace{m''_1(p)}^{1} m'_1(p) + \overbrace{m''_1(p)}^{1} m'_1(p \cup \bar{f})$$
$$m_1(p \cap \bar{f}|p) = \overbrace{m''_1(p)}^{1} m'_1(\bar{f}) + \overbrace{m''_1(p)}^{1} m'_1(p \cap \bar{f})$$

The information $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$ implies

$$\text{Bel}_1(\bar{f}|p) = m_1(\bar{f}|p) + m_1(p \cap \bar{f}|p) = 1 - \epsilon_1$$

Since $m_1(p|p) + m_1(p \cap \bar{f}|p) = 1$, one has necessarily $m_1(\bar{f}|p) = 0$ and thus from previous equation $m_1(\bar{f} \cap p|p) = 1 - \epsilon_1$, which implies both

$$m_1(p|p) = \epsilon_1$$

$$m_1(p \cap \bar{f}|p) = \overbrace{m_1''(p)}^{1} m_1'(\bar{f}) + \overbrace{m_1''(p)}^{1} m_1'(p \cap \bar{f}) = m_1'(\bar{f}) + m_1'(p \cap \bar{f}) = 1 - \epsilon_1$$

Applying the MCP, it results that one must choose

$$m_1'(\bar{f}) = 1 - \epsilon_1 \quad \text{and} \quad m_1'(p \cap \bar{f}) = 0$$

The sum of remaining masses of $m_1'(.)$ must be then equal to ϵ_1, i.e.

$$m_1'(p) + m_1'(p \cup \bar{f}) = \epsilon_1$$

Applying again the MCP on this last constraint, one gets naturally

$$m_1'(p) = 0 \quad \text{and} \quad m_1'(p \cup \bar{f}) = \epsilon_1$$

Finally the belief assignment $m_1(.|p)$ relative to the source $\mathcal{B}_1$ and compatible with the constraint $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$, holds the same numerical values as within the DST analysis (see eqs. (12.20)-(12.21)) and is given by

$$m_1(p \cap \bar{f}|p) = 1 - \epsilon_1$$

$$m_1(p|p) = \epsilon_1$$

but results here from the DSm combination of the two following assignments (i.e. $m_1(.) = [m_1' \oplus m_1''](.) = [m_1'' \oplus m_1'](.)$)

$$\begin{cases} m_1'(\bar{f}) = 1 - \epsilon_1 \quad \text{and} \quad m_1'(p \cup \bar{f}) = \epsilon_1 \\ m_1''(p) = 1 \end{cases} \tag{12.34}$$

In a similarly manner and working on $\Theta_2 = \{b, f\}$ for source $\mathcal{B}_2$ with the condition $\text{Bel}_2(f|b) = 1 - \epsilon_2$, the mass $m_2(.|b)$ results from the internal DSm combination of the two following assignments

$$\begin{cases} m_2'(f) = 1 - \epsilon_2 \quad \text{and} \quad m_2'(b \cup f) = \epsilon_2 \\ m_2''(b) = 1 \end{cases} \tag{12.35}$$

Similarly and working on $\Theta_3 = \{p, b\}$ for source $\mathcal{B}_3$ with the condition $\text{Bel}_3(b|p) = 1 - \epsilon_3$, the mass $m_3(.|p)$ results from the internal DSm combination of the two following assignments

$$\begin{cases} m_3'(b) = 1 - \epsilon_3 \quad \text{and} \quad m_3'(b \cup p) = \epsilon_3 \\ m_3''(p) = 1 \end{cases} \tag{12.36}$$

It can be easily verified that these (less specific) basic belief assignments generates the conditions $\text{Bel}_1(\bar{f}|p) = 1 - \epsilon_1$, $\text{Bel}_2(f|b) = 1 - \epsilon_2$ and $\text{Bel}_3(b|p) = 1 - \epsilon_3$.

Now let's examine the result of the fusion of all these masses based on DSmT, i.e by applying the DSm rule of combination of the following basic belief assignments

$$m_1(p \cap \bar{f}|p) = 1 - \epsilon_1 \quad \text{and} \quad m_1(p|p) = \epsilon_1$$

$$m_2(b \cap f|b) = 1 - \epsilon_2 \quad \text{and} \quad m_2(b|b) = \epsilon_2$$

$$m_3(p \cap b|p) = 1 - \epsilon_3 \quad \text{and} \quad m_3(p|p) = \epsilon_3$$

Note that these basic belief assignments turn to be identical to those drawn from DST framework analysis done in previous section for this specific problem because of integrity constraint $f \cap \bar{f} = \emptyset$ and the MCP, but result actually from a slightly different and simpler analysis here drawn from DSmT. So we attack the TP2 with the same information as with the analysis based on DST, but we will show that a coherent conclusion can be drawn with DSm reasoning.

Let's emphasize now that one has to deal here with the hypotheses/elements p, b, f and $\bar{f}$ and thus our global frame is given by $\Theta = \{b, p, f, \bar{f}\}$. Note that Θ doesn't satisfy Shafer's model since the elements of Θ are not all exclusive. This is a major difference between the foundations of DSmT with respect to the foundations of DST. But because only f and $\bar{f}$ are truly exclusive, i.e. $\bar{f} \cap f = \emptyset$, we are face to a quite simple hybrid DSm model $\mathcal{M}$ and thus the hybrid DSm fusion must apply rather than the classic DSm rule. We recall briefly here (a complete derivation, justification and examples can be found in chapter 4) the hybrid DSm rule of combination associated to a given hybrid DSm model for $k \geq 2$ independent sources of information is defined for all $A \in D^\Theta$ as:

$$m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A)\Big[S_1(A) + S_2(A) + S_3(A)\Big] \tag{12.37}$$

where $\phi(A)$ is the *characteristic non emptiness function* of the set A, i.e. $\phi(A) = 1$ if $A \notin \boldsymbol{\emptyset}$ ($\boldsymbol{\emptyset} \triangleq \{\emptyset, \boldsymbol{\emptyset}_{\mathcal{M}}\}$ being the set of all relatively and absolutely empty elements) and $\phi(A) = 0$ otherwise, and

$$S_1(A) \triangleq \sum_{\substack{X_1, X_2, \ldots, X_k \in D^\Theta \\ (X_1 \cap X_2 \cap \ldots \cap X_k) = A}} \prod_{i=1}^{k} m_i(X_i) \tag{12.38}$$

$$S_2(A) \triangleq \sum_{\substack{X_1, X_2, \ldots, X_k \in \boldsymbol{\emptyset} \\ [\mathcal{U} = A] \vee [(\mathcal{U} \in \boldsymbol{\emptyset}) \wedge (A = I_t)]}} \prod_{i=1}^{k} m_i(X_i) \tag{12.39}$$

$$S_3(A) \triangleq \sum_{\substack{X_1, X_2, \ldots, X_k \in D^\Theta \\ (X_1 \cup X_2 \cup \ldots \cup X_k) = A \\ (X_1 \cap X_2 \cap \ldots \cap X_k) \in \boldsymbol{\emptyset}}} \prod_{i=1}^{k} m_i(X_i) \tag{12.40}$$

with $\mathcal{U} \triangleq u(X_1) \cup u(X_2) \cup \ldots \cup u(X_k)$ where $u(X)$ is the union of all singletons θ_i that compose X and $I_t \triangleq \theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$ is the total ignorance defined on the frame $\Theta = \{\theta_1, \ldots, \theta_n\}$. For example, if X is a singleton then $u(X) = X$; if $X = \theta_1 \cap \theta_2$ or $X = \theta_1 \cup \theta_2$ then $u(X) = \theta_1 \cup \theta_2$; if $X = (\theta_1 \cap \theta_2) \cup \theta_3$ then $u(X) = \theta_1 \cup \theta_2 \cup \theta_3$; by convention $u(\emptyset) \triangleq \emptyset$.

The first sum $S_1(A)$ entering in the previous formula corresponds to mass $m_{\mathcal{M}^f(\Theta)}(A)$ obtained by the classic DSm rule of combination based on the free DSm model $\mathcal{M}^f$ (i.e. on the free lattice D^Θ). The second sum $S_2(A)$ entering in the formula of the hybrid DSm rule of combination (12.37) represents the mass of all relatively and absolutely empty sets which is transferred to the total or relative ignorances. The third sum $S_3(A)$ entering in the formula of the hybrid DSm rule of combination (12.37) transfers the sum of relatively empty sets to the non-empty sets in the same way as it was calculated following the DSm classic rule.

To apply the hybrid DSm fusion rule formula (12.37), it is important to note that $(p\cap \bar{f})\cap(b\cap f)\cap p \equiv p\cap b\cap f\cap \bar{f} = \emptyset$ because $f\cap \bar{f} = \emptyset$, thus the mass $(1-\epsilon_1)(1-\epsilon_2)\epsilon_3$ is transferred to the hybrid proposition $H_1 \triangleq (p\cap \bar{f}) \cup (b\cap f) \cup p \equiv (b\cap f)\cup p$; similarly $(p\cap \bar{f}) \cap (b\cap f) \cap (p\cap b) \equiv p\cap b\cap f\cap \bar{f} = \emptyset$ because $f\cap \bar{f} = \emptyset$ and therefore its associated mass $(1-\epsilon_1)(1-\epsilon_2)(1-\epsilon_3)$ is transferred to the hybrid proposition $H_2 \triangleq (p\cap \bar{f}) \cup (b\cap f) \cup (p\cap b)$. No other mass transfer is necessary for this Tweety Penguin Triangle Problem and thus we finally get from hybrid DSm fusion formula (12.37) the following result for $m_{123}(.|p\cap b) = [m_1 \oplus m_2 \oplus m_3](.)$ (where $\oplus$ symbol corresponds here to the DSm fusion operator and we omit the conditioning term $p\cap b$ here due to space limitation):

$$
\begin{aligned}
m_{123}((b\cap f)\cup p|p\cap b) &= (1-\epsilon_1)(1-\epsilon_2)\epsilon_3 \\
m_{123}((p\cap \bar{f})\cup(b\cap f)\cup(p\cap b)|p\cap b) &= (1-\epsilon_1)(1-\epsilon_2)(1-\epsilon_3) \\
m_{123}(p\cap b\cap \bar{f}|p\cap b) &= (1-\epsilon_1)\epsilon_2\epsilon_3 + (1-\epsilon_1)\epsilon_2(1-\epsilon_3) = (1-\epsilon_1)\epsilon_2 \\
m_{123}(p\cap b\cap f|p\cap b) &= \epsilon_1(1-\epsilon_2)\epsilon_3 + \epsilon_1(1-\epsilon_2)(1-\epsilon_3) = \epsilon_1(1-\epsilon_2) \\
m_{123}(p\cap b|p\cap b) &= \epsilon_1\epsilon_2\epsilon_3 + \epsilon_1\epsilon_2(1-\epsilon_3) = \epsilon_1\epsilon_2
\end{aligned}
$$

We can check all these masses add up to 1 and that this result is fully coherent with the *rational intuition* especially when $\epsilon_3 = 0$, because non null components of $m_{123}(.|p\cap b)$ reduces to

$$
\begin{aligned}
m_{123}((p\cap \bar{f})\cup(b\cap f)\cup(p\cap b)|p\cap b) &= (1-\epsilon_1)(1-\epsilon_2) \\
m_{123}(p\cap b\cap \bar{f}|p\cap b) &= (1-\epsilon_1)\epsilon_2 \\
m_{123}(p\cap b\cap f|p\cap b) &= \epsilon_1(1-\epsilon_2) \\
m_{123}(p\cap b|p\cap b) &= \epsilon_1\epsilon_2
\end{aligned}
$$

which means that from our DSm reasoning there is a strong uncertainty (due to the conflicting rules of our rule-based system), when ϵ_1 and ϵ_2 remain small positive numbers, that a penguin-bird animal is either a penguin-nonflying animal or a bird-flying animal. The small value $\epsilon_1\epsilon_2$ for $m_{123}(p \cap b|p \cap b)$ expresses adequately the fact that we cannot commit a strong basic belief assignment only to $p \cap b$ knowing $p \cap b$ just because one works on $\Theta = \{p, b, f, \bar{f}\}$ and we cannot consider the property $p \cap b$ solely because the"birdness" or "penguinness" property endow necessary either the flying or non-flying property.

Therefore the belief that the particular observed penguin-bird animal Tweety (corresponding to the particular mass $m_o(T = (p \cap b)) = 1$) can be easily derived from the DSm fusion of all our prior summarized by $m_{123}(.|p \cap b)$ and the available observation summarized by $m_o(.)$ and we get

$$m_{o123}(T = (p \cap b \cap \bar{f})|T = (p \cap b)) = (1 - \epsilon_1)\epsilon_2$$
$$m_{o123}(T = (p \cap b \cap f)|T = (p \cap b)) = \epsilon_1(1 - \epsilon_2)$$
$$m_{o123}(T = (p \cap b)|T = (p \cap b)) = \epsilon_1\epsilon_2$$
$$m_{o123}(T = (b \cap f) \cup p|T = (p \cap b)) = (1 - \epsilon_1)(1 - \epsilon_2)\epsilon_3$$
$$m_{o123}(T = (p \cap \bar{f}) \cup (b \cap f) \cup (p \cap b)|T = (p \cap b)) = (1 - \epsilon_1)(1 - \epsilon_2)(1 - \epsilon_3)$$

From the DSm reasoning, the belief that Tweety can fly is then given by

$$\mathrm{Bel}(T = f|T = (p \cap b)) = \sum_{x \in D^\Theta, x \subseteq f} m_{o123}(T = x|T = (p \cap b))$$

Using all the components of $m_{o123}(.|T = (p \cap b))$, one directly gets

$$\mathrm{Bel}(T = f|T = (p \cap b)) = m_{o123}(T = (f \cap b \cap p)|T = (p \cap b))$$

and finally

$$\mathrm{Bel}(T = f|T = (p \cap b)) = \epsilon_1(1 - \epsilon_2) \tag{12.41}$$

In a similar way, one will get for the belief that Tweety cannot fly

$$\mathrm{Bel}(T = \bar{f}|T = (p \cap b)) = \epsilon_2(1 - \epsilon_1) \tag{12.42}$$

So now for both cases the beliefs remain very low which is normal and coherent with analysis done in section 12.3.2. Now let's examine the plausibilities of the ability for Tweety to fly or not to fly. These are given by

$$\mathrm{Pl}(T = f|T = (p \cap b)) \triangleq \sum_{x \in D^\Theta, x \cap f \neq \emptyset} m_{o123}(T = x|T = (p \cap b))$$
$$\mathrm{Pl}(T = \bar{f}|T = (p \cap b)) \triangleq \sum_{x \in D^\Theta, x \cap \bar{f} \neq \emptyset} m_{o123}(T = x|T = (p \cap b))$$

which turn to be after elementary algebraic manipulations

$$\text{Pl}(T = f|T = (p \cap b)) = (1 - \epsilon_2) \tag{12.43}$$

$$\text{Pl}(T = \bar{f}|T = (p \cap b)) = (1 - \epsilon_1) \tag{12.44}$$

So we conclude, as reasonably/rationally expected, that we can't decide on the ability for Tweety of flying or of not flying, since one has

$$[\text{Bel}(f|p \cap b), \text{Pl}(f|p \cap b)] = [\epsilon_1(1 - \epsilon_2), (1 - \epsilon_2)] \approx [0, 1]$$

$$[\text{Bel}(\bar{f}|p \cap b), \text{Pl}(\bar{f}|p \cap b)] = [\epsilon_2(1 - \epsilon_1), (1 - \epsilon_1)] \approx [0, 1]$$

Note that when setting $\epsilon_1 = 0$ and $\epsilon_2 = 1$ (or $\epsilon_1 = 1$ and $\epsilon_2 = 0$), i.e. one forces the full consistency of the initial rules-based system, one gets coherent result on the certainty of the ability of Tweety to not fly (or to fly respectively).

This coherent result (radically different from the one based on Dempster-Shafer reasoning but starting with exactly the same available information) comes from the hybrid DSm fusion rule which transfers some parts of the mass of empty set $m(\emptyset) = (1 - \epsilon_1)(1 - \epsilon_2)\epsilon_3 + (1 - \epsilon_1)(1 - \epsilon_2)(1 - \epsilon_3) \approx 1$ onto propositions $(b \cap f) \cup p$ and $(p \cap \bar{f}) \cup (b \cap f) \cup (p \cap b)$.

It is clear however that the high value of $m(\emptyset)$ in this TP2 indicates a high conflicting fusion problem which proves that the TP2 is a true almost impossible problem and the fusion result based on DSmT reasoning allows us to conclude on the true undecidability on the ability for Tweety of flying or of not flying. In other words, the fusion based on DSmT can be applied adequately on this almost impossible problem and concludes correctly on its indecibility. Another simplistic solution would consist to say naturally that the problem has to be considered as an impossible one just because $m(\emptyset) \geq 0.5$.

12.6 Conclusion

In this chapter we have proposed a deep analysis of the challenging Tweety Penguin Triangle Problem. The analysis proves that the Bayesian reasoning cannot be mathematically justified to characterize the problem because the probabilistic model doesn't hold, even with the help of acceptance of the principle of indifference and the conditional independence assumption. Any conclusions drawn from such representation of the problem based on a hypothetical probabilistic model are based actually on a fallacious Bayesian reasoning. This is a fundamental result. Then one has shown how the Dempster-Shafer reasoning manages in what we feel is a wrong way the uncertainty and the conflict in this problem. We then proved that the DSmT can deal properly with this problem and provides a well-founded and reasonable conclusion about the undecidability of its solution.

Acknowledgments

Authors are grateful to Dr. Roy Streit, Naval Undersea Warfare Center, Newport, RI, U.S.A., to have them introduced and encouraged to work on this nice exciting and challenging problem during the Fusion 2003 International Conference on Information Fusion, Cairns, Australia, July 2003. Authors want to thank Professor Smets for his discussions, suggestions and for pointing out important references for the improvement of this chapter.

12.7 References

[1] Benferhat S., Saffioti A., Smets Ph., *Belief functions and default reasoning*, Artificial Intelligence 122, pp. 1-69, 2000.

[2] Buchanan B.G., Shortliffe E.H., *Rule Bases Expert Systems - The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison Wesley, 1984.

[3] Dezert J., Smarandache F., *Partial ordering of hyper-power sets and matrix representation of belief functions within DSmT*, Proc. of Fusion 2003 Conf., Cairns, Australia, July 8-11, 2003.

[4] Dezert J., Smarandache F., *On the generation of hyper-power sets for the DSmT*, Proceedings of the 6th International Conference on Information Fusion, Cairns, Australia, July 8-11, 2003.

[5] Dubois D., Prade H., *Fuzzy sets and systems, theory and applications*, Academic Press, 1980.

[6] Dubois D., Prade H., *Théorie des possibilités, application à la représentation des connaissances en informatique*, Masson, (2nd edition), 1987.

[7] Goodman I.R., Mahler R.P.S., Nguyen H.T., *Mathematics of Data Fusion*, Kluwer Acad. Press, 1997.

[8] Klawonn F., Smets Ph., *The Dynamic of Belief in the Transferable Belief Model and Specialization-Generalization Matrices*, Proc. of the 8th Conf. on Uncertainty in AI. Dubois D., Wellman M.P., D'Ambrosio B. and Smets Ph. (eds). Morgan Kaufmann Publ., San Mateo, CA, pp. 130-137, 1992.

[9] Kleene, S. C., *Mathematical Logic*, New York,Dover, 2002.

[10] Mendelson, E.,*The Propositional Calculus*, Ch. 1 in Introduction to Mathematical Logic, 4th ed. London: Chapman & Hall, pp. 12-44, 1997.

[11] Pearl J., *Probabilistic reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, San Mateo, CA, 1988.

[12] Pearl J., *Reasoning with belief functions and analysis of compatibility*, Int. Journal of Approximate Reasoning, 4, pp. 363-390, 1990.

[13] Lewis D., *Probabilities of conditonals and conditional probabilities*, Philsophical Review, 85, 1976.

[14] Poole D.L., *On the comparison of theories: preferring the most specific explanation*, Proc. 9th Int. Joint Conf. on Artif. Intelligence, Los Angeles, pp. 465-474, 1985.

[15] Nidditch, P. H, *Propositional Calculus*, New York, Free Press of Glencoe, 1962.

[16] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[17] Shafer G., *Belief functions and parametric models*, J. Roy. Statist. Soc., B., 44, pp. 322-352, 1982.

[18] Smarandache F., *Deducibility Theorems in Mathematics logics*, Ann. Univ. Timisoara, Seria ST., Matematica, Vol. XVII, Fasc. 2, pp. 163-168, 1979 (available in [19])

[19] Smarandache F., *Collected Papers, Vol. I*, Editura Tempus Romania SRL, 1995, pp. 232-239, available at `http://www.gallup.unm.edu/~smarandache/CP1.pdf`.

[20] Smarandache F., Dezert J. (Editors), *Advances and Applications of DSmT for Information Fusion (Collected works)*, American Research Press, June 2004.

[21] Smets Ph., *Un modèle mathématico-statistique simulant le processus de diagnostic médical*, Ph.D. Thesis, Université Libre de Bruxelles (available through University Microfilm International, 30-32 Mortimer Street London, W1N 7RA, Thesis 80-70,003), 1978.

[22] Smets Ph., Hsia Y.T., *Default reasoning and the transferable belief model*, In M. Henrion, L.N. Kanal and J.F. Lemmer (Editors), Uncertainty in Artificial Intelligence 5, Elsevier Science Publishers, pp. 495-504, 1990.

[23] Smets Ph., *Resolving misunderstandings about belief functions: A response to the many criticisms raised by Judea Pearl*, Int. J. Approximate Reasoning, 6, pp. 321-344, 1992.

[24] Smets Ph., *Belief functions: The disjunctive rule of combination and the Generalized Bayesian Theorem*, Int. J. Approximate Reasoning, 9, pp. 1-35, 1993.

[25] Zadeh L., *Fuzzy sets as a basis for a theory of possibility*, Fuzzy sets and systems 1, pp. 3-28, 1978.

Chapter 13

Estimation of Target Behavior Tendencies using DSmT

Albena Tchamova, Tzvetan Semerdjiev
Institute for Parallel Processing
Bulgarian Academy of Sciences
Sofia, Bulgaria

Jean Dezert
ONERA
29 Av. de la Division Leclerc
92320 Châtillon, France

Abstract: *This chapter presents an approach for target behavior tendency estimation (Receding, Approaching). It is developed on the principles of Dezert-Smarandache theory (DSmT) of plausible and paradoxical reasoning applied to conventional sonar amplitude measurements, which serve as an evidence for corresponding decision-making procedures. In some real world situations it is difficult to finalize these procedures, because of discrepancies in measurements interpretation. In these cases the decision-making process leads to conflicts, which cannot be resolved using the well-known methods. The aim of the performed study is to present and to approve the ability of DSmT to finalize successfully the decision-making process and to assure awareness about the tendencies of target behavior in case of discrepancies in measurements interpretation. An example is provided to illustrate the benefit of the proposed approach application in comparison of fuzzy logic approach, and its ability to improve the overall tracking performance.*

This chapter is based on a paper [7] presented during the International Conference on Information Fusion, Fusion 2003, Cairns, Australia, in July 2003 and is reproduced here with permission of the International Society of Information Fusion. This work has been partially supported by MONT grants I-1205/02, I-1202/02 and by Center of Excellence BIS21 grant ICA1-2000-70016

13.1 Introduction

Angle-only tracking systems based on sonars are poorly developed topic due to a number of complications. These systems tend to be less precise than those based on active sensors, but one important advantage is their vitality of being stealth. In a single sensor case only direction of the target as an axis is known, but the true target position and behavior (approaching or descending) remain unknown. Recently, the advances of computer technology lead to sophisticated data processing methods, which improve sonars capability. A number of developed tracking techniques operating on angle-only measurement data use additional information. In our case we utilize the measured emitter's amplitude values in consecutive time moments. This information can be used to assess tendencies in target's behavior and, consequently, to improve the overall angle-only tracking performance. The aim of the performed study is to present and to approve the ability of DSmT to finalize successfully the decision-making process and to assure awareness about the tendencies of target behavior in case of discrepancies of angle-only measurements interpretation. Results are presented and compared with the respective results, but drawn from the fuzzy logic approach.

13.2 Statement of the Problem

In order to track targets using angle-only measurements it is necessary to compensate the unknown ranges by using additional information received from the emitter. In our case we suppose that in parallel with measured local angle the observed target emits constant signal, which is perceived by the sensor with a non-constant, but a varying strength (referred as amplitude). The augmented measurement vector at the end of each time interval $k = 1, 2, \dots$ is $Z = \{Z_\theta, Z_A\}$, where: $Z_\theta = \theta + \nu_\theta$ denotes the measured local angle with zero-mean Gaussian noise $\nu_\theta = \mathcal{N}(0, \sigma_{\nu_\theta})$ and covariance σ_{ν_θ}; $Z_A = A + \nu_A$ denotes corresponding signal's amplitude value with zero-mean Gaussian noise $\nu_A = \mathcal{N}(0, \sigma_{\nu_A})$ and covariance σ_{ν_A}. The variance of amplitude value is because of the cluttered environment and the varying unknown distance to the object, which is conditioned by possible different modes of target behavior (approaching or descending). Our goal is, utilizing received amplitude feature measurement, to predict and to estimate the possible target behavior tendencies.

Figure 13.1 represents a block diagram of the target's behavior tracking system. Regarding to the formulated problem, we maintain two single-model-based Kalman-like filters running in parallel using two models of possible target behavior - *Approaching* and *Receding*. At initial time moment k the target is characterized by the fuzzified amplitude state estimates according to the models $A^{\text{App}}(k|k)$ and $A^{\text{Rec}}(k|k)$. The new observation $Z_A(k+1) = A(k+1) + \nu_A(k+1)$ is assumed to be the true value, corrupted by additive measurement noise. It is fuzzified according to the chosen fuzzification interface.

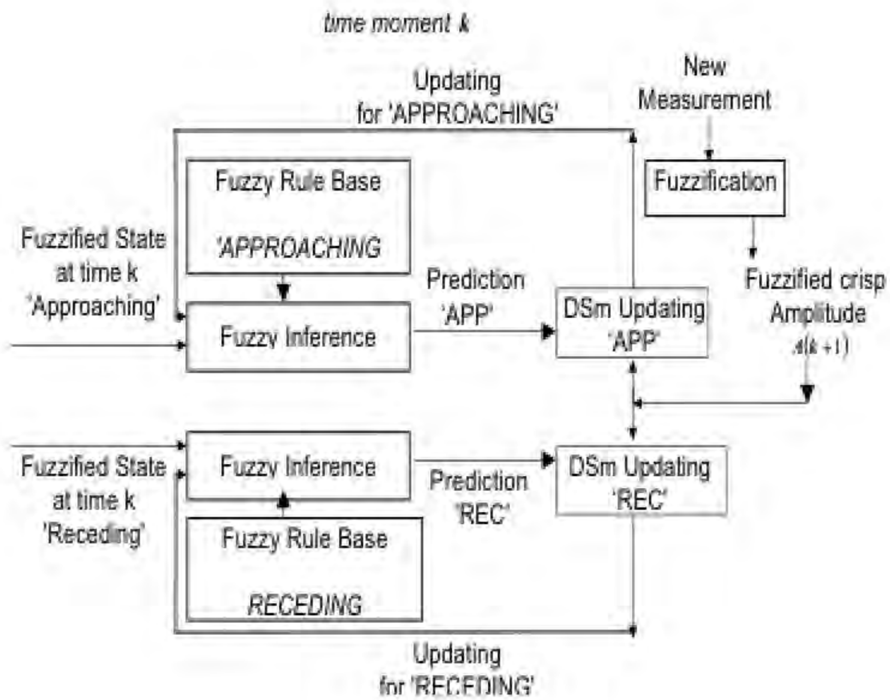


Figure 13.1: Block diagram of target's behavior tracking system

The tendency prediction approach is based on Zadeh compositional rule. The updating procedure uses Dezert-Smarandache classical combination rule based on the free DSm model to estimate target behavior states. Dezert-Smarandache Theory assures a particular framework where the frame of discernment is exhaustive but not necessarily exclusive and it deals successfully with rational, uncertain or paradoxical data. In general this diagram resembles the commonly used approaches in standard tracking systems [1, 2], but the peculiarity consists in the implemented particular approaches in the realizations of the main steps.

13.3 Approach for Behavior Tendency Estimation

There are a few particular basic components in the block diagram of target's behavior tracking system.

13.3.1 The fuzzification interface

A decisive variable in our task is the transmitted from the emitter amplitude value $A(k)$, received at consecutive time moments $k = 1, 2, \ldots$. We use the fuzzification interface (fig. 13.2), that maps it into two fuzzy sets defining two linguistic values in the frame of discernment $\Theta = \{S \triangleq \text{Small}, B \triangleq \text{Big}\}$. Their membership functions are not arbitrarily chosen, but rely on the inverse proportion dependency between the measured amplitude value and corresponding distance to target.

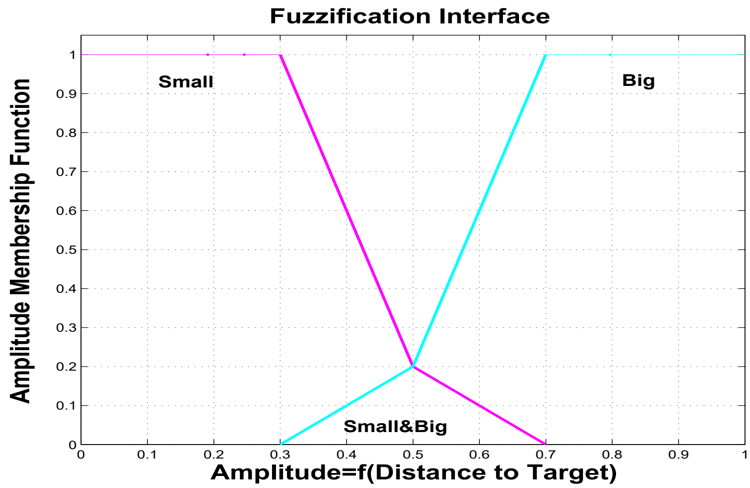


Figure 13.2: Fuzzification Interface

The length of fuzzy sets' bases provide design parameter that we calibrate for satisfactory performance. These functions are tuned in conformity with the particular dependency $A \approx f(1/\delta D)$ known as a priori information The degree of overlap between adjacent fuzzy sets reflects amplitude gradients in the boundary points of specified distance intervals.

13.3.2 The behavior model

In conformity with our task, fuzzy rules' definition is consistent with the tracking of amplitude changes tendency in consecutive time moments $k = 1, 2, \ldots$. With regard to this a particular feature is that considered fuzzy rules have one and the same antecedents and consequents. We define their meaning by using the prespecified in paragraph linguistic terms and associated membership functions (according to paragraph 13.3.1). We consider two essential models of possible target behavior:

Approaching Target - it's behavior is characterized as a stable process of gradually amplitude value increasing, i.e. the transition $S \to S \to B \to B$ is held in a timely manner;

Receding Target - it's behavior is characterized as a stable process of gradually amplitude value decreasing, i.e. the transition $B \to B \to S \to S$ is held in a timely manner.

To comprise appropriately these models the following rule bases have to be carried out:

Behavior Model 1: Approaching Target:

Rule 1: IF $A(k) = S$ THEN $A(k+1) = S$

Rule 2: IF $A(k) = S$ THEN $A(k+1) = B$

Rule 3: IF $A(k) = B$ THEN $A(k+1) = B$

Behavior Model 2: Receding Target:

Rule 1: IF $A(k) = B$ THEN $A(k+1) = B$

Rule 2: IF $A(k) = B$ THEN $A(k+1) = S$

Rule 3: IF $A(k) = S$ THEN $A(k+1) = S$

The inference schemes for these particular fuzzy models are conditioned on the cornerstone principle of each modeling process. It is proven [4], that minimum and product inferences are the most widely used in engineering applications, because they preserve cause and effect. The models are derived as fuzzy graphs:

$$g = \max_i(\mu_{A_i \times B_i}(u,v)) = \max_i(\mu_{A_i}(u) \cdot \mu_{B_i}(v)) \tag{13.1}$$

in which $\mu_{A_i \times B_i}(u,v) = \mu_{A_i}(u) \cdot \mu_{B_i}(v)$ corresponds to the Larsen product operator for the fuzzy conjunction, $g = \max_i(\mu_{A_i \times B_i})$ is the maximum for fuzzy union operator and

$$\mu_{B'}(y) = \max_{x_i}(\min(\mu_{A'}(x_i), \mu_{A \times B}(x_i, y_i)))$$

is the *Zadeh max-min operator for the composition rule.*

The fuzzy graphs related to the two models are obtained in conformity with the above described mathematical interpretations, by using the specified membership functions for linguistic terms *Small*, *Big*, and taking for completeness into account all possible terms in the hyper-power set $D^\Theta = \{S, B, S \cap B, S \cup B\}$:

$k \to k+1$	S	$S \cap B$	B	$S \cup B$
S	1	0	1	0
$S \cap B$	0	0	0	0
B	0.2	0	1	0
$S \cup B$	0	0	0	0

Relation 1: Approaching Target

$k \to k+1$	S	$S \cap B$	B	$S \cup B$
S	1	0	0.2	0
$S \cap B$	0	0	0	0
B	1	0	1	0
$S \cup B$	0	0	0	0

Relation 2: Receding Target

13.3.3 The amplitude state prediction

At initial time moment k the target is characterized by the fuzzified amplitude state estimates according to the models $\mu_{A^{\text{App}}}(k|k)$ and $\mu_{A^{\text{Rec}}}(k|k)$. Using these fuzzy sets and applying the *Zadeh max-min compositional rule* [4] to relation 1 and relation 2, we obtain models' conditioned amplitude state predictions for time $k+1$, i.e. $\mu_{A^{\text{App}}}(k+1|k)$ is given by $max(\min(\mu_{A^{\text{App}}}(k|k), \mu_{\text{App}}(k \rightarrow k+1)))$ and $\mu_{A^{\text{Rec}}}(k+1|k)$ by $max(\min(\mu_{A^{\text{Rec}}}(k|k), \mu_{\text{Rec}}(k \rightarrow k+1)))$.

13.3.4 State updating using DSmT

The classical DSm combinational rule is used here for state updating. This procedure is realized on the base of fusion between predicted states according to the considered models (Approaching, Receding) and the new measurement. Since D^Θ is closed under $\cup$ and $\cap$ operators, to obey the requirements to guarantee that $m(.) : D^\Theta \mapsto [0,1]$ is a proper general information granule, it is necessarily to transform fuzzy membership functions representing the predicted state and new measurement into mass functions. It is realized through their normalization with respect to the unity interval. Models' conditioned amplitude state prediction vector $\mu_{\text{pred}}^{\text{App/Rec}}(.)$ is obtained in the form:

$$[\mu_{\text{pred}}^{\text{A/R}}(S), \mu_{\text{pred}}^{\text{A/R}}(S \cap B), \mu_{\text{pred}}^{\text{A/R}}(B), \mu_{\text{pred}}^{\text{A/R}}(S \cup B)] \tag{13.2}$$

In general the terms, contained in $\mu_{\text{pred}}^{\text{App/Rec}}$ represent the possibilities that the predicted amplitude behavior belongs to the elements of hyper-power set D^Θ and there is no requirement to sum up to unity. In order to use the classical DSm combinational rule, it is necessary to make normalization over $\mu_{\text{pred}}^{\text{App/Rec}}$ to obtain respective generalized basic belief assigments (gbba) $\forall C \in D^\Theta = \{S, S \cap B, B, S \cup B\}$:

$$m_{\text{pred}}^{\text{App/Rec}}(C) = \frac{\mu_{\text{pred}}^{\text{App/Rec}}(C)}{\sum_{A \in D^\Theta} \mu_{\text{pred}}^{\text{App/Rec}}(A)} \tag{13.3}$$

The equivalent normalization has to be made for the received new measurement before being fused with the DSm rule of combination.

Example

Let's consider at scan 3 the predicted vector for the model *Approaching* $\mu_{\text{pred}}^{\text{App/Rec}}(4|3)$ with components $\mu(S) = 0.6$, $\mu(S \cap B) = 0.15$, $\mu(B) = 0.05$ and $\mu(S \cup B) = 0.0$, then the normalization constant is $K = 0.6 + 0.15 + 0.05 + 0.0 = 0.8$ and after normalization, one gets the resulting gbba

$$m_{\text{pred}}^{\text{App/Rec}}(S) = \frac{0.6}{K} = 0.75 \qquad m_{\text{pred}}^{\text{App/Rec}}(S \cap B) = \frac{0.15}{K} = 0.1875$$

$$m_{\text{pred}}^{\text{App/Rec}}(B) = \frac{0.05}{K} = 0.0625 \qquad m_{\text{pred}}^{\text{App/Rec}}(S \cup B) = \frac{0.0}{K} = 0.0$$

That way one can obtain $m_{\text{pred}}^{\text{App/Rec}}(.)$ as a general (normalized) information granule for the prediction of the target's behavior.

The target behavior estimate $m_{\text{upd}}^{\text{App/Rec}}(.)$ at measurement time is then obtained from $m_{\text{pred}}^{\text{App/Rec}}(.)$ and the amplitude belief assignment $m_{\text{mes}}(B)$ (built from the normalization of the new fuzzyfied crisp amplitude measurement received) by the DSm rule of combination, i.e.

$$m_{\text{upd}}^{\text{App/Rec}}(C) = [m_{\text{pred}}^{\text{App/Rec}} \oplus m_{\text{mes}}](C) = \sum_{A,B \in D^{\Theta}, A \cap B = C} m_{\text{pred}}^{\text{App/Rec}}(A) m_{\text{mes}}(B) \tag{13.4}$$

Since in contrast to the DST, DSmT uses a frame of discernment, which is exhaustive, but in general case not exclusive (as it is in our case for $\Theta = \{S, B\}$), we are able to take into account and to utilize the paradoxical information $S \cap B$ although being not precisely defined. This information relates to the case, when the moving target resides in an overlapping intermediate region, when it is hard to predict properly the tendency in its behavior. Thus the conflict management, modeled that way contributes to a better understanding of the target motion and to assure awareness about the behavior tendencies in such cases.

13.4 The decision criterion

It is possible to build for each model M = (A)pproaching, (R)eceding a subjective probability measure $P_{\text{upd}}^{\text{M}}(.)$ from the bba $m_{\text{upd}}^{\text{M}}(.)$ with the generalized pignistic transformation (GPT) [3, 6] defined $\forall A \in D^{\Theta}$ by

$$P_{\text{upd}}^{\text{M}}\{A\} = \sum_{C \in D^{\Theta} | A \cap C \neq \emptyset} \frac{\mathcal{C}_{\mathcal{M}^f}(C \cap A)}{\mathcal{C}_{\mathcal{M}^f}(C)} m_{\text{upd}}^{\text{M}}(C) \tag{13.5}$$

where $\mathcal{C}_{\mathcal{M}^f}(X)$ denotes the DSm cardinal of proposition X for the free DSm model $\mathcal{M}^f$ of the problem under consideration here. The decision criterion for the estimation of correct model M is then based on the evolution of the Pignistic entropies, associated with updated amplitude states:

$$H_{pig}^{\text{M}}(P_{\text{upd}}^{\text{M}}) \triangleq -\sum_{A \in \mathcal{V}} P_{\text{upd}}^{\text{M}}\{A\} \ln(P_{\text{upd}}^{\text{M}}\{A\}) \tag{13.6}$$

where $\mathcal{V}$ denotes the parts of the Venn diagram of the free DSm model $\mathcal{M}^f$. The estimation $\hat{\text{M}}(k)$ of correct model at time k is given by the most informative model corresponding to the smallest value of the pignistic entropy between $H_{pig}^{\text{A}}(P_{\text{upd}}^{\text{A}})$ and $H_{pig}^{\text{R}}(P_{\text{upd}}^{\text{R}})$.

13.5 Simulation study

A non-real time simulation scenario is developed for a single target trajectory (fig.13.3) in plane coordinates X,Y and for constant velocity movement. The tracker is located at position $(0km, 0km)$. The target's starting point and velocities are: $(x_0 = 5km, y_0 = 10km)$, with following velocities during the two part of the trajectory $(\dot{x} = 100m/s, \dot{y} = 100m/s)$ and $(\dot{x} = -100m/s, \dot{y} = -100m/s)$.

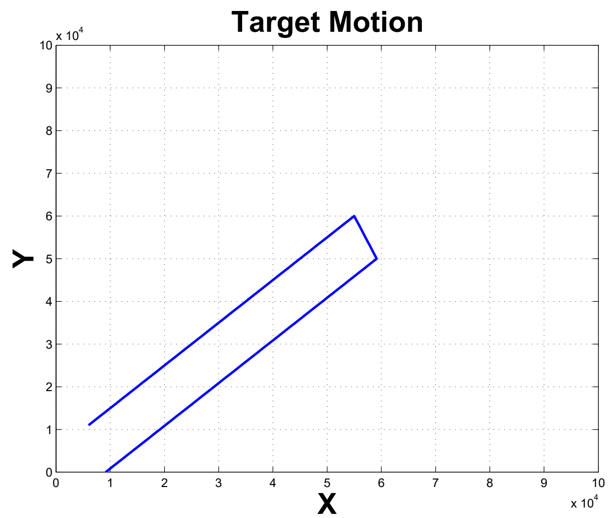


Figure 13.3: Target trajectory.

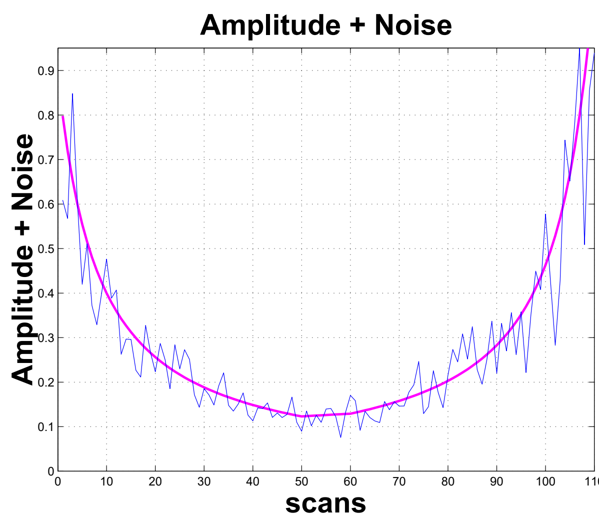


Figure 13.4: Measurements statistics.

The time sampling rate is $T = 10s$. The dynamics of target movement is modeled by equations:

$$x(k) = x(k-1) + \dot{x}T \quad \text{and} \quad y(k) = y(k-1) + \dot{y}T$$

The amplitude value $Z_A(k) = A(k) + \nu_A(k)$ measured by sonar is a random Gaussian distributed process with mean $A(k) = 1/D(k)$ and covariance $\sigma_A(k)$ (fig. 13.4). $D(k) = \sqrt{x^2(k) + y^2(k)}$ is the distance to the target, $(x(k), y(k))$ is the corresponding vector of coordinates, and $\nu_A(k)$ is the measurement noise. Each amplitude value (true one and the corresponding noisy one) received at each scan is processed according to the block diagram (figure 13.1).

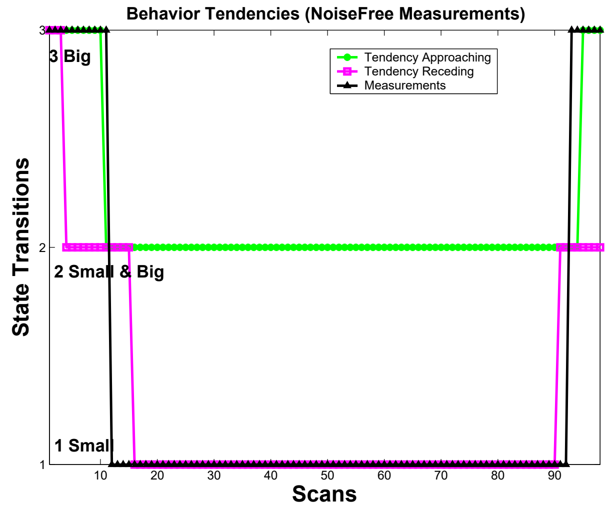


Figure 13.5: Behavior tendencies (Noise-free measurements).

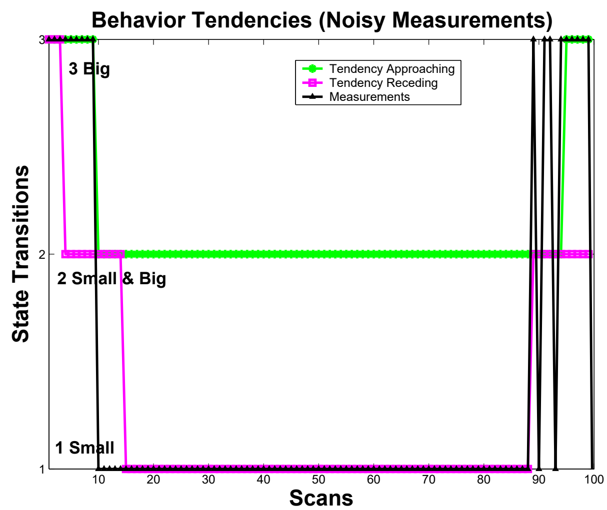


Figure 13.6: Behavior Tendencies (Noisy measurements).

Figures 13.5 and 13.6 show the results obtained during the whole motion of the observed target. Figure 13.5 represents the case when the measurements are without noise, i.e. $Z(k) = A(k)$. Figure 13.6 represents the case when measured amplitude values are corrupted by noise. In general the presented graphics show the estimated tendencies in target behavior, which are described via the scan consecutive transitions of the estimated amplitude states.

Figure 13.7 represents the evolution of pignistic entropies associated with updated amplitude states for the Approaching and Receding models in case of noisy measurements; the figure for the noise-free measurement is similar. It illustrates the decision criterion used to choose the correct model. If one takes a look at the figure 13.5 and figure 13.7, it can be seen that between scans 1st and 15th the target motion is supported by *Approaching* model, because that mode corresponds to the minimum entropies values, which means that it is the more informative one.

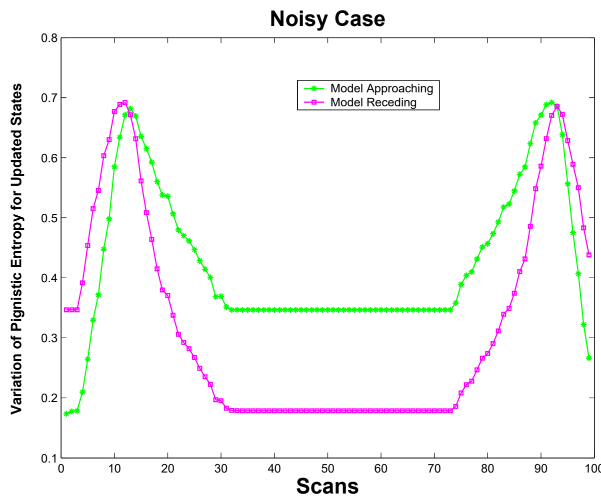


Figure 13.7: Evolution of the pignistic entropy for updated states.

The *Approaching* model is dominant, because the measured amplitude values during these scans stable reside in the state *Big*, as it is obvious from the fuzzification interface (fig.13.2). In the same time, *Receding* model supports the overlapping region $S \cap B$, which is transition towards the state *Small*. Between scans 16th and 90th the *Receding* model becomes dominant since the variations of amplitude changes are minimal and their amplitude values stable support the state *Small*. During these scans *Approaching* model has a small reaction to the measurement statistics, keeping paradoxical state $S \cap B$.What it is interesting and important to note is that between scans 16th and 30th the difference of entropies between *Approaching* and *Receding* models increases, a fact, that makes us to be increasingly sure that the *Receding* mode is becoming dominant. Then, between scans 75th and 90th the difference of these entropies is decreasing, which means that we are less and less sure, that *Receding* model remain still dominant. After switching scan 91th the *Approaching* model becomes dominant one, until scan 100th. In general the reaction of the considered models to the changes of target motion is not immediate, because the whole behavior estimation procedure deals with vague propositions *Small*, *Big*, and sequences of amplitude values at consecutive scans often reside stable in one and the same states.

Comparing the results in figure 13.6 with the results in figure 13.5, it is evident, that although some disorder in the estimated behavior tendencies, one can make approximately correct decision due to the possibility of DSmT to deal with conflicts and that way to contribute for a better understanding of target behavior and evaluation of the threat.

13.6 Comparison between DSm and Fuzzy Logic Approaches

The objective of this section is to compare the results received by using DSm theory and respective results but drawn from the Fuzzy Logic Approach (FLA) [4, 8, 9], applied on the same simulation scenario. The main differences between the two approaches consist in the domain of considered working propositions and in the updating procedure as well. In present work, we use DSm combination rule to fuse the predicted state and the new measurement to obtain the estimated behavior states, while in the fuzzy approach state estimates are obtained through a fuzzy set intersection between these entities. It is evident from the results, shown in figures 13.8 and 13.9, that here we deal with only two propositions $\Theta = \{\text{Small}, \text{Big}\}$. There is no way to examine the behavior tendencies in the overlapping region, keeping into considerations every one of possible target's movements: from $S \cap B$ to B or from $S \cap B$ to S.

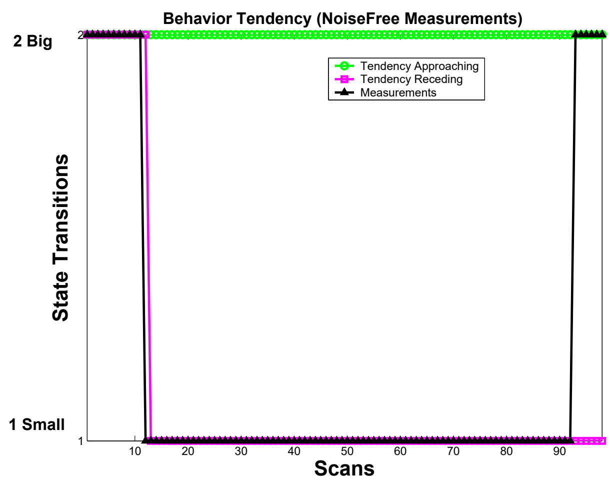


Figure 13.8: Behavior Tendencies drawn from FLA (NoisyFree Measurements).

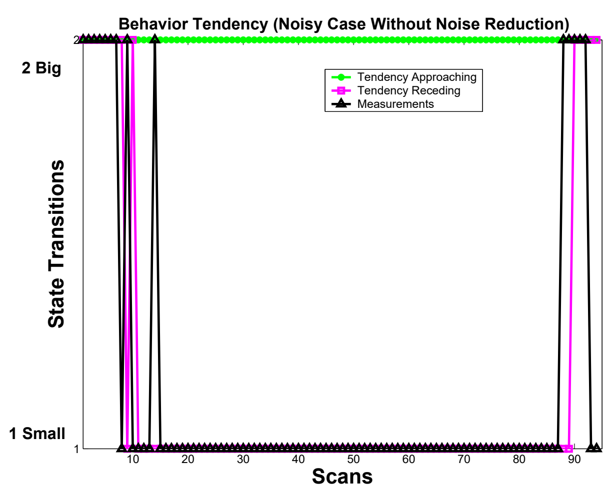


Figure 13.9: Behavior Tendencies without Noise Reduction drawn from FLA (Noisy Case).

Figure 13.8 shows the noise-free measurement case. It could be seen that between scan 10 and 90 target motion is supported by the correct for that case Receding model, while Approaching one has no reaction at all. If we compare corresponding figure 13.5 (DSm case) and present figure 13.8, we can see, that in the case of DSm approach Receding model reacts more adequately to the true target tendency , because there is a possibility to deal with the real situation – the tendency of the target to make a movement from B to the overlapping region $B \cap S$. In the FLA case there is no such opportunity and because of that between scan 1st and 10th Receding model has no reaction to the real target movement towards the $B \cap S$. Figure 13.9 represents the case when the measured amplitude values are corrupted by noise. It is difficult to make proper decision about the behavior tendency, especially after scan 90th., because it is obvious, that here the model *Approaching* coincide with the model *Receding*. In order to reduce the influence of measurement noise over tendency estimation, an additional noise reduction procedure has to be applied to make the measurements more informative. Its application improves the overall process of behavior estimation. Taking in mind all the results drawn from DSmT and FLA application, we can make the following considerations:

- DSmT and FLA deal with a frame of discernment, based in general on imprecise/vague notions and concepts $\Theta = \{S, B\}$. But DSmT allows us to deal also with uncertain and/or paradoxical data, operating on the hyper-power set $D^\Theta = \{S, S \cap B, B, S \cup B\}$. In our particular application it gives us an opportunity for flexible tracking the changes of possible target behavior during the overlapping region $S \cap B$.

- DSmT based behavior estimates can be characterized as a noise resistant, while FLA uses an additional noise reduction procedure to produce 'smoothed' behavior estimates.

13.7 Conclusions

An approach for estimating the tendency of target behavior was proposed. It is based on Dezert-Smarandache theory applied to conventional sonar measurements. It was evaluated using computer simulation. The provided example illustrates the benefits of DSm approach in comparison of fuzzy logic one. Dealing simultaneously with uncertain and paradoxical data, an opportunity for flexible and robust reasoning is realized, overcoming the described limitations relative to the fuzzy logic approach. It is presented and approved the ability of DSmT to ensure reasonable and successful decision-making procedure about the tendencies of target behavior in case of discrepancies of angle-only measurements interpretation. The proposed approach yields confident picture for complex and ill-defined engineering problems.

13.8 References

[1] Bar-Shalom Y. (Ed.), *Multitarget-Multisensor Tracking: Advanced Applications*, Norwood, MA, Artech House, 1990.

[2] Blackman S., Popoli R., *Design and Analysis of Modern Tracking*, Artech House, 1999.

[3] Dezert J., Smarandache F., Daniel M., *The Generalized Pignistic Transformation*, Proc. of Fusion 2004 Conf., Stockholm, Sweden, June, 2004.

[4] Mendel J.M., *Fuzzy Logic Systems for Egineering: A Tutorial*, Proc. of the IEEE, pp. 345-377, March 1995.

[5] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[6] Smarandache F. Dezert J. (Editors), *Advances and Applications of DSmT for Information Fusion (Collected Works)*, American Research Press, 2004.

[7] Tchamova A., Semerdjiev T., Dezert J., *Estimation of Target Behavior Tendencies using Dezert-Smarandache Theory*, Proc. of Fusion 2003 Conf., Cairns, Australia, July 8-11, 2003.

[8] Zadeh L., *Fuzzy Sets as a Basis for a Theory of Possibility*, Fuzzy Sets and Syst. 1, pp.3-28, 1978.

[9] Zadeh L., *From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions*, IEEE Trans. on Circuits and Systems 45,1, pp.105-119, Jan. 1999.

Chapter 14

Generalized Data Association for Multitarget Tracking in Clutter

A. Tchamova, T. Semerdjiev, P. Konstantinova
Institute for Parallel Processing
Bulgarian Academy of Sciences
"Acad. G. Bonchev" Str.,bl.25-A, 1113 Sofia
Bulgaria

J. Dezert
ONERA
29 Av. de la Division Leclerc
92320 Châtillon
France

Abstract: *The objective of this chapter is to present an approach for target tracking in cluttered environment, which incorporates the advanced concept of generalized data (kinematics and attribute) association (GDA) to improve track maintenance performance in complicated situations (closely spaced and/or crossing targets), when kinematics data are insufficient for correct decision making. It uses Global Nearest Neighbour-like approach and Munkres algorithm to resolve the generalized association matrix. The main peculiarity consists in applying the principles of Dezert-Smarandache theory (DSmT) of plausible and paradoxical reasoning to model and process the utilized attribute data. The new general Dezert-Smarandache hybrid rule of combination is used to deal with particular integrity constraints associated with some elements of the free distributive lattice. The aim of the performed study is to provide coherent decision making process related to generalized data association and to improve the overall tracking performance. A comparison with the corresponding results, obtained via Dempster-Shafer theory is made.*

This work has been partially supported by MONT grants I-1205/02, I-1202/02 and by Center of Excellence BIS21 grant ICA1-2000-70016

14.1 Introduction

One important function of each radar surveillance system in cluttered environment is to keep and improve targets' tracks maintenance performance. It becomes a crucial and challenging problem especially in complicated situations of closely spaced, and/or crossing targets. The design of a modern multitarget tracking (MTT) algorithms in a such real-life stressful environment motivates the incorporation of the advanced concepts for generalized data association. In order to resolve correlation ambiguities and to select the best observation-track pairings, in this study, a particular generalized data association (GDA) approach is proposed and incorporated in a MTT algorithm. It allows the introduction of target attribute into the association logic, based on the general Dezert-Smarandache rule for combination, which is adapted to deal with possible integrity constraints on the problem under consideration due to the true nature of the elements involved into it. This chapter extends recent research work published in [15] which was limited to target tracking in clutter-free environment.

14.2 Basic Elements of Tracking Process

The tracking process consists of two basic elements: *data association* and *track filtering*. The first element is often considered as the most important. Its goal is to associate observations to existing tracks.

14.2.1 Data Association

To eliminate unlikely observation-to-track pairing at the begining a validation region (gate) is formed around the predicted track position. The measurements in the gate are candidates for association to the corresponding track.

14.2.1.1 Gating

We assume zero-mean Gaussian white noise for measurements. The vector difference between received measurement vector $\mathbf{z}_j(k)$ and predicted measurement vector $\hat{\mathbf{z}}_i(k|k-1)$ of target i is defined to be residual vector (called innovation)

$$\tilde{\mathbf{z}}_{ij}(k) = \mathbf{z}_j(k) - \hat{\mathbf{z}}_i(k|k-1)$$

with residual covariance matrix $\mathbf{S} = \mathbf{HPH}' + \mathbf{R}$, where $\mathbf{P}$ is the state prediction covariance matrix, $\mathbf{H}$ is the measurement matrix and $\mathbf{R}$ is the measurement covariance matrix [2, 3, 4, 5]. The scan indexes k will be dropped for notational convenience. The norm (normalized distance function) of the innovation is evaluated as:

$$d_{ij}^2 = \tilde{\mathbf{z}}'_{ij}\mathbf{S}^{-1}\tilde{\mathbf{z}}_{ij}$$

One defines a threshold constant for gate γ such that correlation is allowed if the following relationship is satisfied

$$d_{ij}^2 \leq \gamma \tag{14.1}$$

Assume that the measurement vector size is M. The quantity d_{ij}^2 is the sum of the squares of M independent Gaussian random variables with zero means and unit standard deviations. For that reason d_{ij}^2 will have χ_M^2 distribution with M degrees of freedom and allowable probability of a valid observation falling outside the gate. The threshold constant γ can be defined from the table of the chi-square (χ_M^2) distribution [3].

14.2.1.2 Generalized Data Association (GDA)

If a single observation is within a gate and if that observation is not within a gate of any other track, the observation can be associated with this track and used to update the track filter. But in a dense target environment additional logic is required when an observation falls within the gates of multiple target tracks or when multiple observations fall within the gate of a target track.

When attribute data are available, the generalized probability can be used to improve the assignment. In view of independence of the kinematic and attribute measurement errors, the generalized probability for measurement j originating from track i is:

$$P_{\mathrm{gen}}(i,j) = P_{\mathrm{k}}(i,j)P_{\mathrm{a}}(i,j)$$

where $P_{\mathrm{k}}(i,j)$ and $P_{\mathrm{a}}(i,j)$ are kinematic and attribute probability terms respectively.

Our goal is to choose a set of assignments $\{\chi_{ij}\}$, for $i = 1, \ldots n$ and $j = 1, \ldots, m$, that assures maximum of the total generalized probability sum. To find it, we use the solution of the assignment problem

$$\min \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\chi_{ij}$$

where:

$$\chi_{ij} = \begin{cases} 1 & \text{if measurement } j \text{ is assigned to track } i \text{ according to assignment problem solution} \\ 0 & \text{otherwise} \end{cases}$$

If, in the attempt to maximize the number of assignments, the assignment algorithm chooses a pairing that does not satisfy the gate, the assignment is later removed.

Because our probabilities vary $0 \leq P_{\mathrm{k}}(i,j), P_{\mathrm{a}}(i,j) \leq 1$ and to satisfy the condition to be minimized, the elements of the particular assignment matrix are defined as :

$$a_{ij} = 1 - P_{\mathrm{gen}}(i,j) = 1 - P_{\mathrm{k}}(i,j)P_{\mathrm{a}}(i,j)$$

14.2.2 Filtering

The used tracking filter is the first order extended Kalman filter [7] for *target state vector* $\mathbf{x} = [x\,\dot{x}\,y\,\dot{y}]'$, where x and y are Cartesian coordinates and $\dot{x}$ and $\dot{y}$ are velocities along Cartesian axes and *measurement vector* $\mathbf{z} = [\beta\,D]'$, where β is the azimuth (measured from the North), and D is the distance from the observer to the target under consideration.

The measurement function $\mathbf{h}(.)$ is (assuming the sensor located at position (0,0)):

$$\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x})\,h_2(\mathbf{x})]' = [\arctan(\frac{x}{y})\,\sqrt{x^2+y^2}]'$$

and the Jacobian [3]:

$$\mathbf{H} = [H_{ij}] = [\partial h_i/\partial x_j] \quad i = 1,2 \quad j = 1,\ldots,4$$

We assume constant velocity target model. The process noise covariance matrix is: $\mathbf{Q} = \sigma_v^2 \mathbf{Q}_T$, where T is the sampling/scanning period, σ_v is standard deviation of the process noise and $\mathbf{Q}_T$ is given by [8]:

$$\mathbf{Q}_T = diag(\mathbf{Q}_{2\times 2}, \mathbf{Q}_{2\times 2}) \quad \text{with} \quad \mathbf{Q}_{2\times 2} = \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} \\ \frac{T^3}{2} & T^2 \end{bmatrix}$$

The measurement error matrix is $\mathbf{R} = diag(\sigma_\beta^2, \sigma_D^2)$ where σ_β and σ_D are the standard deviations of measurement errors for azimuth and distance.

The track initiation is performed by two-point differencing [7]. After receiving observations for first two scans the initial state vector is estimated by $\hat{\mathbf{x}} = [x(2)\,\frac{x(2)-x(1)}{T}\,y(2)\,\frac{y(2)-y(1)}{T}]'$ where $(x(1), y(1))$ and $(x(2), y(2))$ are respectively the target positions at the first scan for time stamp $k = 1$, and at the second scan for $k = 2$. The initial (starting at time stamp $k = 2$) state covariance matrix $\mathbf{P}$ is evaluated by:

$$\mathbf{P} = diag(\mathbf{P}^x_{2\times 2}, \mathbf{P}^y_{2\times 2}) \quad \text{with} \quad \mathbf{P}^{(.)}_{2\times 2} = \begin{bmatrix} \sigma^2_{(.)} & \frac{\sigma^2_{(.)}}{T} \\ \frac{\sigma^2_{(.)}}{T} & \frac{2\sigma^2_{(.)}}{T^2} \end{bmatrix}$$

where the index (.) must be replaced by either x or y indexes with $\sigma_x^2 \approx \sigma_D^2 \sin^2(z_\beta) + z_D^2 \sigma_\beta^2 \cos^2(z_\beta)$ and $\sigma_y^2 \approx \sigma_D^2 \cos^2(z_\beta) + z_D^2 \sigma_\beta^2 \sin^2(z_\beta)$. z_β and z_D are the components of the measurement vector received at scan $k = 2$, i.e. $\mathbf{z} = [z_\beta\,z_D]' = \mathbf{h}(\mathbf{x}) + \mathbf{w}$ with $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$.

14.3 The Attribute Contribution to GDA

Data association with its goal of partitioning observations into tracks is a key function of any surveillance system. An advanced tendency is the incorporation of generalized data (kinematics and attribute) association to improve track maintenance performance in complicated situations, when kinematics data are insufficient for coherent decision making process. Analogously with the kinematic tracking, the attribute tracking can be considered as the process of combining information collected over time from one or more sensors to refine the knowledge about the evolving attributes of the targets. The motivation for attribute fusion is inspired from the necessity to ascertain the targets' types, information, that in consequence has an important implication to enhance the tracking performance. A number of techniques, probabilistic in nature are available for attribute fusion. Their analysis led us to belief, that the theory of Dempster-Shafer is well suited for representing uncertainty, but especially in case of low conflicts between the bodies of evidence. When the conflict increases and becomes high, (case, which often occurs in data association process) the combinational rule of Dempster hides the risk to produce indefiniteness. To avoid that significant risk we consider the form of attribute likelihood function within the context of DSm theory, i.e. the term to be used for computing the probabilities of validity for data association hypotheses. There are a few basic steps, realizing the concept of attribute data association.

14.3.1 The Input Fuzzification Interface

Fuzzification interface (see fig. 14.1) transforms numerical measurement received from a sensor into fuzzy set in accordance with the a priori defined fuzzy partition of input space-the frame of discernments Θ. This frame includes all considered linguistic values related to the chosen particular input variable and their corresponding membership functions. The fuzzification of numerical sensory data needs dividing an optimal membership into a suitable number of fuzzy sets [14]. Such division provides smooth transitions and overlaps among the associated fuzzy sets, according to the particular real world situation.

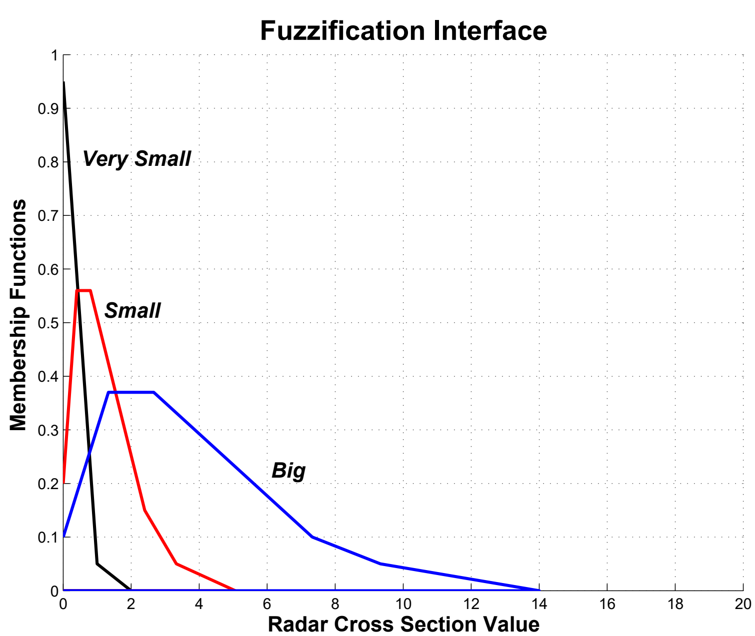


Figure 14.1: Fuzzification interface

The considerable input variable in the particular case is the Radar Cross Section (RCS) of the observed targets. In our work the modeled RCS data are analyzed to determine the target size with the subsequent declaration that the observed target is an aircraft of specified type (Fighter, Cargo) or False Alarms. Taking it in mind, we define two frames of discernments: first one according to the **size of RCS**: $\Theta_1 = \{\text{Very Small (VS)}, \text{Small (S)}, \text{Big (B)}\}$ and the second one determining the corresponding to its **Target Type** $\Theta_2 = \{\text{False Alarms (FA)}, \text{Fighter (F)}, \text{Cargo (C)}\}$.

The radar cross section according to the real targets is modeled as Swerling 3 type, where the density function for the RCS σ is given by:

$$f(\sigma) = \frac{4\sigma}{\sigma_{\text{ave}}^2} \exp[-\frac{2\sigma}{\sigma_{\text{ave}}}]$$

with the average RCS (σ_{ave}) varying between different targets' types [16]. The cumulative distribution function of the radar cross section is given by

$$F(\sigma_0) = P\{0 \leq \sigma \leq \sigma_0\} = 1 - (1 + \frac{2\sigma_0}{\sigma_{\text{ave}}}) \exp[-\frac{2\sigma_0}{\sigma_{\text{ave}}}]$$

Since the probabilities $F(\sigma_0)$ for having different values of radar cross section are uniformly distributed in the interval $[0, 1]$ over time (i.e. these values are uncorrelated in time), a sample of observation of the RCS can be simulated by solving equation:

$$(1 + \frac{2\sigma_0}{\sigma_{\text{ave}}}) \exp[-\frac{2\sigma_0}{\sigma_{\text{ave}}}] = 1 - x$$

where x is a random number that is uniformly distributed between 0 and 1.

The scenario considered in our work deals with targets' types Fighter (F) and Military Cargo (C) with an average RCS :

$$\sigma_{\text{ave}}^{\text{F}} = 1.2\, m^2 \qquad \text{and} \qquad \sigma_{\text{ave}}^{\text{C}} = 4\, m^2$$

The radar cross section according to the False Alarms [1] is modeled as Swerling 2 type, where the density function for the RCS is given by:

$$f(\sigma) = \frac{1}{\sigma_{\text{ave}}} \exp[-\frac{\sigma}{\sigma_{\text{ave}}}] \qquad \text{with} \qquad \sigma_{\text{ave}} = 0.3\, m^2$$

The cumulative distribution function is given by

$$F(\sigma_0) = P\{0 \leq \sigma \leq \sigma_0\} = 1 - \exp[-\frac{\sigma_0}{\sigma_{\text{ave}}}]$$

A sample of observation of the RCS can be computed by solving equation:

$$\exp[-\frac{\sigma_0}{\sigma_{\text{ave}}}] = 1 - x$$

where x is a random number that is uniformly distributed between 0 and 1.

The input fuzzification interface maps the current modeled RCS values into three fuzzy sets: **VerySmall**, **Small** and **Big**, which define the corresponding linguistic values, defining the variable "RCS". Their membership functions are not arbitrarily chosen, but rely on the calculated respective histograms for 10000 Monte Carlo runs. Actually these fuzzy sets form the frame of discernements Θ_1. After fuzzification the new RCS value (rcs) is obtained in the form :

$$\text{rcs} \Rightarrow [\mu_{\text{VerySmall}}(\text{rcs}), \mu_{\text{Small}}(\text{rcs}), \mu_{\text{Big}}(\text{rcs})]$$

In general, the grades $\mu_{\text{VerySmall}}(\text{rcs})$, $\mu_{\text{Small}}(\text{rcs})$, $\mu_{\text{Big}}(\text{rcs})$ represent the possibilities the new RCS value to belong to the elements of the frame Θ_1 and there is no requirement to sum up to unity. Figure 14.2 below shows the way which the new observations for Cargo, Fighter and False Alarms are modeled for 500 Monte Carlo runs, using the corresponding Swerling type functions type 3 and 2. It is evident that they are too much mixed. It influences over the distinction between them. That fact hides the possibility of intrinsic conflicts between the fused bodies of evidence (general basic belief assignment (gbba) of targets' tracks and observations), because of their imprecise belief functions and consequently yields a poor targets tracks' performance. To deal successfuly with such kind of stressful, but real situation, we need DSm theory to process flexibly and adequately these conflicts.

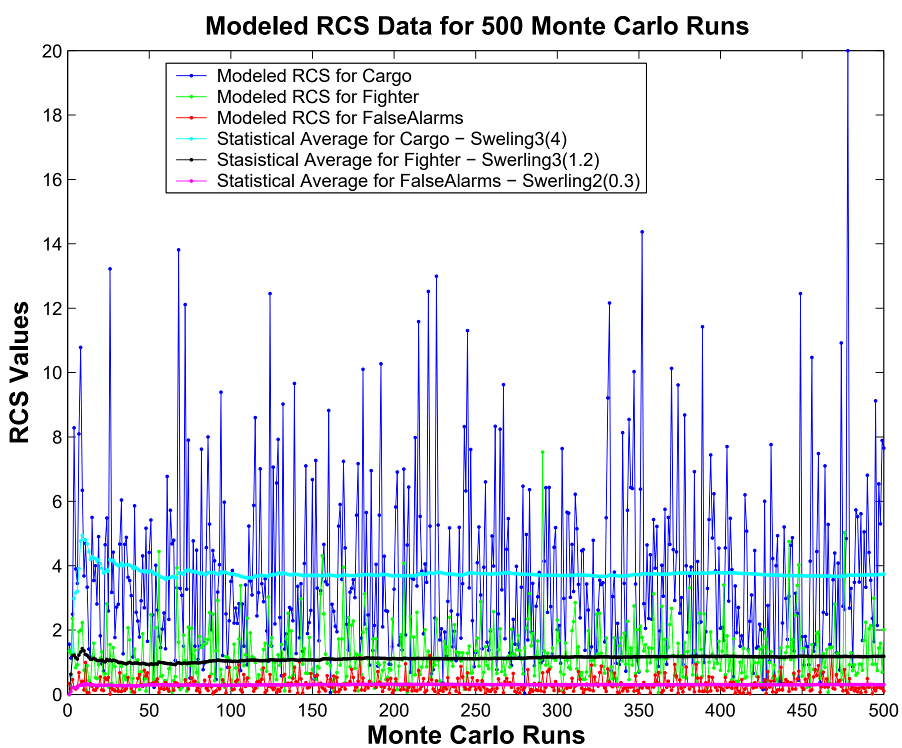


Figure 14.2: Simulation of RCS values over 500 Monte Carlo runs

14.3.2 Tracks' Updating Procedures

14.3.2.1 Using Classical DSm Combinational Rule

After receiving the new observations, detected during the current scan k, to obey the requirements to guarantee that their particular belief assignment $m(.)$ are general information granules, it is necessary to transform each measurement's set of fuzzy membership grades into the corresponding mass function, before being fused. It is realized through normalization with respect to the unity:

$$m_{\text{meas}}(C) = \frac{\mu_C(\text{rcs})}{\sum_{C\in\Theta_1}\mu_C(\text{rcs})}, \quad \forall C \in \Theta_1 = \{\text{VS}, \text{S}, \text{B}\}$$

The general basic belief assignments (gbba) of tracks' histories are described in terms of the hyper-power set :

$$\begin{aligned} D^{\Theta_1} = \{&\emptyset, \text{VS}, \text{S}, \text{B}, \text{VS}\cap\text{S}\cap\text{B}, \text{VS}\cap\text{S}, \text{VS}\cap\text{B}, \text{S}\cap\text{B}, (\text{VS}\cup\text{S})\cap\text{B}, (\text{VS}\cup\text{B})\cap\text{S},\\ &(\text{S}\cup\text{B})\cap\text{VS}, (\text{VS}\cap\text{S})\cup(\text{VS}\cap\text{B})\cup(\text{S}\cap\text{B}), (\text{VS}\cap\text{S})\cup\text{B}, (\text{VS}\cap\text{B})\cup\text{S},\\ &(\text{S}\cap\text{B})\cup\text{VS}, \text{VS}\cup\text{S}, \text{VS}\cup\text{B}, \text{S}\cup\text{B}, \text{VS}\cup\text{S}\cup\text{B}\}\end{aligned}$$

Then DSm classical combinational rule (see chapter 1) is used for tracks' updating:

$$m^{ij}_{\text{upd}}(C) = [m^i_{\text{hist}} \oplus m^j_{\text{meas}}](C) = \sum_{A,B\in D^{\Theta_1}, A\cap B=C} m^i_{\text{hist}}(A)m^j_{\text{meas}}(B)$$

where $m^{ij}_{\text{upd}}(.)$ represents the gbba of the updated track i with the new observation j; m^i_{hist}, m^j_{meas} are respectively gbba vectors of track's i history and the new observation j.

It is important to note, that for us the two considered independent sources of information are the tracks' histories and the new observations with their gbbas maintained in terms of the two hyper-power sets. That way we assure to obtain and to keep the decisions according to the target types during all the scans.

Since, DSmT uses a frame of discernment, which is exhaustive, but in general case not exclusive, we are able to take into account and to utilize the paradoxical information $\text{VS}\cap\text{S}\cap\text{B}$, $\text{VS}\cap\text{S}$,$\text{VS}\cap\text{B}$ and $\text{S}\cap\text{B}$. This information relates to the cases, when the RCS value resides in an overlapping regions, when it is hard to make proper judgement about the tendency of behavior of its value. Actually these nonempty sets and related to it mass assignments contribute to a better understanding of the overall tracking process.

14.3.2.2 Using Hybrid DSm Combinational Rule

As it was mentioned above in our work, RCS data here are used to analyze and subsequently to determine the specified type of the observed targets. Because of this it is maintained the second frame of discernement $\Theta_2 = \{\text{False Alarm (FA)}, \text{Fighter (F)}, \text{Cargo (C)}\}$, in terms of which the decisions according to target types have to be made. Doing this, we take in mind the following correspondencies:

- If rcs is **Very Small** *then* the "target" is **False Alarm**
- If rcs is **Small** *then* the target is **Fighter**
- If rcs is **Big** *then* the target is **Cargo**

We may transform the gbba of updated tracks, formed in D^{Θ_1} into respective gbba in D^{Θ_2}, i.e:

$$m^{ij}_{\text{upd}}(C_{C\in D^{\Theta_2}}) = m^{ij}_{\text{upd}}(C_{C\in D^{\Theta_1}})$$

But let us go deeper into the meaning of the propositions in the second hyper-power set. It should be:

$$\begin{aligned}D^{\Theta_2} = \{&\emptyset, \text{FA}, \text{F}, \text{C}, \text{FA}\cap\text{F}\cap\text{C}, \text{FA}\cap\text{F}, \text{FA}\cap\text{C}, \text{F}\cap\text{C}, (\text{FA}\cup\text{F})\cap\text{C}, (\text{FA}\cup\text{C})\cap\text{F},\\ &(\text{F}\cup\text{C})\cap\text{FA}, (\text{FA}\cap\text{F})\cup(\text{FA}\cap\text{C})\cup(\text{F}\cap\text{C}), (\text{FA}\cap\text{F})\cup\text{C}, (\text{FA}\cap\text{C})\cup\text{F},\\ &(\text{F}\cap\text{C})\cup\text{FA}, \text{FA}\cup\text{F}, \text{FA}\cup\text{C}, \text{F}\cup\text{C}, \text{FA}\cup\text{F}\cup\text{C}\}\end{aligned}$$

In the real life however, it is a proven fact, that the target can not be in one and the same time FalseAlarm and Fighter; FalseAlarm and Cargo; Fighter and Cargo; FalseAlarm and Fighter and Cargo. It leads to the following hybrid DSm model $\mathcal{M}_1(\Theta_2)$, built by introducing the following exclusivity constraints (see chapter 4 for a detailed presentation of the hybrid DSm models and the hybrid DSm rule of combination):

$$\text{FA}\cap\text{F} \overset{\mathcal{M}_1}{\equiv} \emptyset \qquad \text{FA}\cap\text{C} \overset{\mathcal{M}_1}{\equiv} \emptyset \qquad \text{F}\cap\text{C} \overset{\mathcal{M}_1}{\equiv} \emptyset \qquad \text{FA}\cap\text{F}\cap\text{C} \overset{\mathcal{M}_1}{\equiv} \emptyset$$

These exclusivity constraints imply directly the following ones:

$$\begin{aligned}(\text{FA}\cup\text{F})\cap\text{C} &\overset{\mathcal{M}_1}{\equiv} \emptyset \qquad & (\text{FA}\cap\text{F})\cup\text{C} &\overset{\mathcal{M}_1}{\equiv} C\\ (\text{FA}\cup\text{C})\cap\text{F} &\overset{\mathcal{M}_1}{\equiv} \emptyset \qquad & (\text{FA}\cap\text{C})\cup\text{F} &\overset{\mathcal{M}_1}{\equiv} F\\ (\text{F}\cup\text{C})\cap\text{FA} &\overset{\mathcal{M}_1}{\equiv} \emptyset \qquad & (\text{F}\cap\text{C})\cup\text{FA} &\overset{\mathcal{M}_1}{\equiv} FA\end{aligned}$$

and also the more generalized one

$$(\text{FA}\cap\text{F})\cup(\text{FA}\cap\text{C})\cup(\text{F}\cap\text{C}) \overset{\mathcal{M}_1}{\equiv} \emptyset$$

The obtained that way model corresponds actually to Shafer's model, which can be considered as a particular case of the generalized free DSm model.

Therefore, while the corresponding sets in D^{Θ_1} are usually non empty, because of the exclusivity constraints, in the second frame Θ_2, the hyper-power set D^{Θ_2} is reduced to classical power set :

$$D^{\Theta_2}_{\mathcal{M}_1} = \{\emptyset, \mathrm{FA}, \mathrm{F}, \mathrm{C}, \mathrm{FA}\cup\mathrm{F}, \mathrm{FA}\cup\mathrm{C}, \mathrm{F}\cup\mathrm{C}, \mathrm{FA}\cup\mathrm{F}\cup\mathrm{C}\}$$

So, we have to update the previous fusion result, obtained via the classical DSm rule of combination with this new information on the model $\mathcal{M}_1(\Theta_2)$ of the considered problem. It is solved with the hybrid DSm rule (see chapter 4), which transfers the mass of these empty sets to the non-empty sets of $D^{\Theta_2}_{\mathcal{M}_1}$.

14.4 The Generalized Data Association Algorithm

We consider a particular cluster and assume the existence of a set of n tracks at the current scan and a set of m received observations. A validated measurement is one which is either inside or on the boundary of the validation gate of a target. The inequality given in (14.1) is a validation test. It is used for filling the assignment matrix $\mathbf{A}$:

$$\mathbf{A} = [A_{ij}] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \vdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \vdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \vdots & a_{nm} \end{bmatrix}$$

The elements of the assignment matrix $\mathbf{A}$ have the following values [13]:

$$a_{ij} = \begin{cases} \infty & \text{if } d^2_{ij} > \gamma \\ 1 - P_{\mathrm{k}}(i,j)P_{\mathrm{a}}(i,j) & \text{if } d^2_{ij} \leq \gamma \end{cases}$$

The solution of the assignment matrix is the one that minimizes the sum of the choosen elements. We solve the assignment problem by realizing the extension of Munkres algorithm, given in [10]. As a result, it obtains the optimal measurements to tracks association. Because of the considered crossing and/or closely spaced target scenarios, to produce the probability terms P_{k} and P_{a}, the joint probabilistic approach is used [7]. It assures a common base for their defining, making that way them to be compatible.The joint probabilistic data association (JPDA) approach imposes restriction on the problem size because of exponential increasing of the number of generated hypotheses and the time for assignment problem solution. That's why it is advisable to make clustering before solving data association problem. Cluster is a set of closely spaced objects. In our case if two tracks have an observation in their overlapping parts of the gates, the tracks form cluster i.e. their clusters are merged. In such a way the number of clusters are equal or less than the number of tracked tracks. The clustering is usefull at least for two reasons:

1. In such a way the size of assignment matrix and also the time for its solution decreases;

2. The number of hypotheses for JPDA like approach for defining kinematic and attribute probabilities also decreases.

In the worst case when all m measurements fall in the intersection of the validation regions of all n tracks, the number of hypotheses can be obtained as:

$$S(n,m) = \sum_{i=0}^{\min(n,m)} C_m^i A_n^i$$

where

$$C_m^i \triangleq \frac{m!}{i!(m-i)!} \text{ for } 0 \le i \le m \qquad \text{and} \qquad A_n^i \triangleq \frac{n!}{(n-i)!} \text{ for } 0 \le i \le n$$

With these formulae the number of hypotheses for various values of the m and n are computed and are shown in the following table. The enormous increasing of the number of hypothesis can be seen.

	Hyp. #		Hyp. #		Hyp. #
$n=2, m=2$	7	$n=4, m=4$	209	$n=6, m=6$	13327
$n=2, m=3$	13	$n=4, m=5$	501	$n=7, m=8$	394353
$n=3, m=3$	34	$n=5, m=5$	1546	$n=10, m=9$	58941091
$n=3, m=4$	73	$n=5, m=6$	4051	$n=10, m=10$	234662231

Table 14.1: Worst case hypotheses number

As further improvement, first k-best hypotheses can be used [12] as the score of the hypotheses decrease and a big amount of hypotheses practically does not influence the result. Another original frame of hypotheses generation has been considerably optimized in [9] and that way it becomes a practical alternative of Murty's approach.

To define the probabilities for data association for different scenarios with random number of false alarms we implement the following steps on each scan:

1. **Check gating** - using information for the received observations and for tracked targets (at the moment) and for each pair (track i - observation j) check inequality (14.1). As a result an array presents each observation in which track's gates is fallen.

2. **Clustering** – define clusters with tracks and observations fallen in their gates.

3. **For each cluster**:

   3.1 - Generate hypotheses following Depth First Search (DFS) procedure with certain constraints [17]. In the JPDAF approach, the two constraints which have to be satisfied for a feasible event are:

   (a) each observation can have only one origin (either a specific target or clutter), and

(b) no more than one observation originates from a target.

As a result of hypotheses generation for each hypothesis is defined a set of numbers representing the observations assigned to the corresponding tracks, where the zero represents the assignment of no observation to a given track.

3.2 - Compute hypothesis probabilities for kinematic and attribute contributions (detailed in the next paragraphs).

3.3 - Fill assignment matrix, solve assignment problem and define observation to track association.

14.4.1 Kinematics probability term for generalized data association

On the basis of defined hypotheses, the kinematic probabilities are computed as:

$$P'(H_l) = \beta^{N_M-(N_T-N_{nD})}(1-P_d)^{N_{nD}} {P_d}^{(N_T-N_{nD})} \prod_{i\neq 0, j\neq 0 | (i,j)\in H_l} g_{ij}$$

N_M being the number of observations in cluster, N_T the number of targets, N_{nD} the number of not detected targets. $(i,j) \in H_l$ involved in the product represents all the possible observation to track associations involved in hypothesis H_l.The likelihood function g_{ij}, associated with the assignment of observation j to track i is:

$$g_{ij} = \frac{e^{-d_{ij}^2/2}}{(2\pi)^{M/2}\sqrt{|\mathbf{S}_i|}}$$

P_d is the probability of detection and β is the extraneous return density, that includes probability density for new tracks and false alarms:

$$\beta = \beta_{NT} + \beta_{FA}$$

The normalized probabilities are computed as:

$$P_{\mathrm{k}}(H_l) = \frac{P'(H_l)}{\sum_{k=1}^{N_H} P'(H_k)}$$

where N_H is the number of hypotheses. To compute the probability $P_{\mathrm{k}}(i,j)$ that observation j should be assigned to track i, a sum is taken over the probabilities $P_{\mathrm{k}}(.)$ from those hypotheses H_l, in which this assignment occurs.

As an particular example for a cluster with two tracks and two new observations, see Fig. 14.3, detected during the moment of their closely spaced movement, where $P1$ and $P2$ are the tracks' predictions and $O1$, $O2$ are the received observations. The table 14.2 shows the particular hypotheses for the alternatives with respect to targets tracks and associated probabilities.

Hyp. #	Track 1	Track 2	Hyp. proba. $P'(H_l)$
H_1	0	0	$(1-P_d)^2\beta^2$
H_2	1	0	$g_{11}P_d(1-P_d)\beta$
H_3	2	0	$g_{12}P_d(1-P_d)\beta$
H_4	0	1	$g_{21}P_d(1-P_d)\beta$
H_5	0	2	$g_{22}P_d(1-P_d)\beta$
H_6	1	2	$g_{11}g_{22}P_d^2$
H_7	2	1	$g_{12}g_{21}P_d^2$

Table 14.2: Target-oriented hypothesis based on kinematics.

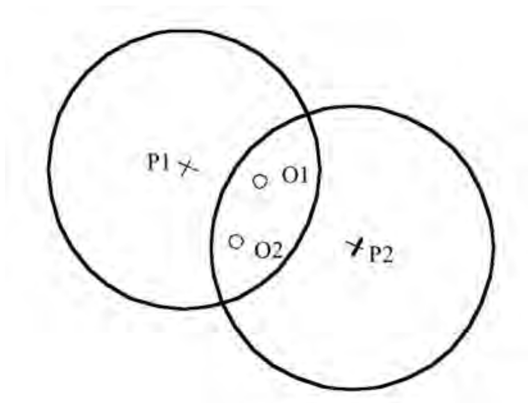


Figure 14.3: Scenario with two tracks and two observations

14.4.2 Attribute probability terms for generalized data association

The way of calculating the attribute probability term follows the joint probabilistic approach.

$$P''(H_l) = \prod_{i\neq 0, j\neq 0 | (i,j)\in H_l} d_e(i,j)$$

where

$$d_e(ij) = \sqrt{\sum_{C\in D^{\Theta_1}} [m^i_{\text{hist}}(C) - m^{i,j}_{\text{CandHist}}(C)]^2}$$

where $m^{i,j}_{\text{CandHist}}(C)$ is a candidate history of the track - result, obtained after the fusion via DSm classical rule of combination between the new received attribute observation j and predicted track's attribute state of the track i (the confirmed track history from the previous scan).

In the case of existence of two tracks and two new observations, considered in previous section and on the basis of the hypotheses matrix, one can obtain the probabilities of the hypotheses according to the following table:

Hyp. #	Track 1	Track 2	Closeness measure
H_1	0	0	$P''(H_1) = d_e(0,0) = 0$
H_2	1	0	$P''(H_2) = d_e(1,1)$
H_3	2	0	$P''(H_3) = d_e(1,2)$
H_4	0	1	$P''(H_4) = d_e(2,1)$
H_5	0	2	$P''(H_5) = d_e(2,2)$
H_6	1	2	$P''(H_6) = d_e(1,1)d_e(2,2)$
H_7	2	1	$P''(H_7) = d_e(1,2)d_e(2,1)$

Table 14.3: Target-oriented hypothesis based on attributes.

The corresponding normalized probabilities of association drawn from attribute information are obtained as:

$$P_{\mathrm{a}}(H_l) = \frac{P''(H_l)}{\sum_{k=1}^{N_H} P''(H_k)}$$

where N_H is the number of association hypotheses.

To compute the probability $P'_{\mathrm{a}}(i,j)$ that observation j should be assigned to track i, a sum is taken over the probabilities $P_{\mathrm{a}}(.)$ from those hypotheses H_l, in which this assignment occurs. Because the Euclidean distance is inversely proportional to the probability of association, the probability term $P_{\mathrm{a}}(i,j) = 1 - P'_{\mathrm{a}}(i,j)$ is used to match the corresponding kinematics probability.

14.5 Simulation scenarios

14.5.1 Simulation scenario1: Crossing targets

The simulation scenario consists of two air targets (Fighter and Cargo) and a stationary sensor at the origin with $T_{\mathrm{scan}} = 5$ sec., measurement standard deviations 0.3 deg and 60 m for azimuth and range respectively. The targets movement is from West to East with constant velocity of 250 m/sec. The headings of the fighter and cargo are 225 deg and 315 deg from North respectively. During the scan 11th-14th the targets perform maneuvers with 2.5g. Their trajectories are closely spaced in the vicinity of the two crossing points. The target detection probabilities have been set to 0.99 for both targets and the extraneous return density β to 10^{-6}. In our scenario we consider the more complicated situations, when the false alarms are available. The number of false alarms are Poisson distributed and their positions are uniformly distributed in the observation space.

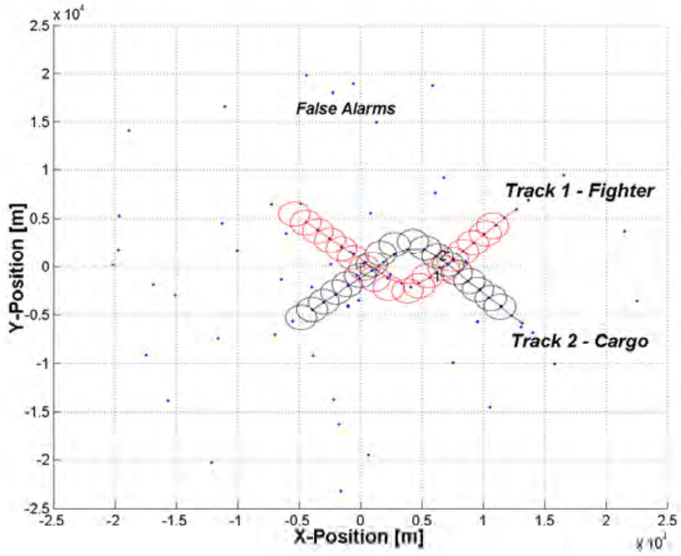


Figure 14.4: Typical simulation for scenario 1 - Two Crossing Targets' Tracks

14.5.2 Simulation scenario 2: Closely spaced targets

The second simulation scenario is influenced by the recent works of Bar-Shalom, Kirubarajan and Gokberk [6], which considers a case of closely spaced ground targets, moving in parallel. Our case consists of four air targets (alternating Fighter,Cargo, Fighter,Cargo) moving with constant velocity of 100 m/sec. The heading at the begining is 155 [deg] from North. The targets make maneuvers with 0.85g - (right, left , right turns). The sensor parameters and the false alarms are the same as in the first scenario.

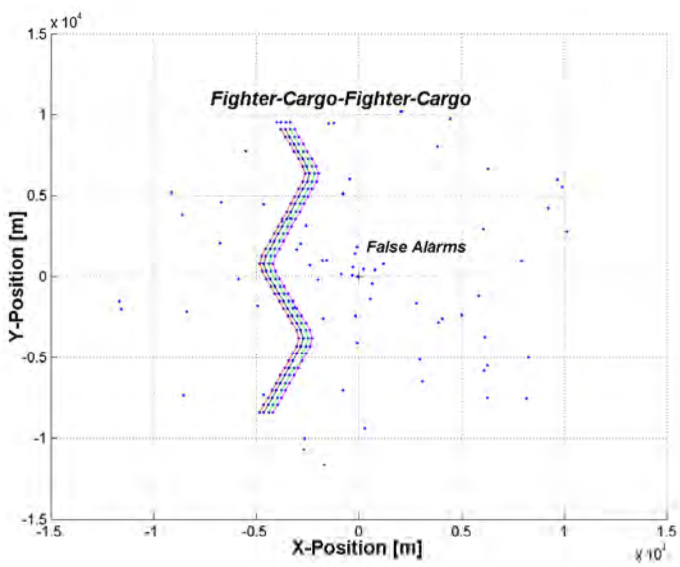


Figure 14.5: Typical simulation for scenario 2 - Four Closely Spaced Air Targets' Tracks

14.6 Simulation results

In this section the obtained simulation results, based on 100 Monte Carlo runs are presented. The goal is to demonstrate how the attribute measurements contribute for the improvement of the track performance, especially in critical cases, when the tracks are crossing and/or closely spaced.

14.6.1 Simulation results: Two crossing targets

In the case when only kinematics data are available for data association (see fig. 14.6), it is evident that after scan 15 (the second crossing moment for the targets), the tracking algorithm loses the proper targets' direction.

Here the *Tracks' Purity* performance criterion is used to examine the ratio of the right associations. Track purity is considered as a ratio of the number of correct observation-target associations (in case of detected target) over the total number of available observations during tracking scenario.

The results from table 14.4 show the proper (observation-track) associations in that case. Here "missed" is used for the case when in the track's gate there is no observation, and "FA" is used for the case, when the track is associated with the false alarm.

	Obs. 1	Obs. 2	Missed	FA
Track 1	**0.7313**	0.2270	0.0304	0.0113
Track 2	0.2409	**0.7035**	0.0426	0.0130

Table 14.4: Tracks'Purity in case of Kinematics Only Data Association (KODA).

Table 14.5 shows the result, when attribute data are utilized in the generalized data association algorithm in order to improve the tracks' maintenance performance. The hybrid DSm rule is applied to produce the attribute probability term in generalized assignment matrix. As a result it is obvious that the tracks' purity increases

	Obs. 1	Obs. 2	Missed	FA
Track 1	**0.8252**	0.1496	0.0165	0.0087
Track 2	0.1557	**0.8243**	0.0165	0.0035

Table 14.5: Tracks' Purity in case of Generalized Data Association based on DSmT.

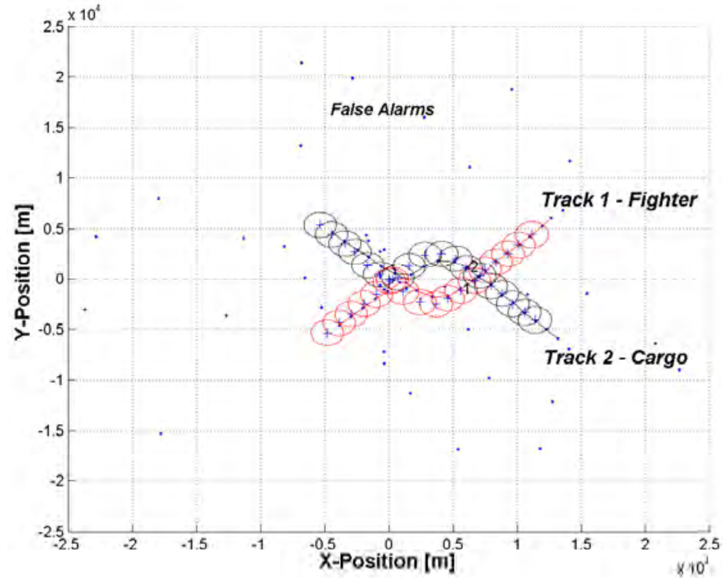


Figure 14.6: Performance of Tracking Algorithm with Kinematics Only Data Association

14.6.2 Simulation results: Four closely spaced targets

Figure 14.7 shows the performance of the implemented tracking algorithm with kinematics only data association. One can see that the four closely spaced moving in parallel targets lose the proper directions and the tracks switch.

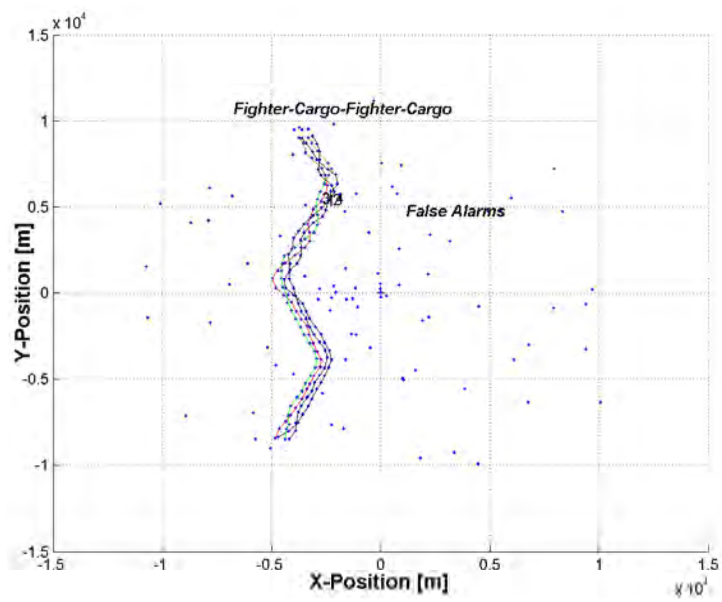


Figure 14.7: Performance of Tracking Algorithm with Kinematics Only Data Association

The results in table 14.6 show the proper (observation-track) associations in that case. The corresponding results in case of GDA based on DSmT are described in table 14.7.

	Obs. 1	Obs. 2	Obs. 3	Obs. 4	Missed	FA
Track 1	**0.5874**	0.3321	0.0558	0.0216	0.0021	0.0011
Track 2	0.2895	**0.5411**	0.1126	0.0521	0.0021	0.0026
Track 3	0.1089	0.0874	**0.5084**	0.2916	0.0021	0.0016
Track 4	0.0126	0.0332	0.3168	**0.6337**	0.0005	0.0032

Table 14.6: Tracks' Purity in case of Kinematics Only Data Association.

	Obs. 1	Obs. 2	Obs. 3	Obs. 4	Missed	FA
Track 1	**0.7026**	0.2437	0.0037	0.0216	0.0026	0.0005
Track 2	0.2253	**0.5826**	0.0584	0.0521	0.0016	0.0000
Track 3	0.0511	0.0853	**0.6047**	0.2563	0.0011	0.0016
Track 4	0.0189	0.0853	0.2121	**0.6805**	0.0016	0.0016

Table 14.7: Tracks' Purity with GDA based on DSmT.

14.6.3 Simulation results of GDA based on Dempster-Shafer theory

The results based on Dempster-Shafer theory for attribute data association are described in the tables below. For scenario 1 (**two crossing targets**), the tracks' purity is obtained in table 14.8. For scenario 2 (**four closely spaced targets**), the tracks' purity performance is obtained in table 14.9.

	Obs. 1	Obs. 2	Missed	FA
Track 1	**0.7548**	0.1609	0.0643	0.0200
Track 2	0.2209	**0.7548**	0.0174	0.0070

Table 14.8: Tracks' Purity with GDA based on Dempster-Shafer Theory (two crossing targets).

	Obs. 1	Obs. 2	Obs. 3	Obs. 4	Missed	FA
Track 1	**0.5874**	0.2679	0.1211	0.0174	0.0016	0.0047
Track 2	0.2511	**0.5489**	0.1268	0.0700	0.0011	0.0021
Track 3	0.1216	0.1058	**0.5926**	0.1742	0.0021	0.0037
Track 4	0.0374	0.0711	0.1505	**0.6563**	0.0005	0.0042

Table 14.9: Tracks' Purity with GDA based on Dempster-Shafer Theory (four closely spaced targets).

14.7 Comparative analysis of the results

It is evident from the simulation results presented in previous sections, that in general the incorporated advanced concept of generalized data association leads to improving of the tracks' maintenance performance especially in complicated situations (closely spaced and/or crossing targets in clutter). It influenced over the obtained tracks' purity results. In the same time one can see, that the tracks' purity in case of using Dezert-Smarandache theory increases in comparison with the obtained one via Dempster-Shafer theory. Analysing all the obstacles making these simulations, it can be underlined that :

- Dezert-Smarandache theory makes possible to analize, process and utilize flexibly all the paradoxical information - case, which is peculiar to the problem of multiple target tracking in clutter, when the conflicts between the bodies of evidence (tracks' attribute histories and corresponding attribute measurements) often become high and critical. That way it contributes to a better understanding of the overall tracking situation and to producing an adequate decision. Processing the paradoxes (propositions, which are more specific than the others in the hyper-power set), the estimated entropy in the confirmed (via the right track-observation association) tracks' attribute histories decreases during the consecutive scans. It can be seen on the last figure 14.8, where the Pignistic entropy (i.e the Shannon entropy based on pignistic probabilities derived from the resulting belief mass [15, 11]) is estimated in the frame of $\Theta_1 = \{\text{Very Small (VS)}, \text{Small (S)}, \text{Big (B)}\}$ and the corresponding hyper-power set D^{Θ_1} (blue color curve on the top subfigure). Simulation steps show, that source of evidence here is a hybrid one - paradoxical and uncertain. In the same time the entropy of the track's attribute history, described in the second frame $\Theta_2 = \{\text{False Alarm (FA)}, \text{Fighter (F)}, \text{Cargo (C)}\}$ (red color curve on the bottom subfigure) increases. It can be explained with the applied here hybrid DSm model $\mathcal{M}_1(\Theta_2)$, built by introducing the exclusivity constraints, imposed by the real life requirements (section 14.3.2.2). The obtained that way model corresponds actually to Shafer's model, which is a particular case of hybrid DSm model (the most constrained one). Therefore, while the corresponding sets in D^{Θ_1} are usually non empty, because of the exclusivity constraints, in the second frame Θ_2, the hyper-power set is reduced to:

$$D^{\Theta_2}_{\mathcal{M}_1} = \{\emptyset, \mathrm{FA}, \mathrm{F}, \mathrm{C}, \mathrm{FA} \cup \mathrm{F}, \mathrm{FA} \cup \mathrm{C}, \mathrm{F} \cup \mathrm{C}, \mathrm{FA} \cup \mathrm{F} \cup \mathrm{C}\}$$

So, it is obvious, in that frame, the track's attribute history represents uncertain source of information. Here the entropy increases with the uncertainty during the consequtive scans, because all the masses assigned to the empty sets in D^{Θ_2} are transferred to the non-empty sets, in our case actually to the uncertainty.

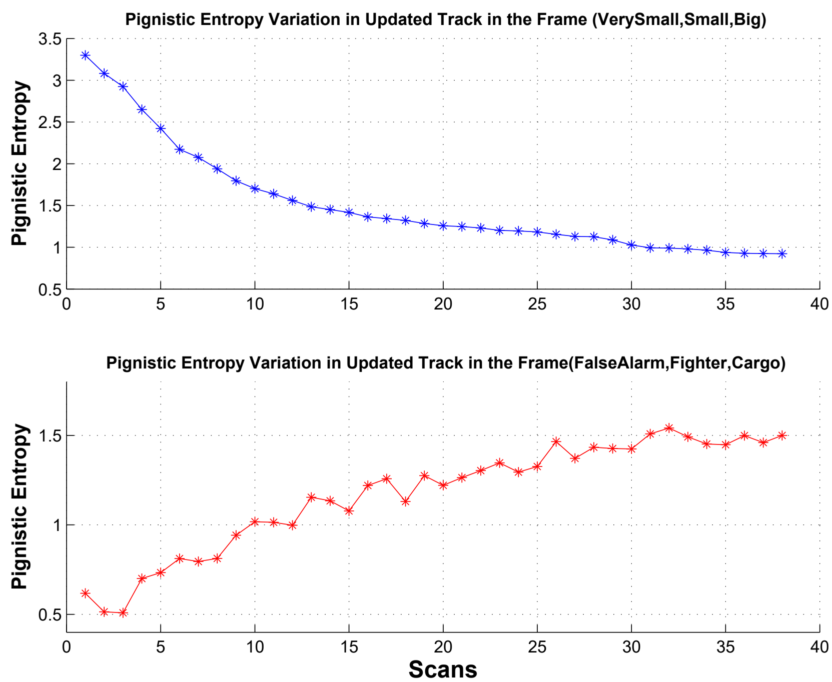


Figure 14.8: Variation of Pignistic Entropy in Track's Attribute History in the two frames Θ_1 and Θ_2

- Because of the Swerling type modelling, the observations for False Alarms, Fighter and Cargo are too much mixed. That fact causes some conflicts between general basic beliefs assignments of the described bodies of evidence. When the conflict becomes unity, it leads to indefiniteness in Dempster's rule of combination and consequently the fusion process can not be realized. From the other side, if the received modeled measurement leads to track's attribute update, in which the unity is assigned to some particular elementary hypothesis, after that point, the combinational rule of Dempster becomes indifferent to any other measurements in the next scans. It means the track's attribute history remains the same, regardless of the received observations. It naturally leads to non coherent and non adequate decisions according to the right observation-to-tracks associations.

14.8 Conclusions

In this work an approach for target tracking, which incorporates the advanced concept of generalized data (kinematics and attribute) association is presented. The realized algorithm is based on Global Nearest Neighbour-like approach and uses Munkres algorithm to resolve the generalized association matrix. The principles of Dezert-Smarandache theory of plausible and paradoxical reasoning to utilize attribute data are applied. Especially the new general hybrid DSm rule of combination is used to deal with particular integrity constraints associated with some elements of the free distributive lattice. A comparison with the corresponding results, obtained via Dempster-Shafer theory is made. It is proven, that Dempster-Shafer theory is well suited for representing uncertainty, but only in the cases of low conflicts between the bodies of evidence, while Dezert-Smarandache theory contributes to improvement of track maintenance performance in complicated situations (crossing and/or closely spaced targets), assuring a flexible and coherent decision-making, when kinematics data are insufficient to provide the proper decisions.

14.9 References

[1] Angelova D., Vassileva B., Semerdjiev Tz., *Monte Carlo-based Filter for Target Tracking with Feature Measurement*, Proceedings of the Fifth International Conference on Information Fusion, July 8-11, Annapolis, Maryland, USA, pp. 1499-1505, 2002.

[2] Bar-Shalom Y., *Multitarget-Multisensor Tracking: Advanced Applications*, Artech House,1990.

[3] Bar-Shalom Y., Fortmann T., *Tracking and Data Association*, Academic Press, 1988.

[4] Bar-Shalom Y., Li X.-R., *Estimation and Tracking : Principles, Techniques and software*, Artech House, 1993.

[5] Bar-Shalom Y., Li X.-R., *Multitarget-multisensor tracking:Principles and techniques*, YBS, 1995.

[6] Bar-Shalom Y., Kirubarajan T., Gokberk C., *Classification-Aided Sliding-Window MHT: Solution with Multidimensional Assignment*, ESP Lab, University of Connecticut (Presentation Slides HT Workshop: A Tribute to Sam Blackman, San Diego, CA, May 2003 /May 20, 2003).

[7] Blackman S., *Multitarget tracking with Radar Apllications*, Artech House, 1986.

[8] Blackman S., Popoli R., *Design and Analysis of Modern Tracking Systems*, Artech House,1999.

[9] Bojilov, L.V., *An improved version of a Multiple Targets Tracking Algorithm*, Information & Security, An International Journal, vol. 9, 2002, pp. 130-140.

[10] Bourgeois F., Lassalle J.C., *An Extension of the Munkres Algorithm for the Assignment Problem to Rectangular Matrices*, Communications of the ACM, Vol.14, Dec.1971, pp.802-806.

[11] Dezert J., Smarandache F., Daniel M., *The Generalized Pignistic Transformation*, Proceedings of Fusion 2004, Stockholm, Sweden, June 28th-July 1th, 2004.

[12] Murty K.G., *An Algorithm for Ranking All the Assignment in Order of Increasing Cost*, Operations Research 16, pp. 682-687, 1968.

[13] Popp R.L., Pattipati K.R., Bar-Shalom Y., *Dynamically Adaptable m-Best 2-D Assignment Algorithm and Multilevel Parallelization*, IEEE Trans.on AES, Vol.35, No.4, Oct.1999.

[14] Mendel J.M., *Fuzzy Logic Systems for Egineering: A Tutorial*, Proceedings of the IEEE, March 1995, pp.345-377.

[15] Tchamova A., Dezert J., Semerdjiev Tz., Konstantinova P., *Target Tracking with Generalized Data Association based on the General DSm Rule of Combination*, Proceedings of Fusion 2004, Stockholm, Sweden, June 28th-July 1th, 2004.

[16] Watson G., Blair W.D., *Benchmark Problem for Radar Resource Allocation and Tracking Maneuvering Targets in the presence of ECM*, Technical Report NSWCDD/TR-96/10, 1996.

[17] Zhou B., Bose N.K., *Multitarget Tracking in Clutter: Fast Algorithms for Data Association*, IEEE Trans. on Aerospace and Electronic Systems, vol.29, No 2, pp. 352-363, April 1993.

Chapter 15

On Blackman's Data Association Problem

Jean Dezert	Florentin Smarandache	Albena Tchamova
ONERA	Department of Mathematics	Institute for Parallel Processing
29 Av. de la Division Leclerc	University of New Mexico	Bulgarian Academy of Sciences
92320 Châtillon, France	Gallup, NM 8730, U.S.A.	Sofia, Bulgaria

Abstract: *Modern multitarget-multisensor tracking systems involve the development of reliable methods for the data association and the fusion of multiple sensor information, and more specifically the partitioning of observations into tracks. This chapter discusses and compares the application of Dempster-Shafer Theory (DST) and the Dezert-Smarandache Theory (DSmT) methods to the fusion of multiple sensor attributes for target identification purpose. We focus our attention on the paradoxical Blackman's association problem and propose several approaches to outperform Blackman's solution. We clarify some preconceived ideas about the use of degree of conflict between sources as potential criterion for partitioning evidences.*

15.1 Introduction

The association problem is of major importance in most of modern multitarget-multisensor tracking systems. This task is particularly difficult when data are uncertain and are modeled by basic belief masses and when sources are conflicting. The solution adopted is usually based on the Dempster-Shafer Theory (DST) [9] because it provides an elegant theoretical way to combine uncertain information.

This chapter is based on a paper [4] presented during the International Conference on Information Fusion, Fusion 2003, Cairns, Australia, in July 2003 and is reproduced here with permission of the International Society of Information Fusion.

However Dempster's rule of combination can give rise to some paradox/anomaly and can fail to provide the correct solution for some specific association problems. This has been already pointed out by Samuel Blackman in [2]. Therefore more study in this area is required and we propose here a new analysis of Blackman's association problem (BAP). We present in the sequel the original BAP and remind the classical attempts to solve it based on DST (including Blackman's method). In the second part of the chapter we propose and compare new approaches based on the DSmT with the free DSm model. The last part of the chapter provides a comparison of the performances of all the proposed approaches from Monte-Carlo simulation results.

15.2 Blackman's Data Association Problem

15.2.1 Association Problem no. 1

Let's recall now the original Blackman's association problem [2]. Consider only two target attribute types corresponding to the very simple *frame of discernment* $\Theta = \{\theta_1, \theta_2\}$ and the association/assignment problem for a single attribute observation Z and two tracks (T_1 and T_2). Assume now the following two predicted basic belief assignments (bba) for attributes of the two tracks:

$$m_{T_1}(\theta_1) = 0.5 \quad m_{T_1}(\theta_2) = 0.5 \quad m_{T_1}(\theta_1 \cup \theta_2) = 0$$

$$m_{T_2}(\theta_1) = 0.1 \quad m_{T_2}(\theta_2) = 0.1 \quad m_{T_2}(\theta_1 \cup \theta_2) = 0.8$$

We now assume to receive the new following bba drawn from attribute observation Z of the system

$$m_Z(\theta_1) = 0.5 \quad m_Z(\theta_2) = 0.5 \quad m_Z(\theta_1 \cup \theta_2) = 0$$

The problem is to develop a general method to find the correct assignment of the attribute measure $m_Z(.)$ with the predicted one $m_{T_i}(.)$, $i = 1, 2$. Since $m_Z(.)$ matches perfectly with $m_{T_1}(.)$ whereas $m_Z(.)$ does not match with $m_{T_2}(.)$, the optimal solution is obviously given by the assignment $(m_Z(.) \leftrightarrow m_{T_1}(.))$. The problem is to find an unique general and reliable method for solving this specific problem and for solving all the other possible association problems as well.

15.2.2 Association Problem no. 2

To compare several potential issues, we propose to modify the previous problem into a second one by keeping the same predicted bba $m_{T_1}(.)$ and $m_{T_2}(.)$ but by considering now the following bba $m_Z(.)$

$$m_Z(\theta_1) = 0.1 \quad m_Z(\theta_2) = 0.1 \quad m_Z(\theta_1 \cup \theta_2) = 0.8$$

Since $m_Z(.)$ matches perfectly with $m_{T_2}(.)$, the correct solution is now directly given by $(m_Z(.) \leftrightarrow m_{T_2}(.))$. The sequel of this chapter in devoted to the presentation of some attempts for solving the BAP, not only for these two specific problems 1 and 2, but for the more general problem where the bba $m_Z(.)$ does not match perfectly with one of the predicted bba m_{T_i}, $i = 1$ or $i = 2$ due to observation noises.

15.3 Attempts for solutions

We examine now several approaches which have already been (or could be) envisaged to solve the general association problem.

15.3.1 The simplest approach

The simplest idea for solving BAP, surprisingly not reported by Blackman in [2] is to use a classical minimum distance criterion directly between the predictions m_{T_i} and the observation m_Z. The classical L^1 (city-block) or L^2 (Euclidean) distances are typically used. Such simple criterion obviously provides the correct association in most of cases involving perfect (noise-free) observations $m_Z(.)$. But there exists numerical cases for which the optimal decision cannot be found at all, like in the following numerical example:

$$m_{T_1}(\theta_1) = 0.4 \quad m_{T_1}(\theta_2) = 0.4 \quad m_{T_1}(\theta_1 \cup \theta_2) = 0.2$$

$$m_{T_2}(\theta_1) = 0.2 \quad m_{T_2}(\theta_2) = 0.2 \quad m_{T_2}(\theta_1 \cup \theta_2) = 0.6$$

$$m_Z(\theta_1) = 0.3 \quad m_Z(\theta_2) = 0.3 \quad m_Z(\theta_1 \cup \theta_2) = 0.4$$

From these bba, one gets $d_{L^1}(T_1, Z) = d_{L^1}(T_2, Z) = 0.4$ (or $d_{L^2}(T_1, Z) = d_{L^2}(T_2, Z) \approx 0.24$) and no decision can be drawn for sure, although the minimum conflict approach (detailed in next section) will give us instead the following solution ($Z \leftrightarrow T_2$). It is not obvious in such cases to justify this method with respect to some other ones. What is more important in practice [2], is not only the association solution itself but also the attribute likelihood function $P(Z|T_i) \equiv P(Z \leftrightarrow T_i)$. As we know many likelihood functions (exponential, hyper-exponential, Chi-square, Weibull pdf, etc) could be build from $d_{L^1}(T_i, Z)$ (or $d_{L^2}(T_i, Z)$ measures but we do not know in general which one corresponds to the real attribute likelihood function.

15.3.2 The minimum conflict approach

The first idea suggested by Blackman for solving the association problem was to apply Dempster's rule of combination [9] $m_{T_iZ}(.) = [m_{T_i} \oplus m_Z](.)$ defined by $m_{T_iZ}(\emptyset) = 0$ and for any $C \neq \emptyset$ and $C \subseteq \Theta$,

$$m_{T_iZ}(C) = \frac{1}{1 - k_{T_iZ}} \sum_{A \cap B = C} m_{T_i}(A) m_Z(B)$$

and choose the solution corresponding to the minimum of conflict k_{T_iZ}. The sum in previous formula is over all $A, B \subseteq \Theta$ such that $A \cap B = C$. The degree of conflict k_{T_iZ} between m_{T_i} and m_Z is given by

$\sum_{A\cap B=\emptyset} m_{T_i}(A)m_Z(B) \neq 0$. Thus, an intuitive choice for the attribute likelihood function is $P(Z|T_i) = 1-k_{T_iZ}$. If we now apply Dempster's rule for the problem 1, we get the same result for both assignments, i.e. $m_{T_1Z}(.) = m_{T_2Z}(.)$ with $m_{T_iZ}(\theta_1) = m_{T_iZ}(\theta_2) = 0.5$ for $i = 1, 2$ and $m_{TZ}(\theta_1 \cup \theta_2) = 0$, and more surprisingly, the correct assignment ($Z \leftrightarrow T_1$) is not given by the minimum of conflict between sources since one has actually $(k_{T_1Z} = 0.5) > (k_{T_2Z} = 0.1)$. Thus, it is impossible to get the correct solution for this first BAP from the minimum conflict criterion as we firstly expected intuitively. This same criterion provides us however the correct solution for problem 2, since one has now $(k_{T_2Z} = 0.02) < (k_{T_1Z} = 0.1)$. The combined bba for problem 2 are given by $m_{T_1Z}(\theta_1) = m_{T_1Z}(\theta_2) = 0.5$ and $m_{T_2Z}(\theta_1) = m_{T_2Z}(\theta_2) = 0.17347$, $m_{T_2Z}(\theta_1 \cup \theta_2) = 0.65306$.

15.3.3 Blackman's approach

To solve this apparent anomaly, Samuel Blackman has then proposed in [2] to use a relative, rather than an absolute, attribute likelihood function as follows

$$L(Z \mid T_i) \triangleq (1 - k_{T_iZ})/(1 - k^{\min}_{T_iZ})$$

where $k^{\min}_{T_iZ}$ is the minimum conflict factor that could occur for either the observation Z or the track T_i in the case of *perfect* assignment (when $m_Z(.)$ and $m_{T_i}(.)$ coincide). By adopting this relative likelihood function, one gets now for problem 1

$$\begin{cases} L(Z \mid T_1) = \frac{1-0.5}{1-0.5} = 1 \\ L(Z \mid T_2) = \frac{1-0.1}{1-0.02} = 0.92 \end{cases}$$

Using this second Blackman's approach, there is now a larger likelihood associated with the first assignment (hence the right assignment solution for problem 1 can be obtained now based on the max likelihood criterion) but the difference between the two likelihood values is very small. As reported by S. Blackman in [2], *more study in this area is required* and we examine now some other approaches. It is also interesting to note that this same approach fails to solve the problem 2 since the corresponding likelihood functions for problem 2 become now

$$\begin{cases} L(Z \mid T_1) = \frac{1-0.1}{1-0.5} = 1.8 \\ L(Z \mid T_2) = \frac{1-0.02}{1-0.02} = 1 \end{cases}$$

which means that the maximum likelihood solution gives now the incorrect assignment $(m_Z(.) \leftrightarrow m_{T_1}(.))$ for problem 2 as well.

15.3.4 Tchamova's approach

Following the idea of section 15.3.1, Albena Tchamova has recently proposed in [3] to use rather the L^1 (city-block) distance $d_1(T_i, T_iZ)$ or L^2 (Euclidean) distance $d_2(T_i, T_iZ)$ between the predicted bba $m_{T_i}(.)$

and the updated/combined bba $m_{T_iZ}(.)$ to measure the closeness of assignments with

$$d_{L^1}(T_i, T_iZ) = \sum_{A\in 2^\Theta} \mid m_{T_i}(A) - m_{T_iZ}(A) \mid$$

$$d_{L^2}(T_i, T_iZ) = [\sum_{A\in 2^\Theta} [m_{T_i}(A) - m_{T_iZ}(A)]^2]^{1/2}$$

The decision criterion here is again to choose the solution which yields the minimum distance. This idea is justified by the analogy with the steady-state Kalman filter (KF) behavior because if $z(k+1)$ and $\hat{z}(k+1|k)$ correspond to measurement and predicted measurement for time $k+1$, then the well-known KF updating state equation [1] is given by (assuming here that dynamic matrix is identity) $\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(z(k+1) - \hat{z}(k+1|k))$. The steady-state is reached when $z(k+1)$ coincides with predicted measurement $\hat{z}(k+1|k)$ and therefore when $\hat{x}(k+1|k+1) \equiv \hat{x}(k+1|k)$. In our context, $m_{T_i(.)}$ plays the role of predicted state and $m_{T_iZ}(.)$ the role of updated state. Therefore it a priori makes sense that correct assignment should be obtained when $m_{T_iZ}(.)$ tends towards $m_{T_i}(.)$ for some closeness/distance criterion. Monte Carlo simulation results will prove however that this approach is also not as good as we can expect.

It is interesting to note that Tchamova's approach succeeds to provide the correct solution for problem 1 with both distances criterions since $(d_{L^1}(T_1, T_1Z) = 0) < (d_{L^1}(T_2, T_2Z) \sim 1.60)$ and $(d_{L^2}(T_1, T_1Z) = 0) < (d_{L^2}(T_2, T_2Z) \sim 0.98)$, but provides the wrong solution for problem 2 since we will get both $(d_{L^1}(T_2, T_2Z) \sim 0.29) > (d_{L^1}(T_1, T_1Z) = 0)$ and $(d_{L^2}(T_2, T_2Z) \sim 0.18) > d_{L^2}(T_1, T_1Z) = 0)$.

15.3.5 The entropy approaches

We examine here the results drawn from several entropy-like measures approaches. Our idea is now to use as decision criterion the minimum of the following entropy-like measures (expressed in *nats* - i.e. natural number basis with convention $0\log(0) = 0$):

- Extended entropy-like measure:

$$H_{ext}(m) \triangleq - \sum_{A\in 2^\Theta} m(A)\log(m(A))$$

- Generalized entropy-like measure [5, 8]:

$$H_{gen}(m) \triangleq - \sum_{A\in 2^\Theta} m(A)\log(m(A)/|A|)$$

- Pignistic entropy:

$$H_{betP}(m) \triangleq - \sum_{\theta_i\in \Theta} P\{\theta_i\}\log(P\{\theta_i\})$$

where the pignistic(betting) probabilities $P(\theta_i)$ are obtained by

$$\forall \theta_i \in \Theta, \quad P\{\theta_i\} = \sum_{B\subseteq\Theta|\theta_i\in B} \frac{1}{|B|} m(B)$$

It can be easily verified that the minimum entropy criterion (based on H_{ext}, H_{gen} or H_{betP}) computed from combined bba $m_{T1Z}(.)$ or $m_{T2Z}(.)$ are actually unable to provide us correct solution for problem 1 because of indiscernibility of $m_{T1Z}(.)$ with respect to $m_{T2Z}(.)$. For problem 1, we get $H_{ext}(m_{T1Z}) = H_{ext}(m_{T2Z}) = 0.69315$ and exactly same numerical results for H_{gen} and H_{betP} because no uncertainty is involved in the updated bba for this particular case. If we now examine the numerical results obtained for problem 2, we can see that minimum entropy criteria is also unable to provide the correct solution based on H_{ext}, H_{gen} or H_{betP} criterions since one has $H_{ext}(m_{T2Z}) = 0.88601 > H_{ext}(m_{T1Z}) = 0.69315$, $H_{gen}(m_{T2Z}) = 1.3387 > H_{gen}(m_{T1Z}) = 0.69315$ and $H_{betP}(m_{T1Z}) = H_{betP}(m_{T2Z}) = 0.69315$.

These first results indicate that approaches based on absolute entropy-like measures appear to be useless for solving BAP since there is actually no reason which justifies that the correct assignment corresponds to the absolute minimum entropy-like measure just because m_Z can stem from the least informational source. The association solution itself is actually independent of the informational content of each source.

An other attempt is to use rather the minimum of variation of entropy as decision criterion. Thus, the following $\min\{\Delta_1(.), \Delta_2(.)\}$ criterions are examined; where variations $\Delta_i(.)$ for $i = 1, 2$ are defined as the

- variation of extended entropy:

$$\Delta_i(H_{ext}) \triangleq H_{ext}(m_{T_iZ}) - H_{ext}(m_{T_i})$$

- variation of generalized entropy:

$$\Delta_i(H_{gen}) \triangleq H_{gen}(m_{T_iZ}) - H_{gen}(m_{T_i})$$

- variation of pignistic entropy:

$$\Delta_i(H_{betP}) \triangleq H_{betP}(m_{T_iZ}) - H_{betP}(m_{T_i})$$

Only the 2nd criterion, i.e. $\min(\Delta_i(H_{gen}))$ provides actually the correct solution for problem 1 and none of these criterions gives correct solution for problem 2.

The last idea is then to use the minimum of relative variations of pignistic probabilities of θ_1 and θ_2 given by the minimum on i of

$$\Delta_i(P) \triangleq \sum_{j=1}^{2} \frac{|P_{T_iZ}(\theta_j) - P_{T_i}(\theta_j)|}{P_{T_i}(\theta_j)}$$

where $P_{T_iZ}(.)$ and $P_{T_i}(.)$ are respectively the pignistic transformations of $m_{T_iZ}(.)$ and $m_{T_i}(.)$. Unfortunately, this criterion is unable to provide the solution for problems 1 and 2 because one has here in both problems $\Delta_1(P) = \Delta_2(P) = 0$.

15.3.6 Schubert's approach

We examine now the possibility of using a Dempster-Shafer clustering method based on metaconflict function (MC-DSC) proposed in Johan Schubert's research works [6, 8] for solving the associations problems 1 and 2. A DSC method is a method of clustering uncertain data using the conflict in Dempster's rule as a distance measure [7]. The basic idea is to separate/partition evidences by their conflict rather than by their proposition's event parts. Due to space limitation, we will just summarize here the principle of the *classical* MC- DSC method.

Assume a given set of evidences (bba) $E(k) \triangleq \{m_{T_i}(.), i = 1, \ldots, n\}$ is available at a given index (space or time or whatever) k and suppose that a given set $E(k+1) \triangleq \{m_{z_j}(.), j = 1, \ldots, m\}$ of new bba is then available for index $k+1$. The complete set of evidences representing all available information at index $k+1$ is $\chi = E(k) \cup E(k+1) \triangleq \{e_1, \ldots, e_q\} \equiv \{m_{T_i}(.), i = 1, \ldots, n, m_{z_j}(.), j = 1, \ldots, m\}$ with $q = n + m$. The problem we are faced now is to find the optimal partition/assignment of χ in disjoint subsets χ_p in order to combine informations within each χ_p in a coherent and efficient way. The idea is to combine, in a first step, the set of bba belonging to the same subsets χ_p into a new bba $m_p(.)$ having a corresponding conflict factor k_p. The conflict factors k_p are then used, in a second step, at a metalevel of evidence associated with the new frame of discernment $\Theta = \{AdP, \neg Adp\}$ where AdP is short for *adequate partition*. From each subset χ_p, $p = 1, \ldots P$ of the partition under investigation, a new bba is defined as:

$$m_{\chi_p}(\neg AdP) \triangleq k_p \quad \text{and} \quad m_{\chi_p}(\Theta) \triangleq 1 - k_p$$

The combination of all these metalevel bba $m_{\chi_p}(.)$ by Dempster's rule yields a global bba

$$m(.) = m_{\chi_1}(.) \oplus \ldots \oplus m_{\chi_P}(.)$$

with a corresponding *metaconflict factor* denoted $Mcf(\chi_1, \ldots, \chi_P) \triangleq k_{1,\ldots,P}$. It can be shown [6] that the metaconflict factor can be easily calculated directly from conflict factors k_p by the following metaconflict function (MCF)

$$Mcf(\chi_1, \ldots, \chi_P) = 1 - \prod_{p=1}^{P}(1 - k_p) \tag{15.1}$$

By minimizing the metaconflict function (i.e. by browsing all potential assignments), we intuitively expect to find the optimal/correct partition which will hopefully solve our association problem. Let's go back now to our very simple association problems 1 and 2 and examine the results obtained from the

MC-DSC method.

The information available in association problems is denoted $\chi = \{m_{T_1}(.), m_{T_2}(.), m_Z(.)\}$. We now examine all possible partitions of χ and the corresponding metaconflict factors and decision (based on minimum metaconflict function criterion) as follows:

- Analysis for problem 1:
  - the (correct) partition $\chi_1 = \{m_{T_1}(.), m_Z(.)\}$ and $\chi_2 = \{m_{T_2}(.)\}$ yields through Dempter's rule the conflict factors $k_1 \triangleq k_{T_1 Z} = 0.5$ for subset χ_1 and $k_2 = 0$ for subset χ_2 since there is no combination at all (and therefore no conflict) in χ_2. According to (15.1), the value of the metaconflict is equal to

$$\text{Mcf}_1 = 1 - (1-k_1)(1-k_2) = 0.5 \equiv k_1$$

  - the (wrong) partition $\chi_1 = \{m_{T_1}(.)\}$ and $\chi_2 = \{m_{T_2}(.), m_Z(.)\}$ yields the conflict factors $k_1 = 0$ for subset χ_1 and $k_2 = 0.1$ for subset χ_2. The value of the metaconflict is now equal to

$$\text{Mcf}_2 = 1 - (1-k_1)(1-k_2) = 0.1 \equiv k_2$$

  - since $\text{Mcf}_1 > \text{Mcf}_2$, the minimum of the metaconflict function provides the wrong assignment and the MC-DSC approach fails to generate the solution for the problem 1.

- Analysis for problem 2:
  - the (wrong) partition $\chi_1 = \{m_{T_1}(.), m_Z(.)\}$ and $\chi_2 = \{m_{T_2}(.)\}$ yields through Dempter's rule the conflict factors $k_1 \triangleq k_{T_1 Z} = 0.1$ for subset χ_1 and $k_2 = 0$ for subset χ_2 since there is no combination at all (and therefore no conflict) in χ_2. According to (15.1), the value of the metaconflict is equal to

$$\text{Mcf}_1 = 1 - (1-k_1)(1-k_2) = 0.1 \equiv k_1$$

  - the (correct) partition $\chi_1 = \{m_{T_1}(.)\}$ and $\chi_2 = \{m_{T_2}(.), m_Z(.)\}$ yields the conflict factors $k_1 = 0$ for subset χ_1 and $k_2 = 0.02$ for subset χ_2. The value of the metaconflict is now equal to

$$\text{Mcf}_2 = 1 - (1-k_1)(1-k_2) = 0.02 \equiv k_2$$

  - since $\text{Mcf}_2 < \text{Mcf}_1$, the minimum of the metaconflict function provides in this case the correct solution for the problem 2.

From these very simple examples, it is interesting to note that Schubert's approach is actually exactly equivalent (in these cases) to the min-conflict approach detailed in section 15.3.2 and thus will not provide unfortunately better results. It is also possible to show that Schubert's approach also fails if one considers jointly the two observed bba $m_{Z_1}(.)$ and $m_{Z_2}(.)$ corresponding to problems 1 and 2 with $m_{T_1}(.)$ and $m_{T_2}(.)$. If one applies the principle of minimum metaconflict function, one will take the wrong decision since the wrong partition $\{(Z_1, T_2), (Z_2, T_1)\}$ will be declared. This result is in contradiction with our intuitive expectation for the true opposite partition $\{(Z_1, T_1), (Z_2, T_2)\}$ taking into account the coincidence of the respective belief functions.

15.4 DSmT approaches for BAP

As within DST, several approaches can be attempted to try to solve Blackman's Association problems (BAP). The first attempts are based on the minimum on i of new extended entropy-like measures $H^\star_{ext}(m_{T_iZ})$ or on the minimum $H^\star_{betP}(P^\star)$. Both approaches actually fail for the same reason as for the DST-based minimum entropy criterions.

The second attempt is based on the minimum of variation of the new entropy-like measures as criterion for the choice of the decision with the new extended entropy-like measure:

$$\Delta_i(H^\star_{ext}) \triangleq H^\star_{ext}(m_{T_iZ}) - H^\star_{ext}(m_{T_i})$$

or the new generalized pignistic entropy:

$$\Delta_i(H^\star_{betP}) \triangleq H^\star_{betP}(P^\star\{.|m_{T_iZ}\}) - H^\star_{betP}(P^\star\{.|m_{T_i}\})$$

The min. of $\Delta_i(H^\star_{ext})$ gives us the wrong solution for problem 1 since $\Delta_1(H^\star_{ext}) = 0.34657$ and $\Delta_2(H^\star_{ext}) = 0.30988$ while min. of $\Delta_i(H^\star_{betP})$ give us the correct solution since $\Delta_1(H^\star_{betP}) = -0.3040$ and $\Delta_2(H^\star_{betP}) = -0.0960$. Unfortunately, both the $\Delta_i(H^\star_{ext})$ and $\Delta_i(H^\star_{betP})$ criterions fail to provide the correct solution for problem 2 since one gets $\Delta_1(H^\star_{ext}) = 0.25577 < \Delta_2(H^\star_{ext}) = 0.3273$ and $\Delta_1(H^\star_{betP}) = -0.0396 < \Delta_2(H^\star_{betP}) = -0.00823$.

The third proposed approach is to use the criterion of the minimum of relative variations of pignistic probabilities of θ_1 and θ_2 given by the minimum on i of

$$\Delta_i(P^\star) \triangleq \sum_{j=1}^{2} \frac{|P^\star_{T_iZ}(\theta_j) - P^\star_{T_i}(\theta_j)|}{P^\star_{T_i}(\theta_j)}$$

This third approach fails to find the correct solution for problem 1 (since $\Delta_1(P^\star) = 0.333 > \Delta_2(P^\star) = 0.268$) but succeeds to get the correct solution for problem 2 (since $\Delta_2(P^\star) = 0.053 < \Delta_1(P^\star) = 0.066$).

The last proposed approach is based on relative variations of pignistic probabilities conditioned by the correct assignment. The criteria is defined as the minimum of

$$\delta_i(P^\star) \triangleq \frac{|\Delta_i(P^\star|Z) - \Delta_i(P^\star|\hat{Z} = T_i)|}{\Delta_i(P^\star|\hat{Z} = T_i)}$$

where $\Delta_i(P^\star|\hat{Z} = T_i)$ is obtained as for $\Delta_i(P^\star)$ but by forcing $Z = T_i$ or equivalently $m_Z(.) = m_{T_i}(.)$ for the derivation of pignistic probabilities $P^\star_{T_i Z}(\theta_j)$. This last criterion yields the correct solution for problem 1 (since $\delta_1(P^\star) = |0.333 - 0.333|/0.333 = 0 < \delta_2(P^\star) = |0.268 - 0.053|/0.053 \approx 4$) and simultaneously for problem 2 (since $\delta_2(P^\star) = |0.053 - 0.053|/0.053 = 0 < \delta_1(P^\star) = |0.066 - 0.333|/0.333 \approx 0.8$).

15.5 Monte-Carlo simulations

As shown on the two previous BAP, it is difficult to find a general method for solving both these particular (noise-free m_Z) BAP and all general problems involving noisy attribute bba $m_Z(.)$. The proposed methods have been examined only for the original BAP and no *general* conclusion can be drawn from our previous analysis about the most efficient approach. The evaluation of the global performances/efficiency of previous approaches can however be estimated quite easily through Monte-Carlo simulations. Our Monte-carlo simulations are based on 50.000 independent runs and have been done both for the noise-free case (where $m_Z(.)$ matches perfectly with either $m_{T_1}(.)$ or $m_{T_2}(.)$) and for two noisy cases (where $m_Z(.)$ doesn't match perfectly one of the predicted bba). Two noise levels (low and medium) have been tested for the noisy cases. A basic run consists in generating randomly the two predicted bba $m_{T_1}(.)$ and $m_{T_2}(.)$ and an observed bba $m_Z(.)$ according to a random assignment $m_Z(.) \leftrightarrow m_{T_1}(.)$ or $m_Z(.) \leftrightarrow m_{T_2}(.)$. Then we evaluate the percentage of right assignments for all chosen association criterions described in this chapter. The introduction of noise on perfect (noise-free) observation $m_Z(.)$ has been obtained by the following procedure (with notation $A_1 \triangleq \theta_1$, $A_2 \triangleq \theta_2$ and $A_2 \triangleq \theta_1 \cup \theta_2$): $m_Z^{\text{noisy}}(A_i) = \alpha_i m_Z(A_i)/K$ where K is a normalization constant such as $\sum_{i=1}^3 m_Z^{\text{noisy}}(A_i) = 1$ and weighting coefficients $\alpha_i \in [0; 1]$ are given by $\alpha_i = 1/3 \pm \epsilon_i$ such that $\sum_{i=1}^3 \alpha_i = 1$.

The table 1 shows the Monte-Carlo results obtained with all investigated criterions for the following 3 cases: noise-free (NF), low noise (LN) and medium noise (MN) related to the observed bba $m_Z(.)$. The two first rows of the table correspond to simplest approach. The next twelve rows correspond to DST-based approaches.

Assoc. Criterion	NF	LN	MN
Min $d_{L^1}(T_i, Z)$	100	97.98	92.14
Min $d_{L^2}(T_i, Z)$	100	97.90	92.03
Min $k_{T_i Z}$	70.01	69.43	68.77
Min $L(Z\|T_i)$	70.09	69.87	67.86
Min $d_{L^1}(T_i, T_i Z)$	57.10	57.41	56.30
Min $d_{L^2}(T_i, T_i Z)$	56.40	56.80	55.75
Min $H_{ext}(m_{T_i Z})$	61.39	61.68	60.85
Min $H_{gen}(m_{T_i Z})$	58.37	58.79	57.95
Min $H_{betP}(m_{T_i Z})$	61.35	61.32	60.34
Min $\Delta_i(H_{ext})$	57.66	56.97	55.90
Min $\Delta_i(H_{gen})$	57.40	56.80	55.72
Min $\Delta_i(H_{betP})$	71.04	69.15	66.48
Min $\Delta_i(P)$	69.25	68.99	67.35
Min Mcf_i	70.1	69.43	68.77

Table 1 : % of success of association methods

The table 2 shows the Monte-Carlo results obtained for the 3 cases: noise-free (NF), low noise (LN) and medium noise (MN) related to the observed bba $m_Z(.)$ with the DSmT-based approaches.

Assoc. Criterion	NF	LN	MN
Min $H^{\star}_{ext}(m_{T_i Z})$	61.91	61.92	60.79
Min $H^{\star}_{betP}(P^{\star})$	42.31	42.37	42.96
Min $\Delta_i(H^{\star}_{ext})$	67.99	67.09	65.72
Min $\Delta_i(H^{\star}_{betP})$	42.08	42.11	42.21
Min $\Delta_i(P^{\star})$	76.13	75.3	72.80
Min $\delta_i(P^{\star})$	100	90.02	81.31

Table 2 : % of success of DSmT-based methods

15.6 Conclusion

A new examination of Blackman's association problem has been presented in this chapter. Several methods have been proposed and compared through Monte Carlo simulations. Our results indicate that the commonly used min-conflict method doesn't provide the best performance in general (specially w.r.t. the simplest distance approach). Thus the metaconflict approach, equivalent here to min-conflict, does not allow to get the optimal efficiency. Blackman's approach and min-conflict give same performances.

All entropy-based methods are less efficient than the min-conflict approach. More interesting, from the results based on the generalized pignistic entropy approach, the entropy-based methods seem actually not appropriate for solving BAP since there is no fundamental reason to justify them. The min-distance approach of Tchamova is the least efficient method among all methods when abandoning entropy-based methods. Monte Carlo simulations have shown that only methods based on the relative variations of generalized pignistic probabilities build from the DSmT (and the free DSm model) outperform all methods examined in this work but the simplest one. Analysis based on the DSmT and hybrid DSm rule of combination are under investigation.

15.7 References

[1] Bar-Shalom Y., Li X.R., *Estimation and Tracking: Principles, Techniques, and Software*, Artech House, 1993.

[2] Blackman S., *Association and Fusion of Multiple Sensor Data*, Chap. 7 of Multitarget-Multisensor Tracking: Advanced Applications, (Y. Bar-Shalom Editor), Artech House, 1990.

[3] Dezert J., *Fondations pour une nouvelle théorie du raisonnement plausible et paradoxal*, ONERA Tech. Rep. RT 1/06769/DTIM, Jan. 2003.

[4] Dezert J., Smarandache F., Tchamova A., *On the Blackman's association problem*, Proc. of Fusion 2003 Conf., Cairns, Australia, July 8-11, 2003.

[5] Pal N.R., Bezdek J.C., Henasinha R., *Uncertainty measure for evidential reasoning II: A new measure of total uncertainty*, Int. J. of Intell. Syst., 8(1), pp. 1-16, 1993.

[6] Schubert J., *On Nonspecific Evidence*, Int. J. of Intell. Syst. Vol. 8, pp. 711-725, 1993.

[7] Schubert J., *Robust Report Level Cluster-to-Track Fusion*, Proc. of 5th Int. Conf. on Inf. Fusion, pp. 913-918, Annapolis, July 2002.

[8] Schubert J., *Clustering belief functions based on attracting and conflicting metalevel evidence*, Proc. of IPMU conf., Annecy, France, July 1-5, 2002.

[9] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

Chapter 16

Neutrosophic Frameworks for Situation Analysis

Anne-Laure Jousselme and **Patrick Maupin**

Defence R & D Canada - Valcartier

Decision Support Systems

2459, Pie XI Blvd North

Val-Bélair, QC, G3J 1X5, CANADA

Abstract: *In situation analysis, an agent observing a scene receives information from heterogeneous sources of information including for example remote sensing devices, human reports and databases. The aim of this agent is to reach a certain level of awareness of the situation in order to make decisions. For the purpose of applications, this state of awareness can be conceived as a state of knowledge in the classical epistemic logic sense. Considering the logical connection between belief and knowledge, the challenge for the designer is to transform the raw, imprecise, conflictual and often paradoxical information received from the different sources into statements understandable by both man and machines. Situation analysis applications need frameworks general enough to take into account the different types of uncertainty and information present in the situation analysis context, doubled with a semantics allowing meaningful reasoning on situations. The aim of this chapter is to evaluate the capacity of neutrosophic logic and Dezert-Smarandache theory (DSmT) to cope with the ontological and epistemic problems of situation analysis.*

16.1 Introduction

The aim of Situation Analysis (SA) in a decision-making process is to provide and maintain a state of situation awareness for an agent observing a scene [1, 2]. For the purpose of applications, this state of awareness can be conceived as a state of knowledge in the classical epistemic logic sense. Considering the logical connection between belief and knowledge, the challenge for the designer is to transform the raw, imprecise, conflictual and often paradoxical information received from the different sources into statements understandable by both man and machines. Because the agent receives information from heterogeneous sources of information including for example remote sensing devices, human reports and databases, two simultaneous tasks need to be achieved: *measuring the world* and *reasoning about the structure of the world*. A great challenge in SA is the conciliation of both quantitative and qualitative information processing in mathematical and logical frameworks. As a consequence, SA applications need frameworks general enough to take into account the different types of uncertainty and information present in the SA context, doubled with a semantics allowing meaningful reasoning on belief, knowledge and situations. The formalism should also allow the possibility to encompass the case of multiagent systems in which the state of awareness can be distributed over several agents rather than localized.

A logical approach based on a possible worlds semantics for reasoning on belief and knowledge in multiagent context is proposed in [3]. This work by Halpern and Moses can be used as a blueprint considering that it allows to handle numerical evaluations of probabilities, thus treating separately but nevertheless linking belief, knowledge and uncertainty. Related works are those of Fagin and Halpern [4] but also Bundy [5] which extend the probability structure of Nilsson [6] based on possible worlds semantics to a more general one close to the evidence theory developed by Dempster [7] and Shafer [8]. The result is the conciliation of both measures and reasoning in a single framework.

Independently of these works has been introduced *Neutrosophy*, a branch of philosophy which studies neutralities and paradoxes, and relations between a concept and its opposite [9]. Two main formal approaches have emerged from Neutrosophy: *neutrosophic logic*, presented as a unified logic, of which fuzzy logic, classical logic and others are special cases [10, 11]; and *Dezert-Smarandache theory* (DSmT) that can be interpreted as a generalization of Dempster-Shafer theory. On one hand, neutrosophic logic appears as an interesting avenue for SA because (1) indeterminacy is explicitly represented by the means of an indeterminacy assignment, (2) falsity, truth and indeterminacy are represented independently (three distinct assignments), (3) it is a quantified logic, meaning that numerical evaluations of truth, falsity and indeterminacy values are allowed, (4) this quantification is allowed on hyperreals intervals, a generalization of intervals of real numbers given a broader frame for interpretations, (5) many novel connectives are defined (Neut-A, Anti-A, ...). On the other hand, being built on the hyper-power set of the universe of discourse, the DSmT allows to take into account the indeterminacy linked to the very definition of the individual elements of the universe of discourse, relaxing the mutual exclusivity hypothesis imposed by the Dempster-Shafer theory (DST). This framework extends thus the DST by allowing a wider variety of events to be considered when measures become available. Indeed, a particularity of SA is that most of the time it is impossible beforehand to list every possible situation that can occur. The elements of the corresponding universe of discourse cannot, thus, be considered as an exhaustive list of situations. Furthermore, in SA situations are not clearcut elements of the universe of discourse.

The aim of this chapter is to evaluate the potential of neutrosophic logic and Dezert-Smarandache theory (DSmT) to cope with the ontological and epistemic obstacles in SA (section 16.3), *i.e.* problems due to the nature of things and to cognitive limitations of the agents, human or artificial. Section 16.4 exposes four basic principles guiding SA systems design in practice, and highlight the capacity of both neutrosophic logic and DSmT to cope with these principles. After brief formal descriptions of neutrosophic logic and DSmT (section 16.5) we propose in section 16.6 different extensions based on Kripke structures and Demspter-Shafer structures. In particular,a Kripke structure for neutrosophic propositions is presented in section 16.6.2. In the latter section, we assess the ability of neutrosophic logic to process symbolic and numerical statements on belief and knowledge using the possible worlds semantics. Moreover, we investigate the representation of neutrosophic concepts of neutrality and opposite in the possible worlds semantics for situation modelization. In section 16.6.3, after introducing Nilsson and Dempster-Shafer structures, we present a possible extension to DSmT. We also propose an example to illustrate the benefit of using a richer universe of discourse, and thus how DSmT appears as an appropriate modelling tool for uncertainty in SA. We then propose a possible connection between DSmT and neutrosophic logic in the Kripke structures setting (section 16.6.4). Finally, in section 16.7 we conclude on possible research avenues for using DSmT and neutrosophic logic in SA.

16.2 Situation analysis

The term *situation* appears in the mid-fourteenth century derived from medieval Latin *situatio* meaning *being placed into a certain location*. By the middle of the seventeenth century situation is used to discuss the moral dispositions of a person, more specifically the set of circumstances a person lies in, the relations linking this person to its *milieu* or surrounding *environment*. As will be shown below, the latter definition is close to what is meant today in the field of High-Level Data Fusion, where the mental state of *situation awareness* is studied in interaction with the surrounding environment. Common synonyms of situation with a corresponding meaning are *setting*, *case*, *circumstances*, *condition*, *plight*, *scenario*, *state*, *picture*, *state of affairs*.

Although the notion of situation is used informally in everyday language to designate a given state of affairs, a simplified view of the world, and even the position of certain objects, situation is nowadays a central concept in High-Level Data Fusion where it has been given more or less formal definitions. For

Pew [12], a situation is "*a set of environmental conditions and system states with which the participant is interacting that can be characterized uniquely by a set of information, knowledge, and response options*".

16.2.1 Situation awareness as a mental state

For Endsley and Garland [1] *Situation awareness* (SAW) is "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future*". SAW is also defined in [13] as "the active mental *representation* of the status of current cognitive functions activated in the cognitive system in the context of achieving the goals of a specific task". In particular, SAW involves three key tasks: (1) Perception, (2) Comprehension and (3) Projection, in a general multiagent context (Fig. 16.1).

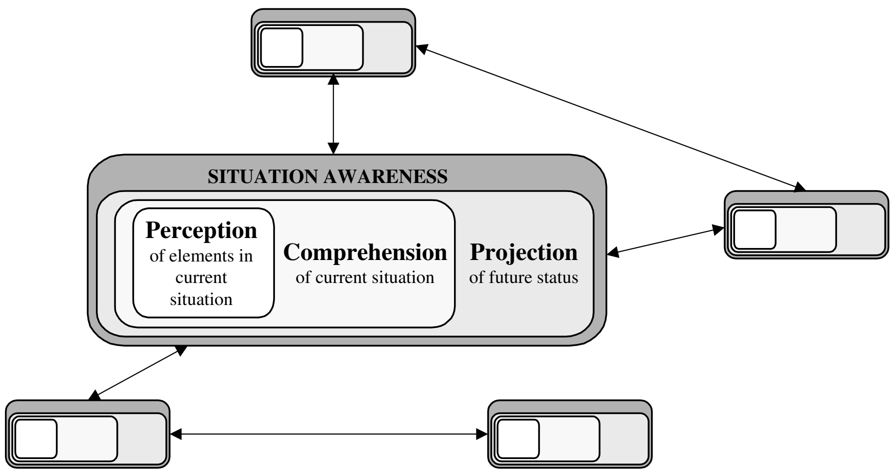


Figure 16.1: The three basic processes of situation awareness according to Endlsey and Garland (modified from [1]), in a multiagent context.

In contemporary cognitive science the concept of *mental representation* is used to study the interface between the external world and mind. Mental states are seen as relations between agents and mental representations. Formally, and following Pitt's formulation [14], for an agent to be in a psychological state Ψ with semantic property Γ is for that agent to be in a Ψ-appropriate relation to a mental representation of an appropriate kind with semantic property Γ. As far as mental states are concerned, purely syntactic approaches are not adequate for representation since semantic concepts need to be modeled.

Explicit reasoning on knowledge and the problems linked to its representation are distinctive features of situation analysis. Our position is to refer to the sources of knowledge usually considered in epistemology, namely, *Perception*, *Memory*, *Reasoning*, *Testimony* and *Consciousness* [15], and extend Endsley's model of situation awareness [1] where perception appears as the only source of knowledge.

16.2.2 Situation Analysis as a process

For Roy [2] "*Situation Analysis is a process, the examination of a situation, its elements, and their relations, to provide and maintain a product, i.e. a state of* Situation Awareness *(SAW) for the decision maker*". For a given situation the SA process creates and maintains a mental representation of the situation. Situation analysis corresponds to the levels 2, 3 and 4 of the JDL data fusion model [16, 17], hence to higher-levels of data fusion. A revisited version of the well-known model is presented on figure 16.2, with classical applications associated to the different levels. A complete situation model must take into account the following tasks of: A. Situation *perception* composed of Situation Element Acquisition, Common Referencing, Perception Origin Uncertainty Management, and Situation Element Perception Refinement as subtasks. B. Situation *comprehension* composed of Situation Element Contextual Analysis, Situation Element Interpretation, Situation Classification, Situation Recognition, and Situation Assessment as subtasks. C. Situation *projection* composed of Situation Element Projection, Impact Assessment, Situation Monitoring. Situation Watch, and Process Refinement [2].

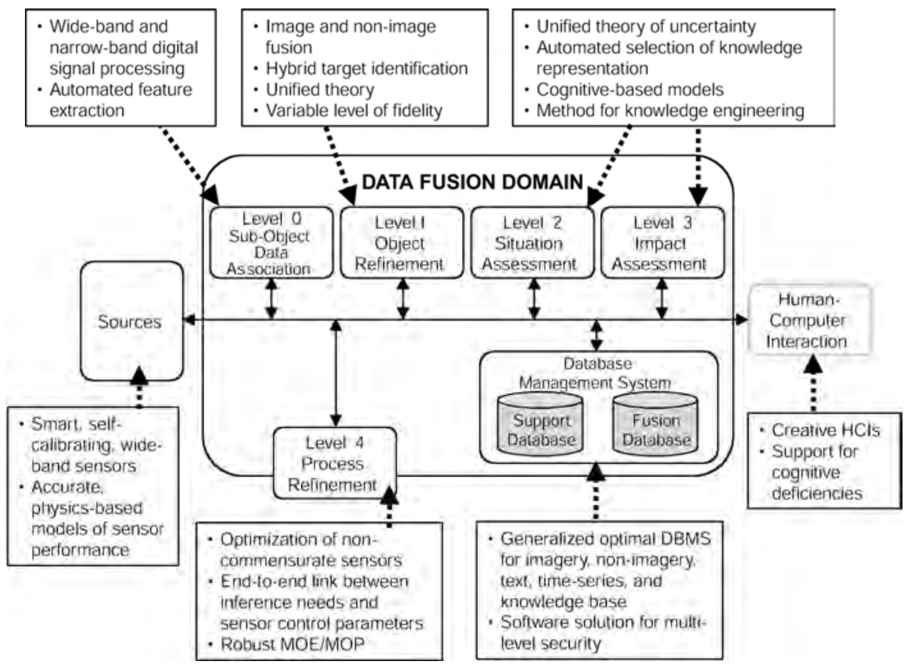


Figure 16.2: Revisited JDL data fusion model and applications [18].

The conception of a system for SA must rely on a mathematical and/or logical formalism capable of translating the mechanisms of the SAW process at the human level. The formalism should also allow the possibility to encompass the case of multiagent systems in which the state of awareness can be distributed over several agents rather than localized. A logical approach based on a possible worlds semantics for reasoning on belief and knowledge is proposed in [3]. This work by Halpern and Moses can be used as a blueprint considering that it allows to handle numerical evaluations of probabilities, thus treating separately but nevertheless linking belief, knowledge and uncertainty.

Furthermore, mathematical and logical frameworks used to model mental states should be able to represent and process autoreference such as beliefs about one's own beliefs, beliefs about beliefs about . . . and so on.

16.2.3 A general model of a distributed system

In 1990, Halpern and Moses proposed a model of distributed knowledge processing [3] that can be used for the purpose of situation analysis, as stated above. Short definitions are given below for the different components of the model:

- A **distributed system** is a finite collection of two or more interacting agents $\mathcal{A}_1, \ldots, \mathcal{A}_n$ (connected by a communication network);
- The **local state** of an agent is the determined by the encapsulation of all the information an agent has access to at a given instant;
- The **state of the environment** is defined as the information relevant to the system but not contained in the state of the agents;
- The **global state** of a system is given by the sum of the agents' local states together with the state of the environment;
- A **run** is a function from time to global states;
- A **point** is a pair (r, m) consisting of a run r and a time m;
- A **system** is defined as a set of runs. A system can also be viewed as a Kripke structure supplemented with a way to assign truth values.

This model is illustrated on figure 16.3 and appears as a sufficient basis for defining the basic concepts of situation analysis. Indeed, the local state of an agent $\mathcal{A}_i$ can also be called its Knowledge-Base (denoted by KB_i) upon which an awareness function delimits these subsets, the latter being particular views of a given situation (see section 16.4.2 on contextualization). From an algebraic point of view, a same agent can generate different views of the same situation, either disjoint or overlapping or nested.

16.3 Sources of uncertainty in Situation Analysis

Situation analysis is experimental by nature. A major obstacle encountered in the process lies in the ubiquity of uncertainty. While in a previous paper [19], we highlighted four main facets of uncertainty:

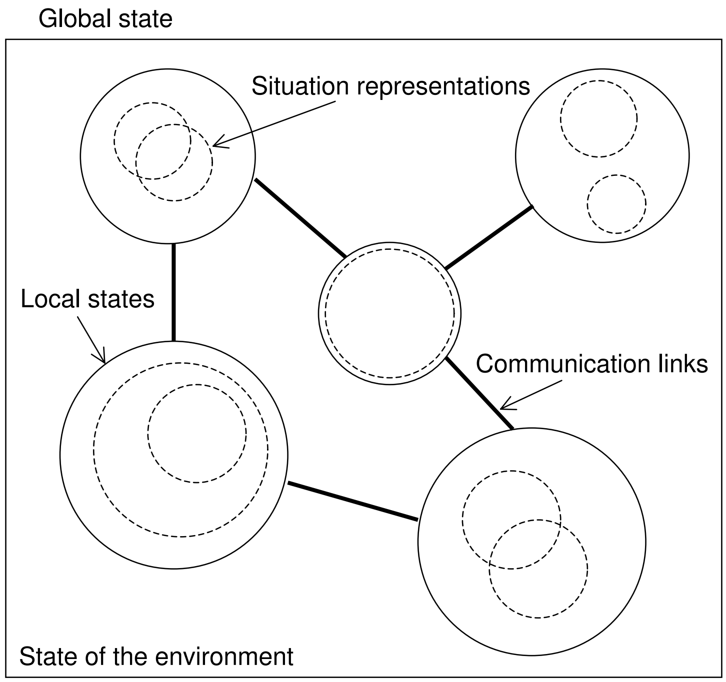


Figure 16.3: The general model of a distributed system proposed by Halpern and Moses in [3] adapted for situation representation.

(1) Meaning (mental sate or property of the information), (2) Interpretation (objective or subjective), (3) Types (fuzziness, non-specificity and discord) and (4) Mathematical representations (quantitative vs. qualitative approaches), in this section, we rather review the potential sources of uncertainty and obstacles arising in a situation analysis context.

Uncertainty has two main meanings in most of the classical dictionaries [19]: Uncertainty as a *state of mind* and uncertainty as a *physical property of information*. The first meaning refers to the state of mind of an agent, which does not possess the needed information or knowledge to make a decision; the agent is in a state of uncertainty: "I'm not sure that this object is a table". The second meaning refers to a physical property, representing the limitation of perception systems: "The length of this table is uncertain" (given the measurement device used).

Sociologists like Gérald Bronner [20] consider uncertainty as a state of mind, this state depending on our power on the uncertainty, and our capacity to avoid it. He distinguishes two types of uncertainty: *uncertainty in finality* (or material uncertainty) and *uncertainty of sense*. Uncertainty in finality is "the state of an individual that, wanting to fulfill a desire, is confronted with the open field of the possibles" ("Will my car start?"). Whereas uncertainty of sense is "the state of an individual when a part, or the whole of its systems of representation is deteriorated or can be". Uncertainty in finality corresponds to the uncertainty in which lies our understanding of the world, while uncertainty of sense bears on the representation of the world. Bronner identifies three types of uncertainty in finality, according to one's power on uncertainty, and the capacity to avoid it:

- Situation of type I: Uncertainty does not depend on the agent and can not be avoided;
- Situation of type II: Uncertainty does not depend on the agent but can be avoided;
- Situation of type III: Uncertainty is generated by the agent and can be avoided.

In situation analysis, agents are confronted to uncertainty of sense (data driven) from the bottom-up perspective and to uncertainty in finality (goal driven) from the top-down perspective. It follows that there are two kinds of limits to state estimation and prediction in Situation Analysis:

1. *Ontological limits* due to the nature of things and
2. *Epistemic limits* due to cognitive limitations of the agents, human or artificial.

Typical obstacles [21] are *anarchy* and *instability* when the situation is not governed by an identifiable law or in the absence of nomic stability. *Chance* and *chaos*, are serious obstacles to state evaluation and prediction as far as an exact estimation is sought for although regularities and determinism are observed. Another typical obstacle is the *vagueness* of concepts. Natural language concepts are inherently vague, meaning that their definition is approximate and borderline cases arise. This is true as well for properties but also for concepts.

Indeterminacy is another unavoidable obstacle. It may arise from paradoxical conclusions to a given inference (*i.e.* Russell's paradox, or sorites paradox), from impossible physical measurements (*i.e.* position *and* speed of an atomic particle) or for practical reasons (*i.e.* NP-complete problems). From a given theoretical stand point (classical vs. quantum mechanics), indeterminacy may nevertheless be proposed as a conclusion to specific unanswerable questions in order to nevertheless allow reasoning using the remaining information.

Ignorance of the underlying laws governing the situation is a major cause of uncertainty. For example not knowing that a given tactical maneuver is possible precludes the possibility to predict its occurrence. Especially present in human affairs *innovation* can be a major obstacle in SA. New kinds of objects (weapons), processes (courses of action) or ideas (doctrines) arise and one has no choice but to deal with it and adapt.

Myopia or data ignorance, is also a typical problem in SA. Data must be available on time in order to assess a situation, meaning that even if the information sources exist circumstances can prevent their delivery. Another case of myopia occurs when data is not available in sufficient detail, as in pattern recognition when classes are only coarsely defined or when sensors have limited spatial resolution. Data is thus accessible through estimations obtained by sampling as in surveys, by the computation of aggregates as in Data Fusion or by the modelization of rough estimates. As a consequence the available data is only imprecise and incomplete and leads most of the time to conflicting choices of decision. A major task of SA is change detection, failure prediction.

Any attempt in the conception of a system is be bounded by *inferential incapacity* of human or artificial agents. Limitations in agents can arise because of a lack of awareness. As far as knowledge is concerned, an agent cannot always give a value to a proposition, for example if it is not even aware of the existence of the concept denoted by the proposition at hand. Agents are *resource bounded* meaning that agents have only limited memorization capabilities, in some cases they have power supply limitations, etc. or have only limited cognitive and computational capabilities. Agents may also have limited visual or auditory acuity. Sometimes, these limitations come from the outside and are situation driven: electronic countermeasures, only a limited amount of time or money is available to do the job, etc. Furthermore agents cannot focus on all issues simultaneously. As Fagin and Halpern puts it in [22] "[...] Even if $\mathcal{A}$ does perfect reasoning with respect to the limited number of issues on which he is focusing in any given frame of mind, he may not put his conclusions together. Indeed, although in each frame of mind agent $\mathcal{A}$ may be consistent, the conclusions $\mathcal{A}$ draws in different frames of mind may be inconsistent." Finally, agents must work with an inconsistent set of beliefs. For example, we know that lying is amoral, but in some case we admit it could be a good alternative to a crisis.

16.4 Ontological principles in Situation Analysis

Given the limitations and the sources of uncertainty involved in Situation Analysis (section 16.3), we state in this section four main ontological principles that should guide SA systems design in practice: (1) allowing statements and reasoning about uncertainty to be made, (3) contextualization, (2) enrichment of the universe of discourse, and (4) allowing autoreference.

16.4.1 Allowing statements and reasoning about uncertainty

We begin with two observations that will guide the discussion of this section:

1. Many concepts are linked to uncertainty: Vagueness, indeterminacy, truth, belief, indiscernibility, ambiguity, non-specificity, incompleteness, imprecision to name a few. Although these concepts are *a priori* distinct, it is common to confuse them and to be unable to talk about one without any reference to the other. The recent development of new theories of uncertainty aims at separating these aspects, and bring clarifications in this direction as it is the case for probability theory and fuzzy logic. Another contribution in this direction is the axiomatization proposed by Fagin and Halpern in [4] which provides a semantical structure to reasoning about both belief and probability, and thus distinguishing these two often confused concepts.

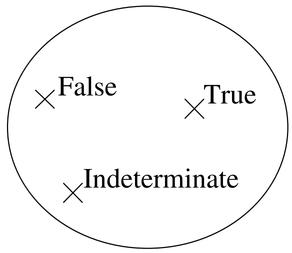
2. Although it is possible to deal with uncertainty in general using purely qualitative notions, the mixture of discrete and continuous objects composing the world has led to introduce degrees.

In a very general sense as written in the previous section (section 16.3), uncertainty is often seen as the result of indeterminacy. As far as formalization is concerned the classical means of reasoning soon exposed their limitations. Propositional Calculus (PC) relies on the principle of bivalence expressing the fact that a proposition is either True or False. Hence, only two truth values are allowed leaving no way to express indeterminacy. The most common way go beyond bivalence is to introduce supplementary truth values in the PC framework. The signification of the supplementary truth value differs from one author to another, from one logic to another. However, it is common to denote truth, falsity and indeterminacy by 1, 0 and $\frac{1}{2}$ respectively.

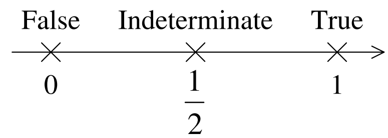
Here the problem of the meaning of the uncertainty arises. For a given type of uncertainty (contingent future events, indetermination, etc.) corresponds a particular interpretation of the set of connectives. If Łukasiewicz was primarily interested with the problem of contingent future event or possibility, Kleene in 1938 [23] proposed three value logics used in recursion theory in order to design stopping criteria and allow for indeterminacy of some propositions. Bochvar (1938) [24] proposed a logic quantifying propositions as sensible and senseless. For him true and false propositions are meaningful, the third truth-value designates meaningless or paradoxical propositions. Bochvar's system of logic, was later rediscovered by HalldÈn in 1949 [25] and used to process vague and nonsensical propositions. In fact, the different meanings of uncertainty are translated in the particular definitions given to logical connectors with respect to common intuition of the terms at hand.

It is important to note that in general the truth values are not ordered and just like in PC the truth values are purely conventional. In this sense, the so-called values of the truth tables can be considered qualitative (see Fig. 16.4-(a)). However, these three truth values can also be ordered, representing then a rough quantitative description of the world (see Fig. 16.4-(b)). But intuition also tells us that things are not always clear cut in the real world and rather appear in tones of gray. A three-valued logic can be generalized to a n-valued logic and by extension to fuzzy logic with an infinite number of truth-values ranging on the real set interval $[0; 1]$. Such an extension introduces thus an order between truth statements (see Fig. 16.4-(c)). Another consequence of this extension is that the notion of uncertainty is now expressed explicitly in terms of truth or falsity. While in a three-valued logic, indeterminacy, possibility or vagueness are expressed as neither True nor False, in Łukasiewicz's or fuzzy logic, to take a more recent example, the uncertainty is expressed by an explicit reference to truth or falsity.

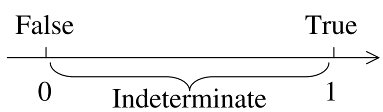
The introduction of degrees imposes then an order between values. The truth becomes then a kind of false and vice-versa, and the qualitative aspect of the three initial truth values is lost, with their independence. Yet another extension which conciliates both qualitative and quantitative aspects of indeterminacy is to consider different independent aspects of uncertainty and represent them on independent axes. This



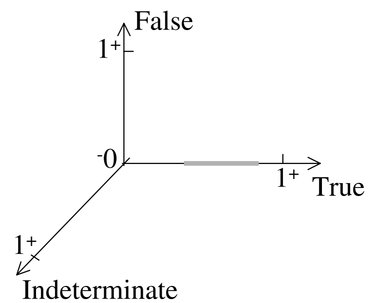
(a) Three non-ordered truth values corresponding to purely qualitative description.



(b) Three ordered truth values corresponding to a rough qualitative description.



(c) Infinite number of ordered truth values ranging from False to True.



(d) Three independent axes for three ordered features of the uncertainty (truth, falsity and indeterminacy).

Figure 16.4: Supplementary truth values for representing indeterminacy.

is the principle developed by Smarandache in the neutrosophic logic [10, 11], where the considered aspects of uncertainty are truth, falsity and indeterminacy (see Fig. 16.4-(d)). Hence, in neutrosophic logic both the qualitative aspect of non-ordered three-valued logics and the quantitative aspect of fuzzy logic are combined. One main benefit of neutrosophic logic is that indeterminacy can be addressed by two different manners: (1) Using the indeterminacy function independently[1] of the truth and falsity functions or (2) using the three previous functions as it is commonly done in fuzzy logic. Moreover, because of the assumed independence of the three concepts of truth, falsity and indeterminacy, NL is able to represent paradoxes, for example something that is completely true, completely false and completely indeterminate. Neutrosophy and neutrosophic logics are introduced respectively in sections 16.5.1 and 16.5.2.

[1]Note however that although truth, falsity and indeterminacy are considered independently in NL, the use of the hyperreals is a means to make them dependent. Indeed, an absolutely True proposition ($T(\phi) = 1^+$) is also absolutely False ($F(\phi) = ^- 0$). This condition is not required for relatively True propositions ($T(\phi) = 1$) [10].

Finally, we remind that although indeterminacy has been discussed from a logical point of view, indeterminacy is also represented in more quantitative approaches. Indeed, in probability theory, assigning a probability value in $]0;1[$ to an event translates the indeterminate state of this event. It has nothing to do with the truth of the event, but rather with its potential occurrence. By extension, Dempster-Shafer theory, possibility theory or Dezert-Smarandache theory are other numerical approaches to deal with indeterminacy. Some of these approaches are briefly discussed in section 16.5.3.

16.4.2 Contextualization

In SA, the operation of contextualization serves many purposes and is at the basis of the abstract notion of *situation* itself as it is understood by defence scientists, software engineers and commanding officers as well. According to Theodorakis [26], in the context of information modelling, "*a context is viewed as a reference environment relatively to which descriptions of real world objects are given. The notion of context may be used to represent real world partitions, divisions, or in general, groups of information, such as situations, viewpoints, workspaces, or versions*". In this sense a context is a mental, thus partial, representation of a real situation. For Theodorakis [26] "*A situation records the state of the world as it is, independently of how it is represented in the mind of an agent. A situation is complete as it records all the state of the world. Whereas, contexts are partial as they represent situations and hence capture different perspectives or record different levels of detail of a particular situation*".

For Brézillon [27] a context can be "*a set of preferences and/or beliefs, a window on a screen, an infinite and only partially known collection of assumptions, a list of attributes, the product of an interpretation, a collection of context schemata, paths in information retrieval, slots in object-oriented languages, buttons which are functional, customizable and shareable, possible worlds, assumptions under which a statement is true or false, a special, buffer-like data structure, an interpreter which controls the system's activity, the characteristics of the situation and the goals of the knowledge use, entities (things or events) related in a certain way, the possibility that permits to listen what is said and what is not said*".

Contextualization is an operation largely applied in artificial intelligence, natural language processing, databases and ontologies, communication, electronic documentation and machine vision. The principal benefits from contextualization are the *modularity of representation*, *context dependent semantics*, and *focused information access* [27]. As far as SA is concerned, a context or if one prefers, a representation of a situation, is a means to encapsulate information while eliminating the unnecessary details, makes it possible to refer to a given representation of the world while allowing different interpretations on the meaning of this precise representation and finally gives a access to a mechanism to focus on details when required.

Using the notation defined earlier (section 16.2.3), a context or a situation s is a view on the global state of an agent $\mathcal{A}$ built on a given database KB. This view can be shared by multiple agents through communication links. As will be shown below, contexts are means to make reasoning local allowing for example an agent to hold incoherent beliefs or to deal with incomplete information and knowledge.

Contextualizations are usually based on criteria such as

- *time*: limits due to real time applications requirements or planning objectives,
- *space*: limits due to range of sensors or territorial frontiers,
- *function*: discrimination according to objects functions or agents social roles,
- *structure*: distinction between cooperative or egoistic behavior.

Agents performing situation analysis are embedded in complex and dynamically changing environments. Many problems arise (1) from the unpredictability and instability of such environments, (2) from the particularities of the SA tasks to accomplish and finally (3) from the agents own limitations, both physical and mental.

1. The unpredictability and instability of the environment will force the agent to concentrate on the most certain information available and leave unmeasured events that are not yet accessible.

   In this case, the result of contextualization is for example the constitution of the σ-algebra used in probability theory (see section 16.5.3). Similarly, the generic operation consisting in the specification of upper and lower bounds over sets of events is also a form of contextualization. This operation is present in different theories such as Demspter-Shafer theory (belief and plausibility measures or lower and upper probabilities) and rough set theory (lower and upper approximations).

2. Depending on the complexity of the environment, the different tasks involved in SA will not require the same level of attention, the same depth of reasoning and nor be subject to the same reaction delays. Consequently the agents will only consider limited time and space frames in order to efficiently answer operational requirements. These limits are imposed voluntarily by designers of SA systems, implemented by experienced game players and but also innate to many biological systems.

   Two models have been proposed for the partition of sets of possibles worlds (see section 16.6.1), the Rantala and sieve models. Rantala models [28] are a modification of the standard Kripke model semantics that incorporate the notion of *impossible worlds*, allowing to distinguish them from possible worlds. In these impossible worlds anything can hold even contradictions. The notion captures the fact that a non-ideal agent may believe in things that are not consistent, false, etc. but are nonetheless considered as epistemic alternatives. Sieve models have been proposed by Fagin and Halpern in 1988 [22] in order to prevent the problem of omniscience by introducing a function that act as a sieve. Instead of introducing nonstandard world or situations, sieve models introduce *segregation between formulas* that can be known or believed and other that cannot. The sieve function indicates in fact if the agent is **aware** of a given formula in a given situation. Being aware amounts at knowing or believing the formula in question.

3. It is a common practice in SA to consider resource bounded agents, even implicitly. In economics the notion of unbounded rationality refers to the consideration of all possible alternatives and choosing the best one often using optimization techniques. The opposite view of rational choice theory, bounded rationality, rather considers that there are finite limits to information and calculations a human brain or a mechanical memory device can hold *i.e.* Bremermann's computational limit. This view also holds that deliberation costs should be included in models, limiting furthermore rationality for the sake of economy.

According to many authors [29, 30, 31], in neutrosophy the attribution of truth values can be bound to specific circumstances making it thus a contextual theory of truth [32]. Unary neutrosophic connectives such as A', Anti-A, Neut-A (see section 16.5.1), seem particularly interesting for the manipulation of contextual concepts.

16.4.3 Enrichment of the universe of discourse

The *universe of discourse* is the set of objects (concrete or abstract) considered in a given context. It could be a set of classes, a set of targets, a set of actions to take, etc, but also a set of possible worlds (*i.e.* of possible states of the world). Let S represent the universe of discourse, the set of all possible outcomes of an experiment:

$$S = \{s_1, s_2, \ldots, s_n\} \tag{16.1}$$

The universe of discourse is in a sense, the result of a contextualization operation (section 16.4.2) since all objects existing in the world are not present in this set; a choice has been made (voluntarily or not). It is then the support for problem-solving situation and represents the objects about which we are able to talk.

However, it represents an ideal model assuming a perfect description. Unfortunately, real world is often different and more complex than expected. Indeed, on one hand the agents have a limited access to knowledge and on the other hand, objects in the real world itself are not clear cut and a perfect description is in general impossible. These features of reality cannot in general be taken into account in the modelization of the problem (*i.e.* in the definition of the universe of discourse). Hence, a solution to deal with the two different kinds of limitations we face to in SA, epistemic limitation (due to cognitive limitations of the agents, human or artificial) and ontological limitation (due to the nature of things), (section 16.3), is to artificially enrich the universe of discourse.

1. The failure of the sources of knowledge of an agent leads mainly to *indiscernibility* (see section 16.3). Indeed, an epistemic limitation implies the necessity of considering other objects than those originally present in S. In particular, the incapacity of a agent to distinguish between two objects s_1 and s_2 at a given time, in a given context is represented by $s_1 \cup s_2$ which is another object, built from S but not explicitly in S. $s_1 \cup s_2$ is then the best answer the agent can give at a given time, even if it knows that the answer is either s_1 **or** s_2.

   In probability theory, because of the axiom of additivity, we cannot refer to $s_1 \cup s_2$ independently of the rest of the universe. Indeed, $\mu(s_1 \cup s_2) = \mu(s_1) + \mu(s_2) - \mu(s_1 \cap s_2)$ if μ is a probability measure over S. Hence, to account for this limitation of the access to knowledge (epistemic limitation), we can enrich the universe of discourse and consider the power set of S, *i.e.* the set of all subsets of S:

$$2^S = \{A | A \subseteq S\} = \{\varnothing, s_1, s_2, \ldots, s_n, (s_1, s_2), \ldots, (s_{n-1}, s_n), \ldots, S\}^2 \tag{16.2}$$

   where $\varnothing$ denotes the empty set. This enrichment of the universe of discourse allows ignorance and uncertainty to be best represented, as well as a supplementary types of conflict to be taken into account. If probability theory is based on the classical set notion, the notion of power set is the basis for Dempster-Shafer theory (see section 16.5.3 for a brief description), possibility theory and rough sets theory. In this context, we can assign measures to every subset of S, independently of the others. Note finally that Dempster-Shafer theory is based on the assumption of a universe of discourse composed by an exhaustive list of mutually exclusive elements [33], a very restrictive constraint in practice.

2. Another limitation is due to the fact that the observable world is more complex than we can describe. This ontological limitation is linked to the properties of the objects and has nothing to do with our perception means. For example, $s_1 \cap s_2$ represents another object composed by both s_1 **and** s_2. It is neither s_1 nor s_2 but something between them. Hence, yet another extension is the construction of the hyper-power set constituted of all the combinations of the union and intersection operators applied to the elements of S:

$$D^S = \{\varnothing, s_1, \ldots, s_n, (s_1 \cup s_2), \ldots, S, (s_1 \cap s_2), \ldots, (s_1 \cap s_2) \cup s_3, \ldots\} \tag{16.3}$$

   If the elements of S are mutually exclusive ($s_i \cap s_j = \varnothing$, for all $i \neq j$), then $D^S = 2^S$. However, considering D^S is a more general case allowing $s_i \cap s_j \neq \varnothing$, *i.e.* allowing objects of the universe of discourse to overlap. An example, is an universe constituted of vague concepts. Extending the definition of the probability measure the hyper-power set is the principle of Dezert-Smarandache theory [33]. In this framework, no initial assumption on the mutually exclusivity on S is imposed,

[2] Here, (s_1, s_2) is used to denote $(s_1 \cup s_2)$.

and the exhaustivity is somewhat delayed since new objects can be constructed on those of S. A brief description of DSmT is proposed in section 16.5.3.

Therefore, we can say that while Dempster-Shafer theory of evidence is an *epistemic theory* since it only represents epistemic limitations, Dezert-Smarandache is basically an *epistemic and ontological theory* since this framework combines both epistemic and ontological view points.

16.4.4 Autoreference

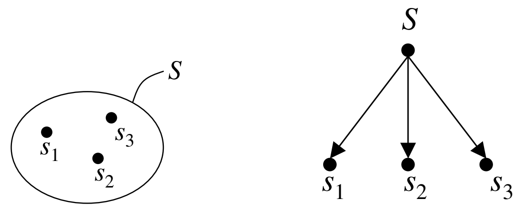
By *autoreference* we mean the capacity of an agent for introspection or selfreference. For example, an agent should be granted the capacity of holding beliefs belief about its own declarations, and not only about the declarations of the other agents.

1. A classical mean for modelling autoreference is by the way of hypersets. The notion of hyperset has been first introduced by Aczel [34] and Barwise and Etchemendy [35] to overcome Russell's paradox[3]. A recursive definition extends the notion of classical set, allowing hypersets to contain themselves, leading to infinitely deep sets (for example, $x = 1 + 1/x$). A *well-founded set* is a set without infinite descending membership sequence, whereas the others are called *non-well-founded sets*.

2. In modal logics, Kripke structures are used as a semantics (see section 16.6.1). In a Kripke structure, an accessibility relation is defined over a set of possible worlds which models either the structure of the world or the agent properties. The desired properties of an agent are then modeled by imposing some properties to the accessibility relation. In particular, if the relation is reflexive and transitive, then the agent possesses the capacity of *positive introspection* (the agent knows that it knows). Also if the relation is an equivalence relation, the agent is capable of formulating declarations about its ignorance (*negative intropection*).

Although these two models, hypersets and Kripke models, are presented here as distinct ones, both are semantics of (multi-agent) modal logics. In [37, 38], it has been proven the equivalence of both semantics. Indeed, with the notion of hyperset comes the graph metaphor which replaces the "container" metaphor used in classical set theory (see figure 16.5). By definition, a *graph* G is a pair (S, R), where S is a set of nodes and R is a relation over S. A *labeled graph* is a triple $\mathcal{S} = \langle S, \mathsf{R}, \pi \rangle = \langle G, \pi \rangle$ where G is a graph and π is a valuation function from P to 2^S, with P being a set of propositional variables, that assigns to each p of P a subset of S. However, a Kripke model can be viewed as a *directed labeled graph*, whose

[3] "*Russell's paradox is the most famous of the logical or set-theoretical paradoxes. The paradox arises within naive set theory by considering the set of all sets that are not members of themselves. Such a set appears to be a member of itself if and only if it is not a member of itself, hence the paradox*" [36].



(a) Using "container" metaphor.

(b) Using graph metaphor.

Figure 16.5: Representation of classical sets. Arrows in figure (b) mean that s_i is a member of S.

nodes are the possible worlds, the link between nodes representing the accessibility relation, labeled by truth assignments[4].

First introduced for modal logics and knowledge logics, the model proposed by Kripke appears as an elegant structure for reasoning about knowledge in a multi-agent context. Moreover, it is based on the notion of *possible world*, which is close to the intuitive notion of situation. Hence, we choose it as the basic structure for situation analysis. In section 16.6, we develop our argumentation to connect Kripke structures with neutrosophic frameworks. After a more formal description of Kripke structures (section 16.6.1), we first extend this structure to neutrosophic logic (section 16.6.2). Then, considering mainly the notion of possible worlds, we extend probability structures to DSm structures (section 16.6.3). And finally, we make the connection between DSmT and neutrosophic logic through Kripke structures (section 16.6.4).

16.5 Neutrosophic frameworks for Situation Analysis

16.5.1 Neutrosophy

Neutrosophy is presented by F. Smarandache as "*a new branch of philosophy, which studies the origin, nature, and scope of neutralities, as well as their interactions with different ideational spectra*" [9]. It is formalized as follows:

> Let A be an idea, a proposition, a theory, an event, a concept, an entity. Then, using different unary operators, we define
>
> - A', a version of A;
> - Anti-A, the opposite of A;

[4]Although the demonstration proposed in [37, 38] is more complex (!) it lies on the previous remark.

- Non-A, what is not A;
- Neut-A, what is neither A nor Anti-A.

Neut-A represents a neutrality in between the two extremes, A and Anti-A. Hence, between A and Anti-A there is a continuum-power spectrum of neutralities Neut-A, A – Neut-A – Anti-A. Note that Non-A is different from Anti-A (Non-$A \neq$ Anti-A), but also that Anti-$A \subset$ Non-A, Neut-$A \subset$ Non-A, $A \cap$ Anti-$A = \varnothing$, $A \cap$ Non-$A = \varnothing$.

We give below an example for multi-agent situation analysis:

> Let's assume a system composed of n agents $\mathcal{A}_1, \ldots, \mathcal{A}_n$. Let call KB_i the Knowledge-Base of agent i, $i = 1, \ldots, n$. Then,
>
> - KB_1 is all the information agent $\mathcal{A}_1$ has access to;
> - KB_1' is another version of KB_1: for example, an update of KB_1, or KB_1 issued from a partition of the sources of information of $\mathcal{A}_1$, hence another view of KB;
> - Anti-KB_1 is all the information agent $\mathcal{A}_1$ has not access to (or the information it did not use for a given representation of the situation);
> - Non-KB_1 is all the information agents $\mathcal{A}_2, \ldots, \mathcal{A}_n$ have access to, but not shared with $\mathcal{A}_1$ plus the information nobody has access to;
> - Neut-KB_1 is all the information agents $\mathcal{A}_2, \ldots, \mathcal{A}_n$ have access to, but not shared with $\mathcal{A}_1$.

The only formal approaches derived from neutrosophy that will be studied in this chapter are: The neutrosophic logic introduced by Smarandache [10, 11] and the Dezert-Smarandache theory proposed by Dezert and Smarandache [33, 39]. In sections 16.5.2 and 16.5.3 we review the basics of these approaches.

16.5.2 Neutrosophic logic

Neutrosophic logic (NL) is a method for neutrosophic reasoning. This non-classical logic is a multiple-valued logic which generalizes, among others, the fuzzy logic. It is the "*(first) attempt to unify many logics in a single field*" [10].

While in classical logic, a concept (proposition) A is either TRUE or FALSE, while in fuzzy logic A is allowed to be more or less TRUE (and consequently more or less FALSE) using truth degrees, in neutrosophic logic, a concept A is $T\%$ TRUE, $I\%$ INDETERMINATE and $F\%$ FALSE, where $(T, I, F) \subset \|^-0, 1^+\|^3$. The interval $\|^-0, 1^+\|$ is an *hyperreal interval*[5], the heigh part of this notation refering to a

[5]Hyperreals - Non-standard reals (hyperreals) have been introduced in 1960. Let $[0, 1]$ be the real standard interval *i.e.* , the set of real numbers between 0 and 1. An extension of this interval is to replace the lower and lower bounds by the non-standard counterparts $^-0$ and 1^+, being respectively $0 - \epsilon$ and $1 + \epsilon$, where $\epsilon > 0$ is an infinitesimal number (*i.e.* such that for all integer $n > 0$, $\epsilon < \frac{1}{n}$).

three-dimensional space. As a general framework, neutrosophic logic corresponds to an extension in three distinct directions:

1. With A, are considered Non-A, Anti-A, Neut-A, and A';
2. The semantics is based on **three independent** assignments, not a single one as it is commonly used in the other logics;
3. These three assignments take their values as **subsets** of the **hyperreal** interval $\|^{-}0, 1^{+}\|$, instead in $[0, 1]$.

A is thus characterized by a triplet of truth-values, called the *neutrosophical value*:

$$\mathrm{NL}(A) = (T(A), I(A), F(A)) \tag{16.4}$$

where $(T(A), I(A), F(A)) \subset \|^{-}0, 1^{+}\|^{3}$.

16.5.3 Dezert-Smarandache theory (DSmT)

Because the theory proposed by Dezert and Smarandache is presented as a generalization of Dempster-Shafer theory, the latter being itself interpreted as a generalization of probability theory, we briefly review the basics of these two theories before introducing DSmT.

A *probability space* is a 3-tuple $\mathcal{P} = \langle S, \chi, \mu \rangle$ where:

- $S = \{s_1, s_2, \ldots, s_n\}$ is the sample space, the set of the elementary events, the set of all outcomes for a given experiment;
- χ is a σ-algebra of S;
- μ is a probability assignment from χ to $[0, 1]$.

To each element of χ is assigned a non-negative real number $\mu(A)$, a *probability measure of* A (or simply probability of A) that must satisfy the following axioms: (1) $\mu(A) \geq 0$; (2) $\mu(S) = 1$; (3) $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ if $A_i \cap A_j = \varnothing$ for $A_i \neq A_j$.

Axiom 3 is also known as the condition of σ-additivity, or simply *axiom of additivity* and plays a crucial role in the theory of probability. Indeed, it imposes a restriction on the measurable sets (*i.e.* the set to which we are able to assign probability measures), since one direct consequence is $\mu(\overline{A}) = 1 - \mu(A)$, where $\overline{A} = S \backslash A$. In other words, $\mu(A)$ does not depend on any $\mu(B)$ such that $B \subset A$.

The theory of evidence has been originally developed by Dempster in 1967 in his work on upper and lower probabilities [7], and later on by Shafer in its famous book *A Mathematical Theory of Evidence* [8], published in 1976. Often interpreted as an extension of the Bayesian theory of probabilities, the theory of evidence offers the main advantage of better representing uncertainty because the measures are defined on the power set of the universe of discourse, instead of the universe itself as the probability theory does. This particularity leads to the relaxation of the additivity axiom of the probability theory by a less restrictive one, a super-additivity axiom.

A *belief function* is defined from 2^S to $[0,1]$, satisfying the following axioms: (1) $\mathrm{Bel}(\varnothing) = 0$; (2) $\mathrm{Bel}(S) = 1$; (3) For every positive integer n, and for every collection $A_1, \ldots, A_n$ of subsets of S, $\mathrm{Bel}(A_1 \cup \ldots \cup A_n) \geq \sum_i \mathrm{Bel}(A_i) - \sum_{i<j} \mathrm{Bel}(A_i \cap A_j) + \ldots + (-1)^{n+1}\mathrm{Bel}(A_1 \cap \ldots \cap A_n)$. Contrary to the probability measure, the belief measure is non-additive and the axiom of additivity for probability theory is replaced by an axiom of *superadditivity*. The main consequence of this axiom is that **every** element of the power set of S is measurable. Hence, we can have $\mathrm{Bel}(A) > \mathrm{Bel}(B)$ if $B \subset A$.

A belief function is often defined using a *basic probability assignment* (or basic belief assignment) m from 2^S to $[0,1]$ that must satisfy the following conditions: (1) $m(\varnothing) = 0$ and (2) $\sum_{A \in 2^S} m(A) = 1$. Then we have $\mathrm{Bel}(A) = \sum_{B \subseteq A, B \in 2^S} m(B)$.

Dezert-Samrandache theory (DSmT) [33, 39] is another extension in this direction since all the elements of the hyper-power set are measurable. Then a *general basic belief mass* is defined from D^S to $[0,1]$, satisfying the following conditions:

$$m(\varnothing) = 0 \text{ and } \sum_{A \in D^S} m(A) = 1 \tag{16.5}$$

Hence, for example elements of the type of $s_i \cap s_j$, $i \neq j$ are allowed to be measured. The *general belief function* is then defined by:

$$\mathrm{Bel}'(A) = \sum_{B \subseteq A, B \in D^S} m(B) \tag{16.6}$$

We note Bel' to distinguish between the belief function in the Shafer sense, Bel.

DSmT is thus a more general framework that deals with both ontological and epistemic uncertainty. However, as most of quantitative approaches it lacks a formal structure for reasoning. In the following section, we propose a way to add such semantics to DSmT.

16.6 Possible worlds semantics for neutrosophic frameworks

The possible world semantics provides an intuitive means for reasoning about situations. It delivers a general approach to providing semantics to logical approaches with applicability to neutrosophic logic (section 16.6.2). Moreover, possible worlds semantics is often borrowed from logical approaches to fill the lack of semantics of numerical approaches, as it will be detailed below (section 16.6.3).

16.6.1 Kripke model

A Kripke model [40] is a mathematical structure that can be viewed as a *directed labeled graph*. The graph's nodes are the possible worlds s belonging to a set S of possible worlds, labeled by truth assignments π. More formally,

> A *Kripke model* is a triple structure $\mathcal{S}_K$ of the form $\langle S, \mathsf{R}, \pi \rangle$ where
>
> - S is a non-empty set (the set of possible worlds);
> - $\mathsf{R} \subseteq S \times S$ is the accessibility relation;
> - $\pi : (S \longrightarrow P) \longrightarrow \{0; 1\}$ is a truth assignment to the propositions per possible world.
>
> where $P = \{p_1, \ldots, p_n\}$ is a set of propositional variables, and $\{0; 1\}$ stands for $\{$True; False$\}$.

A world s is considered *possible* with respect to another world s' whenever there is an edge linking s and s'. This link is defined by an arbitrary binary relation, technically called the *accessibility relation* Figure Fig. 16.6 illustrates the following example:

> An agent is wondering if "it is raining in New York" (ϕ) and if "it is raining in Los Angeles" (ψ). Since this agent has no information at all about the situation, it will consider possible situations (worlds)$S = \{s_1, s_2, s_3, s_4\}$:
>
> - A situation s_1 in which it is both raining in New York and in Los Angeles, *i.e.* $\pi(s_1)(\phi) =$ True and $\pi(s_1)(\psi) =$ True.
> - A situation s_2 in which it is raining in New York but not in Los Angeles, *i.e.* $\pi(s_2)(\phi) =$ True and $\pi(s_2)(\psi) =$ False.
> - A situation s_3 in which it is not raining in New York and raining in Los Angeles, *i.e.* $\pi(s_3)(\phi) =$ False and $\pi(s_3)(\psi) =$ True.
> - A situation s_4 in which it is neither raining in New York nor in Los Angeles, *i.e.* $\pi(s_4)(\phi) =$ False and $\pi(s_4)(\psi) =$ False.

16.6.1.1 modelling the structure of the world

A very interesting feature of Kripke model semantics, is that it is possible to generate axioms for the different systems of modal logic by expressing conditions on the accessibility function defined on $\mathcal{S}_K$. These conditions can be used to express properties or limitations of agents (according to a given model of the world). For example, any epistemic system built upon a Kripke model satisfying a *reflexive* accessibility relation satisfies also the *true knowledge axiom* (T). If the model satisfies a *reflexive* and *transitive* accessibility relation, it satisfies also the axiom of *positive introspection* (4). Satisfaction of the

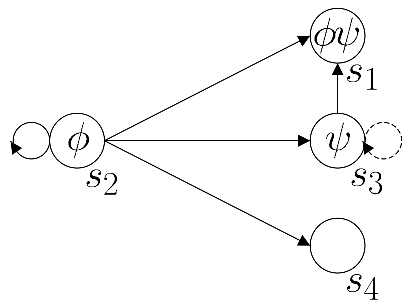


Figure 16.6: Example of a set of possible worlds and their accessibility relations.

axiom of *negative introspection* (5) is given by an *equivalence* relation (see table 16.1). System **K45** is obtained by making *transitive* and *Euclidian* the accessibility function, whereas **KD45** which is sometimes used to model evidential reasoning on Dempster-Shafer structures (see section 16.6.3.2) is obtained by making R *transitive*, *Euclidian* and *serial*. This is summarized in , and explained below.

Table 16.1: Axioms, epistemic logic systems and accessibility relations between possible worlds.

Accessibility relation (R)	Axioms
Reflexive	(T) $K\phi \rightarrow \phi$ (True knowledge)
Reflexive + Transitive	(4) $K\phi \rightarrow KK\phi$ (Positive introspection)
Equivalence	(5) $K\phi \rightarrow K\neg K\phi$ (Negative intropection)

16.6.1.2 Truth assignment

As previously said, to each world $s \in S$, there is an associated truth assignment $\pi(s)$ defined from P to $\{0;1\}$ such that:

$$\pi(s)(p) = \begin{cases} 1 \text{ if } s \vDash p \\ 0 \text{ if } s \nvDash p \end{cases} \tag{16.7}$$

where $p \in P$. $s \vDash p$ means that the world s entails the proposition p, or in other words, that p is TRUE in s.

The assignments $\pi(s)$ are expected to obey to the classical definitions of the connectives so that for example $\pi(s)(p) = S \backslash \pi(s)(p)$, $\pi(s)(p \wedge q) = \pi(s)(p) \cap \pi(s)(q)$, etc.

A *formula* is any composition of some elements of P with the basic connectives $\neg$ and $\wedge$. Let call Φ the set of formulae and ϕ an element of Φ. For example, $\phi_1 = p_1 \wedge \neg p_2$, $\phi_2 = \neg p_1$, $\phi_3 = p_1 \wedge \ldots p_n$, $\phi_i \in \Phi$, $i = 1, \ldots, n$. Hence, the truth assignments $\pi(s)$ are also defined for any formula of Φ, $\pi(s)(\phi)$ being equal to 1 if ϕ is TRUE in s.

To each p of P, there is an associated *truth set* A_p of all the elements of S for which $\pi(s)(p)$ is TRUE:

$$A_p = \{s \in S | \pi(s)(p) = 1\} \tag{16.8}$$

A_p is then the set of possible worlds in which p is TRUE, and can also be noted $A_p = \{s \in S | s \vDash p\}$. By extension, to each formula ϕ is associated a truth set, A_ϕ.

Note that the elements of P are not necessarily mutually exclusive. A way to obtain mutually exclusive elements is to build the set $\mathcal{A}_t$, the set of basic elements, where a basic element[6] is a formula of the (conjunctive) form $\delta = p'_1 \wedge \ldots \wedge p'_n$ with p'_i being either p_i or $\neg p_i$, $p_i \in P$. Any formula $\phi \in \Phi$ can then be written in a disjunctive form as $\phi = \delta_1 \vee \ldots \vee \delta_k$, with $\delta_i \in \mathcal{A}_t$.

To each world s, there is an associated basic element δ of $\mathcal{A}_t$ describing thus the truth values of the propositions of P in S. Whereas many worlds can be associated to the same basic element, a basic element can be associated with any world (see example of section 16.6.3.4). The basic elements are just an alternate way to specify the truth assignment π.

16.6.1.3 Multi-agent context

The definition of $\mathcal{S}_K$ can easily be extended to the multi-agent case. Indeed, if we consider a set of agents $\mathcal{A}_1, \ldots, \mathcal{A}_n$, then on the same set of possible worlds S, and with the same truth assignment π, we can define n accessibility relations R_i, $i = 1, \ldots, n$, one per agent.

The different conditions on the R_is will characterize then the different properties of the $\mathcal{A}_i$s, facing to the same situation.

16.6.2 Kripke structure for neutrosophic propositions

We introduced in section 16.5.2, the basics of neutrosophic logic.

While is classical logic, a formula ϕ is simply characterized by its truth value $\pi(\phi)$ being either 0 or 1 (TRUE or FALSE), in neutrosophic logic ϕ is allowed to be $T\%$ TRUE and $F\%$ FALSE, and $I\%$ INDETERMINATE. ϕ is thus characterized by a triplet of truth-values, called the *neutrosophical value*:

$$\text{NL}(\phi) = (T(\phi), I(\phi), F(\phi)) \tag{16.9}$$

where $(T(\phi), I(\phi), F(\phi)) \subset \|^{-}0, 1^{+}\|^3$, $\|^{-}0, 1^{+}\|$ being an interval of hyperreals.

In an equivalent manner as it is done in quantum logic, where Kripke structures are extended to deal with fuzzy propositions [41], we propose here to extend the Kripke structure to deal with neutrosophic assignments. Hence, we have,

A *Kripke model for neutrosophic propositions* is a triple structure $\mathcal{S}_K^{NL}$ of the form $\langle S, \mathsf{R}, \vec{\pi} \rangle$ where

[6] A basic element is sometimes called an atom.

- S is a non-empty set (the set of possible worlds);
- $\mathsf{R} \subseteq S \times S$ is the accessibility relation;
- $\vec{\pi} = (\pi_T, \pi_I, \pi_F)$ is a neutrosophic assignment to the propositions per possible world, *i.e.* $\pi : (S \longrightarrow P) \longrightarrow \|^{-}0, 1^{+}\|$ with π being either π_T or π_I or π_F.

where $P = \{p_1, \dots, p_n\}$ is a set of propositional variables.

The "truth" assignment π of a classical Kripke model becomes then $\vec{\pi} = (\pi_T, \pi_F, \pi_I)$, a three-dimensional assignment, where π_T is the truth assignment, π_F is the falsity assignment and π_I is the indeterminacy assignment. Hence, in each possible world s of S, a proposition ϕ can be evaluated as $\pi_T(s)(\phi)$ TRUE, $\pi_F(s)(\phi)$ FALSE and $\pi_I(s)(\phi)$ INDETERMINATE. It follows that to ϕ is associated a truth-set A^T_ϕ, a falsity-set A^F_ϕ and an indeterminacy-set A^I_ϕ:

$$A^T_\phi = \{s \in S | \pi_T(s)(\phi) \neq 0\}$$
$$A^F_\phi = \{s \in S | \pi_F(s)(\phi) \neq 0\}$$
$$A^I_\phi = \{s \in S | \pi_I(s)(\phi) \neq 0\}$$

Note that A^T_ϕ, A^T_ϕ and A^T_ϕ are (1) no longer related, (2) fuzzy sets and may overlap.

16.6.2.1 Knowledge and belief

Halpern in [42] gives the following definitions for knowledge and belief in PWS:

- ϕ is **known** if it is TRUE in **all** the possible worlds s of S
- ϕ is **believed** if it is TRUE in **at least one** possible world s of S

On the other hand, Smarandache [10] uses the notion of world and states that $T(\phi) = 1^+$ if ϕ is TRUE in **all** the possible worlds s of S (absolute truth) and $T(\phi) = 1$ if ϕ is TRUE in **at least one** possible world s of S (relative truth) (see Tab. 16.2). Hence, in the neutrosophical framework, we can state the following definitions for knowledge and belief: ϕ is **known** if $T(\phi) = 1^+ \equiv F(\phi) =^- 0$ and ϕ is **believed** if $T(\phi) = 1 \equiv F(\phi) = 0$. Table 16.2 shows several special cases.

Furthermore, one can consider the unary operators of neutrosophic logic (Non-ϕ, Anti-ϕ, Neut-ϕ, ϕ') to model new epistemic concepts but also as a means to represent situational objects, such as neutral situation, environment (to be detailed in the final version).

16.6.3 Probability assignments and structures

Let S be the frame of discernment, s a singleton of S and A any subset of S. In probability theory, measurable objects are singletons s of S. The measures assigned to any subsets A of S are guided by the additivity axiom. Hence, measurable elements belong to a σ-algebra χ of 2^S. In Dempster-Shafer theory,

Table 16.2: Neutrosophical values for special cases (adapted from [10]).

ϕ is ...	in ... poss. world(s)	Neutrosophical value
true	**all**	$T(\phi) = 1^+ \equiv F(\phi) =^- 0$
false		$F(\phi) = 1^+ \equiv T(\phi) =^- 0$
indet.		$I(\phi) = 1^+$
true	**at least one**	$T(\phi) = 1 \equiv F(\phi) = 0$
false		$F(\phi) = 1 \equiv T(\phi) = 1$
indet.		$I(\phi) = 1$
indet.	**no**	$I(\phi) =^- 0$
not indet.	**at least one**	$I(\phi) = 0$

any element of the power set of S, 2^S is measurable. Finally, Dezert-Smarandache theory allows any element of the hyper-power set of S, D^S, to be measured. Apart these extensions to probability theory that rely on the definition set of the probability measure, there exists a clear interest for giving a better semantics to these numerical approaches. For its probabilistic logic, Nilsson uses the possible worlds semantics to build a "semantical generalization of logic", combining logic with probability theory [6] (see section 16.6.3.1). Later on, Fagin and Halpern [4] and also Bundy [43] extend Nilsson's structure for probabilities allowing all elements of the power set to be measurable, leading to a general structure just as Dempster-Shafer theory generalizes probability theory, the Dempster-Shafer structure (see section 16.6.3.2).

In the following, after a brief review of Nilsson and Demspter-Shafer structures, we extend the latter and propose a Dezert-Smarandache structure (section 16.6.3.3), combining the DSmT framework and the possible worlds semantics. To end this part, we propose in section 16.6.3.4 an example of the potential interest of such a structure.

16.6.3.1 Nilsson structure

A *Nilsson structure* is a tuple $\mathcal{S}_N = \langle S, \chi, \mu, \pi \rangle$ where

- $S = \{s_1, s_2, s_3, \ldots\}$, the set of all possible worlds;
- χ, a σ-algebra of subsets of S;
- μ, a probability measure defined on χ;
- $\pi : (S \longrightarrow P) \longrightarrow \{0; 1\}$, is a truth assignment to the propositions per possible world.

with P being a set of propositional variables.

Note that $\langle S, \chi, \mu \rangle$ is a probability space, and Nilsson structure is also called a probabilistic structure. In this kind of structure, the only measurable elements are those of χ. However, if we are interested in any other formula of Φ, the best thing we can do is to compute the *inner* and *outer measures* [4] defined respectively by

$$\mu_*(A) = \sup\{\mu(B)|B \subseteq A, B \in \chi\} \text{ and } \mu^*(A) = \inf\{\mu(B)|B \supseteq A, B \in \chi\}$$

The unknown value $\mu(A_\phi)$ is replaced by the interval:

$$\mu_*(A_\phi) \leq \mu(A_\phi) \leq \mu^*(A_\phi) \tag{16.10}$$

Hence, instead of a single probability measure μ from χ to $[0,1]$, we can compute a pair of probability measures μ_* and μ^*.

Because in a Nilsson structure, μ is defined on χ (the set of measurable subsets) means that χ_π (the image of χ by π) is a sub-algebra of χ to ensure that $\mu(\phi) = \mu(A_\phi)$, for all $\phi \in \Phi$. Dropping this condition is a means to extend μ to 2^S (hence Nilsson structure) and leads to Dempster-Shafer structure as formalized in [4][7] and detailed below. The probability measure μ is then replaced by its inner measure μ_*.

16.6.3.2 Dempster-Shafer structure

Nilsson structure can be extended using the inner measure, *i.e.* allowing all the elements of 2^S to be measurable. Because the inner measure turns to be the belief measure introduced by Shafer in its theory of evidence [8], the resulting structure is called *Dempster-Shafer structure*. Note that χ and π are no longer required to be related in any sense.

A *Dempster-Shafer structure* [4] is a tuple $\mathcal{S}_{DS} = \langle S, 2^S, \text{Bel}, \pi\rangle$ in which

- $S = \{s_1, s_2, s_3, \ldots\}$, the set of all possible worlds;
- 2^S, the powerset of S;
- Bel, a belief measure on 2^S;
- $\pi : (S \longrightarrow P) \longrightarrow \{0;1\}$, is a truth assignment to the propositions per possible world.

with P being a set of propositional variables.

Note that we can simply write $\mathcal{S}_{DS} = \langle S, \text{Bel}, \pi\rangle$, where Bel is a belief function $\text{Bel} : 2^S \longrightarrow [0,1]$, in the Shafer sense (see section 16.5.3).

A Nilsson structure is then a special case of Dempster-Shafer structures, in which

$$\mu_*(A_\phi) = \mu^*(A_\phi) = \mu(A_\phi) \tag{16.11}$$

for any $\phi \in \Phi$.

[7] Another way is to consider a partial mapping π, leading to Bundy's structure of incidence calculus [43].

16.6.3.3 Dezert-Smarandache structure

In [33], the authors propose a generalization of Dempster-Shafer theory defining a belief function on the hyper-power set instead of the power set as Shafer. This theory is called Dezert-Smarandache theory or simply DSmT. In an equivalent manner to the extension of Nilsson's structure to DS structure, the definition of μ can be extended to D^S, allowing all elements of the hyper-power set to be measurable. We obtain then what we can call a *Dezert-Smarandache structure* (DSm structure), an extension of the DS structure in an equivalent way as DSmT is an extension of Dempster-Shafer theory.

A *Dezert-Smarandache structure* is a tuple $\mathcal{S}_{DSm} = \langle S, D^S, \mathrm{Bel}', \pi \rangle$ where

- $S = \{s_1, s_2, s_3, \dots\}$, the set of all possible worlds;
- D^S, the hyper-power set of S;
- Bel', a general belief measure on D^S;
- $\pi : (S \longrightarrow P) \longrightarrow \{0; 1\}$, is a truth assignment to the propositions per possible world.

with P being a set of propositional variables.

Note that we can simply write $\mathcal{S}_{DSm} = \langle S, \mathrm{Bel}', \pi \rangle$ where Bel' is the generalized belief function defined on D^S, as defined by Dezert and Smarandache (see section 16.5.3).

16.6.3.4 Example: Ron suits

This example is proposed in [4] as *Example 2.4*:

> "*Ron has two blue suits and two gray suits. He has a very simple method for deciding what color suit to wear on any particular day: he simply tosses a (fair) coin. If it lands heads, he wears a blue suit and if it lands tails, he wears a gray suit. Once he's decided what color suit to wear, he just chooses the rightmost suit of that color on the rack. Both of Ron's blue suits are single-breasted, while one of Ron's gray suit is single-breasted and the other is double-breasted. Ron's wife, Susan, is (fortunately for Ron) a little more fashion-conscious than he is. She also knows how Ron makes his sartorial choices. So, from time to time, she makes sure that the gray suit she considers preferable is to the right (which depends on current fashions and perhaps on other whims of Susan). Suppose we don't know about the current fashion (or about Susan's current whims). What can we say about the probability of Ron's wearing a single-breasted suit on Monday?* [4]"

Let P be a set of primitive propositions, $P = \{p_1, p_2\}$. Let p_1="The suit is gray" and let p_2="The suit is double-breasted". Then $\mathcal{A}_t$, the corresponding set of basic elements is:

$$\mathcal{A}_t = \{p_1 \wedge p_2, p_1 \wedge \neg p_2, \neg p_1 \wedge p_2, \neg p_1 \wedge \neg p_2\}$$

$\mathcal{A}_t$ is thus a set of mutually exclusive hypotheses: "Ron chooses a gray double-breasted suit", ..., "Ron chooses a blue single-breasted suit".

S is the set of possible states of the world, *i.e.* the set of possible worlds, where a state corresponds in this example to a selection of a particular suit by Ron. To fix the ideas, let number the suits from 1 to 4. Hence, $S = \{s_1, s_2, s_3, s_4\}$, s_i being the world in which Ron chooses the suit i. Table 16.3 lists the possible worlds and their associated meaning and atom[8]. Table 16.4 give some sets of worlds of interest and their associated formula. An alternative to describe the state of a world (*i.e.* the truth values of each

Table 16.3: The 4 states of the worlds and their associated basic element.

World	Meaning	Basic element
s_1	Blue single-breasted suit nb 1	$\neg p_1 \wedge \neg p_2$
s_2	Blue single-breasted suit nb 2	$\neg p_1 \wedge \neg p_2$
s_3	Gray single-breasted suit	$p_1 \wedge \neg p_2$
s_4	Gray double-breasted suit	$p_1 \wedge p_2$

propositions in P) is by using π is a truth assignment defined from P to 2^S. For each s in S, we have a truth assignment $\pi(s)$ defined from P to $\{0; 1\}$, such that $\pi(s)(p) = 0$ if p is false in s, and $\pi(s)(p) = 1$ if p is true in s.

Table 16.4: Some subsets of possible worlds of interest and their associated formula.

World(s)	Meaning	Formula
(s_1, s_2)	A blue suit	$\neg p_1$
(s_3, s_4)	A gray suit	p_1
(s_1, s_2, s_3)	A single-breasted suit	$\neg p_2$

Here, we have only 4 measurable events: $\mu(s_1, s_2) = \mu(s_3, s_4) = \frac{1}{2}$, $\mu(\varnothing) = 0$ and $\mu(S) = 1$. The question of interest here (What is the probability of Ron's wearing a single-breasted suit?) concerns another non-measurable event, *i.e.* (s_1, s_2, s_3). In [4], the authors gave this example to illustrate the utility of attributing values to non-measurable events, and then introduce Demspter-Shafer structures. Their conclusion for this example is then that the best we can say is that $\frac{1}{2} \leq \mu(s_1, s_2, s_3) \leq 1$, based on the inner and outer measures.

modelling the problem with 4 states means that given our prior knowledge, these states correspond to the only possible situations after Ron's selection: He will select one and only one suit among the 4 available. However, suppose that the two parts of the suits may have been mixed so we have two pieces (trousers and jacket) on the same coat-hanger. The 4 possible worlds correspond then to the 4 coat-hangers, and no longer to the 4 distinct suits. Imagining that the trousers is inside the jacket, Ron

[8]Note that the basic element $\neg p_1 \wedge p_2$ is associated with any state, while $\neg p_1 \wedge \neg p_2$ is associated with two states, s_1 and s_2.

will select his suit only on the basis of the color of the jacket. Suppose for example, that the coat-hanger he selects supports a blue jacket and gray trousers. Then, waht is the corresponding state of the world? Clearly, this situation has not been considered in the modelisation of the problem, based on a DS structure. However, using a DSm structure allow the elements of the hyper-power set of S to be measurable. Hence, the state resulting of a selection of a mixed suit corresponds to $s_i \cap s_j$, with $i \neq j$. This means that we are in both worlds s_i and s_j, and that with a single selection, Ron selected in fact two suits. So, we allow other events than those forecast to overcome.

One benefit of the resulting structure for situation analysis, is that it provides an interesting framework for dealing with both vagueness and conflict, combining the logical, semantical and reasoning aspect through the possible worlds semantics, and the measuring, combination aspect through the DSmT.

16.6.4 Connection between DSmT and neutrosophic logic in Kripke structures

Here we describe informally a possible connection between Dezert-Smarandache theory and the neutrosophic logic.

Let $\mathcal{S}_{DSm} = \langle S, \mathrm{Bel}', \pi\rangle$ be a DSm structure, and let $\mathcal{S}_K^{NL} = \langle S, \mathsf{R}, \vec{\pi}\rangle$ be the corresponding Kripke structure for neutrosophic propositions. Hence, we define a general neutrosophic structure to be $\mathcal{S}_N = \langle S, \mathrm{Bel}', \mathsf{R}, \vec{\pi}\rangle$, where:

- $S = \{s_1, s_2, s_3, \ldots\}$, the set of all possible worlds;
- Bel', a general belief measure on D^S, the hyper-power set of S;
- $\mathsf{R} \subseteq S \times S$ is the accessibility relation[9];
- $\vec{\pi} = (\pi_T, \pi_I, \pi_F)$ is a neutrosophic assignment to the propositions per possible world, *i.e.* $\pi : (S \longrightarrow P) \longrightarrow \|^-0, 1^+\|$ with π being either π_T or π_I or π_F.

where $P = \{p_1, \ldots, p_n\}$ is a set of propositional variables.

In order to reason on this structure, we need a set of axioms (as it is for example done in [4] for belief and probability) characterizing valid formulae. This can be achieved by imposing conditions on the accessibility relation R, conditions yielding hopefully to neutrosophic behaving agents.

Hence, the aim of this general structure is to conciliate (1) DSmT as a tool for modelling both epistemic and ontological uncertainty, (2) possible worlds for the representation of situations, (3) neutrosophic logic as a general logical approach to deal independently with truth, falsity and indeterminacy, and (4) Kripke structures as a support for reasoning and modelling the properties of a collection of interacting agents.

We finally note, that although a connection can be found or stated, there is *a priori* no trivial link between the neutrosophic assignments $(\pi_T(s)(\phi),\ \pi_F(s)(\phi),\ \pi_I(s)(\phi))$ that quantify truth, falsity and

[9] We consider here the monoagent case, although the extension to the multiagent case is trivial.

indeterminacy of formulae, and the belief granted to the corresponding sets of possible worlds through the general belief function proposed in DSmT, Bel'.

16.7 Conclusion

In this chapter, we proposed a discussion on neutrosophy and its capacity to tackle the situation analysis challenges. In particular, we underlined and connected to neutrosophy four basic ontological principles guiding the modelization in Situation Analysis: (1) allowing statements about uncertainty to be made, (2) contextualization, (3) enrichment of the universe of discourse, (4) allowing autoreference. The advantages of DSmT and neutrosophic logic were studied with these principles in mind. In particular, we highlighted the capacity of neutrosophic logic to conciliate both qualitative and quantitative aspects of uncertainty. Distinguishing ontological from epistemic obstacles in SA we further showed that being based on the power set, Dempster-Shafer theory appears in fact as an epistemic theory whereas Dezert-Smarandache theory, based on the richer hyper-power set, appears capable to deal with both epistemic and ontological aspects of SA. Putting forward the connection between hypersets and Kripke structures as means to model autoreference, we then focused on Kripke structures as an appropriate device for reasoning in SA. In particular, we showed that it is feasible to build a DSm structure upon the possible worlds semantics, an extension of the classical probabilistic and Dempster-Shafer structures. Considering neutrosophic logic, we showed that is could be possible to extend Kripke structures in order to take into account neutrosophic propositions, *i.e.* triplets of assignments on intervals of hyperreal numbers. We also showed how to represent the concepts of belief and knowledge with hyperreal truth (resp. falsity, indeterminacy) assignments on possible worlds. This allows one to introduce a clear qualitative distinction between *certain belief* and *knowledge*, a distinction that is not clear in traditional epistemic logic frameworks. Finally, we proposed a connection between neutrosophic logic and DSmT in the Kripke semantics setting.

16.8 References

[1] Endsley M. R. Garland D. J. , *Situation Awareness Analysis and Measurement*, Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers, 2000.

[2] Roy J., *From data fusion to situation analysis*, in Fourth International Conference on Information Fusion (ISIF, ed.), vol. II, (Montreal, Canada), pp. ThC2–3 – ThC2–10, 2001.

[3] Halpern J. Y., Moses Y., *Knowledge and common knowledge in a distributed environment*, Journal of the Association for Computing Machinery, vol. 37, no. 3, pp. 549–587, 1990.

[4] Fagin R., Halpern J. Y. , *Uncertainty, belief and probability*, Computational Intelligence, vol. 7, pp. 160–173, 1991.

[5] Bundy A., *Incidence calculus*, The Encyclopedia of AI, pp. 663–668, 1992.

[6] Nilsson N., *Probabilistic logic*, Artificial Intelligence, vol. 28, no. 1, pp. 71–87, 1986.

[7] Dempster A., *Upper and Lower Probabilities Induced by Multivalued Mapping*, *Ann. Math. Statist.*, vol. 38, pp. 325–339, 1967.

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

[9] Smarandache F., *Neutrosophy, a New Branch of Philosophy*, 1991. Personal communication.

[10] Smarandache F., *Neutrosophic Logic - a Unifying Field in Logic*, 1991. Personal communication.

[11] Smarandache F., *A Unifying Field in Logics: Neutrosophic Logic, Neutrosophic Probability, Neutrosophic Set*, tech. rep., Western Section Meeting - Meeting n^{o} 951 of the American Mathematical Society, Santa Barbara, 2000.

[12] Pew R. W., *The state of situation awareness measurement; heading toward the next century*, in Situation Awareness Analysis and Measurement (M. Endsley and D. Garland, eds.), pp. 33–50, Mahwah, New Jersey: Lawrence, Erlbaum Associates, 2000.

[13] . TTCP C3I Information Fusion Action Group (AG-2), Information fusion definitions, concepts and models for coalition situation awareness, tech. rep., february 2002.

[14] Pitt D., *Mental representation*, in The Stanford Encyclopedia of Philosophy (E. N. Zalta, ed.), winter 2002 edition ed., 2002.

[15] Audi R., *Epistemology: A contemporary introduction of the theory of knowledge*, Routledge Comtemporary Introductions to Philosophy, London, UK: Routledge, second ed., 2003.

[16] White F. E. , *Data fusion lexicon*, Joint Directors of Laboratories, Technical Panel for C3, Data Fusion Sub-Panel Naval Ocean Systems Center, San Diego, 1987.

[17] Steinberg A. N., Bowman C. L., White F. E., *Revision to the JDL data fusion model*, in Joint NATO/IRIS Conference, (Quebec City), October 1998.

[18] Steinberg A. N., Bowman C. L., *Revisions to the JDL data fusion model*, in Handbook of Multisensor Data Fusion (D. L. Hall and J. Llinas, eds.), The Electrical Engineering and Applied Signal Processing Series, pp. 2–1–2–19, Boca Raton: CRC Press, 2001.

[19] Jousselme A.-L., Maupin P., Bossé E., *Uncertainty in a situation analysis perspective*, in Proceedings of 6th Annual Conference on Information Fusion, (Cairns, Australia), pp. 1207–1214, July 2003.

[20] Bronner G., *L'incertitude*, vol. 3187 of Que sais-je?, Paris: Presses Universitaires de France, 1997.

[21] Rescher N., *Predicting the Future: An Introduction to the Theory of Forecasting*, State University of New York Press, 1997.

[22] Fagin R., Halpern J. Y., *Belief, awareness, and limited reasoning*, Artificial Intelligence, vol. 34, no. 1, pp. 39–76, 1988.

[23] Kleene S. C., *On a notation for ordinal numbers*, Journal of Symbolic Logic, vol. 3, no. 1, pp. 150–155, 1938.

[24] Bochvar D. A., *Ob odnom tréhznacnom iscislénii i égo priménénii k analizu paradosov klassièéskogo rasširennogo funkcjonal'nogo isèislénia (on a three-valued calculus and its application to analysis of paradoxes of classical extended functional calculus)*, Matématicéskij Sbornik, vol. 4, no. 2, pp. 287–308, 1938.

[25] Halldèn S., *The logic of nonsense*, Uppsala, Finland: Uppsala Universitets Arsskrift, 1949.

[26] Theodorakis M., *Contextualization: An Abstraction Mechanism for Information Modeling*, PhD. Thesis, University of Crete - Department of Computer Science, 2001, 186 pages.

[27] Brézillon P, *Context in problem solving: A survey*, The Knowledge Engineering Review, vol. 14, no. 1, pp. 1–34, 1999.

[28] Rantala V., *Impossible world semantic and logical omniscience*, Acta Philosophica, vol. 33, no. 1, pp. 106–115, 1982.

[29] Lucas C., *A logic of complex values*, in Proceedings of the First International Conference on Neutrosophy, Neutrosophic Logic, Neutrosophic Set, Neutrosophic Probability and Statistics (F. Smarandache, ed.), pp. 121–138, 2001.

[30] Liu F., *Name, denominable and undenominable - On neither* $<A>$ *nor* $<Anti-A>$, in Proceedings of the First International Conference on Neutrosophy, Neutrosophic Logic, Neutrosophic Set, Neutrosophic Probability and Statistics (F. Smarandache, ed.), pp. 107–113, 2001.

[31] Liu F., *Comments on neutrosophy*, in Proceedings of the First International Conference on Neutrosophy, Neutrosophic Logic, Neutrosophic Set, Neutrosophic Probability and Statistics (F. Smarandache, ed.), pp. 129–146, 2001.

[32] Gupta A., *Truth*, in The Blackwell Guide to Philosophical Logic (L. Gobble, ed.), Blackwell Philosophy Guides, pp. 90–114, Oxford, UK: Blackwell Publishers, 2001.

[33] Dezert J., *Foundations for a new theory of plausible and paradoxical reasoning*, Information & Security: An International Journal, vol. 9, pp. 13-57, 2002.

[34] Aczel P., *Lectures on non-well-founded sets*, in CLSI Lecture Notes, vol. 9, Harward University Press, 1987.

[35] Barwise J., Etchemendy J., *The Liar. An Essay on Truth and Circularity*, Oxford, UK: Oxford University Press, 1987.

[36] Irvine A. D., *Russell's paradox*, in The Stanford Encyclopedia of Philosophy (E. N. Zalta, ed.), summer 2003 edition ed., 2003.

[37] Moss L. S., *From Hypersets to Kripke Models in Logics of Announcements*, 1999, JFAK - CD-ROM - `http://www.illc.uva.nl/j50/contribs/moss/`.

[38] Gerbrandy J., *Bisimulations on Planet Kripke*, PhD thesis, ILLC/Department of Philosophy - University of Amsterdam, 1999.

[39] Dezert J., *Fondations pour une nouvelle théorie du raisonnement plausible et paradoxal. application à la fusion d'informations incertaines et conflictuelles*, Tech. Rep. RT 1/06769/DTIM, ONERA - Département traitement de l'information et modélisation, 2003.

[40] Kripke S. A., *Semantical analysis of modal logic I - Normal modal propositional calculi.*, Zeitschrift für mathematische Logik und Grundlagen der Mathematik, vol. 9, pp. 67–96, 1963.

[41] Dalla Chiara M. L., Giuntini R., *Quantum logics*, in Handbook of Philosophical Logic (D. Gabbay, ed.), Dordrecht: Kluwer, 2001.

[42] Halpern J. Y., *Reasoning about knowledge: A survey*, in Handbook of Logic in Artificial Intelligence and Logic Programming (D. Gabbay, C. J. Hogger, and J. A. Robinson, eds.), vol. 4, pp. 1–34, Oxford, UK: Oxford University Press, 1995.

[43] Bundy A., *Incidence calculus: A mechanism for probabilistic reasoning*, Journal of Automated Reasoning, vol. 1, pp. 263–283, 1985.

Chapter 17

Application of DSmT for Land Cover Change Prediction

Samuel Corgne, Laurence Hubert-Moy
COSTEL, CNRS UMR LETG 6554, Univ. Rennes 2
Place du recteur Henri Le Moal
35043 Rennes, France.

Grégoire Mercier
TAMCIC, CNRS FRE 2658 team TIME
GET/ENSTB, Technopole Brest-Iroise
CS 83818, F-29238 Brest Cedex, France.

Jean Dezert
ONERA
29 Av. de la Division Leclerc
92320 Châtillon, France.

Abstract: *This chapter presents an environmental application of DSmT for the land cover prediction. The spatial prediction of land cover at the field scale in winter is useful to reduce the bare soils in agricultural intensive regions. Fusion process with the Dempster-Shafer theory (DST) proved to have limitations with the increase of conflict between the sources of evidence that support land cover hypotheses. Several modifications may be used such as source weighting or the hedging methods, but with no benefit in the considered case studied since the conflict may not explain by itself all the bad decisions. Actually, sources of evidence may induce all together a wrong decision. Then, it is necessary to introduce paradoxical information. Nevertheless, sources of evidence that are in use, are defined according to hypothesis "covered soil" or "bare soil" in the frame of DST. We investigate several points of view to define the belief assignments of the hyper-power set of the DSmT from the initial power set of DST. So, smart belief assignments induce a better prediction of bare soils.*

Samuel Corgne is also affiliated with TAMCIC, CNRS FRE 2658, team TIME, GET/ENST Bretagne, France.

17.1 Introduction

In intensive agricultural areas, water quality may be improved by reducing bare soil surfaces during the winter months. In this context, the knowledge of the spatio-temporal variations of the land use and cover as well as the spatial prediction of the land cover at the field scale appear essential for the issue of bare soils reduction. Land-cover prediction, that is useful for stakeholders that manage water-quality programs in focusing on the areas where the probability to find a bare soil is high, requires the identification and characterization of the driving factors of observed land-cover changes. The high variability of the driving factors that motivate land-cover changes between two successive winters induces the integration of uncertainty in the modelling of the prediction process.

Several short-term predictions have been simulated with the Dempster-Shafer (DS) theory in previous studies to assess land-cover distribution in winter on a relatively intensive farming watershed of 61.5km^2 [1]. This study area, located in western France, produces significant amounts of nitrogen before winter infiltration of water. Fusion process with the DS theory proved to have limitations with the increase of conflict between the sources of evidence that support land cover hypotheses. Several modifications may be used (such as source weighting or the Hedging methods) but with no benefit in our application. It appears that conflict may not explain by itself all the bad decisions. Actually, each sources of evidence may induce all together a wrong decision. Then, paradoxical information was introduced to improve the prediction accuracy.

A first application of the Dezert-Smarandache theory on the study area has pointed some results a little bit better than the DS, but the rate for the hypothesis "bare soil" was still inferior to 40% of good prediction. An improvement of the fusion process must be performed specially for this hypothesis. In this application, sources of evidence that are in use, are still defined according to hypothesis "Covered soil" or "Bare soil" in the frame of the Dempster-Shafer theory. Mass functions assignment determined from statistical analysis and expert knowledge are defined to support the hypotheses but the high level of conflict between sources requires a finest mass attribution and a "contextual" fusion process to manage the uncertainty and the paradoxical.

This chapter focuses on the application of the Dezert-Smarandache theory for the land-cover prediction in winter, and more precisely on the transfer from evidence to plausible and paradoxical reasoning. Our objective is to improve the land-cover prediction scores in investigating several points of view to define the belief assignments of the hyper-powerset of the Dezert-Smarandache theory from the initial powerset of the Dempster-Shafer theory. A first part concerns the identification and hierarchization of the driving factors that drive the land cover changes on the studied watershed for their transformation in pieces of evidences for the selected working hypothesis. The other one presents the process of the land cover modelling with the Dezert-Smarandache theory comparatively to the Dempster-Shafer theory and its adaptation for this specific environmental study.

17.2 Determination of information sources

The land cover in winter has been classified from remote sensing images in two land cover categories, "Bare soil" and "Covered soil" that correspond to the two hypotheses of work. The determination of the information sources for each hypothesis for the fusion process consists in identifying and hierarchizing the factors that motivate the land cover changes between winters for the studied period (1996-2003).

17.2.1 Identification of the driving factors of land cover change

The land-cover changes between winters in intensive agricultural regions are characterized by an high spatio-temporal variability depending on factors of several origin (economical, social, political, physics constraints) that need to be carfully defined in the modelling process. The identification of the driving factors of land-cover changes requires to study the land use on a quite long period. A set of 10 satellite images (9 SPOT images and 1 IRS-LISS III —2 per year over 5 years since 1996—) has been acquired, pre-processed and classified. Winter land cover change trajectories were produced by merging successively all classifications [2]. All this data have been integrated in a GIS (Geographic Information System) to identify the crop successions spatially and the land-cover changes between winters on the field scale. A statistical analysis and a meeting with the agricultural experts provided four main driving factors of land-cover changes, namely the field size, the crop successions, the agro-environmental actions and the distance of the fields from farm buildings. All this factors explain the winter land-cover distribution in the categories "Bare soil" or "Covered soil". Then, a hierarchization of the identified driving factors of land-cover change was needed in the fusion process to predict the future land-cover (Mass belief assignment to the sources of evidence), to assess the respective "weight" of each explicative factors.

17.2.2 Hierarchization of the factors of land cover change

The mutual information between the variables has been used to hierarchize the explicative factors of land-cover change. The mutual information analysis is based on the information theory [3]. It is used to outline relations between the variables [4]. For this study, three indicators have been chosen to characterize the relationship between variables that may explicit the land cover evolution between the winters.

- Entropy $\mathcal{H}$: the main property of the information concept is that the quantity of information is maximum when the events are distributed uniformly. It allows to calculate the information quantity between the set of events.

$$\mathcal{H} = \sum_{i=1}^{N} p_i \log p_i,$$

with N number of possible events and p_i probability of event i.

- Mutual Information $\mathcal{I}$: it represents the mutual information between two variables X and Y; it is obtained through the difference between the entropy $\mathcal{H}$ of X, Y and the joint entropy $\mathcal{H}(X,Y)$ as follows.

$$\mathcal{I}(X,Y) = \mathcal{H}(X) + \mathcal{H}(Y) - \mathcal{H}(X,Y).$$

- Redundancy $\mathcal{R}$: It is issued from the entropy and the mutual information. It measures the heterogeneity rate of two variables X, Y.

$$\mathcal{R} = \frac{\mathcal{I}(X,Y)}{\mathcal{H}(Y)}.$$

The process provides a hierarchization of the information quantity for the explicative variables with the variable to explain. The results of the mutual information test (Table 17.1) show that the most representative variable is "Crop successions (1996-2002)", followed by "Size of the fields", "Agro-environmental actions" and "Distance from farm buildings" in decreasing representative order. These results allow to optimise the mass belief assignment for the hypotheses "Bare soil" and "Covered soil", in comparison with an empirical "expert knowledge" method.

	Classes	$N_F(\%)$	$\mathcal{R}$	$\mathcal{I}$
Distance from farm buildings	1: < 1.25	1255 (67.6 %)	0.14 %	0.0006
	2: > 1.25	601 (32.4 %)		
Agro-environmental actions	1: without	1619 (87.2 %)	0.2 %	0.0008
	2: with	237 (12.8 %)		
Field size	1: < 1.5 ha	1517 (81.7 %)	0.97 %	0.0039
	2: > 1.5 ha	339 (18.3 %)		
Crop rotation (1996–2002)	1: (SC W)	1046 (56.4 %)	5.19 %	0.0211
	2: (BS 1W)	301 (16.2 %)		
	3: (BS 2W)	186 (10 %)		
	4: (BS 3W)	179 (9.64 %)		
	5: (BS 4W)	89 (4.8 %)		
	6: (BS 5W)	55 (2.96 %)		

Table 17.1: Explicative variables hierarchization with the mutual information analysis.

Column $N_F(\%)$ of the Table 17.1 indicates the numbers N_F of fields (and their percentage). Column 5 of the table indicates the values of redundancy $\mathcal{R}$ and column 6 the values of mutual information $\mathcal{I}$. In the last row (i.e. crop rotations during 1996-2000) of Table 17.1, six cases have been identified and correspond to

1. (SC W) : soils covered during all winters
2. (BS 1W) : bare soil during one winter
3. (BS 2W) : bare soil during two winters
4. (BS 3W) :bare soil during three winters
5. (BS 4W) : bare soil during four winters
6. (BS 5W) : bare soil during five winters

17.3 Land cover prediction with the Dempster-Shafer Theory

The theory of evidence proposed by Dempster was developed by Shafer in 1976 and the basic concepts of this theory have often been exposed [5, 6]. Detailed applications of the Dempster-Shafer theory can be found in [7]. Previous applications of the DS theory for our study [1] showed that 45% of the information sources were highly conflicting and generate misprediction results. Performances decrease when the conflict between the evidences is rising ($k < 0.6$). In our case, only 75% of the fields concerned by a high degree of conflict are correctly predicted. On the contrary, results become clearly better (91% of right prediction) when the conflict is low ($k < 0.2$).

Several methods that attempt to make the fusion operators more reliable in considering the different sources of conflict may be found in [8, 9, 10, 11]. No optimal techniques exist yet, even if an approximate adjustment of the fusion threshold can be successful for some applications. In order to deal with the conflict between the information sources, we have applied here a method based on the source weakness.

17.3.1 Basic belief assignment

The assignment of basic beliefs (membership function shape) on the selected indicators is assigned by experts and from the evidence image distribution (Fig. 17.1). They are adjusted and validated with past-observed data and expert's knowledge. Table 17.2 illustrates this stage in including the uncertainty through mass function affectation. For each evidences, denoted B for "bare soil", C for "covered soil", and $B \cup C$ for "Bare soil or covered soil", classes are defined in order to support one of the hypotheses B, C or $B \cup C$.

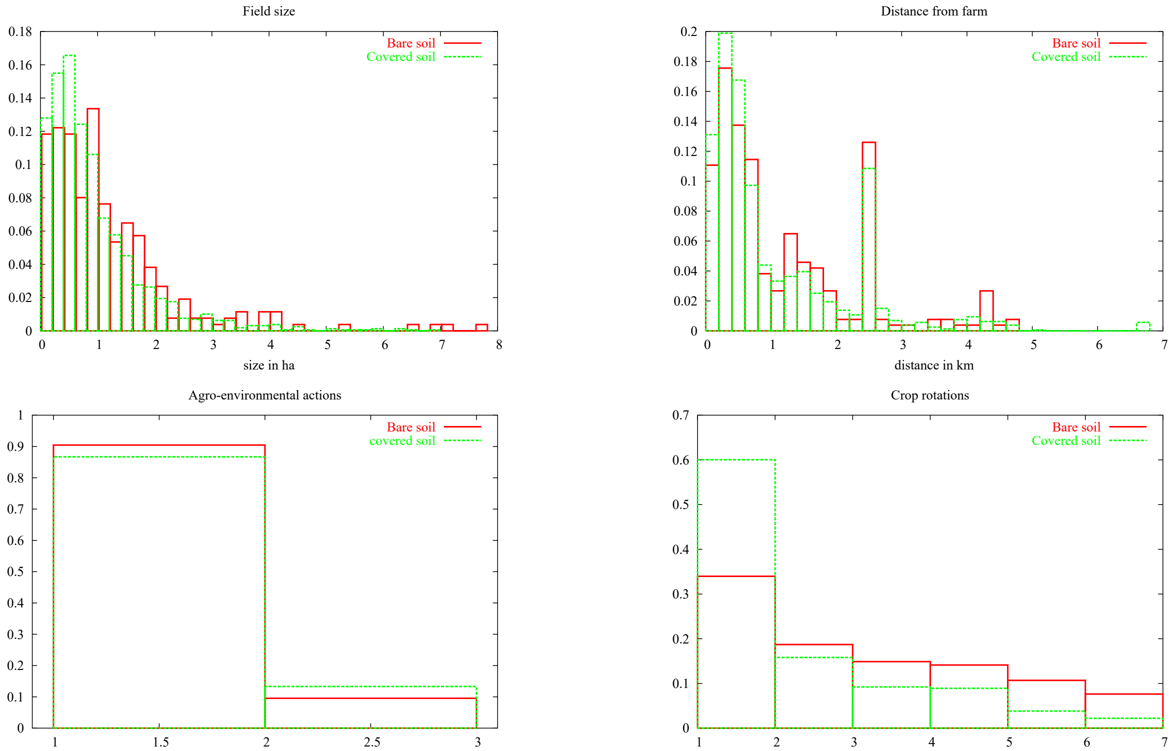


Figure 17.1: Evidence image distribution for each hypothesis.

	Classes	hyp. B	hyp. C	hyp. $B \cup C$
Distance from farm buildings	1: < 1 km	0.3	0.5	0.2
	2: > 1 km	0.6	0.2	0.2
Agro-environmental actions	1: without	0.6	0.3	0.1
	2: with	0.005	0.95	0.045
Field size	1: < 1.5 ha	0.2	0.5	0.3
	2: > 1.5 ha	0.65	0.2	0.15
Crop rotation (1996–2002)	1: (SC W)	0.005	0.95	0.045
	2: (BS 1W)	0.01	0.9	0.09
	3: (BS 2W)	0.25	0.7	0.05
	4: (BS 3W)	0.45	0.4	0.15
	5: (BS 4W)	0.65	0.3	0.05
	6: (BS 5W)	0.85	0.1	0.05

Table 17.2: Affectation of the belief masses for the DS theory.

17.3.2 Conflict managing with the source weakness

17.3.2.1 Principle

Sources weakness method (i.e. discounting technique presented in chapter 1 consists in taking in account the reliability of the evidences by using reliability factor α for each source as a value such as $0 \leqslant \alpha \leqslant 1$. This way, a source may be considered as totally reliable if $\alpha = 1$, or on the contrary completely unreliable if $\alpha = 0$. Damping rule is defined as follows:

$$\begin{cases} m'(A) & = \alpha\, m(A) \qquad \forall A \neq \Theta \\ m'(\Theta) & = (1-\alpha) + \alpha\, m(\Theta). \end{cases}$$

The weakness process is performed when the conflict is too high (relatively to a threshold, such as $k < 0.4$). Two rules have been investigated:

- α is set to a value so that the source does not interfere in the decision process. Then,

$$\begin{cases} m'(\theta_{\text{bare soil}}) & = 0.01 \\ m'(\theta_{\text{covered soil}}) & = 0.01 \\ m'(\theta_{\text{bare soil}} \cup \theta_{\text{covered soil}}) & = 0.98. \end{cases}$$

- α is set to a value linked to the conflict level k. So that the more the conflict, the more the weakness. We remind the conflict between two sources is defined as:

$$k = \sum_{A \cap B \neq \varnothing} m_1(A) m_2(B).$$

17.3.2.2 Results and partial conclusion

The results provided with this method are a little better than the simple application of the DS theory for the hypothesis “bare soil” since 84 fields are correctly predicted against 73 for the DS. But the analysis of the results showed that the conflict does not necessary take place in the mispredictions for the “bare soil” hypothesis. Also, Plausibility-Belief interval can not be helpful for the accuracy of the predictions. Then, an ambiguity between the sources must be taken into consideration in the process. Than is why, prediction process has been moved to the DSm theory in order to deal with paradoxical.

17.4 Land cover prediction with DSmT

The Dezert-Smarandache theory (DSmT) can be considered as a generalization of the Dempster-Shafer. In this new theory, the rule of combination takes into account both uncertain and paradoxical information, see chapter 1 of this book and [12]. Let be the simplest frame of discernment $\Theta = \{\theta_{\text{bare soil}}, \theta_{\text{covered soil}}\}$ involving only two elementary hypotheses with no more additional assumptions on $\theta_{\text{bare soil}}$ and $\theta_{\text{covered soil}}$.

DSm theory deals with new basic belief assignments $m(\cdot) \in [0,1]$ in accepting the possibility for paradoxical information such that:

$$m(\theta_{\text{bare soil}}) + m(\theta_{\text{covered soil}}) + m(\theta_{\text{bare soil}} \cup \theta_{\text{covered soil}}) + m(\theta_{\text{bare soil}} \cap \theta_{\text{covered soil}}) = 1.$$

Recently, a hybrid rule of combination issued of the DSm theory has been developed by the authors of the theory, see chapter 4 of this book. The fusion of paradoxical and uncertain evidences with the hybrid DSm rule of combination combines several masses of independent sources of information and takes into consideration the dynamics of data sets. Thus, hybrid DSm model can be considered as an intermediary model between the DS and the DSm theory. The capacity to deals with several hyper-power set makes the hybrid model an interesting alternative in various fusion problems.

17.4.1 Mass belief assignment

17.4.1.1 Fuzzy mass belief assignment

the mass belief assignment follows the same process as the DS theory. Nevertheless, a fuzzy mass belief assignment is here applied for two sources of evidence: "size of fields" and "distance from farm buildings" because of their specific characteristics (Fig. 17.1). For the variable "Size of fields" for example, the size evolves to 0.05 to 7.7 ha. Then, a continuous mass belief affectation appears pertinent for fusion process, by integrating paradoxical information when experts had introduced threshold instead. It is achieved by smoothing the actual bi-level assignment (Fig. 17.2).

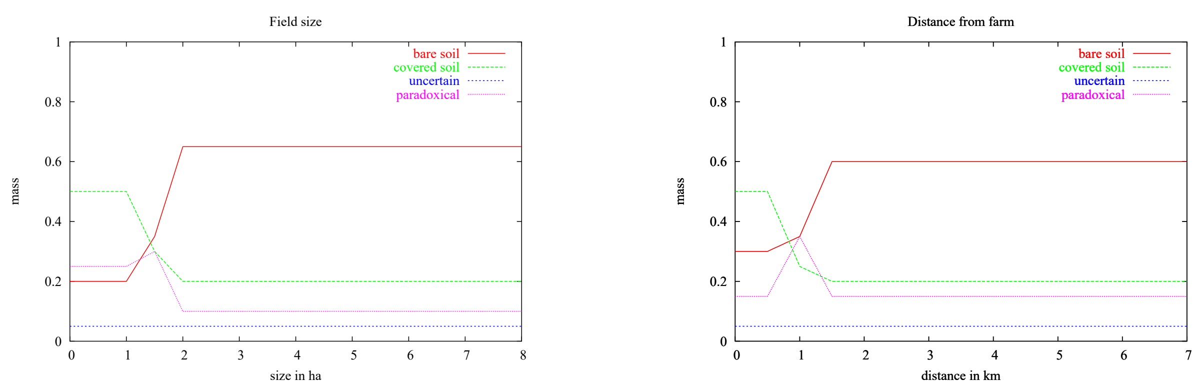


Figure 17.2: Fuzzy mass belief assignment for the evidences "Distance" and "Field size".

17.4.1.2 Contextual damping of source of evidence

since the conflit level between sources is not necessary involved in the misprediction for the "bare soil" hypothesis, a contextual damping strategy is applied depending on the decision that is about to be taken. Actually, we consider that when the decision is about to be taken to the "bare soil" hypothesis, distance to farm and field size are completely paradoxical when crop rotation belongs to class 1 or 2. Furthermore, when the decision is to be taken to the “covered soil” hypothesis, all the sources become paradoxical when crop rotation is greater than 3 (bare soil during two winters at least).

In order to make sources of evidence paradoxical, a partial damping is applied as follows:

$$\begin{cases} m'(\theta_{\text{bare soil}}) & = \alpha\, m(\theta_{\text{bare soil}}) \\ m'(\theta_{\text{covered soil}}) & = \beta\, m(\theta_{\text{covered soil}}) \\ m'(\theta_{\text{bare soil}} \cup \theta_{\text{covered soil}}) & = m(\theta_{\text{bare soil}} \cup \theta_{\text{covered soil}}) \\ m'(\theta_{\text{bare soil}} \cap \theta_{\text{covered soil}}) & = 1 - \alpha\, m(\theta_{\text{bare soil}}) - \beta\, m(\theta_{\text{covered soil}}) - m(\theta_{\text{bare soil}} \cup \theta_{\text{covered soil}}). \end{cases}$$

The couple (α, β) allows to remove the mass of an hypothesis to the benefit of the paradoxical. Here, $(\alpha, \beta) = (0.1, 1)$ is applied when the decision “bare soil” is about to be taken with crop rotation of 1 or 2 (bare soil during no more than one winter). Also, (α, β) is set to $(1, 0.1)$ when deciding a “covered soil” while crop rotation is greater than 3 (bare soil during 2 winters at least).

Here, this contextual partial damping allows the DSm rule to take into consideration a kind of contional mass assignment.

17.4.2 Results

The application of a contextual DSm rule of combination provides better results for the hypotheses “bare soil”. 121 fields (Table 17.4.2) are correctly predicted against 73 with the DS and 84 with the source weakness process. The “bare soil” hypothesis still generates a high level of mispredictions, which is not the case for the “covered soil” hypothesis. Several factors can explain the weak rate of right prediction for the hypothesis “Bare soils”. It is strongly linked to the high spatio-temporal variability of the land-use. Actually, an important number of fields covered with meadows during four or five years are ploughed in autumn and re-integrated in a cycle of crop successions. This kind of change is difficult to model since it can be due to unexpected individual human decisions, or exceptional and isolated weather-events. The spatial distribution of the results can be analyzed on the Fig. 17.3. The west part of the watershed corresponds to more intensive system farming than the east part. In the context of intensive system, the variability of land cover changes is higher than the others systems, it depends mostly on economics constraints that are difficult to model. On the contrary, the south part of the watershed is characterized by dairy milk production system. In this part of the watershed, the land cover evolution is better known and highly depends of the crop successions. Its integration into DSm theory is easier and the prediction process yields finest results.

	Land use for winter 2001/2002 (from remote sensing data)	Prediction (rate)
bare soils	266 fields	121 (0.46 %)
covered soils	1588 fields	1239 (0.78 %)
Total	1856 fields	1360 (0.73 %)

Table 17.3: Performance of hybrid DSm rule for land prediction

.

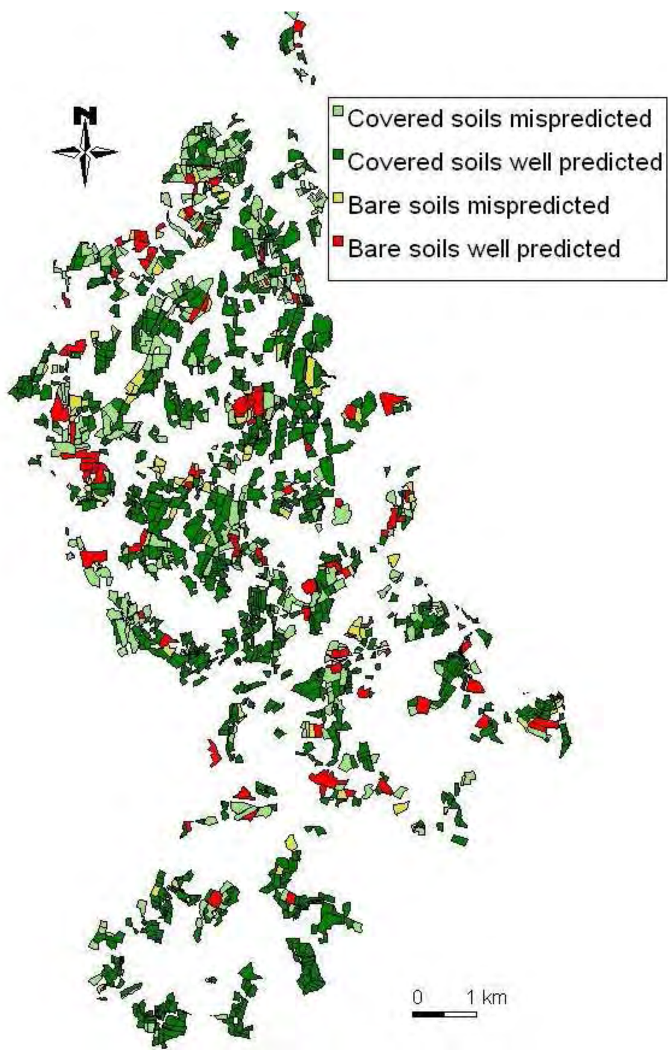


Figure 17.3: Prediction performance with the hybrid DSm rule on the Yar watershed (Brittany).

17.5 Conclusion

Two studies have been analyzed in this chapter for the prediction of land cover on a watershed subject to environmental problems. The land cover prediction with DS proved to have limitations with the increase of conflict between the sources of evidence that support land cover hypotheses. Several modifications may be used such as source weighting or the Hedging methods, but with no benefit in our case. To manage the conflict, the DSm has been applied with a little improvement of the accuracy of predictions. Actually conflict may not explain by itself all the bad decisions since the sources of evidence may induce all together a wrong decision. That is why, a contextual fusion rule appeared necessary for this environmental problem where information sources can be paradoxical or/and uncertain. This new fusion process required first the identification of the driving factors of land cover changes. Then, a mass belief assignment is built for the two hypotheses "covered soil" and "bare soil" through expert knowledge and a mutual information analysis that yield a hierarchization of the source of evidences. A fuzzy affectation is performed for two of the information sources and a "contextual" combination rule is applied to manage the uncertainty and the paradoxical characteristics of the information sources into the DSm decision process. The results for the "bare soil" hypothesis, which still generates too many mispredictions, are better than the prediction through DS decision rule (46% of correct "bare soil" predictions against 36% issued from the previous study). The hypothesis "covered soil" yields 78% of right prediction; this difference between the hypotheses can be explained with the weak rate of bare soil on the watershed and especially with the high variability of the land cover changes that characterized the intensive farm systems located on the north-west part of the watershed. Nevertheless, the fusion process appears to be robust and doesn't require specifics data as input. Thus, prediction system developed with the DSm theory can be apply on different watersheds in Brittany and provides a useful tool for assessing and planning land use. The knowledge of land use is one of the key for restoring water quality intensive agricultural regions.

Acknowledgements

This work is funded by the PNTS (*Programme National de Télédétection Spatiale*) and the *Conseil Régional de Bretagne*. We thank the *Chambre d'Agriculture des Côtes d'Armor* and the *Communauté d'agglomération de Lannion* for their help in identifying factors of change and validating results.

17.6 References

[1] Corgne S., Hubert-Moy L., Dezert J., Mercier G., *Land cover change prediction with a new theory of plausible and paradoxical reasoning*, Proc. of the Information Fusion Conference, Fusion 2003, Cairns, Australia, 8-11 July 2003, pp. 1141–1148.

[2] Hubert-Moy L., Quêté Y., Billot M., Corgne S., *Trajectories of land-cover-change in intensive agricultural regions : a short-time predictive approach with the dempster-shafer theory*, Proceedings of the First International Workshop on the Analysis of Multi-temporal Remote Sensing Images, World Scientific Co. Pte. Ltd, 2001, pp. 403–414.

[3] Shanon C., Weawer W., *The mathematical theory of communication.* University of Illinois Press, 1949.

[4] Baudry J., Al Jallad A., *La théorie de l'information : un outil d'analyse des liaisons entre les variables*, Rapport CERESA, 1992.

[5] Richards J., *Remote Sensing Digital Image Analysis: an introduction*, 2nd ed. Berlin: Springer-Verlag, 1993.

[6] Srinivasan A., Richards J., *Knowledge-based techniques for multi-source classification*, International Journal of Remote Sensing, vol. 11, no. 3, pp. 505–525, 1990.

[7] Mertikas P., Zervakis M., *Exemplifying the theory of evidence in remote sensing image classification, international journal of remote sensing classification*, International Journal of Remote Sensing, vol. 22, no. 6, pp. 1081–1095, 2001.

[8] Yager R.R., *Entropy and specificity in a mathematical theory of evidence*, Int. J. General Systems, vol. 9, pp. 249–260, 1983.

[9] Smets Ph., *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in Artificial Intelligence, M. Henrion, R. Shachter, L. Kanal, and J. Lemmer, Eds., 1990, pp. 29–40.

[10] Cholvy L., *Towards another logical interpretation of theory of evidence and a new combination rule*, Proc. of the IPMU, 1-5 July 2002.

[11] Dubois D., Prade H., *Representation and combination of uncertainty with belief functions and possibility measures*,Computational Intelligence, vol. 4, pp. 244–264, 1988.

[12] Dezert J., *Fondations pour une nouvelle théorie du raisonnement plausible et paradoxal*, ONERA Tech. Rep. RT 1/06769/DTIM, 2003.

Chapter 18

Power and Resource Aware Distributed Smart Fusion

Shubha Kadambe

HRL Laboratories, LLC

3011 Malibu Canyon Rd., Malibu, CA 90265, USA

Abstract: *Large distributed sensor networks (DSN) with disparate sensors, processors and wireless communication capabilities are being developed for a variety of commercial and military applications. Minimizing power consumption of the nodes is a critical issue to their good functioning during the mission or application, to reduce their size and weight, and their cost so that their deployment is economically viable. In this chapter, we describe a robust, flexible, and distributed smart fusion algorithm that provides high decision accuracy and minimizes power consumption through efficient use of network sensing, communication, and processing resources. Our approach, developed on information theory-based metrics, determines what network resources (sensors, platforms, processing, and communication) are necessary to accomplish mission tasks, then uses only those necessary resources. It minimizes the network power consumption and combines valuable information at features and decision level using DSmT. We demonstrate the proposed optimal, fully autonomous, smart distributed fusion algorithm for target detection and classification using a DSN. Our experimental results show that our approach significantly improves the detection and classification accuracy using the required high quality sensors and features, and valuable fused information.*

The part of this work was funded by DARPA/IXO under the contract # F30602-01-C-0192. The authors would like to thank in particular Dr. Sri Kumar for funding this work under the SensIT program.

18.1 Introduction

Spatially distributed network of inexpensive, small and smart nodes with multiple onboard sensors is an important class of emerging networked systems for various defense and commercial applications. Since this network of sensors has to operate efficiently in adverse environments using limited battery power and resources, it is important that appropriate sensors process information hierarchically and share information only if it is valuable in terms of improving the decision accuracy such that highly accurate decision is made progressively. One way to address this problem is to activate only those sensors that provide missing and relevant information, to assess the quality of information obtained from the activated sensors (this helps in determining the sensor quality), to assess the value of obtained information in terms of improving the decision (e.g., target detection/track) accuracy, to communicate only relevant, high quality and valuable information to the neighboring nodes and to fuse only valuable information that aid in progressive decisions. Information theoretic approaches provide measures for relevance, utility, missing information, value of information, etc. These measures help in achieving hierarchical extraction of relevant and high quality of information that enable in selection/actuation of relevant sensors and dynamically discard information from noisy or dead sensors and, progressive improvement of decision accuracy and confidence by utilizing only valuable information while fusing information obtained from neighboring nodes. In this chapter, we describe a minmax entropy based technique for missing information (feature) and information type (sensor) discovery, within class entropy based technique for sensor discrimination (i.e., quality assessment), mutual information for features quality assessment and, mutual information and other measures for assessing the value of information in terms of improvement in decision accuracy. In addition, we briefly describe how high quality, relevant and valuable information is fused using a new theory – DSmT which provides rules for combining two or more masses of independent sources of information that is dynamically changing in real time which is essential in the network of disparate sensors that is considered here.

To the best knowledge of this author there is no study on sensor discrimination using within class entropy metric is reported even though, there is one study on using mutual information for selecting a subset of features from a bigger set that is described in [2]. The technique described in this chapter uses within class entropy as a metric to assess the quality (good vs. bad) of a sensor. Unlike our technique, the technique in [2] is static in nature and cannot handle the case where the dimensionality of the feature set varies. In [15], the author shows that in general by fusing data from selective sensors the performance of a network of sensors can be improved. However, in this study, no specific novel metrics for the feature discovery and feature/sensor discrimination were developed unlike in this chapter. In [10], techniques to represent Kalman filter state estimates in the form of information – Fisher and Shannon entropy are provided. In such a representation it is straightforward to separate out what is new information from what is either prior knowledge or common information. This separation procedure is used in decentralized data fusion algorithms that are described in [10]. However, to the best knowledge of this author no study has been reported on using minmax entropy principle for the feature and information type discovery. Furthermore, to our knowledge the proposed value of information based fusion is not studied by others and is another significant contribution of this chapter. In addition, the significance of this study is the application of feature discovery and sensor discrimination in awakening the required sensor and in the formation of a cluster of distributed sensors to reduce the power consumption, to improve the decision accuracy and to reduce the communication bandwidth requirements. This chapter is a comprehensive of our studies reported in [6, 7, 8] with the addition of application of DSmT for fusion at both feature and decision levels.

In the next section, proposed techniques are described. The simulation description and experimental results are provided in section 18.3. Conclusions and future research directions are provided in section 18.4.

18.2 Description of proposed research

18.2.1 Discovery of missing information

In the case of applications of a distributed network of disparate sensors such as (a) target detection, identification and tracking, (b) classification, (c) coalition formation, etc., the missing information could correspond to feature discovery. This helps in only probing (awakening) the sensor node that can provide the missing information and thus save power and processing by not arbitrarily activating nodes and by letting the unused sensor be in the sleep mode. We apply the minmax entropy principle described in [9] for the feature discovery. The details of estimation of missing information in other words feature discovery and information type using the minmax entropy principle are as follows.

18.2.1.1 Minmax entropy principle

Let N given values corresponds to n different information types. Let z_{ij} be the j-th member of i-th information type (where the information type is defined as a sensor type that gives similar information measures) so that

$$i = 1, 2, \ldots, n \quad j = 1, 2, \ldots, m_i \qquad \sum_{i=1}^{n} m_i = N \tag{18.1}$$

Then the entropy for this type of classes of information is:

$$H = -\sum_{i=1}^{n}\sum_{j=1}^{m_i} \frac{z_{ij}}{T} \ln \frac{z_{ij}}{T} \qquad \text{where} \quad T = \sum_{i=1}^{n}\sum_{j=1}^{m_i} z_{ij} \tag{18.2}$$

Let $T_i = \sum_{j=1}^{m_i} z_{ij}$. Using this, H can be written as:

$$H = \sum_{i=1}^{n} \frac{T_i}{T} H_i - \sum_{i=1}^{n} \frac{T_i}{T} \ln \frac{T_i}{T} = H_W + H_B \tag{18.3}$$

where $H_i = -\sum_{j=1}^{m_i} \frac{z_{ij}}{T} \ln \frac{z_{ij}}{T}$ is the entropy of values that belong to information i.

In the equation above, H_W and H_B are entropy of *within classes* (information types) and *between classes*, respectively. We would like types of information to be as distinguishable as possible and we would like the information within each type to be as homogenous as possible. The entropy is high if the values belonging to a type (class) represent similar information and is low if they represent dissimilar information. Therefore, we would like H_B to be as small as possible and H_W as large as possible. This is the principle of minmax entropy.

18.2.1.2 Application of minimax entropy principle for feature discovery

Let z be the missing value (feature). Let T be the total of all known values such that the total of all values is $T + z$. Let T_1 be the total of values that belong to information type to which z may belong. $T_1 + z$ then is the total of that particular type of information. This leads to:

$$\begin{cases} H = -\sum' \frac{z_{ij}}{T+z} \ln \frac{z_{ij}}{T+z} - \frac{z}{T+z} \ln \frac{z}{T+z} \\ H_B = -\sum'' \frac{T_i}{T+z} \ln \frac{T_i}{T+z} - \frac{T_1+z}{T+z} \ln \frac{T_1+z}{T+z} \end{cases} \tag{18.4}$$

Here $\sum'$ denotes the summation over all values of i, j except that correspond to the missing information and $\sum''$ denotes over all values of i except for the type to which the missing information belongs, respectively.

We can then estimate z by minimizing H_B/H_W or $H_B/(H - H_B)$ or H_B/H, or by maximizing $(H - H_B)/H_B$ or H/H_B. The estimates of z provide the missing information values (features) and information (sensor) type. From the above discussion, we can see that we will be able to discover features as well as type of sensor from which these features can be obtained. This has the advantage of probing the appropriate sensor in a DSN. The transfer of information and probing can be achieved in such a network by using network routing techniques. Before trying to use the newly acquired feature set from the estimated information type i.e., sensor, it is advisable to check the quality of the sensor to make sure that the sensor from which we are seeking the information is not noisy (not functioning properly) or "dead" to reduce the cost of processing. In a DSN this has an added advantage of reducing the communication cost. We measure (see next section) the quality (i.e. discriminate a good sensor vs. bad sensor) by using an information theoretic measure - the within class entropy.

18.2.2 Measure of consistency

We measure relevance by measuring consistency. For this we have developed a metric based on within class entropy that is described in this section. Let there are N events (values) that can be classified in to m classes and let an event x_{ij} be the j-th member of i-th where $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n_i$ and $\sum_{i=1}^{m} n_i = N$. The entropy for this classification is:

$$\begin{aligned}
H &= \sum_{i=1}^{m} \sum_{j=1}^{n_i} p(i)p(x_{ij}) \log(\frac{1}{p(i)p(x_{ij})}) \\
&= -\sum_{i=1}^{m} \sum_{j=1}^{n_i} p(i)p(x_{ij} \log(p(i)p(x_{ij}) \\
&= -\sum_{i=1}^{m} p(i) \sum_{j=1}^{n_i} p(x_{ij}) \log(p(x_{ij}) - \sum_{i=1}^{m} p(i) \log(p(i)) \sum_{j=1}^{n_i} p(x_{ij}) \\
&= \sum_{i=1}^{m} p(i)H_i - \sum_{i=1}^{m} p(i) \log(p(i)) \\
&= H_W + H_B
\end{aligned}$$

The penultimate equality comes from the definition of $H_i = -\sum_{i=1}^{m} p(i) \sum_{j=1}^{n_i} p(x_{ij}) \log(p(x_{ij})$ representing the entropy of a class i and the total probability theorem, i.e. $\sum_{j=1}^{n_i} p(x_{ij}) = 1$. H_W is called the entropy *within* classes and H_B is called the entropy *between* classes.

The entropy H_W is high if the values or events belonging to a class represent similar information and is low if they represent dissimilar information. This means H_W can be used as a measure to define consistency. That is, if two or more sensor measurements are similar then their Hw is greater than if they are dissimilar. Therefore, this measure can be used in sensor discrimination. Note that even though the definitions of within class and between class entropy here are slightly different from section 18.2.1, they are similar in concept. Note also that the minmax entropy measure that uses both within and between class entropies was used earlier in the estimation of missing information; but here, within class entropy is defined as a consistency measure that can be used in sensor discrimination or selection. These two metrics have different physical interpretations and are used for different purposes.

18.2.3 Feature discrimination

After making sure about the quality of sensor (the information type) from which missing information can be obtained, it is necessary to make sure that the observations (features) from that sensor does help in gaining information as far as the required decision is concerned. This step doubly makes sure that the estimated missing information is indeed needed. For this, we have developed metrics based on conditional entropy and mutual information which are described in the following two subsections.

18.2.3.1 Conditional entropy and mutual information

Entropy is a measure of uncertainty. Let $H(x)$ be the entropy of previously observed x events. Let y be a new event. We can measure the uncertainty of x after including y by using the conditional entropy which is defined as:

$$H(x|y) = H(x,y) - H(y) \tag{18.5}$$

with the property $0 \leq H(x|y) \leq H(x)$. The conditional entropy $H(x|y)$ represents the amount of uncertainty remaining about x after y has been observed. If the uncertainty is reduced then there is information gained by observing y. Therefore, we can measure the importance of observing estimated y by using conditional entropy. Another measure that is related to conditional entropy that one can use is the mutual information $I(x,y)$ which is a measure of uncertainty that is resolved by observing y and is defined as:

$$I(x,y) = H(x) - H(x|y) \tag{18.6}$$

To explain how this measure can be used to measure the importance of estimated missing information (e.g., features) which is referred to as feature discrimination, an example is provided below.

18.2.3.2 Example of feature discrimination based on entropy metrics

Let $A = \{a_k\}$, $k = 1, 2, \ldots$ be the set of features from sensor 1 and let $B = \{b_l\}$, $l = 1, 2, \ldots$ be the set of features from sensor 2. Let $p(a_k)$ be the probability of feature a_k and $p(b_l)$ the probability of feature b_l. Let $H(A)$, $H(B)$ and $H(A|B)$ be the entropy corresponding to sensor 1, sensor 2 and sensor 1 given sensor 2, respectively, and they are defined as [9]:

$$\begin{cases} H(A) = \sum\limits_k p(a_k)\log(\dfrac{1}{p(a_k)}) \\ H(A|B) = H(A,B) - H(B) = \sum\limits_l p(b_l)H(A|b_l) = \sum\limits_l p(b_l)\sum\limits_k p(a_k|b_l)\log(\dfrac{1}{p(a_k|b_l)}) \end{cases} \tag{18.7}$$

Here, the entropy $H(A)$ corresponds to the prior uncertainty and the conditional entropy $H(A|B)$ corresponds to the amount of uncertainty remaining after observing features from sensor 2. The mutual information $I(A,B) = H(A)–H(A|B)$ corresponds to uncertainty that is resolved by observing B in other words features from sensor 2. From the definition of mutual information, it can be seen that the uncertainty that is resolved basically depends on the conditional entropy. Let us consider two types of sensors at node 2. Let the set of features of these two sensors be B_1 and B_2, respectively and let the set of features estimated by the minmax entropy principle described in the previous section be B_1. If $H(A|B_1) < H(A|B_2)$ then $I(A,B_1) > I(A,B_2)$. This implies that the uncertainty is better resolved by observing B_1 as compared to B_2. This further implies that indeed the estimated B_1 corresponds to features that help in gaining information that aid in the decision process of sensor 1 and B_2 does not and hence, should not be considered.

Note that even though in the above example only two sensor nodes are considered for simplicity, this measure or metric can be used in a network of more than two sensors. In such a case, A would be a set of features that a node already has from other sensors in a cluster that it is a member of and B would be a new feature set that it receives from a different sensor type that it has not already received from and it may be a member or not a member of that particular cluster. If the mutual information increases by including the set of features B then we make a decision of including that sensor as part of this particular cluster if it is not a member. In case it is a member and the mutual information does not increase then it would be discarded from that particular cluster.

18.2.4 Measures of value of information

This section describes the measures of value of information that we have developed to determine when to fuse information from disparate sources. The value is in terms of improving the decision accuracy. Even though the mathematics of the metrics described below are not novel, the usage of metrics in the context of verifying value of information with respect to improving the decision accuracy (e.g., classification accuracy, detection accuracy) is new.

18.2.4.1 Mutual information

Mutual information defined in section 18.2.3.1 can also be used as a measure of value.

18.2.4.2 Euclidean Distance

Unlike mutual information, Euclidean distance does not evaluate the amount of information available from a second source. It does, however, measure the similarity between two feature sets in Euclidean space. This value can then be used to determine when to fuse two sources of information, whether they are from different types of sensors on the same node or from same type of sensors different nodes. A simple measure, Euclidean distance is defined as:

$$d = \sqrt{\sum_i (a_i - b_i)^2} \tag{18.8}$$

where a_i, b_i and i are defined in Section 18.2.3.1.

18.2.4.3 Correlation

Correlation is also a well known measure of similarity. We use the standard measure of correlation as defined by:

$$\rho = \frac{E[(a - \mu_a)(b - \mu_b)]}{E[a - \mu_a]E[b - \mu_b]} \tag{18.9}$$

where μ_a and μ_b are the means of feature sets a and b, respectively. Note that correlation is very closely related to mutual information, $I(x,y)$ because (18.6) can be rewritten as:

$$I(x,y) = \sum_k p(a_k,b_k)\log(\frac{p(a_k,b_k)}{p(a_k)p(b_k)}) \tag{18.10}$$

18.2.4.4 Kullback-Liebler distance

Finally, the Kullback-Liebler (KL) distance is derived from entropy, and again is a measure of the separation of two feature sets. It is defined as:

$$D = \sum_k p(a_k)\log(\frac{p(a_k)}{p(b_k)}) + \sum_k p(b_k)\log(\frac{p(b_k)}{p(a_k)}) \tag{18.11}$$

18.2.5 Fusion using DSmT

Since in a network of disparate sensor nodes as is considered here, the sources of information are independent and changing dynamically based on which sensor and features are selected, for the smart distributed fusion we use the new theory of plausible and paradoxical reasoning – DSmT developed in [5]. This theory provides a hybrid DSm rule which combines or fuses two or more masses of independent sources of information and takes care of restraints i.e., of sets which might become empty at certain time or new sets that might arise at some other time. In a network of sensor nodes these situations arise (sometimes we discard the feature set or decision from the other nodes and sometimes we use features from different type of sensors based on how the scene is changing dynamically) and hence, the application of hybrid DSm rule for fusion is very appropriate. In addition, since fusion is not done at a centralized location but done locally dynamically based on the information received from the neighboring nodes, we propose to extend the decentralized dynamical fusion by combining dynamical fusion using the hybrid DSm rule for the chosen hybrid model $\mathcal{M}$. Specifically, at the feature level fusion at each sensor node the frame under consideration at time t_l will be

$$\Theta(t_l) \triangleq \{\theta_1 = \text{acoustic sensor}, \theta_2 = \text{seismic sensor}, \theta_3 = \text{IR sensor location}\}$$

and at decision level fusion, $\Theta(t_l) = \{\theta_1 = \text{vehicle present}, \theta_2 = \text{vehicle not present}\}$ in the case of a detection application and,

$$\Theta(t_l) \triangleq \{\theta_1 = \text{AAV}, \theta_2 = \text{DW}, \theta_3 = \text{HMMWV}\}$$

where AAV, DW, and HMMWV represent the vehicle types that are being classified) for the decision level fusion in the case of a classification application.

Both detection and classification applications are described in section 18.3.2. We derive basic belief assignments based on the observations (a) from the sensor type for feature level fusion, (b) from the features extracted from the sensors' signals for fusion at the decision level in the case of classification and detection applications. For example, $m_a(\theta_1) = 0$ and $m_a(\theta_2) = 0$, if the feature from the acoustic sensor (a) – energy is well above the threshold level in the case of the detection application. $\Theta(t_l)$ changes as the observation is different based on the above described sensor and feature selection and results in $\Theta(t_{l+1})$. If we discard observations from a sensor (based on the feature discrimination algorithm explained above) then Θ diminishes and we apply the hybrid DSm rule to transfer the masses of empty sets to non-empty sets. If we include observations from a new sensor then we use the classical DSm fusion rule to generate basic belief assignments $m_{t_{l+1}}(.)$. For the decentralized decision level fusion at the current node, consider the $\Theta_p(t_l)$ obtained from the previous node and the $\Theta(t_l)$ of the current node and apply the hybrid DSm rule by taking the integrity constraints in to consideration. These constraints are generated differently for the fusion between the sensors and for the fusion from node to node. The pseudo-codes which generate these constraints are given in section 18.3.2. For example, in the case of node to node fusion for classification application that is described in section 18.3.2.2.1 `fuse_4class=1` will indicate to put the constraint $\theta_1 \cap \theta_2 \overset{\mathcal{M}}{\equiv} \emptyset$, $\theta_1 \cap \theta_3 \overset{\mathcal{M}}{\equiv} \emptyset$, $\theta_1 \cap \theta_2 \cap \theta_3 \overset{\mathcal{M}}{\equiv} \emptyset$ at the current node if the classification at the previous node corresponds to $\theta_1 =$ AAV since if the vehicle at the previous node is AAV, the vehicle at the current node which is very close to the previous node has to be AAV.

18.3 Experimental details and results

Above described algorithms have been applied for the feature discovery, sensor and feature evaluation (discrimination), cluster formation and distributed smart fusion in a network of both simulated radar sensors and a network of real disparate sensors and sensor nodes that are spatially distributed. First, in section 18.3.1, the results obtained using a simple simulated network of radar sensors is provided for the purposes of proving the concepts. In section 18.3.2 however, experimental results obtained by using a real DSN of disparate sensors is provided.

18.3.1 Simulated network of radar sensors

This network of sensors is used for tracking multiple targets. Each sensor node has a local and global Kalman filter based target trackers. These target trackers estimate the target states - position and velocity in Cartesian co-ordinate system. The local tracker uses the local radar sensor measurements to estimate the state estimates while the global tracker fuses target states obtained from other sensors if it is consistent and improves the accuracy of the target tracks.

For the purposes of testing the proposed algorithms of this chapter, a network of three radar sensors and a single moving target with constant velocity are considered. Two sensors are considered as good and one as bad. A sensor is defined as bad if its measurements were corrupted with high noise (for example SNR = -6 dB) or is biased. In the first set of examples the SNR of a good sensor is set to be 10 dB.

In the case of simulation of a biased sensor, the bias was introduced as the addition of a random number to the true position of a target. The bias was introduced this way because the biases in azimuth and range associated with a radar sensor translate into measured target position that is different from the true target position. In addition, in our simulations, we assume that the sensors are measuring the target's position in the Cartesian co-ordinate system instead of the polar co-ordinate system. The amount of bias was varied by multiplying the random number by a constant k i.e.,measured position = (true position + k·randn) + measurement noise.

First, the minmax entropy principle was applied to find the missing information, the appropriate sensor was probed to obtain that information, then the consistency measure – within class entropy was applied to check whether the new sensor type and the information obtained from that particular sensor is consistent with the other sensors.

In the following two figures, within class entropy is plotted for features discovered from two unbiased sensors and, one biased and one unbiased sensor. The measurement noise level was kept the same for all three sensors. However, the bias k was set to 1.0 in Figure 18.1 and was set to 2 in Figure 18.2. The within class entropy was computed for different iterations using the definition provided in the previous section. The probability values needed in this computation were estimated using the histogram approach which is elaborated below. From these two figures, it can be seen that the within class entropy of two unbiased sensors is greater than the within class entropy of one biased and one unbiased sensors. This indicates that the within class entropy can be used as a measure to discriminate between sensors or to assess the quality of sensors (to select sensors).

Next, the conditional entropy and mutual information measures described in the previous section are used to make sure the estimated features obtained from the selected sensors indeed aid in the decision process.

For this, the target states that were estimated from the measurements of a simulated radar at each sensor node using the local Kalman filter algorithm is used as feature sets. The estimated target states at each sensor node were transmitted to other nodes. For this simulation, only estimated position was considered for simplicity.

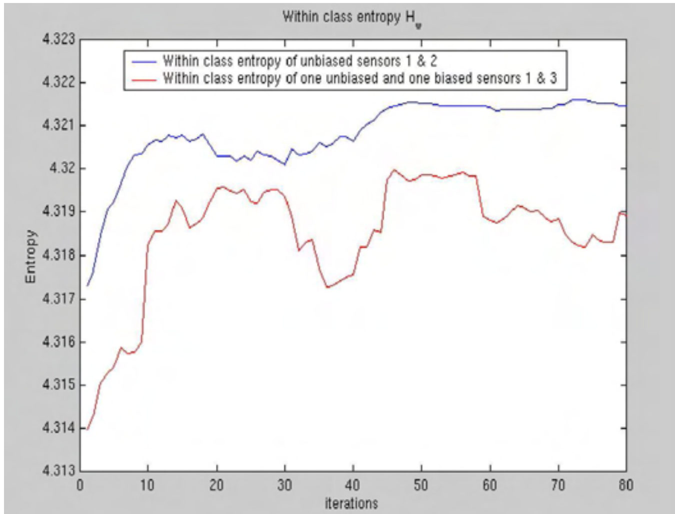


Figure 18.1: The plot of within class entropy of sensors 1 & 2 (unbiased sensors) and, 1 (unbiased) and 3 (biased). Bias constant $k = 1$

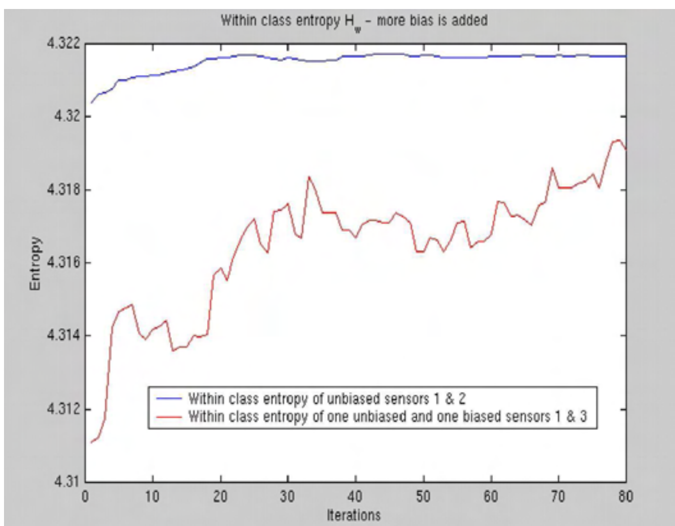


Figure 18.2: The plot of within class entropy of sensors 1 & 2 (unbiased sensors) and, 1 (unbiased) and 3 (biased). Bias constant $k = 2$

We considered the estimated state vector as the feature set here. Since the goal of this simulation is proof of concept, the feature discrimination algorithm was implemented at sensor node 1 with the assumption it is a good sensor. Let the state estimate outputs of this node be A_g. Let the state estimate outputs of a second sensor correspond to Bg and a third sensor correspond to B_b.

For the computation of entropy, the probability values are needed as seen from the equation above. To obtain these values, ideally, one would need probability distribution functions (pdfs). However, in practice it is hard to obtain closed form pdfs. In the absence of knowledge of actual pdfs it is a general practice to estimate them by using histograms [11]. Researchers in signal and image processing use this technique most commonly [13]. Another practical solution to estimate the probability and conditional probabilities is by using the counting or frequency approach [12]. However, it is well known that the estimates of probabilities and conditional probabilities are more accurate if they are estimated by using the pdfs that are approximated from the histograms. Therefore, we use the histogram approach here. In order to obtain the histograms, initially, we need some data (features) to know how it is distributed. For this purpose, it was assumed that initially N state estimate vectors were accumulated at each sensor node and this accumulated vector was transmitted to other nodes. Note also that the accuracy of probability estimates using the histogram approach depends on the amount of accumulated (training) data. Also for non-stationary features, it depends on how often the histograms are updated. In practice, since the training data is limited we have set N to 10 in this simulation. To take care of the non-stationarity of the features, initially, we wait till N estimates are obtained at each node. From then on we update the histograms every time instant using the new state estimate and previous nine state estimates. At each time instant we discard the oldest feature (oldest state estimate).

To get the probability of occurrence of each feature vector, first the histogram was computed. For this, bin size N_{bin} of 5 was used. The center point of each bin was chosen based on the minimum and maximum feature values. In this simulation the bin centers were set as:

$$\min(\text{feature values}) + (0 : N_{\text{bin}} - 1) \cdot \frac{\max(\text{feature values}) - \min(\text{feature values})}{N_{\text{bin}}} \tag{18.12}$$

Since the histogram provides the number of elements in a given bin, it is possible to compute the probabilities from the histogram. In particular it is computed as:

$$\frac{\text{Number of elements in a particular bin}}{Total number of elements}$$

Hence, from these histograms, probabilities were computed. Similarly, conditional probabilities of $p(A_g|B_g)$ and $p(A_g|B_b)$ were computed from the conditional histograms and these conditional probabilities are plotted in Figures 18.3 and 18.4, respectively.

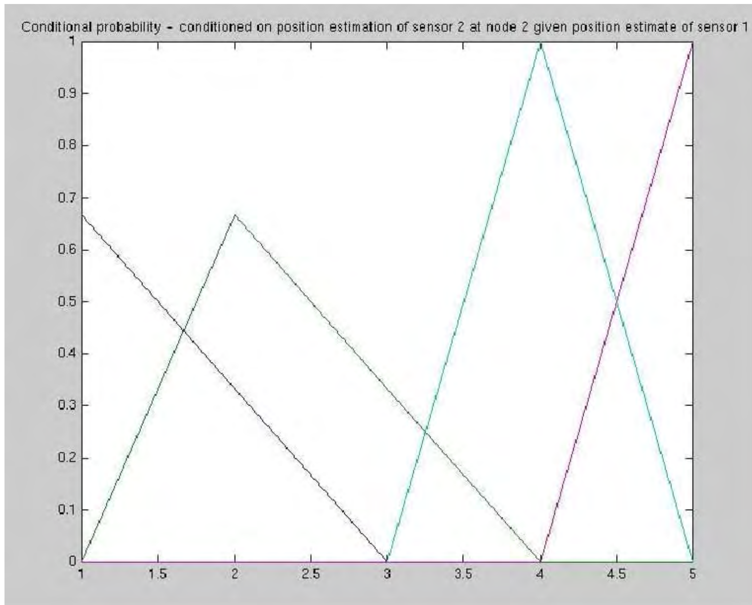


Figure 18.3: Conditional probability of position estimates of sensor 2 at node 2 given position estimates of sensor 1 at node 1

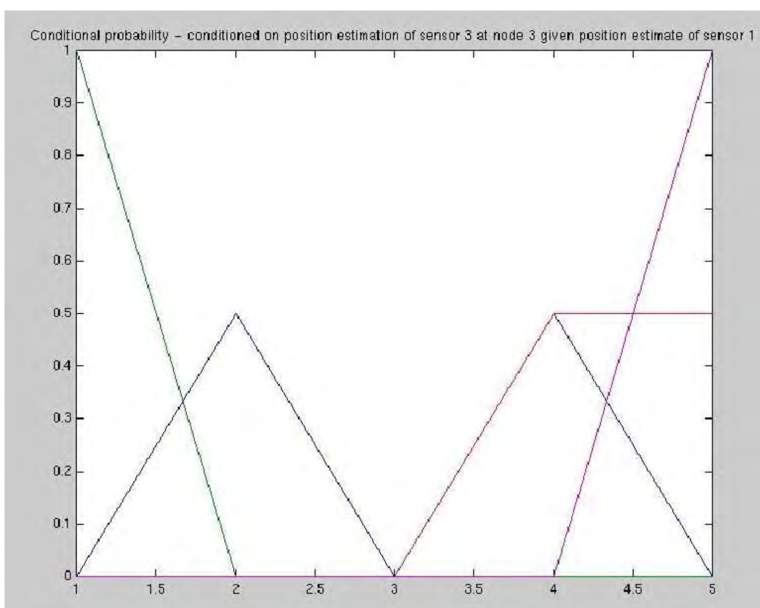


Figure 18.4: Conditional probability of position estimates of sensor 3 at node 3 given position estimates of sensor 1 at node 1

Each colored line in these two plots represents one conditional probability distribution function. Note that both A and B are vectors and there would be one pdf for each member of set A. Since we have chosen bin size as 5 there would be 5 members in set A and hence, there are 5 subplots in Figures 18.3 and 18.4.

Using these probabilities, conditional entropies $H(A_g|B_g)$ and $H(A_g|B_b)$, and mutual information $I(A_g, B_g)$ and $I(A_g, B_b)$ were computed using the equations mentioned above for one set of features from sensor at node 2 and node 3. After this kind of initial computation of probabilities, conditional entropy and mutual information, whenever a sensor estimates a new feature it is replaced by the oldest feature in the feature set and transmitted to other sensors. Subsequently, histograms, probabilities, conditional entropy and mutual information were computed using this updated feature set. This would take care of the non-stationarties of features. Thus each new feature can be verified to make sure it is relevant in terms of aiding in the decision process (e.g., track accuracy) and it is obtained from a good sensor. Therefore, this technique is dynamic in nature.

18.3.1.1 Versatility of the algorithm

To verify the versatility of this algorithm we considered a different feature sets namely, the sensor measurements itself instead of the position estimates and the first difference in position estimates. We performed similar simulation that is described above using these two types of feature sets and the associated histograms for the probability, entropy and mutual information computations. In these two cases also we always obtained $I(A_g, B_g) > I(A_g, B_b)$ for all the 100 runs of Monte Carlo simulations.

18.3.1.2 Sensitivity of the algorithm for sensor discrimination

Next, noise level at sensor 2 and 3 were varied to determine the sensitivity of the sensor discrimination algorithm. The SNR at sensor 1 was fixed at 10 dB. The algorithm was able to discriminate between good and bad sensor 100 % of the time when the noise level at sensor 2 is 8 dB and at sensor 3 is 3 dB. The algorithm was able to discriminate about 80 % of the time if the noise level at sensor 3 is 5 dB when the noise level at sensor 2 is fixed at 8 dB. If the noise level at both sensor 1 and 2 is 10 dB then the algorithm was able to discriminate 100 % of the time when the noise level at sensor 3 is 5 dB. However, when the noise level at sensor 3 was changed to 7 dB, the percentage of correct discrimination was dropped to 82 %. *Therefore, if the minimum difference between the noise level at sensor 2 and 3 is 5 dB then the discrimination accuracy is 100 %.* If the noise level at both sensor 2 and 3 is close (a difference of 1 dB) then the algorithm cannot discriminate as expected.

18.3.1.3 Mutual information versus track accuracy

To check indeed when mutual information metric is used to evaluate the information gain by observing the estimated missing features (information) and it aids in the improvement of the accuracy of decision (e.g., track accuracy), the following experiment was conducted.

As before, mutual information $I(A_g, B_g)$ and $I(A_g, B_b)$ was computed using measurements as feature set. If $I(Ag, Bg) > I(Ag, Bb)$ then the state estimates from the good sensor was fused with sensor 1 using the global Kalman filter algorithm and the DSm combination rule that is described in section 18.2.5. The position estimation error was computed by comparing the fused state estimate with the true position. To compare the track accuracies, the state estimates from the bad sensor and good sensor were also fused. The position estimation error was then computed the same way as explained above.

In Fgure 18.5, the position estimation error using the fused state estimates of sensor 1 & a good sensor (blue plot) and sensor 1 & a bad sensor (red plot) are plotted. From this figure, it can be seen that the track accuracy after fusing state estimates from good sensors (1 & 2) is much better than fusing state estimates from a good sensor and a bad sensor (1 & 3). This implies that better mutual information correlates to better track accuracy.

In Figure 18.6, the position error is plotted for the case when the noise level at sensor 2 and 3 differs by 5 dB. In this case also it can be seen that the track accuracy is better when the state estimates from good sensors is fused as compared to the track accuracy of fused state estimates of a good sensor and a bad sensor.

We then form a cluster of sensors that are consistent and apply the mutual information metric. We have shown above that by fusing information from sensors when the mutual information increases, the decision accuracy improves. We transmit the fused decision (which requires much lower bandwidth compared to the transmission of decision of each sensor to every other in the network) to other clusters of sensors and thus reduce the communication bandwidth requirement.

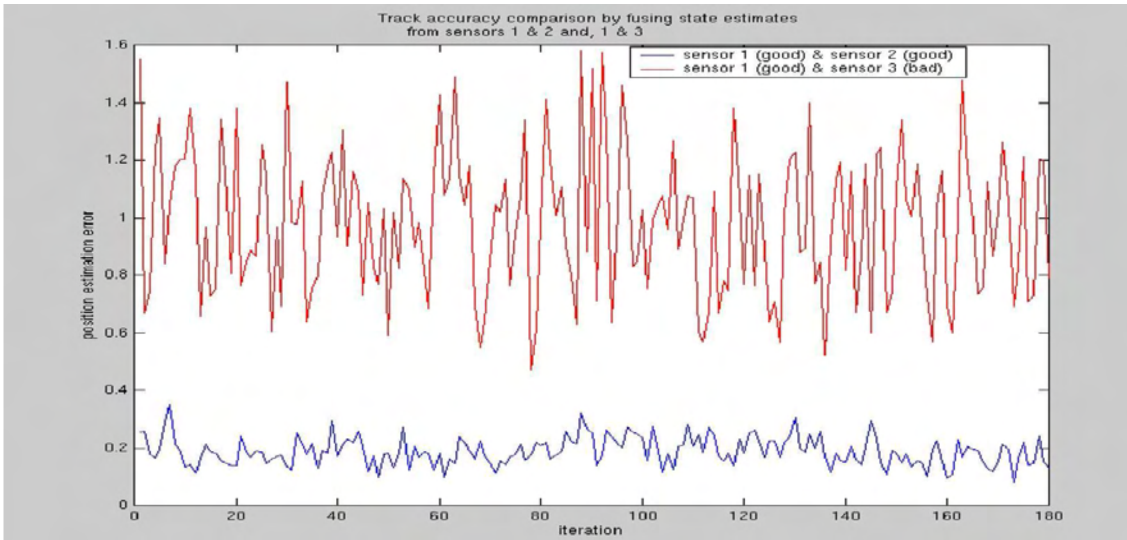


Figure 18.5: Track accuracy comparison - Noise level at sensor 1 and 2 = 10dB and at sensor 3=0dB

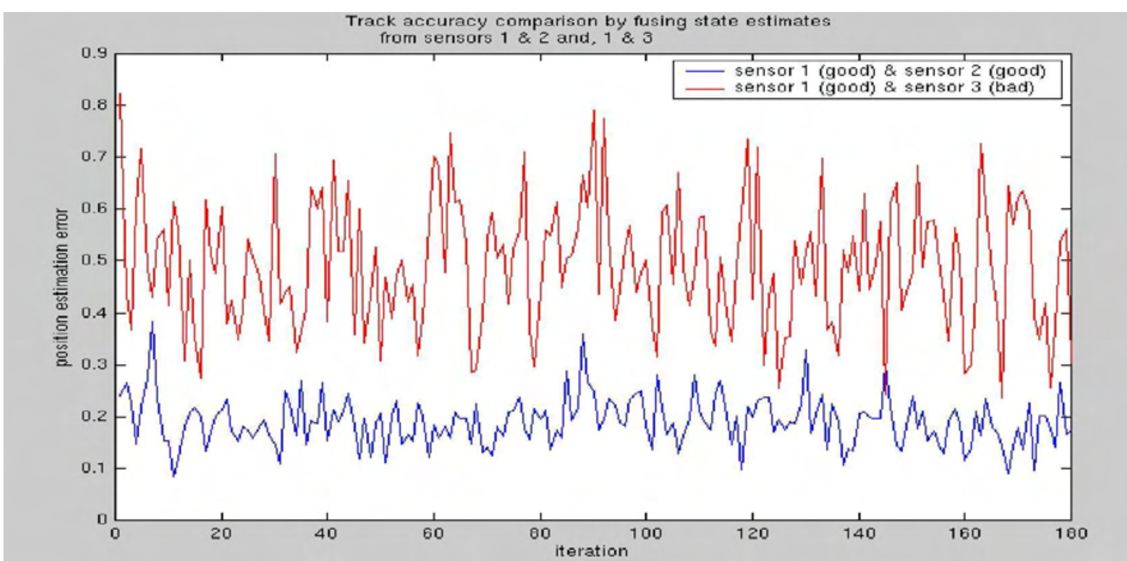


Figure 18.6: Track accuracy comparison - Noise level at sensor 1 and 2 = 10dB and at sensor 3=0dB

18.3.2 A real network of spatially DSN with disparate sensors

The proposed algorithms described in section 18.2 was implemented on sensor nodes that consists of multiple sensors, a communication radio and a Sharc processor. These sensor nodes were distributed in a rough terrain such as a desert. This network was used in detecting, tracking and classifying targets. Even though we verified the algorithms that estimate the missing information, sensor selection, sensor and feature assessment in this network of sensor node, in the following subsections, we are concentrating on the value of information based smart fusion that is described in sections 18.2.4 and 18.2.5 since the experimental results for the other algorithms are provided in the last section. We provide the experimental details and the results. We begin with the review of detection and classification algorithms that were used in this context.

18.3.2.1 Review of algorithms used to check the value of information based smart fusion

The metrics described in section 2.4 are used to measure the value of information obtained from other sources such as multiple sensors on a single node and from the neighboring nodes in the context of target detection and classification. For target detection, energy based detector was used and for classification, maximum likelihood based classifier was used. As mentioned before the value of information is in terms of improvement in the decision accuracy which corresponds to classification accuracy for a classifier and detection accuracy or probability of detection for a detector. Note that in this study, we did not develop a classifier or a detector; however, used those developed by others since the goal of this part of the study is to develop measures of value of information and verify them in terms of improvement in decision accuracy when they were used to make a decision of whether to fuse information obtained from the other source or not. In the following two sections we review the classifier and the detector that were used in this study.

18.3.2.1.1 Maximum likelihood based classifier The classifier we used for the verification of measures of value of information in terms of improving the decision accuracy is a maximum likelihood based classifier developed by the University of Wisconsin [16] as part of DARPA's sensor information technology (SensIT) program. For a given training features and target labels a Gaussian mixture model is determined during the training phase of the classifier. During testing the distance between the test feature vector and ith class Gaussian mixture is computed. This corresponds to negative log likelihood. Then a priori probability is used to obtain the maximum a posterior classification. The features' set that is used here consists of twenty features from the power spectral density. This is computed using 1024 FFT. The feature set is collected by summing up the values over equal length segments of the power spectrum. For the acoustic and seismic sensors the maximum frequency used was 1000 and 200 Hz, respectively.

18.3.2.1.2 Energy based detector An energy based detector is also used for the verification of improvement in decision accuracy when the value of information based fusion architecture is used. This detector is developed by BAE, Austin [3]; also as part of the SensIT program. A brief description of this detector is provided below.

For every block of a given signal the energy of the down sampled version of the power spectral density is computed. For the computation of the power spectral density, 1024 point FFT is used. This energy is compared with a threshold value. Whenever the energy is above the threshold it was declared that the target was detected. The threshold value is adaptively changed based on the background energy.

18.3.2.2 Experimental details

The above described classifier and detector, and measures of value of information and the fusion algorithm which uses these measures while deciding when to and when not to fuse information were implemented and were tested using real data that was obtained by distributing sensor nodes along the east-west and south-north road at Twentynine Palms, CA during one of the field tests (SITEX'02) as shown in Figure 18.7. These sensor nodes are manufactured by Sensoria. On each sensor node, three sensors - acoustic, seismic and IR sensors, a four channel data acquisition board and a processing board are available. These nodes also have communication capabilities. For more details on the sensor node, refer to [14].

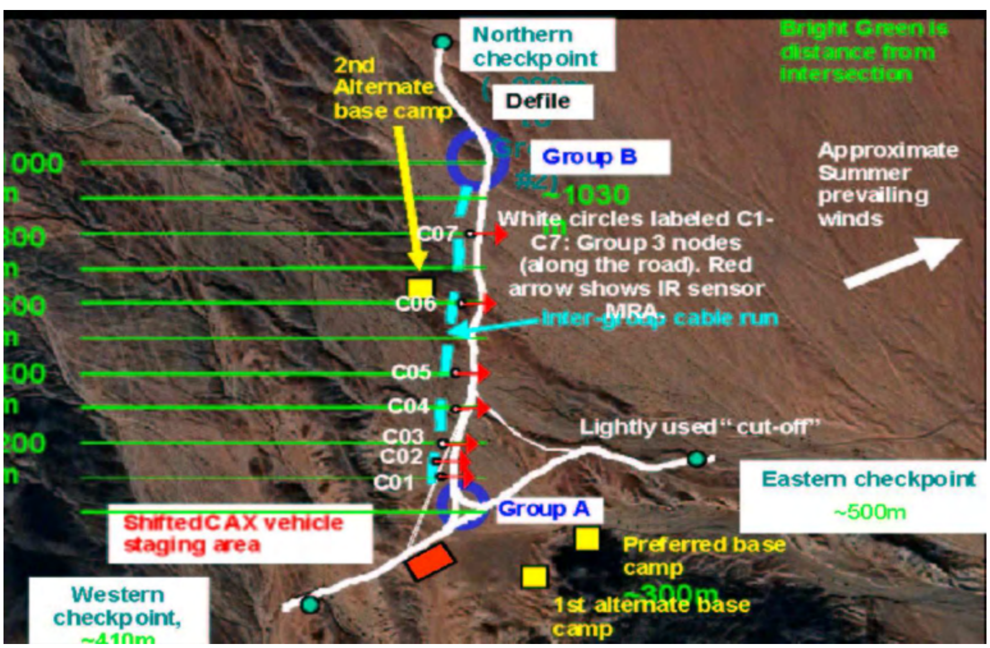


Figure 18.7: Sensor node distribution at Twenty nine Palms, CA

Three vehicles – AAV, Dragon Wagon (DW) and HMMWV were driven along the east-west and north-south road as shown in Figure 18.7 while conducting the experiments. In this figure, nodes placements are also provided. Totally twenty four nodes were considered in our experiments. We used both seismic and acoustic data from these nodes when it is appropriate. In the next section, the classification experimental details and the results are provided and in section 18.3.2.2.2 the detection experiments and the results are provided. In both these sections experimental details and results are provided with and without value of information based fusion technique that was developed in this study.

18.3.2.2.1 Classification experiments First, acoustic data from each node is considered. The maximum likelihood classifier is trained using only acoustic data from individual nodes. The challenges in the classification experiments are threefold: 1) when to reject a source of data, 2) when to propagate data between sequential nodes, and 3) when to share individual sensor data within the same node. Using only acoustic data, we investigated the effectiveness of the four measures of value of information outlined in Section 18.2.4 - mutual information, Euclidean distance, correlation, and Kullback-Liebler distance.

In addition, we investigated two methods of using these measures. When evaluating the effectiveness of fusing two sources of data, is it better to compare the two sources with each other or with the stored training data? To answer this question, we devised several similarity measures to measure the closeness of two data sources. We calculated these measures between data at all sequential nodes. Then for each similarity measure, we computed its correlation with correct classification performance at each node. We call this the performance correlation. The average performance correlation over all nodes for each class of data using previous node similarity measures is shown in Figure 18.8. Next, we calculated the same similarity measures between the data at each node and the data stored in the training sets. Again, for each similarity measure, we computed its correlation with correct classification performance at each node.

The average performance correlation over all nodes for each class of data using training set similarity measures is shown in Figure 18.9.

Inspection of Figures 18.8 and 18.9 show that the similarity measures Euclidean distance and correlation are more closely aligned with correct classification performance than either mutual information or Kullback-Liebler distance. In practice, however, we found that the Euclidean distance outperformed correlation as the determining factor in fusion decisions. Furthermore, comparing Figures 18.8 and 18.9 shows that using the training set for similarity measures is more effective than using the data from the previous node in the network. We found this to be true in practice as well. Subsequent work with the seismic data echoed the findings of the acoustic data. Note that even though we use the training data to make the fusion decision, we perform the actual data fusion with current and previous node data.

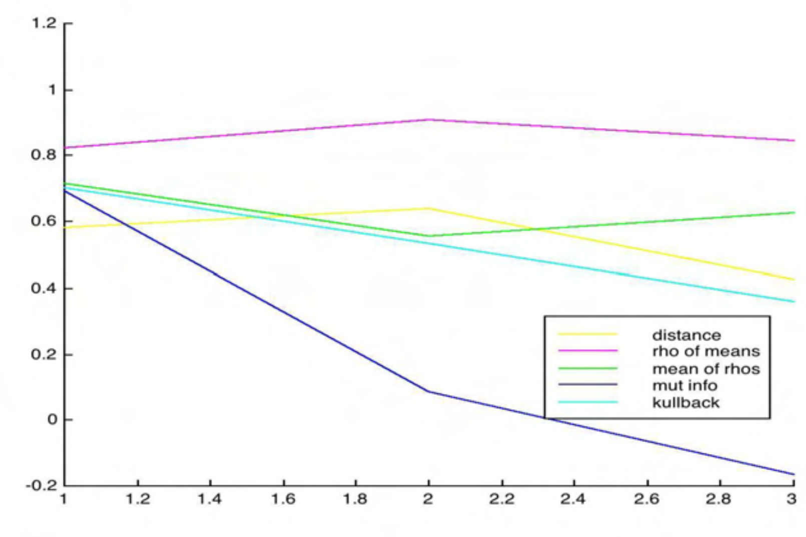


Figure 18.8: Performance correlation of previous node data

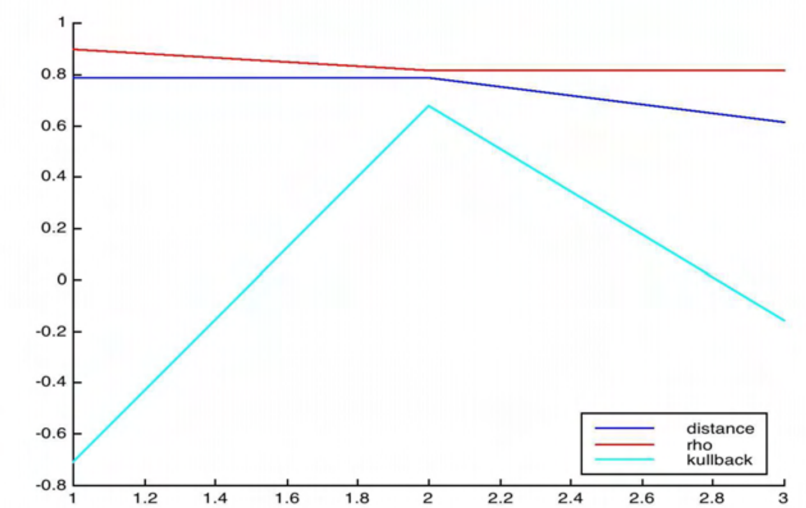


Figure 18.9: Performance correlation of training class data

Rejection of bad data Sometimes one node or one sensor can have bad data, in which case we prefer to reject this data rather than classify with poor results. The feature discrimination algorithm is used for this. By rejecting the data, we did not fuse it with any other data, pass it on to any other node, nor even compute a classification at that source. Our method resulted in the rejection of several sources of bad data, thus improving the overall classification results as shown in Figures 18.10 and 18.11.

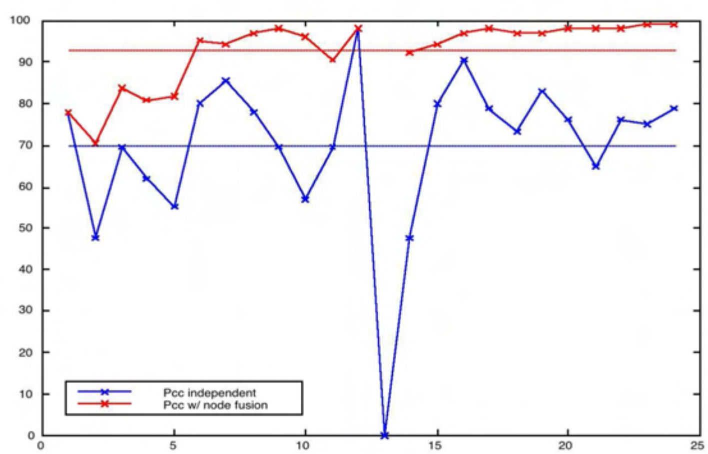


Figure 18.10: Performance of node fusion for the AAV with acoustic sensor data

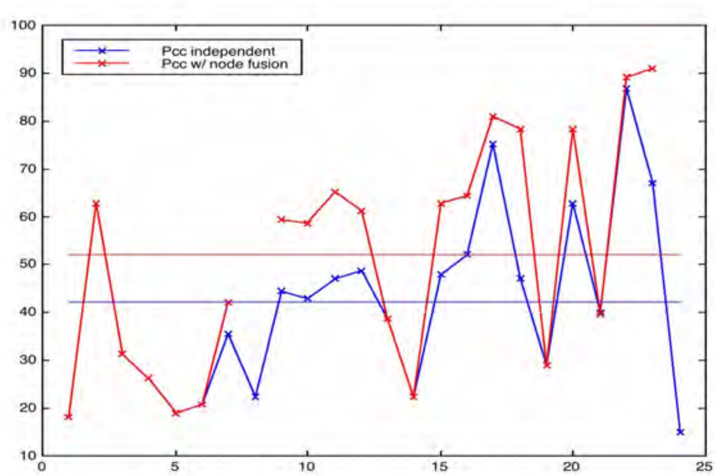


Figure 18.11: Performance of node fusion for the DW with seismic sensor data

Node to node fusion The fusion decision can be made with a threshold, i.e. if the distance between two features sets is below some value, then fuse the two feature sets. The threshold value can be predetermined off-line or adaptive. We sidestep the threshold issue, however, by basing the fusion decision on relative distances. To do so, we initially assume the current node belonged to the same class (aka the target class) as the previous node and employ the following definitions. Let x_n be the mean vector of the current node data. Let x_{nf} be the mean vector of the fused data at the current node. Let x_{c_1} be the mean vector of the target training class data. Let x_{c_2}, x_{c_3} be the mean vectors of the remaining training classes. A Euclidean distance ratio is defined as:

$$r_{dist} = d_{c_1} / \min(d_{c_2}, d_{c_3}) \tag{18.13}$$

where d_{c_i} is the Euclidean distance (18.8) between x_n and x_{c_i}. We then use the following pseudocode to make our fusion decisions.

```
if (r_dist <= 1.0)
        fuse_4class = 1;    fuse_4carry = 1;
        class_ind = classify x_n;
        if (class_ind >= 70%)  check class_fuse;
        end
else
        fuse_4class = 0;    fuse_4carry = 0;
        if {(d_c1 <= 3s_c1) & (d_c2 <= 3s_c2) & (d_c2 <= 3s_c2)}
                class_ind = classify x_n;
                if (class_ind == target class)  fuse_4class = 1;
                        if (class_ind >= 70%)
                                fuse_4carry = 1;
                                class_fuse = classify x_nf;
        if (class_ind > class_fuse)
                                        class_fuse = class_ind;
        end
                        end
                end
        else
                reject this data;
        end
end
```

There are two outcomes to the fusion decision. First we decide whether or not to fuse the data at the current node. If the current node has bad data, fusion can pull up the performance, however, we may not want to carry the bad data forward to the next node (the second fusion decision outcome). `fuse_4class` is a flag indicating whether or not to fuse for the current classification. `fuse_4carry` is a flag indicating whether or not to include data from the current node in the fused data that is carried forward. Based on this decision, the fusion of classification decision is achieved by applying the fusion algorithm described in section 18.2.5. In Figures 18.10 and 18.11 we show the correct classification improvement gained by fusing from node to node for the acoustic and seismic sensors, respectively. For the acoustic sensor we show classification results from the AAV data, while using DW data for the seismic sensor results. In the case of the acoustic data, the mean correct classification performance across all nodes increases from 70% for independent operation to 93% with node to node fusion across the network. Similarly, the seismic correct classification performance increases from 42% to 52%.

Fusion between sensors After fusion from node to node of the individual sensors, we look at the benefit of fusing the acoustic and seismic sensor data at the same node. To do so, we employ the following definitions. Let r_{dist} be defined as in (18.13) but with the new data types (a - acoustic, s - seismic, and as – a concatenated acoustic/seismic vector). Let xa be the mean vector of the current node acoustic data after fusion from node to node. Let xs be the mean vector of the current node seismic data after fusion from node to node. Let $x_{as} = x_a$ concatenated with x_s (dumb fusion). Let x_{asf} = smart fusion of x_a with x_s. Let x_{in} be the data input to the classifier. Now, we employ two steps in the sensor fusion process as shown in the pseudocode below. In this case also for the fusion of features from two independent sources such as acoustic and seismic, DSm based technique described in section 18.2.5 is applied. First we employ a smart sensor fusion routine:

```
indx = min(r_a_dist, r_s_dist, r_as_dist)
if (indx == 1)  x_in = x_a;
elseif (indx == 2)  x_in = x_s;
elseif (indx == 3)  x_in = x_as;
end
```

Next, we employ a final fusion routine:

```
class_acst = classify x_a;
class_seis = classify x_s;
class_as_dumb = classify x_as;
class_as_smart = classify x_asf;
if { (class_acst >= 70%) | (class_seis >= 70%) | (class_as_ind >= 70%) }
        class_final_fuse = max (class_acst, class_seis,    class_as_dumb, class_as_smart)
end
```

Figure 18.12 shows the results of fusion at each stage in the form of a bar plot. The classification performance is averaged over all the nodes for each vehicle class. The correct classification performance improves at each stage of fusion processing as shown in Table 18.1. The results indicate that the fusion based on value of information helps in improving the decision accuracy at each node significantly.

	AAV	DW	HMMV
Acoustic independent	70%	58%	46%
Seismic independent	72%	42%	24%
Acoustic fusion	93%	80%	69%
Seismic fusion	93%	52%	31%
Acoustic & seismic, independent	76%	55%	58%
Acoustic & seismic with fusion	95%	90%	77%

Table 18.1: Summary of classification performance

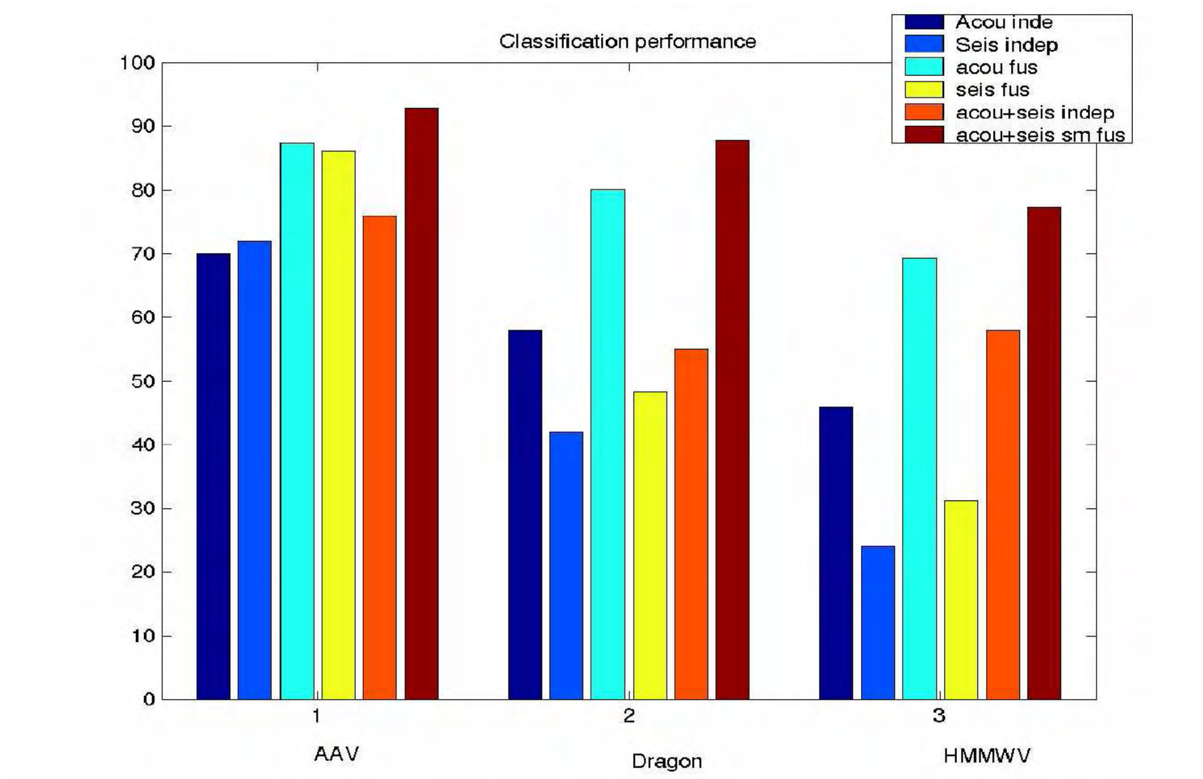


Figure 18.12: Average correct classification performance at each step in the fusion process

18.3.2.2.2 Detection experiments For the detection experiments also both acoustic and seismic data were considered. First, only acoustic data from individual nodes were used. A threshold value was initially set which was varied adaptively based on the background energy. The power spectral density of acoustic data was computed using 1024 point FFT and it was downsampled by 8. The energy of the downsampled version of the power spectral density was computed. This energy was compared with the threshold value. If the energy was above the threshold value, it was decided that the target was detected. The time of detection and the confidence on detection were also calculated. The detection and time of detection were compared with the ground truth. If the target was detected when it is supposed to be and if the time of detection is within the region of interest then it was counted towards calculating the probability of detection. If the detection time is outside the region of interest (missed detection) and if a target was detected when it should not have been (false alarm) it was counted towards computing the probability of false alarm. The probability of detection and false alarm using only acoustic data from individual nodes without any fusion for AAV, DW and HMMWV are: 0.8824, 0.8677, 0.8382 and 0.1176, 0.1323, 0.1618, respectively. Similarly, the probability of detection and false alarm using only seismic data from individual nodes without any fusion for AAV, DW and HMMWV are: 0.8030, 0.7910, 0.5735 and 0.1970, 0.2090, 0.4265, respectively.

Next, the mutual information based value of information measure was used on the energy of power spectral density to make a decision of fusing data between sensors - acoustic and seismic on each individual node. The detector was tested using the fused data on each node. The probability of detection and false alarm were computed as described above. The probability of detection of this intelligently fused data for AAV, DW and HMMWV is: 0.9394, 0.9105 and 0.8529, respectively. The probability of false alarm is not provided here because it is equal to 1 − probability of detection since both false alarm and missed detections are combined together. These results are summarized in Figure 18.13 in the form of a bar graph. From this, it can be seen that the intelligent sensor data fusion based on value of information and DSmT significantly improves the detection accuracy. This type of fusion especially helps in difficult data as in the case of HMMWV.

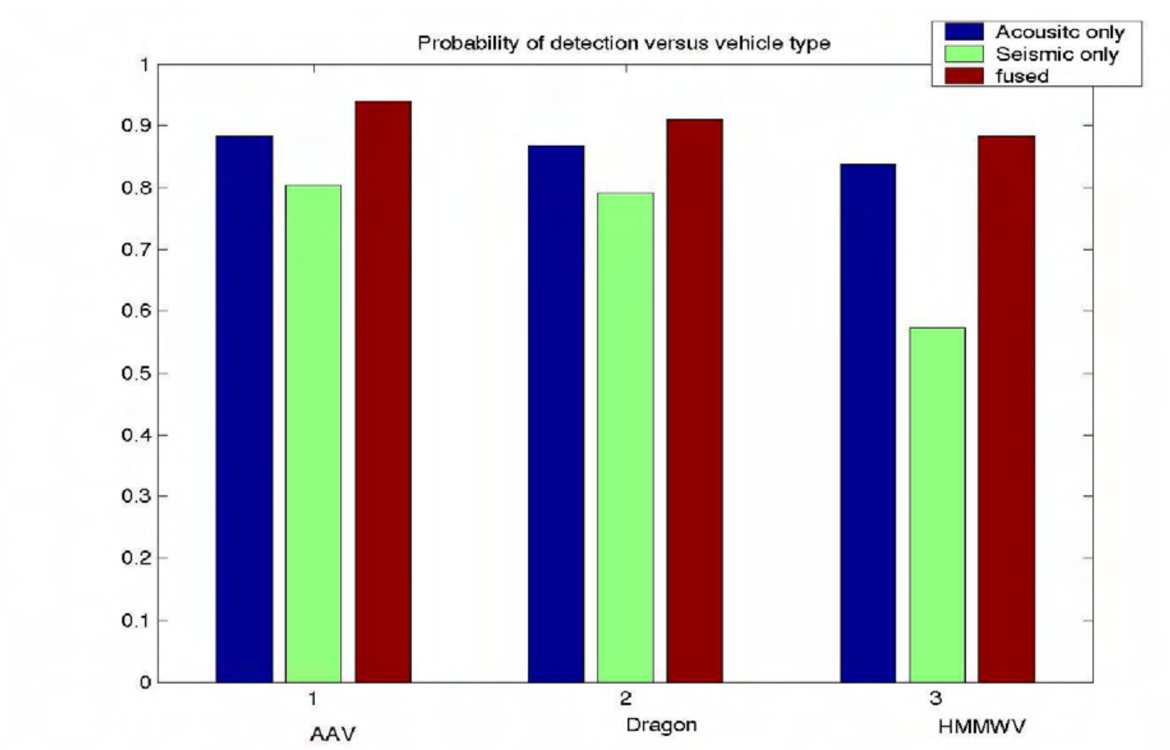


Figure 18.13: Performance of a detector

18.4 Conclusions

In this chapter, we have described how minmax entropy principle can be used in feature (missing information) discovery and the type of sensor (information type) from which this missing information can be obtained. Further, a consistency measure is defined and it has been shown that this measure can be used in discriminating or assessing the quality of sensors. Next, conditional entropy and mutual information measures are defined and it has been shown that these two measures can be used in making sure that the estimated missing information or new feature set indeed help in gaining information and aid in decision process. Further more, we have introduced several measures for value of information. We have used these measures in deciding when to fuse information. For the fusion we have developed an algorithm using DSmT. We have proven the concept of all the measures and fusion by first considering a simulated network of radar sensors and then by considering a real network of spatially distributed sensor nodes which have multiple sensors on each sensor node. The experimental results indicate that (a) the minmax entropy principle can be used in estimating the missing information and information type and it can be used in the cluster formation; (b) the constancy measure based on within class entropy can be used in sensor discrimination; (c) the mutual information can be used in feature quality assessment and in evaluating the value of information; (d) the measures of value of information helps in smart fusion; (e) the distributed smart fusion significantly improves the decision accuracy. All these measures help in probing (awakening) the required sensor for the required missing information, only transmitting the valuable information when and where it is needed and fusing only valuable information. Thus, power and, computing and communication resources can be efficiently utilized.

18.5 References

[1] Bar-Shalom Y., Fortmann T.E., *Tracking and data association*, Academic press, San Diego, 1988.

[2] Battti R., *Using mutual information for selecting features in supervised neural net learning*, IEEE Trans. On Neural Network, Vol. 5, No. 4, July 1994, pp. 537-550.

[3] Beck S., *Energy based detector*, SensIT PI meeting, Jan 15-17, 2002, Santa Fe, New Mexico.

[4] Darbellay G.A., Vajda I., *Estimation of the information by an adaptive partitioning of the observation space*, IEEE Transactions on Information Theory, vol. 45, no. 4, pp. 1315-1321, May 1999.

[5] Dezert J., Smarandache F., *Combination of beliefs on hybrid DSm models*, Advances and Applications of DSmT for Information Fusion, American Research Press, Rehoboth, June 2004, Chap. 4.

[6] Kadambe S., *Information theoretic based sensor discrimination for information fusion and cluster formation in a network of distributed sensors*, in Proc. of 4th International Conference on Information fusion, Montréal, Québec, Canada, August 7-10, 2001, pp. ThC1-19-ThC1-25.

[7] Kadambe S., *Feature discovery and sensor discrimination in a network of distributed sensors for target tracking*, in Proc. of IEEE workshop on Statistical signal processing, Singapore, August 6-8, 2001, pp. 126-129.

[8] Kadambe S., Daniell C., *Sensor/Data Fusion Based on Value of Information*, in Proc. of 6th International Conference on Information Fusion, Fusion 2003, Cairns, Australia, July 2003, pp. 25-32.

[9] Kapur J. N., *Measures of information and their applications*, John Wiley, Eastern Limited, 1994.

[10] Manyika J., Durrant-Whyte H., *Data fusion and sensor management: An information theoretic approach*, Prentice Hall, 1994.

[11] Papoulis A., *Probability, Random variables and Stochastic Processes*, Second edition, McGraw Hill, 1984, pp. 500-567.

[12] Press W.H., Teukolsky S. A., Vetterling W. T., Flannery B. P. , *Numerical Recipes in C*, Cambridge University Press, 1992, pp. 632-635.

[13] Rabiner L. R., Juang B-H. , *Fundamentals of Speech Recognition*, Prentice Hall, 1993, Chapter 6.

[14] `http://www.sensoria.com`

[15] Thomopoulos S. C. A. , *Sensor selectivity and intelligent data fusion*, Proc. Of the IEEE MIT'94, October 2-5, 1994, Las Vegas, NV, pp. 529-537.

[16] Hen Wu Y., *Maximum likelihood based classifier*, SensIT PI meeting, Jan 15-17, 2002, Santa Fe, New Mexico.

Biographies of contributors

Samuel Corgne is currently PhD student at the University of Rennes 2, France (COSTEL, CNRS UMR 6554 LETG). His current research area is on the remote sensing and (high and very high resolution) image processing applied to the land cover modelling and prediction in Brittany, France. E-mail: `samuel.corgne@Uhb.Fr`.

Frédéric Dambreville studied mathematics, logic, signal and image processing. He received the Ph.D. degree in signal processing and optimization, from the university of Rennes, France, in 2001. He worked as a postdoctorate in the Naval Postgraduate School (U.S.A.) in 2001. In 2003, he joined the department image, perception and robotic of the CTA laboratory (Délégation Générale pour l'Armement), France. His main interests are in optimization/planning, data/sensor fusion, Markov models/Bayesian networks. E-mail: `Frederic.Dambreville@dga.defense.gouv.fr`.

Milan Daniel was born in Prague in 1962. He graduated in the Faculty of Mathematics and Physics of Charles University Prague in 1985. He defended his PhD thesis in the Institute of Computer Science of the Academy of Sciences of the Czech Republic in 1993. His research activities have been always related to the Institute, the Department of Theoretical Computer Science, formerly the Department of Knowledge Based Systems, see `www.cs.cas.cz`. Author's current main scientific interests are belief functions, namely combination of belief functions and probabilistic transformations of belief functions. The other interests are uncertainty processing, fuzzy logic and knowledge based systems E-mail: `milan.daniel@cs.cas.cz`.

Denis De Brucq is honorary Professor of Rouen University, Rouen, France and Member of the "Académie des Sciences Belles Lettres et Arts de Rouen". He has developed finite filtering for special probability evolution laws and also extended Kalman filtering approximations for non-gaussian and non-linear models in discrete time evolutions. But from applications in handwritten descriptions and melanoma detection, he has recently proposed fusion constructions to take care of varied information sources. Web page:`denis.debrucq.free.fr/DBMA.htm`, E-mail: `denis.debrucq@free.fr`.

Jean Dezert was born in l'Hay les Roses, France, on August 25, 1962. He received the electrical engineering degree from the École Française de Radioélectricité Électronique and Informatique (EFREI), Paris, in 1985, the D.E.A. degree in 1986 from the University Paris VII (Jussieu), and his Ph. D. from the University Paris XI, Orsay, in 1990, all in Automatic Control and Signal Processing. During 1986-1990 he was with the Systems Department at the Office National d'Études et de Recherches Aérospatiales (ONERA), Châtillon, France, and did research in tracking. During 1991-1992, he visited the Department of Electrical and Systems Engineering, University of Connecticut, Storrs, U.S.A. as an European Space Agency (ESA) Postdoctoral Research Fellow. During 1992-1993 he was teaching assistant in Electrical Engineering at the University of Orléans, France. Since 1993, he is senior research scientist in the Image Estimation and Decision (IED) Research Lab. with the Information and Modelling and Processing Department (DTIM) at ONERA. His current research interests include autonomous navigation, estimation theory, stochastic systems theory and its applications to multisensor-multitarget tracking (MS-MTT), information fusion and plausible reasoning. Dr. Jean Dezert is developing since 2001 with Professor Smarandache a new theory of plausible and paradoxical reasoning for information fusion (DSmT) and has edited the first textbook (collected works) devoted to this new emerging research field published by American Research Press, Rehoboth in 2004. He owns one international patent in the autonomous navigation field and has published several papers in international conferences and journals. He coauthored one chapter in Multitarget-Multisensor Tracking: Applications and Advances, Vol.2 (Y. Bar-Shalom Editor). He is member of IEEE and of Eta Kappa Nu, serves as reviewer for different International Journals, teaches a MS-MTT course at the French ENSTA Engineering School, collaborates for the development of the International Society of Information Fusion (ISIF) since 1998, and has served as Local Arrangements Organizer for the Third International Conference on Information Fusion, FUSION 2000, July 10-13, in Paris. He has been involved in the Technical Program Committees of Fusion 2001-2004 International Conferences. Since 2001, he is a member of the board of the International Society of Information Fusion (`http://www.isif.org`) and served as secretary for ISIF since 2001. He serves as executive vice-president of ISIF in 2004. In 2003, he organized with Professor Smarandache, the first special session devoted to plausible and paradoxical reasoning for information fusion at the International conference on Information Fusion, Fusion 2003, Cairns, Australia and also a panel discussion and a special session on DSmT at Fusion 2004, Stockholm in June 2004. E-mail:`jean.dezert@onera.fr,jdezert@yahoo.com`.

Mohamad Farooq received his BscEng and MTech degrees from Punjab Engineering College, Chandigarh, India and the Indian Institute of Technology, Delhi in 1965 and 1967, respectively, and his PhD from the University of New Brunswick, Canada, in 1974, all in Electrical Engineering. In March 1980, he joined the Royal Military College of Canada (RMC), Kingston, Ontario, where he is currently a professor in the Department of Electrical and Computer Engineering. During 1998-1999, he was on Sabbatical at the Computing Devices Canada Ltd. (CDC), a General Dynamics Company, where he was involved in Technology Transfer between RMC and CDC in the areas of Target Tracking and Sensor Data Fusion, with possible applications to the Swedish's coastal sonar Defense systems, a contract recently awarded to CDC. During his Sabbatical in 1991-1992 at Thomson-CSF Canada, he was instrumental in preparing proposals for a number of research projects in the area Multisensor Multitarget Tracking and securing the necessary funding on these projects. He has organized as well as served as technical chair on a number of conferences on Applications of Advance Technologies to Canadian Forces, and served as a Co-Editor of the resulting Conference Proceedings. He continues to be a member of the Steering Committee of the Midwest Symposium on Circuits and Systems. He has served on a number of Program Committees on various International Conferences, such as the Conference Editorial Board of the IEEE Control Society, the IEEE Industrial Electronics Conferences, the Internal Society for Optical Engineering (SPIE Conferences), Fusion '98, '99, '2000, '2002 and '2003 to name a few. In 1998, he was nominated as a candidate to the First Board of Directors of the International Society of Information Fusion. He has also been appointed to the Editorial Board of the International Journal on Multisensor Information Fusion published by Elservier Science S.A. He was a member of the Defense Advisory Board (Canada) which drafted a report on Fusion Technology in Canada. He has presented a number of invited lectures on Sensor Fusion Technology and its Applications. He has served as a Consultant to a number of Defense Consulting Companies in Canada. He has been instrumental in establishing a research group on Multisensor Multitarget Tracking at RMC. He holds a number of research contracts on various aspects on Target Tracking, Robotics, Motor Control and Drives and Sensor Data Fusion and has published numerous technical papers in these areas. E-mail:`farooq@rmc.ca`.

Laurence Hubert-Moy holds her Ph.D. degree in 1989 from the University of Rennes in Geography. She is an Associate Professor at the University of Rennes, Department of Geography. She also served as a research scientist in 2000-2002 at CNRS (French National Center for Scientific Research) and a visiting scientist in the Department of Geography, Center for Remote Sensing, Boston University, Massachusetts. Dr. Hubert-Moy's research has extensively focused in the area of agriculture applications of remote sensing data. Specifically, she has been focusing her efforts in various critical areas such as: the local and regional land cover and land cover change mapping in intensive farming areas; wetlands identification and characterization with remote sensing data; spatial modeling and spatial statistics as applied to image understanding and use. Dr. Hubert-Moy is the recipient of the Bronze medal of the French National Center for Scientific Research in 2003. E-mail: `Laurence.Hubert@Uhb.Fr`.

Anne-Laure Jousselme received a PhD degree in Signal Processing from Laval University (Québec, Canada) and the "Institut National Polytechnique de Grenoble" (France) in 1997. In 2002, she joined the Defence Research and Development Canada in Valcartier in the Decision Support System Section for the Situation Analysis Support Systems Group. Since 2000, she is Adjunct Professor at the Laval University (Québec city, Canada) in the Electrical Engineering and Computer Engineering Department, where she supervises PhD and master students through a collaborative grant whose purpose is "Quality assessment and unified information fusion for object recognition". This NSERC collaborative project with Lockheed Martin Canada focuses specifically on the study of general frameworks for information fusion such as random sets and fuzzy evidence theory, as well as the modelisation of imperfect information representing uncertainty, reliability, and utility. Her current research interests include information fusion, reasoning under uncertainty, generalized information theory, situation analysis, fusion of classifiers, and performance measures of target identification and classification systems. E-mail: `Anne-Laure.Jousselme@drdc-rddc.gc.ca`.

Shubha Kadambe obtained her PhD EE from the Univ. of Rhode Island in May 1991. Since then she has had held various research positions at A. I. DuPont Institute, At & T Bell Laboratories and Atlantic Aerospace Electronics Corporation. Currently she is a senior research scientist at HRL Laboratories, LLC. She has extensively worked in the areas of sensor processing, speech, time-frequency representations, advanced signal and image processing, fusion and bio-medical signal processing. She has over sixty technical papers, one video tutorial, three invited chapters and one monogram (to be published by John Wiley) in these areas. Her biography has been included in Who's-who in America and who's-who of American women published by Marquis in 1998 and 2002, respectively. She has served as an associated editor for IEEE trans. On signal processing and as a member of technical committee of several international conferences. She has also chaired several technical sessions in various international conferences. Her research interests include time-frequency/time-scale and their applications to speech, sonar and radar, smart fusion, speech, neural networks, image processing and non-linear signal processing. She has recently been elected to the grade of senior member of IEEE by its officers and board of directors in recognition of professional standing. She is an adjunct professor at the University of Maryland Baltimore campus (UMBC). She has taught (teaches) graduate and undergraduate courses at UCLA, UMBC and California Institute of Technology. She has guest lectured at the University of Southern California. E-mail: `skadambe@hrl.com`.

Mohammad Khoshnevisan is born in Iran and he partly grew up in the United States. Dr Khoshnevisan is a lecturer (Assistant Professor) in Finance at Griffith University Australia. He has completed his B.A in Mathematics at CBU in Memphis and he has obtained his M.Com in Wollongong, Australia.

In 2000 he has completed his Ph.D from the University of Melbourne, Department of Computer Science and Software Engineering. Dr. Khoshnevisan is worldwide known for his expertise in Financial Eng. and fuzzy systems. Dr. Khoshnevisan is affiliated with BISC at UC-Berkeley and he has been invited by both Harvard University and UC-Berkeley as a visiting Professor. Dr. Khoshnevisan is the author of over 34 papers and five of his best papers have been cited by the American Mathematical Society. E-mail: `m.khoshnevisan@griffith.edu.au`

Pavlina Dimitrova Konstantinova received M.Sc. and Ph.D. degrees in Computers and Computer-Aided Transient and Steady State Analysis of Nonlinear Electronic Circuits from Technical University of Sofia in 1967 and 1987, respectively. From 1971 to 1989 she was research assistant at the Radio electronics Research Institute. Since 1989 she is assistant research professor at the Department of Mathematical Methods for Sensor Information Processing, Institute for Parallel Processing - Bulgarian Academy of Sciences. She is a member of the Bulgarian Union of Automation and Information, member of Image Recognition Society, BAS. Her research interests include sensor data processing, data association, mathematical modeling, parallel algorithms and object-oriented programming. E-mail: `pavlina@bas.bg`.

Patrick Maupin joined in 2001 the Decision Support System Section and the Situation Analysis Support Systems Group of Canada's Defence Research and Development Canada establishment in Valcartier. Between 2000 and 2002 he taught Neural Networks and Fuzzy Systems as well as Multivariate Statistical Analysis at Ecole de technologie supérieure in Montréal (Canada) where he is still a research associate and supervises graduate students involved in human-centered pattern recognition projects. From 1998 until 2001 he worked as a consultant in information processing, in particular for Québec's Public Health Authority and as a research assistant at McGill University on hyperspectral data processing. The Natural Sciences and Engineering Research Council of Canada and the Canadian Space Agency supported his graduate studies. His research interests include information fusion, reasoning under uncertainty, epistemic logic, situation analysis, human-centered pattern recognition and optimization of ensembles of classifiers. `Patrick.Maupin@drdc-rddc.gc.ca, pmaupin@livia.etsmtl.ca`.

Grégoire Mercier graduated from the Institut National des Télécommunications (INT), Evry, France in 93 and received the Ph.D. degree from the University of Rennes, Rennes, France in 99. He is currently associate professor at the Ecole Nationale Supérieure des Télécommunications de Bretagne (ENST Bretagne), Brest, France, where his current research are the remolety sensed image processing. More precisely, he focuses mostly his topics on the compression and the segmentation of SAR and hyperspectral data. E-mail: `gregoire.mercier@enst-bretagne.fr`.

Tzvetan Atanasov Semerdjiev was born on 5th of July 1949 in Pazardjik, Bulgaria. M.Sc. (Avionics) – 1973, Ph.D. (Air Traffic Control Systems) - 1977, Zhukovsky Air Force Engineering Academy, Moscow Russia and D.Sc. (C3I) - 1985, "G.S. Rakovski" Defense and Staff College. During 1985 – 1988, he served as Director of the Special Electronics Institute and head of Advanced Research and Development Department. During 1988 – 1992 he was the Head of Laboratory "Signal" and Deputy Director of the Center of Informatics and Computer Technology, Bulgarian Academy of Sciences. Since 1992, he is Professor in the Central Laboratory for Parallel Processing, Bulgarian Academy of Sciences, head of "Mathematical Methods for Sensor data Processing" department. Since 1999, he is Professor in "G. S. Rakovski" Defense and Staff College and Head of National and International Security department. Professional affiliations: IEEE-AES, AFCEA (life membership), ISIF, AIPT (Int. Academy of Information Processes and Technologies) - Moscow, Russia. Teaching: "G. S. Rakovski" Defense and Staff College, Department of Air Transport of Sofia Technical University, Academy of Ministry of Interior. Fields of Research - Radar and Navigation, Model Estimation and Identification, Target Tracking, Modeling and Simulation, Multisensor Data Fusion, Information Security. Supervisor of 8 Ph.D. thesis's and 10 international R&D projects. Over 200 publications, including books, papers, R&D reports and documents in Bulgarian, Russian and English. E-mail: `signal@bas.bg`

Florentin Smarandache was born in Balcesti, Romania, in 1954. He got a M. Sc. Degree in both Mathematics and Computer Science from the University of Craiova in 1979, received a Ph. D. in Mathematics from the State University of Kishinev in 1997, and continued postdoctoral studies at various American Universities (New Mexico State University in Las Cruces, Los Alamos National Laboratory) after emigration. In 1988 he escaped from his country, pasted two years in a political refugee camp in Turkey, and in 1990 emigrated to USA. In 1996 he became an American citizen. Dr. Smarandache worked as a professor of mathematics for many years in Romania, Morocco, and United States, and between 1990-1995 as a software engineer for Honeywell, Inc., in Phoenix, Arizona. In present, he teaches mathematics at the University of New Mexico, Gallup Campus. Very prolific, he is the author, co-author, and editor of 62 books, over 85 scientific notes and articles, and contributed to about 50 scientific and 100 literary journals from around the world (in mathematics, informatics, physics, philosophy, rebus, literature, and arts). He wrote in Romanian, French, and English. Some of his work was translated into Spanish, German, Portuguese, Italian, Arabic, Esperanto. He was so attracted by contradictions that, in 1980s, he set up the "Paradoxism" avant-garde movement in literature, philosophy, art, even science, which made many advocates in the world, and it's based on excessive use of antitheses, antinomies, paradoxes in creation - making an interesting connection between mathematics, philosophy, and literature [`http://www.geocities.com/charlestle/paradoxism.html`] and led him to coining the neutrosophic logic, a logic able to deal with paradoxes. In mathematics there are several entries named

Smarandache Functions, Sequences, Constants, and especially Paradoxes in international journals and encyclopedias. He organized the 'First International Conference on Neutrosophics' at the University of New Mexico, 1-3 December 2001 [http://www.gallup.unm.edu/~smarandache/FirstNeutConf.htm]. Small contributions he had in physics and psychology too. Much of his work is held in "The Florentin Smarandache Papers" Special Collections at the Arizona State University, Tempe, and Texas State University, Austin (USA), also in the National Archives (Rm. Vâlcea) and Romanian Literary Museum (Bucharest), and in the Musée de Bergerac (France). In 2003, he organized with Dr. Jean Dezert, the first special session devoted to plausible and paradoxical reasoning for information fusion at the Fusion 2003 International conference on Information Fusion in Cairns, Australia. Web info: http://www.gallup.unm.edu/~smarandache/. E-mail:smarand@unm.edu.

Hongyan Sun received her B.S. and M.S. degree in Mathematics Dept. at Harbin Institute of Technology in China in 1987 and 1992, respectively, and her Ph.D. degree in Dept. of Electrical Engineering at Beijing University of Aeronautics and Astronautics in China in 1996. Then she worked at State Key Lab of Intelligent Technology & Systems and Dept. of Computer Science & Technology at Tsinghua University in China and at CIM Lab of Dept. of Mechanical Industrial Engineering at University of Toronto in Canada as a post-doctoral fellowship. Now she works in Dept. of Electrical & Computer Engineering at Royal Military College of Canada as a research associate. Her main research interests are theory and application of Multi-source information processing, Communications, AI, and Mathematics application to engineering. E-mail: hongyansun_11@yahoo.com.

Albena Plamenova Tchamova is assistant research professor at the Department of Mathematical Methods for Sensor Information Processing, Institute for Parallel Processing, Bulgarian Academy of Sciences. She received M.Sc. degree in Microelectronics and Ph.D. degree in Radiolocation and Radionavigation (Target Attributes Identification and Classification on the base of Evidence Reasoning Theory and Fuzzy Logic) from the Technical University of Sofia, Bulgaria, in 1978 and 1998 respectively. She worked as a researcher in numerous academic and industry projects (Bulgarian and international) in the area of Multisensor Multitarget Tracking (MS-MTT) and Data Fusion. She is a member of the Association of European Operational Research Societies on Fuzzy Sets (EUROFUSE), member of Image Recognition Society, BAS. Dr. Tchamova's current research interests include MS-MTT, Target Identification, Information Fusion, DSm Theory of Plausible and Paradoxical Reasoning, Decision Making under Uncertainty. E-mail: tchamova@bas.bg.

This book presents the recent theoretical advances and applications of the Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning for information fusion. DSmT proposes a new mathematical framework to deal with the combination of uncertain, imprecise and highly conflicting sources of information expressed in terms of generalized basic belief functions. DSmT works beyond the limits of the Dempster-Shafer Theory and proposes a new general rule of combination which does not require a normalization step and works with any models (free DSm, hybrid DSm and Shafer's model) whatever the degree of conflict between sources is. The DSmT is well adapted for static and dynamic fusion problematics and allows to work on any finite frames (discrete, continuous and/or hybrid). Therefore it can combine belief functions even if the refinement of the frame of discernment is inaccessible because of the vague, relative, and imprecise intrinsic nature of its elements. Part 1 of this book presents in details the last investigations on DSmT but also some related theoretical works with other approaches for information fusion. Part 2 of the book presents interesting applications of DSmT for solving both academic fusion problems and real-world fusion problems.

$22.50

ISBN 1-931233-82-9

52250>

9 781931 233828