Approved for Public Release; Distribution Unlimited
Case # 08-0173

Algebra of Dempster-Shafer evidence accumulation

Robert H. Enders [1]*[a], Andrzej K. Brodzik [a], Michael R. Pellegrini [a]
[a]The MITRE Corporation, Burlington Road, Bedford, MA 01730

ABSTRACT

In this work we focus on the relationship between the Dempster-Shafer (DS) and Bayesian evidence accumulation. While it is accepted that the DS theory is, in a certain sense, a generalization of the probability theory, the approaches vary in several important respects, including the treatment of uncertain information and the way the evidence is combined, making direct comparison of results of the two analyses difficult. In this work we ameliorate these difficulties by proposing a mathematical framework within which the relationship between the two methods can be made precise. The findings of the investigation elucidate the role uncertainty plays in the DS theory and enable evaluation of relative fitness of the two techniques for practical data fusion scenarios.

Keywords: data fusion, evidence accumulation, Dempster-Shafer theory, Bayesian inference, uncertainty, semigroup, semigroup homomorphism

1. INTRODUCTION

Dempster-Shafer (DS) theory is one of the main tools for reasoning about data obtained from multiple sources, subject to uncertain information [12], [16]. The principal task of such reasoning is data fusion, or evidence accumulation. The goals of data fusion are to identify the most likely of a set of alternatives and to decrease uncertainty by accumulating evidence from multiple sources. The DS approach has applications in several areas, including sensor fusion, medical diagnostics, biometrics, and decision support. [2], [17], [18], [23], [24]

Despite the ubiquity of the DS technique in science and engineering, several problems remain unsolved, making an effective translation of theory into practice difficult. Among these problems are: (1) lack of model-based rules for mass assignment, (2) lack of theoretical justification for the evidence combination rule, (3) lack of an appropriate formalism that would allow one to interpret evidence combination results in probabilistic terms, (4) high asymptotic complexity of the evidence combination computation, and (5) doubts about the treatment of incompatible evidence. [9], [13], [30], [31]

Many alternative approaches to evidence accumulation have been proposed to address these issues. [1], [20], [21], [22], [27] [28], [29] While these approaches seem to address some of the aforementioned concerns to a certain degree, they often increase the complexity of the analysis and generally yield different results. Since no consensus on which of these methods is to be preferred has emerged, it appears that, despite its shortcomings, the original DS approach still remains the standard reference.

In this work we focus on the relationship between the DS and Bayesian evidence accumulation. There are several reasons for doing so. First, this relationship encapsulates the key ideas of data fusion and links all of the aforementioned problems. Second, since the latter technique is both more familiar and better understood, it is reasonable to expect that specification of the relationship between the two analyses will produce new insights into the DS theory. While it is accepted that the DS theory is, in a certain sense, a generalization of the probability theory [2][26], the approaches vary in several important respects, including the treatment of uncertain information and the way the evidence is combined, making direct comparison of results of the two analyses difficult. In this work we ameliorate these difficulties by proposing a mathematical framework within which the relationship between the two methods can be made precise and computations performed by the two analyses can be compared. The findings of the investigation elucidate the role

1 *renders@mitre.org; phone 1 782 271-3588; fax 1 781 271-7045; www.mitre.org

uncertainty plays in the DS theory and enable evaluation of the relative fitness of the two techniques for practical data fusion scenarios.

The approach chosen for the analysis of the DS evidence accumulation is based on semigroup theory. This is appropriate as, properly defined, the DS set is a semigroup with respect to the DS evidence combining operation. The abstract algebra approach allows one to access the DS theory at the most general level, highlighting in the process its most essential properties. Focusing on the semigroup-theoretic structures of the fundamental DS concepts reveals key relationships between certain special cases of the DS analyses and, most importantly, between the DS and Bayes analyses. In particular, it is shown that operation on certain subsets of the DS set is equivalent to combination of Bayesian probabilities, and that, in certain cases, up to the uncertainty value, the results of the two analyses are identical.[4]

The content of the paper is as follows.
Section 2 reviews Dempster-Shafer evidence accumulation.
Section 3 reviews Bayesian evidence accumulation
Section 4 provides a resume of some elementary facts of group and semigroup theory.
Section 5 investigates the semigroup structure of Dempster-Shafer and states the main results relating it to Bayes.
Section 6 gives concluding remarks.

.

2. DEMPSTER-SHAFER EVIDENCE ACCUMULATION

Denote by $\Omega = \{\omega_1, \ldots \omega_n\}$ the non-empty set of *n* possible outcomes (alternatives) of the event of interest, and let 2^Ω denote the power set of Ω (the collection all possible subsets of Ω). A Dempster-Shafer mass assignment is a mapping a: $2^\Omega \mapsto [0,1]$ such that $a_\varnothing = a(\varnothing) = 0$ and

$$\sum_{A \in 2^\Omega} a_A = 1$$

We call a_A the mass of *A*. By an abuse of notation we will also write $a = \{a_A\}_{A \in 2^\Omega}$ and refer to *a* as a *mass assignment* or a Dempster-Shafer element. Finally, we will call the pairing of each set *A* with its corresponding masses a_A the *body of evidence* of *a*. We will distinguish a special type of assignment, in which mass is assigned only to singletons or to the entire set Ω, by calling it singleton Dempster-Shafer.

The key difference between probability and mass is that probability is a measure and therefore it satisfies the additivity condition, that is, given a finite sequence A_i of disjoint subsets of Ω,

$$P\left(\bigcup A_i\right) = \sum P(A_i).$$

In general, this additivity condition is not satisfied by a mass assignment.

Removing the additivity constraint can be convenient, as it permits inclusion of subjective judgments in the DS information fusion system, but it also has the undesirable consequence of making the interpretation of results of such fusion problematic. In particular, when considered together with the DS rule of combination, it is not always clear when mass can be made consistent with the standard probability evaluation.

A key feature of the Dempster-Shafer theory is the rule for combining bodies of evidence. Let *a* and *b* be two distinct bodies of evidence. The DS rule for combining the masses of *a* and *b* into a resultant body of evidence *c* is given by

$$c_C = \frac{1}{1-\kappa} \sum_{A \cap B = C} a_A b_B, \quad C \in 2^\Omega$$

where $\kappa = \sum_{A_i \cap B_j = \varnothing} a_i b_j \neq 1$ is the *conflict coefficient*. When the renormalization constant $1-\kappa = 0$, the assignments *a* and *b* are said to be *incompatible* and the result of their combination will be set to an algebraic zero, 0. The combination of a zero with any other assignment is also a zero. (Zero will be discussed further in section 4, on semigroups.)

Apart from mass, two other concepts play a key role in the DS theory: balance and plausibility. *Balance* (also variously termed, *belief*, *support*, or *lower probability*) of a set *A* is the sum of the masses assigned to all subsets of *A*., i.e.,

$$P_{*A} = \sum_{B \subseteq A} a_B \ .$$

Plausibility (also termed *upper probability*) of a set *A* is the sum of the masses of all sets having non-empty intersection with *A*, i.e.,

$$P_A^* = \sum_{B \cap A \neq \varnothing} a_B$$

Balance and plausibility, like mass, are mappings from the power set of Ω to the unit interval.

It is immediate from the definitions that. $P_{*A} \leq P_A^*$.

Most practitioners of DS focus on balance, but the plausibility on the singleton sets turns out to be related to Bayesian evidence accumulation (see section 5).

3. BAYESIAN EVIDENCE ACCUMULATION

The conditional probability of a measurement *B*, given an alternative ω_i is

$$P(B \mid \omega_i) = \frac{P(B \cap \omega_i)}{P(\omega_i)}, \quad i = 1, \ldots, n$$

Given a measurement *B*, the probability of each alternative ω_i is given by Bayes rule:

$$P(\omega_i \mid B) = \frac{P(B \mid \omega_i) P(\omega_i)}{P(B)}$$

Although the probabilities conditioned on different alternative do not generally sum to one, we can renormalize them:

$$b_i \equiv \frac{P(B \mid \omega_i)}{\sum_{i=1}^n P(B \mid \omega_i)} \ .$$

Then both $a_i \equiv P(\omega_i)$ and the b_i function as probability assignments that combine by component-wise multiplication and renormalization, when $\Sigma \equiv \sum_{i=1}^n b_i a_i \neq 0$, as:

$$c_i \equiv P(A_i \mid B) = \frac{b_i a_i}{\Sigma}, \quad i = 1, \ldots, n \ .$$

When the renormalization constant Σ = 0, the combination is set to an algebraic zero and assignments are said to be incompatible. The combination of zero with any other assignment is always zero. (Zero is discussed further in section 4.) This, for our purposes, is the Bayesian method of evidence combination.

4. SEMIGROUP PRIMER

The main algebraic structures considered in this paper are abelian semigroups, monoids and groups. We will review the most elementary properties of these structures. More advanced concepts are covered in the standard texts, including [10], [14], [15].

Suppose S is a non-empty set and $\circ$ is a binary operation on S. If $\circ$ is *associative*, i.e., $a \circ (b \circ c) = (a \circ b) \circ c$ for all $a, b, c \in S$, then S is a *semigroup*. If, additionally, $\circ$ is *commutative*, i.e., $a \circ b = b \circ a$ for all $a, b \in S$, then the semigroup S is an *abelian* semigroup.

A semigroup S that contains the *identity* element 1_S i.e., $1_S \circ a = a \circ 1_S = a$ for all $a \in S$, is a *monoid*.

A monoid S such that for every $a \in S$ there is an *inverse* element $a^{-1} \in S$, i.e., $a^{-1} \circ a = 1_S$ for all $a \in S$, is a *group*. A partial semigroup is a non-empty set with associative binary operation defined on *some* pairs of its elements. The algebra of partial operations, including partial semigroups, is described in detail in [19]. It is possible to develop an algebraic theory of Dempster-Shafer and Bayesian evidence accumulation without defining zeros by working with partial semigroups, but it is more cumbersome and we will not take that approach.

Suppose S is a semigroup and T is a nonempty subset of S. If T is a semigroup under the operation in S, then T is a subsemigroup of S. The same convention is used for monoids and groups.

A subset J of S is called an *ideal*, or an *absorbing set* of S if for all j in J and a in S, the combinations $j \circ a$ and $a \circ j$ are both in J. An ideal that consists of a single element is a zero, denoted 0, and has the usual zero-multiplication property, namely, for all $a \in S$, zero annihilates, i.e., $0 \circ a = a \circ 0 = 0$.

Of fundamental importance in theory and applications are mappings between sets that preserve algebraic structures.
One of these mappings is a semigroup (monoid, group) homomorphism.

Suppose $(S, \circ)$ and $(T, \text{å})$ are semigroups. A mapping $\phi : S \to T$ is a *semigroup homomorphism* if and only if

$$\phi(a \circ b) = \phi(a) \,\text{å}\, \phi(b) \quad \text{for all} \quad a, b \in S.$$

If $(S, \circ)$ and $(T, \text{å})$ are monoids with the identity elements 1_S and 1_T, respectively, the semigroup homomorphism $\phi : S \to T$ is a *monoid homomorphism* if and only if $\phi(1_S) = 1_T$. Any homomorphism partitions its domain into disjoint sets called equivalence classes.[11]

Group homomorphism is defined identically to semigroup homomorphism, since for groups conservation of identity elements follows from conservation of group operation.

5. HOMOMORPHIC AND CONVERGENT RELATIONSHIP BETWEEN DS AND BAYES

Suppose the non-zero mass assignments in a and b are made entirely to singleton sets. We can denote the mass that a assigns to alternative ω_i by a_i and the mass assigned by b to alternative ω_i by b_i. Then the DS combination of a and b is

$$a \circ b = \left(\frac{a_1 b_1}{\Sigma}, \dots, \frac{a_n b_n}{\Sigma} \right),$$

where the normalizing constant Σ is the sum of the numerators. (As in sections 2 and 3 the result will be 0 if $\Sigma = 0$.) This is the Bayesian combination of *a* and *b*, so DS combination becomes Bayesian combination in the special case in which mass is assigned only to singleton sets. We will return to this point shortly. First we show that Dempster Shafer combination is commutative and associative.

Commutativity: Set Suppose Dempster-Shafer assignments *a*, *b*, are *compatible*, that is, $a \circ b \neq 0$. Then $a \circ b = b \circ a$ since

$$a \circ b = \left\{ \frac{1}{\Sigma} \sum\nolimits_{A \cap B = C} a_A b_B \right\}_{C \in 2^\Omega} = \left\{ \frac{1}{\Sigma} \sum\nolimits_{A \cap B = C} b_B a_A \right\}_{C \in 2^\Omega} = b \circ a$$

by the commutativity of ordinary multiplication and the commutativity of set intersection. If $a \circ b = 0$ then *a* and b assign mass only to collections of sets that do not intersect each other. So $b \circ a = 0 = a \circ b$.

D-S Associativity: We show that $(a \circ b) \circ c = a \circ (b \circ c)$.

Proof: Suppose that for Dempster-Shafer assignments *a*, *b*, and *c* the combinations $(a \circ b) \circ c$ and $a \circ (b \circ c)$ are compatible. The mass assigned by $(a \circ b) \circ c$ to the set *D* is given

$$\begin{aligned} \left[(a \circ b) \circ c \right]_D &= \frac{1}{\Sigma'} \sum_{C \cap E = D} c_C \frac{1}{\Sigma} \sum_{A \cap B = E} a_A b_B \\ &= \frac{1}{\Sigma' \Sigma} \sum_{A \cap B \cap C = D} a_A b_B c_C \end{aligned}$$

The mass assigned the set *D* by $a \circ (b \circ c)$ is given

$$\begin{aligned} \left[a \circ (b \circ c) \right]_D &= \frac{1}{\Sigma'''} \sum_{E \cap A = D} a_A \frac{1}{\Sigma''} \sum_{A \cap B = E} c_C b_B \\ &= \frac{1}{\Sigma''' \Sigma''} \sum_{A \cap B \cap C = D} a_A b_B c_C \end{aligned}$$

Since the re-normalization constant must be equal to sum of the un-normalized terms $\sum_D \sum_{A \cap B \cap C = D} a_A b_B c_C$, it follows that $\Sigma' \Sigma = \sum_D \sum_{A \cap B \cap C = D} a_A b_B c_C = \Sigma''' \Sigma''$ and the assignments $\left[(a \circ b) \circ c \right]_\upsilon$ and $\left[a \circ (b \circ c) \right]_\upsilon$ are equal. If any of the combinations forms a zero then the collections of sets assigned non-zero mass do not intersect, so $(a \circ b) \circ c = 0 = a \circ (b \circ c)$. □

It follows that the set of Dempster-Shafer assignments over *n* alternatives is a commutative semigroup under Dempster-Shafer combination. *A fortieori*, the set of Bayesian assignments over *n* alternatives is also a commutative semigroup. Both sets have identity elements. The identity element for DS combination assigns all mass to the entire set of alternatives:

$$1_D = \begin{cases} 1 & A = \Omega \\ 0 & \mathrm{A} \neq \Omega \end{cases}$$

If $a = \{a_A\}_{A \in 2^\Omega}$ is a DS assignment, then the combination $1_D \circ a = a$, since $A \cap \Omega = A$ and $1 \cdot a_A = a_A$. The element 1_D is not a Bayesian assignment since it assigns mass to a non-singleton set.

The Bayesian identity assigns equal mass, $\frac{1}{n}$, to all *n* alternatives:

$$1_B = \left(\tfrac{1}{n},\ldots,\tfrac{1}{n}\right)$$

If $b = \left(b_1,\ldots,b_n\right)$ is a Bayesian assignment, then the combination

$$1_B \circ b = \left(\frac{b_1/n}{\Sigma} \quad ,\ldots, \quad \frac{b_n/n}{\Sigma}\right) = \left(b_1,\ldots,b_b\right) = b\,.$$

Since both the DS set and the Bayes set have identity elements they are both monoids.

The DS combination of the Bayesian identity with an arbitrary DS assignment *a* is a Bayesian element (assignment of mass only to singletons):

$$1_B \circ a = \left(\frac{\sum_{A:\omega_1 \in A} a_A}{\Sigma} \quad ,\ldots, \quad \frac{\sum_{A:\omega_n \in A} a_A}{\Sigma}\right).$$

as the intersection of any set *A* with a singleton set $\left\{\omega_i\right\}$ is either empty if $\omega_i \notin A$, or else is $\left\{\omega_i\right\}$ if $\omega_i \in A$. By inspection, the *n* numerators are the plausibilities of the *n* alternatives.

We now define a function ϕ from DS to Bayesian assignments by

$$\phi\left(a\right) = 1_B \circ a \;\;,$$

which, from the previous equation, consists of the plausibilities on the *n* alternatives, normalized to sum to one. The following result is then readily apparent:

Theorem: ϕ is a homomorphism of the DS set onto the Bayes set.
Proof: $\phi\left(a \circ b\right) = 1_B \circ a \circ b = 1_B \circ 1_B \circ a \circ b = 1_B \circ a \circ 1_B \circ b = \phi\left(a\right) \circ \phi\left(b\right)$. $\square$

In particular, the homomorphism maps the DS identity to the Bayes identity $\phi\left(1_D\right) = 1_B$, so the homomorphism is a monoid homomorphism.

This homomorphism between Dempster-Shafer and Bayesian representations, which has been overlooked until our investigations, has many interesting implications for evidence accumulation. Every DS assignment has a corresponding Bayesian assignment, and the combination of any two DS assignments results in a DS element in the pre-image of the Bayesian combination of the corresponding Bayesian assignments. In a sense, every DS combination is governed by its corresponding Bayesian combination, and the DS evidence accumulation process is governed by Bayesian evidence accumulation.

As with any homomorphism, ϕ partitions all DS assignments into sets (equivalence classes) according to their Bayesian images: For distinct elements *b* and *c* in the Bayes set, the DS sets $D_b = \left\{d \in \text{DS}{:}\phi\left(d\right) = b\right\}$ and $D_c = \left\{d \in \text{DS}{:}\phi\left(d\right) = c\right\}$ are disjoint. The structure of this partition for singleton Dempster-Shafer is explored in [4]. In [8] the implications of the homomorphism approach are explored for other evidence approaches, including un-normalized DS and Smets' Transferable Belief Model. Here we emphasize the implications of the homomorphism to assignments in applications, especially the *convergence to Bayes* that is implied.

The convergence of DS evidence accumulation to Bayesian evidence accumulation is immediate if any element in the accumulation process is Bayesian because the Bayes set is an ideal in the DS set, as follows. If g is Bayesian then $g = 1_B \circ g$ and

$$\begin{aligned} a \circ b \circ \ldots \circ g \circ \ldots \circ z &= a \circ b \circ \ldots \circ 1_B \circ g \circ \ldots \circ z \\ &= \phi(a \circ \ldots \circ z) \\ &= \phi(a) \circ \ldots \circ \phi(z). \end{aligned}$$

So in this case we can replace each DS assignment by its image under the homomorphism (its Bayesian *homomorph*) and do Bayesian combination, which is computationally more efficient.

Even if none of the assignments is Bayesian, there will be convergence toward a Bayesian result, as intersecting sets always produce sets with equal or fewer members than the smaller of the two sets, and as mass accumulates on the resulting smaller sets. The following Monte Carlo result illustrates the point. Based on a set of three alternatives with mass assigned to Ω randomly over the interval from 0.2 to 0.8, with the rest of the mass randomly assigned among the three alternatives, it shows the rapid disappearance in the composite of the mass assigned to "uncertainty" i.e., to Ω . .

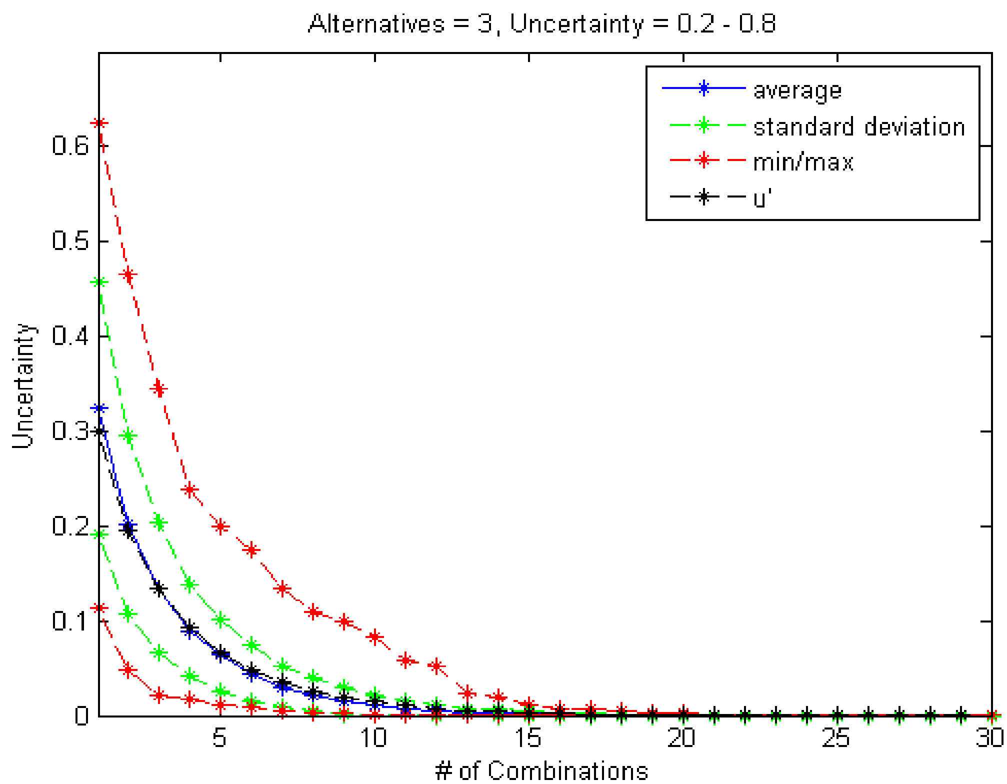

Fig. 1. This MonteCarlo result illustrates the rapid convergence of Dempster-Shafer evidence accumulation to a Bayesian result. Dempster-Shafer mass is assigned to Ω ("uncertainty") uniformly and randomly in the interval from 0.2 to 0.8, with the remaining mass randomly assigned among the three alternatives. The average value of uncertainty after the number of combinations given by the ordinate appears as the middle curve with a fit-curve beneath it. The curves immediately above and below the average represent the average plus or minus one standard deviation, and the outer-most curves represent the maximal and minimal levels of mass assigned to uncertainty seen across the Monte Carlo runs (400 runs). If lower values of uncertainty were allowed the convergence would be still more rapid.

For Dempster-Shafer evidence accumulation to provide a richer insight than Bayesian accumulation we must have three factors present: there must be no Bayesian assignments in the evidence stream, the number of combinations must be small enough that Bayesian convergence is not too closely approached, and yet the evidence must not be so uncertain that the process reaches no useful decisions. Because it may be difficult to achieve all these conditions, and because the Dempster-Shafer combination is potentially $2^n/n$ times more computationally more expensive than Bayesian combination, most practical applications will find Bayesian evidence accumulation more suitable.

6. CONCLUSIONS

Developments of this paper reveal an underlying link between computations taking place in Dempster-Shafer and Bayesian evidence accumulation. This link is made precise by the homomorphism theorem, which shows that Dempster-Shafer combinations are governed by Bayesian combinations. Since the set of Bayesian assignments is an ideal (absorbing set) in the set of Dempster-Shafer assignments, the combination of a DS assignment with a Bayesian assignment yields a Bayesian result This last property permits viewing the DS analysis as a monoid extension of the Bayes analysis. When all the evidence assignments are non-Bayesian, replacement of the DS computation with the corresponding Bayesian computation, results in loss of the uncertainty estimate, as homomorphism distributes uncertainty among all components of the composite mass. This likely will not be a significant drawback in situations when uncertainty is either very small or very large. In the former case the results of DS and Bayes evidence accumulations are immediately convergent. In the latter case (much less likely to occur as combination of evidence reduces uncertainty very quickly) it is unlikely that either approach will produce effective decisions. However, a loss of potentially useful information might possibly occur when uncertainty assumes an intermediate value. This loss needs to be evaluated, for a given application, and compared with the gains brought about by the replacement of the DS approach with the Bayes approach--reduction of computational complexity of the combining operation and simplification of the analysis--before the Dempster-Shafer can be decisively ruled out.

REFERENCES

[1] Basir O., Karray F. and Zhu H, “Connectionist-based Dempster-Shafer evidential reasoning for data fusion,” IEEE Trans. Neural Networks, Vol. 16, No. 6, pp 1513-1530, (2005).

[2] Blackman S. and Popoli R, Design and analysis of modern tracking systems, Artech House, Norwood, MA, (1999).

[3] Bloch I, “Information combination operators for data fusion: a comparative review with classification,” IEEE Trans. Syst. Man, Cybern., Vol. 26, pp 52-67, (1996).

[4] Brodzik A.K. and Enders R.H., “Semigroup structure of singleton Dempster-Shafer evidence accumulation,” submitted to IEEE Trans. IT. (2007).

[5] Clifford A.H. and Preston G.B., The algebraic theory of semigroups, American Mathematical Society, (1967)

[6] Delmotte F. and Smets P., “Target identification based on the transferable belief model interpretation of Dempster-Shafer model,” IEEE Trans. Syst. Man, Cybern., Vol. 34, No. 4, pp 457-471, (2004).

[7] Dempster A. P., Upper and lower probabilities induced by a multi-valued mapping, Ann. Math. Statist., Vol. 38, pp 325-329, (1967).

[8] Enders R. H., and Brodzik, A. K., Algebra of evidence accumulation, (2008).

[9] Gelman A., The boxer, the wrestler, and the coin flip: a paradox of robust Bayesian inference and belief functions, American Statistician, Vol. 60, No. 2, pp 146-150, (2006).

[10] Grillet P. A., Commutative semigroups, Kluwer, (2001).

[11] Hall Jr. M., Combinatorial theory, Wiley, (1998).

[12] Halpern, J.Y. Reasoning about Uncertainty, MIT Press, (2003).

[13] Henkind S.J. and Harrison M.C., An analysis of four uncertainty calculi, IEEE Trans. Syst. Man, Cybern., Vol. 18, No. 5, pp 700-714, (1988).

[14] Howie J. M., Fundamentals of semigroup theory, Oxford Univ. Press, (2003).

[15] Hungerford T. W., Algebra, Springer, (2003).

[16] Klir, G., Uncertainty and Information: Foundations of Generalized Information Theory, Wiley-Interscience, (2006).

[17] Laha A., Pal N.R. and Das J., “Land cover classification using fuzzy rules and aggregation of contextual information through evidence theory,” IEEE Trans. Geoscience and Remote Sensing, Vol. 44, No. 6, pp 1633-1641, (2006).
[18] Le Geharat-Mascle S., Bloch I. and Vidal-Madjar D., “Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing,” IEEE Trans. Geoscience and Remote Sensing, Vol. 35, No. 4, pp 1018-1031, (1997).
[19] Ljapin E. S. and Evseev A.E., The theory of partial algebraic operations, Springer, (1997).
[20] Lucas C. and Araabi B.N., “Generalization of the Dempster-Shafer theory: a fuzzy-valued measure,” IEEE Trans. Fuzzy Systems, Vol. 7, No. 3, pp 255-270, (1999).
[21] Fixsen D. and Mahler R.P.S., “The modified Dempster-Shafer approach to classification,” IEEE Trans. Syst. Man, Cybern., Vol. 27, No. 1, pp 96-104, (1997).
[22] Mahler R.P.S., Combining ambiguous evidence with respect to ambiguous a priori knowledge, I: Boolean logic,” IEEE Trans. Syst. Man, Cybern., Vol. 26, No. 1, pp 27-41, (1996).
[23] Pearl J., Causality: models, reasoning and inference, Cambridge Univ. Press, Princeton, (2000).
[24] Rey M., Tunaley J.K.E. and Sibbald T., “Use of the Dempster-Shafer algorithm for the detection of SAR ship wakes,” IEEE Trans. Geoscience and Remote Sensing, Vol. 31, No. 5, pp 1114-1118, (1993).
[25] Rota G.-C., “Theory of Möbius functions,” Z. Wahrscheinlichkeitstheorie und Gebiete, 2, pp 340-368, (1964).
[26] Shafer G., A mathematical theory of evidence, Princeton Univ. Press, Princeton, JN, (1976).
[27] Shi S., Hull M.E.C. and Bellm D.A., “Updating belief in Gordon-Shartliff's hierarchical structure,” in Proc. 2nd Austral. New Zealand Conf. Intelligent Information Systems, pp 402-406, (1994).
[28] Smets, P. “The Combination of Evidence in the Transferable Belief Model,” IEEE Trans on Pattern Analysis and Machine Intelligence. Vol. 11. No. *5*. May, 447-458 1990.
[29] Wu Y., Yang J., Liu K., and Liu L., “On the evidence inference theory,” Information Science, Vol 89, pp 245-260, (1996).
[30] Yager R.R., Fedrizzi, M. and Kacprzyk, J. (eds), Advances in the Dempster-Shafer theory of evidence, Wiley, New York, (1994).
[31] Zadeh L., “On the validity of Dempster's rule of combination of evidence”, Memo M79/24, Univ. of California, Berkeley, USA, (1979).