An attempt to implement Compositional Models in Dempster-Shafer theory of Evidence

Václav Kratochvíl
Inst. of Information Theory and Automation, Academy of Science, and
Faculty of Management of University of Economics
Czech Republic
velorex@utia.cas.cz

Abstract

It has been published recently that some of the ideas for representation of multidimensional distributions in probability theory can be transferred into Dempster-Shafer theory of Evidence [7], [8]. Namely, they showed that multidimensional basic assignments can be rather efficiently represented in a form of so-called compositional models. These models are based on the iterative application of the operator of composition, whose definition for basic assignments has been introduced in [5]. It appears that a software tool supporting computations within compositional model is necessary for additional theoretical research in this framework. In this paper we will familiarize the reader with our first attempts and basic problems of the implementation itself.

1 Introduction

Plenty of applications of Artificial intelligence in the field of quantitative reasoning and decision under uncertainty is dominated by probabilistic models like Bayesian networks and their variants. It these models a multidimensional probability distribution is used to represent the real world problem and capture and represent uncertainty. We can distinguish two types of uncertainty. The first is variability that arises from environmental stochasticity, inhomogenity of materials, fluctuations in time, variation in space, or heterogenity or other differences among components or individuals. This variability is sometimes called *aleatory* uncertainty to emphasize its relation to the randomness in gambling and games of chance. The second kind of uncertainty is the incertitude that comes from scientific ignorance, measurement uncertainty, inobservability, censoring, or other lack of knowledge. This is sometimes called *epistemic* uncertainty.

For situations in which the uncertainty about quantities is purely aleatory, probability theory is usually preferred and it is sufficient for this purpose. When the gasp in our knowledge involve both aleatory and epistemic uncertainty, several competing approaches have been suggested: The common practice is to use probability theory as well. As another example we would like to mention *probability boxes* [21] and especially *Dempster-Shafer theory of Evidence* (D-S) [1] [15] which we will deal within this paper.

There is one problem when using probability framework to handle uncertainty. Assume that we have no information concerning behavior of a variable. Using probability theory, one might assume equal priors and distribute the weight of evidence equally among all possible states of the variable. But, as Shafer pointed out, here one will fail to distinguish between uncertainty (or lack of knowledge), and equal certainty. And it is this kind of uncertainty that can be easily captured in the framework of D-S theory.

In this paper we will deal with D-S theory, especially we will work with the notion of *Compositional models*. Compositional models were originally introduced in the probability framework. The intention was to create an algebraic alternative to the well-known Markov graphical models like Bayesian networks. The important advantage of compositional models is that they can be generalized in the framework of possibility theory as well as D-S theory by introducing a special operator of composition [8]. The recent research [7] [8] revealed the necessity of an software tool supporting compositional models in D-S theory.

The intention of this paper is nothing more than to summarize our initial problems when attempting to implement such a software tool. Here we describe our first steps, ideas and preliminary solutions.

2 Notation

For an index set $N = \{1, 2, \ldots, N\}$ let $\{X_i\}_{i\in N}$ be a finite set of finite valued variables, each X_i having its values in $\mathbf{X}_i$. In this paper we deal with multidimensional frame of discernment $\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n$, and its subframes (for $K \subseteq N$) $\mathbf{X}_K = \times_{i\in K}\mathbf{X}_i$. The symbol X_K will denote a group of variables $\{X_i\}_{i\in K}$. A projection of $x = (x_1, x_2, \ldots, x_n) \in \mathbf{X}_N$ into $\mathbf{X}_K$ will be denoted $x^{\downarrow K}$, i.e. for $K = \{i_1, i_2, \ldots, i_k\}$

$$x^{\downarrow K} = (x_{i_1}, x_{i_2}, \ldots, x_{i_k}) \in \mathbf{X}_K.$$

Analogously, for $M \subset K \subseteq N$ and $A \subset \mathbf{X}_K$, $A^{\downarrow M}$ will denote a *projection* of A into $\mathbf{X}_M$:

$$A^{\downarrow M} = \{y \in \mathbf{X}_M | \exists x \in A : y = x^{\downarrow M}\}.$$

In addition to the projection, in this text we will need also an opposite operation, which will be called a *join*[1]. By a join of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ $(K, L \subseteq N)$ we will understand a set

$$A \bowtie B = \{x \in \mathbf{X}_{K\cup L} : x^{\downarrow K} \in A \ \ \& \ \ x^{\downarrow L} \in B\}.$$

[1]This term and notation are taken form the theory of relational databases

Let use note that if K and L are disjoint, then $A \bowtie B = A \times B$, and if $K = L$ $A \bowtie B = A \cap B$.

The symbol $\mathcal{P}(\mathbf{X}_K)$ will denote the powerset of $\mathbf{X}_K$, i.e. the set of all subsets of $\mathbf{X}_K$.

2.1 Basic assignments

The role played by a probability distribution in probability theory is replaced by that of a set function in D-S theory: belief function, plausibility function, commonality function, or basic (*probability* or *belief*) assignment. Knowing one of them, one can derive the remaining three. In this paper we will use almost exclusively basic assignments.

If $m(A) > 0$, then A is said to be a *focal element* of m. The set of focal elements will be denoted by S. A *basic assignment* (bpa) m in $\mathbf{X}_K$ ($K \subseteq N$) is a function

$$m : \mathcal{P}(\mathbf{X}_K) \rightarrow [0, 1],$$

for which

$$\sum_{\emptyset \neq A \subseteq \mathbf{X}_K} m(A) = 1.$$

The quantity $m(A)$ is a measure of that portion of the total belief committed exactly to A, where A is an element of $\mathcal{P}(\mathbf{X}_K)$ and the total belief is 1. The portion of belief cannot be further subdivided among the subsets of A and does not include portions of belief committed ti subsets of A. Since belief in a subset certainly entails belief in subsets, containing that subset, it would be useful to define a function that computes a total amount of belief in A. Such a function is called *belief function*.

On the contrary, *plausible function* characterizes the degree in which a proposal A is plausible based on available evidence B expressed by each basic assignment that contributes to realization of A. *Commonality function* doesnt have a simple interpretation but it allows a simple statement of Dempsters combination rule [1].

2.2 Operator of composition

Compositional models theory has been introduced in the framework of probability theory [6] as an algebraic alternative to well known and widely used Bayesian networks for efficient representations of multidimensional measures more than twelve years ago. Compositional models are based on recurrent application of an operator of composition. Later, the operator of composition was introduced also within the framework of D-S theory in [5]:

Definition 2.1. *For two arbitrary bpa m_1 on $\boldsymbol{X}_K$ and m_2 on $\boldsymbol{X}_L$ $(K, L \neq \emptyset)$, a composition $m_1 \triangleright m_2$ is defined for each $C \subseteq \boldsymbol{X}_{K \cup L}$ by one of the following expressions:*

a) if $m_2^{\downarrow K\cap L}(C^{\downarrow K\cap L} > 0$ *and* $C = C^{\downarrow K} \bowtie C^{\downarrow L}$ *then*

$$(m_1 \triangleright m_2)(C) = \frac{m_1(C^{\downarrow K}) \cdot m_2(C^{\downarrow L})}{m_2^{\downarrow K\cap L}(C^{\downarrow K\cap L})}$$

b) if $m_2^{\downarrow K\cap L}(C^{\downarrow K\cap L}) = 0$ *and* $C = C^{\downarrow K} \times \boldsymbol{X}_{L\setminus K}$ *then*

$$(m_1 \triangleright m_2)(C) = m_1(C^{\downarrow K});$$

c) in all other cases $(m_1 \triangleright m_2)(C) = 0$.

In the D-S theory, there exists several way how to combine different sources of evidence and the above defined operator of composition seems to be one of them. But this is not the case. The classical way is represented by Dempster's combination rule [1]. A criticism of this rule appeared later caused by its behavior when combining two conflicting evidences and several additional combination rules were designed. Recall for example *Yager's rule* [23], *Inagakis rule* [4], *Zhangs rule* [25], or *Dubois and Prades Disjunctive Consensus* [2]. However, the intention of the operator of composition is not to be another combination rule and combine different sources of evidence. Its intention is completely different.

Despite the success of D-S theory of evidence as a well founded and general model of human reasoning under uncertainty, belief functions are rarely used in concrete applications. One of the most significant arguments raised against using belief functions in practice is their relatively high computational complexity, especially in comparison with methods based on classical probability theory. E.g. combining evidence using relatively simple Dempster's rule of combination is known to be #P-complete in the number of evidential sources. Recall that bpa (as well as belief function, plausibility function, and commonality function) is a set function. We work with the powerset of possible events and the number of sets that can be focal elements of a bpa can be superexponential within the number of involved variables.

To overcome these computational limitations, different approximation methods have been proposed. Previous work can be divided into two categories [3]. The first category consists of *Monte-Carlo* techniques [22]. The idea is to estimate exact values of belief and plausibility by ratios of different outcomes relative to randomly generated samples. The second category consist of simplification procedures. They are motivated by the fact that the most algorithms involving belief functions have a complexity polynomial in the number of focal elements. The underlying idea is therefore to restrict in different ways the number of focal elements. A simple method is called *Bayesian approximation* [20], where only singletons are allowed - which corresponds to the restriction on probability distributions only. Other methods like *k-l-x approximation* [19], *summarization* [10], and others try to reduce the number of focal elements by taking the first k-most important assignments. The sum of the omitted assignments is then redistributed in different ways depending on the respective method.

The idea of operator of composition goes in a different way: Practically all methods for efficient computations with multidimensional models take advantage of the fact that the model in question in a way factorizes. It means that it is possible to decompose the model into its low-dimensional parts, each of which can be defined with a reasonable number of parameters. This is the basic idea for computation with probabilistic Graphical Markov Models. Such a factorization not only decreases the storage requirements for representation of a multidimensional distribution but it usually also induces possibility to employ efficient computational procedures.

Since we need efficient methods for representation of probabilistic distributions, which require exponential number of parameters, the more we need of efficient methods for representation of an evidence, which cannot be represented by a point function. For such a representation we need a set function, and thus its space requirements are superexponential.

2.2.1 Compositional models

The factorizable evidence will be then represented in a form of the so-called *compositional model*. Assume a system of low-dimensional basic assignments $m_1, m_2, \ldots, m_n$ defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \ldots, \mathbf{X}_{K_n}$, respectively. Composing them together by multiple application of operator of composition, one get multidimensional basic assignment on $\mathbf{X}_{K_1 \cup K_2 \cup \ldots \cup K_n}$. Note that the operator of composition is neither commutative, nor associative. By "composing them together" we understand that the operator of composition is performed successfully from left to right and $m_1 \triangleright_2 \triangleright \ldots \triangleright m_n = (\ldots((m_1 \triangleright m_2) \triangleright m_3) \triangleright \ldots) \triangleright m_n$.

2.2.2 New Concept of Conditional Independence

For belief functions, two type of factorization were designed in the literature. One is based on various combination rules mentioned above, the other use an operator of composition [5]. It has been shown in [7] that approach concerning Dempster's rule and the operator of compositions are equivalent each other in case of unconditional factorization.

The idea of factorization is closely related to the notion of (un)conditional independence in probabilistic modeling. However, as pointed out by Studený, the original definition of conditional independence (published in [24]) was not consistent with marginalization. That is why a new definition of conditional independence was introduced in D-S theory in [7]:

Definition 2.2. *Let m be a basic assignment on $\boldsymbol{X}_N$ and $K, L, M \subset N$ be disjoint, $K, L \neq \emptyset$. We say that groups of variables X_K and X_L are conditionally independent given X_M with respect to m (and denote it by $K \perp\!\!\!\perp L | M[m]$), if the equality*

$$m^{\downarrow K \cup L \cup M}(A) \cdot m^{\downarrow M}(A^{\downarrow M}) = m^{\downarrow K \cup M}(A^{\downarrow K \cup M}) \cdot m^{\downarrow K \cup L}(A^{\downarrow K \cup L})$$

holds for any $A \subseteq \boldsymbol{X}_{K \cup L \cup M}$ such that $A = A^{\downarrow K \cup M} \bowtie A^{\downarrow L \cup M}$, and $m(A) = 0$

otherwise. If $M = \emptyset$ then we say that groups of variables X_K and X_L are independent with respect to m (in symbol $K \perp\!\!\!\perp L[m]$).

Above that, it has been shown in [8] that the above defined conditional independence satisfies semigraphoid properties and that there is a link between operator of composition and conditional independence:

Theorem 2.3. *Let m be a joint basic assignment on $\boldsymbol{X}_M$, $K, L \subseteq M$. Then $(K \setminus L) \perp\!\!\!\perp (L \setminus K) | (K \cap L)[m]$ iff $m^{\downarrow K \cup L}(A) = (m^{\downarrow K} \triangleright m^{\downarrow L})(A)$ for any $A \subseteq \boldsymbol{X}_{K \cup L}$.*

This theorem justifies the usage of the operator of composition when factorizing an evidence.

3 Implementation

To evaluate various hypotheses and support accelerate further theoretical research, it is necessary to create an experimental tool for calculations with compositional models in the framework of D-S theory. In this section we would like to describe several problems when attempting to implement such a tool. The tool itself is developed as an extension package for R-Project[2] and it is available at http://dar1.utia.cas.cz/mudim altogether with another tool supporting compositional models in the probability framework.

During our survey of existing implementation of D-S theory we found out that there is no successful universal tool supporting theoretical research. The majority of existing implementations is usually single purpose and base on restricted assumptions. One can find not very up-to-date, but exhausted overview of applications of D-S theory in [14].

The key problem of the implementation is the representation of belief structures. Restrict ourselves to finite sets. For a finite set P of possible outcomes ($P \subseteq \mathbf{X}_N$) with cardinality $|P|$, there are at most $2^{|P|}$ unique basic probability assignments. We assume that rarely is a full set of $2^{|P|}$ unique bpa used in practice. It corresponds to the limited sources of information. Within the research literature there exists four common subclasses of bpa for finite sets [9]:

1. The trivial case of total ignorance where $m(P) = 1$ and $m(A) = 0$ iff $A \neq P$. This is highlighted as a more accurate representation of total ignorance when compare to traditional probability theory, which must apply Laplace's principle of indifference in these circumstances.

2. Every assignment is made to a singleton of the set P. This corresponds to a traditional probability measure on the set P.

[2] R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To download R, please visit http://www.r-project.org/

3. Every assignment is made to a nested set. In other words, for every two sets A and B such that $m(A) > 0$ and $m(B) > 0$, then $A \subset B$ or $B \subset A$. This arrangement of is known as possibility theory.

4. Every assignment can be made to an arbitrary set.

In our case we focus on the most general case - the last one. However, we step aside the fact that involved variables can be either contiguous (discrete or continuous), or categorical. Each of these data types requires a special representation in a computer memory. In a survey of real-world application of D-S theory to infinite sets [16] it has been published that the contiguous frame of elements assigns basic probability statements to the closed intervals $[x_{i_a}, x_{i_b}]$ as a rule. Thus, in case of contiguous variables, we will store the interval boundaries. See the overview of various data types in Table 1.

data type	*example*	*implementation*
continuous finite data	age - integer	interval boundaries
cont. infinite data	sensor output - real number	interval boundaries
categorical data	sex {male, female}	set of elements

Table 1: Implementation of various variable types

3.0.3 Problems

Let $A \subseteq \mathbf{X}_K$, $B \subseteq \mathbf{X}_L$, bpa m on $\mathbf{X}_K$, and S set of focal elements of m. The key problem of the implementation itself is the fact that we have to store every focal element $A \in S$ of m and pair it with the value $m(A)$. This does not sound very difficult unless we realize that S is a set of sets of vectors and that every set of vectors $A \in S$ is of various cardinality. The implementation of data structure will will have an enormous impact on overall system performance.

The most basic operation which will be instantly used is the checking whether a set A is a focal element, i.e. whether $A \in S$. It is logical to assume that if the data structure will be optimized with respect to this operation, then the system performance allows to add additional functionalities like operator of composition etc.

There are multiple ways of implementing set (and map) functionality, that is:

- ordered (e.g. tree-based) approaches, and
- unordered (e.g. hash-based) approaches

Here we propose the unordered (hash-based) approach, which naturally builds on top of the value-indexed array technique. The problem here is that we have set of sets of vectors, which significantly complicates the implementation of respective hash function.

However, in this attempt, we simply store the evidence in multidimensional arrays (tables) of vectors and we implement the search of a set in a set of sets simply as a full table scan - i.e. the algorithm gradually passes through all elements of S and compares them with set A. Such a comparison is described in Algorithm 1. A, B are two sets of possible outputs. Note that we employ the definition $A = B \Leftrightarrow A \subseteq B \& B \subseteq A$. Then, in case of checking e.g. $A \subseteq B$ (Algorithm 2) we simply check whether $\exists b \in B$ such that $a = b$ for all $a \in A$. In the worst case scenario, the complexity is $2 \cdot |A| \cdot |B|$ of vector comparisons. The improvement of this will have an enormous impact on the efficiency of the tool. Our idea is either implement a specific hash function, or to use embedded relational database [17] with optimized index-based search algorithms.

Algorithm 1 MySetEqual($A, B \subseteq \mathbf{X}_K$): boolean

if MySubset(A,B) **and** MySubset(B,A) **then**
    return TRUE;
else
    return FALSE;
end if

Algorithm 2 MySubset($A, SubA \subseteq \mathbf{X}_K$): boolean

found: flag if the corresponding element is found in the other set
for $i = 1$ to $|SubA|$ **do**
    found=FALSE;
    for $j = 1$ to $|A|$ **do**
        if $SubA[i] == A[j]$ **then**
            found=TRUE; $\{SubA[i] \in A\}$
            break; {additional search is useless}
        end if
    end for
    if not found **then**
        return FALSE; $\{SubA[i] \notin A \Rightarrow SubA \not\subseteq A\}$
    end if
end for
return TRUE;

The other operations that have to be considered when designing a data structure are:

- marginalization $A^{\downarrow M} = \{y \in \mathbf{X}_M | \exists x \in A : y^{\downarrow M} = x\}$
- join operation $A \bowtie B = \{x \in \mathbf{X}_{K \cup L} : x^{\downarrow K} \in A \ \ \& \ \ x^{\downarrow L} \in B\}$

Using above defined functions one can implement the operation of composition specified in Definition 2.1. See Algorithm 3 for the pseudo-code of the implementation. Here two auxiliary boolean flags are employed - *found* and *marginalComputed*. The first one decides between cases a) and b) of Definition 2.1. The second one highlights whether the respective marginal $m_2^{\downarrow K \cap L}(C^{\downarrow K \cap L})$ from Definition 2.1 has been already computed or not.

The careful reader notices that the loop on lines 18-22 of the previous algorithm may be performed for the same set $C^{\downarrow K \cap L}$ several times. This could be easily improved. Let us define an auxiliary vector $m_2 marginal$ to store computed marginal of m_2 and index it in the same way as m_2. Then it is enough to use the value $m_2 marginal[k]$ on line 20 if $k < l$ and break respective cycle (lines 18-22).

Conclusion

Recently, it has been published that some of the ideas for representation of multidimensional distributions in probability theory can be transferred into Dempster-Shafer theory of Evidence [7], [8]. Namely, they showed that multidimensional basic assignments can be rather efficiently represented in a form of so-called compositional models. However, only an application of the theory can show which parts still need to be improved. Our goal is to develop not only an interesting theory but also an efficient tool based on these theoretical results. In other words, we intend to create a software tool which could be used for experiments and additional theoretical research.

In this paper we have described our problems when implementing compositional models in the framework of Dempster-Shafer theory of Evidence. The tool is implemented as an extension package for R-Project and one can find it, altogether with another tool supporting compositional models in probability framework, at http://dar1.utia.cas.cz/mudim. In this paper we described our first steps and basic problems which we faced during implementation.

The paper contains just a preliminary ideas and gives answers only to very simple questions. So there are many more that remain to be answered. For example:

- Does it exist an efficient representation of sets of vectors?
- How does an effective hash function for a set of sets of vectors look like?
- Can be an embedded SQL database used for representation of focal elements?
- Let $\mathbf{X}_2 = \{a_2, \bar{a}_2\}$. Is it reasonable to combine two elements (a_1, a_2), $(a_1, \bar{a}_2)$ into $(a_1, \mathbf{X}_2)$ and store the information in this way?

Acknowledgement

This work was supported by institutional research support of Faculty of Management, University of Economics, Prague.

Algorithm 3 $\triangleright$ operation of composition: $m_3 = m_1 \triangleright m_2$

```
input $S_1$: set of focal elements of $m_1, S_1 \subseteq \mathcal{P}(\mathbf{X}_K)$
input $m_1$: set of basic probability assignments $m_1[i] = m_1(S_1[i])$
input $S_2$: set of focal elements of $m_2, S_2 \subseteq \mathcal{P}(\mathbf{X}_L)$
input $m_2$: set of basic probability assignments $m_2[i] = m_2(S_2[i])$
output $S_3$: set of focal elements of $m_3 = m_1 \triangleright m_2, S_3 \subseteq \mathcal{P}(\mathbf{X}_{K \cup L})$
output $m_3$: set of basic probability assignments $m_3 = m_1 \triangleright m_2$
$l = 1$;
$S_3 = \emptyset$;
$m_3 = \emptyset$;
for $i = 1$ to $|S_1|$ do
  marginalComputed = FALSE;
  found = FALSE;
  marginalValue = 0;
  for $j = 1$ to $|S_2|$ do
    if MySetEqual($(S_1[i])^{\downarrow K \cap L}, (S_2[i])^{\downarrow K \cap L}$) then
      found = TRUE; {i.e. $m_2((S_2[i])^{\downarrow K \cap L}) > 0$}
      if not marginalComputed then
        for $k = 1$ to $|S_2|$ do
          if MySetEqual($S_2[k], S_2[j]$) then
            marginalValue = marginalValue + $m_2[k]$;
          end if
        end for
        marginalComputed = TRUE;
      end if
      $S_3[l] = S_1[i] \bowtie S_2[j]$; {case $a$ of the Definition 2.1}
      $m_3[l] = (m_1[i] \cdot m_2[j])$/marginalValue;
      $l = l + 1$;
    end if
  end for
  if not found then
    $S_3[l] = S_1[i] \times \mathbf{X}_{L \setminus K}$; {case $b$ of the Definition 2.1}
    $m_3[l] = m_1[i]$;
    $l = l + 1$;
  end if
end for
return $m_3, S_3$;
```

References

[1] A.P. Dempster: *Upper and lower probabilities induced by a multi-valued mapping.* Annals of Mathematical Statistics 38, 325-339 (1967)

[2] D. Dubois and H. Prade: *On the combination of evidence in various mathematical frameworks.* Reliability Data Collection and Analysis. J. Flamm and T. Luisi. Brussels, ECSC, EEC, EAFC: 213-241 (1992)

[3] R. Haenni and N. Lehmann: *Resource bounded and anytime approximation of belief function computations*, International Journal of Approximate Reasoning (2002)

[4] T. Inagaki: *Interdependence between Safety-Control Policy and Multiple-Sensor Schemes Via Dempster-Shafer Theory.* IEEE Transactions on Reliability 40(2): 182-188 (1991)

[5] R. Jiroušek, J. Vejnarová, M. Daniel: *Compositional models for belief functions* in Proc. of 5th Int. Symposium on Imprecise Probability: Theories and Applications, Eds: De Cooman G., Vejnarová J., Zaffalon M., Praha(2007)

[6] R. Jiroušek: *Foundations of compositional model theory*, International Journal of General Systems - Volume 40, Issue 6, (2011), pp. 623-678.

[7] R. Jiroušek and P.P. Shenoy: *A Note on Factorization of Belief Functions*, Proceedings of 14th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty CJS 2011 , Eds: Barták Roman (2011)

[8] R. Jiroušek, J. Vejnarová: *Compositional models and conditional independence in evidence theory*, International Journal of Approximate Reasoning vol.52, 3 (2011), p. 316-334 (2011)

[9] I. Kramosil: *Probabilistic Analysis of Belief Functions.* Kluwer Academic/Plenum Publishers, New York, (2001)

[10] H. Lowrance, T. Garvey, T. Strat: *A framework for evidential-reasoning systems* T. Kehler, S. Rosenschein (Eds.), Proceedings of the 5th National Conference on Artificial Intelligence, vol. 2Morgan Kaufmann, CA (1986), pp. 896903

[11] J. Pearl: *Probabilistic Reasoning in Intelligent systems: Networks of Plausible Inference*, Margan Kaufmann, San Mateo, CA (1988)

[12] J. Pearl: *Reasoning with Belief Functions: An Analysis of Compatibility.* The International Journal of Approximate Reasoning 4 (5/6): 363389 (1990)

[13] R Development Core Team: *R - A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org (2008)

[14] K. Sentz and S. Ferson: *Combination of Evidence in DempsterShafer Theory*, Sandia National Laboratories SAND 2002-0835 (2002)

[15] G. Shafer G: *A Mathematical Theory of Evidence*, Princeton University Press (1976)

[16] M. Spiegel: *A proposal for computing with imprecise probabilities: A framework for multiple representations of uncertainty in simulation software* A Disertation Proposal, University of Virginia (2007)

[17] *SQL.* In Wikipedia. Retrieved July 12, 2012, from http://en.wikipedia.org/wiki/SQL

[18] M. Studený: *Probabilistic Conditional Independence Structures*, Springer, London, (series Information Science and Statistics), (2005).

[19] B. Tessem: *Approximations for efficient computation in the theory of evidence.* Artificial Intelligence, 61 (2) (1993), pp. 315329

[20] F. Voorbraak: A *computationally efficient approximation of Dempster-Shafer theory.* International Journal of Man-Machine Studies, 30 (5) (1989), pp. 525536

[21] R.C. Williamson and T. Downs. *Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds.* International Journal of Approximate Reasoning 4: 89158 (1990)

[22] N. Wilson, S. Moral: *Fast Markov chain algorithms for calculating DempsterShafer belief*, in: ECAI96; 12th European Conference on Artificial Intelligence, Wiley, pp. 672678 (1996)

[23] R. Yager: *On the Dempster-Shafer Framework and New Combination Rules.* Information Sciences 41: 93-137 (1987)

[24] B.B. Yaghlane, Ph. Smets, K. Mellouli: *Belief functions independence: II. The conditional case*, International Journal of Approximate Reasoning 31 (2002) 3175.

[25] L. Zhang: *Representation, independence, and combination of evidence in the Dempster-Shafer theory.* Advances in the Dempster-Shafer Theory of Evidence. R. R. Yager, J. Kacprzyk and M. Fedrizzi. New York, John Wiley & Sons, Inc.: 51-69 (1994)