International Journal of Approximate Reasoning 56 (2015) 122–136

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

# The choice of generalized Dempster–Shafer rules for aggregating belief functions [☆]

Andrey Bronevich [a,*], Igor Rozenberg [b]

[a] National Research University Higher School of Economics, Pokrovskiy Boulevard, 11, 109028 Moscow, Russia
[b] JSC "Research, Development and Planning Institute for Railway Information Technology, Automation and Telecommunication", Orlikov per. 5, building 1, 107996 Moscow, Russia

ARTICLE INFO

*Article history:*
Received 19 May 2014
Received in revised form 22 September 2014
Accepted 14 October 2014
Available online 22 October 2014

*Keywords:*
Generalized Dempster–Shafer rules
Belief functions
Imprecision indices
Uncertainty in information

ABSTRACT

In the paper we investigate the criteria of choosing generalized Dempster–Shafer rules for aggregating sources whose information is represented by belief functions. The approach is based on measuring various types of uncertainty in information and we use for this purpose in particular linear imprecision indices. Some results concerning properties of such rules are also presented.



## 1. Introduction

In the literature there are many generalizations of Dempster–Shafer (D–S) rule, see for instance [20]. Many of them [23,14,18] were the answer to the critique of this combination rule by Zadeh [24]. In this paper we support the idea of Smets [22], in which the whole family of rules of combination is divided on various types like conjunctive and disjunctive rules, and the optimal rule should be chosen according to each application. Many critiques [18,19] also concern the case when sources of information are conflicting[1] (or contradictory) and the measure of contradiction of two sources of information based on D–S rule is not adequate. There are also works, see for example [1], where you can find argumentation that the classical D–S rule is not justified in probability theory.

In this paper we investigate the generalized Dempster–Shafer (GD–S) rules that were firstly introduced by Dubois and Yager [15], where each GD–S rule is defined as follows. In the D–S theory each source of information can be described by a random set. If we assume that two sources of information are independent, then we get the D–S rule by taking the intersection of these two sets. In GD–S rules the joint probability distribution of random sets is not known and the choice of the rule can be based on the least commitment principle [10], the principle of the minimal contradiction between sources of information [4] and others [11,12,5]. Some authors try to define combination rules for dependent sources of information to preserve the associativity. In [8] Denoeux proposes conjunctive combination of belief functions for dependent sources

☆ This study (Research Grant No. 14-01-0015) was supported by The National Research University Higher School of Economics Academic Fund Program in 2014/2015.
* Corresponding author.
[1] In the rest of the paper we choose the term "contradiction" instead of "conflict" because we use this term for uncertainty described by probability theory.

http://dx.doi.org/10.1016/j.ijar.2014.10.002
0888-613X/

of information. This rule uses the special representation of bodies of evidences, and guarantees the associativity, but this rule can give sometimes strange results (see for example Proposition 7 in [8]). The associativity is also guaranteed for combination rules proposed by Cattaneo [5]. Notice that these rules are not GD–S rules which are the subject of our article.

In the paper we thoroughly investigate and generalize well-known approaches of choosing GD–S rules and find their probabilistic interpretation. We come to the conclusion that any GD–S rule can be conceived as an approximation from above of two belief functions by using the partial order on belief functions usually called specialization. Using this order we propose several approaches to find such approximations that generalize well-known ones. The paper has the following structure. We introduce first the basic notions of D–S theory, and how the D–S rule of combination can be generalized. In Section 3 we analyze what kind of uncertainty are described by belief functions and remind the justified functionals for measuring uncertainty. Based on this analysis, one can make a conclusion that non-normalized belief functions can describe three types of uncertainty: conflict, non-specificity and contradiction. The last type of uncertainty is due to a strictly positive mass of the empty set. In Section 4 we introduce imprecision indices for measuring non-specificity which are later used for defining optimization problems for choosing GD–S rules in Section 6. In Section 5 we remind about the specialization order on belief functions and give also some of its properties. In Section 6 we prove Theorem 2 that shows how an optimal choice of GD–S rule can be formulated with the help of this order. Section 7 is devoted to analyzing properties of measuring contradiction by the classical D–S rule and by GD–S rules. One can come to the conclusion that measuring contradiction between sources of information based on GD–S rules is more justified than by using the classical D–S rule because it has clearer probabilistic interpretation. In Section 8 we give properties of optimal GD–S rules or rather of Pareto optimal GD–S rules w.r.t. the order of specialization. Because the solution of the optimization problem can be not unique, the result of applying optimal GD–S rules can be conceived as a subset of belief measures. This implies that it is necessary to investigate other possible ways of evaluating GD–S rules quality. For example, according to Theorem 2, an optimal GD–S rule should give the best approximation from above of the maximum of combining belief functions. This idea is exploited in Section 9, where we show that the choice of an optimal GD–S rule is equivalent to the choice of GD–S rule based on the special linear imprecision index if we use some metric on the set of belief functions. Other idea is used in Section 10, where we try to analyze the contradiction of the chosen Pareto optimal rule to other possible choices.

## 2. Evidence theory and generalized Dempster–Shafer rules

Let $X$ be a finite set and let $2^X$ be the powerset of $X$. One can say that the *body of evidence* is given on $2^X$ if a non-negative set function $m: 2^X \to [0,1]$ is defined with $\sum_{A\in 2^X} m(A) = 1$. Through the body of evidence the following functions are also introduced $Bel(B) = \sum_{A\subseteq B} m(A)$, $Pl(B) = \sum_{A\cap B\neq\emptyset} m(A)$, which are called *belief function* and *plausibility function* respectively. The function $m$ is usually called the *basic belief assignment* (bba). We accept here the *transferable belief model* [22], where $m(\emptyset) = Bel(\emptyset)$ shows the degree of contradiction in information. If the contradiction in information is equal to zero, then the corresponding belief function is called *normalized*. In the next we will use the following notations and definitions.

(1) A set $A \in 2^X$ is called *focal* if $m(A) > 0$.
(2) A belief function is called *categorical* if the body of evidence contains only one focal element $B \in 2^X$. This belief function is denoted by $\eta_{\langle B\rangle}$ and obviously $\eta_{\langle B\rangle}(A) = \begin{cases} 1, & B\subseteq A, \\ 0, & \text{otherwise.} \end{cases}$ Using categorical belief functions, we can express any belief function by the formula $Bel = \sum_{B\in 2^X} m(B)\eta_{\langle B\rangle}$.
(3) A belief measure is called a *probability measure* if $m(A) = 0$ for $|A| > 1$.[2]
(4) We denote correspondingly by $\bar{M}_{bel}$ and $\bar{M}_{pr}$ the families of all belief functions and probability measures on $2^X$, and if these families are normalized we denote them by $M_{bel}$ and $M_{pr}$.
(5) For any set functions $\mu_1$, $\mu_2$ on $2^X$ we write $\mu_1 \leqslant \mu_2$ if $\mu_1(A) \leqslant \mu_2(A)$ for all $A \in 2^X$.

Let us consider the probabilistic interpretation of the transferable belief model based on random sets. A random set $\xi$ is a random variable taking its values in $2^X$. Any such random variable can be defined by probabilities $P(\xi = A)$ being identified with values $m(A)$ in the theory of evidence. Given two random sets $\xi_1$ and $\xi_2$ with values in $2^X$, if we assume that these random sets are independent, then

$$P(\xi_1 = A, \xi_2 = B) = P(\xi_1 = A)P(\xi_2 = B).$$

Dempster proposes to aggregate these sources of information by a new random set $\xi$ defined by

$$P(\xi = C) = \sum_{A\cap B = C} P(\xi_1 = A, \xi_2 = B).$$

Let us notice that if we assume that the sources of information are independent,[3] then we get the original D–S rule defined by:

[2] Here we allow that a probability measure $P$ may be non-normalized. In this case $P(\emptyset) > 0$.

[3] Recently Nakama and Ruspini in [21] soften the independence requirement in D–S rule by changing it to the conditional independence to sharing knowledge, but we will not discuss this approach in our paper.