İstatistikçiler Dergisi: İstatistik & Aktüerya

Journal of Statisticians: Statistics and Actuarial Sciences

IDIA 13, 2020, 2, 78-100

Geliş/Received:05.12.2020, Kabul/Accepted: 26.12.2020

www.istatistikciler.org

Araştırma Makalesi / Research Article

A new combination method in mathematical theory of evidence "Analytic Fusion Process"

Murat Büyükyazıcı

Hacettepe Üniversitesi
Fen Fakültesi, Aktüerya Bilimleri Bölümü
06800 Beytepe, Ankara
muratby@hacettepe.edu.tr
orcid.org/*0000-0002-8622-4659*

Meral Sucu

Hacettepe Üniversitesi
Fen Fakültesi, Aktüerya Bilimleri Bölümü
06800 Beytepe, Ankara
msucu@hacettepe.edu.tr
orcid.org/*0000-0002-7991-1792*

Abstract

In this paper, we are interested in a consensus generator which combine two belief functions obtained from equally reliable and independent sources of information. The independence mentioned here is between occurrences of the sources of information. We propose a new consensus generator called "Analytic Fusion Process" which satisfy the idempotent and commutative law. Furthermore, this method also produces a measure of conflict shows whether the original beliefs were in harmony or in conflict. Another advantage is that the measure of conflict produced by this method reflects both qualitative and quantitative conflict.

Keywords: *Mathematical theory of evidence; Dempster-Shafer theory; Belief function; Data fusion; Consensus generator; Analytic fusion process; Integer programming.*

Özet

Kanıt kuramında yeni bir birleştirme yöntemi "Analitik Birleştirme Süreci"

Bu çalışmada, eşit derecede güvenilir ve bağımsız bilgi kaynaklarından elde edilen iki kanaat fonksiyonunu birleştiren bir uzlaşma oluşturucu ile ilgileniyoruz. Burada bahsedilen bağımsızlık, bilgi kaynaklarının oluşumları arasındadır. Eş kuvvetlilik ve değişme özelliklerini sağlayan "Analitik Birleştirme Süreci" adlı yeni bir uzlaşma oluşturucu öneriyoruz. Bu yöntem aynı zamanda orijinal kanaatlerin uyum içinde mi yoksa çelişki halinde mi olduğunu gösteren bir çelişki ölçüsü üretir. Diğer bir avantaj, bu yöntemle üretilen çelişki ölçüsünün hem nitel hem de nicel çelişkiyi yansıtmasıdır.

Keywords: *Kanıt kuramı, Dempter-Shafer kuramı, Kanaat fonksiyonu, Veri birleştirme, Uzlaşı üretici, Analitik birleştirme süreci, Tam sayılı programlama.*

1. Introduction

The mathematical theory of evidence (MTE), also known as Dempster-Shafer (DS) theory of evidence, has gained relatively wide acceptance as a reasonable tool for the representation, and combination such as revision, updating and conditioning of uncertain knowledge, evidence or information. The seminal work [1] on this subject is carried out by G. Shafer in 1976. This work was an expansion of A. Dempster's study

[2] about upper and lower probabilities dated 1967. In a finite discrete space, MTE can be interpreted as a generalization of probability theory where probabilities are assigned to sets as opposed to mutually exclusive singletons [3]. On the other hand, there are also some theoritical contributions which can be considered as an extension of the classical MTE but includes fundamental differences [4, 5, 6, 7].

Aggregation or fusion of information are basic concerns for all kinds of knowledge-based systems from image processing to decision making, from pattern recognition to machine learning [8]. The basic rule of combination to aggregate or fusion of information from distinct bodies of evidence in the framework of MTE is Dempster's rule of combination (DRC). However, after L. A. Zadeh presented an example for which this method gives counter-intuitive results especially in the case of existence of high conflict between bodies of evidence [9, 10], the success story of DS theory was abruptly slowed down [11]. In literature, also other examples claiming that DRC gives counter-intuitive results can be found [12, 13, 14, 15, 16]. Although R. Haenni and many others state that the counter-intuitive results are not a problem of DRC, but rather a problem of misunderstandings and misapplications [17, 11]; many alternative combination rules have been proposed in the framework of the MTE [18, 19, 20, 21, 13, 14, 22, 23, 24, 25, 26] for situations where DRC is not applicable, where its assumptions are not satisfied. D. Dubois and H. Prade have stated that many alternative rules can potentially occupy a continuum between conjunction (AND-based on set intersection) and disjunction (OR-based on set union) [27]. DRC is a *conjunctive operation,* because in the pairwise combination phase, the intersections of focal elements from distinct bodies of evidence are used. In this phase of some combination rules, the union of focal elements are used. These rules are called *disjunctive operation*. There is also a third type of combination rules called *tradeoff operation* in which task is carried out by using both intersection and union operators [27].

There are some alternative combination rules that have different assumptions from MTE on dependency of the sources of information or on reliability of the sources. R. Haenni and S. Hartmann's paper approaches the problem of independent and partially reliable information source from a very general perspective by using the *theory of probabilistic argumentation* [6] as modeling and DS theory of evidence as the underlying mathematical mechanism [28]. If the independency of the sources of information assumption is questionable, M. E. G. V. Cattaneo suggests using the least specific combination minimizing the conflict among the ones allowed by a simple generalization of Dempster's rule [29], does not propose a new method. In MTE, if the requirement of independence between sources of information is not satisfied, it is prevented a direct application of Dempster's rule. So, in [30] P. A. Monney and M. Chan proposed a method for dealing with this situation. The method relies on the ability of the *theory of hints* [5] to explicitly represent variables for which probabilistic information is avalibale. The fundamental limitation of DRC lies in the assumption that the belief functions combined be based on distinct bodies of evidence. So, in the case of nondistinct bodies of evidence, T. Denœux proposed two new commutative, associative and idempotent operators for belief functions; the *cautious conjunctive rule*, and the *bold disjunctive rule* [31]. These operators rely on the *transferable belief model* [4]. Beside the distinctness assumption, the choice of one operator between two depends on assumptions regarding reliability of the sources. If all are reliable then the cautious conjunctive rule, or if at least one is reliable then the bold disjunctive rule was recommended. In the recently paper, K. Yamada proposed a new model of combination and a new rule of combination called *combination by compromise* as a consensus generator [32]. This combination model for consensus generation is based on the assumptions that the information sources are independent in the sense of occurrance but may collide with each other over their contents, and that the information sources may not be totally reliable.

Experts and researchers in the fields today have no agreement on the superiority on one combination method over the others [33, 17, 34, 15]. To prevent this confusion about the belief combination, it should be clear what it means to combine belief functions; conditioning, revisioning, updating, consensus generating or anything else. Then, it should be stated expressly what kind of dependence or independence assumed. There are many concepts of dependence. In [30], P. A. Monney and M. Chan discriminate between dependence at the sources level, namely at the level of the assumption variables and dependence at the Θ level, namely *distinctness* [15] of bodies of evidence. In two papers [35, 36], B. B. Yaghlane et al. introduce the concept of *doxastic independency*. This concept of independence is once again represented at the Θ level [30]. K. Yamada asserts that the problem comes from confusion of independence between occurences of bodies of evidence with *consistency between contents of the information* [32]. It also should be stated expressly how reliable the sources of information are; partially or totally? Are the reliabilities of the sources of information equally or not?

A set of experts or other sources of information can provide more information than a single expert. Although it is sometimes reasonable to provide a decision maker with only the individual experts' opinions seperately, it is often necessary to combine the opinions into a single one [37]. In many cases, a single belief function is needed for input into a decision model. Even if this is not the case, it can also be informative and effective to generate a *combined belief function* as a summary of the available information. So, a *consensus generator* in the framework of the MTE is needed. A consensus generator is expected to be idempotent and commutative, but associativity is dispensable such as in the case of simple aritmethic mean operator. Also, a *measure of conflict* is useful to show whether the original beliefs were in harmony or in conflict. A consensus generator and measure of conflict are more informative if they are given together as in the case mean and variance. Recently, K. Yamada's rule of combination called combination by compromise as a consensus generator is commutative; however, it does not satisfy the idempotent law nor the associative law [32]. The *weighted average operator* [38] also can be tought as a consensus generator. The weighted average operator also does not satisfy the idempotent law nor the associative law. Furthermore, neither of both does not generate any measure of conflict.

In MTE, there are some rules which also produce a measure of conflict. However, there are no mechanisms to measure the degree of conflict other than using the mass of the combined belief assigned to the emptyset before normalization [39]. Is the mass of the combined belief to the emptyset before normalization real conflict? Under this consideration W. Liu has proposed a method to measure the conflict among beliefs using a pair of values, and then investigated the effect of these measures on deciding when Dempster's rule can be applied [39]. Still, there is not any real single measure of conflict which measures how the original beliefs were in conflict.

In this paper, we are interested in a consensus generator which combines two belief functions obtained from equally reliable and independent sources of information. The independence mentioned here is between occurrences of the sources of information. We propose a new consensus generator called *Analytic Fusion Process* (AFP) which satisfies the idempotent and commutative law. Furthermore, this method also produces a measure of conflict which shows how the original beliefs were in conflict. Another advantage of the AFP is that the measure of conflict produced by AFP reflects both qualitative and quantitative conflict. Section 2 presents a background on MTE and DRC. Some definitions and principles of the AFP are given in Section 3, and Section 4 introduces the three stages of the AFP: definition, matching, and combination stages. Section 5 gives two examples, one of them is for a general case example and the other for a special case example. A comparison among the combination by compromise, the weighted average operator and the AFP is provided in Section 6, and the conclusion is provided at Section 7.

2. Mathematical Theory of Evidence

In MTE, knowledge is represented by *basic belief assignment (bba)* or *belief function*. These two functions have one-to-one correspondence [1]. The set called *sample space* in the traditional probability theory is called *frame of discernment* in the MTE. The frame of discernment is a finite set of mutually exclusive elements, denoted by Θ hereafter. The set of all subsets of Θ is called the power set of Θ and is denoted by 2^{Θ}. The *bba* is defined as follows:

Definition 1. Let's assume Θ is a frame of discernment, then a function $m: 2^{\Theta} \rightarrow [0,1]$ is called a *bba* whenever it satisfies the following conditions:

(1) $m(\varnothing)=0$

(2) $\sum_{A \subseteq \Theta} m(A)=1$.

The quantity of $m(A)$ is called *basic belief mass* (*bbm*), and it is meant to be the measure of the belief that is committed exactly to *A* (exactly in A in nothing smaller). Each subset $A \subseteq \Theta$ with $m(A)>0$ is called a *focal element* of m and their set is represented by ε. In this case *body of evidence* is represented by $E(\varepsilon, m)$. If the true state of interest is denoted by θ and the set of its possible values by Θ, then the *propositions* of interest are precisely those of the form "the true value of θ is exactly in A" where A is a subset of Θ [1]. In this paper we will use the word "proposition" to refer not only to set A but also to its *bbm* $m(A)$. In this case, if we denote the true state (or value) of interest by θ and the set of its possible values by Θ, any focal element A of m and its *bbm* $m(A)$ can also thought as a proposition which asserts that "the true state of θ is exactly in A with a quantity of belief $m(A)$". Thus, a proposition is composed of two parts: a focal element, which we call the *qualitative part* of the proposition because it gives the content of the proposition, and a *bbm*, which we call the *quantitative part* of the proposition because it gives the quantity of the proposition.

The belief function measures how much the information from a body of evidence supports the belief in a set specified elements as the right answer [18]. A definition of the belief function which can be equivalently represented by m is given as follows:

Definition 2. A function $Bel: 2^{\Theta} \rightarrow [0,1]$ is called a belief function over Θ if it is given for some *bba* $m: 2^{\Theta} \rightarrow [0,1]$ as follows:

$$Bel(A)=\sum_{B \subseteq A} m(B).$$

The quantity $Bel(A)$ is called *A*'s *degree of belief*, and it is meant to be the measure of the total belief that is committed to *A*. If all the focal elements are singletons than this is the *Bayesian belief function* [1].

Let Bel_1 and Bel_2 be two belief functions from two distinct bodies of evidence $E_1(\varepsilon_1, m_1)$ and $E_2(\varepsilon_2, m_2)$ over the same frame of discernment Θ. If Bel_1's focal elements are denoted by $\varepsilon_1=\{A_1,\ldots,A_n\}$ and Bel_2's focal elements by $\varepsilon_2=\{B_1,\ldots,B_m\}$, then the combination result (denoted here by index *D*) is given by the following equation.

$$m_D(X)=\frac{\sum\limits_{\substack{i,j \\ A_i \cap B_j=X}} m_1(A_i) \times m_2(B_j)}{1-\sum\limits_{\substack{i,j \\ A_i \cap B_j=\varnothing}} m_1(A_i) \times m_2(B_j)} \quad , \quad X \neq \varnothing \text{ and } X \subseteq \Theta$$

$m_D(X)$ is a proper *bba*, for all non-empty $X \subset \Theta$, if and only if the denominator in above equation is non-zero. The *degree of conflict* between the bodies of evidence $E_1(\varepsilon_1, m_1)$ and $E_2(\varepsilon_2, m_2)$ is defined by

$$\kappa_D = \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i) \times m_2(B_j).$$

Shafer explained in [1] at pg. 66 what the combination means if it is obtained by DRC as follows:

> "*Dempster's rule of combination permits a simple description of how the assimilation of new evidence should change our beliefs: if our initial beliefs are expressed by a belief function Bel_1 over Θ, and the new evidence alone determines a belief function Bel_2 over Θ, then after assimilating the new evidence we should have the beliefs given by $Bel_1 \oplus Bel_2$. This description avoids the doctrine that a body of evidence can always be cast in the form of a single proposition known with certainty.*"

Shafer revealed that if the effect of the new evidence on the frame of discernment Θ is to establish any particular subset with certainty then the result of DRC is similar to the result of Bayes' rule of conditioning. When the new evidence occurs in the form of a certainty, he called this special case of DRC as *Dempster's rule of conditioning* [1]. So it is obvious to say that DRC is a conditioning operation which generalize Bayes' rule of conditioning to the case where it is not necessary for the new evidence to occur in the form of certainty.

3. Foundations of the Analytic Fusion Process

Mathematical aggregation methods range from simple summary measures such as arithmetic or geometric means of probabilities to procedures based on axiomatic approaches or on various models of the information aggregation process requiring inputs regarding characteristics such as the quality of and dependence among the experts' probabilities [37]. C. Genest and J. V. Zidek declared that the "logarithmic opinion pool" involves many advantages over the "linear opinion pool" when finding the consensus distribution of the subjective probability distributions. For details, you can see [40]. If all the weights are equal, as in the case of equally reliable sources of information, the consensus distribution is proportional to the geometric mean of the individual distributions. So, as a consensus generation process, AFP is a geometric mean based analytical method that operates on the individual belief functions to produce a single combined belief function when the individual belief functions comes from equally reliable and occurrence independent sources of information.

It is now appropriate to give some definitions of AFP before it is explained in detail. In any context in this paper even it is not denoted, belief functions which will be combined are from equally reliable and occurrence independent sources of information.

3.1. Pairwise combination

Let $E_1(e_1, m_1)$ and $E_2(e_2, m_2)$ be two occurrence independent bodies of evidence defined over the same frame of discernment Θ. In this way, a belief function Bel_1 with a focal element set e_1 and *bba* m_1 and another belief function Bel_2 with a focal element set e_2 and *bba* m_2 are given. If the number of focal elements of the Bel_1 is taken as n, each focal element of this function is shown by $e_{1i},\ i = 1,\dots,n$. Similarly, if the number of focal elements of the Bel_2 is taken as m, each focal element of this function is shown by $e_{2j},\ j = 1,\dots,m$. In other words, there is n propositions for Bel_1 and m for Bel_2. So, it is obvious that reciprocal propositions of two *bba*s defined above can produce $n \times m$ pairs. In a consensus generation process, since we are looking the answer of how can we generate a new *compromised proposition* which embody the joint effect of the two propositions, one needs a kind of summarizing process for each one of the $n \times m$ pairs. The summarizing process of a pair of propositions is called *pairwise combination*. A pairwise combination process is expected to generate a compromised proposition which has its own qualitative parts and quantitative parts. In Section 3.1.1 generating qualitative parts and in Section 3.1.3 calculating quantitative parts of the compromised proposition will be given.

3.1.1. Compromised focal element

We called the qualitative parts of compromised proposition as *compromised focal element (cfe)*. According to the concept of reliability of the information sources, three ideas might be possible to generate *cfe*s if information sources are equally reliable. If information sources are completely reliable $e_{1i} \cap e_{2j}$ should be chosen as in DRC [1], if at least an unknown one of the sources reliable $e_{1i} \cup e_{2j}$ should be chosen as in Dubois and Prade's combination rule [20]. K. Yamada proposes a third approach, combination by compromise, as a natural consensus [32]. In this paper we use this third approach to generate a *cfe.* According to this approach *cfe* composed of three subsets of $e_{1i} \cup e_{2j}$: $e_{1i} \cap e_{2j}$, $e_{i\backslash j}^{1\backslash 2}$, and $e_{i\backslash j}^{2\backslash 1}$. It is obvious to see that the first subset $e_{1i} \cap e_{2j}$ is just an intersection set. The second subset $e_{i\backslash j}^{1\backslash 2}$ is the set of all elements which are members of e_{1i}, but not members of e_{2j}. Similarly, the third subset $e_{i\backslash j}^{2\backslash 1}$ is the set of all elements which are members of e_{2j}, but not members of e_{1i}. The reason why *1\2* is used as a superscript is that so as to exterminate the confusion which may occur under the condition that *i* and *j* focal element numbers given in subscript are the same. In AFP additionally, we called the $e_{1i} \cap e_{2j}$ as *agreement se*t, $e_{i\backslash j}^{1\backslash 2}$ as *conflict set one,* and $e_{i\backslash j}^{2\backslash 1}$ as *conflict set two*. Consequently, the *cfe* can be represented as follows:

$$(e_{i\backslash j}^{1\backslash 2})((e_{1i} \cap e_{2j}))(e_{i\backslash j}^{2\backslash 1}) . \qquad (1)$$

3.1.2. Region sharing of compromised focal element

The focal elements, namely qualitative parts of two propositions, e_{1i}, e_{2j} and the *bbm*s of the focal elements, namely quantitative parts of the same propositions, $m_1(e_{1i})$, $m_2(e_{2j})$ are given. A pairwise combination of two propositions includes two issues: how to make *qualitative summarizing*, and how to make *quantitative summarizing*. Producing the *cfe* releases the first issue. The second issue can be rewritten as how to share the region of $m_1(e_{1i}) \times m_2(e_{2j})$, namely the size of *pairwise combination region,* among subsets $e_{1i} \cap e_{2j}$, $e_{i\backslash j}^{1\backslash 2}$, and $e_{i\backslash j}^{2\backslash 1}$. The justification of using the word "region" is simple: When the *bbm*s of m_1 and m_2 are depicted graphically on line segments on y-axis and x-axis respectively with both sizes one, the size of $m_1(e_{1i}) \times m_2(e_{2j})$ of any pairwise combination is equal to a region in the total area equal to one (See Fig. 2 in Section 3.2). In each pairwise combination region, we called the parts as *agreement region* belonging to agreement set elements $e_{1i} \cap e_{2j}$ and the remaining parts as *conflict region one* and *conflict region two* belonging to conflict set elements $e_{i\backslash j}^{1\backslash 2}$ and $e_{i\backslash j}^{2\backslash 1}$, respectively. The region sharing can be made proportionally with respect to both their *bba*s assigned by the information sources and the cardinalities of the subsets. Thus, if the size of agreement region is given as r_{ij}^{a} whereas the size of conflict regions is given as r_{ij}^{b}, and r_{ij}^{d} respectively, the pairwise combination region sharing can be made as in Fig. 1.

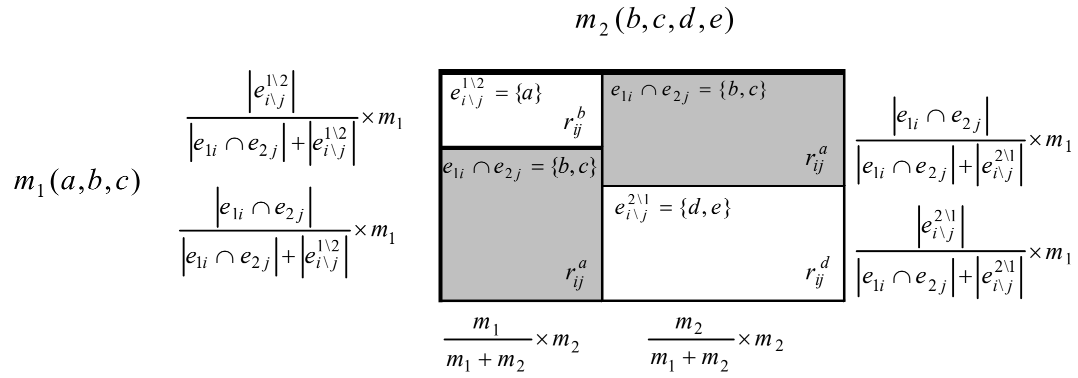

Fig. 1. Illustration of region sharing.

In Fig. 1, four small rectangles represent how a pairwise combination region can be allocated to the subsets of *cfe* proportionally with respect to both their *bba*s and the cardinalities of the subsets. Firstly, In Fig.1, the vertical line shares the biggest rectangle two parts proportionally with respect to their *bba*s. Secondly, the two horizontal lines share the two separated parts proportionally with respect to their subset's cardinalities. So, the region sharing can be made by the following equations:

$$r_{ij}^{a}=\left(\frac{m_1(e_{1i})}{m_1(e_{1i})+m_2(e_{2j})}\times\frac{\left|e_{1i}\cap e_{2j}\right|}{\left|e_{1i}\cap e_{2j}\right|+\left|e_{i\setminus j}^{1\setminus 2}\right|}+\frac{m_2(e_{1i})}{m_1(e_{1i})+m_2(e_{2j})}\times\frac{\left|e_{1i}\cap e_{2j}\right|}{\left|e_{1i}\cap e_{2j}\right|+\left|e_{i\setminus j}^{2\setminus 1}\right|}\right)\times m_1(e_{1i})\times m_2(e_{2j}),$$

$$r_{ij}^{b}=\frac{m_1(e_{1i})}{m_1(e_{1i})+m_2(e_{2j})}\times\frac{\left|e_{i\setminus j}^{1\setminus 2}\right|}{\left|e_{1i}\cap e_{2j}\right|+\left|e_{i\setminus j}^{1\setminus 2}\right|}\times m_1(e_{1i})\times m_2(e_{2j}),$$

$$r_{ij}^{d}=\frac{m_1(e_{1i})}{m_1(e_{1i})+m_2(e_{2j})}\times\frac{\left|e_{i\setminus j}^{2\setminus 1}\right|}{\left|e_{1i}\cap e_{2j}\right|+\left|e_{i\setminus j}^{2\setminus 1}\right|}\times m_1(e_{1i})\times m_2(e_{2j}).$$

3.1.3. Mass sharing of compromised focal element

It is important in the case of consensus generation to emphasize that the geometric mean of *bbm*s should be used for quantitative summarizing not just the area of the pairwise combination region as in the case of conditioning with DRC. In conditioning case, multiplying of two *bbm*s is reasonable, but in the case of consensus generation averaging is expected. So, after sharing out of the region $m_1(e_{1i})\times m_2(e_{2j})$ between the agreement region, the conflict region one and the conflict region two, it is need to take square roots of distributed region as in the case of two probabilities' geometric mean. We called the parts as *mass of agreement* belonging to agreement set elements $e_{1i}\cap e_{2j}$, *mass of conflict one* belonging to conflict set $e_{i\setminus j}^{1\setminus 2}$, and *mass of conflict two* belonging to conflict set $e_{i\setminus j}^{2\setminus 1}$. If the mass of agreement is given by $a_{ij}(e_{1i}\cap e_{2j})$ or shortly a_{ij}, the mass of conflict one and two are given by $b_{ij}(e_{i\setminus j}^{1\setminus 2})$ or shortly b_{ij}, and $d_{ij}(e_{i\setminus j}^{2\setminus 1})$ or shortly d_{ij} respectively, the mass sharing can be made by the following equations:

$$a_{ij}=\sqrt{r_{ij}^{a}}\ , \tag{2}$$

$$b_{ij} = \sqrt{r_{ij}^{b}}\,, \tag{3}$$

$$d_{ij} = \sqrt{r_{ij}^{d}}\,, \tag{4}$$

Sum of the mass of conflict one and the mass of conflict two, that is $b_{ij} + d_{ij}$, is called *mass of conflict*.

3.1.4. Qualitative and quantitative conflicts

In the phase of pairwise combination, the elements which are involved in one of the focal elements which will be combined but not involved in the other one cause a *qualitative conflict*. As a measure of qualitative conflict, we use the proportion of the mass of conflict in sum of the mass of agreement and the mass of conflict for each pairwise combination. Since this is a proportion, it takes value on $[0, 1]$. So, if the size of qualitative conflict is given by c_{ij} , it can be formulated as follows:

$$c_{ij} = (b_{ij} + d_{ij}) / (a_{ij} + b_{ij} + d_{ij}) \tag{5}$$

Whether there is a qualitative conflict or not, there can be a quantitative conflict. *Quantitative conflict* is a conflict which stems from inequality of $m_1(e_{1i})$ and $m_2(e_{2j})$ *bbms*. It has been stated that mass of agreement and mass of conflicts are geometric means of parts of $m_1(e_{1i})$ and $m_2(e_{2j})$ *bbms* which belong to agreement set elements and conflict sets elements. As a measure of quantitative conflict we use the measure of deviation from geometric mean g_{ij} of $m_1(e_{1i})$ and $m_2(e_{2j})$ *bbms*, namely geometric variance v_{ij} . When $n = 2$, geometric mean and geometric variance values can be obtained as in Eq. (6) and Eq. (7).

$$g_{ij} = \sqrt{m_{1i}(e_{1i}) \times m_{2j}(e_{2j})} \tag{6}$$

$$v_{ij} = 1 - \frac{\max\left\{\dfrac{m_{1i}(e_{1i})}{g_{ij}}, \dfrac{g_{ij}}{m_{1i}(e_{1i})}\right\}}{\dfrac{\max\{m_{1i}(e_{1i}), m_{2j}(e_{2j})\}}{\min\{m_{1i}(e_{1i}), m_{2j}(e_{2j})\}}} \tag{7}$$

In Eq. (7) geometric variance is developed for AFP as being defined on $[0, 1]$. It represents the size of quantitative conflict. This value is closer to 0 in situations where the difference between *bbms* is low and it is closer to 1 in situations where the difference between *bbms* is high.

3.2. Matching the most suitable propositions

In this section, to determine the matching the most suitable propositions principle of the AFP, we will start making analogy between the simple case of averaging two Bayesian belief functions and the general case of averaging two belief functions.

Example 1. Given two Bayesian belief functions, Bel_1 and Bel_2 , over the same frame of discernment $\Theta = \{a, b, c\}$ from two bodies of evidence $E_1(\varepsilon_1, m_1)$ and $E_2(\varepsilon_2, m_2)$ as follows:

$$E_1(\varepsilon_1, m_1) = \{\, m_1(a) = 0.1\,;\; m_1(b) = 0.2\,;\; m_1(c) = 0.7 \,\}$$

$$E_2(\varepsilon_2, m_2) = \{ m_2(a) = 0.2 \, ; \; m_2(b) = 0.3 \, ; \; m_2(c) = 0.5 \}.$$

In this simple case, to obtain a single combined belief function from given these two Bayesian belief functions, we simply match the propositions which have the same focal elements reciprocally and average the *bbm*s of matched propositions. Since we use the geometric mean as an averaging operator, it is needed normalization. Normalization is carried out by calculating the portions of these geometric means in their sum. The combined belief function obtained with matching, averaging and normalization process is as follows: $m_{geo}(a) = 0.1446$, $m_{geo}(b) = 0.2505$, $m_{geo}(c) = 0.6049$.

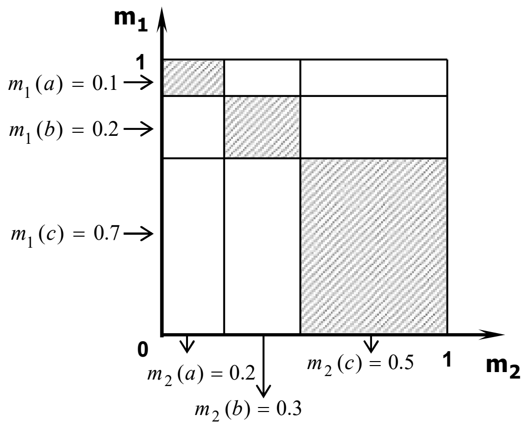

Fig. 2. Illustration of pairwise combinations of propositions.

The matching situations can be seen graphically in Fig. 2. The propositions $m_1(a) = 0.1$, $m_1(b) = 0.2$, and $m_1(c) = 0.7$ of Bel_1 were matched and pairwise combined with the propositions $m_2(a) = 0.2$, $m_2(b) = 0.3$, and $m_2(c) = 0.5$ of Bel_2, respectively. In the figure, when the shaded regions correspond to *matched pairwise combinations*, the clear regions correspond to the unmatched pairwise combinations.

It is important in the case of consensus generation to emphasize that the clear regions in this example neither represent the conflict mass nor have any other sense. So, the right way of matching a proposition from a body of evidence would be to match it with only one proposition in the other body of evidence. In other words, proposition $m_1(a) = 0.1$ of Bel_1 has to be matched with most suitable one $m_2(a) = 0.2$ of Bel_2. The propositions $m_1(a) = 0.1$ and $m_2(a) = 0.2$ must not be pairwise combined with others. Therefore, in Fig. 2, only one matching must be made in each row and column. So, a proposition from a body of evidence must be matched and pairwise combined with the most suitable proposition from the other body of evidence. The size of qualitative conflict, and the size of quantitative conflict lead determination of the most agreeable matching among propositions. This will be explained in Section 4.2.

4. Stages of the Analytic Fusion Process

AFP is composed of three stages. These three stages are called *definition*, *matching*, and *combination* stages respectively. In the definition stage, a general combination table which lays outs the outputs produced by pairwise combinations the propositions of the *bba*s which will be combined. In the matching stage, the propositions are matched so as to minimize the conflict by an integer programming problem for conflict minimization. In the combination stage, the combined *bba* is obtained based on the matches of the previous stage.

4.1. Definition stage

A *bba* having n focal elements and a *bba* having *m* focal elements are to be combined. In the definition stage of the process, the pairwise combinations are performed $n \times m$ times; then *cfe*, mass of agreement a_{ij}, mass of conflicts b_{ij} and d_{ij}, the size of qualitative conflict c_{ij}, and the size of quantitative conflict v_{ij} are obtained in the result of each combination as explained in Section 3.1. So, all results are achieved for each of $n \times m$ combinations. The table in which the results obtained through $n \times m$ numbered pairwise combinations of propositions which belong to two *bba*s is called *general combination table* (See Table 1 in Section 5.1). In the i th row and the j th column of this table, pairwise combinations results of the the i th proposition of the first *bba* and the j th proposition of the second *bba* are given.

4.2. Matching stage

A *linear programming* [41] problem is an optimization problem for which we attempt to maximize or minimize a linear function of the decision variables. The function that is to be maximized or minimized is called the *objective function*. The values of the decision variables must satisfy a set of *constraints*. Each constraint must be a linear equation or linear inequality. In any linear programming model, the *decision variables* should completely describe the decisions to be made [41]. As a special case of the linear programming problem, if all decision variables are restricted to be integers the problem is called integer programming, if all decision variables are restricted to 0 or 1 values the problem is called *0-1 integer programming* [41, 42].

In Section 3.2 with Example 1 we explained why it is needed to match the most suitable propositions of the bbas. So, the propositions of two *bba*s which are closest to each other in the sense of quality and quantity aspects should be matched. The combined *bba* should be calculated based on the results of these matching. In matching stage, we propose to make the matching with the assistance of a 0-1 integer programming problem, shortly we said *matching problem*. Decision variables, constraints and objective function of matching problem will be given in Section 4.2.1, Section 4.2.2, and Section 4.2.3, respectively. In Section 4.2.4 the matching problem will be given as a whole.

4.2.1. Decision variables

Taking the general combination table with *n* row and *m* column into consideration, the matching of the propositions is made by valuing the matching decision variable of the cells as *1* where the matching is agreed, and valuing the matching decision variable of the cells as *0* where the matching is not agreed. Thus, x_{ij} the *matching decision variable* is defined as:

$$x_{ij} = \begin{cases} 1, & \textit{if the } i\text{th}\textit{ focal element of the first bba is matched} \\ & \textit{with the } j\text{th}\textit{ focal element of the second bba} \\ 0, & \textit{otherwise} \end{cases} \tag{8}$$

After the definition of the decision variables about the matching of the propositions, the constraints and the objective function should be determined using the decision variables.

4.2.2. Constraints

In the case that the numbers of propositions of the two *bba*s to be combined, *n* and *m,* are equal, each proposition *bba* should have been matched with a proposition of the other *bba* or vice versa. However, in the case that the numbers of propositions, *n* and *m* are not equal, it is natural that some of the propositions will be out of matching. When a generalization is made by considering this situation, the sum of matching decision variables in each row and also each column of the general combination table can be at most *1*. So, the constraints related to each row cells are as follows,

$$\sum_{j=1}^{m} x_{ij} \leq 1 \qquad i = 1,...,n. \tag{9}$$

The constraints related to each column cells are as follows,

$$\sum_{i=1}^{n} x_{ij} \leq 1 \qquad j = 1,...,m. \tag{10}$$

The sum of matching decision variables in each row and each column does not exceed *1* and this has been acquired with the constraints given in Eq. (9) and Eq. (10).

However, in such a case, all the matching decision variables would be *0* in a conflict minimization problem. By adding the constraint that the total of all the matching decision variables will be $\min\{n, m\}$, the smallest value of the numbers of row or column is matched. To this end, the constraint

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij} = \min\{n, m\} \tag{11}$$

is added to the problem. Lastly, the constraints showing that the matching decision variables can be only 0 or 1,

$$x_{ij} = \{0,1\} \qquad i = 1,...,n, \ j = 1,...,m \tag{12}$$

are added to the problem. These constraints have the generalized forms for the situations that the numbers of propositions, *n* and *m,* are equal or not.

4.2.3. Objective function

As the matching decision variables and the related constraints have been determined, the fact that under which circumstances the matching decision variables will take the value 0 or 1 based on constraints Eq. (9), Eq. (10), Eq. (11) and Eq. (12) should be determined. Following the pairwise combination phase a_{ij} mass of agreement, c_{ij} size of qualitative conflict and v_{ij} size of quantitative conflict for each of the pairwise combinations have been calculated and all of these results have been shown in the general combination table.

Taking only the size of qualitative conflict or the size of quantitative conflict in the objective function into account would be not enough. So, the objective function should be a form of minimizing a function of both. Such an objective function can be given as shown in Eq. (13) in order to minimize the total of the *generalized means* [37] of the size of qualitative conflict c_{ij}, and the size of quantitative conflict v_{ij}.

$$\min \sum_{i=1}^{n}\sum_{j=1}^{m} \left(\left(\frac{c_{ij}^2 + v_{ij}^2}{2} \right)^{1/2} \times x_{ij} \right) \tag{13}$$

Either the arithmetic or the geometric mean of the c_{ij} and v_{ij} could have been used in Eq. (13). However, the generalized mean has been preferred here as it always yields in a bigger value than the arithmetic and geometric mean. So, it has been made harder to make matching between propositions have bigger qualitative and/or quantitative conflict.

4.2.4. Matching problem

The problem as how to match the $\min\{n,m\}$ pair from $n \times m$ probable pairwise combinations under constraints given in Eq. (9), Eq. (10), Eq. (11) and Eq. (12) with objective function given in Eq. (13) can be overcome through optimal solution of 0-1 integer programming problem. Thus; matching problem of the AFP given as follows in Eq. (14):

$$\begin{aligned}
&\min \sum_{i=1}^{n}\sum_{j=1}^{m}\left(\frac{c_{ij}^2+v_{ij}^2}{2}\right)^{1/2} \times x_{ij} \\
&\qquad \sum_{j=1}^{m} x_{ij} \le 1, \qquad\qquad i=1,...,n \\
&\qquad \sum_{i=1}^{n} x_{ij} \le 1, \qquad\qquad j=1,...,m \\
&\qquad \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij} = \min\{n,m\} \\
&\qquad x_{ij} = 0,1 \quad i=1,...,n,\ j=1,...,m.
\end{aligned} \tag{14}$$

The $n \times m$ cells of the general combination table had $n \times m$ matching decision variables x_{ij}, in a matching problem. In linear programming, a *feasible solution* is a choice of values for decision variables that satisfies all constraints. For a minimization problem as in the case matching problem, the *optimal solution* is a feasible solution that gives the smallest objective function value. If an optimization problem has more than one optimal solution, it is said that it has *alternative optimal solutions* [41]. In an optimal solution of a model of Eq. (14) created in matching stage; $\min\{n,m\}$ of them get *1* and the rest get 0. As stated in Section 4.2.1, matching decision variables which have value *1* in optimal solution determine the matched pairwise combinations. The optimal solution method of Eq. (14) is not subject of this paper. Detailed information about solving this mathematical model can be found at [41, 42]. There are also many computer programs solving this kind of optimization problem.

4.3. Combination stage

In the calculation of *combined bba*; the matched pairwise combinations will be used as base and if necessary, the other non-base pairwise combinations will be considered with a lower weight than the weight of the matched pairwise combinations. A pairwise combination of matched propositions is called *base of pairwise combination*. The cell corresponding to this pairwise combination in the general combination table will be called *base cell of pairwise combination*. In other words, the cells corresponding to the matching decision variables which get *1* in the general combination table are called base cells of pairwise combination. The weights of the cells of pairwise combination on the general combination result are shown with w_{ij} at the first row of every cell of pairwise combination in the general combination table. The weight of base cell of pairwise combination will be shown with w_{ij}^b, the size of qualitative conflict with c_{ij}^b and the size of quantitative conflict with v_{ij}^b. The general combination table in which the weights are also shown is called *weighted general combination table*. Table 2 in Section 5.1 gives a weighted general combination table for Example 2.

The weights of the base cells of pairwise combination on the general combination result are calculated as shown in Eq. (15).

$$w_{ij}^{b} = 1 - \left(\frac{(c_{ij}^{b})^2 + (v_{ij}^{b})^2}{2} \right)^{1/2} \tag{15}$$

Then, firstly the one with the biggest weight of the base cells of pairwise combination is considered. The weight of the non-base cells of pairwise combination that is positioned in the same row or column with the base cell of pairwise combination with the biggest weight is calculated as shown in Eq. (16).

$$w_{ij} = 1 - w_{ij}^{b} \tag{16}$$

If there is neither qualitative nor quantitative conflict ($c_{ij}^{b} = 0$, $v_{ij}^{b} = 0$) between two *bba*s of the base cell of pairwise combination; Eq. (15) and Eq. (16) provide that the combination weight of the base cell of pairwise combination is $w_{ij}^{b} = 1$. Thus, the combination weight of the other cells of pairwise combination that are in the same row or column as this base cell of pairwise combination are zero as in the case of Example 1. If there is any measure of qualitative or quantitative conflict, Eq. (15) and Eq. (16) provide that the weight of the base cell of pairwise combination on the combination is measured by subtracting the generalized mean of this qualitative or quantitative measure from *1*. In such a case, the combination weight of other pairwise combinations that are in the same row or column with this base cell of pairwise combination equals to the generalized mean of conflicts in the base cell of pairwise combination.

A non-base cell of pairwise combination can be in the same row with a base cell of pairwise combination and in the same column with another base cell of pairwise combination. If this is the case, how can the combination weight for a non-base cell of pairwise combination be achieved? Which one of the base cells of pairwise combination is affecting the non-base cell of pairwise combination? It will be wise firstly to calculate combination weights of the non-base cells of pairwise combination, which are in the same row or column with the base cell of pairwise combination with the biggest combination weight of all base cells of pairwise combination. Once the weight of a non-base cell of pairwise combination is determined, it will not change then.

In the general combination table, after the weights for all the cells of pairwise combination are calculated the general combination results are calculated as in Eq. (17),

$$m_{AFP}(X) = \frac{\sum\limits_{\substack{i,j \\ e_{i \cap j} = X}}\sum w_{ij} \times a_{ij}(X) + \sum\limits_{\substack{i,j \\ e_{i \setminus j}^{1\setminus 2} = X}}\sum w_{ij} \times b_{ij}(X) + \sum\limits_{\substack{i,j \\ e_{i \setminus j}^{2\setminus 1} = X}}\sum w_{ij} \times d_{ij}(X)}{\sum\limits_{\substack{i,j \\ A \in 2^{\Theta}}}\sum w_{ij} \times a_{ij}(A) + \sum\limits_{\substack{i,j \\ A \in 2^{\Theta}}}\sum w_{ij} \times b_{ij}(A) + \sum\limits_{\substack{i,j \\ A \in 2^{\Theta}}}\sum w_{ij} \times d_{ij}(A)}, \quad X \neq \varnothing \tag{17}$$

and then the *combined bba* and so the *combined belief function* is achieved. Eq. (17) shows the combined *bba* $m_{AFP}(X)$, which is given for any of the subsets $X \in 2^{\Theta}$ products of agreement or conflict sets achieved through pairwise combination phases. It can be seen in the numerator of Eq. (17) that the mass of agreement or conflict (a_{ij}, b_{ij}, d_{ij}) committed to the sets with X agreement or conflict sets ($e_{i \cap j} = X$, $e_{i \setminus j}^{1\setminus 2} = X$, or $e_{i \setminus j}^{2\setminus 1} = X$) of all the cells of pairwise combination are added by multiplying with their weights (w_{ij}). In the denominator the mass of agreement and conflict committed to all the agreement and conflict sets ($A \in 2^{\Theta}$) are added by multiplying with their weights. Indeed, this is normalization. So, the value N in Eq. (18) is called *normalization constant*.

$$N = \sum_{\substack{i,j \\ A \in 2^{\Theta}}}\sum w_{ij} \times a_{ij}(A) + \sum_{\substack{i,j \\ A \in 2^{\Theta}}}\sum w_{ij} \times b_{ij}(A) + \sum_{\substack{i,j \\ A \in 2^{\Theta}}}\sum w_{ij} \times d_{ij}(A) \tag{18}$$

Then, given Eq. (18), Eq. (17) can be rewritten as Eq. (19).

$$m_{AFP}(X) = \left(\sum_{\substack{i,j \\ e_{i \cap j} = X}}\sum w_{ij} \times a_{ij}(X) + \sum_{\substack{i,j \\ e_{i \setminus j}^{1 \setminus 2} = X}}\sum w_{ij} \times b_{ij}(X) + \sum_{\substack{i,j \\ e_{i \setminus j}^{2 \setminus 1} = X}}\sum w_{ij} \times d_{ij}(X) \right) \Big/ N, \quad X \neq \varnothing \tag{19}$$

The normalization term N given in Eq. (18) of the combined *bba* is composed of three terms. In the first term the mass of agreement committed to all of the agreement sets are added by multiplying with their weights. In the second and third terms the mass of conflict committed to all of the conflict sets are added by multiplying with their weights. In the Eq. (18) if the first term is shown by N_a, and the total of the second term and the third term is shown by N_c Eq. (18) can be rewritten as Eq. (20). The total *measure of conflict* κ as an indicator of the effect of the conflict sets over the general combination results is achieved by the Eq. (21).

$$N = N_a + N_c \tag{20}$$

$$\kappa = N_c / (N_a + N_c) = N_c / N \tag{21}$$

It is said that as the total measure of conflict approaches *1* the conflict between the *bba*s which will be combined is high and as it approaches *0* the conflict between the *bba*s which will be combined is low. The fact that the conflict is high does not make the result meaningless but it stresses the existence of the high conflict between the data from two evidence sources.

5. Numerical examples

In this section we give two numerical examples. In these examples, we are interested in two individual belief functions to produce a single combined belief function when the individual belief functions come from equally reliable and occurrence independent sources of information. In Section 5.1, an example is illustrated to show the use of the AFP for general case. On the other hand, in Section 5.2, we chosen a special case example which have alternative optimal solution for its matching problem.

5.1. General case example

Example 2 has an optimal solution in the matching stage, so it serves showing the use of the AFP for general case.

Example 2. Given two belief functions, Bel_1 and Bel_2, over the same frame of discernment $\Theta = \{a,b,c\}$ from two bodies of evidence $E_1(\varepsilon_1, m_1)$ and $E_2(\varepsilon_2, m_2)$ as follows:

$$E_1(\varepsilon_1, m_1) = \{ m_1(a,b,c) = 0.3 ; m_1(c) = 0.1 ; m_1(c,d) = 0.2 ; m_1(a,b,c,d) = 0.4 \}$$

$$E_2(\varepsilon_2, m_2) = \{ m_2(a,b) = 0.4 ; m_2(c,d) = 0.6 \}.$$

Definition stage

The focal element set of the first *bba* is e_1 and the number of focal elements is $n = |e_1| = 4$. The focal element set of the second *bba* is e_2 and the number of focal elements is $m = |e_2| = 2$.

The agreement set of the *cfe* set to be obtained as a result of pairwise combination of the first focal element (*i=1*) of the first *bba,* namely $e_{11} = (a,b,c)$ and the first focal element (*j=1*) of the second *bba*, namely $e_{21} = (a,b)$; is

$$e_{1\cap 1} = e_{11} \cap e_{21} = \{a,b,c\} \cap \{a,b\} = \{a,b\} .$$

The conflict sets of the *cfe*, are obtained as;

$$e_{1\backslash 1}^{1\backslash 2} = e_{11} - \mathrm{e}_{21} = \{a,b,c\} - \{\mathrm{a,b}\} = \{\mathrm{c}\}$$

$$e_{1\backslash 1}^{2\backslash 1} = e_{21} - \mathrm{e}_{11} = \{a,b\} - \{\mathrm{a,b,c}\} = \{\varnothing\}$$

So with respect to the conflict set one is $\{\mathrm{c}\}$, the agreement set is $\{a,b\}$, and the conflict set two is $\{\varnothing\}$; by using the Eq. (1), the *cfe* of the first pairwise combination can be represented in the first cell of Table 1 as follows:

$$(c)\,((ab))\,(\ ) .$$

By using the Eq. (2), Eq. (3), and Eq. (4) for mass sharing, the mass of agreement, mass of conflict one and mass of conflict two are calculated as follows:

$a_{11} = a_{11}(a,b) = \sqrt{r_{11}^a} = \sqrt{0.103} \cong 0.321$, $b_{11} = b_{11}(c) = \sqrt{r_{11}^b} = \sqrt{0.017} \cong 0.131$, and

$$d_{11} = d_{11}(\varnothing) = \sqrt{r_{11}^d} = 0 .$$

By using the Eq. (5) the size of qualitative conflict is calculated as,

$$c_{11} = \frac{b_{11} + d_{11}}{a_{11} + b_{11} + d_{11}} = \frac{0.131 + 0}{0.321 + 0.131 + 0} = 0.290 .$$

By using the Eq. (6) and Eq. (7) the size of quantitative conflict is calculated as,

$$g_{11} = \sqrt{0.3 * 0.4} = \sqrt{0.12} \cong 0.346 ,$$

$$v_{11} = 1 - \frac{\max\left\{\frac{m_{11}(e_{11})}{g_{11}}, \frac{g_{11}}{m_{11}(e_{11})}\right\}}{\frac{\max\{m_{11}(e_{11}), m_{21}(e_{21})\}}{\min\{m_{11}(e_{11}), m_{21}(e_{21})\}}} = 1 - \frac{\max\left\{\frac{0.3}{0.346}, \frac{0.346}{0.3}\right\}}{\frac{\max\{0.3, 0.4\}}{\min\{0.3, 0.4\}}} \cong 0.134 .$$

Up to here, only the results of the first pairwise combination have been obtained but the rest of the *4*2-1=7* pairwise combinations will not be given here in detail. The results of all pairwise combinations are given in Table 1.

Table 1. General combination table for Example 2

	m2(ab)=0.400	m2(cd)=0.600
m1 (abc) 0.300	(c)((ab))() b11 a11 d11 0.131 0.321 0.000 c11=0.290 v11=0.134	(ab)((c))(d) b12 a12 d12 0.200 0.283 0.245 c12=0.611 v12=0.293
m1 (c) 0.100	(c)(())(ab) b21 a21 d21 0.089 0.000 0.179 c21=1.000 v21=0.500	()((c))(d) b22 a22 d22 0.000 0.185 0.160 c22=0.464 v22=0.592
m1 (cd) 0.200	(cd)(())(ab) b31 a31 d31 0.163 0.000 0.231 c31=1.000 v31=0.293	()((cd))() b32 a32 d32 0.000 0.346 0.000 c32=0.000 v32=0.423
m1 (abcd) 0.400	(cd)((ab))() b41 a41 d41 0.200 0.346 0.000 c41=0.366 v41=0.000	(ab)((cd))() b42 a42 d42 0.219 0.438 0.000 c42=0.333 v42=0.184

Matching stage

In the matching stage for this example, the matching problem formulation by utilizing Eq. (14) is obtained as:

$$\min \quad 0.226x_{11} + 0.479x_{12} + 0.791x_{21} + 0.532x_{22} + 0.737x_{31} + 0.299x_{32} + 0.259x_{41} + 0.269x_{42}$$

$$x_{11} + x_{12} \le 1,\; x_{21} + x_{22} \le 1,\; x_{31} + x_{32} \le 1,\; x_{41} + x_{42} \le 1$$

$$x_{11} + x_{21} + x_{31} + x_{41} \le 1, \quad x_{12} + x_{22} + x_{32} + x_{42} \le 1$$

$$x_{11} + x_{12} + x_{21} + x_{22} + x_{31} + x_{32} + x_{41} + x_{42} = 2$$

$$x_{ij} = 0,1 \;, \qquad i = 1,\ldots,4 \quad j = 1,2$$

In the optimal solution of the matching problem stated above, the values of the two matching decision variables are obtained as $x_{11}^* = 1$, $x_{42}^* = 1$ and the others are obtained as *0*. Thus, the *1*st proposition of the first *bba* and the *1*st proposition of the second *bba* are matched and also the *4*th proposition of the first *bba* and the *2*nd proposition of the second *bba* are matched.

Combination stage

In conformity with the above stated matching, the cells *(1,1)* and *(4,2)* of the general combination table become base cells of pairwise combination. Thus, the weights of the base cells of pairwise combination from the Eq: (15) are:

$$w_{11}^b \cong 1 - 0.226 = 0.774\,, \qquad w_{42}^b \cong 1 - 0.269 = 0.731$$

As the largest weight between the base cells of pairwise combination is $w_{11}^b = 0.774$, the weights of the cells *(1,2)* existing in the same row with the base cell *(1,1)* and the weights of the cells *(2,1)*, *(3,1)* and *(4,1)* existing in the same column with the base cell *(1,1)* are calculated by the Eq (16) as:

$$w_{21} = w_{31} = w_{42} = 0.226$$

As the second largest weight between the base cells of pairwise combination is $w_{42}^{b} = 0.731$ the weights of the cells *(2,2)* and *(3,2)*, existing in the same column with the base cell *(4,2)* are (using the Eq (16)) obtained as:

$$w_{22} = w_{32} = 0.269$$

If these obtained weight values are placed in the general combination table, the weighted general combination table will be obtained as given in Table 2.

Table 2. Weighted general combination table for Example 2

	m2(ab)=0.400	m2(cd)=0.600
m1 (abc) 0.300	w11=0.774 (c)((ab))() b11 a11 d11 0.131 0.321 0.000 c11=0.290 v11=0.134	w12=0.226 (ab)((c))(d) b12 a12 d12 0.200 0.283 0.245 c12=0.611 v12=0.293
m1 (c) 0.100	w21=0.226 (c)(())(ab) b21 a21 d21 0.089 0.000 0.179 c21=1.000 v21=0.500	w22=0.269 ()((c))(d) b22 a22 d22 0.000 0.185 0.160 c22=0.464 v22=0.592
m1 (cd) 0.200	w31=0.226 (cd)(())(ab) b31 a31 d31 0.163 0.000 0.231 c31=1.000 v31=0.293	w32=0.269 ()((cd))() b32 a32 d32 0.000 0.346 0.000 c32=0.000 v32=0.423
m1 (abcd) 0.400	w41=0.226 (cd)((ab))() b41 a41 d41 0.200 0.346 0.000 c41=0.366 v41=0.000	w42=0.731 (ab)((cd))() b42 a42 d42 0.219 0.438 0.000 c42=0.333 v42=0.184

By using the Table 2, Eq. (18), and Eq. (20) the normalization coefficient *N* is obtained as:

$$N_a = 0.854\,,\ N_c = 0.600\,,\ N = 0.854 + 0.600 = 1.454\,.$$

By using the Table 2 and the Eq. (19) the combined *bba* is obtained as:

$$m_{AFP}(c) = \left(w_{12}a_{12} + w_{22}a_{22} + w_{11}b_{11} + w_{21}b_{21}\right)/N = 0.162$$

$$m_{AFP}(d) = \left(w_{12}d_{12} + w_{22}d_{22}\right)/N = 0.068$$

$$m_{AFP}(a,b) = \left(w_{11}a_{11} + w_{41}a_{41} + w_{12}b_{12} + w_{42}b_{42} + w_{21}d_{21} + w_{31}d_{31}\right)/N = 0.429$$
$$m_{AFP}(c,d) = \left(w_{32}a_{32} + w_{42}a_{42} + w_{31}b_{31} + w_{41}b_{41}\right)/N = 0.341\,.$$

The total measure of conflict from the Eq. (21) is obtained as,

$$\kappa = 0.510/1.330 \cong 0.413\,.$$

5.2. Special case example

As a special case, some of the matching problems may have alternative optimal solutions. Example 3 has alternative optimal solutions in the matching stage, so it serves showing how can be obtained a final combined *bba* from alternative combined *bba*s.

Example 3. Given two belief functions, Bel_1 and Bel_2, over the same frame of discernment $\Theta = \{a,b,c\}$ from two bodies of evidence $E_1(\varepsilon_1, m_1)$ and $E_2(\varepsilon_2, m_2)$ as follows:

$$E_1(\varepsilon_1, m_1) = \{ m_1(a,b,c) = 1 \}$$

$$E_2(\varepsilon_2, m_2) = \{ m_2(a,b) = 0.5 \,;\; m_2(b,c) = 0.5 \}.$$

Definition stage

The general combination table in which the results obtained through $1 \times 2 = 2$ numbered pairwise combinations of propositions which belong to two *bba*s is given in Table 3.

Table 3. General combination table for Example 3

```
            |      m2(ab)=0.500      |      m2(bc)=0.500      |
----------- ---------------------- ----------------------
    m1      |      (c)((ab))()       |      (a)((bc))()       |
   (abc)    |    b11   a11   d11     |    b12   a12   d12     |
   1.000    | 0.333  0.624  0.000    | 0.333  0.624  0.000    |
            | c11=0.348  v11=0.293   | c12=0.348  v12=0.293   |
----------- ---------------------- ----------------------
```

Matching stage

In the matching stage for this example, the matching problem formulation by utilizing Eq. (14) is obtained as:

$$\min \quad 0.308x_{11} + 0.308x_{12}$$

$$x_{11} + x_{12} \le 1,\; x_{11} \le 1,\; x_{12} \le 1$$

$$x_{11} + x_{12} = 1$$

$$x_{1j} = 0,1 \;, \qquad j = 1,2$$

The matching problem stated above has alternative optimal solutions. One of the alternative solutions is $x_{11}^{*} = 1$, $x_{12}^{*} = 0$, and the other is $x_{11}^{*} = 0$, $x_{12}^{*} = 1$.

Combination stage

The weighted general combination tables according to the two alternative optimal solutions are given in Table 4a and Table 4b, respectively.

Table 4a. Weighted general combination table for first alternative optimal solution of Example 3

```
           |     m2(ab)=0.500       |     m2(bc)=0.500       |
----------- ---------------------- ----------------------
           |      w11=0.678         |      w12=0.322         |
   m1      |     (c)((ab))()        |     (a)((bc))()        |
  (abc)    |    b11   a11   d11     |    b12   a12   d12     |
  1.000    | 0.333   0.624   0.000  | 0.333   0.624   0.000  |
           | c11=0.348   v11=0.293  | c12=0.348   v12=0.293  |
----------- ---------------------- ----------------------
```

Table 4b. Weighted general combination table for second alternative optimal solution of Example 3

```
           |     m2(ab)=0.500       |     m2(bc)=0.500       |
----------- ---------------------- ----------------------
           |      w11=0.322         |      w12=0.678         |
   m1      |     (c)((ab))()        |     (a)((bc))()        |
  (abc)    |    b11   a11   d11     |    b12   a12   d12     |
  1.000    | 0.333   0.624   0.000  | 0.333   0.624   0.000  |
           | c11=0.348   v11=0.293  | c12=0.348   v12=0.293  |
----------- ---------------------- ----------------------
```

So according to two weighted general combination tables given in Table 4a and 4b, two alternatives combined bba are obtained. Then, as doing in Example 1 in Section 3.2, the geometric mean of these two alternatives combined bba can be used as the final combined bba. The results are given as follows:

The first alternative combined bba	*The second alternative combined bba*	*The final combined bba with geometric mean*
$m_{AFP}(a) = 0.112$	$m_{AFP}(a) = 0.236$	$m_{AFP}(a) = 0.174$
$m_{AFP}(c) = 0.236$	$m_{AFP}(c) = 0.112$	$m_{AFP}(c) = 0.174$
$m_{AFP}(a,b) = 0.442$	$m_{AFP}(a,b) = 0.210$	$m_{AFP}(a,b) = 0.326$
$m_{AFP}(b,c) = 0.210$	$m_{AFP}(b,c) = 0.442$	$m_{AFP}(b,c) = 0.326$
$\kappa = 0.348,$	$\kappa = 0.348,$	$\kappa = 0.348.$

6. Comparison of three methods

This section describes four examples that compare weighted average operator (WAO), combination by compromise (CBC) and the AFP. The definitions of the WAO and the CBC can be seen in [38] and [32], respectively.

Example 4: Zadeh's example

In the literature, Zadeh's example appears in different but essentially equivalent versions of disagreeing experts. We will present Zadeh's example by the story of that a patient examined by two doctors [9, 10]. Assume that the first doctor diagnosis is that patient has either meningitis, with probability 0.99, or brain tumor, with probability 0.01. The second doctor agrees with the first one that the probability of brain tumor is 0.01, but believes that it is the probability of concussion rather than meningitis that is 0.99. So, they provide the following diagnosis:

$m_1(m) = 0.99 \quad m_1(t) = 0.01$ and $m_2(c) = 0.99 \quad m_2(t) = 0.01$.

The combination results are given in Table 5. In the columns for the WAO and AFP, it is observed that the *bbm* of the agreed *t* increases and the *bbm* of conflicted *m* and *c* decreases. However, with CBC, the *bbm* of the agreed *t* decreases. According to the AFP, the measure of conflict for Zadeh's example is *0.976*. The other methods did not give a measure of conflict.

Table 5. Comparison of methods in Zadeh's example

	m_1	m_2	WAO	CBC	AFP
m	0.99	0.00	0.494951	0.499851	0.488097
t	0.01	0.01	0.010099	0.000298	0.023805
c	0.00	0.99	0.494951	0.499851	0.488097
conflict, κ			-	-	0.976000

Example 5: Zadeh's modified example

When introducing a small amount of uncertainty in the doctor's opinions, the *bbm*s and the results of applying the methods are given in Table 6. In this case, only in the column for the AFP, it is observed that the *bbm* of the agreed *t* and the *bbm* of the agreed Θ increases. According to the AFP, the measure of conflict for Zadeh's modified example is *0.953*. When comparing with the measure of conflict of Zadeh's example, there is a decrease in the measure of conflict. This decrease is also a consistent result.

Table 6. Comparison of methods in Zadeh's modified example

	m_1	m_2	WAO	CBC	AFP
m	0.98	0.00	0.490000	0.494801	0.476522
t	0.01	0.01	0.010100	0.000398	0.023478
c	0.00	0.98	0.490000	0.494801	0.476522
m,t	0.00	0.00	0.000000	0.004900	0.000000
m,c	0.00	0.00	0.000000	0.000100	0.000000
t,c	0.00	0.00	0.000000	0.004900	0.000000
Θ	0.01	0.01	0.009900	0.000100	0.023478
conflict, κ			-	-	0.953000

Example 6: Bayesian belief functions

Let m_1 and m_2 belong to two Bayesian belief functions over the $\Theta = \{a, b, c\}$. The *bbm*s and the results of applying the methods are given in Table 7. In this example, the results are expected to be between $(0.10, 0.20)$ for singleton a, $(0.20, 0.30)$ for singleton b, and $(0.50, 0.70)$ for singleton c. The results produced by CBC are not corresponding to these expectations. The WAO produce results in expected intervals, however the results are very close to the borders. In the columns for the AFP the results are consistent with expectations. Furthermore, the results are consistent with geometric means between m_1's and m_2's *bbm*s of same singletons. According to the AFP, the measure of conflict for this example is *0.242*. It is reasonable since the existence of the quantitative conflict between *bbm*s.

Table 7. Comparison of methods in Bayesian belief functions

	m_1	m_2	WAO	CBC	AFP
a	0.10	0.20	0.105500	0.086944	0.152085
b	0.20	0.30	0.202500	0.194071	0.264078
c	0.70	0.50	0.692000	0.718984	0.583837
conflict, κ			-	-	0.242000

Example 7: Identical Bayesian belief functions

Let m_1 and m_2 belong to two identical Bayesian belief functions over the $\Theta = \{a, b, c\}$. The *bbm*s and the results of applying the methods are given in Table 8. In this case, according to the idempotency rule, the results are expected to be the same with these identical *bba*s. Only the AFP produced the expected results. According to the AFP, the measure of conflict for this example is zero. It is reasonable since there is no conflict between two bbas with respect both qualitative and quantitative.

Table 8. Comparison of methods in identical Bayesian belief functions

	m_1	m_2	WAO	CBC	AFP
a	0.10	0.10	0.056000	0.040833	0.100000
b	0.20	0.20	0.132000	0.128889	0.200000
c	0.70	0.70	0.812000	0.830278	0.700000
conflict, κ			-	-	0.000000

7. Conclusions

We propose a new consensus generator called "Analytic Fusion Process" in the framework of mathematical theory of evidence. The proposed method is a geometric mean based analytical method that operate on the individual belief functions to produce a single combined belief function when the individual belief functions comes from equally reliable and occurrence independent sources of information. This method satisfies the idempotent and commutative law. Furthermore, this method also produces a measure of conflict shows whether the original beliefs were in harmony or in conflict. Another advantage is that the measure of conflict produced by this method reflects both qualitative and quantitative conflict. Unfortunately, Analytic Fusion Process is not associative. However, an *n*-ary version of the method can be developed, and combining *n* basic belief assignments simultaneously can be a practical substitute for associativity in many real world application. The other disadvantage of the proposed method is it needs many calculation process. However, using a computer program to accomplish the whole calculation process of the methods simplifies the task. We have written such a program that accomplish the whole calculation process of the Analytic Fusion Process.

The proposed method, the weighted average operator, and the combination by compromise method are compared with each other using four examples. The results show that the Analytic Fusion Process produce results that are much more convincing than the others. Furthermore, it is seen that measure of conflict produced by Analytic Fusion Process is really reasonable. To our knowledge it is the only real single measure of conflict which measures how the original beliefs were in conflict in the framework of mathematical theory of evidence.

Acknowledgements

The authors thank the anonymous reviewer for providing valuable remarks and suggestions which helped to improve quality of the paper. We also thank Prof. Glen Meeden, Yasemin Gençtürk, and Gürkan Özel for providing language help.

References

[1] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, London, 1976.

[2] A.P. Dempster, Upper and lower probabilities induced by a multi-valued mapping, Ann. Mathematic Statistics, (1967) Vol. 38, pp. 325-339.

[3] K. Sentz, S. Ferson, Combination of Evidence in Dempster-Shafer Theory, Sandia National Laboratories Report, SAND2002-0835, California, 2002.

[4] Ph. Smets, R. Kennes, The Transferable Belief Model, Artificial Intelligence, Vol. 66 (1994), pp. 191-234.

[5] J. Kohlas, P.A. Monney, *A Mathematical Theory of Hints: An Approach to the Dempster-Shafer Theory of Evidence*, Vol. 425 of Lecture Notes in Economics and Mathematical Systems, Springer –Verlag, 1995.

[6] R. Haenni, J. Kohlas, N. Lehman, Probabilistic argumentation systems, J. Kohlas, S. Moral (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Vol. 5 of Algorithms for Uncertainty and Defeasible Reasoning, pp. 221-288, Kluwer Academic Publishers, 2000.

[7] F. Smarandache, J. Dezert (Eds.), *Advances and Applications of DSmT for Information Fusion, Smarandache*, Vol 2. American Research Press, Rehoboth, 2006.

[8] M. Detyniecki, Fundamentals on Aggregation Operators, http://www.cs.berkeley.edu/~marcin/agop.pdf , 2001.

[9] L.A. Zadeh, On the validty of Dempster's rule of combination of evidence. Technical Report 79/24, University of California, Berkely, 1979.

[10] L.A. Zadeh, A mathematical theory of evidence (book review), AI Mag. 5(3) (1984), pp. 81-83.

[11] R. Haenni, Shedding new light on Zadeh's Criticism of Dempster's rule of Combination, Proceedings of Information Fusion 2005, Philadelphia, July 2005.

[12] H.Y. Hau, R.L. Kashyap, On the Robustness of Dempster' s Rule of Combination, IEEE International Workshop on Tools for Artificial Intelligence, Fairfax, VA, Oct. (1989), pp. 578-582.

[13] E. Lefevre, O. Colot, P. Vannoorenberghe, Belief Function Combination and Conflict Management, Information Fusion, Vol. 3 (2002), pp. 149-162.

[14] C.K. Murphy, Combining Belief Functions when Evidence Conflicts, Decision Support Systems, Vol. 29 (2000), pp. 1-9.

[15] Ph. Smets, Analyzing the Combination of Conflicting Belief Functions, Information Fusion, Vol. 8, (2007), pp. 387-412.

[16] F. Voorbraak, On the justification of Dempster's rule of combination, Artificial Intelligence, 48, (1991), pp. 171-197.

[17] R. Haenni, Are alternatives to Dempster's rule of combination real alternatives? Comments on "About the belief function combination and the conflict management problem"-Lefevre et al, Information Fusion, Vol. 3 (2002), pp. 237-239.

[18] F. Campos, S. Cavalcante, An Extended Approach for Dempster-Shafer Theory, IEEE International Conference on Information Reuse and Integration (IRI 2003), Las Vegas, USA, October, 2003.

[19] M. Daniel, Associativity in Combination of belief functions; a derivation of minC combination, Soft Computing, 7(5), (2003), pp. 288–296.

[20] D. Dubois, H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, Computational Intelligence, Volume 4, Issue 3, (1988), pp. 244-264.

[21] A. Josang, The consensus operator for combining beliefs, Artificial Intelligence, 141 (2002), pp. 157-170.

[22] R. R. Yager, On the Dempster-Shafer framework and new combination rules, Information Sciences, (1987), 41 (2), p.93-137.

[23] Fuyuan Xiao, Evidence combination based on prospect theory for multi-sensor data fusion, ISA Transactions, Volume 106, (2020), Pages 253-261,

[24] Jiang W., Zhan J., A modified combination rule in generalized evidence theory, Appl. Intell., (2017), 46 (3) pp. 630-640.

[25] Jian W., Kuoyuan Q., Zhiyong Z., An improvement for combination rule in evidence theory, Future Generation Computer Systems, (2019), Volume 91, Pages 1-9,

[26] Wenjun M., Yuncheng J., Xudong L., A flexible rule for evidential combination in Dempster–Shafer theory of evidence, Applied Soft Computing, (2019), Volume 85.

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

[28] R. Haenni, S. Hartmann, Modeling partially reliable information source: A general approach based on Dempster-Shafer theory, Information Fusion, Vol. 7 (2006), pp. 361-379.

[29] M.E.G.V. Cattaneo, Combining belief functions issued from dependent sources, In: J.M. Bernard, T. Seidenfeld, M. Zaffalon (Eds.), Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications (ISIPTA'03), 2003, Carleton Scientific, Lugano, Switzerland. pp. 133-147.

[30] P.A. Monney, M. Chan, Modelling dependency in Dempster-Shafer theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 15, No. 1 (2007), pp. 93-114.

[31] T. Denœux, Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence, Artificial Intelligence, 172, (2008), 234-264.

[32] K. Yamada, A new combination of evidence based on compromise, Fuzzy Sets and Systems, 159, (2008), 1689-1708.

[33] S. Ferson, V. Kreinovich, L. Ginzburg, D.S. Myers, K. Sentz, Constructing Probability Boxes and Dempster-Shafer Structures, Sandia National Laboratories Report, SAND2002-4015, California, 2003.

[34] F. Smarandache, J. Dezert, Proportional Conflict Redistribution Rules for Information Fusion, arXiv Archives, Los Alamos National Laboratory; the Abstract and the whole paper are available at http://arxiv.org/PS_cache/cs/pdf/0408/0408064.pdf , 2005.

[35] B.B. Yaghlane, P. Smets, K. Mellouli, Belief function independence: I. The marginal case, International Journal of Approximate Reasoning, Vol. 29, (2002), pp. 47-70.

[36] B.B. Yaghlane, P. Smets, K. Mellouli, Belief function independence: II. The conditional case, International Journal of Approximate Reasoning, Vol. 31, (2002), pp. 31-75.

[37] R.T. Clemen, R.L. Winkler, Combining Probability Distributions From Experts in Risk Analysis, Risk Analysis, (1999) Vol. 19, No. 2 pp. 187-203.

[38] A. Josang, M. Daniel, P. Vannoorenberghe, Strategies for combining conflicting dogmatic beliefs, Applications of plausable, paradoxical, and neutrosophical reasoning for information fusion (The sixth international conference on information fusion), Cairns, Queensland, Australia, July 8-11, 2003.

[39] W. Liu, Analyzing the degree of conflict among belief functions, Artificial Intelligence, Vol. 170 (11) (2006), pp. 909-924.

[40] C. Genest, J.V. Zidek, Combining probability distributions: A critique and a annotated bibliography, Statistical Science, Vol. 1, No. 1, (1986), pp.114-148.

[41] W.L. Winston, *Operations Research Aplications and Algorithms*, Duxbury Press, Belmont, 1994.

[42] L.A. Wolsey, *Integer Programming*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience Publication, NewYork, 1998.