Multiple-Attribute Decision Making Under Uncertainty: The Evidential Reasoning Approach Revisited

Van-Nam Huynh, Yoshiteru Nakamori, *Member, IEEE*, Tu-Bao Ho, and Tetsuya Murai, *Member, IEEE*

***Abstract*—In multiple-attribute decision making (MADM) problems, one often needs to deal with decision information with uncertainty. During the last decade, Yang and Singh (1994) have proposed and developed an evidential reasoning (ER) approach to deal with such MADM problems. Essentially, this approach is based on an evaluation analysis model and Dempster's rule of combination in the Dempster–Shafer (D–S) theory of evidence. This paper reanalyzes the ER approach explicitly in terms of D–S theory and then proposes a general scheme of attribute aggregation in MADM under uncertainty. In the spirit of such a reanalysis, previous ER algorithms are reviewed and two other aggregation schemes are discussed. Theoretically, it is shown that new aggregation schemes also satisfy the synthesis axioms, which have been recently proposed by Yang and Xu (2002) for which any rational aggregation process should grant. A numerical example traditionally examined in published sources on the ER approach is used to illustrate the discussed techniques.**

***Index Terms*—Assessment, evidence combination, evidential reasoning (ER), multiple-attribute decision making (MADM), uncertainty.**

I. Introduction

PRACTICALLY, decision makers are often required to choose between several alternatives or options, where each option exhibits a range of attributes of both quantitative and qualitative in nature. A decision may not be properly made without fully taking into account all attributes concerned [2], [13], [21], [27], [33], [38]. In addition, in many multiple-attribute decision making (MADM) problems, one also frequently needs to deal with decision knowledge represented in forms of both qualitative and quantitative information with uncertainty.

So far, many attempts have been made to integrate techniques from artificial intelligence (AI) and operational research (OR) for handling uncertain information, e.g., [1], [3], [5], [8], [14], [16], [20], [29], [32], and [40]. During the last decade or so, an evidential reasoning (ER) approach has been proposed and developed for MADM under uncertainty in [33], [34], and [36]–[38]. Essentially, this approach is based on an evaluation analysis model [42] and the evidence combination rule of the Dempster–Shafer (D–S) theory [25], which has in turn been one of the major techniques for dealing with uncertainty in AI [9]. The ER approach has been applied to a range of MADM problems in engineering and management, including motorcycle assessment [34], general cargo ship design [23], system safety analysis and synthesis [28], and retro-fit ferry design [35], among others.

The kernel of the ER approach is an ER algorithm developed on the basis of a multiattribute evaluation framework and Dempster's rule of combination in the D–S theory of evidence [33]. Basically, the algorithm makes use of Dempster's rule of combination to aggregate attributes of a multilevel structure. Due to a need of developing theoretically sound methods and tools for dealing with MADM problems under uncertainty, recently, Yang and Xu [38] have proposed a system of four synthesis axioms within the ER assessment framework with which a rational aggregation process needs to satisfy. It has also been shown that the original ER algorithm only satisfies these axioms approximately. At the same time, exactly guided by the aim, the authors have proposed a new ER algorithm that satisfies all the synthesis axioms precisely.

Interestingly enough, the D–S theory of evidence on one hand allows us to coarse or refine the data by changing to a higher or lower level of granularity (or attribute in the context of a multilevel structure) accompanied with a powerful evidence combination rule. This is an essential feature for multiple-attribute assessment systems based on a multilevel structure of attributes. On the other hand, one of the major advantages of the D–S theory over conventional probability is that it provides a straightforward way of quantifying ignorance and is therefore a suitable framework for handling incomplete uncertain information. This is especially important and useful for dealing with uncertain subjective judgments when multiple basic attributes (also called factors) need to be considered simultaneously [33].

It would be worth mentioning that the underlying basis of using Dempster's rule of combination is the independent assumption of information sources to be combined. Here, the authors would like to distinguish between the independent

Manuscript received August 9, 2004; revised February 5, 2005. This paper was supported by a Monbukagakusho 21st COE Program and the Japan Advanced Institute of Science and Technology (JAIST) International Joint Research Grant. This paper is a substantially expanded and revised version of the paper presented at ASIAN'04—Ninth Asian Computing Science Conference. This paper was recommended by Associate Editor K. W. Hipel.

V.-N. Huynh, Y. Nakamori, and T.-B. Ho are with the Japan Advanced Institute of Science and Technology, Tatsunokuchi, Ishikawa 923-1292, Japan (e-mail: huynh@jaist.ac.jp; nakamori@jaist.ac.jp; bao@jaist.ac.jp).

T. Murai is with the Graduate School of Information Science and Technology, Hokkaido University, Kita-ku, Sapporo 060-0814, Japan (e-mail: murahiko@main.eng.hokudai.ac.jp).

Digital Object Identifier 10.1109/TSMCA.2005.855778

assumption among uncertain assessments of sub-attributes used to derive the assessment of an aggregated attribute in a multilevel structure of attributes and the assumption of utility independence imposed on assessment grades of a decision problem (see Sections IV-F and V). Regarding the former, in practice, it is rarely verified and difficult to check (see, e.g., [7] and [15]). In other words, in situations of multiple-attribute assessment based on a multilevel structure of attributes, the independent assumption of the basic attributes' uncertain evaluations is not always appropriate. Another important issue concerning the rule is that it may yield counterintuitive results especially when a high conflict between information sources to be combined arises. This problem of completely ignoring conflict caused by normalization in Dempster's rule was originally pointed out in [41] and discussed in more detail in [7]. Consequently, this has highly motivated researchers to propose a number of other combination rules in the literature to address the problem, e.g., [7], [29], and [31] (see [24] for a recent survey).

This paper deals with the attribute aggregation problem in the ER approach to MADM under uncertainty developed in, e.g., [33] and [38]. First, the authors reanalyze the previous ER approach in terms of the D–S theory so that the attribute aggregation problem in MADM under uncertainty can be generally formulated as a problem of evidence combination. Then, they propose several new aggregation schemes and simultaneously examine their theoretical features. For the purpose of the present paper, only qualitative attributes of an MADM problem with uncertainty were taken into account, though quantitative attributes would be also included in a similar way as considered in [33] and [34].

To proceed, it is first necessary to briefly recall the basic notions on the MADM problem with uncertainty, the basic evaluation framework, and the D–S theory of evidence. This is undertaken in Section II and followed in Section III by a discussion of the ER approach to MADM under uncertainty proposed previously. Section IV then explores the attribute aggregation problem in detail, and Section V examines a motorcycle performance assessment problem taken from [38]. Finally, Section IV presents some concluding remarks.

II. Background

A. Problem Description

This subsection describes an MADM problem with uncertainty through a tutorial example taken from [38]. As mentioned above, for the purpose of this paper, only qualitative attributes of the problem are taken into account. For more details, see [33] and [34].

To subjectively evaluate qualitative attributes (or features) of alternatives (or options), a set of evaluation grades may first be supplied as

$$\mathcal{H} = \{H_1, \ldots, H_n, \ldots, H_N\}$$

where H_ns are called evaluation grades to which the state of a qualitative attribute y may be evaluated. That is, $\mathcal{H}$ provides a complete set of distinct standards for assessing qualitative attributes in question. In accomplishing this objective, an important aspect to analyze is the level of discrimination among different countings of evaluation grades or, in other words, the cardinality of the set $\mathcal{H}$ used to express the information. The cardinality of the set must be small enough so as not to impose useless precision on the users and must be rich enough in order to allow discrimination of the assessments in a limited number of degrees. According to the observation in [18], in practice, human beings can reasonably manage to bear in mind seven or so terms. In addition, although different attributes may have different sets of evaluation grades, for the sake of simplicity, in this paper, we assume the same set $\mathcal{H}$ for all attributes of concern. Further, without loss of generality, it is assumed that H_{n+1} is preferred to H_n.

Let us turn to a problem of motorcycle evaluation [10]. To evaluate the quality of the operation of a motorcycle, the set of distinct evaluation grades is defined as

$$\mathcal{H} = \{poor\ (H_1), indifferent\ (H_2)\\ average\ (H_3), good\ (H_4), excellent\ (H_5)\}. \quad (1)$$

Because operation is a general technical concept and is not easy to evaluate directly, it needs to be decomposed into detailed concepts such as $handling, transmission,$ and $brakes.$ Again, if a detailed concept is still too general to assess directly, it may be further decomposed into more detailed concepts. For example, the concept of $brakes$ is measured by $stopping\ power, braking\ stability,$ and $feel\ at\ control,$ which can probably be directly evaluated by an expert and therefore referred to as basic attributes (or basic factors).

Generally, a qualitative attribute y may be evaluated through a hierarchical structure of its sub-attributes. For instance, the hierarchy for evaluation of the $operation$ of a motorcycle is depicted as in Fig. 1.

In the evaluation of qualitative attributes, judgments could be uncertain. For example, in the problem of evaluating different types of motorcycles, the following types of uncertain subjective judgments for the brakes of a motorcycle, say "Yamaha," was frequently used [10], [38].

1) Its $stopping\ power$ is $average$ with a confidence degree of 0.3 and it is $good$ with a confidence degree of 0.6.
2) Its $braking\ stability$ is $good$ with a confidence degree of 1.
3) Its $feel\ at\ control$ is evaluated to be $good$ with a confidence degree of 0.5 and to be $excellent$ with a confidence degree of 0.5.

In the above statements, the confidence degrees represent the uncertainty in the evaluation. Note that the total confidence degree in each statement may be smaller than 1 as the case of the first statement. This may be due to the incompleteness of available information.

In a similar fashion, all basic attributes in question could be evaluated. The problem now is how to generate an overall assessment of the operation of a motorcycle by aggregating all the uncertain judgments of its basic attributes in a rational way. The ER approach developed in [33], [34], and [38] has provided a means based on Dempster's rule of combination for dealing with such an aggregation problem.

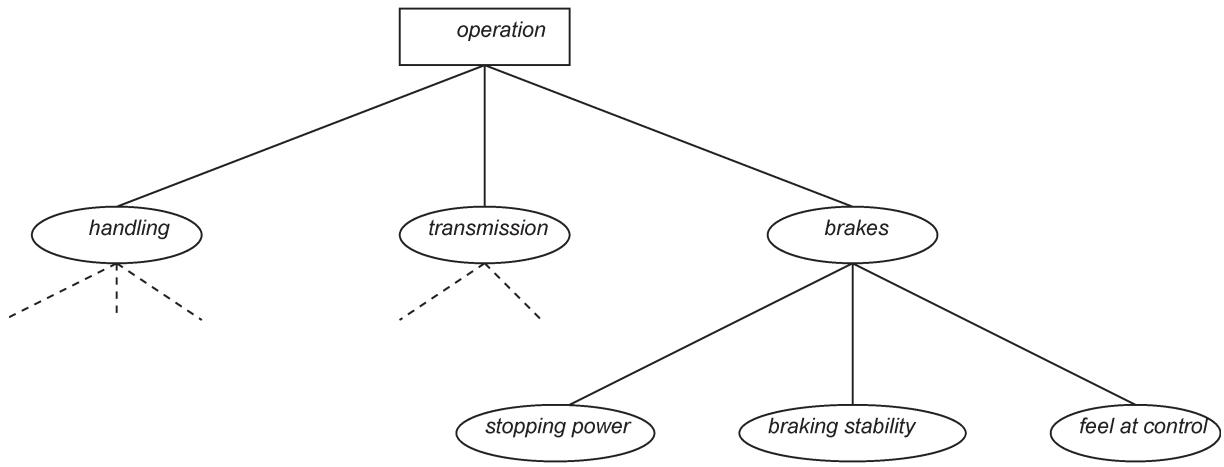


Fig. 1. Evaluation hierarchy for operation [38].

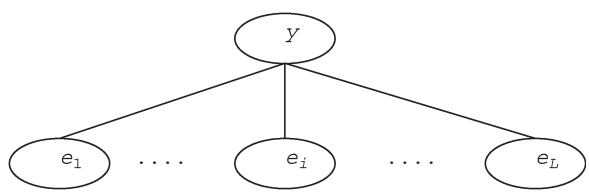


Fig. 2. Two-level hierarchy.

B. Evaluation Analysis Model

The evaluation analysis model was proposed in [42] to represent uncertain subjective judgments, such as statements specified in the preceding subsection, in a hierarchical structure of attributes.

To begin with, let us suppose a simple hierarchical structure consisting of two levels with a general attribute, denoted by y, at the top level and a finite set E of its basic attributes at the bottom level (graphically shown in Fig. 2). Let

$$E = \{e_1, \ldots, e_i, \ldots, e_L\}$$

and assume the weights of basic attributes are given by $W = (w_1, \ldots, w_i, \ldots, w_L)$, where w_i is the relative weight of the ith basic attribute (e_i) with $0 \leq w_i \leq 1$. Attribute weights essentially play an important role in multiattribute decision models. Because the elicitation of weights can be difficult, several methods have been proposed for reducing the burden of the process [19].

Given the set of evaluation grades

$$\mathcal{H} = \{H_1, \ldots, H_n, \ldots, H_N\}$$

designed as distinct standards for assessing an attribute, then an assessment for e_i of an alternative can be mathematically represented in terms of the distribution [38]

$$S(e_i) = \{(H_n, \beta_{n,i}) | n = 1, \ldots, N\}, \qquad \text{for } i = 1, \ldots, L \tag{2}$$

where $\beta_{n,i}$ denotes a degree of belief satisfying $\beta_{n,i} \geq 0$, and $\sum_{n=1}^{N} \beta_{n,i} \leq 1$. An assessment $S(e_i)$ is called complete (respectively, incomplete) if $\sum_{n=1}^{N} \beta_{n,i} = 1$ (respectively, $\sum_{n=1}^{N} \beta_{n,i} < 1$).

For example, the three assessments 1)–3) given in the preceding subsection can be represented in the form of distributions defined by (2) as

$$S(stopping\ power) = \{(H_3, 0.3), (H_4, 0.6)\}$$
$$S(braking\ stability) = \{(H_4, 1)\}$$
$$S(feel\ at\ control) = \{(H_4, 0.5), (H_5, 0.5)\}$$

where only grades with nonzero degrees of belief are listed in the distributions.

Let us denote by β_n the degree of belief to which the general attribute y is assessed to the evaluation grade of H_n. The problem now is to generate β_n, for $n = 1, \ldots, N$, by combining the assessments for all associated basic attributes e_i $(i = 1, \ldots, L)$ as given in (2). However, before continuing the discussion, it is necessary to briefly review the basis of the D–S theory of evidence in the next subsection.

C. D–S Theory of Evidence

In the D–S theory, a problem domain is represented by a finite set Θ of mutually exclusive and exhaustive hypotheses called the frame of discernment [25]. In the standard probability framework, all elements in Θ are assigned a probability, and when the degree of support for an event is known, the remainder of the support is automatically assigned to the negation of the event. On the other hand, in the D–S theory, mass assignments are carried out for events as they know, and committing support for an event does not necessarily imply that the remaining support is committed to its negation. Formally, a basic probability assignment (BPA, for short) is a function $m : 2^{\Theta} \to [0, 1]$, verifying

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{A \in 2^{\Theta}} m(A) = 1.$$

The quantity $m(A)$ can be interpreted as a measure of the belief that is committed exactly to A, given the available evidence. A subset $A \in 2^{\Theta}$ with $m(A) > 0$ is called a focal element of m.

A BPA m is said to be vacuous if $m(\Theta) = 1$ and $m(A) = 0$ for all $A \neq \Theta$.

Two evidential functions derived from the BPA are the belief function Bel and the plausibility function Pl defined as

$$\mathrm{Bel}(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad \text{and} \quad \mathrm{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B).$$

The difference between $m(A)$ and $\mathrm{Bel}(A)$ is that while $m(A)$ is our belief committed to the subset A excluding any of its proper subsets, $\mathrm{Bel}(A)$ is our degree of belief in A as well as all of its subsets. Consequently, $\mathrm{Pl}(A)$ represents the degree to which the evidence fails to refute A. Note that all the three functions are in a one-to-one correspondence with each other.

Two useful operations that play a central role in the manipulation of belief functions are discounting and Dempster's rule of combination [25]. The discounting operation is used when a source of information provides a BPA m, but one knows that this source has probability α of reliability. Then, one may adopt $(1-\alpha)$ as one's discount rate, which results in a new BPA m^α defined by

$$m^\alpha(A) = \alpha m(A), \qquad \text{for any } A \subset \Theta \tag{3}$$

$$m^\alpha(\Theta) = (1-\alpha) + \alpha m(\Theta). \tag{4}$$

Consider now two pieces of evidence on the same frame Θ represented by two BPAs m_1 and m_2. Dempster's rule of combination is then used to generate a new BPA, denoted by $(m_1 \oplus m_2)$ (also called the orthogonal sum of m_1 and m_2), defined as

$$(m_1 \oplus m_2)(\emptyset) = 0$$

$$(m_1 \oplus m_2)(A) = \frac{1}{1-k} \sum_{B \cap C = A} m_1(B) m_2(C) \tag{5}$$

where

$$\kappa = \sum_{B \cap C = \emptyset} m_1(B) m_2(C). \tag{6}$$

Note that the orthogonal sum combination is only applicable to such two BPAs that meet the condition $\kappa < 1$.

As we will partially see in the following sections, these two operations essentially play an important role in the ER approach to MADM under uncertainty developed in, e.g., [33], [34], and [38], though the discounting operation has not been mentioned explicitly in these published sources.

III. ER Approach

Let us return to the two-level hierarchical structure with a general attribute y at the top level and a finite set $E = \{e_1, \ldots, e_i, \ldots, e_L\}$ of its basic attributes at the bottom level. Let us be given weights w_i $(i = 1, \ldots, L)$ of basic attributes e_i $(i = 1, \ldots, L)$, respectively. Denote by β_n the degree of belief to which the general attribute y is assessed to the evaluation grade of H_n, for $n = 1, \ldots, N$.

A. Original ER Algorithm

The original ER algorithm proposed in [33] has been used for the purpose of obtaining β_n $(n = 1, \ldots, N)$ by aggregating the assessments of basic attributes given in (2). The summary of the algorithm in this subsection is taken from [38].

Given the assessment $S(e_i)$ of a basic attribute e_i $(i = 1, \ldots, L)$, let $m_{n,i}$ be a basic probability mass representing the belief degree to which the basic attribute e_i supports the hypothesis that the attribute y is assessed to the evaluation grade H_n. Let $m_{\mathcal{H},i}$ be the remaining probability mass unassigned to any individual grade after all the N grades have been considered for assessing the general attribute y as far as e_i is concerned. These quantities are defined as

$$m_{n,i} = w_i \beta_{n,i}, \qquad \text{for } n = 1, \ldots, N \tag{7}$$

$$m_{\mathcal{H},i} = 1 - \sum_{n=1}^{N} m_{n,i} = 1 - w_i \sum_{n=1}^{N} \beta_{n,i}. \tag{8}$$

Let $E_{I(i)} = \{e_1, \ldots, e_i\}$ be the subset of first i basic attributes. Let $m_{n,I(i)}$ be a probability mass defined as the belief degree to which all the basic attributes in $E_{I(i)}$ support the hypothesis that y is assessed to H_n. Let $m_{\mathcal{H},I(i)}$ be the remaining probability mass unassigned to individual grades after all the basic attributes in $E_{I(i)}$ have been assessed. The quantities $m_{n,I(i)}$ and $m_{\mathcal{H},I(i)}$ can be generated by combining the basic probability masses $m_{n,j}$ and $m_{\mathcal{H},j}$ for all $n = 1, \ldots, N$ and $j = 1, \ldots, i$.

With these notations, the key step in the original ER algorithm is to inductively calculate $m_{n,I(i+1)}$ and $m_{\mathcal{H},I(i+1)}$ as

$$m_{n,I(i+1)} = K_{I(i+1)} \left(m_{n,I(i)} m_{n,i+1} + m_{n,I(i)} m_{\mathcal{H},i+1} + m_{\mathcal{H},I(i)} m_{n,i+1} \right) \tag{9}$$

$$m_{\mathcal{H},I(i+1)} = K_{I(i+1)} \left(m_{\mathcal{H},I(i)} m_{\mathcal{H},i+1} \right) \tag{10}$$

for $n = 1, \ldots, N$, $i = 1, \ldots, L-1$, and $K_{I(i+1)}$ is a normalizing factor defined by

$$K_{I(i+1)} = \left[1 - \sum_{t=1}^{N} \sum_{\substack{j=1 \\ j \neq t}}^{N} m_{t,I(i)} m_{j,i+1} \right]^{-1}. \tag{11}$$

Then, we obtain

$$\beta_n = m_{n,I(L)}, \qquad \text{for } n = 1, \ldots, N$$

$$\beta_{\mathcal{H}} = m_{\mathcal{H},I(L)} = 1 - \sum_{n=1}^{N} \beta_n. \tag{12}$$

B. Synthesis Axioms and Modified ER Algorithm

Inclined to developing theoretically sound methods and tools for dealing with MADM problems under uncertainty, Yang and Xu [38] have recently proposed a system of four synthesis axioms within the ER assessment framework with

which a rational aggregation process needs to satisfy. These axioms are symbolically stated as follows.

Axiom 1 Independence. If $\beta_{n,i} = 0$ for all $i = 1, \ldots, L$, then $\beta_n = 0$.

Axiom 2 Consensus. If $\beta_{k,i} = 1$ and $\beta_{n,i} = 0$ for all $i = 1, \ldots, L$ and $n = 1, \ldots, N$, $n \neq k$, then $\beta_k = 1$ and $\beta_n = 0$ for $n = 1, \ldots, N$, $n \neq k$.

Axiom 3 Completeness. Assume $\mathcal{H}^+ \subset \mathcal{H}$ and denote $J^+ = \{n | H_n \in \mathcal{H}^+\}$. If $\beta_{n,i} > 0$ for $n \in J^+$ and $\sum_{n \in J^+} \beta_{n,i} = 1$, for all $i = 1, \ldots, L$, then $\beta_n > 0$ for $n \in J^+$ and $\sum_{n \in J^+} \beta_n = 1$ as well.

Axiom 4 Incompleteness. If there exists $i \in \{1, \ldots, L\}$ such that $\sum_{n=1}^{N} \beta_{n,i} < 1$, then $\sum_{n=1}^{N} \beta_n < 1$.

It is easily seen from (9)–(12) that the original ER algorithm naturally follows the independency axiom. Concerning the second axiom, the following theorem is due to Yang and Xu [38].

Theorem 1: If β_n and $\beta_{\mathcal{H}}$ are calculated using (12), then the consensus axiom holds if and only if

$$\prod_{i=1}^{L} (1 - w_i) = 0. \tag{13}$$

Note that the only constraint imposed on the weights w_i $(i = 1, \ldots, L)$ in the ER approach is $0 \leq w_i \leq 1$. By Theorem 1, it implies that if $w_i = 1$, then e_i dominates the assessment of y, i.e., other basic attributes with smaller weights play no role in the assessment. To resolve this dilemma, the following scheme for weight normalization has been considered in [33]–[35]:

$$\overline{w}_i = \alpha \frac{w_i}{\max\limits_{i=1,\ldots,L} 1\{w_i\}} \tag{14}$$

and α is a constant determined by satisfying

$$\prod_{i=1}^{L} \left(1 - \alpha \frac{w_i}{\max\limits_{i=1,\ldots,L} \{w_i\}}\right) \leq \delta$$

where δ is a small constant representing the degree of approximation in aggregation. By considering normalized weights $\overline{w}_i$'s instead of w_i's, it means the consensus axiom could only be satisfied approximately. However, this weight normalization has still another shortcoming, that the most important attribute may play a dominating role in the assessment of y. Further, it has also been shown in [38] that the original ER algorithm does not satisfy the completeness axiom.

Under such a consideration, Yang and Xu [38] proposed a new ER algorithm that satisfies all the synthesis axioms. Its main features are summarized as follows.

1) Weight normalization. In the new ER algorithm, the weights w_i $(i = 1, \ldots, L)$ of basic attributes are normalized such that $0 \leq w_i \leq 1$, and

$$\sum_{i=1}^{L} w_i = 1. \tag{15}$$

2) Aggregation process. First, the probability mass $m_{\mathcal{H},i}$ given in (8) is decomposed into two parts: $m_{\mathcal{H},i} = \tilde{m}_{\mathcal{H},i} + \overline{m}_{\mathcal{H},i}$, where

$$\overline{m}_{\mathcal{H},i} = 1 - w_i \quad \text{and} \quad \tilde{m}_{\mathcal{H},i} = w_i \left(1 - \sum_{n=1}^{N} \beta_{n,i}\right). \tag{16}$$

Then, with the notations as in the preceding section, the process of aggregating the first i assessments with the $(i+1)$th assessment is recursively carried out as

$$m_{n,I(i+1)} = K_{I(i+1)} \left[m_{n,I(i)} m_{n,i+1} + m_{n,I(i)} m_{\mathcal{H},i+1} + m_{\mathcal{H},I(i)} m_{n,i+1}\right] \tag{17}$$

$$m_{\mathcal{H},I(i)} = \tilde{m}_{\mathcal{H},I(i)} + \overline{m}_{\mathcal{H},I(i)}, \qquad n = 1, \ldots, N$$

$$\tilde{m}_{\mathcal{H},I(i+1)} = K_{I(i+1)} \left[\tilde{m}_{\mathcal{H},I(i)} \tilde{m}_{\mathcal{H},i+1} + \overline{m}_{\mathcal{H},I(i)} \tilde{m}_{\mathcal{H},i+1} + \tilde{m}_{\mathcal{H},I(i)} \overline{m}_{\mathcal{H},i+1}\right] \tag{18}$$

$$\overline{m}_{\mathcal{H},I(i+1)} = K_{I(i+1)} \left[\overline{m}_{\mathcal{H},I(i)} + \overline{m}_{\mathcal{H},i+1}\right] \tag{19}$$

where $K_{I(i+1)}$ is defined as in (11).

For assigning the assessment $S(y)$ for the general attribute y, after all L assessments of basic attributes have been aggregated, the algorithm finally defines

$$\beta_n = \frac{m_{n,I(L)}}{1 - \overline{m}_{\mathcal{H},I(L)}}, \quad \text{for } n = 1, \ldots, N \tag{20}$$

$$\beta_{\mathcal{H}} = \frac{\tilde{m}_{\mathcal{H},I(L)}}{1 - \overline{m}_{\mathcal{H},I(L)}} \tag{21}$$

and then

$$S(y) = \{(H_n, \beta_n),\ n = 1, \ldots, N\}. \tag{22}$$

The following theorems that are taken for granted to develop the new ER algorithm above are due to Yang and Xu [38].

Theorem 2: The degrees of belief defined by (20) and (21) satisfy

$$0 \leq \beta_n, \beta_{\mathcal{H}} \leq 1, \quad n = 1, \ldots, N$$

$$\sum_{n=1}^{N} \beta_n + \beta_{\mathcal{H}} = 1.$$

Theorem 3: If $\beta_{k,i} = 1$ and $\beta_{n,i} = 0$ for all $n = 1, \ldots, N$ with $n \neq k$ and $i = 1, \ldots, L$, then $\beta_k = 1$ and $\beta_n = 0$ for all $n = 1, \ldots, N$ with $n \neq k$ and $\beta_{\mathcal{H}} = 0$.

Theorem 4: Let $\mathcal{H}^+ \subset \mathcal{H}$ and $\mathcal{H}^- = \mathcal{H} \setminus \mathcal{H}^+$ and denote $J^+ = \{n | H_n \in \mathcal{H}^+\}$ and $J^- = \{j | H_j \in \mathcal{H}^-\}$. If $\beta_{n,i} > 0$ $(n \in J^+)$ and $\sum_{n \in J^+} \beta_{n,i} = 1$ and $\beta_{j,i} = 0$ $(j \in J^-)$ for all $i = 1, \ldots, L$, then $\sum_{n \in J^+} \beta_n = 1$ and $\beta_j = 0$ $(j \in J^-)$.

Theorem 5: Assume that $0 < w_i < 1$ for all $i = 1, \ldots, L$. If there exists an i such that $\sum_{n=1}^{N} \beta_{n,i} < 1$, then $\beta_{\mathcal{H}} > 0$.

In [38], the authors have given direct proofs of these theorems, which are somehow complicated. In the next section, by analyzing the ER approach in terms of the D–S theory, we however show that these theorems follow quite naturally.

IV. Reanalysis of the ER Approach

Let us reexamine the available information given to an assessment problem in the two-level hierarchical structure, as depicted in Fig. 2:

1) the assessments $S(e_i)$ for basic attributes e_i $(i = 1, \ldots, L)$, and
2) the weights w_i of the basic attributes e_i $(i = 1, \ldots, L)$.

Given the assessment $S(e_i)$ of a basic attribute e_i $(i = 1, \ldots, L)$, we now define a corresponding BPA, denoted by m_i, which quantifies the belief about the performance of e_i for any $H \subseteq \mathcal{H}$ as

$$m_i(H) = \begin{cases} \beta_{n,i}, & \text{if } H = \{H_n\} \\ \left(1 - \sum\limits_{n=1}^{N} \beta_{n,i}\right), & \text{if } H = \mathcal{H} \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

For the sake of simplicity, we will write $m_i(H_n)$ instead of $m_i(\{H_n\})$ as usual. The quantity $m_i(H_n)$ represents the belief degree that supports the hypothesis that e_i is assessed to the evaluation grade H_n, while $m_i(\mathcal{H})$ is the remaining probability mass unassigned to any individual grade after all evaluation grades have been considered for assessing e_i. If $S(e_i)$ is a complete assessment, m_i is a probability distribution, i.e., $m_i(\mathcal{H}) = 0$. Otherwise, $m_i(\mathcal{H})$ quantifies the ignorance.

As such, with L basic attributes e_i, we obtain L corresponding BPAs m_i as quantified beliefs of the assessments for basic attributes. The problem now is how to generate an assessment for y, i.e., $S(y)$, represented by a BPA m, from m_i and w_i $(i = 1, \ldots, L)$. Formally, we aim at obtaining a BPA m that combines all m_i's while taking weights w_i's into account in the general form of

$$m = \bigoplus_{i=1}^{L} (w_i \otimes m_i) \tag{24}$$

where $\otimes$ is a product-type operation, and $\oplus$ is a sum-type operation in general.

With this general rule of weighted combination, by applying different particular operations for $\otimes$ and $\oplus$, we may have different aggregation schemes for obtaining the BPA m representing the generated assessment $S(y)$. However, before exploring any new aggregation schemes, we first interestingly reinterpret the original ER approach in the spirit of the new formulation.

A. Discounting-and-Orthogonal Sum Scheme

Let us first consider $\otimes$ as the discounting operation and $\oplus$ as the orthogonal sum in the D–S theory. Then, for each $i = 1, \ldots, L$, we have $(w_i \otimes m_i)$ as a BPA [refer to (3) and (4)] defined by (25) shown at the bottom of the page, for any $H \subseteq \mathcal{H}$ and $n = 1, \ldots, N$.

With this formulation, we consider each m_i as the belief quantified from the information source $S(e_i)$ and the weight w_i as the "degree of trust" of $S(e_i)$ supporting the assessment of y as a whole. As mentioned in [25], an obvious way to use discounting with Dempster's rule of combination is to discount all BPAs m_i $(i = 1, \ldots, L)$ at corresponding rates $(1 - w_i)$ $(i = 1, \ldots, L)$ before combining them.

Thus, Dempster's rule of combination now allows us to combine BPAs $m_i^{w_i}$ $(i = 1, \ldots, L)$ under the independent assumption of information sources for generating the BPA m for the assessment of y. Namely

$$m = \bigoplus_{i=1}^{L} m_i^{w_i} \tag{26}$$

where, with an abuse of the notation, $\oplus$ stands for the orthogonal sum.

At this juncture, we can see that the aggregation processes in the original ER approach reviewed above formally follow this discounting-and-orthogonal sum scheme. In other words, despite the differences in interpretation, the two approaches lead essentially to mathematically equivalent formulations. In addition, it is of interest to note that, by definition, in this aggregation scheme, it would not be necessary to require the process of weight normalization satisfying the constraint

$$\sum_{i=1}^{L} w_i = 1.$$

This remark implies that the aggregation scheme could be applied to different normalization strategies only satisfying $0 \le w_i \le 1$. Furthermore, in our opinion, in the hierarchical representation of an aggregated attribute y as shown in Fig 2, we have at least two possible interpretations. In the first one, the performance of an alternative regarding the attribute y depends on one regarding all the sub-attributes simultaneously; the second interpretation is the performance of an alternative regarding the attribute y depending on that of it regarding only one of the sub-attributes, but we do not know which one actually plays the role. While the mutual and exclusive assumption of information sources is appropriate for the latter interpretation, it may not be the case in the former. Therefore, by relaxing the constraint on weights, we may be able to avoid the mutual and exclusive assumption of information sources supporting the assessment for y in some situations.

$$(w_i \otimes m_i)(H) \triangleq m_i^{w_i}(H) = \begin{cases} w_i \beta_{n,i}, & \text{if } H = \{H_n\} \\ (1 - w_i) + w_i \left(1 - \sum\limits_{n=1}^{N} \beta_{n,i}\right) = 1 - w_i \sum\limits_{n=1}^{N} \beta_{n,i}, & \text{if } H = \mathcal{H} \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

It would be worth noting that two BPAs $m_i^{w_i}$ and $m_j^{w_j}$ are combinable, i.e., $(m_i^{w_i} \oplus m_j^{w_j})$ does exist, if and only if

$$\kappa = \sum_{t=1}^{N} \sum_{\substack{n=1 \\ n \neq t}}^{N} m_i^{w_i}(H_n) m_j^{w_j}(H_t) < 1.$$

For example, assume that we have two basic attributes e_1 and e_2 with

$$S(e_1) = \{(H_1, 0), (H_2, 0), (H_3, 0), (H_4, 1), (H_5, 0)\}$$
$$S(e_2) = \{(H_1, 0), (H_2, 0), (H_3, 1), (H_4, 0), (H_5, 0)\}$$

and both are equally important, or $w_1 = w_2$. In the extreme, setting $\delta = 0$ and the weights w_1 and w_2 being normalized using (14) lead to $w_1 = w_2 = 1$, resulting in $(m_1^{w_1} \oplus m_2^{w_2})$, which does not exist.

Another issue with the orthogonal sum operation is in using the total probability mass κ (called the degree of conflict [29]) associated with conflict, as defined in the normalization factor. Consequently, applying it in an aggregation process may yield counterintuitive results in the face of significant conflict in certain contexts, as pointed out in [41]. Fortunately, in the context of the aggregated assessment based on a hierarchical evaluation model, by discounting all BPAs m_i $(i = 1, \ldots, L)$ at corresponding rates $(1 - w_i)$ $(i = 1, \ldots, L)$, we actually reduce the conflict between the various basic assessments before combining them.

Note further that, by definition, the focal elements of each $m_i^{w_i}$ are either singleton sets or a whole set $\mathcal{H}$. It is easy to see that m also verifies this property if applicable. Interestingly, the commutative and associative properties of Dempster's rule of combination with respect to a combinable collection of BPAs $m_i^{w_i}$ $(i = 1, \ldots, L)$ and the mentioned property essentially form the basis for the ER algorithms developed in [33] and [38]. In other words, the original ER algorithm summarized in (9) and (10) has been implemented for calculation of the BPA m. More particularly, with the same notations as in the preceding section, and denoting further

$$m_{I(i)} = \bigoplus_{j=1}^{i} m_j^{w_j}$$

for $i = 1, \ldots, L$, we have

$$m_{I(i)}(H_n) = m_{n,I(i)}, \quad \text{for } n = 1, \ldots, N \tag{27}$$
$$m_{I(i)}(\mathcal{H}) = m_{\mathcal{H},I(i)}. \tag{28}$$

Further, by a simple induction, we easily see that the following holds.

Lemma 1: With the quantity $\overline{m}_{\mathcal{H},I(i)}$ inductively defined by (19), we have

$$\overline{m}_{\mathcal{H},I(i)} = K_{I(i)} \prod_{j=1}^{i} (1 - w_j) \tag{29}$$

where $K_{I(i)}$ is inductively defined by (11).

By now, it is obviously clear that, except the weight normalization, the key difference between the original ER algorithm and the modified ER algorithm is in the way of assignment of β_n $(n = 1, \ldots, N)$ and $\beta_{\mathcal{H}}$ after obtaining m. That is, in the original ER algorithm, the BPA m is directly used to define the assessment for y by assigning

$$\beta_n = m(H_n) = m_{n,I(L)}, \quad \text{for } n = 1, \ldots, N \tag{30}$$
$$\beta_{\mathcal{H}} = m(\mathcal{H}) = m_{\mathcal{H},I(L)}. \tag{31}$$

While in the modified ER algorithm, after obtaining the BPA m, instead of using m to define the assessment for y, as in the original ER algorithm, it defines a BPA m' derived from m as

$$m'(H_n) = \frac{m(H_n)}{1 - \overline{m}_{\mathcal{H},I(L)}}, \quad \text{for } n = 1, \ldots, N \tag{32}$$
$$m'(\mathcal{H}) = \frac{\left(m(\mathcal{H}) - \overline{m}_{\mathcal{H},I(L)}\right)}{1 - \overline{m}_{\mathcal{H},I(L)}} = \frac{\tilde{m}_{\mathcal{H},I(L)}}{1 - \overline{m}_{\mathcal{H},I(L)}}. \tag{33}$$

Note that in this case we must have $w_i < 1$ for all $i = 1, \ldots, L$.

Then, the assessment for y is defined by assigning

$$\beta_n = m'(H_n), \quad \text{for } n = 1, \ldots, N \tag{34}$$
$$\beta_{\mathcal{H}} = m'(\mathcal{H}). \tag{35}$$

By (32) and (33), Theorem 2 straightforwardly follows as m is a BPA. Further, the following lemma holds.

Lemma 2: If all assessments $S(e_i)$ $(i = 1, \ldots, L)$ are complete, we have

$$m(\mathcal{H}) = \overline{m}_{\mathcal{H},I(L)} = K_{I(L)} \prod_{i=1}^{L} (1 - w_i) \tag{36}$$

i.e., $\tilde{m}_{\mathcal{H},I(L)} = 0$; and, consequently, $S(y)$ defined by (34) is also complete.

As if $w_i = 0$, then the BPA $m_i^{w_i}$ immediately becomes the vacuous BPA and consequently plays no role in the aggregation. Thus, without any loss of generality, we assume that $0 < w_i < 1$ for all $i = 1, \ldots, L$. Under this assumption, it is easy to see that if the assumption of Theorem 4 holds, then

$$\mathcal{F}_{m_i^{w_i}} = \left\{\{H_n\} | n \in I^+\right\} \cup \{\mathcal{H}\}, \quad \text{for } i = 1, \ldots, L \tag{37}$$

where $\mathcal{F}_{m_i^{w_i}}$ denotes the family of focal elements of $m_i^{w_i}$. Hence, by a simple induction, we also have

$$\mathcal{F}_m = \left\{\{H_n\} | n \in I^+\right\} \cup \{\mathcal{H}\}. \tag{38}$$

Note that the assumption of Theorem 3 is the same as that given in Theorem 4 with $|I^+| = 1$.

Therefore, Theorems 3 and 4 immediately follow from Lemma 2 along with (32)–(35), and (38).

It is also easily seen that

$$m(\mathcal{H}) = K_{I(L)} \prod_{i=1}^{L} m_i^{w_i}(\mathcal{H}) = K_{I(L)} \prod_{i=1}^{L} [w_i m_i(\mathcal{H}) + (1 - w_i)] \tag{39}$$

and, in addition, if there is an incomplete assessment, say $S(e_j)$, then $w_j m_j(\mathcal{H}) > 0$, resulting in

$$w_j m_j(\mathcal{H}) \prod_{\substack{i=1 \\ i \neq j}}^{L} (1 - w_i) > 0.$$

This directly implies $m'(\mathcal{H}) > 0$. Consequently, Theorem 4 follows as (34) and (35).

B. Discounting-and-Yager's Combination Scheme

To address the issue of conflict as mentioned above, Yager proposed in [29] a modification of Dempster's rule of combination by adding the total probability mass associated with conflict to the frame of discernment instead of using it for normalization. That is, given two BPAs m_1 and m_2 over a frame Θ, Yager's rule of combination yields a BPA denoted by m^Y as

$$m_1 \oplus m_2(A) \triangleq m^Y(A) = \begin{cases} 0, & \text{if } A = \emptyset \\ m_1(\Theta)m_2(\Theta) + \sum\limits_{B \cap C = \emptyset} m_1(B)m_2(C), & \text{if } A = \Theta \\ \sum\limits_{B \cap C = A} m_1(B)m_2(C), & \text{otherwise.} \end{cases} \tag{40}$$

As such, in Yager's rule of combination, the total probability mass associated with conflict between the two BPAs to be combined is attributed to the frame Θ and, consequently, enlarges the degree of ignorance.

In the context of a multiattribute assessment framework, after discounting the BPA m_i $(i = 1, \ldots, L)$ obtained from a basic assessment for e_i at a discount rate of $(1 - w_i)$, we would now like to apply Yager's rule of combination for obtaining an aggregated BPA for the assessment of the general attribute y. As Yager's rule of combination is not associative, we cannot combine $m_i^{w_i}$ $(i = 1, \ldots, L)$ in a recursive manner, as in the case of Dempster's rule, but apply a multiple piece of evidence version defined in [29]. This rule is suitable for an aggregation process (but not an updating process), as in the multiple-attribute aggregation.

Particularly, we define m^Y as a combination of BPAs $m_i^{w_i}$ $(i = 1, \ldots, L)$

$$m^Y(H) = \begin{cases} 0, & \text{if } H = \emptyset \\ \prod\limits_{i=1}^{L} m_i^{w_i}(\mathcal{H}) + \sum\limits_{\substack{H^i \subseteq \mathcal{H} \\ \bigcap_{i=1}^{L} H^i = \emptyset}} \prod\limits_{i=1}^{L} m_i^{w_i}(H^i), & \text{if } H = \mathcal{H} \\ \sum\limits_{\substack{H^i \subseteq \mathcal{H} \\ \bigcap_{i=1}^{L} H^i = H}} \prod\limits_{i=1}^{L} m_i^{w_i}(H^i), & \text{otherwise.} \end{cases} \tag{41}$$

Recall, by definition, that the focal elements of each $m_i^{w_i}$ are either singleton sets or the whole set $\mathcal{H}$. For $i = 1, \ldots, L$, let us denote

$$\mathcal{F}_i = \{\{H_n\} | H_n \in \mathcal{H} \wedge w_i \beta_{n,i} > 0\} \cup \{\mathcal{H}\}. \tag{42}$$

With this notation, if $w_i = 1$ and $S(e_i)$ is complete, the family of focal elements of $m_i^{w_i}$ $(i = 1, \ldots, L)$ is $\mathcal{F}_i^+ = \mathcal{F}_i \setminus \{\mathcal{H}\}$. Otherwise, the family of focal elements of $m_i^{w_i}$ $(i = 1, \ldots, L)$ is $\mathcal{F}_i$. For simplicity, we use H_n instead of $\{H_n\}$ without danger of confusion.

Then, we get

$$m^Y(H_n) = \sum_{\substack{H^i \in \mathcal{F}_i \\ \bigcap_{i=1}^{L} H^i = H_n}} \prod_{i=1}^{L} m_i^{w_i}(H^i) \tag{43}$$

$$m^Y(\mathcal{H}) = \prod_{i=1}^{L} m_i^{w_i}(\mathcal{H}) + K \tag{44}$$

where K is a constant defined as

$$K = \sum_{\substack{H^i \in \mathcal{F}_i^+ \\ \bigcap_{i=1}^{L} H^i = \emptyset}} \prod_{i=1}^{L} m_i^{w_i}(H^i) \tag{45}$$

representing the total degree of conflict.

Now, by this aggregation scheme, we can define

$$\beta_n = m^Y(H_n), \quad \text{for } n = 1, \ldots, N \tag{46}$$

$$\beta_{\mathcal{H}} = m^Y(\mathcal{H}) = 1 - \sum_{n=1}^{N} m^Y(H_n) \tag{47}$$

for generating the aggregated assessment for y.

Let us denote

$$I = \{1, 2, \ldots, L\} \tag{48}$$

$$I_n^+ = \{i \in I | w_i \beta_{n,i} > 0\}, \quad \text{for } n = 1, \ldots, N \tag{49}$$

and $\mathcal{P}(I_n^+)$ the power set of I_n^+. Then, β_n $(n = 1, \ldots, N)$ is calculated algorithmically as

$$m^Y(H_n) = \sum_{\emptyset \neq \tau \in \mathcal{P}(I_n^+)} \prod_{i \in \tau} w_i \beta_{n,i} \prod_{j \in I \setminus \tau} \left(1 - w_j \sum_{n=1}^{N} \beta_{n,j}\right). \tag{50}$$

We are now ready with the synthesis axioms. Obviously, the first independency axiom is followed as there is at least one H^i in (43) being H_n; thus, we have $\beta_n = 0$ if $\beta_{n,i} = 0$ for all i. Similar to the original ER algorithm, we have the following theorems.

Theorem 6: If β_n and $\beta_{\mathcal{H}}$ are calculated using (46) and (47), then the consensus axiom holds if and only if

$$\prod_{i=1}^{L}(1 - w_i) = 0.$$

Proof: Assume that $\beta_{k,i} = 1$ for all $i = 1, \ldots, L$, and $\beta_{n,i} = 0$ for $n = 1, \ldots, N$, $n \neq k$, and $i = 1, \ldots, L$.

By definition, we have

$$m_i^{w_i}(H_n) = \begin{cases} w_i, & \text{if } n = k \\ 0, & \text{if } n \neq k \end{cases}$$

and $m_i^{w_i}(\mathcal{H}) = (1 - w_i)$, for $i = 1, \ldots, L$. Consequently, $\mathcal{F}_i^+ = \{H_k\}$. This directly implies from (45) that $K = 0$. So, we obtain

$$\beta_{\mathcal{H}} = \prod_{i=1}^{L}(1 - w_i).$$

Then, the consensus axiom and $\mathcal{F}_i^+ = \{H_k\}$ immediately imply

$$\beta_{\mathcal{H}} = \prod_{i=1}^{L}(1 - w_i) = 0.$$

Conversely, if $\prod_{i=1}^{L}(1 - w_i) = 0$, then the consensus axiom is trivially satisfied. This concludes the proof. ■

It is not surprising that, as in the case of the discounting-and-orthogonal sum scheme above, the discounting-and-Yager's combination scheme does not directly yield a generated assessment for y exactly satisfying the completeness axiom. This can be overcome by modifying the assignment of β_n and $\beta_{\mathcal{H}}$ from m^Y as shown in the following.

C. Modified Discounting-and-Yager's Combination Scheme

As we have seen, the direct use of the discounting-and-Yager's combination scheme for defining the assessment for y makes it fail to desirably satisfy the synthesis axioms. This is caused mainly by the fact that an aggregated rate of discount is attributed to $\beta_{\mathcal{H}}$ as a part of the unassigned probability mass. Yet a so-called degree of disagreement as a part of the conflict factor K also plays a role. In this part, instead of committing these factors to the unassigned probability mass, they are used for the normalization before assigning for β_n's and $\beta_{\mathcal{H}}$. However, before doing so, we must first be clear on what these factors are.

Denote

$$\mathcal{F}^+ = \bigcap_{i=1}^{L} \mathcal{F}_i^+ \tag{51}$$

where $\mathcal{F}_i^+ = \mathcal{F}_i \setminus \{\mathcal{H}\}$ $(i = 1, \ldots, L)$. That is, $\mathcal{F}^+$ consists of common singleton focal elements of all BPAs $m_i^{w_i}$. In other words, all basic attributes e_i $(i = 1, \ldots, L)$ are assessed to all evaluation grades in $\mathcal{F}^+$ to variously positive degrees of belief. As a part of conflict that arises during the aggregation process of basic assessments, we define the degree of disagreement, denoted by κ_1, among the various basic assessments at evaluation grades in $\mathcal{F}^+$ as

$$\kappa_1 = \sum_{\substack{H^i \in \mathcal{F}^+ \\ \bigcap_{i=1}^{L} H^i = \emptyset}} \prod_{i=1}^{L} m_i^{w_i}(H^i). \tag{52}$$

Note that κ_1 is also a constant and is a part of the degree of conflict K defined by (45), as shown in the Appendix.

Due to the multiplicative nature of the combination rule, we define the aggregated rate of discount, denoted by κ_2, as

$$\kappa_2 = \prod_{i=1}^{L}(1 - w_i). \tag{53}$$

Also, κ_2 is a constant and is a part of $\prod_{i=1}^{L} m_i^{w_i}(\mathcal{H})$.

The assignment of a constant amount of $(\kappa_1 + \kappa_2)$ to $\beta_{\mathcal{H}}$ as part of an unassigned probability mass may cause the aggregated assessment to be incomplete even when all basic assessments are not. Therefore, in the modified discounting-and-Yager's combination scheme, this amount of $(\kappa_1 + \kappa_2)$ is assigned proportionally back to all individual grades and $\mathcal{H}$ using the normalization process

$$\beta_n = \frac{m^Y(H_n)}{1 - (\kappa_1 + \kappa_2)}, \quad \text{for } n = 1, \ldots, N \tag{54}$$

$$\beta_{\mathcal{H}} = \frac{m^Y(\mathcal{H}) - (\kappa_1 + \kappa_2)}{1 - (\kappa_1 + \kappa_2)}. \tag{55}$$

As remarked above on κ_1 and κ_2, it is easily seen, by definition, that the following holds.

Proposition 1: The degrees of belief generated using (54) and (55) satisfy

$$0 \leq \beta_n, \beta_{\mathcal{H}} \leq 1, \quad \text{for } n = 1, \ldots, N$$

$$\sum_{n=1}^{N} \beta_n + \beta_{\mathcal{H}} = 1.$$

Regarding the synthesis axioms, we have, desirably, the following.

Theorem 7: The aggregated assessment for y defined as in (54) and (55) exactly satisfies all four synthesis axioms.

Proof: See the Appendix. ■

D. Discounting-and-Averaging Scheme

In the aggregation schemes above, we have defined, for each $i = 1, \ldots, L$, $(w_i \otimes m_i)$ as the BPA $m_i^{w_i}$ by discounting m_i by a factor $(1 - w_i)$ [refer to (24)]. Then, Dempster's and Yager's rules of combination are, respectively, applied for obtaining the BPA m of the assessment for y. Here, we also assume that $0 < w_i \leq 1$ for all $i = 1, \ldots, N$.

In this subsection, instead of applying these combination operations after discounting m_is, we apply the averaging operation over L BPAs $m_i^{w_i}$ $(i = 1, \ldots, L)$ to obtain a BPA $\overline{m}$ defined by

$$\overline{m}(H) = \frac{1}{L} \sum_{i=1}^{L} m_i^{w_i}(H) \tag{56}$$

for any $H \subseteq \mathcal{H}$.

Due to (24), we have

$$\overline{m}(H) = \begin{cases} \frac{1}{L} \sum\limits_{i=1}^{L} w_i \beta_{n,i}, & \text{if } H = \{H_n\} \\ \frac{1}{L} \sum\limits_{i=1}^{L} \left(1 - w_i \sum\limits_{n=1}^{N} \beta_{n,i}\right), & \text{if } H = \mathcal{H} \\ 0, & \text{otherwise.} \end{cases} \tag{57}$$

After obtaining the aggregated BPA $\overline{m}$, the problem now is to use $\overline{m}$ for generating the aggregated assessment for the general attribute y. Naturally, we can assign

$$\beta_n = \overline{m}(H_n) = \frac{1}{L} \sum_{i=1}^{L} w_i \beta_{n,i}, \quad \text{for } n = 1, \ldots, N \tag{58}$$

$$\beta_{\mathcal{H}} = \overline{m}(\mathcal{H}) = \frac{1}{L} \sum_{i=1}^{L} \left(1 - w_i \sum_{n=1}^{N} \beta_{n,i}\right). \tag{59}$$

Then, the assessment for y is defined by

$$S(y) = \{(H_n, \beta_n) | n = 1, \ldots, N\}. \tag{60}$$

Regarding the synthesis axioms, we easily see that the first axiom holds for the assessment (60). For the next two axioms, we have the following.

Theorem 8: The assessment (60) defined via (58) and (59) satisfies the consensus axiom and/or the completeness axiom if and only if $w_i = 1$ for all $i = 1, \ldots, L$.

Proof: For the consensus axiom, the proof is straightforward. Now we rewrite $\beta_{\mathcal{H}}$ defined by (59) as

$$\begin{aligned} \beta_{\mathcal{H}} &= \frac{1}{L} \sum_{i=1}^{L} \left(1 - w_i \sum_{n=1}^{N} \beta_{n,i}\right) \\ &= \frac{1}{L} \sum_{i=1}^{L} w_i \left(1 - \sum_{n=1}^{N} \beta_{n,i}\right) + \left(1 - \frac{\sum_{i=1}^{L} w_i}{L}\right). \end{aligned} \tag{61}$$

Thus, if all $w_i = 1$ and the assumption of the completeness axiom holds, we have $\beta_{\mathcal{H}} = 0$, and the conclusion of the axiom follows easily. Inversely, if the completeness axiom is satisfied, we must have

$$1 - \frac{\sum\limits_{i=1}^{L} w_i}{L} = 0$$

that directly implies $w_i = 1$ for all i. ■

The assessment for y according to this aggregation scheme also satisfies the incompleteness axiom trivially due to the nature of discounting-and-averaging.

Unfortunately, the requirement of $w_i = 1$ for all i to satisfy the consensus axiom and the completeness axiom would not be appropriate in general. This is due to the allocation of the average of discount rates

$$\overline{\alpha} \triangleq \left(1 - \frac{\sum\limits_{i=1}^{L} w_i}{L}\right)$$

to $\mathcal{H}$ as a part of an unassigned probability mass. This dilemma can be resolved in a similar way as in the modified algorithms above. Interestingly, this modification leads to the weighted sum scheme, as shown in the following.

E. Weighted Sum as Modified Discounting-and-Averaging Scheme

By applying the discounting-and-averaging scheme, we obtain the BPA $\overline{m}$ as defined by (57). Now, guided by the synthesis axioms, instead of making direct use of $\overline{m}$ in defining the generated assessment $S(y)$ (i.e., allocating the average discount rate $\overline{\alpha}$ to $\beta_{\mathcal{H}}$ as a part of unassigned probability mass) as above, we define a new BPA denoted by $\overline{m}'$ derived from $\overline{m}$ by making use of $(1 - \overline{\alpha})$ as a normalization factor. More particularly, we define

$$\overline{m}'(H_n) = \frac{\overline{m}(H_n)}{1 - \overline{\alpha}}, \quad \text{for } n = 1, \ldots, N \tag{62}$$

$$\overline{m}'(\mathcal{H}) = \frac{\overline{m}(\mathcal{H}) - \overline{\alpha}}{1 - \overline{\alpha}}. \tag{63}$$

Combining (57) and (61), we interestingly obtain

$$\overline{m}'(H_n) = \sum_{i=1}^{L} \overline{w}_i \beta_{n,i}, \quad \text{for } n = 1, \ldots, N \tag{64}$$

$$\overline{m}'(\mathcal{H}) = \sum_{i=1}^{L} \overline{w}_i \left(1 - \sum_{n=1}^{N} \beta_{n,i}\right) \tag{65}$$

where

$$\overline{w}_i = \frac{w_i}{\sum\limits_{i=1}^{L} w_i}, \quad \text{for } i = 1, \ldots, L.$$

Let us turn back to the general scheme of combination given in (25). Under the view of this general scheme, the above BPA

$\overline{m}'$ is nothing but an instance of it by simply considering $\otimes$ as the multiplication and $\oplus$ as the weighted sum. Namely, we have

$$\overline{m}'(H_n) = \sum_{i=1}^{L} \overline{w}_i m_i(H_n), \quad \text{for } n = 1, \ldots, N \tag{66}$$

$$\overline{m}'(\mathcal{H}) = \sum_{i=1}^{L} \overline{w}_i m_i(\mathcal{H}) \tag{67}$$

where relative weights $\overline{w}_i$ are normalized as above so that $\sum_i \overline{w}_i = 1$. It is of interest to note that the possibility of using such an operation has previously been mentioned in, for example, [7] and [30]. Especially, the weighted sum operation of two BPAs has been used for the integration of distributed databases for purposes of data mining [17].

Now we quite naturally define the assessment for y by assigning

$$\beta_n = \overline{m}'(H_n) = \sum_{i=1}^{L} \overline{w}_i m_i(H_n), \quad \text{for } n = 1, \ldots, N \tag{68}$$

$$\beta_{\mathcal{H}} = \overline{m}'(\mathcal{H}) = \sum_{i=1}^{L} \overline{w}_i m_i(\mathcal{H}). \tag{69}$$

Appealingly simple as it is, we can see quite straightforwardly that the following theorem holds.

Proposition 2: The degrees of belief generated using (68) and (69) satisfy

$$0 \le \beta_n, \beta_{\mathcal{H}} \le 1, \quad \text{for } n = 1, \ldots, N$$

$$\sum_{n=1}^{N} \beta_n + \beta_{\mathcal{H}} = 1.$$

Furthermore, concerning the synthesis axioms, we have the following theorem.

Theorem 9: The aggregated assessment for y defined as in (68) and (69) exactly satisfies all four synthesis axioms.

Proof: The proof is trivial. ■

F. Expected Utility in the ER Approaches

In the tradition of decision making under uncertainty [22], the notion of expected utility has been mainly used to rank alternatives in a particular problem. That is, one can represent the preference relation $\succeq$ on a set of alternatives X with a single-valued function $u(x)$ on X, called expected utility, such that for any $x, y \in X$, $x \succeq y$ if and only if $u(x) \ge u(y)$. Maximization of $u(x)$ over X provides the solution to the problem of selecting x.

When the D–S theory is considered in the perspective of decision analysis under uncertainty, as a BPA does not in general provide a unique probability distribution but a set of compatible probabilities bounded by the corresponding belief and plausibility functions, it is strictly nothing equivalent to the notion of expected utility, leading to a single decision. However, it is possible to meaningfully define the interval utility as done in the ER approach in the following way [38], [39] (or in decision analysis with costs, one can define the interval expected costs, as mentioned in [4] and [6]).

In the ER approach, we assume a utility function

$$u' : \mathcal{H} \to [0, 1]$$

satisfying

$$u'(H_{n+1}) > u'(H_n), \text{ if } H_{n+1} \text{ is preferred to } H_n.$$

This utility function u' may be determined using the probability assignment method [14] or using other methods as in [33] and [38].

If all assessments for basic attributes are complete, Lemma 2 shows that the assessment for y is also complete, i.e., $\beta_{\mathcal{H}} = 0$. Then, the expected utility of an alternative on the attribute y is defined by

$$u(y) = \sum_{n=1}^{N} \beta_n u'(H_n). \tag{70}$$

An alternative a is strictly preferred to another alternative b if and only if $u(y(a)) > u(y(b))$.

Due to incompleteness, in general, in basic assessments, the assessment for y may be incomplete. In such a case, in [38], the authors define three measures, called the minimum, maximum, and average expected utilities, as

$$u_{\max}(y) = \sum_{n=1}^{N-1} \beta_n u'(H_n) + (\beta_N + \beta_{\mathcal{H}}) u'(H_N) \tag{71}$$

$$u_{\min}(y) = (\beta_1 + \beta_{\mathcal{H}}) u'(H_1) + \sum_{n=2}^{N} \beta_n u'(H_n) \tag{72}$$

$$u_{\text{avg}}(y) = \frac{u_{\max}(y) + u_{\min}(y)}{2} \tag{73}$$

where, without loss of generality, suppose H_1 is the least preferred grade having the lowest utility and H_N the most preferred grade having the highest utility.

The ranking of two alternatives a and b on y is carried out by

- $a \succ_y b$ if and only if $u_{\min}(y(a)) > u_{\max}(y(b))$
- $a \sim_y b$ if and only if $u_{\min}(y(a)) = u_{\min}(y(b))$ and $u_{\max}(y(a)) = u_{\max}(y(b))$.

If these are not the case, the average expected utility can be used to generate a ranking (see, e.g., [38] for more details).

Alternatively, guided by the Generalized Insufficient Reason Principle, we are also able to define a probability function P_m on $\mathcal{H}$ derived from m for the purpose of making decisions via the pignistic transformation [26]. Namely

$$P_m(H_n) = m(H_n) + \frac{1}{N} m(\mathcal{H}), \quad \text{for } n = 1, \ldots, N. \tag{74}$$

That is, as in the two-level language of the so-called transferable belief model [26], the aggregated BPA m itself representing the belief is entertained based on available evidence at the credal level, and when a decision must be made, the belief at the credal level induces the probability function P_m defined by (74)

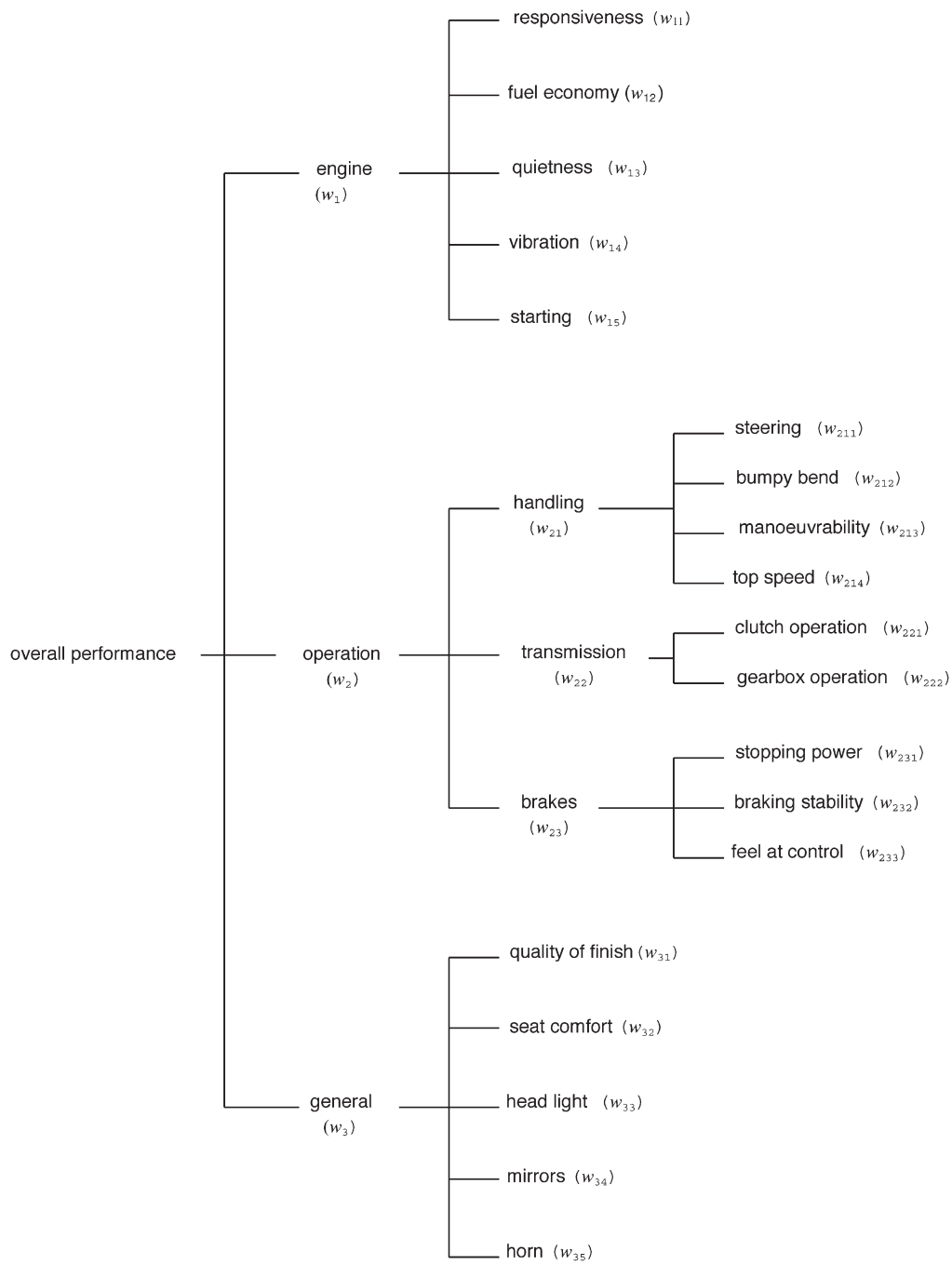


Fig. 3. Evaluation hierarchy for motorcycle performance assessment [38].

for decision making. Particularly, the approximate assessment for y for the purpose of decision making is then defined as

$$\beta'_n = P_m(H_n) = \beta_n + \frac{1}{N}\beta_{\mathcal{H}}, \quad \text{for } n = 1, \ldots, N. \tag{75}$$

Therefore, the expected utility of an alternative on the attribute y is straightforwardly defined by

$$u(y) = \sum_{n=1}^{N} \beta'_n u'(H_n) = \sum_{n=1}^{N} \left(\beta_n + \frac{1}{N}\beta_{\mathcal{H}}\right) u'(H_n). \tag{76}$$

As such, while the amount of belief $\beta_{\mathcal{H}}$ (due to ignorance) is allocated either to the least preferred grade H_1 or to the most preferred grade H_N to define the expected utility interval in Yang's approach [38], in our approach, it is uniformly allocated to every evaluation grade H_n, guided by the Generalized Insufficient Reason Principle [26], to define an approximate assessment for y and, hence, a single-valued expected utility function.

In the following section, we examine a tutorial example taken from [38] to illustrate how the difference between the various aggregation schemes as well as the respective results yielded.

TABLE I
GENERALIZE DECISION MATRIX FOR MOTORCYCLE ASSESSMENT [38]

General attributes			Basic attributes	types of motor cycle (alternatives)			
				Kawasaki (a_1)	*Yamaha* (a_2)	*Honda* (a_3)	*BMW* (a_4)
Overall performance	engine		responsiveness	$E(0.8)$	$G(0.3)\ E(0.6)$	$G(1.0)$	$I(1.0)$
			fuel economy	$A(1.0)$	$I(1.0)$	$I(0.5)\ A(0.5)$	$E(1.0)$
			quietness	$I(0.5)\ A(0.5)$	$A(1.0)$	$G(0.5)\ E(0.3)$	$E(1.0)$
			vibration	$G(1.0)$	$I(1.0)$	$G(0.5)\ E(0.5)$	$P(1.0)$
			starting	$G(1.0)$	$A(0.6)\ G(0.3)$	$G(1.0)$	$A(1.0)$
	operation	handling	steering	$E(0.9)$	$G(1.0)$	$A(1.0)$	$A(0.6)$
			bumpy bends	$A(0.5)\ G(0.5)$	$G(1.0)$	$G(0.8)\ E(0.1)$	$P(0.5)\ I(0.5)$
			maneuverability	$A(1.0)$	$E(0.9)$	$I(1.0)$	$P(1.0)$
			top speed stability	$E(1.0)$	$G(1.0)$	$G(1.0)$	$G(0.6)\ E(0.4)$
		transmission	clutch operation	$A(0.8)$	$G(1.0)$	$E(0.85)$	$I(0.2)\ A(0.8)$
			gearbox operation	$A(0.5)\ G(0.5)$	$I(0.5)\ A(0.5)$	$E(1.0)$	$P(1.0)$
		brakes	stopping power	$G(1.0)$	$A(0.3)\ G(0.6)$	$G(0.6)$	$E(1.0)$
			braking stability	$G(0.5)\ E(0.5)$	$G(1.0)$	$A(0.5)\ G(0.5)$	$E(1.0)$
			feel at control	$P(1.0)$	$G(0.5)\ E(0.5)$	$G(1.0)$	$G(0.5)\ E(0.5)$
	general		quality of finish	$P(0.5)\ I(0.5)$	$G(1.0)$	$E(1.0)$	$G(0.5)\ E(0.5)$
			seat comfort	$G(1.0)$	$G(0.5)\ E(0.5)$	$G(0.6)$	$E(1.0)$
			headlight	$G(1.0)$	$A(1.0)$	$E(1.0)$	$G(0.5)\ E(0.5)$
			mirrors	$A(0.5)\ G(0.5)$	$G(0.5)\ E(0.5)$	$E(1.0)$	$G(1.0)$
			horn	$A(1.0)$	$G(1.0)$	$G(0.5)\ E(0.5)$	$E(1.0)$

TABLE II
AGGREGATED ASSESSMENTS FOR FOUR TYPES OF MOTORCYCLES USING THE MODIFIED ER METHOD [38]

	Poor(P)	*Indifference*(I)	*Average*(A)	*Good*(G)	*Excellent*(E)	*Unknown*(U)
Kawasaki	0.0547	0.0541	0.3216	0.4452	0.1058	0.0186
Yamaha	0.0	0.1447	0.1832	0.5435	0.1148	0.0138
Honda	0.0	0.0474	0.0621	0.4437	0.4068	0.0399
BMW	0.1576	0.0792	0.1124	0.1404	0.5026	0.0078

V. EXAMPLE: MOTORCYCLE ASSESSMENT PROBLEM

The problem is to evaluate the performance of four types of motorcycles, namely, Kawasaki, Yamaha, Honda, and BMW.

The overall performance of each motorcycle is evaluated based on three major attributes, which are $quality\ of\ engine$, $operation$, and $general\ finish$. These attributes are all general and difficult to assess directly. So these attributes are correspondingly decomposed into more detailed sub-attributes to facilitate the assessment. The process of attribute decomposition for the evaluation problem of motorcycles results in an attribute hierarchy graphically depicted in Fig. 3, where the relative weights of attributes at a single level associated with the same upper level attribute are defined by w_i, w_{ij}, and w_{ijk} for the attributes at levels 1, 2, and 3, respectively.

Using the five-grade evaluation scale as given in (1), the assessment problem of motorcycles is given in Table I, where P, I, A, G, and E are the abbreviations of poor, indifferent, average, good, and excellent, respectively, and a number in brackets denoted the degree of belief to which an attribute is assessed to a grade. For example, $E(0.8)$ means "excellent to a degree of 0.8."

Further, all relevant attributes are assumed to be of equal relative importance [38]. That is

$$w_1 = w_2 = w_3 = 0.3333$$
$$w_{11} = w_{12} = w_{13} = w_{14} = w_{15} = 0.2$$
$$w_{21} = w_{22} = w_{23} = 0.3333$$
$$w_{211} = w_{212} = w_{213} = w_{214} = 0.25$$
$$w_{221} = w_{222} = 0.5$$
$$w_{231} = w_{232} = w_{233} = 0.3333$$
$$w_{31} = w_{32} = w_{33} = w_{34} = w_{35} = 0.2.$$

In the sequel, for the purpose of comparison, we generate three different results of aggregation corresponding to Yang and Xu's modified ER method and the other two developed in this paper.

By applying the modified ER method, the distributed assessments for overall performance of four types of motorcycles are given in Table II. These four distributions and their

TABLE III
APPROXIMATE ASSESSMENTS FOR FOUR TYPES OF MOTORCYCLES USING THE MODIFIED ER METHOD

	Poor(P)	*Indifference*(I)	*Average*(A)	*Good*(G)	*Excellent*(E)
Kawasaki	0.05842	0.05782	0.32532	0.44892	0.10952
Yamaha	0.00276	0.14746	0.18596	0.54626	0.11756
Honda	0.00798	0.05538	0.07008	0.45168	0.41478
BMW	0.15916	0.08076	0.11396	0.14196	0.50416

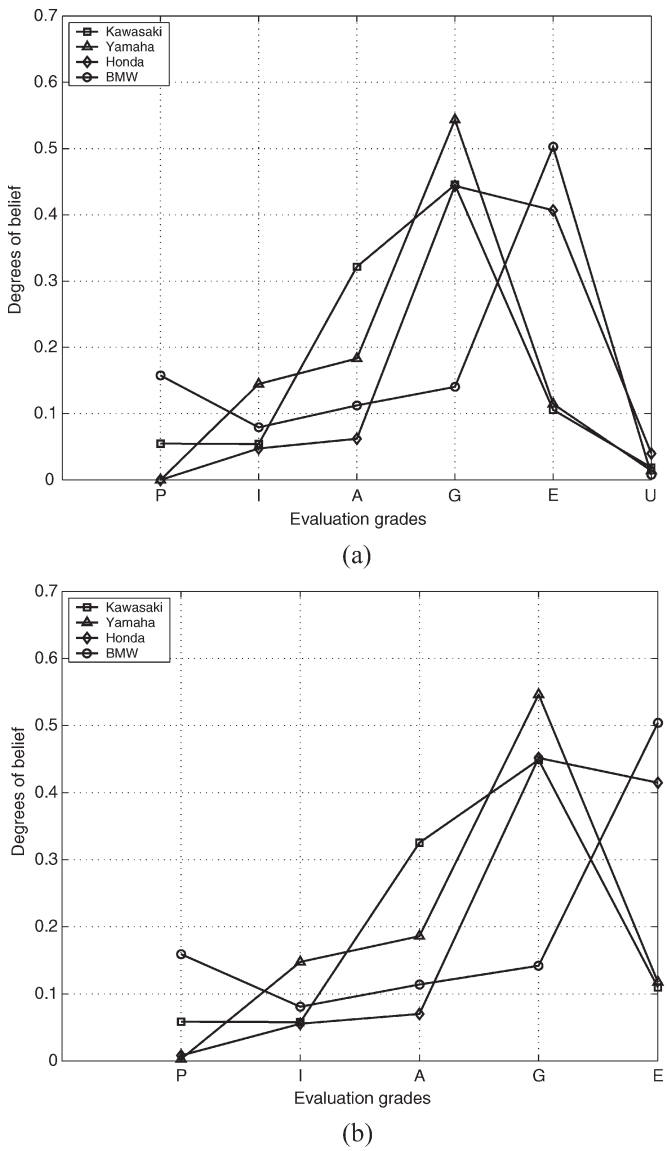


Fig. 4. Overall evaluation of motorcycles via the modified ER method. (a) Aggregated assessment. (b) Approximate assessment via the pignistic transformation.

approximations via the pignistic transformation (Table III) are graphically shown in Fig. 4.

At the same time, by applying the weighted sum aggregation scheme (shortly, WS method), we easily obtain the distributed assessments for overall performance of four types of motorcycles, as shown in Table IV [graphically depicted in Fig. 5(a)]. The pignistic transformation applied to these aggregated assessments yields the approximate assessments for overall performance of motorcycles, as given in Table V [graphically, Fig. 5(b)].

It would be worth noticing that, though there is not so much difference between the results obtained by the modified ER algorithm and that obtained by the weighted sum method [especially where the behavior of corresponding assessment distributions is almost the same as shown in Figs. 4(a) and 5(a)], the two methods have very different behaviors. That is, while the weighted sum method clearly has a linear behavior, the modified ER method exhibits a quasi-linear behavior with equal weights and strongly nonlinear behavior with unequal weights [37]. In fact, the two sets of results generated in the example look largely similar because of the assumption of equal weights applied to the same group of attributes. This is theoretically also in relation to the relationship between the discounting-and-orthogonal sum scheme and the averaging combination scheme regarding conflicting information, as established in [25]. More particularly, when combining a number L of equally reliable sources using Dempster's rule on equally discounted BPAs, one gets a similar result as by trading off between the L sources with equal reliability weights, i.e., averaging operation, if the sources are highly conflicting and have been sufficiently discounted.

However, as we see in the following, the result yielded by the modified Yager's combination method (shortly, MY method) is relatively different from those obtained by the above methods. This is unsurprising as we were attributing a factor of conflict to $\mathcal{H}$ as "unknown" in the aggregated assessment.

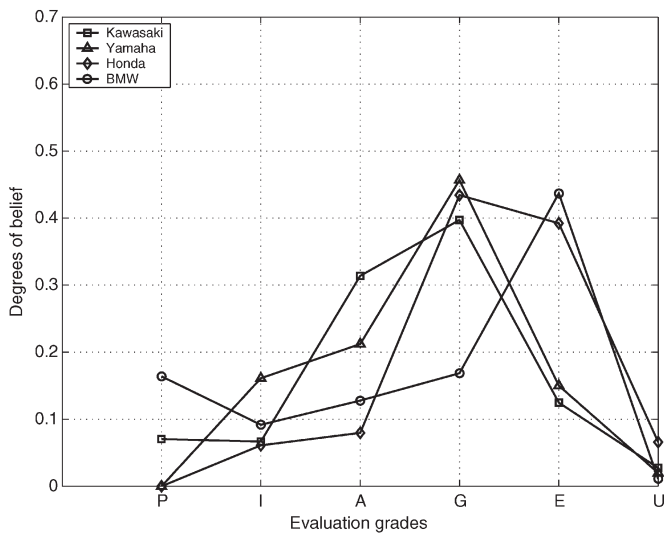
For generating the assessment for an attribute y at a higher level in the hierarchy of attributes shown in Fig. 3, all the BPAs of its direct sub-attributes are first aggregated via (50), and the generated assessment for y is then obtained using the normalization process represented in (54) and (55). This process is carried out upward from the bottom level to the top of the hierarchy in order to obtain the overall assessment. With this method of aggregation, we obtain the distributed assessments for overall performance of four types of motorcycles, as shown in Table VI, which are graphically depicted in Fig. 6(a).

From the obtained result, it is interesting to observe that, although a total degree of incompleteness in basic assessments of Honda is 1.25 compared to those of the other three, which in turn are 0.5 for both Kawasaki and Yamaha and 0.4 for BMW, the unassigned probability mass of the generated assessment for Honda is smaller than those of the remainders. This is due to a lower conflict between basic assessments of Honda compared to those of the others.

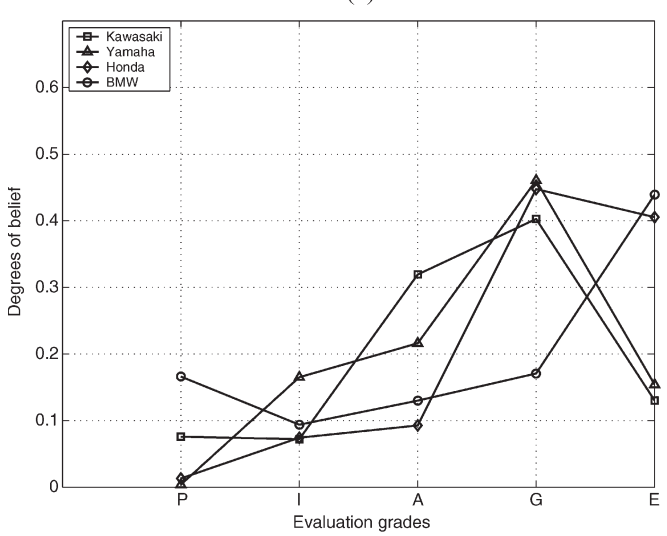
For the purpose of decision making, we apply the pignistic transformation to the aggregated assessments in order to obtain the approximate assessments for overall performance of motorcycles, as shown in Table VII and depicted graphically in Fig. 6(b).

TABLE IV
AGGREGATED ASSESSMENTS FOR FOUR TYPES OF MOTORCYCLES USING THE WEIGHTED SUM COMBINATION METHOD

	Poor(P)	*Indifference*(I)	*Average*(A)	*Good*(G)	*Excellent*(E)	*Unknown*(U)
Kawasaki	0.0703	0.0667	0.3139	0.3972	0.1247	0.0272
Yamaha	0.0	0.1611	0.2122	0.4567	0.1501	0.0198
Honda	0.0	0.0611	0.0796	0.4344	0.3922	0.0659
BMW	0.1639	0.0917	0.1278	0.1685	0.437	0.0111



(a)



(b)

Fig. 5. Overall evaluation of motorcycles via the weighted sum method. (a) Aggregated assessment. (b) Approximate assessment via the pignistic transformation.

We are now ready to assume a utility function $u' : \mathcal{H} \to [0, 1]$ defined in [38] as

$$u'(P) = 0, \quad u'(I) = 0.35,$$
$$u'(A) = 0.55, \quad u'(G) = 0.85, \quad u'(E) = 1.$$

Using (76), we easily obtain the expected utility of four types of motorcycles according to the various methods given in Table VIII.

Consequently, the ranking of the four types of motorcycles is given in Table IX.

Note that the same ranking result for all methods could also be obtained by the expected utility interval and the ranking scheme by Yang and Xu [38] as mentioned above. As we have seen, although the solution to the problem of selecting the best alternative is the same for all the three methods of aggregation, the ranking order between the alternatives is different. More particularly, while Yamaha is preferred over BMW according to the results of the first two methods, BMW is preferred over Yamaha according to the third method. This is because, by the third method of aggregation, the former is assessed to good and excellent to a total degree of 0.5797, while the latter is 0.54872.

VI. CONCLUDING REMARKS

In this paper, the authors have re-analyzed the ER approach to MADM under uncertainty. Interestingly, the analysis provides a general formulation for the attribute aggregation problem in MADM under uncertainty. Under such a generalization, several various aggregation schemes have been examined, including the previous one. The theoretical properties of new schemes regarding the synthesis axioms proposed in [38] were also explored. In addition, this reformulation of the attribute aggregation problem has practically shown that the aggregation scheme based on the weighted sum operation could also be considered for the aggregation process in the context of MADM under uncertainty, especially when the assumption regarding the independence of attributes' uncertain evaluations is not appropriate.

For the purpose of decision making, an approximate method of uncertain assessments based on the so-called pignistic transformation [26] has been applied to define the expected utility function instead of using the expected utility interval proposed previously. A tutorial example has been examined to illustrate the discussed techniques.

In summary, by the results obtained in this paper, the authors do hope to support further aggregation schemes for the attribute aggregation problem in MADM under uncertainty. This is especially helpful in decision-making situations where a single method of aggregation would be inapplicable or not sufficient.

The approach proposed in this paper can also be extended to apply to uncertain assessments where belief degrees may be assigned not only to singleton sets but also to non-singleton ones. For further work, the authors are planning to study problems of hybrid MADM under uncertainty [34] in the spirit of the proposed approach. Also, the decision model

TABLE V
APPROXIMATE ASSESSMENTS FOR FOUR TYPES OF MOTORCYCLES USING THE WEIGHTED SUM COMBINATION METHOD

	Poor(P)	*Indifference*(I)	*Average*(A)	*Good*(G)	*Excellent*(E)
Kawasaki	0.07574	0.07214	0.31934	0.40264	0.13014
Yamaha	0.00396	0.16506	0.21616	0.46066	0.15406
Honda	0.01318	0.07428	0.09278	0.44758	0.40538
BMW	0.16612	0.09392	0.13	0.17072	0.43922

TABLE VI
AGGREGATED ASSESSMENTS FOR FOUR TYPES OF MOTORCYCLES USING THE MODIFIED YAGER'S COMBINATION METHOD

	Poor(P)	*Indifference*(I)	*Average*(A)	*Good*(G)	*Excellent*(E)	*Unknown*(U)
Kawasaki	0.0344	0.0369	0.2114	0.276	0.0653	0.3761
Yamaha	0	0.099	0.127	0.3143	0.0843	0.3753
Honda	0	0.028	0.0375	0.2835	0.3331	0.3177
BMW	0.0855	0.0464	0.0628	0.1025	0.3268	0.376

recently proposed in [12] for multiexpert decision-making problems under linguistic assessments may be applied to MADM problems under uncertainty as well, but further research is required.

APPENDIX

In this Appendix, we give the proof of Theorem 7 on the synthesis axioms for the modified discounting-and-Yager's combination scheme. Clearly, the independency axiom is immediately followed from (43) as the case of the discounting-and-Yager's combination scheme. Note that we assume here that weights w_i are normalized so that $0 < w_i < 1$ for all $i \in I$. First, we need some preparations. Recall that

$$\beta_n = \frac{m^Y(H_n)}{1-(\kappa_1+\kappa_2)}, \quad \text{for } n = 1, \ldots, N$$

$$\beta_{\mathcal{H}} = \frac{m^Y(\mathcal{H}) - (\kappa_1+\kappa_2)}{1-(\kappa_1+\kappa_2)}$$

$$m^Y(\mathcal{H}) = \Delta + K$$

where

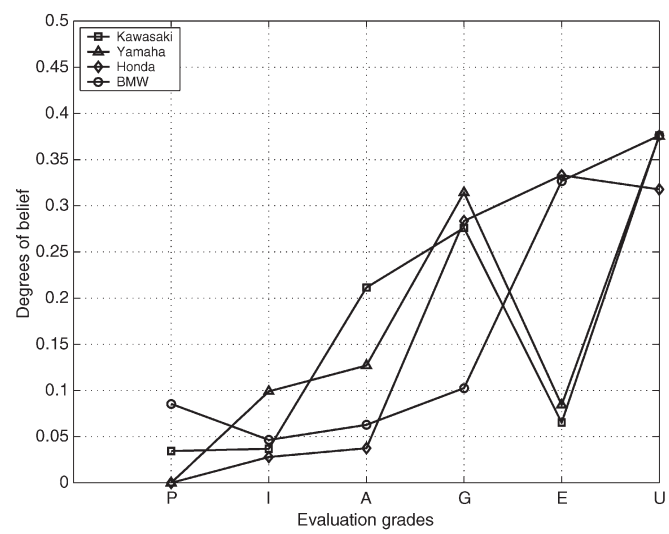
$$\Delta = \prod_{i=1}^{L} [w_i m_i(\mathcal{H}) + (1-w_i)]$$

$$K = \sum_{H^i \in \mathcal{F}_i^+ : \bigcap_{i=1}^{L} H^i = \emptyset} \prod_{i=1}^{L} w_i m_i(H^i)$$

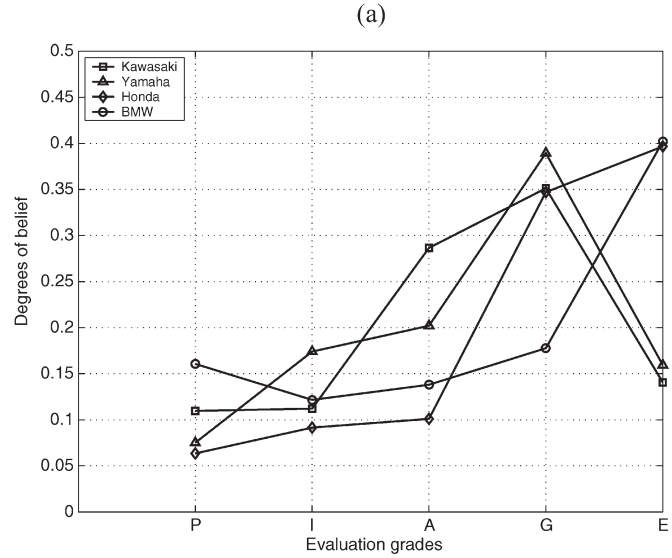
$$\mathcal{F}_i^+ = \{H_n | H_n \in \mathcal{H} \wedge w_i \beta_{n,i} > 0\}.$$

Let us denote

$$\mathbb{F} = \mathcal{F}_1^+ \times \mathcal{F}_2^+ \times \cdots \times \mathcal{F}_L^+$$
$$\mathbb{F}^+ = \mathcal{F}^+ \times \mathcal{F}^+ \times \cdots \times \mathcal{F}^+$$



(a)



(b)

Fig. 6. Overall evaluation of motorcycles via the modified Yager's combination method. (a) Aggregated assessment. (b) Approximate assessment via the pignistic transformation.

TABLE VII
APPROXIMATE ASSESSMENTS FOR FOUR TYPES OF MOTORCYCLES USING THE MODIFIED YAGER'S COMBINATION METHOD

	Poor(P)	*Indifference*(I)	*Average*(A)	*Good*(G)	*Excellent*(E)
Kawasaki	0.10962	0.11212	0.28662	0.35122	0.14052
Yamaha	0.07506	0.17406	0.20206	0.38936	0.15936
Honda	0.06354	0.09154	0.10104	0.34704	0.39664
BMW	0.1607	0.1216	0.138	0.1777	0.402

TABLE VIII
EXPECTED UTILITY OF FOUR TYPES OF MOTORCYCLES

	Kawasaki	*Yamaha*	*Honda*	*BMW*
ER method	0.69026	0.73577	0.85664	0.71577
WS method	0.67327	0.72228	0.86285	0.68871
MY method	0.63594	0.66237	0.77923	0.67151

TABLE IX
RANKING OF FOUR TYPES OF MOTORCYCLES

Method	**Ranking order**
ER method	*Honda* $\succ$ *Yamaha* $\succ$ *BMW* $\succ$ *Kawasaki*
WS method	*Honda* $\succ$ *Yamaha* $\succ$ *BMW* $\succ$ *Kawasaki*
MY method	*Honda* $\succ$ *BMW* $\succ$ *Yamaha* $\succ$ *Kawasaki*

where $\times$ denotes the Cartesian product. For

$$H = (H^1, \ldots, H^L) \in \mathbb{F} \quad (\text{or } \mathbb{F}^+)$$

by $\cap H$, we mean $\cap_{i=1}^{L} H^i$. We can now decompose Δ into two parts as $\Delta = \Delta' + \kappa_2$, where

$$\Delta' = \sum_{\emptyset \neq \tau \in \mathcal{P}(I)} \prod_{i \in \tau} w_i m_i(\mathcal{H}) \prod_{i \in I \setminus \tau} (1 - w_i). \tag{77}$$

Similarly, K is decomposed as $K = K' + \kappa_1$ with

$$K' = \sum_{\substack{H \in (\mathbb{F} \setminus \mathbb{F}^+) \\ \cap H = \emptyset}} \prod_{i=1}^{L} w_i m_i(H^i). \tag{78}$$

Proof for the consensus axiom: Suppose that $\beta_{k,i} = 1$ for all $i \in I$, and $\beta_{n,i} = 0$ for $k \neq n = 1, \ldots, N$, $n \neq k$, and $i \in I$.

Then, we have

$$m_i^{w_i}(H_n) = \begin{cases} w_i, & \text{if } n = k \\ 0, & \text{if } n \neq k \end{cases}$$

and $m_i^{w_i}(\mathcal{H}) = (1 - w_i)$, for $i \in I$. Thus

$$\mathcal{F}_i^+ = \{H_k\}, \quad \text{for all } i = 1, \ldots, L.$$

This directly implies that $K = 0$. Further, we have $\Delta' = 0$ since $m_i(\mathcal{H}) = 0$ for all $i \in I$. Hence, we get

$$m^Y(\mathcal{H}) = \kappa_2 = \prod_{i=1}^{L} (1 - w_i)$$

and it immediately follows that $\beta_{\mathcal{H}} = 0$.

Inductively, we have

$$\sum_{\tau \in \mathcal{P}(I)} \prod_{i \in \tau} w_i \prod_{j \in I \setminus \tau} (1 - w_j) = 1.$$

Therefore

$$\sum_{\emptyset \neq \tau \in \mathcal{P}(I)} \prod_{i \in \tau} w_i \prod_{j \in I \setminus \tau} (1 - w_j) = 1 - \kappa_2.$$

From (49) and the assumption, we have $I_k^+ = I$. Thus, the last equation and (50) imply that $\beta_k = 1$. This completes the proof for the consensus axiom. ■

Proof for the completeness axiom: Assume $\mathcal{H}^+ \subset \mathcal{H}$ and denote $J^+ = \{n | H_n \in \mathcal{H}^+\}$. We now prove the following statement.

If $\beta_{n,i} > 0$ for $n \in J^+$ and $\sum_{n \in J^+} \beta_{n,i} = 1$ for all $i \in I$, then $\beta_n > 0$ for $n \in J^+$ and $\sum_{n \in J^+} \beta_n = 1$ as well.

Since $0 < w_i < 1$ for all $i \in I$, we have

$$m_i^{w_i}(H_n) = \begin{cases} w_i \beta_{n,i} > 0, & \text{if } n \in J^+ \\ 0, & \text{otherwise} \end{cases}$$

and hence

$$\mathcal{F}_i^+ = \{H_n | n \in J^+\}, \quad \text{for all } i \in I.$$

Therefore, $\mathbb{F}^+ = \mathbb{F}$, which directly follows $K' = 0$. Further, from (49) we get $I_n^+ = I$ for all $n \in J^+$. Using (50), we obtain

$$m^Y(H_n) = \sum_{\emptyset \neq \tau \in \mathcal{P}(I)} \prod_{i \in \tau} w_i \beta_{n,i} \prod_{j \in I \setminus \tau} (1 - w_j) \tag{79}$$

for any $H_n \in \mathcal{H}^+$ (i.e., $n \in J^+$).

On the other hand, we have $m_i(\mathcal{H}) = 0$ for all i. So, we also have $\Delta' = 0$. By definition, we get

$$m^Y(\mathcal{H}) = \kappa_1 + \kappa_2.$$

Thus

$$\beta_{\mathcal{H}} = \frac{m^Y(\mathcal{H}) - (\kappa_1 + \kappa_2)}{1 - (\kappa_1 + \kappa_2)} = 0.$$

From (43) and the assumption $\beta_{n,i} = 0$, if $n \notin J^+$ for all $i \in I$, we easily deduce that

$$m^Y(H_n) = 0, \quad \text{for any } n \in \{1, \ldots, N\} \setminus J^+.$$

Immediately, it follows that

$$\beta_n = 0, \quad \text{for any } n \in \{1, \ldots, N\} \setminus J^+.$$

Again, since $0 < w_i < 1$ for all i, it follows $\kappa_2 > 0$. This implies from (79) that $m^Y(H_n) > 0$ for all $n \in J^+$. Thus, $\beta_n > 0$ for all $n \in J^+$. Finally, the desired equation

$$\sum_{n \in J^+} \beta_n = 1$$

is followed as $\sum_{n=1}^{N} \beta_n + \beta_{\mathcal{H}} = 1$. This concludes the proof for the completeness axiom. ■

Proof for the incompleteness axiom: Now we give proof for the last axiom. Assuming there is an index $i_0 \in I$ such that $\sum_{n=1}^{N} \beta_{n,i_0} < 1$, we must prove that

$$\sum_{n=1}^{N} \beta_n < 1, \text{ or equivalently } \beta_{\mathcal{H}} > 0.$$

By definition, we have

$$\beta_{\mathcal{H}} = \frac{m^Y(\mathcal{H}) - (\kappa_1 + \kappa_2)}{1 - (\kappa_1 + \kappa_2)} = \frac{\Delta' + K'}{1 - (\kappa_1 + \kappa_2)}.$$

So it is sufficient to show either $\Delta' > 0$ or $K' > 0$, say $\Delta' > 0$. Indeed, since

$$\sum_{n=1}^{N} \beta_{n,i_0} < 1$$

we have $m_{i_0}(\mathcal{H}) > 0$. This follows

$$w_{i_0} m_{i_0}(\mathcal{H}) \prod_{i \in I \setminus \{i_0\}} (1 - w_i) > 0$$

as $0 < w_i < 1$ for all $i \in I$. Thus, from (77), we easily deduce

$$\Delta' > 0$$

which we desired. This completely concludes the proof of the theorem. ■

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their insightful and constructive comments that have helped to significantly improve the presentation of this paper.

REFERENCES

[1] G. Balestra and A. Tsoukias, "Multicriteria analysis represented by artificial intelligence techniques," *Eur. J. Oper. Res.*, vol. 41, no. 5, pp. 419–430, 1990.

[2] V. Belton and T. J. Stewart, *Multiple Criteria Decision Analysis: An Integrated Approach*. Norwell, MA: Kluwer, 2002.

[3] B. G. Buchanan and E. H. Shortliffe, *Rule-Based Expert Systems*. Reading, MA: Addison-Wesley, 1984.

[4] W. F. Caselton and W. Luo, "Decision making with imprecise probabilities: Dempster–Shafer theory and application," *Water Resour. Res.*, vol. 28, no. 12, pp. 3071–3081, 1992.

[5] S. J. Chen and C. L. Hwang, *Fuzzy Multiple Attribute Decision Making—Methods and Applications*. Berlin, Germany: Springer-Verlag, 1992.

[6] T. Denoeux, "A k-nearest neighbor classification rule based on Dempster–Shafer theory," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, no. 5, pp. 804–813, May 1995.

[7] D. Dubois and H. Prade, "Representation and combination of uncertainty with belief functions and possibility measures," *Comput. Intell.*, vol. 4, no. 4, pp. 244–264, 1988.

[8] D. Dubois, H. Fargier, P. Perny, and H. Prade, "A characterization of generalized concordance rules in multicriteria decision making," *Int. J. Intell. Syst.*, vol. 18, no. 7, pp. 751–774, 2003.

[9] R. Fagin, J. Y. Halpern, and N. Megiddo, "A logic for reasoning about probabilities," *Inf. Comput.*, vol. 87, no. 1–2, pp. 78–128, 1990.

[10] T. Isitt, "The sports tourers," *Mot. Cycle Int.*, vol. 64, pp. 19–27, Sep. 1990.

[11] V. N. Huynh, Y. Nakamori, and T. B. Ho, "Assessment aggregation in the evidential reasoning approach to MADM under uncertainty: Orthogonal sum versus weighted sum," in *Advances in Computer Science—ASIAN 2004: Higher-Level Decision Making*, M. Maher, Ed. New York: Springer-Verlag, 2004, pp. 109–127. LNCS 3321.

[12] V. N. Huynh and Y. Nakamori, "A satisfactory-oriented approach to multi-expert decision-making under linguistic assessments," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 35, no. 2, pp. 184–196, Apr. 2005.

[13] C. L. Hwang and K. Yoon, *Multiple Attribute Decision Making Methods and Applications*. New York: Springer-Verlag, 1981.

[14] R. L. Keeney and H. Raiffa, *Decisions With Multiple Objectives: Preferences and Value Tradeoffs*. New York: Wiley, 1976.

[15] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, Mar. 1998.

[16] L. Levrat, A. Voisin, S. Bombardier, and J. Bremont, "Subjective evaluation of car seat comfort with fuzzy set techniques," *Int. J. Intell. Syst.*, vol. 12, no. 12, pp. 891–913, 1997.

[17] S. McClean and B. Scotney, "Using evidence theory for the integration of distributed databases," *Int. J. Intell. Syst.*, vol. 12, no. 10, pp. 763–776, 1997.

[18] G. A. Miller, "The magical number seven or minus two: Some limits on our capacity of processing information," *Psychol. Rev.*, vol. 63, no. 2, pp. 81–97, 1956.

[19] R. Roberts and P. Goodwin, "Weight approximations in multi-attribute decision models," *J. Multi-Criteria Decis. Anal.*, vol. 11, no. 6, pp. 291–303, 2002.

[20] M. Roubens, "Fuzzy sets and decision analysis," *Fuzzy Sets Syst.*, vol. 90, no. 2, pp. 199–206, 1997.

[21] T. L. Saaty, *The Analytic Hierarchy Process*. New York: McGraw-Hill, 1980.

[22] L. J. Savage, *The Foundations of Statistics*. New York: Wiley, 1954.

[23] P. Sen and J. B. Yang, "Multiple criteria decision making in design selection and synthesis," *J. Eng. Des.*, vol. 6, no. 3, pp. 207–230, 1995.

[24] K. Sentz and S. Ferson, "Combination of evidence in Dempster–Shafer theory," Sandia Nat. Lab. Albuquerque, NM, SAND Rep. SAND2002-0835, Apr. 2002.

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

[26] P. Smets and R. Kennes, "The transferable belief model," *Artif. Intell.*, vol. 66, no. 2, pp. 191–234, 1994.

[27] T. J. Stewart, "A critical survey on the status of multiple criteria decision making theory and practice," *OMEGA Int. J. Manag. Sci.*, vol. 20, no. 5/6, pp. 569–586, 1992.

[28] J. Wang, J. B. Yang, and P. Sen, "Multi-person and multi-attribute design evaluations using evidential reasoning based on subjective safety and cost analysis," *Reliab. Eng. Syst. Saf.*, vol. 52, no. 2, pp. 113–127, 1996.

[29] R. R. Yager, "On the Dempster–Shafer framework and new combination rules," *Inf. Sci.*, vol. 41, no. 2, pp. 93–137, 1987.

[30] ——, "Non-monotonic reasoning with belief structures," in *Advances in the Dempster–Shafer Theory of Evidence*, R. R. Yager, M. Fedrizzi, and J. Kacprzyk, Eds. New York: Wiley, 1994, pp. 534–554.

[31] *Advances in the Dempster–Shafer Theory of Evidence*, R. R. Yager, M. Fedrizzi, and J. Kacprzyk, Eds. New York: Wiley, 1994.

[32] R. R. Yager, "Decision-making under various types of uncertainties," *J. Intell. Fuzzy Syst.*, vol. 3, no. 4, pp. 317–323, 1995.

[33] J. B. Yang and M. G. Singh, "An evidential reasoning approach for multiple attribute decision making with uncertainty," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 1, pp. 1–18, Jan. 1994.

[34] J. B. Yang and P. Sen, "A general multi-level evaluation process for hybrid multiple attribute decision making with uncertainty," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 10, pp. 1458–1473, Oct. 1994.

[35] ——, "Multiple attribute design evaluation of large engineering products using the evidential reasoning approach," *J. Eng. Des.*, vol. 8, no. 3, pp. 211–230, 1997.
[36] J. B. Yang, "Rule and utility based evidential reasoning approach for multiple attribute decision analysis under uncertainty," *Eur. J. Oper. Res.*, vol. 131, no. 1, pp. 31–61, 2001.
[37] J. B. Yang and D. L. Xu, "Nonlinear information aggregation via evidential reasoning in multiattribute decision analysis under uncertainty," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 32, no. 3, pp. 376–393, May 2002.
[38] ——, "On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 32, no. 3, pp. 289–304, May 2002.
[39] J. B. Yang, Y. M. Wang, D. L. Xu, and K. S. Chin, "The evidential reasoning approach for MCDA under both probabilistic and fuzzy uncertainties," *Eur. J. Oper. Res.*, 2005, to be published.
[40] J. Yen, "Generalizing the Dempster–Shafer theory to fuzzy sets," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, no. 3, pp. 559–570, May–Jun. 1990.
[41] L. A. Zadeh, "Reviews of books: A mathematical theory of evidence," *Artif. Intell. Mag.*, vol. 5, no. 3, pp. 81–83, 1984.
[42] Z. J. Zhang, J. B. Yang, and D. L. Xu, "A hierarchical analysis model for multiobjective decision making," in *Analysis, Design and Evaluation of Man–Machine Systems*, selected papers from the 4th IFAC/IFIP/IFORS/IEA Conf., Xian, China, Sep. 1989. Oxford, U.K.: Pergamon, 1990, pp. 13–18.

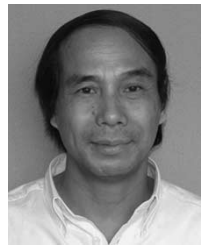
Van-Nam Huynh received the B.S. degree in mathematics from the University of Quinhon, Quinhon, Vietnam, in 1990 and the Ph.D. degree in applied mathematics from the Institute of Information Technology, Vietnamese Academy of Science and Technology, Hanoi, Vietnam, in 1999.

From 1990 to 2003, he worked as an Instructor and a Lecturer at the Departments of Mathematics and Computer Science, University of Quinhon. From 2001 to 2002, he was a Postdoctoral Fellow at the Japan Advanced Institute of Science and Technology, Ishikawa, Japan, under a fellowship awarded by the Inoue Foundation for Science. He is currently a Research Associate at the School of Knowledge Science, Japan Advanced Institute of Science and Technology. His main research interests include fuzzy logic and approximate reasoning, rough-fuzzy hybridization, decision analysis, and data mining.

Yoshiteru Nakamori (M'92) received the B.S., M.S., and Ph.D. degrees in applied mathematics and physics from Kyoto University, Kyoto, Japan, in 1974, 1976, and 1980, respectively.

He served at the Department of Applied Mathematics, Faculty of Science, Konan University, Kobe, Japan, as an Assistant Professor from 1981 to 1986, an Associate Professor from 1986 to 1991, and a Professor since 1991. From September 1984 to December 1985, he was with the International Institute for Applied Systems Analysis, Luxenberg, Austria, where he joined the Regional Water Policies Project. He joined the Japan Advanced Institute of Science and Technology, Ishikawa, Japan, in April 1998 as a Professor of the School of Knowledge Science. Since October 2003, he has been the Leader of a 21st Century Center of Excellence (COE) program on theory and practice of technology creation based on knowledge science, with the goals of program being to create a world-class center of excellence in the areas of theoretical research and practical research in knowledge science. His fields of research interest cover identification and measurement optimization of large-scale complex systems, modeling and control of environmental systems, and methodology and software of decision support systems. Recent activities include development of modeling methodology based on hard as well as soft data and support systems for soft thinking around hard data. Current topics include modeling and simulation for large-scale complex systems, system development for environmental policy-making support, and systems methodology based on Japanese intellectual tradition.

Dr. Nakamori is a member of the Society of Instrument and Control Engineers of Japan, the Institute of Systems, Control, and Information Engineers, the Japan Society for Fuzzy Theory and Systems, the Japan Association of Simulation and Gaming, and the Society of Environmental Science of Japan.



Tu-Bao Ho received the B.Tech. degree in applied mathematics from Hanoi University of Technology, Hanoi, Vietnam, in 1978, the M.S. and Ph.D. degrees in computer science from Pierre and Marie Curie University, Paris, France, in 1984 and 1987, respectively, and the Habilitation degree from Paris Dauphine University, in 1998.

He is currently a Professor at the School of Knowledge Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan. His current research interests include knowledge-based systems, machine learning, knowledge discovery, and data mining, in particular, bioinformatics and scientific data mining.

Dr. Ho has served on the Editorial Board of several international journals and program committees of various international conferences.

Tetsuya Murai (M'92) received the B.Sc. degree from the Department of Mathematics, Faculty of Science, Hokkaido University, Sapporo, Japan, in 1987 and the M.Eng. and Dr. Eng. degrees in mathematical information engineering from the Graduate School of Engineering, Hokkaido University, in 1987 and 1994, respectively.

From 1987 to 1992, he was a Lecturer at the School of Allied Health Professions, Sapporo Medical College. From 1992 to 1995, he was an Associate Professor at the Hokkaido University of Education, Hakodate, Japan. Since 1995, he has been an Associate Professor at Hokkaido University. In 1998, he was a Visiting Professor at the Catholic University, Brescia, Italy, supported by il Consiglio Nazionale delle Richerche (CNR) in Italy and the Japan Society for the Promotion of Sciences (JSPS) for collaboration with Prof. G. Resconi. His current research interests include modal logic, rough set theory and their applications to artificial intelligence, data mining, and granular computing.