Introduction to the *ER* Rule for Evidence Combination

Jian-Bo Yang and Dong-Ling Xu

Manchester Business School
The University of Manchester
Manchester M15 6PB, UK
{jian-bo.yang,ling.xu}@mbs.ac.uk

Abstract. The Evidential Reasoning (*ER*) approach has been developed to support multiple criteria decision making (*MCDM*) under uncertainty. It is built upon Dempster's rule for evidence combination and uses belief functions for dealing with probabilistic uncertainty and ignorance. In this introductory paper, following a brief introduction to Dempster's rule and the *ER* approach, we report the discovery of a new generic *ER* rule for evidence combination [16]. We first introduce the concepts and equations of a new extended belief function and then examine the detailed combination equations of the new *ER* rule. A numerical example is provided to illustrate the new *ER* rule.

Keywords: Evidential reasoning, Belief function, Evidence combination, Dempster's rule, Multiple criteria decision making.

1 Basic Concepts of Evidence Theory

The evidence theory was first investigated in 1960's [2] and formalised in 1970's [7]. It has since been further developed and found widespread applications in many areas such as artificial intelligence, expert systems, pattern recognition, information fusion, database and knowledge discovery, multiple criteria decision making (*MCDM*), audit risk assessment, etc. [1, 3, 5, 9–15]. In this section, the basic concepts of belief function and Dempster's combination rule of the evidence theory are briefly introduced as a basis for introduction of the Evidential Reasoning (*ER*) approach in the next section.

Suppose $H = \{H_1, \ldots, H_N\}$ is a set of mutually exclusive and collectively exhaustive propositions, referred to as the frame of discernment. A basic probability assignment (bpa) is a belief function $m : \Theta \to [0, 1]$, satisfying:

$$m(\emptyset) = 0, \text{ and } \sum_{C \in \Theta} m(c) = 1 \tag{1}$$

with $\emptyset$ being an empty set, C any subset of H, and Θ the power set of H, consisting of all the 2^N subsets of H, or

$$\Theta = \{\emptyset, \{H_1\}, \ldots, \{H_N\}, \{H_1, H_2\}, \ldots, \{H_1, H_N\}, \ldots, \{H_1, \ldots, H_{N-1}\}, H\} \tag{2}$$

Y. Tang, V.-N. Huynh, and J. Lawry (Eds.): IUKM 2011, LNAI 7027, pp. 7–15, 2011.

A basic probability mass $m(C)$ measures the degree of belief exactly assigned to a proposition C and represents how strongly the proposition is supported by evidence. Probabilities assigned to all the subsets of H are summed to unity and there is no belief left to the empty set. A probability assigned to H, or $m(H)$, is referred to as the degree of **global ignorance**. A probability assigned to any subset of H, except for any individual proposition $H_n (n = 1, \ldots, N)$ and H, is referred to as the degree of **local ignorance**. If there is no local or global ignorance, a belief function reduces to a conventional probability function.

Associated with each bpa to C are a belief measure, $Bel(C)$, and a plausibility measure, $Pl(C)$, defined by the following equations:

$$Bel(C) = \sum_{B \subseteq C} m(B) \text{ and } Pl(C) = \sum_{B \cap C \neq \emptyset} m(B) \tag{3}$$

$Bel(C)$ represents the exact support to C and its subsets, and $Pl(C)$ represents all possible support to C and its subsets. The interval $[Bel(C), Pl(C)]$ can be seen as the lower and upper bounds of support to C. The two functions can be connected by the following equation

$$Pl(C) = 1 - Bel(\overline{C}) \tag{4}$$

where $\overline{C}$ denotes the complement of C. The difference between the belief and plausibility measures of C describes the degree of ignorance in assessment to C [7].

The core of the evidence theory is Dempster's rule for evidence combination by which evidence from different sources is combined. The rule assumes that information sources are independent and uses the so-called orthogonal sum to combine multiple belief functions:

$$m = m_1 \oplus m_2 \oplus \ldots \oplus m_L \tag{5}$$

where $\oplus$ is the orthogonal sum operator. With two pieces of evidence m_1 and m_2, Dempster's rule for evidence combination is given as follows:

$$[m_1 \oplus m_2](\theta) = \begin{cases} 0, & \theta = \emptyset \\ \frac{\sum_{B \cap C = \theta} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B) m_2(C)}, & \theta \neq \emptyset \end{cases} \tag{6}$$

Note that Dempster's rule provides a non-compensatory process for aggregation of two pieces of evidence and can lead to irrational conclusions in the aggregation of multiple pieces of evidence in conflict [4, 6, 8], in particular in cases where multiple pieces of evidence are mutually compensatory in nature. By a compensatory process of evidence combination, it is meant that any piece of evidence is not dominating but plays a relative role, which is related to its relative importance. On the other hand, the *ER* approach [9, 11–15] introduced in the next section provides a compensatory evidence aggregation process, which is different from Dempster's rule in that it treats basic probability assignments as weighted belief degrees, embraces the concept of the degree of indecisiveness caused due to evidence weights, and adopts a normalisation process for combined probability masses without leaving any belief to the empty set.

2 The Main Steps of the *ER* Approach for *MCDM*

In the *ER* approach, a *MCDM* problem is modelled using a belief decision matrix. Suppose M alternatives $(A_l, l = 1, \ldots, M)$ are assessed on L criteria $e_i (i = 1, \ldots, L)$ each on the basis of N common evaluation grades (proposition) $H_n (n = 1, \ldots, N)$, which are required to be mutually exclusive and collectively exhaustive. If an alternative A_l is assessed to a grade H_n on a criterion e_i with a belief degree of $\beta_{n,i}(A_l)$, this assessment can be denoted by a belief function with global ignorance $S_i(A_l) = S(e_i(A_l)) = \{(H_n, \beta_{n,i}(A_l)), n = 1, \ldots, N, (H, \beta_{H,i}(A_l))\}$, with $\beta_{H,i}(A_l)$ used to measure the degree of global ignorance, $\sum_{i=1}^{N} \beta_{n,i}(A_l) + \beta_{H,i}(A_l) = 1$, $\beta_{n,i}(A_l) \geq 0$ $(n = 1, \ldots, N)$ and $\beta_{H,i}(A_l) \geq 0$. The individual assessments of all alternatives each on every criterion can be represented by a belief decision matrix, defined as follows:

$$D = (S_i(A_l))_{L \times M} \tag{7}$$

Suppose ω_i is the relative weight of the i^{th} criterion, normalised by

$$0 \leq \omega_i \leq 1 \text{ and } \sum_i \omega_i = 1 \tag{8}$$

The *ER* approach has both the commutative and associative properties and as such can be used to combine belief functions in any order. The *ER* aggregation process can be implemented recursively [11–13], summarised as the following main steps.

Step 1: *Assignment of basic probability masses*

Suppose the basic probability masses for an assessment $S_1(A_l)$ are given by:

$$\begin{aligned} m_{n,1} &= \omega_1 \beta_{n,1}(A_l) \text{ for } n = 1, \ldots, N, \\ m_{H,1} &= \omega_1 \beta_{H,1}(A_l), \text{ and} \\ m_{\Theta,1} &= 1 - \omega_1 \left(\textstyle\sum_{n=1}^{N} \beta_{n,1}(A_l) + \beta_{H,1}(A_l) \right) = 1 - \omega_1 \end{aligned} \tag{9}$$

In the evidence theory, $m_{n,1}$ may be interpreted as discounted belief. In *MCDM*, it should be interpreted as weighted belief or individual support to the assessment of A_l to H_n, as it means that in assessing an alternative A_l the 1^{st} criterion only plays a limited role that is proportional to its weight. $m_{H,1}$ represents the weighted global ignorance of the assessment. $m_{\Theta,1}$ is referred as to the degree of indecisiveness left by $S_1(A_l)$, representing the amount of belief that is not yet assigned to any individual or any subset of grades by $S_1(A_l)$ alone but needs to be jointly assigned in accordance with all other assessments in question.

Similarly, the basic probability masses for another assessment $S_2(A_l)$ are given by

$$\begin{aligned} m_{n,2} &= \omega_2 \beta_{n,2}(A_l) \text{ for } n = 1, \ldots, N, \\ m_{H,2} &= \omega_2 \beta_{H,2}(A_l), \text{ and} \\ m_{\Theta,2} &= 1 - \omega_2 \left(\textstyle\sum_{n=1}^{N} \beta_{n,2}(A_l) + \beta_{H,2}(A_l) \right) = 1 - \omega_2 \end{aligned} \tag{10}$$

Step 2: *Combination of basic probability masses*

The basic probability masses for $S_1(A_l)$ and $S_2(A_l)$ can be combined using the following *ER* algorithm:

$$\{H_n\}: \; m_{n,12} = k(m_{n,1}m_{n,2} + m_{n,1}(m_{H,2} + m_{\Theta,2}) + (m_{H,1} + m_{\Theta,1})m_{n,2}) \quad (11)$$

for $n = 1, \ldots, N$

$$\{H\}: \; m_{H,12} = k(m_{H,1}m_{H,2} + m_{H,1}m_{\Theta,2} + m_{H,2}m_{\Theta,1}) \quad (12)$$

$$\{\Theta\}: \; m_{\Theta,12} = k(m_{\Theta,1}m_{\Theta,2}) \quad (13)$$

$$k = \left(1 - \sum_{n=1}^{N} \sum_{t=1;t\neq n}^{N} m_{n,1}m_{t,2}\right)^{-1} \quad (14)$$

In the above *ER* algorithm, $m_{n,12}$ and $m_{H,12}$ measure the relative magnitudes of the total beliefs in the individual grade H_n and the frame of discernment H, respectively, generated by combining the two belief functions $S_1(A_l)$ and $S_2(A_l)$. $m_{\Theta,12}$ is the degree of indecisiveness left by both $S_1(A_l)$ and $S_2(A_l)$, representing the amount of belief that needs to be re-assigned back to all subsets of grades proportionally after the combination process is completed, so that no belief is assigned to the empty set. k measures the degree of conflict between $S_1(A_l)$ and $S_2(A_l)$.

Step 3: *Generation of total belief degrees*

If there are more than two assessments, Step 2 can be repeated to combine an uncombined assessment with the previously-combined assessment given by $m_{n,12}$ ($n = 1, \ldots, N$), $m_{H,12}$ and $m_{\Theta,12}$. After all assessments are combined recursively, the finally combined probability masses need be normalised to generate the total belief degrees $\beta_{n,12}$ and $\beta_{H,12}$ (for L=2) by proportionally re-assigning $m_{\Theta,12}$ back to all subsets of grades as follows:

$$\{H_n\}: \; \beta_{n,12} = \frac{m_{n,12}}{1 - m_{\Theta,12}}, \; n = 1, \ldots, N \quad (15)$$

$$\{H\}: \; \beta_{H,12} = \frac{m_{H,12}}{1 - m_{\Theta,12}} \quad (16)$$

The combined assessment for A_l is then given by the following belief function:

$$S(A_l) = \{(H_1, \beta_{1,12}), (H_2, \beta_{2,12}), \ldots, (H_N, \beta_{N,12}), (H, \beta_{H,12})\} \quad (17)$$

The above belief function provides a panoramic view about the combined assessment of the alternative A_l with the degrees of strength and weakness explicitly measured by the belief degrees.

3 Introduction to the *ER* Rule for Evidence Combination

In Section 2, a belief function with global ignorance was represented by $S_i(A_l) = S(e_i(A_l)) = \{(H_n, \beta_{n,i}(A_l)), n = 1, \ldots, N, (H, \beta_{H,i}(A_l))\}$, with $(H_n, \beta_{n,i}(A_l))$ referred to as a **focal element** of $S(e_i(A_l))$ if $\beta_{n,i}(A_l) > 0$. $m_{n,1} = \omega_1 \beta_{n,1}(A_l)$ given in Equation (9) represents the individual support of the evidence $S(e_1(A_l))$ to the hypothesis that A_l is assessed to H_n. Similarly, $m_{n,2} = \omega_2 \beta_{n,2}(A_l)$ given in Equation (10) represents the individual support of the evidence $S(e_2(A_l))$ to the same hypothesis. As such, $m_{n,1} m_{n,2}$ represents the joint support of both $S(e_1(A_l))$ and $S(e_2(A_l))$ to the same hypothesis.

Generally, suppose a piece of evidence $S(e_i)$ with the weight ω_i is represented by the following conventional belief function with $\sum_{\theta \in \Theta} \beta_{\theta,i} = 1$

$$S(e_i) = \{(\theta, \beta_{\theta,i}), \forall \theta \in \Theta\} \tag{18}$$

We can now show the extension of the above conventional belief function to include a special element $(\Theta, (1 - \omega_i))$ for constructing a new extended belief function for $S(e_i)$ as follows [16]:

$$m_i = \{(\theta, m_{\theta,i}), \forall \theta \in \Theta, (\Theta, m_{\Theta,i})\} \tag{19}$$

with

$$m_{\theta,i} = \omega_i \beta_{\theta,i}, \forall \theta \in \Theta, \text{ and } m_{\Theta,i} = 1 - \omega_i \tag{20}$$

Note that the following relationships between a conventional belief function and its extended belief function are always true [16]:

$$\beta_{\theta,i} = \frac{m_{\theta,i}}{1 - m_{\Theta,i}}, \forall \theta \in \Theta \tag{21}$$

We are now in a position to introduce the new *ER* rule [16] as follows. Let two pieces of independent evidence $S(e_1)$ and $S(e_2)$ with the relative weights ω_1 and ω_2 be represented by the conventional belief functions defined by Equation (18) with $\omega_1 + \omega_2 = 1$, $m_{\theta,1} = \omega_1 \beta_{\theta,1}$ and $m_{\theta,2} = \omega_2 \beta_{\theta,2}$ for all $\theta \subseteq H$. Then, $S(e_1)$ and $S(e_2)$ can be combined by the following *ER* rule which can be used recursively for aggregating multiple pieces of evidence [16]:

$$m_{\theta,12} = \frac{\widetilde{m}_{\theta,12}}{\sum\limits_{\theta \subseteq H} \widetilde{m}_{\theta,12} + \widetilde{m}_{\Theta,12}}, \forall \theta \subseteq H, \text{ and } m_{\Theta,12} = \frac{\widetilde{m}_{\Theta,12}}{\sum\limits_{\theta \subseteq H} \widetilde{m}_{\theta,12} + \widetilde{m}_{\Theta,12}} \tag{22}$$

$$\beta_{\theta,12} = \frac{\widetilde{m}_{\theta,12}}{\sum\limits_{\theta \subseteq H} \widetilde{m}_{\theta,12}}, \forall \theta \subseteq H, \tag{23}$$

$$\widetilde{m}_{\theta,12} = [(1 - \omega_2) m_{\theta,1} + (1 - \omega_1) m_{\theta,2}] + \sum_{B,C \subseteq H; B \cap C = \theta} m_{B,1} m_{C,2} \tag{24}$$

$$\widetilde{m}_{\Theta,12} = m_{\Theta,1} m_{\Theta,2} \tag{25}$$

The combined extended and conventional belief functions can then be represented as follows:

$$m_1 \oplus m_2 = \{(\theta, m_{\theta,12}), \forall\theta \in \Theta, (\Theta, m_{\Theta,12})\} \tag{26}$$

$$S(e_1) \otimes S(e_2) = \{(\theta, \beta_{\theta,12}), \forall\theta \in \Theta\} \tag{27}$$

where $\oplus$ is the orthogonal sum operator composed of Equations (22), (24) and (25) for generating combined extended belief functions, which can be applied recursively, and $\otimes$ is the *ER* operator consisting of Equations (23) and (24) for generating combined conventional belief functions, which can be used after extended belief functions are combined.

The new *ER* rule results from the innovation of implementing Dempster's rule on the new extended belief functions. It can be shown that the current *ER* approach as summarized in section 2 is a special case of the new *ER* rule. The new *ER* rule provides a generic process for generating total beliefs from combination of multiple pieces of independent evidence under the normal condition that each piece of evidence plays a role equal to its relative weight. The *ER* rule can be applied in areas where the above normal condition is satisfied, for example in multiple criteria decision making.

It is important to note that the combined belief generated by using the *ER* rule to aggregate two pieces of evidence is composed of two parts: the bounded average of the individual support, which is the first bracketed term in Equation (24), and the orthogonal sum of the joint support, which is the last term in Equation (24). This is in contract to the partial belief generated by using Dempster's rule on conventional belief functions, including only the orthogonal sum to count for joint support, with individual support either abandoned or assigned to the empty set, either of which is irrational.

4 Illustration of the *ER* Rule

We now examine a simple example to illustrate how the *ER* rule can be implemented and explain whether the results it generates are rational. Suppose three pieces of evidence of equal importance are given as the following three belief functions each with only its focal elements listed:

$$S(e_1) = \{(A, 0.99), (B, 0.01)\} \tag{28}$$

$$S(e_2) = \{(B, 0.01), (C, 0.99)\} \tag{29}$$

$$S(e_3) = \{(B, 0.01), (\{A, C\}, 0.99)\} \tag{30}$$

with $\omega_1 = \omega_2 = \omega_3 = 0.3333$. Suppose they each play a role equal to their relative weights. Note that

$$H = \{A, B, C\} \text{ and } \Theta = \{\emptyset, A, B, C, \{A, B\}, \{A, C\}, \{B, C\}, \{A, B, C\}\}$$

in this example. The extended belief functions corresponding to Equations (28)-(30) are given using Equations (19) and (20) as follows

$$m_1 = \{(A, 0.33), (B, 0.0033), (\Theta, 0.6667)\} \tag{31}$$

$$m_2 = \{(B, 0.0033), (C, 0.33), (\Theta, 0.6667)\} \tag{32}$$

$$m_3 = \{(B, 0.0033), (\{A, C\}, 0.33), (\Theta, 0.6667)\} \tag{33}$$

The calculations of the *ER* rule for the above example are shown in Table 1. The *ER* rule is applied recursively (Equations (22), (24) and (25)) to Equations (31)-(33). The results of the last row, generated using Equation (23) after the second iteration, show the final combined conventional belief function, complementary to the extended belief function, shown in the last but one row, which is generated by aggregating all the three extended belief functions shown in rows 5-7.

In the final results, $\beta_{A,123} = \beta_{C,123} = 0.3718$ are the highest total belief, which makes sense as the first evidence supports the proposition A and the second evidence supports the proposition C with the same magnitude, while the third evidence supports the proposition $\{A, C\}$ with no discrimination between the two individual propositions A and C. $\beta_{\{A,C\},123} = 0.2487$ is generated rightly as the second highest total belief as the third evidence supports $\{A, C\}$, so the significant local ignorance in $\{A, C\}$ should remain in the final results. The proposition B is assessed to be unlikely by all the three pieces of evidence, so it makes sense that the total belief in this proposition should also be rather small. The total belief in each of the other propositions ($\emptyset$, $\{A, B\}$, $\{B, C\}$ and $\{A, B, C\}$) is zero as it should be.

Table 1. Illustration of the *ER* Rule

Belief	A	B	C	$\{A, B\}$	$\{A, C\}$	$\{B, C\}$	$\{A, B, C\}$	Θ
$\beta_{\theta,1}$	0.99	0.01	0	0	0	0	0	
$\beta_{\theta,2}$	0	0.01	0.99	0	0	0	0	
$\beta_{\theta,3}$	0	0.01	0	0	0.99	0	0	
$m_{\theta,1}$	0.3300	0.0033	0.0000	0.0000	0.0000	0.0000	0.0000	0.6667
$m_{\theta,2}$	0.0000	0.0033	0.3300	0.0000	0.0000	0.0000	0.0000	0.6667
$m_{\theta,3}$	0.0000	0.0033	0.0000	0.0000	0.3300	0.0000	0.0000	0.6667
$\widetilde{m}_{\theta,12}$	0.2200	0.0045	0.2200	0.0000	0.0000	0.0000	0.0000	0.4444
$m_{\theta,12}$	0.2475	0.0050	0.2475	0.0000	0.0000	0.0000	0.0000	0.5000
$\beta_{\theta,12}$	0.495	0.01	0.495	0.0000	0.0000	0.0000	0.0000	
$\widetilde{m}_{\theta,123}$	0.2467	0.0050	0.2467	0.0000	0.1650	0.0000	0.0000	0.3333
$m_{\theta,123}$	0.2475	0.0050	0.2475	0.0000	0.1655	0.0000	0.0000	0.3344
$\beta_{\theta,123}$	**0.3718**	**0.0076**	**0.3718**	**0.0000**	**0.2478**	**0.0000**	**0.0000**	

5 Conclusion

In this paper, following the discussion of Dempster's rule and the *ER* approach, we reported the discovery of the new *ER* rule that provides a general process for combining multiple pieces of independent evidence in form of belief functions under the normal condition that every piece of evidence plays a limited role equivalent to its relative weight. The *ER* rule generates the total beliefs from combination of every two pieces of evidence as the addition of the bounded average of the individual support from each of the two pieces of evidence and the orthogonal sum of the joint support from the two pieces of evidence, which reveals that the orthogonal sum of the joint support from two pieces of evidence is only part of their total combined belief. A numerical example was examined in some detail to illustrate this general yet rational and rigorous *ER* rule for evidence combination. The new *ER* rule can be applied for combination of independent evidence in any cases where the normal condition is satisfied.

Acknowledgments. This work was supported by the UK Engineering and Physical Science Research Council under the Grant No: EP/F024606/1 and the Natural Science Foundation of China under the Grant No: 60736026.

References

1. Beynon, M.: DS/AHP method: A mathematical analysis, including an understanding of uncertainty. European Journal of Operational Research 140(1), 148–164 (2002)
2. Dempster, A.P.: Upper and lower probabilities induced by a multi-valued mapping. Annals of Mathematical Statistics 38, 325–339 (1967)
3. Denoeux, T., Zouhal, L.M.: Handling possibilistic labels in pattern classification using evidential reasoning. Fuzzy Sets and Systems 122(3), 409–424 (2001)
4. Haenni, R.: Shedding new light on zadeh's criticism of Dempster's rule of combination. In: The 7th International Conference on Information Fusion (FUSION) (2005)
5. McClean, S., Scotney, B.: Using evidence theory for knowledge discovery and extraction in distributed databases. International Journal of Intelligent Systems 12, 763–776 (1997)
6. Murphy, C.K.: Combining belief functions when evidence conflicts. Decision Support Systems 29, 1–9 (2000)
7. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
8. Wang, Y.M., Yang, J.B., Xu, D.L.: The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees. European Journal of Operational Research 175(1), 35–66 (2006)
9. Xu, D.L., Yang, J.B., Wang, Y.M.: The ER approach for multi-attribute decision analysis under interval uncertainties. European Journal of Operational Research 174(3), 1914–1943 (2006)
10. Yager, R.R.: Decision making using minimization of regret. International Journal of Approximate Reasoning 36, 109–128 (2004)

11. Yang, J.B., Singh, M.G.: An evidential reasoning approach for multiple attribute decision making with uncertainty. IEEE Transactions on Systems, Man, and Cybernetics 24(1), 1–18 (1994)
12. Yang, J.B.: Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties. European Journal of Operational Research 131, 31–61 (2001)
13. Yang, J.B., Xu, D.L.: On the evidential reasoning algorithm for multiattribute decision analysis under uncertainty. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans 32(3), 289–304 (2002)
14. Yang, J.B., Wang, Y.M., Xu, D.L.: The Evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties. European Journal of Operational Research 171(1), 309–343 (2006)
15. Yang, J.B., Xu, D.L., Xie, X.L., Maddulapalli, A.K.: Multicriteria evidential reasoning decision modelling and analysis - prioritising voices of customer. Journal of the Operational Research Society 62, 1638–1654 (2011)
16. Yang, J.B., Xu, D.L.: The evidential reasoning rule for evidence combination. Artificial Intelligence (submitted to, 2011)