A Geometrical Approach to Evidential Reasoning

Chua-Chin Wang and Hon-Son Don
Department of Electrical Engineering
State University of New York
Stony Brook, NY 11794

Abstract- **A model for evidential reasoning is proposed, in which the belief function of a piece of evidence is modeled as a probability density function, which can be a continuous or discrete form. We consider a polar notation of mutual dependency relationship between the evidence, in which the dependency between two interrelated pieces of evidence is described by an angle. This method can resolve the conflict resulting from either the mutual dependency among many pieces of evidence or the structural dependency in an inference network due to the evidence combination order. Belief conjunction, belief combination, belief propagation procedures and AND/OR operations of an inference network based on the proposed model are all presented. Some examples are given to demonstrate the advantages of this method over the conventional approaches.**

I. INTRODUCTION

Evidential reasoning, which has been an essential part of many computational system, is a task to assess a certain hypothesis when certain pieces of evidence are given, and the result is a new piece of evidence representing the consensus of the original pieces of evidence, which in turn can be applied to another stage of reasoning. There are three major frameworks of evidential reasoning in the literature, i.e., the Dempster-Shafer theory of evidence, the fuzzy set theory, and the Bayesian probability theory [4], [5], [6]. The advantages and weaknesses of these three frameworks have been discussed in [1], [2], [7], [8], and [9]. The application of Shafer's belief function to manage uncertainty of information in a rule-based system has attracted much attention in artificial intelligence research. The non-robustness of this model has been discussed in [3], [10]. Besides this drawback, the basic probability assignment (BPA) of a belief function is in a form of discrete type function which can not always provide a precise description of any evidence for all the situations. The possible quantization problem caused by thresholding a continuous region for the weight of evidence has been discussed in [1]. The continuous form of belief function, which is a more general representation, is more appropraite for the expression of the vagueness of an evidence in many situations. With regard to the dependency relationship between many pieces of evidence, Hau proposed a coefficient between maximally dependent and independent cases to indicate the degree of dependency [3]. However, the bilateral mutual dependency relationship has not been discussed. Such a dependency relationship in an inference network has never been seriously addressed in the literature.

In this paper, we propose a novel method which focuses the efforts on achieving conflict resolution of belief combination resulting from mutual dependency of many pieces of evidence in an inference network. It can also handle the information aggregation based upon the continuous belief functions, which is closer to the human reasoning process.

II. THEORY OF THE POLAR MODEL

A. Representation of Evidence and Rule

In the belief function introduced by Shafer [6] and Hau [3], two parameters, i.e., lower bound and upper bound, are employed to indicate *credibility* and *plausibility*. For the sake of clarity, the *belief function* is borrowed to represent the probability density function associated with a piece of evidence in the following text, and the belief function proposed by Shafer, [6], is named as *Shafer's belief function*. As discussed earlier, the probability assignment strategy for Shafer's belief function has its inherent drawback. For example, if a piece of evidence is to emphasize that the closer it is to the truth, the stronger it is, then that evidence can be conveniently modeled by a continuous function, which is a probability density function,

$$Bel(\theta) = k \cdot \theta, \qquad (1)$$

where θ is in the interval $[0,1]$ indicating the truth of the evidence, and k is a constant. We can hardly find any significant thresholds to quantize the associated belief function into a discrete form which can be handled

ISSN# 0-7803-0233-8/91 $1.00©1991 IEEE

by either Dempster-Shafer theory [6] or Hau's modified Demsper's rule [3]. Therefore, a more general representation of evidence is needed to represent such kind of uncertainty.

Definition 1 : A piece of evidence or a rule in a rule-based system is represented by a subset A of the frame of discernment Θ, and a belief function associated with A is represented by a probability density function $p_A(\theta)$, where θ is a variable indicating the degree of truth for the evidence. $\overline{A}$ denotes the complement of A. 1 is used to denote the truth of the evidence and 0 is used to denote the falsity of the evidence. θ is a numerical value in the interval $[0, 1]$.

For the sake of simplicity, 0 is used to denote the complement of the authenticity of a piece of evidence, and 1 is used to denote the truth of the piece of evidence.

Definition 2 : A rule R in a rule-based system conveying uncertainty is represented by the following format

$$R : E \longrightarrow H, \qquad \text{with } p_{E\rightarrow H}(\theta),$$

where E is a piece of evidence, H is a hypothesis implied by E, and $p_{E\rightarrow H}(\theta)$ is a probability density function which describes the degree of truth of the rule. The rule $R : E \longrightarrow H$ in the above definition is interpreted as logic implication, where E is called the ***antecedent*** and H is called the *consequent*.

B. Polar Model and Belief Conjunction

By definition, belief conjunction refers to the deduction of the belief associated with $(A \cap B)$ from the belief associated with the two pieces of evidence A and B, respectively. That is, given two frames of discernment Θ_A and Θ_B, a compatibility relation between Θ_A and Θ_B is the Cartesian product of them, which is represented as

$$\Theta_A \times \Theta_B \longrightarrow \Theta_{A\cap B} \tag{2}$$

We can apply the logic operation symbol to represent the above definition,

$$Conj(A, B) = A \cap B \tag{3}$$

where $Conj(A, B)$ represents the result of belief conjunction of A and B.

When we emphasize that one piece of evidence A is dependent on another piece of evidence B, we are also aware of that the conjunction or combination result of A and B is dominated by the evidence B from the analog of human reasoning. That is, in human reasoning, if the degree of dependency of an evidence A on another evidence B is getting stronger, then the belief conjunction result is naturally biased toward the latter one, B. In the extreme case, if the evidence A is totally dependent on B, then the conjunction result should be the same as the evidence B. Henceforth, an appropriate representation of dependency between two pieces of evidence is needed. Referring to Fig. 1, suppose we have two pieces of evidence A and B conveying, respectively, belief functions $p_A(\theta)$ and $p_B(\theta)$. Motivated from (2), we place these two pieces of evidence, respectively, on the two axes μ and ν of the Cartesian coordinates. Hence, a 2-D surface will be formed by the product operation in (2),

$$p_{surface}(\mu, \nu) = p_A(\mu) \cdot p_B(\nu) \tag{4}$$

The above expression is a 2-D function forming a 2-D surface which should be somehow transformed into 1-D form for the belief function of the conjunction result. In other words, the whole 2-D function has to be properly projected onto a single line on the μ-ν plane. This line is called the *conjunction line* on which the probability density function will denote the belief function of the conjunction of the two given pieces of evidence. We suggest to choose the following line as the conjunction line :

$$L_{conj} : \qquad \nu = m \cdot \mu = \tan\alpha \cdot \mu \tag{5}$$

where α is the angle between L_{conj} and μ-axis. Since the mutual dependency between two pieces of evidence can be interpreted as the relative degree of the conjunction result biased towards individual evidence. Therefore, the angle of the conjunction line forms a natural index of the degree of the dependency. For example, if α is 0, then the conjunction line will reside on the μ-axis in Fig. 1, i.e., the evidence B completely depends on A; if α is $\frac{\pi}{2}$, then the evidence A completely depends on B; if α is $\frac{\pi}{4}$, the conjunction result is not biased towards either of the two pieces of evidence, which indicates the nature of independence. Henceforth, we can define the degree of mutual dependency as follows.

Definition 3 : If two pieces of evidence A and B are arranged as the configuration in Fig. 1, the tangent of the angle α between the conjunction line and the μ-axis is called the degree of dependency of evidence B upon evidence A, which is denoted by

$$\rho_{BA} \;=\; \tan\alpha \;(=\; m).$$

The dependency relationship between two pieces of evidence will be treated as directed links, which means ρ_{AB} is not equal to ρ_{BA}. However, the following relationship holds,

$$\rho_{AB} = \frac{1}{\rho_{BA}}.$$

Note that the range of α is $[0, \frac{\pi}{2}]$ and that the degree of dependency defined in the above definition is in the range $[0, \infty]$, instead of the conventional range $[0, 1]$. Let the variable ω denote the degree of belief of the conjunction result of two pieces of evidence A and B. If evidence B depends on evidence A with a degree of dependency $m = \tan\alpha$, then for a certain value of ω it implies that the strength of a certain degree of belief of evidence B is the same as one mth of the degree of belief of evidence A. In other words, all of the probability density product $p_A(\mu)\cdot p_B(\nu)$, where $\nu + \frac{1}{m}\mu = \omega$, should be attributed to $p_{conj}(\omega)$, i.e.,

$$p_{conj}(\omega) = \int_{l.b.}^{u.b.} p_A(\mu)\, p_B(\omega - \frac{1}{m}\mu)d\mu \qquad (6)$$

where *l.b.* and *u.b.* denote the *lower bound* and the *upper bound* of the region for the integral. Note that $\nu + \frac{1}{m}\mu = \omega$ forms a line, which is called the *cumulative line* associated with ω. Referring to Fig. 2, which are top views of the model in Fig. 1, we can derive the *l.b.* and the *u.b.* for different ω.

$$u.b. = \min(m\omega, 1), \; l.b. = \max(m(\omega - 1), 0) \quad (7)$$

The following property of conservation of probability for the belief function under belief conjunction can be proved.

Lemma 1 :

$$\int_0^{1+\frac{1}{m}} p_{conj}(\omega)d\omega = 1 \qquad (8)$$

It is noted that the range of ω is $[0, 1 + \frac{1}{m}]$, instead of $[0, 1]$. This means a normalization step is needed to normalize the result $p_{conj}(\omega)$ in (6) to an eligible belief function so that it can be utilized in another stage of conjunction operation.

$$p_{conj}^N(\theta) = (1 + \frac{1}{m})p_{conj}((1 + \frac{1}{m})\theta), \quad 0 \le \theta \le 1, \quad (9)$$

where p_{con}^N denotes the normalized belief function of the conjunction result.

C. Dependency Propagation Problem

Consider the inference network shown in Fig. 3. Let evidence B and C be conjuncted first, which will produce an *intermediate evidence*, D, which possesses the same properties as an evidence. Since an intermediate evidence is conjuncted by two pieces of evidence, the dependency relationship of the conjuncted intermediate evidence to other evidence can be assumed to be linear interpolation of the individual dependency relationship of each terminal evidence, which complys with the nature of human reasoning. And the linear interpolation coefficients are the relative weights of the two pieces of evidence with respect to the intermediate evidence. Hence, before evidence D and E are conjuncted to assess the final uncertainty of the belief conjunction of the three pieces of evidence B, C, and E, the dependency relationship between the intermediate evidence D and the evidence E must be determined basing upon the given informations which are dependency relationships of any two of three pieces of evidence, B, C, and E. That is, the dependency relationship between the given pieces of evidence, B and C, will propagate through the inference network to build up the dependency relationship of the consequent intermediate evidence D.

Suppose we are given the following mutual dependency relationships,

$$\begin{aligned} \rho_{CB} &= m_1 = \tan\alpha_1, \qquad (10) \\ \rho_{EB} &= m_2 = \tan\alpha_2, \\ \rho_{EC} &= m_3 = \tan\alpha_3. \end{aligned}$$

The dependency degree of the evidence B on C is $\tan(\frac{\pi}{2} - \alpha_1)$, which is equal $\frac{1}{m_1}$. Similarly we can also get ρ_{BE}, and ρ_{CE}. Therefore, the mutual dependency relationship between E and D can be determined by,

$$\begin{aligned} m_4 &= \frac{\frac{1}{m_1}}{\frac{1}{m_1} + m_1} \cdot m_2 + \frac{m_1}{\frac{1}{m_1} + m_1} \cdot m_3 \\ &= \frac{1}{m_1^2 + 1} \cdot m_2 + \frac{m_1^2}{m_1^2 + 1} \cdot m_3, \end{aligned}$$

where $\rho_{DE} = m_4 = \tan\alpha_4$. Hence the following proposition will hold,

Dependency Propagation Proposition : Given the configuration shown in Fig. 3 and the degree of dependency in (10), the degree of dependency of evidence D on E is given by

$$\alpha_4 = \tan^{-1}(\frac{1}{\tan^2\alpha_1 + 1}\tan\alpha_2 + \frac{\tan^2\alpha_1}{\tan^2\alpha_1 + 1}\tan\alpha_3) \qquad (11)$$

and $\rho_{DE} = m_4 = \tan\alpha_4$,

where ρ_{DE} is the degree of evidence D depending on E.

D. Belief Combination

Belief combination refers to the belief conjunction of several pieces of evidence or facts supporting the same goal hypothesis. Since the basic probability assignment method discussed in [3] and [8] will introduce a conflict $(A \cap \overline{A})$, a conflict resolution approach is needed to achieve the consistency of belief combination. In our proposed model, if both pieces of evidence A and B are supporting hypothesis C, then the probability density function projected on the conjunction line is the belief function of C. Therefore, Eqns (2) to (11) can also be applied to the combination of two pieces of evidence, except that the two pieces of evidence must support the same hypothesis. The deficiency and nonrobustness of Dempster's rule have been discussed in detail in Hau's work [3]. However, both Hau and Shafer ignored the vagueness of the dependency relationship between the two pieces of evidence. The mutual dependency relationship of pieces of evidence always introduces significant conflict in the reasoning result of an inference network. Referring to Fig. 4, if there are three pieces of evidence, M, N, and K, supporting a hypothesis L, then in a sequential rule-based system, two of them have to be combined first, and then the result is combined with the third evidence. Two cases are shown in Fig. 5. In the example of Section III, we will demonstrate the conflict resulting from the combination order in an inference network.

E. Belief Propagation

Belief propagation refers to the aggregation of the uncertainty associated with the evidence or fact to fire a rule and the uncertainty of the rule itself so as to deduce the uncertainty of the goal hypothesis of the rule. Given a rule $R : P \longrightarrow Q$ and an evidence P, we are interested in exploring the belief function of the conclusion Q supported by the evidence P, which is defined as the belief propagation result of the evidence P and the rule R. Because it is impossible to obtain the exact belief function of the consequent Q from the given evidence and rule, what we can expect is an assessment of the *maximum bound* and *minimum bound* of the belief function associated with Q. If one piece of evidence A is covered by another evidence B, i.e., the information of A is contained in that of B, then we denote their relationship by $A \subseteq B$. Using this notation, the relationship among the maximum bound Q_{max}, the minimum bound Q_{min}, and Q can be expressed as

$$Q_{min} \subseteq Q \subseteq Q_{max}. \tag{12}$$

Let $p_{P \rightarrow Q}(\theta)$ be the belief function associated with the rule R and $p_P(\theta)$ be the belief function associated with the evidence P. The conjunction of the rule and the evidence is

$$(P \rightarrow Q) \cap P = (Q \cap P) \tag{13}$$

The conjunction result $(Q \cap P)$ means the consequent Q holds when it is supported by the antecedent P. Therefore, the belief function of this conjunction provides the minimum bound of the belief function of Q, i.e., $Q_{min} = (Q \cap P)$. This result meets the definition which we have explored for the belief propagation. Therefore, by applying the conjunction procedure of Section II.B to a piece of evidence and a rule, we will get the result of belief propagation. In other words, the belief function obtained by the conjunction of the belief functions of the antecedent and the rule is actually the minimum bound of the belief function of the consequent.

However, we are also interested in assessing the maximum bound of the belief function of the consequent Q. Referring to [3], the smallest range of $\overline{Q}$ is $(P \cap \overline{Q})$ which can also be derived from basic logic operations. Hence, we conclude that

$$Q_{max} = \overline{(P \cap \overline{Q})} = P \rightarrow Q \tag{14}$$

which implies the maximum bound of the belief function of Q is exactly the same as the belief function of the given rule despite what the belief function of P is. In other words, if the maximum bound is employed as the belief propagation result, then the uncertainty of the antecedent will not have any impact on the belief propagation. This is not true at all. Note that many researchers have proposed various approaches to recover and approximate the belief function of the consequent, and provided many explanations to their methods, e.g., [3]. However, we are only interested in the impact provided by the antecedent to the consequent which shows the degree of the antecedent supporting the consequent. Henceforth, we adopt only the minimum bound of the belief function of the consequent in a belief propagation procedure to assess the uncertainty aggregation of belief propagation.

In propositional logic, the logic implication, $A \rightarrow C$, can be synthesized by the conjunction of other logic implications, for example, the rules $A \rightarrow B$ and $B \rightarrow C$. It is easy to derive the following "*chaining syllogism*".

Lemma 3 : Given two rules

$$R_1 : \ A \rightarrow B, \text{ and } R_2 : \ B \rightarrow C \tag{15}$$

with the belief functions $p_{R_1}(\theta)$ and $p_{R_2}(\theta)$, respectively. The belief function $p_{R_3}(\theta)$ associated with the new rule, $R_3 : A \rightarrow C$, is the conjunction of the two belief functions, $p_{R_1}(\theta)$ and $p_{R_2}(\theta)$.

F. AND/OR operation

In an inference network, the function of each node is either AND or OR operation. In order to analyze the uncertainty aggregation of an inference network, it is necessary to consider these two operations for belief function. In Section II.B, we mentioned that the conjunction of two pieces of evidence are deemed as the AND operation of the two pieces of evidence, which is described as $Conj(A, B) = A \cap B$. Therefore, all of the theory of the belief conjunction given in Section II.B can be applied to the AND operation.

An OR operation for two pieces of evidence can be defined as $Union(A, B) = A \cup B$. From basic logic theory, we know

$$A \cup B = A + B - A \cap B,$$

which means the following equation holds,

$$p_{Union}(\theta) = p_A(\theta) + p_B(\theta) - p_{Conj}(\theta) \qquad (16)$$

where $p_{Conj}(\theta)$ can be derived according to the model introduced in Section II.B.

Since an inference network can be viewed as an AND/OR graph, if all of the pieces of evidence, rules, and the mutual dependency relationships are given, the belief function of the hypothesis can be derived by the proposed model.

III. SIMULATION EXAMPLE

Example 1. Considering the two cases in Fig. 5, which have different structures. Assume

$$Cr(M) = 0.98, \quad Pl(M) = 0.99, \quad \rho_{NK} = 0.5,$$
$$Cr(N) = 0.01, \quad Pl(N) = 0.02, \quad \rho_{NM} = 0.1,$$
$$Cr(K) = 0.01, \quad Pl(K) = 0.99, \quad \rho_{KM} = 0.9.$$

The inconsistent results of these two cases computed by Hau's method are tabulated in Table 1.

	Cr_L	Θ_L	$1 - Pl_L$
Case 1	0.006903	0.012857	0.980239
Case 2	0.003064	0.033357	0.963579

Table 1: The results of Hau's approach applied to cases of Fig. 5.

According to Table 1, the resulted credibility of case 1 is more than twice of that of case 2. On the contrary, the plausibility of case 1 is only half of that of case 2. These results indicate that Hau's method is easily subject to the combination order of the pieces of evidence. This is not consistent with the intuition of human reasoning. To a human, if these three pieces of evidence are given, and a belief function of the combination of these three pieces of evidence, then intuitively there should not be such a serious conflict no matter what the combination orders are. Therefore, we conclude that the results obtained by Hau's method are not consistent with the human reasoning. The reason why the conflict happens in this example is the mutual dependency relationship between evidence will propagate through the inference network during the multi-stage belief combination.

Given the same data, we use our method, i.e., Eqns. (2) to (11) to process each combination in the two cases in Fig. 5. The final results are listed in Table 2.

	Cr_L	Θ_L	$1 - Pl_L$
Case 1	0.010000	0.019700	0.970300
Case 2	0.010000	0.019700	0.970300

Table 2: The results of proposed model applied to either case of Fig. 5.

By applying the dependency propagation proposition in (11), the degree of dependency, ρ_{KG}, in Case 1 of Fig. 5 is 0.9109, and the degree of dependency, ρ_{NG}, in Case 2 of Fig. 5 is 0.2790. In Table 2, it is obvious that both cases have the same credibility Cr_L and plausibility $Cr_L + \Theta_L$, which means our model will not be seriously influenced by the order of the belief combination, and thus meets the intuition of human reasoning. The confliction appearing in Example 1 is resolved by our model. Here we have introduced a dependency propagation proposition, which turns out to be superior to the conventional scalar form for the dependency between pieces of evidence in resolving conflict caused by the dependency problem.

IV. CONCLUSION

A novel approach to manage uncertainty in rule-based systems is presented in this paper. Our model offers several advantages over previous works. First, the conflict caused by the combination order and the dependencies among many pieces of evidence in an inference network can be solved. Second, not only the discrete belief functions, but also the arbitrary continuous belief functions can be processed, which has not been explored up to date. Third, the dependency propagation problem of an intermediate evidence has been fully solved. Fourth, the conflict of belief propagation caused by the mutual dependency relationship of the antecedent of the rule and the rule itself is solved. The simulation results of the proposed model also turn out to be very appealing.

References

[1] Y. Cheng and R. L. Kashyap, "A study of associative evidential Reasoning," *IEEE Trans. Pattern Anal. Machine Intell.*, vol.11, no.6, June 1989.

[2] P. R. Cohen, *Heuristic reasoning about uncertainty: an artificial intelligence approach,* Morgan Kaufmann Publishers, Inc., Los Altos, California, 1985.

[3] H. Y. Hau and R. L. Kashyap, "Belief combination and propagation in a lattice-structured inference network," *IEEE Trans. Systems Man Cybernet,* vol. 20, no.1, Jan./Feb. 1990.

[4] J. Ihara, "Extension of conditional probability and measures of belief and disbelif in hypothesis based on uncertain evidence," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-9, no. 4, Jul. 1987

[5] J. Pearl, "Bayesian and belief-functions formalisms for evidential reasoning: a conceptual analysis," *Proc., 5th Israeli Symposium on Artificial Intelligence,* Tel-Aviv, Dec. 1988.

[6] G. Shafer and J. Pearl, *Readings in uncertain reasoning,* Morgan Kaufmann Publishers, Inc., San Mateo, California, 1990.

[7] H. Tahani and J. M. Keller, "Information fusion in computer vision using the fuzzy integral," *IEEE Trans. Systems Man Cybernet,* vol. 20, no.3, May/June 1990.

[8] L. E. Widman, K. A. Loparo and N. R. Nielsen, *Artificial intelligence, simulation, and modeling,* John Wiley & Sons Inc., 1989.

[9] J. Yen, "Generalizing the Dempster-Shafer theory to fuzzy sets," *IEEE Trans. Systems Man Cybernet,* vol. 20, no.3, May/June 1990.

[10] L. A. Zadeh, "A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination," *Fuzzy Set Syst.,* vol. 11, 1983

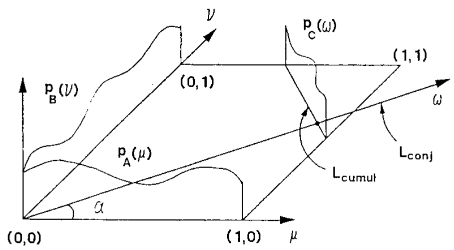

Fig. 1 Configuration of the polar model.

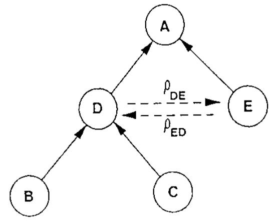

Fig. 3 An inference network for demonstrating the mutual dependency relationship.

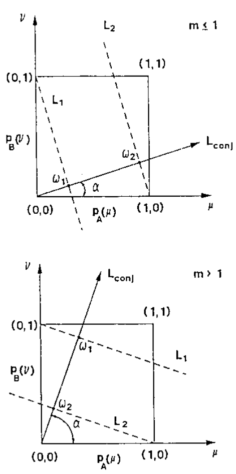

Fig. 2 The top views of Fig. 1.

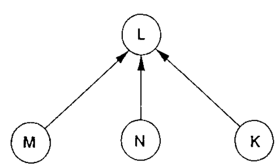

Fig. 4 Combination of three pieces of evidence to support a hypothesis.

Case 1:

Case 2:

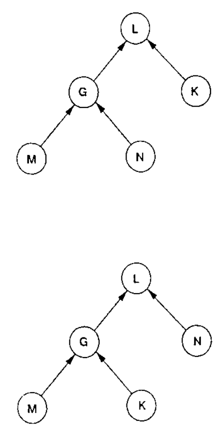

Fig. 5 Two different structures of combination of three pieces of evidences.