Control and Cybernetics

vol. **25** (1996) No. 2

An approach to system control based on combination-update of belief

by

Shengli Shi and D.A. Bell

University of Ulster
Northern Ireland, BT37 0QB, U.K.
{s.shi, da.bell}@ulst.ac.uk

Abstract: Researchers concerned with the control of various complex systems have become quite enthusiastic about fuzzy approaches. The need to transform observed data into fuzzy numbers and then to defuzzify the fuzzy numbers into control actions causes some problems when a general design paradigm for fuzzy controllers is sought. Now, a control system is essentially a knowledge-based system and its knowledge is uncertain, often being described in probabilistic terms. This prompts us to consider application of any of the available approaches to handling uncertain knowledge to control. In particular, approaches based on probability theory are attractive.

The Dempster-Shafer theory of evidence has proved powerful for handling uncertainty in other domains, and yet little work has been done in applying it in control. We have previously proposed a method, called the *Combination-Update* (*CU* for short) method, for handling uncertainty based on the theory. We also outlined a way to apply the *CU* method to system control.

In this paper we indicate a way of applying *non-truth functional methods* in complex system control. In our method, instead of using fuzzy logic as the decision making system, *Belief Logic*, which has been developed for the capture of the logical properties of the Dempster-Shafer theory of evidence , is adopted. Since Belief Logic is consistent with the two-valued logics our method will not suffer the documented problems associated with fuzzy logic.

Keywords: control system, uncertainty, probabilistic reasoning, belief logic.

1. Introduction

Fuzzy control (for survey see Yamakawa and Hirota, 1989; Lee, 1990) is now a very active research area in the application of fuzzy set theory, Zadeh (1965). It has been shown to be successful, in Japan especially, for the control of simple electric devices and in car components, for example. Now in addition to a knowledge base/database storing expert knowledge which is essentially uncertain, and a reasoning engine which is based on fuzzy logic, a typical fuzzy controller also consists of a *fuzzifier* and a *defuzzifier*. These are required because knowledge representation in a fuzzy controller is based on fuzzy sets and fuzzy logic, while the behaviour of the system to be controlled is usually measured by statistics or probabilities. The observed data has to be transformed into fuzzy numbers by a fuzzifier, and control commands which are fuzzy numbers have to be mapped to control actions by a defuzzifier.

The facts that knowledge in a control system is uncertain and the behaviour of controlled systems are often measured by statistics or probabilities suggests that other approaches to handling uncertainty based on probability theory can be candidates in complex system control. An example is the Dempster-Shafer theory of evidence , Shafer (1976), Guan, Bell (1991).

On the other hand, logical systems are necessary for making decisions in a control system. An approach for handling uncertainty to be applied to system control must be accompanied by a logical component. When extending standard two-valued logics to enable them to handle uncertainty, we would hope that the extended system is consistent with standard two-valued logics. That is, if all propositions are certain then the logical system should be identical to the standard two-valued logic. If so, the logical system is said to be *consistent* with the standard two-valued logic.

In more detail, if there exists a function f defined on all formulas of a logical system that can be decomposed with logical connectives then the logical system is said to be *truth functional* with respect to f. One way to ensure that a logical system handling uncertainty is consistent with the standard logical system is to define such a function to represent uncertainties of propositions[1], i.e. to make the system truth functional with respect to the uncertainty measure.

But there is a difficulty here. Bundy (1985) has proved a general result that the simple combination of any purely numerical method for handling uncertainties and standard two-valued logic may result in inconsistencies, i.e., "the logical connectives cannot be truth functional with respect to probabilities, or any other purely numeric uncertainty measure." So a direct extension of standard logical systems with the mechanism of handling uncertainty cannot give the desired result — that two-valued logics should be trivial cases. This fact gives the motivation for the present study.

In this paper we indicate a way of applying non-truth functional methods in complex system control. The Dempster-Shafer theory of evidence is adopted as a basis for our approach. This has been proved to provide a powerful tool for handling uncertainty, Guan, Bell (1991), and some knowledge-based systems based on the Dempster-Shafer theory of evidence have been established, see for example Kak (1990), Saffiotti, Umkehrer (1991).

[1] A particular case is when f is a set membership function.

An important feature of the Dempster-Shafer theory of evidence is that it permits an explicit representation of *ignorance* and does not assume that degrees of beliefs in a set of propositions are uniformly allocated to the individual propositions. The investigation of the logical properties of ignorance has resulted in a new logical system, *Belief Logic*, Bell, Shi, Hull (1995), which can express ignorance. This offers the possibility of expressing uncertainty without conforming to the requirement that a complete specification of uncertainties must be given before conducting reasoning, Pearl (1988). Furthermore, it allows our attention to be focussed on interesting propositions by ignoring all other factors. Since Belief Logic is not truth functional, it promises a solution to the problem that fuzzy set theory and probability theory have left unsolved.

Some researchers have tried to apply the Dempster-Shafer theory of evidence to control systems, see Inagaki (1991), Murphy (1991). We have previously suggested this also: Bell, Shi, Hull (1994), Shi, Hull, Bell (1994). In this paper we supplement the approach based on our previous work on dealing with uncertainty, which is called the *Combination-Update* (*CU* for short) method, Shi, Hull, n this method, the control processes are considered as repeatedly *combining* and *updating* beliefs of an agent (or controller) using Dempster's rule of combination – Shafer (1976) – and the *proportional sum* which is proposed in Shi, Hull, Bell (1994). Control decisions are made within the framework of Belief Logic. An immediate advantage of adopting the Dempster-Shafer theory of evidence in system control is that the observed data of controlled systems is consistent with the controller's internal representation, since the Dempster-Shafer theory of evidence is a generalization of probability theory.

The rest of this paper is organized as follows. In the next section we give a brief introduction to the Dempster-Shafer theory of evidence and Belief Logic and our rule for updating beliefs. In section 3 we describe how to apply the CU method in control systems.

2. Representing beliefs

In this section we give a description of Belief Logic, Bell, Hull, Shi (1995). First we introduce some basic concepts of the Dempster-Shafer theory of evidence . An updating rule, the proportional sum, Shi, Hull, Bell (1994), is also presented. For more details of the Dempster-Shafer theory of evidence the reader may consult Shafer (1976), or Guan, Bell (1991).

2.1. Mass and belief functions

The Dempster-Shafer theory of evidence generalizes the Bayesian method and gives an approach to inference with uncertainty in artificial intelligence, see Goodman, Nguyen (1985), Abel (1988). Each given piece of evidence supports a set of propositions which are represented as subsets of a set, called a *frame of discernment*. Informally, a frame of discernment is a set of all the possible truth-values that some propositions of interest can take. Each proposition is related to one subset of the frame of discernment. The basic propositions are exhaustive and mutually exclusive. The power set of a frame of discernment is called *an evidential space.* If a subset of a frame of discernment is considered as a set of *possible worlds* in which the corresponding propositions are true, then the subset represents a *proof* of the related proposition, where each possible world is an element of the proof. Because of the correspondence between a proof and a proposition, we may consider a proof as a proposition.

It is because a proof may not establish a proposition completely that the uncertainty arises. From the relationship of the proposition to the proof we have a subjective degree of belief for the proposition. To express this the *mass* of a proposition is introduced, which represents the *weightiness* of a proof of the associated proposition. The masses of all propositions of an evidential space are given by a *mass function.*

Let Θ be a frame of discernment. A *mass function* m on frame Θ is a mapping from evidence space 2^Θ to the real interval $[0,1]$ with the following restriction:

(m1) $m(\emptyset) = 0$,

(m2) $\sum_{A \in 2^\Theta} m(A) = 1$

We are interested only in propositions whose proofs have non-zero masses. We call such proofs *focal proofs* (or *focal elements*) of a mass function. The union of all focal proofs is called the *core* of the mass function. When we present a mass function we give values to focal proofs only. We define $\mathcal{M}_\Theta$ to be the set of all mass functions defined on the frame of discernment Θ, and use $\mathcal{P}(m)$ to denote the set of all focal elements (or propositions) of the mass function m, i.e., suppose Θ is a frame of discernment, then

$$\mathcal{P}(m) = \{A \mid A \in 2^\Theta \text{ and } m(A) > 0\}$$

Some other functions are used when conducting reasoning with uncertainty. Examples are the belief function and the plausibility function. These functions are collectively called *evidential functions.*

Let Θ be a frame of discernment and m be a mass function on Θ. The belief function *bel* and plausibility function *pls* induced by m are defined as follows. For each proposition A of Θ

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

$$pls(A) = 1 - bel(\overline{A}) = 1 - \sum_{B \subseteq \overline{A}} m(B)$$

PROPOSITION 2.1 (Shafer, 1976) *Let* $bel : 2^\Theta \to [0,1]$ *be a mapping.* bel *is a belief function if it satisfies the following conditions:*

(b1) $bel(\emptyset) = 0$,

(b2) $bel(\Theta) = 1$,
(b3) *for any collection* $\{A_1, A_2, ..., A_n\}$ $(n \geq 1)$ *of subsets of* Θ,

$$bel(A_1 \cup A_2 \cup ... \cup A_n) \geq \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} bel(\cap_{i \in I} A_i).$$

Intuitively, a belief function defines a lower bound on the uncertainty of a proposition, while the corresponding plausibility function gives an upper bound. Therefore we can use the ordered pair $(bel(A), pls(A))$ to describe the uncertainty of the proposition A.

2.2. Representing ignorance explicitly

Ignorance arises when drawing conclusions from an incomplete set of propositions. Given a set of propositions, if all possibilities are shared by the propositions, then no ignorance exists, since we know all possibilities. A special case is where there are only two propositions: A and its negation $\neg A$. A belief function bel may not satisfy $bel(A) + bel(\neg A) = 1$, because ignorance may be present.

Let Θ be a frame of discernment. Suppose m is a mass function on Θ. For any subsets $A_1,..., A_n$ of Θ, $n \geq 2$, the *ignorance* $\delta_{A_1,...,A_n}$ of $A_1,..., A_n$ is defined as follows:

$$\delta_{A_1,...,A_n} = \sum_{B \subseteq A_1 \cup ... \cup A_n; B \not\subseteq A_1,...,A_n} m(B)$$

The following proposition shows that if we add ignorance to the right hand of (b3), we then obtain an equality.

PROPOSITION 2.2 *Let* Θ *be a frame of discernment. Suppose bel is a belief function on* Θ. *For any subsets* $A_1,..., A_n$ *of* Θ, $n \geq 2$, *we have*

$$bel(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} bel(\cap_{i \in I} A_i) + \delta_{A_1,...,A_n}$$

By this proposition, we can rewrite (b3) as follows,
(b3') $bel(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} bel(\cap_{i \in I} A_i) + \delta_{A_1,...,A_n}$.
A belief function therefore is an extension of probability measure function by introducing ignorance explicitly, and hence, the Dempster-Shafer theory of evidence is an extension of probability theory. The above definition of ignorance is not particularly clear. The following proposition which is a special case of Proposition 2.2 can allow a interpretation to the ignorance.

PROPOSITION 2.3 *Let* Θ *be a frame of discernment. Suppose bel is a belief function on* Θ. *Then*

$$bel(A) + bel(\neg A) + \delta_{A,\neg A} = 1$$

Ignorance arises from interactions among some events. For a particular subset A, a belief function reflects all information about it. Because of ignorance we cannot derive $bel(\neg A)$ merely from $bel(A)$[2]. By the proposition 2.3 we have

$$bel(\neg A) = 1 - bel(A) - \delta_{A,\neg A}$$

So, if we know the ignorance associated with events A and $\neg A$ we can derive $bel(\neg A)$ from $bel(A)$ and the ignorance. On the other hand, in probability theory if we know the probability that A occurs then we also know the probability of event $\neg A$. It is easy to see that the following proposition holds.

PROPOSITION 2.4 *Let Θ be a frame of discernment. Suppose bel is a belief function on Θ. If for any subsets $A_1, \ldots, A_n$ of Θ, $n \geq 2$, $\delta_{A_1,\ldots,A_n} = 0$, then bel is a probability measure.*

2.3. Combining and updating evidence

In the Dempster-Shafer theory of evidence each mass function defines degrees of belief for a set of propositions. Given two mass functions on a frame which stem from two pieces of evidence we need a method to combine them. That is, we require to combine two mass functions so that we can get new, combined degrees of belief for the propositions. Dempster's rule of combination allows us to do this.

Let Θ be a frame and m_1 and m_2 be two mass functions on Θ. If $\sum_{B \cap A \neq \emptyset} m_1(A) \cdot m_2(B) \neq 0$ then m_1 and m_2 are said to be *combinable.*

Let m_1 and m_2 be two combinable mass functions on frame of discernment Θ. The *combination* $m_1 \oplus m_2$ is defined as follows. For each subset $A \neq \emptyset$ of Θ,

$$(m_1 \oplus m_2)(A) = \frac{\sum_{A_i, B_j \subseteq \Theta; A_i \cap B_j = A} m_1(A_i) \cdot m_2(B_j)}{\sum_{A_i, B_j \subseteq \Theta; A_i \cap B_j \neq \emptyset} m_1(A_i) \cdot m_2(B_j)}$$

This is called *Dempster's rule of combination.* There are different interpretations of it corresponding to different interpretations of the Dempster-Shafer theory of evidence , Halpern, Fagin (1992). The first viewpoint is of Dempster-Shafer theory as a generalization of Bayesian statistics, where the belief is given as generalized objective probability. The second is as a way of representing evidence, where the belief represents credibility.

Dempster's rule of combination gives a way to combine two agents' beliefs. On the other hand we also need to consider another case: where an agent's beliefs are revised in the light of another agent's beliefs. When we say that one agent's belief is *updated*[3] by another agent's belief, this means that all focal proofs related to the first agent remain unchanged, but the degrees of belief

[2]This indicates the absence of truth-functionality in this case

[3]We use the term *update* in this sense here. It should be noted that this term is sometimes reserved for a different concept — to reflect changes in the "world" being modelled.

of the propositions may be modified and therefore some focal proofs may no longer be focussed since their masses may be changed to zero. On the other hand, when *combining* two agents' beliefs, this means that a new set of focal proofs is obtained, which is an aggregated result for two agents, and the degrees of belief of focal proofs are also changed. This suggests a new rule for updating (rather than combining) evidence in the Dempster-Shafer theory of evidence . Since it uses a proportion to represent a combined mass, we call the new rule the *proportional sum.*

Let Θ be a frame of discernment. Let m_1 and m_2 be two combinable mass functions on Θ. Then the *proportional rule of update* or the *proportional sum*, $m_1 \oslash m_2$, of m_1 and m_2 is defined as follows: For all $A \in \mathcal{P}(m_1)$,

$$(m_1 \oslash m_2)(A) = \frac{\sum_{B \in \mathcal{P}(m_2); A \cap B \neq \emptyset} m_1(A) \cdot \frac{|A \cap B|}{|A|} \cdot m_2(B)}{c}$$

where

$$c = \sum_{A_i \in \mathcal{P}(m_1), B_j \in \mathcal{P}(m_2); A_i \cap B_j \neq \emptyset} m_1(A_i) \cdot \frac{|A_i \cap B_j|}{|A_i|} \cdot m_2(B_j)$$

and we say that m_1 is updated by m_2.

The intuitive idea of the proportional sum is clear: when updating a mass function, we only consider the masses of subsets which remain and we reallocate masses according to *the insufficient reasoning principle*, a principle which is conspicuously not used for belief combination. This states that if there is no other information about the probabilities of a set of events, we assume that these events occur with equal probabilities. It is not difficult to verify that when using a mass function m_2 to update another mass function m_1, all focal proofs of the updated result $m_1 \oslash m_2$ are also focal proofs of m_1.

Jeffrey (1983) considers the generalization of conditional probability, the Bayesian rule, which gives a way to update probabilities when more than one events occur.

Suppose $\langle M, S, \mu \rangle$ is a finite probability space, where M is a non-empty set, called *universe*; S is a σ–algebra consisting of subsets of M and μ is a complete probability measure[4]. Let $\langle M, S', \mu' \rangle$ be a finite probability space, where S' is a sub-algebra of S. Let $\mathcal{B} = \{A_1, ..., A_n\}$ be a *base* of S'. Jeffrey defines his generalized rule of conditionalization as follows:

$$\mu_{\mathcal{B}}(B) = \sum_{A_i \in \mathcal{B}} \mu(A_i \cap B) \cdot \mu'(A_i)/\mu(A_i)$$

We use the notation $\mu \bar{\oslash} \mu'$ instead of $\mu_{\mathcal{B}}$ in the above rule and we say that $\mu \bar{\oslash} \mu'$ is the result of updating μ with μ'. As we have shown in Bell, Hull, Shi

[4]If μ is not a complete probability measure, we can handle non-measurable subsets using the techniques of inner and outer probabilities, see Halpern, Fagin (1992).

(1995), if there is no ignorance for two mass functions m_1 and m_2, i.e. where m_1 and m_2 are probability measures, the proportional sum of m_1 and m_2 turns out to be Jeffrey's rule of conditionalization. That is, when there is no ignorance $\oslash$ and $\bar{\oslash}$ give the same result.

PROPOSITION 2.5 *Suppose that m_1 and m_2 are probability measures. Then*

$$(\mu \bar{\oslash} \mu')(A) = (m_1 \oslash m_2)(A)$$

This proposition is important in system control when there is enough information to establish a probability measure for the behaviour of a controlled system. In this case techniques based on probability theory can be applied. In cases where there is insufficient information, we need to include ignorance and hence use $\oplus$.

2.4. Belief logic

Since classical logics are not suitable for representing some features of knowledge such as uncertainty, imprecision and fuzziness in artificial intelligence, various extensions have been proposed, e.g. fuzzy logic, probabilistic logic, Nilsson (1984), and *Incidence calculus*, Bundy (1985). A weakness of systems based on probability theory is that all subsets must be assigned subjective probabilities in such a way that all axioms of probability theory are satisfied. In other words, it requires a complete specification of probabilities of all subsets. Although we can use some methods to reduce the number of assignments of probabilities, we are not permitted to concentrate only on some interesting propositions and ignore all other propositions.

Belief Logic can be considered as an extension of Incident Calculus which is based on the Dempster-Shafer theory of evidence . Therefore, it allows the explicit expression of ignorance. In Bell, Shi, Hull (1995) we have shown that if no ignorance exists or if we have enough information to establish a probability structure, beliefs turns out to be probabilities.

Let Θ be a frame of discernment. A *Belief Logic* is a structure:

$$BL = < \Theta, bel, \mathcal{L}, \mathcal{A}_{\mathcal{L}}, t >$$

where $\mathcal{L}$ is a logic language, *bel* is a belief function, t is a *truth assignment* defined on a set of propositions of $\mathcal{L}$, t assigns a subsets of Θ to each proposition.

(11) For any $\alpha \in \mathcal{L}$, if $t(\alpha)$ is defined then $t(\alpha) \subseteq \Theta$.
(12) $t(true) = \Theta$, $t(false) = \emptyset$
(13) $t(\neg\alpha) = \Theta - t(\alpha)$
(14) $t(\alpha \wedge \beta) = t(\alpha) \cap t(\beta)$
(15) $t(\alpha \vee \beta) = t(\alpha) \cup t(\beta)$
(16) $t(\alpha \longrightarrow \beta) = t(\neg\alpha) \cup t(\beta)$

Generally a truth assignment is not necessarily a total function. For example, it can be defined on the set of axioms used in Incidence Calculus, Bundy (1985).

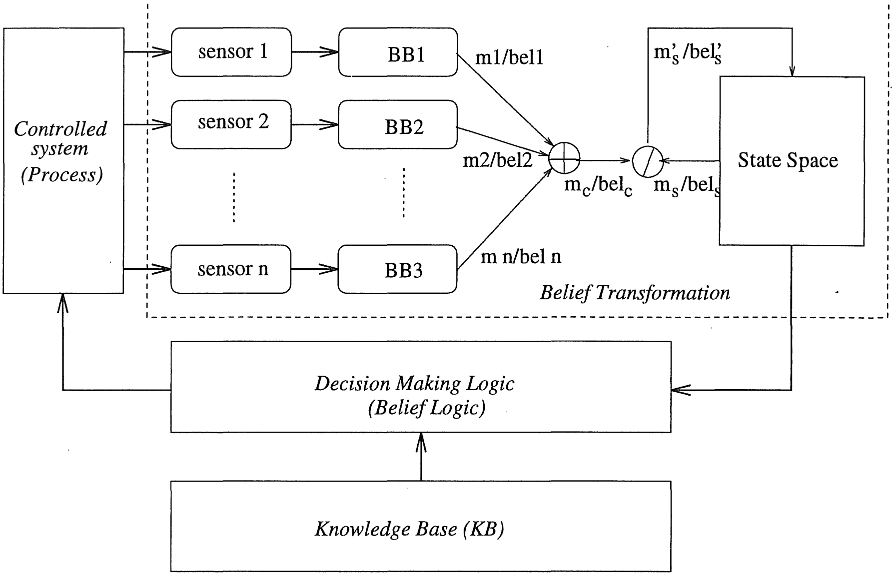

Figure 1. Basic diagram of *CU* controllers.

Belief Logic retains the main advantage of Incidence Calculus — that logical connectives are truth functional with respect to truth assignment. It should be noted that, in Belief Logic, logical connectives are not truth functional with respect to uncertainty measure or the belief function. The method of finding a legal assignment in Incidence Calculus, Bundy (1986) can then be used in Belief Logic.

3. CU controllers

In this section we describe the basic ideas of applying the Dempster-Shafer theory of evidence in system control. The controllers are called *CU controllers*, since the reasoning process is a repeated application of the combining and updating rules described in Section 2.

3.1. A general scheme

The application of the Dempster-Shafer theory of evidence to system control is inspired by the *CU* method, Shi, Hull, Bell (1994), for handling uncertainty in Knowledge Based systems. A control system is considered as a special Knowledge Based system (*CU controller*).

Figure 1 gives a basic configuration of a *CU controller* which, apart from the controlled system (process), consists of three components: *belief transformation, Knowledge Base* and *decision making logic.*

The behaviour of a controlled system is captured by sensors and then transformed into mass/belief functions by *belief builders* $BB_1,...,\ BB_n$. All these mass/belief functions are combined together using Dempster's rule of combination. The combined result is then used to update the system state which is also represented as a mass/belief function. Knowledge of application domain is stored in the Knowledge Base. The core of a *CU controller* is the decision making logic which is based on Belief Logic.

Notice that representation of the system states requires 3 evidence-handling operations — transformation by uncertain rules (each having a rule strength) is required as well as the combination and update discussed above. The transformation process is application domain based. The rules represent knowledge about the system. Some of them permit translation to be made between evidence items referring to different entities. These can be used to homogenise sensor evidence. For example, if a controlled system state transition depends on temperature change, but pressure (rather than temperature) is monitored by the sensors, then a rule relating temperature to pressure can be used to translate the pressure evidence to temperature evidence. The evidence qualities might be further compromised by the use of this (uncertain) rule. A simple function of rules for transformation is to adjust the amount of belief (confidence) that we assign to some evidence from a sensor which is itself not entirely trustworthy. For example, if a sensor registers "vehicles in any monitoring range" with a certain degree of confidence, it is surely interesting if that sensor has only a 90% reliability. The degree of confidence must be adjusted or transformed in the light of this imperfection.

The reasoning process is as follows. From sensors a set of pieces of evidence is obtained and from each piece of evidence a mass/belief function is derived, i.e. $m_1/bel_1,\ ...,\ m_n/bel_n$. Beliefs represented by $m_1/bel_1,\ ...,\ m_n/bel_n$ are *combined* by Dempster's rule of combination. A combined result which is also a mass/belief function is obtained, i.e.

$$m_c = m_1 \oplus m_2 \cdots \oplus m_n.$$

or

$$bel_c = bel_1 \oplus bel_2 \cdots \oplus bel_n.$$

The (initial) state of system is represented as a mass/belief function m_s/bel_s which is updated by the combined belief m_c/bel_c. The system state is transformed, by rules in the Knowledge Base. Actions are decided by the Decision Making Logic Unite — used to control the system. The new behaviour of the controlled system is continually captured by sensors and then the above *combining-updating* process is repeated.

3.2. Belief transformation

The essential part of a *CU controller* is the Belief Transformation which comprises three parts: sensors, belief builders and state space as shown in Figure

1. Observed data related to the behaviour of the controlled system is often expressible in a statistical or probabilistic form. A belief builder associates the observed data with beliefs or masses. Since the Dempster-Shafer theory of evidence permits an explicit representation of ignorance and does not assume that degrees of beliefs in a set of propositions are uniformly allocated to the individual propositions, a sensor's attention can be concentrated by the corresponding belief builder on some interesting states.

EXAMPLE 3.1 *Many factors such as temperature, wind direction and wind strength will affect a sensor in detecting the speed of a flying object. However, if the wind direction and strength are the main factors, we might want to focus on them by ignoring the others. A frame of discernment is formed.*

$$\Theta = \{same - direction, strong - wind, high - temperature\}$$

where $same - direction$ *represents that the flying object has the same direction as the wind; and the other factors,* $strong - wind$ *and* $high - temperature$ *have the obvious meanings. A mass function which focusses as described above can be given as follows.*

$$m(\{same - direction\}) = 0.45, \; m(\{strong - wind\}) = 0.35, \; m(\Theta) = 0.2$$

Note that the joint effect of temperature, wind direction and wind strength is expressed by mass of Θ *— the ignorance associated with* m.

A CU controller allows multiple evidence, i.e. more than one sensor can be used to detect the same parameter of the controlled system. This is particularly useful for critical system control, where the reliability of observed data is enhanced by multiple sensors. Dempster's rule of combination presents a way to combine the beliefs given by such a collection of sensors. Suppose that bel_1, ..., bel_n are belief functions produced by belief builders BB_1,..., BB_n. Then

$$bel = bel_1 \oplus \cdots bel_n$$

gives the combined results of observed data.

The system state is also expressed by a mass/belief function, which can be considered as the current state of the system. The reason for introducing the system state is that, in general, beliefs of rules in Knowledge Bases are not allowed to be changed, i.e. non-monotonic reasoning is not permitted. On the other hand, the uncertainty of knowledge of controlled systems requires modification of the confidence strengths or beliefs. This transitory knowledge, as opposed to the relatively "timeless" knowledge in the Knowledge Base is knowledge of the controlled system's state.

A new system state is produced by updating the previous state using the combined result of observed data. The system state triggers rules in the Knowl-

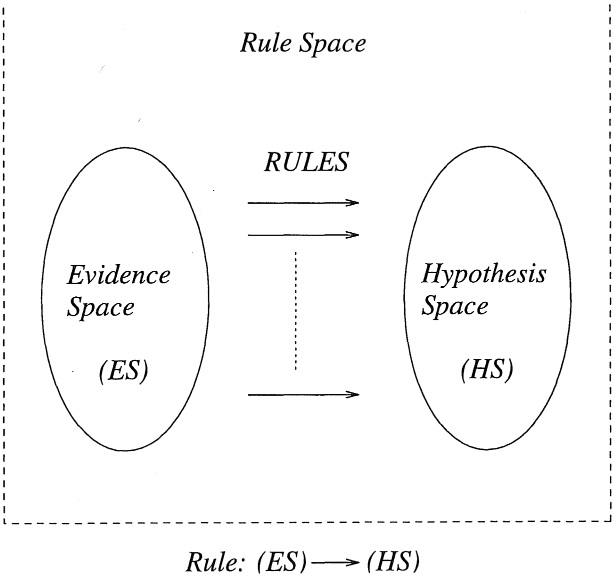

Figure 2. Rule space.

edge Base, and causes the system to make decisions on actions to control the system[5].

As illustrated in the last section, when bel_c and bel_s become probability measures, Belief Logic will become a probabilistic logic, see Nilsson (1984), Bundy (1985). This indicates a way of dealing with observed data in probability theory.

3.3. Representing uncertain knowledge

When applying rule-based methods to represent knowledge in knowledge-based systems, we form two spaces: one is the evidence space which can be considered as a set of premises; and the other is the hypothesis space. Then a rule is a mapping from the evidence space to the hypothesis space. Evidence associated with rules is used to represent "rule strength" as in Figure 2.

As indicated in Guan, Bell, Lesser (1990) it is difficult to apply the Dempster-Shafer theory of evidence directly to rule-based systems since the application of Dempster's rule of combination to combine beliefs of two spaces may result in inconsistencies. Yen (1989) introduces a probabilistic mapping between the evidence space and hypothesis space and uses conditional probabilities to represent uncertain relationships between evidence and hypotheses. Conditional probability can only deal with a simple case, i.e. where one event occurs. How-

[5]Of course, in an ideal world, rules in the Knowledge Base could be updated as well, and we are interested in this problem (e.g. in data mining, Piatetsky-Shapiro, Frawley 1991), but we exclude that possibility here, for clarity of the main ideas of the CU controller.

ever, Jeffrey's generalization – Jeffrey (1983) – of conditional probability can be applied to arbitrarily many events. Another problem is that the mixture of two theories is not convenient since we have two interpretations to evidence and hypothesis spaces. Guan, Bell and Lesser (1990) generalize Yen's method from probability theory to the Dempster-Shafer theory by giving an evidential mapping that uses mass functions to express these uncertain relationships.

We follow these ideas and use the proportional sum in Dempster-Shafer theory of evidence , which corresponds to Jeffrey's generalization of conditional probability, to represent uncertain relationships between evidence and hypotheses. In our method a rule has the following form:

$$\textbf{If } e_1 \textbf{ then } h_{11}, ..., h_{1k_1} \textbf{ with } m_1 \textbf{ else}$$

$$\dots$$

$$\textbf{If } e_n \textbf{ then } h_{n1}, ..., h_{nk_m} \textbf{ with } m_n$$

where e_1, ...,e_n are pieces of evidence which form an evidence space and for each i, h_{i1}, ..., h_{ik_i} are hypotheses represented by subsets of a frame of discernment Θ, and m_i is a mass function on h_{i1}, ..., h_{ik_i}. A Knowledge Base is a set of such rules that have the same frame of discernment.

The following proposition gives a method for transforming beliefs using rules.

PROPOSITION 3.1 *Given an uncertain rule as above and a mass function m on e_1, ...,e_n, the function $m' : e_1, ..., e_n \longrightarrow [0, 1]$*

$$m'(h_j) = \sum_{i=1}^{n} m(e_i) \cdot m_i(h_j)$$

for $j = 1, ..., n$, is a mass function.

3.4. Inference mechanism

A system state represented as a mass/belief function is transformed into a mass/belief function by use of above propositions. Actions for controlling a system are associated with propositions which are represented as subsets of a frame of discernment. In Example 1, the following propositions

(1) *Flight is in the same direction as the wind*

(2) *Wind is strong*

represented as subsets

$$\{same - direction\} \text{ and } \{strong - wind\}$$

respectively, are associated with actions

(1) *Reduce flying speed*, and

(2) *Change flying direction.*

Then the proposition "*flight is in the same direction as the wind*" or "*wind is strong*" is represented by subset $\{\{same-direction\}, \{strong-wind\}\}$ which is associated with actions *Reduce flying speed* or *Change flying direction.*

Decision making is performed in Belief Logic. Since Belief Logic is truth functional with respect to truth assignment, only a set of propositions (called axioms) are assigned to subsets of the frame of discernment. Propositions which are the most believable, i.e. those where the corresponding subsets which have the largest associated beliefs, are chosen and the associated actions are performed to control the system.

4. Discussion

In this paper we have proposed a new method, called the *CU* method, for control systems based upon the Dempster-Shafer theory of evidence and our previous work on dealing with uncertainty. The motivation for this work stems from the inconvenient fact that the behaviour of a system to be controlled, usually measured by statistics or probabilities and the observed data, has to be transformed into fuzzy numbers by a fuzzifier, and control commands which are fuzzy numbers have to be mapped to control actions by a defuzzifier. The fact that knowledge in a control system is uncertain suggests the hypothesis for our study that other approaches for handling uncertainty based on probability theory can be candidates in complex system control.

There are some advantages of applying the Dempster-Shafer theory of evidence in system control. First, the Dempster-Shafer theory of evidence is based on probability theory and, therefore, the behaviour of a controlled system can be represented directly. This overcomes the fuzzifying/defuzzifying inefficiency noted above. Second, the Dempster-Shafer theory of evidence permits an explicit representation of ignorance and does not assume that degrees of beliefs in a set of propositions are uniformly allocated to the individual propositions. This allows a sensor to be concentrated by the corresponding Belief Builder on to some interesting states. Furthermore, the Decision Logic, which is based on Belief Logic, retains the main advantage of Incidence Calculus — that logical connectives are truth functional with respect to truth assignment. The method of finding a legal assignment in Incidence Calculus can then be used in Belief Logic.

Further work is required on the efficiency of the algorithms of combining and updating beliefs. In general, the computation of Dempster's rule of combination and the proportional sum on arbitrary structures of a frame of discernment is exponentially complex. Care must be taken when using these operations in general, because they are not commutative. One way to reduce the complexity is to find a simpler structure to represent evidence which still retains applicabability in practice. For example, in some applications, a simple hierarchy, Gordon, Shortliff (1985), which has been used to represent knowledge in Knowledge Based systems, can be adopted to represent uncertainty. Algorithms of combining and updating beliefs, Guan, Bell (1992), Shi, Hull, Bell (1994), which have linear complexities have been proposed. Parallelism is another possible route to efficiency. The techniques of applying the approach proposed in this paper to practical system control should be also investigated.

References

ABEL, S. (1988) The sum-and lattice-points method based on an evidential-reasoning system applied to the vehicle guidance problem. In *Uncertainty in Artificial Intelligence 2*, 365–370.

BELL, D.A., SHI, SHENGLI and HULL, M.E.C. (1994) Applying evidential theory to control systems. In *Proceedings of 1994 International Fuzzy Systems and Intelligence Control Conference, Louisville, U.S.A.*.

BELL, D.A., SHI, SHENGLI and HULL, M.E.C. (1995) Representing ignorance in Belief Logic. In *Proceedings of the 4th Golden West International Conference On Intelligent Systems, June, 1995, San Francisco. U.S.A.*

BUNDY, A. (1985) Incidence calculus: a mechanism for probability reasoning. *J. of Automated Reasoning*, **3**, 1, 263–283.

BUNDY, A. (1986) Correctness criteria of some algorithms for uncertain reasoning using incidence calculus. *J. of Automated Reasoning*, **2**, 2, 109–126.

GOODMAN, I.R. and NGUYEN, H.T. (1985) *Uncertainty Models for Knowledge-Based Systems.* North-Holland Press, Amsterdam.

GORDON, J. and SHORTLIFF, E.H. (1985) A method for managing evidential reasoning in a hierarchical hypothesis space. *Artificial Intelligence*, **26**, 323 - 357.

GUAN, J.W. and BELL, D.A. (1991) *Evidence Theory and Its Applications.* Volume 1. North-Holland.

GUAN, J.W. and BELL, D.A. (1992) *Evidence Theory and Its Applications.* Volume 2. North-Holland.

GUAN, J.W., BELL, D.A. and LESSER, V.R. (1990) Evidential reasoning and rule strengths in expert systems. In *Proceedings of the 3rd Irish Conference on Artificial Intelligence and Cognitive Science*, Belfast.

HALPERN, J.Y. and FAGIN, R. (1992) Two views of belief: belief as generalized probability and belief as evidence. *Artificial Intelligence*, **54**, 275–314.

INAGAKI, T. (1991) Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory. *IEEE Transactions on Reliability*, **2**, 40, 182–188.

JEFFREY, R.C. (1983) *The Logic of Decision.* University of Chicago Press.

KAK, A.C. et al. (1990) Hierarchical evidence accumulation in the PSEIKI system. In M. Henrion et al., eds., *Uncertainty in Artificial Intelligence 5*. Elsevier Science Publishers, North-Holland.

LEE, R.T.C. (1990) Fuzzy logic and the resolution principle. *JACM*, **1**, 19, 109 - 119.

MURPHY, R.R. (1991) The application of Dempster-Shafer theory to a novel control scheme for sensor fusion. In S. S. Chen, eds., *Stochastic and Neural Methods in Signal Processing, Image Processing, and Computer Vision*, volume 1569, chapter 47, pages 55–68.

NILSSON, N. (1984) Probabilistic logic. Technical report, SRI International, Menlo Park, Calif.

PEARL, J. (1988) *Probabilistic reasoning in intelligent systems, networks of plausible inference.* San Mateo, Calif., Morgan Kaufmann Publishers.

PIATETSKY-SHAPIRO, G. and FRAWLEY, W., eds. (1991) *Knowledge Discovery in Databases.* Cambridge, MA:AAAI/MIT Press.

SAFFIOTTI, A. and UMKEHRER, E. (1991) PULCINELLA: A general tool for propagating uncertainty in valuation networks. In B. D'Ambrosio, P. Smets, and P. Bonissone, eds., *Proc. 7th Conference on Uncertainty in AI*, 323–331. Morgan Kaufmann.

SHAFER, G. (1976) *A Mathematical Theory of Evidence.* Princeton University Press.

SHI, SHENGLI, HULL, M.E.C. and BELL, D.A. (1994) Combination-update methodology for handling uncertainty. In *Second European Congress on Fuzzy and Intelligent Technologies*, 1736–1740.

SHI, SHENGLI, HULL, M.E.C. and BELL, D.A. (1994) CU – A new approach to control systems. In *3rd International Conference on Fuzzy Logic Neural Nets and Soft Computing (IIZUKA'94), Iizuka, Japan*, 587–588.

SHI, SHENGLI, HULL, M.E.C. and BELL, D.A. (1994) A new rule for updating evidence. In *8th International Symposium on Methodologies for Intelligent Systems (ISMIS'94)*, 95–104.

SHI, SHENGLI, HULL, M.E.C. and BELL, D.A. (1994) Updating belief in Gordon-Shortliffe's hierarchical structure. In *IEEE 2nd Australian and New Zealand Conference on Intelligence Information System, Australia.*

YAMAKAWA, T. and HIROTA, K., eds. (1989) Special issue on applications of fuzzy logic control to industry. *Fuzzy Sets and System*, **2**, 32.

YEN, J. (1989) GERTIS: A Dempster-Shafer approach to diagnosing hierarchical hypotheses. *Communications of the ACM*, **5**, 32, 573–585.

ZADEH, L.A. (1965) Fuzzy sets. *Inform. Contr.*, **8**, 338–353.