Context-dependent Combination of Sensor Information in Dempster-Shafer Theory for BDI*

Sarah Calderwood · Kevin McAreavey ·
Weiru Liu · Jun Hong

Abstract There has been much interest in the Belief-Desire-Intention (BDI) agent-based model for developing scalable intelligent systems, e.g. using the AgentSpeak framework. However, reasoning from sensor information in these large-scale systems remains a significant challenge. For example, agents may be faced with information from heterogeneous sources which is uncertain and incomplete, while the sources themselves may be unreliable or conflicting. In order to derive meaningful conclusions, it is important that such information be correctly modelled and combined. In this paper, we choose to model uncertain sensor information in Dempster-Shafer (DS) theory. Unfortunately, as in other uncertainty theories, simple combination strategies in DS theory are often too restrictive (losing valuable information) or too permissive (resulting in ignorance). For this reason, we investigate how a context-dependent strategy originally defined for possibility theory can be adapted to DS theory. In particular, we use the notion of largely partially maximal consistent subsets (LPMCSes) to characterise the context for when to use Dempster's original rule of combination and for when to resort to an alternative. To guide this process, we identify existing measures of similarity and conflict for finding LPMCSes along with quality of information heuristics to ensure that LPMCSes are formed around high quality information. We then propose an *intelligent sensor* model for integrating this information into the AgentSpeak framework which is responsible for applying evidence propagation to construct compatible information, for performing context-dependent combination and for deriving beliefs for revising an agent's belief base. Finally, we present a power grid scenario inspired by a real-world case study to demonstrate our work.

---

* This is a significantly revised and extended version of [8].

S. Calderwood, K. McAreavey, W. Liu and J. Hong
School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, United Kingdom
E-mail: {scalderwood02, kevin.mcareavey, w.liu, j.hong}@qub.ac.uk

1 Introduction and Related Work

Supervisory control and data acquisition (SCADA) systems [5,10] have been applied in a wide variety of industries including power [1,26,27], manufacturing [38] and water treatment [34]. In these settings, SCADA systems deploy large numbers of sensors to collect information for use by control mechanisms, security infrastructure, etc. In the electric power industry, for example, information from sensors may be used to identify faults, preempt outages and manage participation in energy trading markets [39]. However, given this abundance of information from heterogeneous and imperfect sensors, an important challenge is how to accurately model and combine (merge) this information to ensure well-informed decision making. While this already applies to traditional SCADA systems, it is of particular importance to intelligent and scalable SCADA systems which are gaining increased interest, e.g. those based on autonomous multi-agent frameworks [26,27]. Therefore, it is also necessary to understand how this information can be used by these more advanced SCADA models.

Within these large systems, sensors represent independent sources of information. For example, in electric power systems, sensors might independently gather information about features of the environment such as voltage, amperage, frequency, etc. However, the information obtained by these sensors may be uncertain or incomplete (e.g. due to noisy measurements) while the sensors themselves may be unreliable (e.g. due to malfunctions, inherent design limitations, etc). Moreover, information from different sensors may be conflicting and so we need some way to combine this information to find a representative and consistent model of the underlying sources. This combined information can then be used for higher-level decision making. In this paper, we apply the Dempster-Shafer (DS) theory of evidence [31] to model imperfect sensor information since it can be used to model common types of sensor uncertainty (e.g. ignorance and imprecise information) and, more importantly, has mature methods for combining information.

Using DS theory, information from a sensor can be modelled as a mass function. The original and most common method of combining mass functions is using Dempster's rule of combination [31], which assumes that the mass functions have been obtained from reliable and independent sources. While sensors may be unreliable, there exist methods in DS theory which allow us to first discount unreliable information [31] such that we can then treat this discounted information as fully reliable [25]. Given that our mass functions are independent and can be treated as fully reliable, it would seem that this is a sufficient solution. However, the main problem arises when combining conflicting information. A classical example, introduced by Zadeh [37], illustrates the effect of a high degree of conflict when applying Dempster's rule. Suppose patient P_1 is independently examined by two doctors A and B who have the same level of expertise and are equally reliable. A's diagnosis is that P_1 has meningitis with probability 0.99 or a brain tumour with probability 0.01. B agrees with A that the probability of a brain tumour is 0.01 but B believes that P_1 has a concussion with probability 0.99. By applying Dempster's rule, we would conclude that P_1 has a brain tumour with probability 1. Clearly, this result is counter-intuitive since it implies complete support for a diagnosis that both A and B considered highly improbable.

Since Zadeh first introduced this counter-example, there has been extensive work on handling conflicts between mass functions. This has resulted in a broad

range of approaches which fundamentally differ in how they deal with conflict. We might mention, for example, Smet's Transferable Belief Model [33] which extends DS theory with an open-world assumption by allowing a non-zero mass to be assigned to the empty set (representing "none of the above"). On the other hand, a popular alternative within standard DS theory—Dubois & Prade's disjunctive consensus rule [13] being a notable example—is to incorporate all conflict during combination. For example, by applying Dubois & Prade's rule in Zadeh's example, we would intuitively conclude that P_1 has a brain tumour with probability 0.0001. Unfortunately, both approaches tend to increase ignorance when faced with conflict, which is problematic since our goal is to improve support for decision-making. For example, another patient P_2 is examined by doctors A and B. A's diagnosis is that P_2 has meningitis with probability 0.99 or either meningitis, a brain tumour or concussion with probability 0.01. B's diagnosis is that P_2 has either meningitis, a brain tumour or concussion with probability 1. By applying Dubois & Prade's rule, we would only conclude that P_2 has meningitis, a brain tumour or concussion with probability 1. Clearly, this result is counter-intuitive since it implies complete ignorance about the three possibilities, even though A believes it is highly probable that P_2 has meningitis. Conversely, by applying Dempster's rule, we would intuitively conclude that P_2 has meningitis with probability 0.99.

While there are those who maintain that Dempster's rule is the only method of combination that should ever be used [16], we argue that no rule is suitable for every situation. In fact, comparable problems have been identified in other uncertainty theories for which adaptive combination strategies have been proposed [12, 14, 15, 17]. In the literature, the majority of the adaptive combination strategies have been proposed in the setting of possibility theory. In general, these approaches focus on deciding when to use a *conjunctive rule* and when to use a *disjunctive rule*, where the former is applied to information from reliable and consistent sources aiming to reinforce commonly agreed opinions from multiple sources while the latter is applied to information from conflicting sources aiming to not dismiss any opinions from any sources. An early proposal was Dubois & Prade's adaptive combination rule [12] which combines two sources by applying both a conjunctive and disjunctive rule while taking into account the agreement between sources. However, this rule is not associative meaning that it does not reliably extend to combining information from a set of sources. For this reason, there has been more interest in subset-based context-dependent combination strategies. An early example of this approach was proposed in [15] where the combination rule assumes there are n sources from which an unknown subset of j sources are reliable. In this case, a conjunctive rule is applied to all subsets with cardinality j where the cardinality j represents the context for using a conjunctive rule. The conjunctively combined information is then combined using a disjunctive rule. This work was further developed in [14] by proposing a method for deciding on the value of j. Specifically, j was chosen as the cardinality of the largest subset for which the application of a conjunctive rule would result in a normalized possibility distribution. This approach more accurately reflects the specified conditions (i.e. the context) for applying a conjunctive rule in possibility theory than previous adaptive strategies. However, since all the subsets with cardinality j are combined using a conjunctive rule, most of these subsets may contain conflicting information. Thus, the approach does not completely reflect the context for applying a conjunctive rule. Moreover, as highlighted in [17], this approach is not appropriate

for dealing with disjoint consistent subsets of different sizes. To address this issue, Hunter & Liu proposed another context-dependent combination strategy in [17] based on finding a partition of sources, rather than a set of fixed-size subsets. They refer to this partition as the set of *largely partially maximal consistent subsets* (LPMCSes) and, again, each LPMCS represents the context for applying a conjunctive rule. However, rather than requiring that each combined LPMCS be a normalized possibility distribution as in [14], they instead define a relaxation of this rule in order to find subsets which are "mostly" in agreement.

With regards to DS theory, sensor fusion is actively used in applications such as robotics [28] and engine diagnostics [2]. Specifically, in [28], the authors propose an approach to derive mass functions from information obtained from sensor observations and domain knowledge by propagating the evidence into an appropriate frame of features. Once mass functions are derived they are then combined using a single combination rule, e.g. Dempster's rule. A robot will then obtain a belief where they can proceed or terminate a task. Alternatively, in [2], an engine diagnostic problem requires a frame of discernment to represent fault types, mass functions and Dempster's rule for combination. The authors propose methods of calculating mass functions as well as rational diagnosis decision making rules and entropy of evidence to improve and evaluate the performance of the combination. They demonstrate through a scenario that decision conflicts can be resolved and the accuracy of fault diagnosis can be improved by combining information from multiple sensors.

Within DS theory, the most relevant work is that proposed by Browne et al. in [7]. In this paper, the authors do not actually propose an adaptive combination strategy. Instead they use the variant of DS theory known as Dezert-Smarandache (DSm) theory to address some of the limitations with standard DS theory in terms of handling conflict and handling non-independent hypotheses. In particular, they propose a notion of maximal consistent subsets which they use to characterise how mass functions should be discounted prior to combination in DSm theory, given a number of possible discounting techniques. Once mass functions have been discounted, they then use the Proportional Conflict Redistribution Rule no. 5 (PCR5) from DSm theory to combine evidence. As far as we are aware, there have been no other approaches to context-dependent combination in DS theory.

To date, the work in [17] is arguably the most accurate reflection of the context for applying a conjunctive rule. For this reason, we propose an approach to combine mass functions based on the context-dependent strategy for combining information in possibility theory. In this way, we aim to characterise the *context* for when to use Dempster's rule and for when to resort to an alternative (in this paper we focus on Dubois & Prade's rule). In particular, when combining a set of mass functions, we first identify a partition of this set using a measure of conflict in DS theory. This accurately reflects the accepted context for Dempster's rule since, even if a set of mass functions is highly conflicting, it is often possible to find subsets with a low degree of conflict. Likewise, in [17], the authors use the most appropriate concepts in possibility theory to determine the context for using the combination rule. Each element in this partition is called a LPMCS and identifies a subset which should be combined using Dempster's rule. Once each LPMCS has been combined using Dempster's rule, we then resort to Dubois & Prade's rule to combine the information which is highly conflicting. Moreover, we utilise heuristics on the quality and similarity of mass functions to ensure that each

LPMCS is formed around higher quality information. Similarly, in [17], the authors use the most appropriate measures in possibility theory for measuring the quality of information. When compared to traditional approaches, our context-dependent strategy allows us to derive more useful information for decision making while still being able to handle a high degree of conflict between sources.

Turning again to the issue of SCADA systems, we consider a particular multi-agent formulation [35, 36] of SCADA based on the Belief-Desire-Intention (BDI) framework [6]. Recently, there has been interest in using the BDI framework for this purpose due to, among other reasons, its scalability [18]. In this paper, we focus on the AgentSpeak [29] language (for which there are a number of existing implementations [4]) which encodes a BDI agent by a set of predefined plans used to respond to new event-goals. The applicability of these plans is determined by preconditions which are evaluated against the agent's current beliefs. As such, our new context-dependent combination rule ensures that the agent can reason about a large number of sensors which, in turn, ensures that they are well-informed about the current state of the world during plan selection. Importantly, we describe how combined uncertain sensor information can be integrated into this agent model to improve decision making. In addition, we devise a mechanism to identify when information should be combined which ensures the agent's beliefs are up-to-date while minimising the computational cost associated with combining information.

In summary, the main contributions of this work are as follows:

(i) We adapt a context-dependent combination rule from possibility theory to DS theory. In particular, we identify suitable measures from DS theory for determining quality and contextual information used to select subsets for which Dempster's rule should be applied.

(ii) We propose to handle uncertain sensor information in the BDI framework using an *intelligent sensor*. This is responsible for constructing compatible mass functions using evidence propagation, performing context-dependent combination and deriving beliefs for revising an agent's belief base.

(iii) Finally, we present a power grid scenario which demonstrates our framework. In particular, we describe the whole process from modelling uncertain sensor information as mass functions to deriving suitable beliefs for an AgentSpeak agent's belief base along with the effect this has on decision making.

The remainder of the paper is organised as follows. In Section 2, we introduce preliminaries on DS theory. In Section 3, we describe our new context-dependent combination rule for mass functions using LPMCSes. In Section 4, we demonstrate how to handle uncertain combined sensor information in a BDI agent. In Section 5, we present a scenario based on a power grid SCADA system modelled in the BDI framework to illustrate our approach. Finally, in Section 6, we draw our conclusions.

2 Preliminaries

DS theory [31] is well-suited for dealing with epistemic uncertainty and sensor data fusion.

Definition 1 Let Ω be a set of exhaustive and mutually exclusive hypotheses, called a *frame of discernment*. Then a function $\mathsf{m} : 2^{\Omega} \to [0,1]$ is called a *mass*

function[1] over Ω if $\mathsf{m}(\emptyset) = 0$ and $\sum_{A \subseteq \Omega} \mathsf{m}(A) = 1$. Also, a *belief* function and a *plausibility* function from m, denoted Bel and Pl, are defined for each $A \subseteq \Omega$ as:

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

$$\mathsf{Pl}(A) = \sum_{A \cap B \neq \emptyset} \mathsf{m}(B).$$

Any $A \subseteq \Omega$ such that $\mathsf{m}(A) > 0$ is called a *focal element* of m. Intuitively, $\mathsf{m}(A)$ is the proportion of evidence that supports A, but none of its strict subsets. Similarly, $\mathsf{Bel}(A)$ is the degree of evidence that the true hypothesis belongs to A and $\mathsf{Pl}(A)$ is the maximum degree of evidence supporting A. The values $\mathsf{Bel}(A)$ and $\mathsf{Pl}(A)$ represent the lower and upper bounds of belief, respectively. To reflect the reliability of evidence we can apply a discounting factor to a mass function using Shafer's discounting technique [31] as follows:

Definition 2 Let m be a mass function over Ω and $\alpha \in [0, 1]$ be a discount factor. Then a *discounted* mass function with respect to α, denoted m^α, is defined for each $A \subseteq \Omega$ as:

$$\mathsf{m}^\alpha(A) = \begin{cases} (1 - \alpha) \cdot \mathsf{m}(A), & \text{if } A \subset \Omega, \\ \alpha + (1 - \alpha) \cdot \mathsf{m}(A), & \text{if } A = \Omega. \end{cases}$$

The effect of discounting is to remove mass assigned to focal elements and to then assign this mass to the frame. When $\alpha = 0$ the source is completely reliable and when $\alpha = 1$ the source is completely unreliable. Once a mass function has been discounted, it is then treated as fully reliable.

Definition 3 Let M be a set of mass functions over Ω. Then a function $\mathsf{K} : M \times M \to [0, 1]$ is called a *conflict* measure and is defined for $\mathsf{m}_i, \mathsf{m}_j \in M$ as:

$$\mathsf{K}(\mathsf{m}_i, \mathsf{m}_j) = \sum_{B \cap C = \emptyset} \mathsf{m}_i(B)\mathsf{m}_j(C).$$

Understandably it has been argued that K is not a good measure of conflict between mass functions [22]. For example, given a mass function m, then it is not necessarily the case that $\mathsf{K}(\mathsf{m}, \mathsf{m}) = 0$. Indeed, if m has any non-nested focal elements then it will always be the case that $\mathsf{K}(\mathsf{m}, \mathsf{m}) > 0$. Conversely, mass functions derived from evidence that is more dissimilar (e.g. by some measure of distance) does not guarantee a higher degree of conflict. However, the reason we consider this measure of conflict rather than more intuitive measures from the literature, is that this measure is fundamental to Dempster's rule [31]:

Definition 4 Let m_i and m_j be mass functions over Ω from independent and reliable sources. Then the combined mass function using *Dempster's rule of combination*, denoted $\mathsf{m}_i \oplus \mathsf{m}_j$, is defined for each $A \subseteq \Omega$ as:

$$(\mathsf{m}_i \oplus \mathsf{m}_j)(A) = \begin{cases} c \sum_{B \cap C = A} \mathsf{m}_i(B)\mathsf{m}_j(C), & \text{if } A \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases}$$

where $c = \frac{1}{1 - \mathsf{K}(\mathsf{m}_i, \mathsf{m}_j)}$ is a normalisation constant.

---

[1] A mass function is also known as a basic belief assignment (BBA) and $\mathsf{m}(\emptyset) = 0$ is not required.

The effect of c is to redistribute the mass value that would otherwise be assigned to the empty set. As mentioned in Section 1, Dempster's rule is only well-suited to combine evidence with a low degree of conflict [30]. Importantly, when referring to the "degree of conflict", we are specifically referring to the $\mathsf{K}(\mathsf{m}_i, \mathsf{m}_j)$ value. It is this measure that we will use to determine the context for using Dempster's rule. In [13], an alternative to Dempster's rule was proposed as follows:

Definition 5 Let m_i and m_j be mass functions over Ω from independent sources. Then the combined mass function using *Dubois & Prade's (disjunctive consensus) rule*, denoted $\mathsf{m}_i \otimes \mathsf{m}_j$, is defined for each $A \subseteq \Omega$ as:

$$(\mathsf{m}_i \otimes \mathsf{m}_j)(A) = \sum_{B \cup C = A} \mathsf{m}_i(B)\mathsf{m}_j(C).$$

Dubois & Prade's rule incorporates all conflict and does not require normalisation. As such, this rule is suitable for combining evidence with a high degree of conflict [30]. Importantly, each combination rule $\odot \in \{\oplus, \otimes\}$ satisfies the following mathematical properties: $\mathsf{m}_i \odot \mathsf{m}_j = \mathsf{m}_j \odot \mathsf{m}_i$ (*commutativity*); and $\mathsf{m}_i \odot (\mathsf{m}_j \odot \mathsf{m}_k) = (\mathsf{m}_i \odot \mathsf{m}_j) \odot \mathsf{m}_k$ (*associativity*). This means that combining a set of mass functions using either rule, in any order, will produce the same result. For this reason, given a set of mass functions $M = \{\mathsf{m}_1, \ldots, \mathsf{m}_n\}$ over Ω, we will use:

$$\bigodot_{\mathsf{m} \in M} \mathsf{m}$$

to denote the combined mass function $\mathsf{m}_1 \odot \ldots \odot \mathsf{m}_n$. It is worth noting however, that these rules do not satisfy $\mathsf{m} \odot \mathsf{m} = \mathsf{m}$ (*idempotency*).

The ultimate goal in representing and reasoning about uncertain information is to draw meaningful conclusions for decision making. To translate mass functions to probabilities a number of existing transformation models can be applied such as those in [9,11,32]. One of the most widely used models is Smet's pignistic model [32] which allows decisions to be made from a mass function on individual hypotheses as follows (note that we use $|A|$ to denote the cardinality of a set A):

Definition 6 Let m be a mass function over Ω. Then a function $\mathsf{BetP}_\mathsf{m} : \Omega \to [0, 1]$ is called the *pignistic* probability distribution over Ω with respect to m and is defined for each $\omega \in \Omega$ as:

$$\mathsf{BetP}_\mathsf{m}(\omega) = \sum_{A \subseteq \Omega, \omega \in A} \frac{\mathsf{m}(A)}{|A|}.$$

3 Context-dependent Combination

Various proposals have been suggested in the literature for adaptive combination rules [13,14,24]. In this section, we explore ideas presented in [17] for the setting of possibility theory and adapt these to DS theory. In particular, we find suitable measures to replace those used in their work for determining contextual information, including a conflict measure (i.e. the K measure from Dempster's rule), quality measures (i.e. measures of *non-specificity* and *strife*) and a similarity measure (i.e. a *distance* measure). However, before we explain how to adapt their

algorithm to DS theory, we provide the following intuitive summary. First, we identify the highest quality mass function using quality of information heuristics[2]. This mass function is used as an initial reference point. Second, we find the most similar mass function to this reference. Third, we use Dempster's rule to combine the reference mass function with the most similar mass function. Fourth, we repeat the second and third steps—using the combined mass function as a new reference point—until the conflict with the next most similar mass function exceeds some specified threshold. The set of mass functions prior to exceeding this threshold constitutes an LPMCS. Finally, we repeat from the first step (on the remaining mass functions) until all mass functions have been added to an LPMCS. Each LPMCS represents a subset of mass functions with similar opinions and a "low" degree of conflict and thus the context for using Dempster's rule. On the other hand, the set of combined LPMCSes represents highly conflicting information and thus the context for resorting to Dubois & Prade's rule.

3.1 Preferring High Quality Information

The most common methods for characterising the quality of information in DS theory is using measures of non-specificity (related to ambiguity) and strife (i.e. internal conflict). One such measure of non-specificity was introduced in [21] as follows:

Definition 7 Let M be a set of mass functions over Ω. Then a function $\mathsf{N} : M \to \mathbb{R}$ is called a *non-specificity* measure and is defined for each $\mathsf{m} \in M$ as:

$$\mathsf{N}(\mathsf{m}) = \sum_{A \subseteq \Omega \text{ s.t. } A \neq \emptyset} \mathsf{m}(A) \log_2 |A|.$$

A mass function m is completely specific when $\mathsf{N}(\mathsf{m}) = 0$ and is completely non-specific when $\mathsf{N}(\mathsf{m}) = \log_2 |\Omega|$. However, since a frame is a set of mutually exclusive hypotheses, disjoint focal elements imply some disagreement in the information modelled by a mass function. It is this type of disagreement which is characterized by a measure of strife. One such measure was introduced in [21] as follows:

Definition 8 Let M be a set of mass functions over Ω. Then a function $\mathsf{S} : M \to \mathbb{R}$ is called a *strife* measure and is defined for each $\mathsf{m} \in M$ as:

$$\mathsf{S}(\mathsf{m}) = -\left(\sum_{A \subseteq \Omega \text{ s.t. } A \neq \emptyset} \mathsf{m}(A) \log_2 \sum_{B \subseteq \Omega \text{ s.t. } B \neq \emptyset} \mathsf{m}(B) \frac{|A \cap B|}{|A|} \right).$$

A mass function m has no internal conflict when $\mathsf{S}(\mathsf{m}) = 0$ and has complete internal conflict when $\mathsf{S}(\mathsf{m}) = \log_2 |\Omega|$. For example, given a mass function m over $\Omega = \{a, b\}$ such that $\mathsf{m}(\{a\}) = \mathsf{m}(\{b\}) = 0.5$, then m has the maximum strife value $\mathsf{S}(\mathsf{m}) = 1 = \log_2 |\Omega|$. Using these measures for non-specificity and strife allows us to assess the quality of information and, thus, to rank mass functions based on their quality. By adapting Definition 5 from [17], we can define a quality ordering over a set of mass functions as follows:

---

[2] We assume a strict preference ordering over sources when the selection is not unique.

Definition 9 Let M be a set of mass functions over Ω. Then a total order over M, denoted $\succeq_Q$, is called a *quality order* and is defined for $\mathsf{m}_i, \mathsf{m}_j \in M$ as:

$$\mathsf{m}_i \succ_Q \mathsf{m}_j \Leftrightarrow (\mathsf{N}(\mathsf{m}_i) < \mathsf{N}(\mathsf{m}_j)) \vee (\mathsf{N}(\mathsf{m}_i) = \mathsf{N}(\mathsf{m}_j) \wedge \mathsf{S}(\mathsf{m}_i) < \mathsf{S}(\mathsf{m}_j)),$$
$$\mathsf{m}_i \sim_Q \mathsf{m}_j \Leftrightarrow \mathsf{N}(\mathsf{m}_i) = \mathsf{N}(\mathsf{m}_j) \wedge \mathsf{S}(\mathsf{m}_i) = \mathsf{S}(\mathsf{m}_j).$$

Moreover, $\mathsf{m}_i \succeq_Q \mathsf{m}_j$ if $\mathsf{m}_i \succ_Q \mathsf{m}_j$ or $\mathsf{m}_i \sim_Q \mathsf{m}_j$. Finally, m_i is said to be of *higher quality* than m_j if $\mathsf{m}_i \succ_Q \mathsf{m}_j$.

Intuitively, a mass function is of higher quality when it is more specific or when it is equally specific but less internally conflicting. Thus, by definition, a lower non-specificity value is preferred over a lower strife value. Obviously this quality ordering is not a strict order and so does not guarantee a unique highest quality source. We denote the set of *highest quality* mass functions in M as $\max(M, \succeq_Q) = \{\mathsf{m} \in M \mid \nexists \mathsf{m}' \in M, \mathsf{m}' \succ_Q \mathsf{m}\}$. To handle the situation where there are multiple highest quality mass functions, we make the assumption of a prior strict preference ordering over sources, denoted $\succ_P$. Given a set of mass functions M over Ω, then for all $\mathsf{m}_i, \mathsf{m}_j \in M$ we say that m_i is *more preferred* than m_j if $\mathsf{m}_i \succ_P \mathsf{m}_j$. This preference ordering may be based on the reliability of sources, the significance of sources, the order in which sources have been queried, etc. We denote the (unique) *most preferred* mass function in M as $\max(M, \succ_P) = \mathsf{m}$ such that $\mathsf{m} \in M$ and $\nexists \mathsf{m}' \in M, \mathsf{m}' \succ_P \mathsf{m}$.

Definition 10 Let M be a set of mass functions over Ω. Then the mass function $\mathsf{Ref}(M) = \max(\max(M, \succeq_Q), \succ_P)$ is called the *reference* mass function in M.

m_i	$\mathsf{m}_i(\{a\})$	$\mathsf{m}_i(\{b,c\})$	$\mathsf{N}(\mathsf{m}_i)$	$\mathsf{S}(\mathsf{m}_i)$
m_1	0.85	0.15	0.15	0.61
m_2	0.7	0.3	0.3	0.881
m_3	0.55	0.45	0.45	0.993

Table 1: Non-specificity and strife values for a set of mass functions $M = \{\mathsf{m}_1, \mathsf{m}_2, \mathsf{m}_3\}$ over $\Omega = \{a, b, c\}$.

Example 1 Given the set of mass functions M as well as the non-specificity and strife values for M as shown in Table 1, then the quality ordering over M is $\mathsf{m}_1 \succ_Q \mathsf{m}_2 \succ_Q \mathsf{m}_3$. Thus, regardless of the preference ordering over M, the reference mass function in M is $\mathsf{Ref}(M) = \mathsf{m}_1$.

A reference mass function is a mass function with the maximum quality of information. When this is not unique, we select the most preferred mass function from the set of mass functions with the maximum quality of information. As we will see in Section 3.2, a reference mass function is a mass function around which an LPMCS is formed. Thus, defining a reference mass function in this way allows us to prefer higher quality information when finding LPMCSes.

3.2 Defining the Context for Combination Rules

While our motivation in finding a reference mass function is to prefer higher quality information, it is possible that high quality mass functions are also highly conflicting. As such, it is not advisable to rely on the quality ordering for choosing the next mass function to consider since this may result in smaller LPMCSes than is necessary. On the other hand, as mentioned in Section 2, the conflict measure K may assign a low degree of conflict to very dissimilar information. Therefore, relying on the conflict measure for choosing the next mass function to consider does not reflect our stated aim of forming LPMCSes around higher quality information. For this reason, we resort to a distance measure since this is a more appropriate measure of similarity. In DS theory, distance measures have been studied extensively in the literature and have seen theoretical and empirical validation [20]. One of the most popular distance measures was proposed by Jousselme et al. [19] as follows:

Definition 11 Let M be a set of mass functions over Ω. Then a function $\mathsf{d} : M \times M \to \mathbb{R}$ is called a *distance* measure and is defined for $\mathsf{m}_i, \mathsf{m}_j \in M$ as:

$$\mathsf{d}(\mathsf{m}_i, \mathsf{m}_j) = \sqrt{\frac{1}{2} \left(\overrightarrow{\mathsf{m}_i} - \overrightarrow{\mathsf{m}_j} \right)^T D \left(\overrightarrow{\mathsf{m}_i} - \overrightarrow{\mathsf{m}_j} \right)},$$

where $\overrightarrow{\mathsf{m}}$ is a vector representation of mass function m, $\overrightarrow{\mathsf{m}}^T$ is the transpose of vector $\overrightarrow{\mathsf{m}}$ and D is a $2^\Omega \times 2^\Omega$ similarity matrix whose elements are $D(A, B) = \frac{|A \cap B|}{|A \cup B|}$ such that $A, B \subseteq \Omega$.

Importantly, for any mass functions m, we have that $\mathsf{d}(\mathsf{m}, \mathsf{m}) = 0$. We can formalise the semantics of this distance measure over a set of mass functions as follows:

Definition 12 Let M be a set of mass functions over Ω and m be a mass function over Ω. Then a total order over M, denoted $\preceq_{\mathsf{d}}^{\mathsf{m}}$, is called a *distance order* with respect to m and is defined for $\mathsf{m}_i, \mathsf{m}_j \in M$ as:

$$\mathsf{m}_i \preceq_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_j \Leftrightarrow \mathsf{d}(\mathsf{m}, \mathsf{m}_i) \leq \mathsf{d}(\mathsf{m}, \mathsf{m}_j).$$

Moreover, $\mathsf{m}_i \sim_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_j$ if $\mathsf{m}_i \preceq_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_j$ and $\mathsf{m}_j \preceq_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_i$. Also, $\mathsf{m}_i \prec_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_j$ if $\mathsf{m}_i \preceq_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_j$ and $\mathsf{m}_j \not\preceq_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_i$. We say that m_i is *closer* to m than m_j is to m if $\mathsf{m}_i \prec_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}_j$.

We denote the set of *closest* mass functions in M to m as $\min(M, \preceq_{\mathsf{d}}^{\mathsf{m}}) = \{\mathsf{m}' \in M \mid \nexists \mathsf{m}'' \in M, \mathsf{m}'' \prec_{\mathsf{d}}^{\mathsf{m}} \mathsf{m}'\}$. Using this distance ordering and the previously defined reference mass function, we can then define a preferred sequence for combining a set of mass functions with Dempster's rule by adapting Definition 9 from [17] as follows:

Definition 13 Let $M = \{\mathsf{m}_1, \ldots, \mathsf{m}_n\}$ be a set of mass functions over Ω such that $n \geq 1$. Then the sequence $(\mathsf{m}_1, \ldots, \mathsf{m}_n)$ is called the *preferred sequence* for combining M if one of the following conditions is satisfied: (i) $n = 1$; or (ii) $n > 1$, $\mathsf{Ref}(M) = \mathsf{m}_1$ and for each $i \in 2, \ldots, n$ we have that:

$$\mathsf{m}_i = \max \left(\min \left(\{\mathsf{m}_i, \ldots, \mathsf{m}_n\}, \preceq_{\mathsf{d}}^{\mathsf{m}_1 \oplus \cdots \oplus \mathsf{m}_{i-1}} \right), \succ_P \right).$$

m_i	$d(m_i, m_1)$	$d(m_i, m_2)$	$d(m_i, m_3)$	$K(m_i, m_1)$	$K(m_i, m_2)$	$K(m_i, m_3)$
m_1	0	0.15	0.3	0.255	0.36	0.465
m_2	0.15	0	0.15	0.36	0.42	0.48
m_3	0.3	0.15	0	0.465	0.48	0.495
$m_1 \oplus m_2$	0.08	0.23	0.38	0.199	0.328	0.457
$m_1 \oplus m_3$	0.024	0.174	0.324	0.238	0.35	0.463
$m_2 \oplus m_3$	0.11	0.04	0.19	0.332	0.404	0.476

Table 2: Distance and conflict values for the set of mass functions M from Table 1.

Example 2 (Continuing Example 1) Given the distance values for M as shown in Table 2, then the preferred sequence for combining M is (m_1, m_2, m_3).

Again we resort to the prior preference ordering over mass functions if there does not exist a unique mass function which is closest to the previously combined mass functions. From this preferred sequence of combination, we can then identify an LPMCS directly. Specifically, given a set of mass functions, then an LPMCS is formed by applying Dempster's rule to the first m elements in the sequence until the conflict between the combined m mass functions and the next mass function exceeds a specified threshold. This threshold reflects the degree of conflict which we are willing to tolerate and is thus domain specific. By adapting Definition 10 from [17] we have the following:

Definition 14 Let M be a set of mass functions over Ω, $(m_1, \ldots, m_n)$ be the preferred sequence for combining M such that $n = |M|$ and $n \geq 1$ and ϵ_K be a conflict threshold. Then a set of mass functions $\{m_1, \ldots, m_m\} \subseteq M$ such that $1 \leq m \leq n$ is called the *largely partially maximal consistent subset* (LPMCS) with respect to ϵ_K, denoted $L(M, \epsilon_K)$, if one of the following conditions is satisfied:

(i) $n = 1$;
(ii) $m = 1$, $n > 1$ and $K(m_1 \oplus \ldots \oplus m_m, m_{m+1}) > \epsilon_K$;
(iii) $m > 1$, $m = n$ and for each $i \in 2, \ldots, m$ then $K(m_1 \oplus \ldots \oplus m_{i-1}, m_i) \leq \epsilon_K$;
(iv) $m > 1$, $m < n$ and for each $i \in 2, \ldots, m$ then $K(m_1 \oplus \ldots \oplus m_{i-1}, m_i) \leq \epsilon_K$ but $K(m_1 \oplus \ldots \oplus m_m, m_{m+1}) > \epsilon_K$.

Example 3 (Continuing Example 2) Given the conflict values for M as shown in Table 2 and a conflict threshold $\epsilon_K = 0.4$, then the LPMCS in M is $L(M, 0.4) = \{m_1, m_2\}$ since $K(m_1, m_2) = 0.36 \leq 0.4$ but $K(m_1 \oplus m_2, m_3) = 0.457 > 0.4$.

In this way, an LPMCS is the set of mass functions corresponding to the first maximal subsequence (from the preferred sequence of combination) where the conflict is deemed (by the conflict threshold) to be sufficiently low enough to apply Dempster's rule. The complete set of LPMCSes can then be recursively defined as follows:

Definition 15 Let M be a set of mass functions over Ω, ϵ_K be a conflict threshold and $L(M, \epsilon_K) \subseteq M$ be the LPMCS in M. Then the set of LPMCSes in M, denoted $L^*(M, \epsilon_K)$, is a partition of M defined as:

$$L^*(M, \epsilon_K) = \begin{cases} \{L(M, \epsilon_K)\} \\ \quad \cup L^*(M \setminus L(M, \epsilon_K), \epsilon_K), & \text{if } M \setminus L(M, \epsilon_K) \neq \emptyset; \\ \{L(M, \epsilon_K)\}, & \text{otherwise.} \end{cases}$$

Example 4 (Continuing Example 3) Given that $\mathsf{L}(M, \epsilon_{\mathsf{K}}) = \{\mathsf{m}_1, \mathsf{m}_2\}$, then by the first condition in Definition 15 we have that $\mathsf{L}^*(M, \epsilon_{\mathsf{K}}) = \{\{\mathsf{m}_1, \mathsf{m}_2\}\} \cup \mathsf{L}^*(\{\mathsf{m}_3\}, \epsilon_{\mathsf{K}})$, since $M \backslash \{m1, m2\} = \{m3\} \neq \emptyset$. Now, by condition (i) in Definition 14, we have that $\mathsf{L}(\{\mathsf{m}_3\}, \epsilon_{\mathsf{K}}) = \{\mathsf{m}_3\}$. Thus $\mathsf{L}^*(\{\mathsf{m}_3\}, \epsilon_{\mathsf{K}}) = \{\{\mathsf{m}_3\}\}$ by the second condition in Definition 15 since $\{\mathsf{m}_3\} \backslash \{\mathsf{m}_3\} = \emptyset$. Therefore, the set of LPMCSes in M is $\mathsf{L}^*(M, \epsilon_{\mathsf{K}}) = \{\{\mathsf{m}_1, \mathsf{m}_2\}\} \cup \{\{\mathsf{m}_3\}\} = \{\{\mathsf{m}_1, \mathsf{m}_2\}, \{\mathsf{m}_3\}\}$.

Having defined the set of LPMCSes, we can now define our context-dependent combination rule using both Dempster's rule and Dubois & Prade's rule as follows:

Definition 16 Let M be a set of mass functions from independent and reliable sources and ϵ_{K} be a conflict threshold. Then the combined mass function from M with respect to ϵ_{K} using the *context-dependent combination rule* is defined as:

$$\bigotimes_{M' \in \mathsf{L}^*(M, \epsilon_{\mathsf{K}})} \bigoplus_{\mathsf{m} \in M'} \mathsf{m}.$$

In other words, the combined mass function is found by combining each LPMCS using Dempster's rule and then combining the set of combined LPMCSes using Dubois & Prade's rule.

	Context-dependent ($\epsilon_{\mathsf{K}} = 0.4$)	Dempster's rule	Dubois & Prade's rule
$\mathsf{m}(\{a\})$	0.511	0.942	0.327
$\mathsf{m}(\{b, c\})$	0.032	0.058	0.02
$\mathsf{m}(\{a, b, c\})$	0.457	0	0.652

Table 3: Combined mass functions from the set of mass functions M from Table 1.

Example 5 (Continuing Example 4) Given a conflict threshold $\epsilon_{\mathsf{K}} = 0.4$, then the context-dependent combined mass function from M with respect to ϵ_{K} is shown in Table 3. For comparison, the combined mass functions using only Dempster's rule and only Dubois & Prade's rule are also shown in Table 3.

It is evident from Example 5 that the context-dependent combination rule finds a balance between the two other combination rules: neither throwing away too much information (as is the case with Dempster's rule); nor incorporating too much conflict such that it is difficult to make decisions (as is the case with Dubois & Prade's rule).

3.3 Computational Aspects

Algorithm 1 provides a method for executing context-dependent combination of a set of mass functions M, given a conflict threshold ϵ_{K} and strict preference ordering $\succ_P$. On line 1, the set of LPMCSes is initialised. On line 2, a list of mass functions representing the combined quality and preference orderings can be constructed with respect to the non-specificity and strife values of each mass

---

Algorithm 1: Context-dependent combination of a set of mass functions.

Input: Set of mass functions $M \neq \emptyset$, conflict threshold ϵ_{K} and preference ordering $\succ_P$
Output: Combined mass function m_r

1   $M' \leftarrow \emptyset$;
2   $\mathsf{m}_r \leftarrow \max\left(\max\left(M, \succeq_Q\right), \succ_P\right)$;
3   $M \leftarrow M \setminus \{\mathsf{m}_r\}$;
4   **while** $M \neq \emptyset$ **do**
5       $\mathsf{m}_c \leftarrow \max\left(\min\left(M, \preceq_{\mathsf{d}}^{\mathsf{m}_r}\right), \succ_P\right)$;
6   **if** $\mathsf{K}(\mathsf{m}_r, \mathsf{m}_c) \leq \epsilon_{\mathsf{K}}$ **then**
7           $\mathsf{m}_r \leftarrow \mathsf{m}_r \oplus \mathsf{m}_c$;
8           $M \leftarrow M \setminus \{\mathsf{m}_c\}$;
9   **else**
10          $M' \leftarrow M' \cup \{\mathsf{m}_r\}$;
11          $\mathsf{m}_r \leftarrow \max\left(\max\left(M, \succeq_Q\right), \succ_P\right)$;
12          $M \leftarrow M \setminus \{\mathsf{m}_r\}$;

13   **for each** $\mathsf{m}'_r \in M'$ **do**
14      $\mathsf{m}_r \leftarrow \mathsf{m}_r \otimes \mathsf{m}'_r$;

15   **return** m_r;

---

function in M and the preference ordering $\succ_P$. The first element in this list would then be selected as the initial reference mass function m_r as in Definition 10. Moreover, this list can be referred to again on line 11. Lines 5–8 describe how to compute an LPMCS with respect to the current reference mass function m_r. Specifically, on line 5, m_c represents the closest and most preferred mass function to m_r as in Definition 13. It is not necessary to compute the full ordering $\succeq_{\mathsf{d}}^{\mathsf{m}_r}$ on line 5, instead we only require one iteration through M (in order of $\succ_P$) to find m_c. This mass function is compared with m_r on line 6 which represents the condition for including a mass function in the current LPMCS as in Definition 14. On line 7, we combine the reference mass function with the closest and most preferred mass function and this combined mass function becomes our new reference. Lines 9–12 describe the behaviour when a complete LPMCS has been found. In particular, the complete LPMCS (which has been combined using Dempster's rule) is added to the set of previously found LPMCSes on line 10. Then, on line 11, we find the reference mass function for a new LPMCS by referring back to the combined quality and preference list constructed on line 2. Finally, when there are no remaining mass functions in M, the current reference m_r represents the final combined LPMCS. However, this LPMCS is not added to the set of LPMCSes M', rather we just begin combining each combined LPMCS in M' with m_r using Dubois & Prade's rule as described on lines 13 and 14. Once this is complete, the mass function m_r represents the final context-dependent combined mass function as in Definition 16 and this is returned on line 15.

Proposition 1 *The complexity of applying the context-dependent combination rule on a set of mass functions M is $\mathcal{O}(n^2)$ such that $n = |M|$.*

Proof In order to select a reference mass function, we only need to compute n pairs of non-specificity and strife values. Thus, the complexity of selecting a reference mass function is $\mathcal{O}(n)$. After selecting a reference mass function, there can be at most $n - 1$ mass functions which have not been added to an LPMCS. In this case,

selecting a closest and most preferred mass function requires the computation of $n - 1$ distance values. Thus, the complexity of selecting a closest and most preferred mass function is also $\mathcal{O}(n)$. In the worst case, where every LPMCS is a singleton, we need to select n reference mass functions of which there are $n - 1$ requiring the selection of a closest and most preferred mass function. Thus, the overall complexity is $\mathcal{O}(n^2)$. $\square$

4 Handling Sensor Information in BDI

In this section we begin by introducing preliminaries on the AgentSpeak framework [29] for BDI agents. We then describe the necessary steps for handling uncertain sensor information in AgentSpeak. In Section 5, we will present a power grid SCADA scenario which demonstrates how sensor information can be modelled and combined using DS theory and then incorporated into a BDI agent defined in AgentSpeak.

4.1 AgentSpeak

AgentSpeak is a logic-based programming language related to a restricted form of first-order logic, combined with notions for *events* and *actions*. In practice, it shares many similarities with popular logic programming languages such as Prolog. The language itself consists of *variables*, *constants*, *predicate symbols*, *action symbols*, *logical connectives* and *punctuation*. Following logic programming convention, variables begin with uppercase letters while constants, predicate symbols and action symbols begin with lowercase letters. As in first-order logic, these constructs can be understood as *terms*, i.e. a variable is a term, a constant is a term and, if p is an n-ary predicate symbol and $t_1, \ldots, t_n$ are terms, then $p(t_1, \ldots, t_n)$ is a term. A term is said to be *ground* or *instantiated* if there are no variables or if each variable is bound to a constant by a *unifier*. We will use $\mathbf{t}$ to denote the terms $t_1, \ldots, t_n$. The supported logical connectives are $\neg$ and $\wedge$ for negation and conjunction, respectively. Also, the punctuation in AgentSpeak includes !, ?, ; and $\leftarrow$ and the meaning of each will be explained later. From [29], the syntax of the AgentSpeak language is defined as follows:

Definition 17 If b is a predicate symbol, then $b(\mathbf{t})$ is a *belief atom*. If $b(\mathbf{t})$ is a belief atom, then $b(\mathbf{t})$ and $\neg b(\mathbf{t})$ are *belief literals*.

Definition 18 If g is a predicate symbol, then $!g(\mathbf{t})$ and $?g(\mathbf{t})$ are *goals* where $!g(\mathbf{t})$ is an *achievement goal* and $?g(\mathbf{t})$ is a *test goal*.

Definition 19 If $b(\mathbf{t})$ is a belief atom and $!g(\mathbf{t})$ and $?g(\mathbf{t})$ are goals, then $+b(\mathbf{t})$, $-b(\mathbf{t})$, $+!g(\mathbf{t})$, $-!g(\mathbf{t})$, $+?g(\mathbf{t})$ and $-?g(\mathbf{t})$ are *triggering events* where $+$ and $-$ denote addition and deletion events, respectively.

Definition 20 If a is an action symbol, then $a(\mathbf{t})$ is an *action*.

Definition 21 If e is a triggering event, $l_1, \ldots, l_m$ are belief literals and $h_1, \ldots, h_n$ are goals or actions, then $e : l_1 \wedge \ldots \wedge l_m \leftarrow h_1; \ldots; h_n$ is a *plan* where $l_1 \wedge \ldots \wedge l_m$ is the *context* and $h_1; \ldots; h_n$ is the *body* such that ; denotes sequencing.

An AgentSpeak agent $\mathbb{A}$ is a tuple $\langle Bb, Pl, A, E, I \rangle^3$ with a belief base Bb, a plan library Pl, an action set A, an event set E and an intention set I. The belief base Bb is a set of belief atoms[4] which describe the agent's current beliefs about the world. The plan library Pl is a set of predefined plans used for reacting to new events. The action set A contains the set of primitive actions available to the agent. The event set E contains the set of events which the agent has yet to consider. Finally, the intention set I contains those plans which the agent has chosen to pursue where each intention is a stack of partially executed plans. Intuitively, when an agent reacts to a new event e, it selects those plans from the plan library which have e as their triggering event (called the *relevant* plans for e). A relevant plan is said to be *applicable* when the context of the plan evaluates to true with respect to the agent's current belief base. For a given event e, there may be many applicable plans from which one is selected and either added to an existing intention or used to form a new intention. Intentions in the intention set are executed concurrently. Each intention is executed by performing the steps described in the body of each plan in the stack (e.g. executing primitive actions or generating new subgoals). Thus, the execution of an intention may change the environment and/or the agent's beliefs. As such, if the execution of an intention results in the generation of new events (e.g. the addition of subgoals), then more plans may be added to the stack for execution. In Section 5 we will consider AgentSpeak programs in more detail but for the rest of this section we are primarily interested in how to derive belief atoms from sensors.

4.2 Deriving Beliefs from Uncertain Sensor Information

By their nature, heterogeneous sensors output different types of sensor information. While this presents some important practical challenges in terms of how to construct mass functions from original sensor information, any solutions will be unavoidably domain-specific. In this paper, we do not attempt to address this issue directly but we can suggest some possibilities. For example, if a sensor returns numerical information in the form of an interval, then we can assign a mass of 1 to the set of discrete readings included in the interval. Similarly, if a sensor returns a reading with a certainty of 80%, then we can assign a mass of 0.8 to the relevant focal set. So, without loss of generality, we assume that an appropriate mass function can be derived from any source. Moreover, we assume that each source has a reliability degree $1 - \alpha$ and that each mass function has been discounted with respect to the α value of the originating source. Thus, for the remainder of this section, we assume that each mass function is actually a discounted mass function m^α which can be treated as fully reliable. However, it is unlikely that all compatible sources (i.e. sources with information which is suitable for combination) will return strictly compatible mass functions (i.e. mass functions defined over the same frame). To address this issue, the notion of an evidential mapping was introduced in [23] as follows:

---

^{3} For simplicity, we omit the selection functions $\mathcal{S}_\mathcal{E}$, $\mathcal{S}_\mathcal{O}$ and $\mathcal{S}_\mathcal{I}$.

^{4} In the AgentSpeak language, $\neg$ denotes negation as failure rather than strong negation since the belief base only contains belief atoms, i.e. positive belief literals.

Definition 22 Let Ω_e and Ω_h be frames. Then a function $\Gamma : \Omega_e \times 2^{\Omega_h} \to [0,1]$ is called an *evidential mapping* from Ω_e to Ω_h if for each $\omega_e \in \Omega_e$, $\Gamma(\omega_e, \emptyset) = 0$ and $\sum_{H \subseteq \Omega_h} \Gamma(\omega_e, H) = 1$.

In other words, for each $\omega_e \in \Omega_e$, an evidential mapping defines a mass function over Ω_h which describes how ω_e relates to Ω_h. In line with DS terminology, a set $H \subseteq \Omega_h$ is called a *projected focal element* of $\omega_e \in \Omega_e$ if $\Gamma(\omega_e, H) > 0$. Importantly, given a mass function over Ω_e, then an evidential mapping allows us to derive a new mass function over Ω_h directly as follows [23]:

Definition 23 Let Ω_e and Ω_h be frames, m_e be a mass function over Ω_e and Γ be an evidential mapping from Ω_e to Ω_h. Then a mass function m_h over Ω_h is called an *evidence propagated* mass function from m_e with respect to Γ and is defined for each $H \subseteq \Omega_h$ as:

$$\mathsf{m}_h(H) = \sum_{E \subseteq \Omega_e} \mathsf{m}_e(E) \Gamma^*(E, H),$$

where:

$$\Gamma^*(E, H) = \begin{cases} \sum_{\omega_e \in E} \frac{\Gamma(\omega_e, H)}{|E|}, & \text{if } H \neq \bigcup H_E \text{ and } \forall \omega_e \in E,\ \Gamma(\omega_e, H) > 0, \\ 1 - \sum_{H' \in H_E} \Gamma^*(E, H'), & \text{if } H = \bigcup H_E \text{ and } \exists \omega_e \in E,\ \Gamma(\omega_e, H) = 0, \\ 1 - \sum_{H' \in H_E} \Gamma^*(E, H') & \text{if } H = \bigcup H_E \text{ and } \forall \omega_e \in E,\ \Gamma(\omega_e, H) > 0, \\ \quad + \sum_{\omega_e \in E} \frac{\Gamma(\omega_e, H)}{|E|}, & \\ 0, & \text{otherwise,} \end{cases}$$

such that $H_E = \{H' \subseteq \Omega_h \mid \omega_e \in E, \Gamma(\omega_e, H') > 0\}$ and $\bigcup H_E = \{\omega_h \in H' \mid H' \in H_E\}$.

The value $\Gamma^*(E, H)$ from Definition 23 actually represents a *complete* evidential mapping for E and H as defined in [23] such that Γ is a *basic* evidential mapping function. In particular, the first condition defines the mass for E and H directly from Γ if each $\omega_e \in E$ implies H to some degree, i.e. if H is a projected focal element of each $\omega_e \in E$. In this case, the actual mass for E is taken as the average mass from the basic mapping for all supporting elements in E. The second condition then recursively assigns the remaining mass that was not assigned by the first condition to the set of possible hypotheses implied by E, i.e. the union of all projected focal elements for all $\omega_e \in E$. The third condition says that when H is a projected focal element of each $\omega_e \in E$ and when $H = \bigcup H_E$ then we need to sum the results from the first two conditions. Finally, the fourth condition just assigns 0 to all other subsets. Note that it is not necessary to compute the complete evidential mapping function in full since we are only interested in focal elements of m_e.

Example 6 Given the evidential mapping from Table 4 and a mass function m_e over Ω_e such that $\mathsf{m}_e(\{e_2\}) = 0.7$ and $\mathsf{m}_e(\Omega_e) = 0.3$, then a mass function m_h over Ω_h is the evidence propagated mass function from m_e with respect to Γ such that $\mathsf{m}_h(\{h_1, h_2\}) = 0.49$, $\mathsf{m}_h(\Omega_h) = 0.3$ and $\mathsf{m}_h(\{h_4\}) = 0.21$.

	$\{h_1, h_2\}$	$\{h_3\}$	$\{h_4\}$
e_1	0.5	0.5	0
e_2	0.7	0	0.3
e_3	0	0	1

(a) Evidential mapping Γ.

	$\{h_1, h_2\}$	$\{h_3\}$	$\{h_4\}$	$\{h_1, h_2, h_4\}$	Ω_h
$\{e_1\}$	0.5	0.5	0.0	0.0	0.0
$\{e_2\}$	0.7	0.0	0.3	0.0	0.0
$\{e_3\}$	0.0	0.0	1.0	0.0	0.0
$\{e_1, e_2\}$	0.6	0.0	0.0	0.0	0.4
$\{e_1, e_3\}$	0.0	0.0	0.0	0.0	1.0
$\{e_2, e_3\}$	0.0	0.0	0.65	0.35	0.0
Ω_e	0.0	0.0	0.0	0.0	1.0

(b) Complete evidential mapping Γ^*.

Table 4: Evidential mapping Γ from $\Omega_e = \{e_1, e_2, e_3\}$ to $\Omega_h = \{h_1, \ldots, h_4\}$.

In terms of integrating uncertain sensor information into AgentSpeak, evidence propagation provides two benefits. Firstly, it allows us to correlate relevant mass functions that may be defined over different frames such that we can derive compatible mass functions which are suitable for combination. Secondly, it allows us to map disparate sensor languages into a standard AgentSpeak language, i.e. where each frame is a set of AgentSpeak belief atoms. For the remainder of this section we will assume that evidence propagation has been applied and that any mass functions which should be combined are already defined over the same frame. However, in terms of decision making (i.e. the selection of applicable plans), classical AgentSpeak is not capable of modelling and reasoning with uncertain information. As such, it is necessary to reduce the uncertain information modelled by a mass function to a classical belief atom which can be modelled in the agent's belief base. For this purpose, we propose the following:

Definition 24 Let m be a mass function over Ω and $0.5 < \epsilon_{\mathsf{BetP}} \leq 1$ be a pignistic probability threshold. Then m *entails* the hypothesis $\omega \in \Omega$, denoted $\mathsf{m} \models_{\epsilon_{\mathsf{BetP}}} \omega$, if $\mathsf{BetP_m}(\omega) \geq \epsilon_{\mathsf{BetP}}$.

When a mass function is defined over a set of belief atoms, the threshold ϵ_{BetP} ensures that at most one belief atom is entailed by the mass function. However, this definition also means that it is possible for no belief atom to be entailed. In this situation, the interpretation is that we have insufficient reason to believe that any belief atom in the frame is true. Thus, if no belief atom is entailed by a mass function then we are ignorant about the frame of this mass function. In AgentSpeak we can reason about this ignorance using negation as failure. While this approach is a drastic (but necessary) step to modelling uncertain sensor information in classical AgentSpeak, an extension to AgentSpeak was proposed in [3] which allows us to model and reason about uncertain beliefs directly. Essentially, for each mass function m, the pignistic probability distibution $\mathsf{BetP_m}$ would be modelled as one of many *epistemic states* in the extended AgentSpeak agent's belief base, now called a *global uncertain belief set* (GUB). Using an extended AgentSpeak language, we

can then reason about uncertain information held in each epistemic state using qualitative *plausibility* operators (e.g. describing that some belief atom is more plausible than another). For simplicity, we will not present this work in detail but instead refer the reader to [3] for the full definition of a GUB.

Obviously there is a computational cost associated with combination. For this reason, we introduce the following conditions to describe when information has changed sufficiently to warrant combination and revision:

Definition 25 Let S be a set of sources, Ω be a frame, $\mathsf{m}_{s,t}$ be a mass over Ω from source $s \in S$ at time t and ϵ_{d} be a distance threshold. Then the context-dependent combination rule is applied at time $t+1$ to the set $\{\mathsf{m}_{s,t+1} \mid s \in S\}$ if there exists a source $s' \in S$ such that $\mathsf{d}(\mathsf{m}_{s',t}, \mathsf{m}_{s',t+1}) \geq \epsilon_{\mathsf{d}}$.

This definition says that the context-dependent combination rule should only be applied when the information obtained from any one source has changed sufficiently (formalised by the distance threshold ϵ_{d}). However, to ensure that smaller incremental changes are not overlooked, it is still necessary to combine and revise information at some specified intervals. Thus, this definition just provides a method to minimize redundant computation.

In summary, integration with an AgentSpeak agent can be handled by a sensor information preprocessor or *intelligent sensor* which can perform the following steps:

(i) Discount each mass function with respect to the source reliability degree $1-\alpha$.
(ii) Apply evidence propagation using evidential mappings to derive compatible mass functions for combination which are defined over AgentSpeak belief atoms.
(iii) Combine relevant mass functions by applying the context-dependent combination rule if any source has returned sufficiently dissimilar information from the information previously recorded for that source.
(iv) Derive a belief atom from the combined mass function and revise the AgentSpeak agent's belief base with this new belief atom.

Note that even if information for some source does not need to be combined with information for another, then steps (i), (ii) and (iv) can still be performed to reflect the reliability of the source and to derive suitable AgentSpeak belief atoms.

5 Scenario: Power Grid SCADA System

This scenario is inspired by a case study on an electric power grid SCADA system used to manage power generation, transmission and distribution. In the power grid, system *frequency* refers to the frequency of oscillations of alternating current (AC) transmitted from the power grid to the end-user. In the UK, the frequency standard is 50 Hz and suppliers are obligated to maintain the frequency within a deviation of $\pm 1\%$. If demand is greater than generation then frequency will decrease but if generation is greater than demand then frequency will increase. Thus, maintaining the correct frequency level is a load balancing problem between power demand and power generation. In this scenario, the power grid is managed by a BDI agent encoded in AgentSpeak with a set of belief atoms, a set of primitive actions and a predefined plan library. With regards to frequency, the agent may believe the

current level is low, normal or high (represented by belief atoms `freq(l)`, `freq(n)` and `freq(h)`, respectively). Examples of AgentSpeak plans in the agent's plan library include the following:

P1 `+!run_combiner : freq(n) <- !run_inverter; distribute_power; ...`
P2 `+freq(l) : true <- !generate_alert; !stop_combiner; !run_windfarm; ...`

The plan **P1** describes the steps that should be taken if the agent obtains the goal `!run_combiner` and also believes that the current frequency level is normal. These steps involve a new sub-goal `!run_inverter` and a primitive action `distribute_power`. On the other hand, the plan **P2** describes the steps that should be taken when the agent obtains the belief that the current frequency level is low. In this case, the steps involve new sub-goals `!generate_alert`, `!stop_combiner` and `!run_windfarm` with the aim of increasing power generation.

	{`freq(l)`}	{`freq(l)`, `freq(n)`}	{`freq(n)`}	{`freq(n)`, `freq(h)`}	{`freq(h)`}
48.7	1	0	0	0	0
48.8	1	0	0	0	0
48.9	1	0	0	0	0
49	0.75	0.25	0	0	0
49.1	0.5	0.5	0	0	0
49.2	0.25	0.75	0	0	0
49.3	0	1	0	0	0
49.4	0	1	0	0	0
49.5	0	1	0	0	0
49.6	0	0.75	0.25	0	0
49.7	0	0.5	0.5	0	0
49.8	0	0.25	0.75	0	0
49.9	0	0	1	0	0
50	0	0	1	0	0
50.1	0	0	1	0	0
50.2	0	0	0.75	0.25	0
50.3	0	0	0.5	0.5	0
50.4	0	0	0.25	0.75	0
50.5	0	0	0	1	0
50.6	0	0	0	1	0
50.7	0	0	0	1	0
50.8	0	0	0	0.75	0.25
50.9	0	0	0	0.5	0.5
51	0	0	0	0.25	0.75
51.1	0	0	0	0	1
51.2	0	0	0	0	1
51.3	0	0	0	0	1

(a) Sensor frame Ω_e^s.

	{`freq(l)`}	{`freq(n)`}	{`freq(h)`}
normal	0	1	0
abnormal	0.5	0	0.5

(b) Expert frame Ω_e^e.

Table 5: Evidential mappings from Ω_e^s and Ω_e^e to Ω_h.

Source	Type	Reliability ($1-\alpha$)	Value
s_1	sensor	0.98	51.2 Hz
s_2	sensor	0.7	51.2 Hz (± 0.1 Hz)
s_3	sensor	0.92	51.3 Hz
s_4	sensor	0.75	51.1 Hz (± 0.2 Hz)
s_5	expert	0.85	abnormal (90% certain)
s_6	sensor	0.8	51 Hz (± 0.2 Hz)
s_7	sensor	0.96	50.9 Hz (± 0.1 Hz)
s_8	sensor	0.9	51.1 Hz
s_9	sensor	0.95	51.1 Hz (± 0.1 Hz)
s_{10}	expert	0.65	abnormal (75% certain)

(a) Source evidence where $s_1 \succ_P s_2 \succ_P \cdots \succ_P s_{10}$.

s_i	m_i	α	m_i^α
s_1	$\{\{51.2\} \mapsto 1\}$	0.02	$\{\{51.2\} \mapsto 0.98, \Omega \mapsto 0.02\}$
s_2	$\{\{51.1, 51.2, 51.3\} \mapsto 1\}$	0.3	$\{\{51.1, 51.2, 51.3\} \mapsto 0.7, \Omega \mapsto 0.3\}$
s_3	$\{\{51.3\} \mapsto 1\}$	0.08	$\{\{51.3\} \mapsto 0.92, \Omega \mapsto 0.08\}$
s_4	$\{\{50.9, 51.0, \ldots, 51.3\} \mapsto 1\}$	0.25	$\{\{50.9, 51.0, \ldots, 51.3\} \mapsto 0.75, \Omega \mapsto 0.25\}$
s_5	$\{\{\text{abnormal}\} \mapsto 0.9, \Omega \mapsto 0.1\}$	0.15	$\{\{\text{abnormal}\} \mapsto 0.765, \Omega \mapsto 0.235\}$
s_6	$\{\{50.8, 50.9, \ldots, 51.2\} \mapsto 1\}$	0.2	$\{\{50.8, 50.9, \ldots, 51.2\} \mapsto 0.8, \Omega \mapsto 0.2\}$
s_7	$\{\{50.8, 50.9, 51.0\} \mapsto 1\}$	0.04	$\{\{50.8, 50.9, 51.0\} \mapsto 0.96, \Omega \mapsto 0.04\}$
s_8	$\{\{51.1\} \mapsto 1\}$	0.1	$\{\{51.1\} \mapsto 0.9, \Omega \mapsto 0.1\}$
s_9	$\{\{51.0, 51.1, 51.2\} \mapsto 1\}$	0.05	$\{\{51.0, 51.1, 51.2\} \mapsto 0.95, \Omega \mapsto 0.05\}$
s_{10}	$\{\{\text{abnormal}\} \mapsto 0.75, \Omega \mapsto 0.25\}$	0.35	$\{\{\text{abnormal}\} \mapsto 0.488, \Omega \mapsto 0.512\}$

(b) Evidence from s_i modelled as mass function m_i and discounted mass function m_i^α.

	m_1^h	m_2^h	m_3^h	m_4^h	m_5^h	m_6^h	m_7^h	m_8^h	m_9^h	m_{10}^h
$\mathsf{m}_i^h(\{\texttt{freq(l)}\})$	0	0	0	0	0.383	0	0	0	0	0.244
$\mathsf{m}_i^h(\{\texttt{freq(n)}, \texttt{freq(h)}\})$	0	0	0	0.113	0	0.24	0.48	0	0.079	0
$\mathsf{m}_i^h(\{\texttt{freq(h)}\})$	0.98	0.7	0.92	0.637	0.383	0.56	0.48	0.9	0.871	0.244
$\mathsf{m}_i^h(\Omega_h)$	0.02	0.3	0.08	0.25	0.235	0.2	0.04	0.1	0.05	0.512

(c) Evidence propagated mass function m_i^h from m_i^α with respect to Table 5.

Table 6: Information obtained from sources.

Obviously, determining an accurate view of the current frequency level is vital for ensuring that the correct level is maintained. In this scenario, there are many sources of information which the agent may consult to determine the current level. For simplicity, we will focus on two main types of sources relevant to frequency. The first is sensors which provide numerical measurements of frequency levels from the set $\Omega_e^s = \{48.7, 48.8, \ldots, 51.3\}$. The second is experts which provide general estimations of frequency levels from the set $\Omega_e^e = \{\text{normal}, \text{abnormal}\}$. Information from both types of sources may be effected by uncertainty. For example, the accuracy of sensors may vary due to design limitations or environmental factors, experts may lack confidence in their estimations, sources may be treated as unreliable due to previously incorrect information, etc. In Table 5 we define evidential mappings from Ω_e^s and Ω_e^e to the set of AgentSpeak belief atoms $\Omega_h = \{\texttt{freq(l)}, \texttt{freq(n)}, \texttt{freq(h)}\}$. These mappings allow us to combine different types of information and to derive suitable belief atoms for revising the agent's belief base. Essentially the sensor mapping defines a fuzzy distribution over the

range of possible sensor values while the expert mapping defines a probabilistic relation between the abnormal estimation and the belief atoms `freq(l)` and `freq(h)` (assigning equal probability to each).

Table 6a represents an illustrative scenario in which a number of sensors and experts have provided information about the current frequency level. Some sensor values have a deviation (e.g. ± 0.1 Hz) while experts have expressed a degree of confidence in their assessment (e.g. 90% certain). For example, source s_2 is a sensor which has determined that the frequency is 51.2 Hz with a deviation ± 0.1 Hz whereas s_5 is an expert which has estimated that the frequency level is abnormal with 90% certainty. Moreover, each source has a prior reliability rating $1 - \alpha$ based on historical data, age, design limitations and faults etc. From Table 6a, we can see that s_1 is the most reliable source whilst s_{10} is the least reliable. In Table 6b, we can see how the information provided in Table 6a is modelled as mass functions and how these correspond to discounted mass functions, given their respective α values. For example, source s_1 has determined that the frequency is 51.2 Hz. By modelling this sensor information as a mass function m_1, we have that a mass of 1 is assigned to the singleton set $\{51.2\}$. Alternatively, source s_{10} has estimated that the frequency level is abnormal with 75% certainty. By modelling this expert information as a mass function m_{10}, we have that a mass of 0.75 is assigned to the singleton set $\{abnormal\}$, while the remaining mass is assigned to the frame Ω. Each mass function is then discounted with respect to the reliability of the original source. For example, the source s_1 has a reliability of 0.98 and so the mass function m_1 is discounted using a discount factor $\alpha = 1 - 0.98 = 0.02$. Finally, Table 6c provides the evidence propagated mass functions over Ω_h with respect to Table 5 representing the compatible mass functions from each source.

In this scenario we are not interested in the information provided by individual sources, rather we are only interested in reasoning about the information as a whole. To achieve this, we must first combine the set of evidence propagated mass functions from Table 6c. In Table 7 we demonstrate how to apply the context-dependent combination rule on this set of mass functions by computing LPMCSes using an arbitrarily chosen conflict threshold $\epsilon_{\mathsf{K}} = 0.3$. Firstly, Table 7a provides the quality of information heuristics used to find a reference mass function. In this case, we select m_1^h as the reference since it has the minimum non-specificity and strife values. This mass function is then used to form a new (partial) LPMCS. Secondly, Table 7b provides the distance values between m_1^h and all other mass functions. From these, we select m_3^h as the next mass function in the preferred sequence of combination since it has the minimum distance to m_1^h. Thirdly, referring to Table 7c we can see that the conflict between m_1^h and m_3^h is under the threshold meaning that m_3^h is added to the LPMCS and $\mathsf{m}_1^h \oplus \mathsf{m}_3^h$ becomes the new reference mass function. This process continues until we reach m_5^h where the conflict between $\mathsf{m}_1^h \oplus \ldots \oplus \mathsf{m}_7^h$ and m_5^h is above the threshold. Thus, the set $\{\mathsf{m}_1^h, \ldots, \mathsf{m}_4^h, \mathsf{m}_6^h, \ldots, \mathsf{m}_9^h\}$ is the first LPMCS with respect to ϵ_{K} and a new reference must be selected from the set $\{\mathsf{m}_5^h, \mathsf{m}_{10}^h\}$. Referring again to Table 7a we find that m_5^h is the highest quality mass function and select this as our new reference. Since m_{10}^h is the only remaining mass function, we just need to measure the conflict between m_5^h and m_{10}^h where we find that the conflict is below the threshold. Thus, the set $\{\mathsf{m}_5^h, \mathsf{m}_{10}^h\}$ is the second LPMCS and the complete set of LPMCSes is $\mathsf{L}^*(\{\mathsf{m}_1^h, \ldots, \mathsf{m}_{10}^h\}, 0.3) = \{\{\mathsf{m}_1^h, \ldots, \mathsf{m}_4^h, \mathsf{m}_6^h, \ldots, \mathsf{m}_9^h\}, \{\mathsf{m}_5^h, \mathsf{m}_{10}^h\}\}$.

	m_1^h	m_2^h	m_3^h	m_4^h	m_5^h	m_6^h	m_7^h	m_8^h	m_9^h	m_{10}^h
$N(m_i^h)$	0.032	0.475	0.127	0.509	0.372	0.557	0.543	0.158	0.158	0.812
$S(m_i^h)$	0.031	0.272	0.11	0.286	0.774	0.288	0.228	0.132	0.133	0.487

(a) Non-specificity and strife values.

m_i^h	m_2^h	m_3^h	m_4^h	m_5^h	m_6^h	m_7^h	m_8^h	m_9^h	m_{10}^h
m_1^h	0.229	0.049	0.251	0.509	0.294	0.351	0.065	0.076	0.584
$m_1^h \oplus m_3^h$	0.244	–	0.266	0.521	0.308	0.363	0.08	0.089	0.598
$m_1^h \oplus m_3^h \oplus m_8^h$	0.245	–	0.267	0.522	0.309	0.364	–	0.09	0.599
$m_1^h \oplus \ldots \oplus m_9^h$	0.245	–	0.267	0.522	0.309	0.364	–	–	0.6
$m_1^h \oplus \ldots \oplus m_2^h$	–	–	0.267	0.523	0.309	0.364	–	–	0.6
$m_1^h \oplus \ldots \oplus m_4^h$	–	–	–	0.523	0.309	0.364	–	–	0.6
$m_1^h \oplus \ldots \oplus m_6^h$	–	–	–	0.523	–	0.364	–	–	0.6
$m_1^h \oplus \ldots \oplus m_7^h$	–	–	–	0.523	–	–	–	–	0.6

(b) Distance values.

m_i	m_j	$K(m_i, m_j)$
m_1^h	m_3^h	0
$m_1^h \oplus m_3^h$	m_8^h	0
$m_1^h \oplus m_3^h \oplus m_8^h$	m_9^h	0
$m_1^h \oplus \ldots \oplus m_9^h$	m_2^h	0
$m_1^h \oplus \ldots \oplus m_2^h$	m_4^h	0
$m_1^h \oplus \ldots \oplus m_4^h$	m_6^h	0
$m_1^h \oplus \ldots \oplus m_6^h$	m_7^h	0
$m_1^h \oplus \ldots \oplus m_7^h$	m_5^h	0.382
m_5^h	m_{10}^h	0.186

(c) Conflict values.

Table 7: Finding LPMCSes in $\{m_1^h, \ldots, m_{10}^h\}$ when $\epsilon_K = 0.3$.

	Context-dependent ($\epsilon_K = 0.3$)	Dempster's rule	Dubois & Prade's rule
$m(\{\texttt{freq(h)}\})$	0.426	1	0.008
$m(\{\texttt{freq(l)}, \texttt{freq(h)}\})$	0.426	0	0.024
$m(\{\texttt{freq(n)}, \texttt{freq(h)}\})$	0	0	0.021
$m(\Omega_h)$	0.148	0	0.947

(a) Combined mass functions.

	Context-dependent ($\epsilon_K = 0.3$)	Dempster's rule	Dubois & Prade's rule
$BetP_m(\texttt{freq(l)})$	0.262	0	0.328
$BetP_m(\texttt{freq(n)})$	0.049	0	0.326
$BetP_m(\texttt{freq(h)})$	0.688	1	0.346

(b) Pignistic probability distributions.

Table 8: Combined results.

The combined mass function after applying the context-dependent combination rule on the set of mass functions, using an arbitrary conflict threshold $\epsilon_K = 0.3$, is provided in Table 8a. In addition, the results for the full range of conflict thresholds $\epsilon_K = 0, 0.01, ..., 1$ are shown graphically in Figure 1. For comparison, the combined mass functions after applying Dempster's rule and Dubois & Prade's rule in full are also provided. Table 8b then includes the results for applying the pignistic transformation on these combined mass functions. Regarding these results, it is apparent that Dempster's rule arrives at a very strong conclusion with complete belief in the belief atom `freq(h)`. On the other hand, Dubois & Prade's rule arrives at almost complete ignorance over the set of possible belief atoms Ω_h. The result of the context-dependent rule is arguably more intuitive since it doesn't throw away differing opinions (as is the case with the former) or incorporate too much conflict (as is the case with the latter). For example, after applying the context-dependent rule we can conclude that `freq(l)` is more probable than `freq(n)`. Conversely, this conclusion is not supported using Dempster's rule while the support is relatively insignificant using Dubois & Prade's rule. Nonetheless, the final stage of this scenario is to attempt to derive a belief atom for revising the agent's belief base. In this case, if we assume a pignistic probability threshold $\epsilon_{BetP} = 0.6$ then we can derive the belief atom `freq(h)`. If `freq(h)` is added to the agent's belief base, then any plans with this belief addition event as their triggering event will be selected as applicable plans. Thus, the agent can respond by reducing power generation in order to return to a normal frequency level.
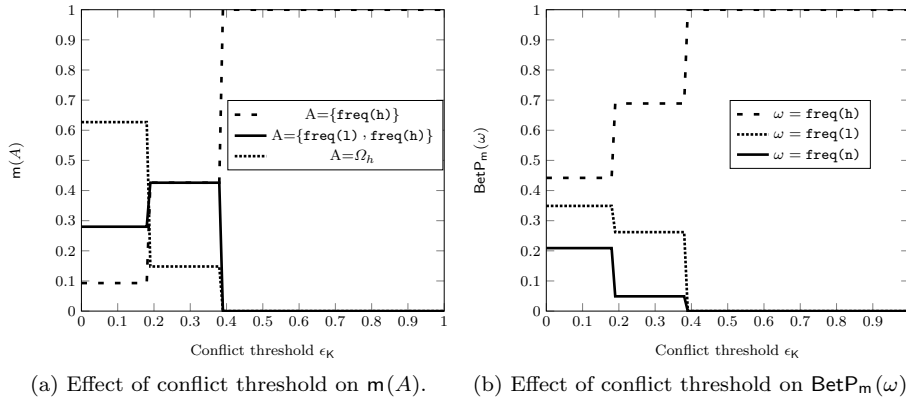

(a) Effect of conflict threshold on $m(A)$.   (b) Effect of conflict threshold on $BetP_m(\omega)$.

Figure 1: The effect of varying the conflict threshold on $m(A)$ and $BetP_m(\omega)$.

In Figure 1, we illustrate the effect of varying the conflict threshold and pignistic probability threshold. Figure 1a shows the mass values of focal elements after combining all relevant mass functions and for each conflict threshold $\epsilon_K = 0, 0.01, ..., 1$. For example, when $\epsilon_K = 0.1$, $m(\{\texttt{freq(h)}, \texttt{freq(l)}\}) = 0.28$, $m(\{\texttt{freq(h)}\}) = 0.093$, $m(\Omega) = 0.627$. When $\epsilon_K \leq 0.18$, the context-dependent combination rule results in three LPMCSes $\{\{m_1^h, \ldots, m_4^h, m_6^h, \ldots, m_9^h\}, \{m_5^h\}, \{m_{10}^h\}\}$, where Ω_h is assigned the highest combined mass value. In contrast, when a

conflict threshold is in the interval $[0.19, 0.38]$, the context-dependent combination rule results in two LPMCSes $\{\{\mathsf{m}_1^h, \ldots, \mathsf{m}_4^h, \mathsf{m}_6^h, \ldots, \mathsf{m}_9^h\}, \{\mathsf{m}_5^h, \mathsf{m}_{10}^h\}\}$. In this case, Ω_h has the lowest combined mass value whilst the other focal sets are assigned equal combined mass values. Furthermore, when $\epsilon_{\mathsf{K}} \geq 0.39$, the context-dependent combination rule results in one LPMCS, i.e. $\{\{\mathsf{m}_1^h, \ldots, \mathsf{m}_{10}^h\}\}$, where $\{\mathtt{freq(h)}\}$ is assigned a combined mass value of 1. In this case, the context-dependent combination rule reduces to Dempster's rule. Conversely, in this scenario, the context-dependent combination rule never reduces to Dubois and Prade's rule since, even when $\epsilon_{\mathsf{K}} = 0$, we never have a set of singleton LPMCSes. Notably, as the conflict threshold increases, the mass value of $\{\mathtt{freq(h)}\}$ increases whilst the mass value of Ω_h decreases. However, the mass value of $\{\mathtt{freq(l)}, \mathtt{freq(h)}\}$ fluctuates (i.e. increasing first before decreasing). In particular, the mass value increases in the interval $[0, 0.38]$, due to the mass value of Ω_h being shifted to strict subsets. By comparing the results from Figure 1a, we can conclude in this scenario that a conflict threshold in the interval $[0, 0.38]$ will offer a balance between both rules such that our rule does not reduce to one rule or the other. In Figure 1b we show the corresponding pignistic probability distribution after applying the pignistic transformation on the combined mass function to demonstrate the effect of varying the conflict threshold on decision making. For example, when $\epsilon_{\mathsf{K}} = 0.1$, $\mathsf{BetP_m}(\mathtt{freq(h)}) = 0.442$, $\mathsf{BetP_m}(\mathtt{freq(l)}) = 0.349$, $\mathsf{BetP_m}(\mathtt{freq(n)}) = 0.209$. As can be seen, when the conflict threshold increases, the probability assigned to $\mathtt{freq(h)}$ increases whilst the probability of $\mathtt{freq(l)}$ and $\mathtt{freq(n)}$ decrease. By comparing the probability distributions in Figure 1b, it indicates $\mathtt{freq(h)}$ is the obvious conclusion for deriving a belief atom for decision-making. However, the choice of thresholds will affect whether a belief atom is actually derived. Specifically, if $\epsilon_{\mathsf{BetP}} \leq 0.7$ and $\epsilon_{\mathsf{BetP}} \geq 0.2$ then $\mathtt{freq(h)}$ will be derived, otherwise nothing will be derived. This demonstrates the benefit of our rule since relying on Dubois and Prade's rule alone would not allow us to derive a belief atom while Dempster's rule loses potentially useful uncertain information.

6 Conclusion

In this paper we presented an approach to handle uncertain sensor information obtained from heterogeneous sources for aiding decision making in BDI. Specifically, we adapted an existing context-dependent combination strategy from the setting of possibility theory to DS theory. As such, we proposed a new combination strategy which balances the benefits of Dempster's rule for reaching meaningful conclusions for decision making and of alternatives, such as Dubois & Prade's rule, for handling conflict. We also proposed an approach to handle this sensor information in the AgentSpeak model of the BDI framework. In particular, we described the formulation of an *intelligent sensor* for AgentSpeak which applies evidence propagation, performs context-dependent combination and derives beliefs for revising the agent's belief base. Finally, in Section 5 we demonstrated the applicability of our work in terms of a scenario related to a power grid SCADA system. An implementation of the work presented in the paper is available online.[5]

---

[5] `http://www.eeecs.qub.ac.uk/~kmcareavey01/context_dependent.html`

Acknowledgements

References

1. N. Arghira, D. Hossu, I. Fagarasan, S.S. Iliesc, and D.R. Costianu. Modern SCADA philosophy in power system operation - A survey. *UPB Scientific Bulletin, Series C: Electrical Engineering*, 73(2):153–166, 2011.
2. O. Basir and X. Yuan. Engine fault diagnosis based on multi-sensor information fusion using Dempster-Shafer evidence theory. *Information Fusion*, 8(4):379–386, 2007.
3. K. Bauters, W. Liu, J. Hong, C. Sierra, and L. Godo. Can(plan)+: Extending the operational semantics for the BDI architecture to deal with uncertain information. In *Proceedings of the 30th International Conference on Uncertainty in Artificial Intelligence*, pages 52–61, 2014.
4. R.H. Bordini, J.F. Hübner, and M. Wooldridge. *Programming multi-agent systems in AgentSpeak using Jason*, volume 8. John Wiley & Sons, 2007.
5. S.A. Boyer. *SCADA: Supervisory Control and Data Acquisition*. International Society of Automation, 2009.
6. M. Bratman. *Intention, Plans and Practical Reason*. Harvard University Press, 1987.
7. F. Browne, N. Rooney, W. Liu, D. Bell, H. Wang, P. Taylor, and Y. Jin. Integrating textual analysis and evidential reasoning for decision making in engineering design. *Knowledge-Based Systems*, 52:165–175, 2013.
8. S. Calderwood, K. Bauters, W. Liu, and J. Hong. Adaptive uncertain information fusion to enhance plan selection in BDI agent systems. In *Proceedings of the 4th International Workshop on Combinations of Intelligent Methods and Applications*, pages 9–14, 2014.
9. B.R. Cobb and P.P. Shenoy. On the plausibility transformation method for translating belief function models to probability models. *International Journal of Approximate Reasoning*, 41(3):314–330, 2006.
10. A. Daneels and W. Salter. What is SCADA? In *Proceedings of the 7th International Conference on Accelerator and Large Experimental Physics Control Systems*, pages 339–343, 1999.
11. J. Dezert and F. Smarandache. A new probabilistic transformation of belief mass assignment. In *Proceedings of the 11th International Conference on Information Fusion*, pages 1–8, 2008.
12. D. Dubois and H. Prade. Combination of fuzzy information in the framework of possibility theory. *Data Fusion in Robotics and Machine Intelligence*, (12):481–505, 1992.
13. D. Dubois and H. Prade. On the combination of evidence in various mathematical frameworks. In J. Flamm and T. Luisi, editors, *Reliability Data Collection and Analysis*, pages 213–241. Springer, 1992.
14. D. Dubois and H. Prade. Possibility theory and data fusion in poorly informed environments. *Control Engineering Practice*, 2(5):811–823, 1994.
15. D. Dubois, H. Prade, and C. Testemale. Weighted fuzzy pattern matching. *Fuzzy Sets and Systems*, 28(3):313–331, 1988.
16. R. Haenni. Are alternatives to Dempster's rule of combination real alternatives? *Information Fusion*, 3(3):237–239, 2002.
17. A. Hunter and W. Liu. A context-dependent algorithm for merging uncertain information in possibility theory. *IEEE Transactions on System, Man and Cybernetics, Part A*, 38(6):1385–1397, 2008.
18. N. Jennings and S. Bussmann. Agent-based control systems: Why are they suited to engineering complex systems? *IEEE Control Systems*, 23(3):61–74, 2003.
19. A. Jousselme, D. Grenier, and É. Bossé. A new distance between two bodies of evidence. *Information Fusion*, 2(2):91–101, 2001.
20. A. Jousselme and P. Maupin. Distances in evidence theory: Comprehensive survey and generalizations. *International Journal of Approximate Reasoning*, 53(2):118–145, 2012.

21. G.J. Klir and B. Parviz. A note on the measure of discord. In *Proceedings of the 8th International Conference on Uncertainty in Artificial Intelligence*, pages 138–141, 1992.
22. W. Liu. Analyzing the degree of conflict among belief functions. *Artificial Intelligence*, 170(11):909–924, 2006.
23. W. Liu, J. Hughes, and M. McTear. Representing heuristic knowledge in DS theory. In *Proceedings of the 8th International Conference on Uncertainty in Artificial Intelligence*, pages 182–190, 1992.
24. W. Liu, G. Qi, and D.A. Bell. Adaptive merging of prioritized knowledge bases. *Fundamenta Informaticae*, 73(3):389–407, 2006.
25. J. Ma, W. Liu, and P. Miller. Event modelling and reasoning with uncertain information for distributed sensor networks. In *Proceedings of the 4th International Conference on Scalable Uncertainty Management*, pages 236–249, 2010.
26. S.D.J. McArthur, E.M. Davdison, V.M. Catterson, A.L. Dimeas, N.D. Hatziargyriou, F. Ponci, and T. Funabashi. Multi-agent systems for power engineering applications - part i: Concepts, approaches and technical challenges. *IEEE Transactions on Power Systems*, 22(4):1743–1752, 2007.
27. S.D.J. McArthur, E.M. Davidson, V.M. Catterson, A.L. Dimeas, N.D. Hatziargyriou, F. Ponci, and T. Funabashi. Multi-agent systems for power engineering applications - part ii: technologies, standards, and tools for building multi-agent systems. *IEEE Transactions on Power Systems*, 22(4):1753–1759, 2007.
28. R.R. Murphy. Dempster-Shafer theory for sensor fusion in autonomous mobile robots. *IEEE Transactions on Robotics and Automation*, 14(2):197–206, 1998.
29. A.S. Rao. AgentSpeak(L): BDI agents speak out in a logical computable language. In *Proceedings of the 7th European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, pages 42–55, 1996.
30. K. Sentz and S. Ferson. Combination of evidence in Dempster-Shafer theory. `http://prod.sandia.gov/techlib/access-control.cgi/2002/020835.pdf`. Technical Report No. SAND2002-0835, Sandia National Laboratories, 2002.
31. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
32. P. Smets. Decision making in the TBM: the necessity of the pignistic transformation. *International Journal of Approximate Reasoning*, 38(2):133–147, 2005.
33. P. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66(2):191–234, 1994.
34. F. Spellman. *Handbook of Water and Wastewater Treatment Plant Operations*. CRC Press, 2013.
35. M. Wooldridge. *Agent Technology: Foundations, Applications, and Markets*. Springer, 1998.
36. M. Wooldridge. *An Introduction to Multiagent Systems*. Wiley, 2002.
37. L. Zadeh. Review of a mathematical theory of evidence. *The AI Magazine*, 5(3):81–83, 1984.
38. L. Zhi, J.S. Qin, T.B. Yu, Z.J. Hu, and Z. Hao. The study and realization of SCADA system in manufacturing enterprises. In *Proceedings of the 3rd World Congress on Intelligent Control and Automation*, pages 3688–3692, 2000.
39. H. Zhou. *The Internet of Things in the Cloud: A Middleware Perspective*. CRC Press, 2012.