Fusion Rules for Merging Uncertain Information

Anthony Hunter*and Weiru Liu†

January 31, 2005

Abstract

In previous papers, we have presented a logic-based framework based on fusion rules for merging structured news reports [Hun00, Hun02b, Hun02a, HS03, HS04]. Structured news reports are XML documents, where the textentries are restricted to individual words or simple phrases, such as names and domain-specific terminology, and numbers and units. We assume structured news reports do not require natural language processing. Fusion rules are a form of scripting language that define how structured news reports should be merged. The antecedent of a fusion rule is a call to investigate the information in the structured news reports and the background knowledge, and the consequent of a fusion rule is a formula specifying an action to be undertaken to form a merged report. It is expected that a set of fusion rules is defined for any given application. In this paper we extend the approach to handling probability values, degrees of beliefs, or necessity measures associated with textentries in the news reports. We present the formal definition for each of these types of uncertainty and explain how they can be handled using fusion rules. We also discuss the methods of detecting inconsistencies among sources.

1 Introduction

Structured news reports are XML documents, where the textentries are restricted to individual words or simple phrases (such as names and domain-specific terminology), dates, numbers and units. We assume that structured news reports do not require natural language processing. In addition, each tag provides semantic information about the textentries, and a structured news report is intended to have some semantic coherence. To illustrate, news reports on corporate acquisitions can be represented as structured news reports using tags including `buyer`, `seller`, `acquisition`, `value`, and `date`. Structured news reports can be obtained from information extraction systems (e.g. [CL96]).

In order to merge structured news reports, we need to take account of the contents. Different kinds of content need to be merged in different ways. In our approach to merging structured news reports [Hun00, Hun02b, Hun02a, HS03, HS04], we draw on domain knowledge to help produce merged reports. The approach is based on fusion rules defined in a logical language. These rules are of the form $\alpha \Rightarrow \beta$, where if α holds, then β is made to hold. So we consider α as a condition to check the information in the structured reports and in the background information, and we consider β as an action to undertake to construct the merged report. To merge a set of structured news reports, we start with the background knowledge and the information in the news reports to be merged, and apply the fusion rules to this information. The application of the fusion rules is then a monotonic process that builds up a set of actions that define how the structured news reports should be merged. To evaluate the approach, we have undertaken a detailed case study of merging weather reports [HS04].

*Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK
†School of Computer Science, Queen's University Belfast, Belfast, Co Antrim BT7 1NN, UK

In this paper, we extend the approach of fusion rules, by including some forms of uncertainty in the structured news reports, and providing ways of merging the uncertainty using fusion rules.

A number of mechanisms for reasoning under uncertainty have been proposed and studied over the past 20 years. Probability theory, Dempster-Shafer theory of evidence (DS theory) [Dem67, Sha76], and possibility theory [DP88a] are the three popular ones that have been widely applied in many different domains. Probability theory, the most traditional method, is one of the first to be used to represent uncertain information in databases (e.g., [CP87, BGP92, NS94]). The restriction of assigning probabilities only to singleton subsets of the sample space led to the investigation of deploying DS theory (e.g., [Lee92]) where uncertainty can be associated with a set of possible events. A mass function m assigns each unit of an agent's belief to a distinct subset of a set of all possible outcomes, with the total sum of all assignments being 1. When all the distinct subsets are singletons, a mass function is reduced to a probability distribution. It is in this sense, that DS theory can be regarded as a generalization of probability theory, and we adopt this view in this paper. Possibility theory is another option to easily express uncertainty in information. A possibility measure Π on a subset A of a set of possible events gives a value in $[0, 1]$ which estimates to what extent A contains the true event. The dual function, necessity measure N with $N(A) = 1 - \Pi(\bar{A})$ where $\bar{A}$ stands for the complementary set of A, evaluates the degree of necessity that A is true.

There are two possible interpretations about the possibility and necessity measures. One is qualitative and another is quantitative. They share the same mathematical definitions for the two measures, but depart when considering merging methods. "Qualitative possibility theory is closely related to non-monotonic reasoning" [DP98b] and the two measures are referred to as graded measures. "Quantitative possibility theory can be related to probability theory and can be viewed as a special case of belief function theory." [DP98b]. In this paper, we take this latter view of possibility theory, since we will deal with the issues of merging uncertain information that is in the form of both mass functions and necessity measures.

We will present approaches to modelling and merging of uncertain information in the form of probability distributions, mass functions and necessity measures respectively, as well as merging heterogeneous uncertain information where two pieces of uncertain information are modelled in different theories.

Since we restrict our textentries in news reports to discrete variables with a fairly small number of possible values, we will not consider the application of fuzzy logic (or fuzzy set theory) in this paper. Though fuzzy databases have also been studied in various papers (e.g., [BP91, Pet95]) and would be particularly useful when a dataset contains variables with continuous values.

To illustrate our approach we consider weather reports such as in Example 1 and Example 2. These kinds of uncertainty arise in many application areas for structured text such as in information/meta-information in bioinformatics and more generally in the emerging area of e-science. The aim of this paper is to address the problems of merging such uncertain information by harnessing well-established uncertainty formalisms in the fusion rule technology. In these examples, we do include the use of intervals, since they constitute a simple and clear choice for merging pairs of reports. However, because of problems that can arise from the lack of associativity with this form of aggregation, we will not consider the use of intervals in detail in the paper. Rather, we will focus on more sophisticated techniques as indicated above.

Example 1 *Consider the following two conflicting weather reports which are for the same day and same city. The* $\langle$`probability`$\rangle$ *tag demarks a probability distribution over textentries* `8°C` *and* `12°C` *for the*

⟨probability⟩ *tag.*

```
⟨report⟩
  ⟨source⟩ TV1 ⟨/source⟩
  ⟨date⟩ 19/3/02 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨outlook⟩ showers ⟨/outlook⟩
  ⟨windspeed⟩ 1 kph ⟨/windspeed⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨prob value = "0.2"⟩8°C⟨/prob⟩
      ⟨prob value = "0.8"⟩12°C⟨/prob⟩
    ⟨/probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

```
⟨report⟩
  ⟨source⟩ TV3 ⟨/source⟩
  ⟨date⟩ 19 March 2002 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨outlook⟩ inclement ⟨/outlook⟩
  ⟨windspeed⟩ 25 kph ⟨/windspeed⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨prob value = "0.4"⟩8°C⟨/prob⟩
      ⟨prob value = "0.6"⟩12°C⟨/prob⟩
    ⟨/probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

We can merge these reports so the source *is* TV1 and TV3*, and* outlook *is* showers and inclement*. Clearly there is a contradiction with regard to* windspeed*. Possibilities for merging include taking the textentry for the prefered source or using the two values as an interval. Also for merging the probabilistic information, we have choices. One possibility is to give intervals. Another possibility is to give a merged distribution - either by taking an average or by using Dempster's rule of combination if we believe the sources are independent. We assume in this and the rest of the examples that both sources use the commonly agreed sets of elements as possible textentries. For instance, for tag* temperature*, the possible elements are integers in* $[-30°C, 40°C]$*. Below we give two possible merged reports that we could obtain for the above input. The report below left has been obtained using Dempster's rule and below right has been obtained using intervals.* [1]

```
⟨report⟩
  ⟨source⟩ TV1 and TV3 ⟨/source⟩
  ⟨date⟩ 19/03/02 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨outlook⟩ showers and inclement⟨/outlook⟩
  ⟨windspeed⟩ 1 − 25 kph ⟨/windspeed⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨prob value = "0.14"⟩8°C⟨/prob⟩
      ⟨prob value = "0.86"⟩12°C⟨/prob⟩
    ⟨probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

```
⟨report⟩
  ⟨source⟩ TV1 and TV3 ⟨/source⟩
  ⟨date⟩ 19/03/02 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨outlook⟩ showers and inclement ⟨/outlook⟩
  ⟨windspeed⟩ 1 − 25 kph ⟨/windspeed⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨probinterval value = "0.2 − 0.4"⟩8°C⟨/prob⟩
      ⟨probinterval value = "0.6 − 0.8"⟩12°C⟨/prob⟩
    ⟨probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

Example 2 *Consider the following two conflicting weather reports which are for the same day and same city. The left report is certain that the temperature is* 8°C*, and the right report gives a probability distribu-*

[1] A brief review on DS theory and Dempster's combination rule is given in Section 2.2. Here we just illustrate how the two pieces of uncertain information given in the two XML documents are actually combined to generate the merged uncertain information on the left.

Each probabilistic piece of information above can be modelled in terms of a mass function in DS theory

$$m_1(\mathsf{Temp} = 8°C) = 0.2, m_1(\mathsf{Temp} = 12°C) = 0.8; m_2(\mathsf{Temp} = 8°C) = 0.4, m_2(\mathsf{Temp} = 12°C) = 0.6$$

After applying Dempster's rule, we have

$$m(\mathsf{Temp} = 8°C) = (0.2 \times 0.4)/0.56 = 0.14, m(\mathsf{Temp} = 12°C) = (0.8 \times 0.6)/0.56 = 0.86$$

tion over the options of 8°C *and* 12°C*.*

```
⟨report⟩
  ⟨source⟩ TV1 ⟨/source⟩
  ⟨date⟩ 19/3/02 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨temperature⟩ 12°C ⟨/temperature⟩
⟨/report⟩
```

```
⟨report⟩
  ⟨source⟩ TV3 ⟨/source⟩
  ⟨date⟩ 19 March 2002 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨prob value = “0.4”⟩8°C⟨/prob⟩
      ⟨prob value = “0.6”⟩12°C⟨/prob⟩
    ⟨/probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

Two possible ways of merging these two reports are given below left and right. The left report is obtained as an average distribution from the input reports, and the right report is given as an interval.

```
⟨report⟩
  ⟨source⟩ TV1 ⟨/source⟩
  ⟨date⟩ 19/3/02 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨prob value = “0.2”⟩8°C⟨/prob⟩
      ⟨prob value = “0.8”⟩12°C⟨/prob⟩
    ⟨/probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

```
⟨report⟩
  ⟨source⟩ TV3 ⟨/source⟩
  ⟨date⟩ 19 March 2002 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨probinterval value = “0.0 − 0.4”⟩8°C⟨/prob⟩
      ⟨probinterval value = “0.6 − 1.0”⟩12°C⟨/prob⟩
    ⟨/probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

Example 3 *Consider the following two conflicting weather reports which are for the same day and same city. Note in the right example, the child of the* ⟨prob value = “...”⟩ *tag is not a textentry. Whilst both fit into our definition for structured news reports, we will restrict consideration in this paper to the structured news reports where the child of the* ⟨prob value = “...”⟩ *tag is a textentry (as in the left but not the right below), e.g., the given probability values are about elementary events. Situations such as on the right in have been discussed in [HL04a, HL04b].*

```
⟨report⟩
  ⟨date⟩ 19/3/02 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨outlook⟩
    ⟨probability⟩
      ⟨prob value = “0.2”⟩sun⟨/prob⟩
      ⟨prob value = “0.8”⟩snow⟨/prob⟩
    ⟨/probability⟩
  ⟨/outlook⟩
  ⟨temperature⟩
    ⟨probability⟩
      ⟨prob value = “0.3”⟩0°C⟨/prob⟩
      ⟨prob value = “0.7”⟩2°C⟨/prob⟩
    ⟨/probability⟩
  ⟨/temperature⟩
⟨/report⟩
```

```
⟨report⟩
  ⟨date⟩ 19 March 2002 ⟨/date⟩
  ⟨city⟩ London ⟨/city⟩
  ⟨probability⟩
    ⟨prob value = “0.4”⟩
      ⟨outlook⟩sun⟨/outlook⟩
      ⟨temperature⟩3°C⟨/temperature⟩
    ⟨/prob⟩
    ⟨prob value“0.6”⟩
      ⟨outlook⟩snow⟨/outlook⟩
      ⟨temperature⟩0°C⟨/temperature⟩
    ⟨/prob⟩
  ⟨/probability⟩
⟨/report⟩
```

We will proceed as follows: in Section 2, we present formal definitions of uncertain information in structured news reports and their constraints. In Section 3, we define fusion rules, and explain how they are executed to generate merged reports in general, and then in Section 4, we consider merging uncertain information in detail. We will focus on merging probability, possibility and mass distributions over textentries. We will also discuss the detection of inconsistency among multiple news reports. In Section 5 we provide methods to merge multiple heterogeneous uncertainty components in news reports. Finally, we compare our work with related research in the final section.

2 Structured news reports

We now provide basic definitions, in Section 2.1, for structured news reports. In Section 2.2, we review the rudiments of DS theory and possibility theory. Then in Section 2.3, we show how to represent uncertain information in structured news reports.

2.1 Basic definition

We use XML to represent structured news reports. So each structured news report is an XML document, but not vice versa, as defined below. This restriction means that we can easily represent each structured news report by a ground term in classical logic.

Definition 1 Structured news report: *If φ is a tagname (i.e an element name), and ϕ is textentry, then $\langle\varphi\rangle\phi\langle/\varphi\rangle$ is a structured news report. If φ is a tagname (i.e an element name), ϕ is textentry, θ is an attribute name, and κ is an attribute value, then $\langle\varphi\ \theta = \kappa\rangle\phi\langle/\varphi\rangle$ is a structured news report. If φ is an tagname and $\sigma_1, .., \sigma_n$ are structured news reports, then $\langle\varphi\rangle\sigma_1..\sigma_n\langle/\varphi\rangle$ is a structured news report.*

The definition for a structured news report is very general. In practice, we would expect a DTD for a given domain. So for example, we would expect that for an implemented system that merges weather reports, there would be a corresponding DTD. One of the roles of a DTD, say for weather reports, would be to specify the minimum constellation of tags that would be expected of a weather report. We may also expect integrity constraints represented in classical logic to further restrict appropriate structured news reports for a domain. We will not consider these issues further in this paper. However, in this paper, we will impose some further constraints on structured news reports, in Section 2.3, to support the handling of uncertainty.

Clearly each structured news report is isomorphic to a tree with the non-leaf nodes being the tagnames and the leaf nodes being the textentries. This isomorphism allows us to give a definition for a branch of a structured news report.

Definition 2 Branch: *Let σ be a structured news report and let ρ be a tree that is isomorphic to σ. A sequence of tagnames $\varphi_1/../\varphi_n$ is a branch of ρ iff (1) φ_1 is the root of ρ and (2) for each i, if $1 \leq i < n$, then φ_i is the parent of φ_{i+1}. Note, the child of φ_n is not necessarily a leaf node. By extension, $\varphi_1/../\varphi_n$ is a branch of σ iff $\varphi_1/../\varphi_n$ is a branch of ρ*

When we refer to a subtree (of a structured news report), we mean a subtree formed from the tree representation of the structured news report, where the root of the subtree is a tagname and the leaves are textentries. We formalize this as follows.

Definition 3 Subtree: *Let σ be a structured news report and let ρ be a tree that is isomorphic to σ. A tree ρ' is a subtree of ρ iff (1) the set of nodes in ρ' is a subset of the set of nodes in ρ, and (2) for each node φ_i in ρ', if φ_i is the parent of φ_j in ρ, then φ_j is in ρ' and φ_i is the parent of φ_j in ρ'. By extension, if σ' is a structured news report, and ρ' is isomorphic to σ', then we describe σ' as a subtree of σ.*

Each structured news report is also isomorphic with a ground term (of classical logic) where each tagname is a function symbol and each textentry is a constant symbol.

Definition 4 News term: *Each structured news report is isomorphic with a ground term (of classical logic) called a news term. This isomorphism is defined inductively as follows: (1) If $\langle\varphi\rangle\phi\langle/\varphi\rangle$ is a structured news report, where ϕ is a textentry, then $\varphi(\phi)$ is a news term that is isomorphic with $\langle\varphi\rangle\phi\langle/\varphi\rangle$; (2) If $\langle\varphi\ \theta = \kappa\rangle\phi\langle/\varphi\rangle$ is a structured news report, where ϕ is a textentry, then $\varphi(\phi, \kappa)$ is a news term that is isomorphic with $\langle\varphi\ \theta = \kappa\rangle\phi\langle/\varphi\rangle$; and (3) If $\langle\varphi\rangle\phi_1..\phi_n\langle/\varphi\rangle$ is a structured news report, and ϕ'_1 is a news term that is isomorphic with ϕ_1,...., and ϕ'_n is a news term that is isomorphic with ϕ_n, then $\varphi(\phi'_1, .., \phi'_n)$ is a news term that is isomorphic with $\langle\varphi\rangle\phi_1..\phi_n\langle/\varphi\rangle$.*

Via this isomorphic relationship, we can refer to a branch of a news term by using the branch of the isomorphic structured news report, and we can refer to a subtree of a news term by using the subtree of the isomorphic structured news report. Next we define two functions that allow us to obtain subtrees and textentries from news terms.

Definition 5 *Let τ be a news term, let τ' be a subtree of τ, and let ϕ be a textentry. If $\varphi_1/../\varphi_n$ is a branch of τ, and the root of τ' is φ_n, then let $\mathsf{Subtree}(\varphi_1/../\varphi_n, \tau) = \tau'$, otherwise let $\mathsf{Subtree}(\varphi_1/../\varphi_n, \tau) = \mathsf{null}$. If $\varphi_1/../\varphi_n$ is a branch of τ, and ϕ is the child of φ_n, then let $\mathsf{Textentry}(\varphi_1/../\varphi_n, \tau) = \phi$, otherwise $\mathsf{Textentry}(\varphi_1/../\varphi_n, \tau) = \mathsf{null}$.*

Example 4 *Consider the following structured news report.*

$$\begin{array}{l}
\langle\mathtt{fieldreport}\rangle \\
\quad\quad \langle\mathtt{log}\rangle\langle\mathtt{station}\rangle\mathtt{Inverness}\langle/\mathtt{station}\rangle\langle\mathtt{date}\rangle\mathtt{12/3/03}\langle/\mathtt{date}\rangle\langle/\mathtt{log}\rangle \\
\quad\quad \langle\mathtt{rainfall}\rangle\mathtt{2.3cm}\langle/\mathtt{rainfall}\rangle \\
\langle/\mathtt{fieldreport}\rangle
\end{array}$$

This can be represented by the following news term:

$$\mathtt{fieldreport}(\mathtt{log}(\mathtt{station}(\mathtt{Inverness}), \mathtt{date}(\mathtt{12/3/02})), \mathtt{rainfall}(\mathtt{2.3cm}))$$

In this news term, $\mathtt{fieldreport/log/station}$ is a branch. If the news term is denoted by τ, we have

$$\begin{array}{r}
\mathsf{Subtree}(\mathtt{fieldreport/log}, \tau) = \mathtt{log}(\mathtt{station}(\mathtt{Inverness}), \mathtt{date}(\mathtt{12/3/03})) \\
\mathsf{Subtree}(\mathtt{fieldreport/log/station}, \tau) = \mathtt{station}(\mathtt{Inverness}) \\
\mathsf{Textentry}(\mathtt{fieldreport/log/station}, \tau) = \mathtt{Inverness}
\end{array}$$

Example 5 *Consider the structured news report in Example 2 (right). Let the isomorphic news term be represented by τ.*

$$\begin{array}{l}
\mathsf{Subtree}(\mathtt{report/precipitation/probability}, \tau) \\
\quad\quad = \mathtt{probability}(\mathtt{prob}(8°\mathtt{C}, 0.4), \mathtt{prob}(12°\mathtt{C}, 0.6))
\end{array}$$

Definition 6 *A* **skeleton** *is of the form $\varphi(\phi_1, .., \phi_n)$ where φ is a tagname and $\phi_1, .., \phi_n$ are each a skeleton. If ϕ_i is a textentry, then it is a skeleton. A skeleton $\varphi(\phi_1, .., \phi_n)$ can be regarded as a tree.*

A skeleton is a equivalent to a structured news report without textentries. It is the underlying structure without the content.

2.2 Brief review of DS theory and possibility theory

2.2.1 Dempster-Shafer theory

Dempster-Shafer theory of evidence offers some significant advantages over probability theory when the degree of uncertainty is not assigned to the singleton set. Furthermore, we can define condition literals for which associativity does hold, and therefore the sequence in which we merge the distributions does not matter.

Let Ω be a finite set containing mutually exclusive and exhaustive solutions to a question. Ω is called the **frame of discernment**.

A **mass function**, also called a **basic probability assignment**, captures the impact of a piece of evidence on subsets of Ω. The empty set, i.e. $\emptyset$ always has 0.0 mass value. Since one of the elements in the frame of discernment has to be true, the total mass value for the frame is 1.0. This is summarised as follows: A mass function m is a function from $\wp(\Omega)$ into $[0, 1]$ such that

$$(1)\ m(\emptyset) = 0$$

$$(2)\ \Sigma_{A \subseteq \Omega}\ m(A) = 1$$

$m(A)$ defines the amount of belief to the subset A exactly, not including any subsets in A. When $m(A) > 0$, A is referred to as a **focal element**. To obtain the total belief in a subset A, i.e. the extent to which all available evidence supports A, we need to sum all the mass assigned to all subsets of A.

For a given mass function m, if each of its focal element A contains only a single element, this mass function is reduced to be a probability distribution. It is with this view that in this paper, we take probability distributions as special cases of mass functions.

A **belief function**, denoted Bel, is defined as follows, where $Bel : \wp(\Omega) \rightarrow [0, 1]$.

$$Bel(A) = \Sigma_{B \subseteq A} m(B)$$

A **plausibility function**, denoted Pl, is defined as follows, where $Pl : \wp(\Omega) \rightarrow [0, 1]$.

$$Pl(A) = 1 - Bel(\bar{A}) = \Sigma_{B \cap A \neq \emptyset}\ m(B)$$

Now we consider Dempster's rule of combination. For this we suppose we have two mass functions m_1 and m_2 from independent sources, and we want to combine them into a combined mass function.

Definition 7 *Let m_1 and m_2 be two mass functions, and let $m_1 \oplus m_2$ be the combined mass function.*

$$m_1 \oplus m_2(C) = \frac{\Sigma_{A \cap B = C}\ (m_1(A) \times m_2(B))}{1 - \Sigma_{A \cap B = \emptyset}\ (m_1(A) \times m_2(B))}$$

When using Dempster's rule, $1 - \Sigma_{A \cap B = \emptyset}\ (m_1(A) \times m_2(B))$ is the normalization coefficient obtained by deducting the mass assigned to the empty set.

2.2.2 Possibility theory

Possibility theory is another useful choice for representing uncertain information ([DP88a, SDK95, BDP97], etc), especially when an agent's knowledge is not sufficient to provide either a probabilistic or a mass assignment.

A possibility measure, a value in $[0,1]$, assigned to a subset of a set of possible solutions, estimates to what extent the true event is believed to be in the subset and a necessity measure, a value in $[0,1]$, evaluates the degree of necessity that the subset is true. Numerical values (or values in a pre-defined partial ordered set) are used to represent levels (grades) of beliefs in propositions.

Even though numerical values are used in possibility theory to express the degrees of possibilities and necessities, the calculation of possibility and necessity measures are not the usual arithmetic calculations. The two main operators used in the theory are comparison functions, `max` and `min`, which obtain either the maximum or minimum values among a set of possibility or necessity measures. Because of this, possibility theory is also described as a "quasi-qualitative" calculus.

Definition 8 *Let Ω be a frame of discernment. A possibility measure and a necessity measure, denoted Π and N respectively, are functions from $\wp(\Omega)$ to $[0,1]$ such that given any two subsets A and B of $\wp(\Omega)$, the following equations hold where $\bar{A}$ is the complementary set of A.*

$$\begin{aligned}
\Pi(\wp(\Omega)) &= 1, \\
\Pi(\emptyset) &= 0, \\
\Pi(A \cup B) &= \mathsf{max}(\Pi(A), \Pi(B)), \\
N(A \cap B) &= \mathsf{min}(N(A), N(B)), \\
\mathsf{max}(\Pi(A), \Pi(\bar{A})) &= 1, \\
\mathsf{min}(N(A), N(\bar{A})) &= 0, \\
N(A) &= 1 - \Pi(\bar{A})
\end{aligned}$$

Both possibility measure and necessity measure can be derived from a more elementary assignment, $\pi : \Omega \rightarrow [0,1]$, which is referred to as a **possibility distribution**. The relationship between Π and π is

$$\Pi(A) = \mathsf{max}(\{\pi(\omega) | \omega \in A\})$$

which satisfies $\Pi(A \cup B) = \mathsf{max}(\Pi(A), \Pi(B))$. The usual condition associated with π is there exists ω_0 such that $\pi(\omega_0) = 1$, and in which case π is said to be normal.

2.3 Representing uncertain information

In order to support the representation of uncertain information in structured news reports, we need some further formalization. First, we assume a set of tagnames that are reserved for representing uncertain information. Second, we assume some constraints on the use of these tags so that we can ensure they are used in a meaningful way with respect to established uncertainty formalisms including probability theory, possibility theory and Dempster-Shafer theory of evidence.

Definition 9 *The set of* **key uncertainty tagnames** *for this paper are* `probability`, `possibility`, *and* `belfunction`. *The set of* **subsidiary uncertainty tagnames** *for this paper are* `prob`, `ness`, `nessitem`, `mass`, *and* `massitem`. *The union of the key uncertainty tagnames and the subsidiary uncertainty tagnames is the set of* **reserved tagnames**.

Definition 10 *The structured news report $\langle\texttt{probability}\rangle\sigma_1,..,\sigma_n\langle/\texttt{probability}\rangle$ is* **probability-valid** *iff each $\sigma_i \in \{\sigma_1,..,\sigma_n\}$ is of the form $\langle\texttt{prob value} = \kappa\rangle\phi\langle/\texttt{prob}\rangle$ where $\kappa \in [0,1]$ and ϕ is a textentry.*

All textentries ϕ_i between $\langle\texttt{prob value}\rangle\phi_i\langle/\texttt{prob}\rangle$ are elements of a pre-defined set (in the background knowledge base) containing mutually exclusive and exhaustive values that the related tagname can take.

Example 6 *The following is a probability-valid news report.*

```
⟨probability⟩
  ⟨prob value = "0.2"⟩8°C⟨/prob⟩
  ⟨prob value = "0.8"⟩12°C⟨/prob⟩
⟨/probability⟩
```

Definition 11 *The structured news report* $\langle\texttt{possibility}\rangle\sigma_1,..,\sigma_n\langle/\texttt{possibility}\rangle$ *is* **possibility- valid** *iff for each* $\sigma_i \in \{\sigma_1,..,\sigma_n\}$*,* σ_i *is of the form* $\langle\texttt{ness value} = \kappa\rangle\sigma_1^i,..,\sigma_m^i\langle/\texttt{ness}\rangle$ *and for each* $\sigma_j^i \in \{\sigma_1^i,..,\sigma_n^i\}$*,* σ_j^i *is of the form* $\langle\texttt{nessitem}\rangle\phi\langle/\texttt{nessitem}\rangle$ *and* $\kappa \in [0,1]$*, and* ϕ *is a textentry.*

Example 7 *The following is a possibility-valid structured news report.*

```
⟨possibility⟩
  ⟨ness value = "0.5"⟩
    ⟨nessitem⟩8°C⟨/nessitem⟩
    ⟨nessitem⟩10°C⟨/nessitem⟩
  ⟨/ness⟩
  ⟨ness value = "0.8"⟩
    ⟨nessitem⟩12°C⟨/nessitem⟩
  ⟨/ness⟩
⟨/possibility⟩
```

In possibility theory, both a possibility measure (Π) and a necessity measure (N) can be assigned to subsets of a set of possible values. In possibilistic logic (e.g. [BDP97, BDKP00]) a weighted formula (ϕ, a) implies that the weight a attached to formula ϕ is interpreted as a lower bound on the degree of necessity $N(\phi)$. In the context of this paper, a weight κ_i attached to a subset $\{\phi_l^1,..,\phi_l^r\}$ is equally interpreted as a lower bound on the degree of necessity that $\{\phi_l^1,..,\phi_l^r\}$ is true. This also explains why we use tagname "ness" instead of "poss".

Definition 12 *The structured news report* $\langle\texttt{belfunction}\rangle\sigma_1,..,\sigma_n\langle/\texttt{belfunction}\rangle$ *is* **belfunction-valid** *iff for each* $\sigma_i \in \{\sigma_1,..,\sigma_n\}$ σ_i *is of the form* $\langle\texttt{mass value} = \kappa\rangle\sigma_1^i,..,\sigma_m^i\langle/\texttt{mass}\rangle$ *and for each* $\sigma_j^i \in \{\sigma_1^i,..,\sigma_n^i\}$*,* σ_j^i *is of the form* $\langle\texttt{massitem}\rangle\phi\langle/\texttt{massitem}\rangle$ *and* $\kappa \in [0,1]$ *and* ϕ *is a textentry.*

The textentries in either a possibility-valid component or a belfunction-valid component are elements of a pre-defined set containing mutually exclusive and exhaustive values for the related tagname as in the case for probability-valid components.

Example 8 *The following is a belfunction-valid structured news report.*

```
⟨belfunction⟩
  ⟨mass value = "0.2"⟩
    ⟨massitem⟩8°C⟨/massitem⟩
    ⟨massitem⟩10°C⟨/massitem⟩
  ⟨/mass⟩
  ⟨mass value = "0.8"⟩
    ⟨massitem⟩12°C⟨/massitem⟩
  ⟨/mass⟩
⟨/belfunction⟩
```

The probability-valid, possibility-valid, and belfunction- valid structured news reports are normally part of larger structured news reports. To describe these larger structured news reports, we use the following definition of uncertainty-valid.

Definition 13 *A structured news report* $\langle\varphi\rangle\sigma_1...\sigma_n\langle/\varphi\rangle$ *is* **uncertainty-valid** *iff one of the following holds for it.*

1. φ *is not a reserved tagname and for all* $\sigma_i \in \{\sigma_1, ..., \sigma_n\}$ σ_i *is uncertainty-valid.*
2. φ *is not a reserved tagname and* $n = 1$ *and* σ_1 *is a textentry.*
3. φ *is the key uncertainty tagname* `probability` *and* $\langle\varphi\rangle\sigma_1...\sigma_n\langle/\varphi\rangle$ *is probability-valid.*
4. φ *is the key uncertainty tagname* `possibility` *and* $\langle\varphi\rangle\sigma_1...\sigma_n\langle/\varphi\rangle$ *is possibility-valid.*
5. φ *is the key uncertainty tagname* `belfunction` *and* $\langle\varphi\rangle\sigma_1...\sigma_n\langle/\varphi\rangle$ *is belfunction-valid.*

Normally, we would expect that for an application, the DTD for the structured news report would exclude a key uncertainty tag as the root of the overall structured news report. In other words, the key uncertainty tags are roots of subtrees nested within larger structured news reports.

Definition 14 *Let* σ *be a structured news report. If* σ *is probability-valid, possibility-valid, or belfunction-valid, then* σ *is an* **uncertainty component***.*

We assume various integrity constraints on the use of the uncertainty components.

Definition 15 *Let* $\langle\texttt{probability}\rangle\sigma_1, .., \sigma_n\langle/\texttt{probability}\rangle$ *be an uncertainty component, and let* $\sigma_i \in \{\sigma_1, .., \sigma_n\}$ *be of the form* $\langle\texttt{prob value} = \kappa_i\rangle\phi_i\langle/prob\rangle$ *This uncertainty component adheres to the* **full probability distribution constraint** *iff the following two conditions hold:*

(1) $\Sigma_i\kappa_i = 1$
(2) *for all* $i, j,$ *if* $1 \leq i \leq n$ *and* $1 \leq j \leq n$ *and* $i \neq j,$ *then* $\phi_i \neq \phi_j$

Definition 16 *Let* $\langle\texttt{possibility}\rangle\sigma_1, .., \sigma_n\langle/\texttt{possibility}\rangle$ *be an uncertainty component, and let* $\sigma_i \in \{\sigma_1, .., \sigma_n\}$ *be of the form* $\langle\texttt{ness value} = \kappa_i\rangle\psi_i^1, .., \psi_i^p\langle/\texttt{ness}\rangle$*, and let* ψ_i^k *be of the form* $\langle\texttt{nessitem}\rangle\varphi_i^k\langle/\texttt{nessitem}\rangle$ *for* $1 \leq k \leq p$*. This uncertainty component adheres to the* **necessity measure constraint** *in possibility theory iff the following conditions hold:*

(1) $\kappa_i \in [0, 1]$
(2) *for all* $i, j,$ *if* $1 \leq i \leq n$ *and* $1 \leq j \leq n$ *and* $i \neq j,$ *then* $\{\phi_i^1, .., \phi_i^p\} \neq \{\phi_j^1, .., \phi_j^q\}$

In possibility theory, it is perfectly acceptable that a possibilistic knowledge base can have both (ϕ, a_1) and (ϕ, a_2) where $a_1 \neq a_2$ are two necessity measures on the same logical sentence. In this case, (ϕ, a_1) subsumes (ϕ, a_2) when $a_1 > a_2$. However, we restrict our discussion in the paper so that for each subset, there is only one necessity degree in XML. This will reduce unnecessary XML segments in XML documents.

Definition 17 *Let* $\langle\texttt{belfunction}\rangle\sigma_1, .., \sigma_n\langle/\texttt{belfunction}\rangle$ *be an uncertainty component, let* $\sigma_i \in \{\sigma_1, .., \sigma_n\}$ *be of the form* $\langle\texttt{mass value} = \kappa_i\rangle\psi_i^1, .., \psi_i^p\langle/\texttt{mass}\rangle$*, and let* ψ_i^k *be of the form* $\langle\texttt{massitem}\rangle\phi_i^k\langle/\texttt{massitem}\rangle$ *for* $1 \leq k \leq p$*. This uncertainty component adheres to the* **full belfunction distribution constraint** *iff the following two conditions hold:*

(1) $\Sigma_i\kappa_i = 1$
(2) *for all* $i, j,$ *if* $1 \leq i \leq n$ *and* $1 \leq j \leq n$ *and* $i \neq j,$ *then* $\{\phi_i^1, .., \phi_i^p\} \neq \{\phi_j^1, .., \phi_j^q\}$

We return to specific issues of merging uncertain information in Section 4.

3 Fusion rules

In this paper we use $\{\mathtt{in_1}, .., \mathtt{in_n}\}$ as the set of report names. When a set of reports is given as input, we need a process of **registration** to assign each report with a report name. In this paper, we make an arbitrary assignment. Once a structured news report has been registered, we can refer to it by name in the fusion rules.

Definition 18 *Let $\varphi_1/../\varphi_n$ be a branch, and let μ be a report name. A* **subtree variable** *is denoted $\mu//\varphi_1/../\varphi_n$, and a* **textentry variable** *is denoted $\mu//\varphi_1/../\varphi_n\#$. A* **schema variable** *is either a subtree variable or a textentry variable. Let the set of schema variables be denoted $\mathcal{S}$.*

In the following definition, we augment the usual definition for a classical logic language with notation for schema variables which are just placeholders to be instantiated with news terms and textentries before logical reasoning. In effect they provide the input for a fusion rule. Logical variables are the other kind of variables that we use in fusion rules. They are just the usual classical variables. As we explain below, after we ground the schema variables, each literal in the antecedent of a fusion rule is a condition to be evaluated with respect to a Prolog knowledgebase, and the logical variable(s) are ground by the Prolog knowledgebase if the condition can be evaluated to "true".

Definition 19 **Atoms:** *We assume the following sets of symbols: (1) $\mathcal{C}$ is a set of constant symbols; (2) $\mathcal{V}$ is a set of variable symbols; (3) $\mathcal{S}$ is a set of schema variables; (4) $\mathcal{F}$ is a set of function symbols; and (5) $\mathcal{P}$ is a set of predicate symbols.*

1. *The set of simple terms $\mathcal{T}_1$ is $\mathcal{C} \cup \{f(d_1, .., d_k) \mid f \in \mathcal{F}$ and $d_1, .., d_k \in \mathcal{C} \cup \mathcal{V} \cup \mathcal{S}\}$.*
2. *The set of schemafree terms $\mathcal{T}_2$ is $\mathcal{C} \cup \{f(d_1, .., d_k) \mid f \in \mathcal{F}$ and $d_1, .., d_k \in \mathcal{C} \cup \mathcal{V}\}$.*
3. *The set of ground terms $\mathcal{T}_3$ is $\mathcal{C} \cup \{f(c_1, .., c_k) \mid f \in \mathcal{F}$ and $c_1, .., c_k \in \mathcal{C}\}$.*
4. *The set of simple atoms $\mathcal{A}_1$ is $\{p(t_1, .., t_k) \mid p \in \mathcal{P}$ and $t_1, .., t_k \in \mathcal{T}_1\}$.*
5. *The set of schemafree atoms $\mathcal{A}_2$ is $\{p(t_1, .., t_k) \mid p \in \mathcal{P}$ and $t_1, .., t_k \in \mathcal{T}_2\}$.*
6. *The set of ground atoms $\mathcal{A}_3$ is $\{p(t_1, .., t_k) \mid p \in \mathcal{P}$ and $t_1, .., t_k \in \mathcal{T}_3\}$.*

We assume that if $\varphi_1/../\varphi_n$ is a branch, then it is a constant symbol in $\mathcal{C}$. Similarly, if $\varphi(\phi_1, .., \phi_n)$ is a skeleton, then it is a constant symbol in $\mathcal{C}$. We also assume that if α is an atom, then α is a literal, and $\neg\alpha$ is a literal. Let $\mathcal{L}_1$ be the literals formed from $\mathcal{A}_1$, $\mathcal{L}_2$ be the literals from $\mathcal{A}_2$, and $\mathcal{L}_3$ be the literals from $\mathcal{A}_3$.

Definition 20 *A* **fusion rule** *is of the following form where $\alpha_1, .., \alpha_n \in \mathcal{L}_1$ and $\beta \in \mathcal{A}_1$.*

$$\alpha_1 \wedge .. \wedge \alpha_n \Rightarrow \beta$$

We call $\alpha_1, .., \alpha_n$ the condition literals and β the action atom.

We regard a fusion rule that incorporates schema variables, as a scheme for one or more classical formulae. These formulae are obtained by grounding schema variables as we explain below. We discuss condition literals in Section 3.1 and action literals in Section 3.2.

Example 9 *The following is a fusion rule that could be used for Example 1 where* X *is a logical variable.*

$$\neg\texttt{Synonymous}(\texttt{in}_1//\texttt{report/outlook}\#, \texttt{in}_2//\texttt{report/outlook}\#)$$
$$\wedge\ \texttt{Coherent}(\texttt{in}_1//\texttt{report/outlook}\#, \texttt{in}_2//\texttt{report/outlook}\#)$$
$$\wedge\ \texttt{Conjunction}(\texttt{in}_1//\texttt{report/outlook}\#, \texttt{in}_2//\texttt{report/outlook}\#, \texttt{X})$$
$$\Rightarrow \texttt{AddText}(\texttt{X}, \texttt{report/outlook})$$

Example 10 *The following is a fusion rule that could be used for Example 1 in the rare case where two distributions are identical. So the condition literal holds when the two distributions obtained from* $\texttt{in}_1$ *and* $\texttt{in}_2$ *are the same.*

$$\texttt{SameDistribution}(\texttt{in}_1//\texttt{report/temperature/probability},$$
$$\texttt{in}_2//\texttt{report/temperature/probability})$$
$$\Rightarrow \texttt{AddTree}(\texttt{in}_1//\texttt{report/temperature/probability}, \texttt{report/temperature})$$

Definition 21 *Let* Φ *be a set of structured news reports to be merged.* ϖ *is a* **valid grounding for a subtree variable** $\mu//\varphi_1/../\varphi_k$ *iff* $\mathsf{Subtree}(\varphi_1/../\varphi_k, \mu) = \varpi$. ϖ *is a* **valid grounding for a textentry variable** $\mu//\varphi_1/../\varphi_k\#$ *iff* $\mathsf{Textentry}(\varphi_1/../\varphi_k, \mu) = \varpi$.

Example 11 *Consider the rule in Example 9. For* $\texttt{in}_1//\texttt{report/outlook}\#$*, the valid grounding is given by* $\mathsf{Textentry}(\texttt{report/outlook}, \texttt{in}_1)$*. This is evaluated to* showers *if we let* $\texttt{in}_1$ *refer to the top left structured news report in Example 1. Similarly, the valid grounding for variable* $\texttt{in}_2//\texttt{report/outlook}\#$ *is* $\mathsf{Textentry}(\texttt{report/outlook}, \texttt{in}_2)$*, which is evaluated to* inclement *if we let* $\texttt{in}_2$ *refer to the top right structured news report in Example 1.*

Definition 22 *A* **schemafree fusion rule** *is a fusion rule with every schema variable replaced by a valid grounding.*

Example 12 *The schemafree fusion rule obtained with the fusion rule given in Example 9 and the news reports in Example 1 is the following where* X *is a logical variable.*

$$\neg\texttt{Synonymous}(\texttt{showers}, \texttt{inclement})$$
$$\wedge\ \texttt{Coherent}(\texttt{showers}, \texttt{inclement})$$
$$\wedge\ \texttt{Conjunction}(\texttt{showers}, \texttt{inclement}, \texttt{X})$$
$$\Rightarrow \texttt{AddText}(\texttt{X}, \texttt{report/outlook})$$

As we discuss next, fusion rules provide a bridge between structured news reports and logical reasoning with background knowledge.

3.1 Condition literals

The condition literals in fusion rules relate the contents of structured news reports to the background knowledge. There are many possible condition literals that we could define that relate one or more features from one or more structured news reports to the background knowledge.

To illustrate, these literals may include the following kinds: $\texttt{SameDate}(\texttt{T}, \texttt{T}')$ where T and $\texttt{T}'$ are news terms with equal date; $\texttt{SameSource}(\texttt{T}, \texttt{T}')$ where T and $\texttt{T}'$ are news terms that refer to the same source; $\texttt{SameCity}(\texttt{T}, \texttt{T}')$ where T and $\texttt{T}'$ are news terms that refer to the same city; $\texttt{Synonymous}(\texttt{T}, \texttt{T}')$ where T and $\texttt{T}'$ are news terms that are synonyms; and $\texttt{Coherent}(\texttt{T}, \texttt{T}')$ where T and $\texttt{T}'$ are news terms that are coherent.

Condition literal with logical variable	Instantiation of the logical query variable
$\texttt{Interval}(\texttt{18C}, \texttt{25C}, \texttt{X})$	$\texttt{18} - \texttt{25C}$
$\texttt{Conjunction}(\texttt{TV1}, \texttt{TV3}, \texttt{X})$	TV1 and TV3
$\texttt{Disjunction}(\texttt{sun}, \texttt{rain}, \texttt{X})$	sun or rain

Table 1: For condition literals, the grounding for the logical variable X is given on the right

Example 13 *Some examples of schemafree condition literals may include the following.*

$$\texttt{SameDate}(\texttt{date}(\texttt{14Nov01}), \texttt{date}(\texttt{14.11.01}))$$
$$\texttt{SameDate}(\texttt{date}(\texttt{day}(\texttt{14}), \texttt{month}(\texttt{11}), \texttt{year}(\texttt{01})), \texttt{date}(\texttt{14.11.01}))$$
$$\texttt{SameCity}(\texttt{city}(\texttt{Mumbai}), \texttt{city}(\texttt{Bombay}))$$
$$\texttt{Coherent}(\texttt{snow}, \texttt{sleet})$$
$$\texttt{Coherent}(\texttt{sun}, \texttt{sunny})$$
$$\texttt{Coherent}(\texttt{showers}, \texttt{inclement})$$
$$\neg\texttt{Coherent}(\texttt{sun}, \texttt{rain})$$
$$\neg\texttt{Coherent}(\texttt{sun}, \texttt{snow})$$
$$\neg\texttt{Synonymous}(\texttt{showers}, \texttt{rain})$$

The condition literals are evaluated by querying the background knowledge in the form of a Prolog knowledgebase. The negation is interpreted as negation-as- failure. If a condition literal incorporates logical variables, these variables are handled by the Prolog knowledgebase. So if the condition is true, a grounding for the variable is returned by the background knowledge. Furthermore, any grounding is used to systematically instantiate further occurrences of that variable in the other literals in the fusion rule.

Example 14 *In the following condition,* X *is a logical variable, and* Interval *is a predicate that captures the interval of textentries, and so* X *will be ground by the prolog knowledgebase with a value that according to the background knowledge is the interval corresponding to the textentries in the input reports. Below the textentries are on the* report/windspeed *branch of the* $\texttt{in}_1$ *and* $\texttt{in}_2$ *reports (given as the top two reports in Example 1).*

$$\texttt{Interval}(\texttt{in}_1//\texttt{report}/\texttt{windspeed}\#, \texttt{in}_2//\texttt{report}/\texttt{windspeed}\#, \texttt{X})$$

After grounding the schema variables, we have the following.

$$\texttt{Interval}(\texttt{1 kph}, \texttt{25 kph}, \texttt{X})$$

Then the grounding for X *returned by the Prolog knowledgebase is* $\texttt{1} - \texttt{25 kph}$ *if we assume appropriate clauses in the Prolog knowledgebase.*

Further examples of condition literals that incorporate logical variables include Conjunction that for the textentries in the argument, the query variable returns the conjunction of them, and Disjunction that for the textentries in the argument, the query variable returns the disjunction of them. In Section 4, we consider in detail condition literals for handling the uncertain information in structured news reports.

3.2 Action atoms

Action atoms specify the structure and content for a merged report. In a ground fusion rule $\alpha' \Rightarrow \beta'$, if the ground literals in the antecedent α' hold, then the merged report should meet the specification represented by the ground atom β'. We look at this in more detail in the next section. We now define a basic set of action atoms. A number of further definitions for action atoms are possible.

Definition 23 *The* **action atoms** *are literals that include the following specifying how the merged report should be constructed.*

1. $\texttt{Initialize}(\varphi(\phi_1, .., \phi_n))$ *where* $\varphi(\phi_1, .., \phi_n)$ *is a skeleton constant. The intended action is to start the construction of the merged structured news report with the basic structure being defined by* $\varphi(\phi_1, .., \phi_n)$*. So the root of the merged report is* φ*.*
2. $\texttt{AddText}(\texttt{T}, \varphi_1/../\varphi_n)$ *where* T *is a textentry, and* $\varphi_1/../\varphi_n$ *is a branch constant. The intended action is to add the textentry* T *as the child to the tagname* φ_n *in the merged report on the branch* $\varphi_1/../\varphi_n$*.*
3. $\texttt{AddTree}(\texttt{T}, \varphi_1/../\varphi_n)$ *where* T *is a news term, and* $\varphi_1/../\varphi_n$ *is a branch constant. The intended action is to add* T *to the merged report so that the tagname for the root of* T *has the parent* φ_n *on the branch* $\varphi_1/../\varphi_n$*.*

The action atoms are specifications that are intended to be made to hold by producing a merged report that satisfies the specification.

Example 15 *Consider the action literal in the consequent of Example 9.*

$$\texttt{AddText}(\texttt{showers and inclement}, \texttt{report/outlook})$$

This specifies that the textentry should be showers and inclement *in the merged report for tagname* outlook *on the branch* report/outlook *.*

Example 16 *An appropriate set of fusion rules and the pair of news reports given in the top of Example 1 together with appropriate background knowledge can give the following action atoms:*

$\texttt{Initialize}(\texttt{report}(\texttt{source}, \texttt{date}, \texttt{city}, \texttt{outlook}, \texttt{windspeed}))$
$\texttt{AddText}(\texttt{TV1 and TV3}, \texttt{report/source})$
$\texttt{AddText}(\texttt{19/03/02}, \texttt{report/date})$
$\texttt{AddText}(\texttt{London}, \texttt{report/city})$
$\texttt{AddText}(\texttt{showers and inclement}, \texttt{report/outlook})$
$\texttt{AddText}(\texttt{1} - \texttt{25Kph}, \texttt{report/windspeed})$
$\texttt{AddTree}(\texttt{temperature}(\texttt{probability}(\texttt{prob}(8^\circ\texttt{C}, 0.14), \texttt{prob}(12^\circ\texttt{C}, 0.86))), \texttt{report})$

These action atoms specify the merged report given in the bottom left of Example 1. For more information on the fusion rules used and how they were executed, see [HS04].

4 Merging uncertain information

In this section, we concentrate on merging news reports with uncertain information (uncertainty components) of the same kind: one of probabilistic, possibilistic, or belief function information. We consider merging pairs of homogeneous distributions in cases of probabilistic and belief function information. If we need to merge more than two, then we apply the process by recursion. So to merge three distributions, we merge the first two, and then merge this output with the third distribution.We leave the topic of merging heterogeneous uncertainty components from multiple news reports until the next section.

When merging two news reports, one with an uncertainty component and one without, we can take the latter as a special case of the former and assign value 1.0 (no matter whether it stands for probability, or possibility, or mass value) to the corresponding textentry (or textentries). Then, these two news reports can be merged using one of the rules defined below.

Before proceeding to the details of fusion rules, we need to emphasize that in this paper any two uncertainty-valid components to be merged are assumed to refer to the same issue (or topic) that is being considered. The method to verify semantically whether two given uncertainty components are eligible for merging is given in [HS04]. In the rest of this paper, whenever we intend to merge two such components, we assume their eligibility has been checked and we will not repeat this prerequisite any further.

4.1 Merge probabilistic uncertainty components

We consider two simple ways of merging probability-valid uncertainty components using condition literals. We start with merging based on taking an average probability distribution.

Definition 24 *Let the following be two probability-valid uncertainty components.*

$$\langle \texttt{probability} \rangle \sigma_1^1, .., \sigma_n^1 \langle /\texttt{probability} \rangle$$
$$\langle \texttt{probability} \rangle \sigma_1^2, .., \sigma_m^2 \langle /\texttt{probability} \rangle$$

Let the **probability average component** *be an uncertainty component that is obtained by merging the above uncertainty components as follows*

$$\langle \texttt{probability} \rangle \sigma_1, .., \sigma_s \langle /\texttt{probability} \rangle$$

where each $\sigma_k \in \{\sigma_1, .., \sigma_s\}$ is of the form $\langle \texttt{prob value} = \kappa_k \rangle \phi \langle /\texttt{prob} \rangle$ and obtained by one of the following three steps.

1. *if there is a $\sigma_i^1 \in \{\sigma_1^1, .., \sigma_n^1\}$ and $\sigma_j^2 \in \{\sigma_1^2, .., \sigma_m^2\}$ where σ_i^1 is $\langle \texttt{prob value} = \kappa_i^1 \rangle \phi \langle /\texttt{prob} \rangle$ and σ_j^2 is $\langle \texttt{prob value} = \kappa_j^2 \rangle \phi \langle /\texttt{prob} \rangle$, then $\kappa_k = (\kappa_i^1 + \kappa_j^2)/2$.*
2. *if there is a $\sigma_i^1 \in \{\sigma_1^1, .., \sigma_n^1\}$ such that σ_i^1 is $\langle \texttt{prob value} = \kappa_i^1 \rangle \phi \langle /\texttt{prob} \rangle$, and there is no $\sigma_j^2 \in \{\sigma_1^2, .., \sigma_m^2\}$ such that σ_j^2 is $\langle \texttt{prob value} = \kappa_j^2 \rangle \phi \langle /\texttt{prob} \rangle$, then $\kappa_k = \kappa_i^1/2$.*
3. *if there is no $\sigma_i^1 \in \{\sigma_1^1, .., \sigma_n^1\}$ such that σ_i^1 is $\langle \texttt{prob value} = \kappa_i^1 \rangle \phi \langle /\texttt{prob} \rangle$, and there is a $\sigma_j^2 \in \{\sigma_1^2, .., \sigma_m^2\}$ such that σ_j^2 is $\langle \texttt{prob value} = \kappa_j^2 \rangle \phi \langle /\texttt{prob} \rangle$, then $\kappa_k = \kappa_j^2/2$.*

Definition 25 *Let the news terms τ_1 and τ_2 each denote a probability-valid uncertainty component and let* `X` *be a logical variable. The condition literal $\texttt{AvProbDist}(\tau_1, \tau_2, \texttt{X})$ is such that* `X` *is evaluated to τ_3 where τ_3 is the news term denoting the average uncertainty component obtained by Definition 24.*

The `AvProbDist` condition literal can be used as a condition in a fusion rule for merging pairs of structured news reports that each incorporate a probability-valid uncertainty component.

Example 17 *Consider the following pair of uncertainty components. Let the left be τ_1 and the right be τ_2.*

$$\begin{array}{ll}
\langle \texttt{probability} \rangle & \langle \texttt{probability} \rangle \\
\quad \langle \texttt{prob value} = \text{“0.2”} \rangle 8^\circ\texttt{C} \langle /\texttt{prob} \rangle & \quad \langle \texttt{prob value} = \text{“0.4”} \rangle 8^\circ\texttt{C} \langle /\texttt{prob} \rangle \\
\quad \langle \texttt{prob value} = \text{“0.8”} \rangle 12^\circ\texttt{C} \langle /\texttt{prob} \rangle & \quad \langle \texttt{prob value} = \text{“0.6”} \rangle 12^\circ\texttt{C} \langle /\texttt{prob} \rangle \\
\langle /\texttt{probability} \rangle & \langle /\texttt{probability} \rangle
\end{array}$$

For this, $\texttt{AvProbDist}(\tau_1, \tau_2, \texttt{X})$ *is such that* $\texttt{X}$ *is evaluated to* τ_3 *where* τ_3 *is the news term denoting the following uncertainty component,*

$$\begin{array}{l} \langle\texttt{probability}\rangle \\ \quad\langle\texttt{prob value} = \text{``}\texttt{0.3}\text{''}\rangle 8^\circ\texttt{C}\langle/\texttt{prob}\rangle \\ \quad\langle\texttt{prob value} = \text{``}\texttt{0.7}\text{''}\rangle 12^\circ\texttt{C}\langle/\texttt{prob}\rangle \\ \langle/\texttt{probability}\rangle \end{array}$$

Whilst in some situations, taking the average distribution is a natural and simple choice, it does suffer from the biases normally associated with taking the mean. More importantly, associativity does not hold, so if we want to merge three uncertainty components τ_1, τ_2 and τ_3 to get a merged uncertainty component (corresponding to the instantiation of $\texttt{Y}$), then the merged component obtained by the following sequence of applications of the condition literal

$$\texttt{AvProbDist}(\tau_1, \tau_2, \texttt{X}) \ \textsf{AND} \ \texttt{AvProbDist}(\texttt{X}, \tau_3, \texttt{Y})$$

is not necessarily the same as the the merged component (corresponding to the instantiation of $\texttt{Y}'$) obtained by the following sequence of applications of the condition literal.

$$\texttt{AvProbDist}(\tau_1, \tau_3, \texttt{X}') \ \textsf{AND} \ \texttt{AvProbDist}(\texttt{X}', \tau_2, \texttt{Y}')$$

4.2 Merge belief functions in Dempster-Shafer theory of evidence

In this subsection, we present procedures to merge two mass functions provided by two distinct bodies of evidence using Dempster's combination rule.

Definition 26 *Let the following be two belfunction-valid uncertainty components*

$$\begin{array}{c} \langle\texttt{belfunction}\rangle\sigma^1_1, .., \sigma^1_p\langle/\texttt{belfunction}\rangle \\ \langle\texttt{belfunction}\rangle\sigma^2_1, .., \sigma^2_q\langle/\texttt{belfunction}\rangle \end{array}$$

where

1. $\sigma^1_i \in \{\sigma^1_1, .., \sigma^1_p\}$ *is of the form* $\langle\texttt{mass value} = \kappa^1_i\rangle\psi^1_i\langle/\texttt{mass}\rangle$
2. ψ^1_i *is of the form* $\langle\texttt{massitem}\rangle\phi^1_{i_1}\langle/\texttt{massitem}\rangle \cdots \langle\texttt{massitem}\rangle\phi^1_{i_x}\langle/\texttt{massitem}\rangle$
3. $\sigma^2_j \in \{\sigma^2_1, .., \sigma^2_q\}$ *is of the form* $\langle\texttt{mass value} = \kappa^2_j\rangle\psi^2_i\langle/\texttt{mass}\rangle$
4. ψ^2_j *is of the form* $\langle\texttt{massitem}\rangle\phi^2_{j_1}\langle/\texttt{massitem}\rangle \cdots \langle\texttt{massitem}\rangle\phi^2_{j_y}\langle/\texttt{massitem}\rangle$

Let the **combined belfunction component** *be an uncertainty component that is obtained by merging the above uncertainty components as follows*

$$\langle\texttt{belfunction}\rangle\sigma_1, .., \sigma_s\langle/\texttt{belfunction}\rangle$$

where each $\sigma_k \in \{\sigma_1, .., \sigma_s\}$ *is of the form* $\langle\texttt{mass value} = \kappa_k\rangle\psi\langle/\texttt{mass}\rangle$ *and*

$$\kappa_k = \frac{\Sigma\kappa^1_i \times \kappa^2_j}{1 - \kappa_\perp}$$

and

$$\kappa_\perp = \Sigma\kappa^1_n \times \kappa^2_m$$

and

1. ψ *is of the form* $\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle\cdots\langle\texttt{massitem}\rangle\phi_z\langle/\texttt{massitem}\rangle$
2. $\{\phi_1, .., \phi_z\} = \{\phi^1_{i_1}, .., \phi^1_{i_x}\} \cap \{\phi^2_{j_1}, .., \phi^2_{j_y}\}$
3. $\sigma^1_n \in \{\sigma^1_1, .., \sigma^1_p\}$ *is of the form* $\langle\texttt{mass value} = \kappa^1_n\rangle\psi^1_n\langle/\texttt{mass}\rangle$
4. ψ^1_n *is of the form* $\langle\texttt{massitem}\rangle\phi^1_{n_1}\langle/\texttt{massitem}\rangle\cdots\langle\texttt{massitem}\rangle\phi^1_{n_v}\langle/\texttt{massitem}\rangle$
5. $\sigma^2_m \in \{\sigma^2_1, .., \sigma^2_q\}$ *is of the form* $\langle\texttt{mass value} = \kappa^2_m\rangle\psi^2_m\langle/\texttt{mass}\rangle$
6. ψ^2_m *is of the form* $\langle\texttt{massitem}\rangle\phi^2_{m_1}\langle/\texttt{massitem}\rangle\cdots\langle\texttt{massitem}\rangle\phi^2_{m_w}\langle/\texttt{massitem}\rangle$
7. $\{\phi^1_{n_1}, .., \phi^1_{n_v}\} \cap \{\phi^2_{m_1}, .., \phi^2_{m_w}\} = \emptyset$

The value $\kappa_\perp = \Sigma\kappa^1_n \times \kappa^2_m$ (that is, $\Sigma_{A\cap B=\emptyset}\ (m_1(A) \times m_2(B))$ indicates how much of the total belief has been committed to the emptyset while combining two pieces of uncertain information. A higher $\kappa_\perp$ value reflects either an inconsistency among the two sources or lower confidence in any of the possible outcomes from both sources. We will discuss this in detail in the next section.

Definition 27 *Let the news terms τ_1 and τ_2 each denote a belfunction-valid component and let* `X` *be a logical variable. The condition literal* $\texttt{Dempster}(\tau_1, \tau_2, \texttt{X})$ *is such that* `X` *is evaluated to τ_3 where τ_3 is the news term denoting the combined belfunction component obtained by Definition 26.*

The `Dempster` condition literal can be used as a condition in a fusion rule for merging pairs of structured news reports that each incorporate a belfunction-valid uncertainty component.

Example 18 *Consider the following belfunction-valid components.*

```
⟨belfunction⟩
  ⟨mass value = “0.2”⟩
    ⟨massitem⟩8°C⟨/massitem⟩
    ⟨massitem⟩10°C⟨/massitem⟩
  ⟨/mass⟩
  ⟨mass value = “0.8”⟩
    ⟨massitem⟩12°C⟨/massitem⟩
  ⟨/mass⟩
⟨/belfunction⟩
```

```
⟨belfunction⟩
  ⟨mass value = “0.4”⟩
    ⟨massitem⟩8°C⟨/massitem⟩
    ⟨massitem⟩10°C⟨/massitem⟩
  ⟨/mass⟩
  ⟨mass value = “0.6”⟩
    ⟨massitem⟩12°C⟨/massitem⟩
  ⟨/mass⟩
⟨/belfunction⟩
```

For this, $\texttt{Dempster}(\tau_1, \tau_2, \texttt{X})$ *is such that* `X` *is evaluated to τ_3 where τ_3 is the news term denoting the following uncertainty component,*

```
⟨belfunction⟩
  ⟨mass value = “0.14”⟩
    ⟨massitem⟩8°C⟨/massitem⟩
    ⟨massitem⟩10°C⟨/massitem⟩
  ⟨/mass⟩
  ⟨mass value = “0.86”⟩
    ⟨massitem⟩12°C⟨/massitem⟩
  ⟨/mass⟩
⟨/belfunction⟩
```

Both sources have a higher confidence in the choice $\{\texttt{12°C}\}$ than in $\{\texttt{8°C}, \texttt{10°C}\}$, so the combined result gives a higher confidence in the choice preferred by both of them and a lower confidence in the less preferred one. This is due to the fact that both sources are in agreement with each other. Therefore, when multiple sources are not in conflict, applying fusion rules will produce a more complete and comprehensive solution than individual sources.

4.3 Detection of inconsistency among sources

Dempster's rule is known to generate counterintuitive results from two almost conflicting pieces of evidence [DP88a, Sme88]. We illustrate this in the next example.

Example 19 *If two belfunction-valid uncertainty components are as below, with corresponding frame of discernment as* $\Omega = \{\phi_1, \phi_2, \phi_3, \phi_4, \phi_5\}$.

```
⟨belfunction⟩
 ⟨mass value = "0.9"⟩
  ⟨massitem⟩$\phi_1$⟨/massitem⟩
  ⟨massitem⟩$\phi_2$⟨/massitem⟩
 ⟨/mass⟩
 ⟨mass value = "0.1"⟩
  ⟨massitem⟩$\phi_3$⟨/massitem⟩
 ⟨/mass⟩
 ⟨mass value = "0.0"⟩
  ⟨massitem⟩$\phi_4$⟨/massitem⟩
 ⟨/mass⟩
⟨/belfunction⟩
```

```
⟨belfunction⟩
 ⟨mass value = "0.0"⟩
  ⟨massitem⟩$\phi_1$⟨/massitem⟩
  ⟨massitem⟩$\phi_2$⟨/massitem⟩
 ⟨/mass⟩
 ⟨mass value = "0.1"⟩
  ⟨massitem⟩$\phi_3$⟨/massitem⟩
 ⟨/mass⟩
 ⟨mass value = "0.9"⟩
  ⟨massitem⟩$\phi_4$⟨/massitem⟩
 ⟨/mass⟩
⟨/belfunction⟩
```

then the condition literal $\texttt{Dempster}(\tau_1, \tau_2, \texttt{X})$ *will result in an uncertainty component as follows.*

```
⟨belfunction⟩
 ⟨mass value = "0.0"⟩
  ⟨massitem⟩$\phi_1$⟨/massitem⟩
  ⟨massitem⟩$\phi_2$⟨/massitem⟩
 ⟨/mass⟩
 ⟨mass value = "1.0"⟩
  ⟨massitem⟩$\phi_3$⟨/massitem⟩
 ⟨/mass⟩
 ⟨mass value = "0.0"⟩
  ⟨massitem⟩$\phi_4$⟨/massitem⟩
 ⟨/mass⟩
⟨/belfunction⟩
```

which confirms the outcome ϕ_3 *with full degree of belief that is hardly supported by either of the sources.*

The degree of conflict among the two sources in this example is well indicated by the mass value assigned to the empty set during the combination, which is 0.99 *out of* 1.0*. However, not all high degrees of belief in the empty set forecast a conflict between two sources, when using Dempster's combination rule. For* *instance, if two sources provide the following two mass distributions to the same frame of discernment,*

$\langle\texttt{belfunction}\rangle$	$\langle\texttt{belfunction}\rangle$
$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$	$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$
$\qquad\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle$	$\qquad\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle$
$\quad\langle/\texttt{mass}\rangle$	$\quad\langle/\texttt{mass}\rangle$
$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$	$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$
$\qquad\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle$	$\qquad\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle$
$\quad\langle/\texttt{mass}\rangle$	$\quad\langle/\texttt{mass}\rangle$
$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$	$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$
$\qquad\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle$	$\qquad\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle$
$\quad\langle/\texttt{mass}\rangle$	$\quad\langle/\texttt{mass}\rangle$
$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$	$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$
$\qquad\langle\texttt{massitem}\rangle\phi_4\langle/\texttt{massitem}\rangle$	$\qquad\langle\texttt{massitem}\rangle\phi_4\langle/\texttt{massitem}\rangle$
$\quad\langle/\texttt{mass}\rangle$	$\quad\langle/\texttt{mass}\rangle$
$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$	$\quad\langle\texttt{mass value} = \text{“0.2”}\rangle$
$\qquad\langle\texttt{massitem}\rangle\phi_5\langle/\texttt{massitem}\rangle$	$\qquad\langle\texttt{massitem}\rangle\phi_5\langle/\texttt{massitem}\rangle$
$\quad\langle/\texttt{mass}\rangle$	$\quad\langle/\texttt{mass}\rangle$
$\langle/\texttt{belfunction}\rangle$	$\langle/\texttt{belfunction}\rangle$

then $\texttt{Dempster}(\tau_1, \tau_2, \texttt{X})$ *will produce a merged uncertainty component that assigns* 0.80 *to the emptyset. Therefore, these two pieces of uncertain information may be wrongly considered as inconsistent if we use this value only to make a judgement. The fact is that these two sources are not inconsistent or conflicting but both have lower confidence in each possible solution they predict.*

We can view the situation in Example 19 in terms of probability theory: The above two mass functions are actually probability distributions, since they both assign mass values to singletons as in the case in probabilities. Therefore, Dempster's rule can be used to combine two probability distributions when they are from independent sources. To do so, we use the following definition to convert a probability-valid component into a belfunction-valid component.

Definition 28 *Let news term* τ *be a probability-valid component* $\langle\texttt{probability}\rangle\sigma_1, .., \sigma_n\langle/\texttt{probability}\rangle$ *and each* $\sigma_i \in \{\sigma_1, .., \sigma_n\}$ *is of the form* $\langle\texttt{prob value} = \kappa\rangle\phi\langle/\texttt{prob}\rangle$ *where* $\kappa \in [0, 1]$ *and* ϕ *is a textentry. Then* τ' *is a news term denoting a belfunction-valid component* $\langle\texttt{belfunction}\rangle\sigma'_1, .., \sigma'_\texttt{n}\langle/\texttt{belfunction}\rangle$ *and each* $\sigma'_i \in \{\sigma'_1, .., \sigma'_n\}$ *is of the form* $\langle\texttt{mass value} = \kappa\rangle\langle\texttt{massitem}\rangle\phi\langle/\texttt{massitem}\rangle\langle/\texttt{mass}\rangle$ *where* $\kappa \in [0, 1]$ *and* ϕ *is a textentry.* τ' *is the* **belfunction conversion** *of* τ*.*

Any pair of probability-valid components, after being converted into belfunction-valid components, can be combined by Dempster's combination rule. The definition below detects whether this pair implies any potentially contradictory imformation.

Definition 29 *Let news terms* τ_1 *and* τ_2 *be two belfunction-valid component that are converted from two probability-valid news terms. If the mass assigned to the empty set, denoted by* κ*, by* $\texttt{Dempster}$ $(\tau_1, \tau_2, \texttt{X})$ *satisfies* $\kappa > \epsilon$ *then* τ_1 *and* τ_2 *are said to be* **potentially inconsistent**, *where* ϵ *is a pre-defined threshold in* $[0, 1]$ *such as* 0.8*. If* τ_1 *and* τ_2 *are potentially inconsistent, then the condition literal* $\texttt{PotentialDempsterInconsistency}(\tau_1, \tau_2, \epsilon)$ *holds.*

Since there does not exist an "absolute meaningful threshold" to measure the inconsistency of two mass functions [AS01], the choice of ϵ is largely subjective. In general, the closer ϵ is to 1.0, the greater the inconsistency becomes. Therefore, the choice of the threshold value ϵ in Definition 29 is sensitive and crucial to the detection of any potential inconsistency, and is application specific. If this value is too high, the above definition will fail to identify some contradictory opinions. For example, if we set ϵ as 0.9, two conflicting mass functions m_1 and m_2 below.

$$m_1(\{\phi_1\}) = 0.1 \;\; m_1(\{\phi_2\}) = 0.8 \;\; m_1(\{\phi_3\}) = 0.1$$

$$m_2(\{\phi_1\}) = 0.1 \;\; m_2(\{\phi_2\}) = 0.1 \;\; m_2(\{\phi_3\}) = 0.8$$

are not regarded as conflicting based on Definition 29. On the other hand, if ϵ is set too low, it rings a false alarm too frequently.

It should be noted that most of the papers in the literature on DS theory refer the amount of mass assigned to the emptyset as **the degree of conflict between the two mass functions** following Shafer's original wordings [Sha76]. And most papers use examples similar to Example 19 to illustrate what a conflict should look like. **A conflict between two sources** therefore can be interpreted as *one source strongly supports one hypothesis and the other strongly supports another hypothesis, and the two hypotheses are not compatible*.

To make sure that an inconsistency among a pair of probability-valid components actually forecast a conflict, when

$$\texttt{PotentialDempsterInconsistency}(\tau_1, \tau_2, \epsilon)$$

holds, we deploy the following definition to double check that the pair indeed has rather different preferences (in terms of high probabilities).

Definition 30 *Let* $\langle\texttt{probability}\rangle\sigma_1^1, .., \sigma_{\texttt{n}}^1\langle/\texttt{probability}\rangle$ *and* $\langle\texttt{probability}\rangle\sigma_1^2, .., \sigma_m^2\langle/\texttt{probability}\rangle$ *be two probability-valid components and be denoted by* τ_1 *and* τ_2*. Let news terms* τ_1' *and* τ_2' *be belfunction-valid components that are converted from them respectively and they are potentially inconsistent (i.e.* $\texttt{PotentialDempsterInconsistency}(\tau_1', \tau_2', \epsilon)$ *holds). The two probability-valid components are* **conflicting** *if the following conditions hold*

1. $\exists\sigma_i^1 \in \{\sigma_1^1, .., \sigma_n^1\}$ *of the form* $\langle\texttt{prob value} = \kappa_i^1\rangle\phi_i^1\langle/\texttt{prob}\rangle$, $\texttt{s.t.}$ *if* $\exists\sigma_j^2 \in \{\sigma_1^2, .., \sigma_m^2\}$ *of the form* $\langle\texttt{prob value} = \kappa_j^2\rangle\phi_j^2\langle/\texttt{prob}\rangle$ *and* $\phi_i^1 = \phi_j^2$*, then* $|\kappa_i^1 - \kappa_j^2| > \epsilon'$*, otherwise* $\kappa_i^1 > \epsilon'$*.*
2. $\exists\sigma_l^2 \in \{\sigma_1^2, .., \sigma_m^2\}$ *of the form* $\langle\texttt{prob value} = \kappa_l^2\rangle\phi_l^2\langle/\texttt{prob}\rangle$, $\texttt{s.t.}$ *if* $\exists\sigma_k^1 \in \{\sigma_1^1, .., \sigma_n^1\}$ *of the form* $\langle\texttt{prob value} = \kappa_k^1\rangle\phi_k^1\langle/\texttt{prob}\rangle$ *and* $\phi_l^2 = \phi_k^1$*, then* $|\kappa_l^2 - \kappa_k^1| > \epsilon'$*, otherwise* $\kappa_l^2 > \epsilon'$*.*

where $\phi_i^1 \neq \phi_l^2$ *and the threshold* $\epsilon' \in [0, 1]$ *is pre-defined, such as* 0.6*. If* τ_1 *and* τ_2 *are conflicting according to this definition, then the following condition literal holds.*

$$\texttt{ProbabilisticDempsterInconsistency}(\tau_1, \tau_2, \epsilon')$$

The value $|\kappa_i^1 - \kappa_j^2|$, such as 0.7, denotes the length of the interval (or the distance of the two values defining the interval) for an outcome. This interval implies the conflict in supporting ϕ_i^1 (or equally ϕ_j^2) from the two sources. In other words, one source supports ϕ_i^1 strongly and another very weakly.

Once again, the choice of the threshold is subject to individual applications. Some applications may require more precise predictions, hence a small interval value would be essential to guarantee the detection of conflicts.

The above two definitions can help to detect whether there is a conflict among a pair of sources, only when both uncertainty components in the two news reports are probability-valid. Since the mass value assigned to the empty set itself can only detect a *possible inconsistent pair*, we provide the following additional definition to further clarify if these two sources are conflicting or not.

Definition 31 *Let news terms τ_1 and τ_2 be belfunction-valid components. If* Dempster $(\tau_1, \tau_2, \mathtt{X})$ *detects a possible conflict among τ_1 and τ_2 (i.e.* PotentialDempsterInconsistency$(\tau_1, \tau_2, \epsilon)$ *holds), and if the following conditions hold for all the uncertainty components σ_i^1 and σ_j^2 in τ_1 and τ_2 respectively,*

$$\sigma_i^1 \textit{ is of the form } \langle \mathtt{mass\ value} = \kappa_i^1 \rangle \psi_i^1 \langle /\mathtt{mass} \rangle \textit{ and } \kappa_i^1 \leq \epsilon'$$

$$\sigma_j^2 \textit{ is of the form } \langle \mathtt{mass\ value} = \kappa_j^2 \rangle \psi_j^2 \langle /\mathtt{mass} \rangle \textit{ and } \kappa_j^2 \leq \epsilon'$$

then the two sources are said to be **not** *in conflict where ϵ' is a pre-defined threshold in $[0, 1]$ such as 0.3. In this case, the condition literal* DempsterInconsistency$(\tau_1, \tau_2, \epsilon')$ *does not hold. Otherwise, the pair is said to be conflicting, and so the condition literal* DempsterInconsistency$(\tau_1, \tau_2, \epsilon')$ *does hold.*

This definition rules out a possible wrong conclusion of contradiction among two sources due to lower mass values initially assigned which have produced a high mass value to the empty set, as illustrated in Example 19.

In this section, we have defined the following predicates which can be used as extra conditions to restrict the application of the Dempster predicate given in Definition 27 in fusion rules.

PotentialDempsterInconsistency
ProbabilisticDempsterInconsistency
DempsterInconsistency

For example for a particular fusion, and for news terms τ_1 and τ_2, we have Dempster(τ_1, τ_2, τ_3) holding, but also the negated literals

$$\neg\mathtt{PotentialDempsterInconsistency}(\tau_1, \tau_2, 0.8)$$

and

$$\neg\mathtt{DempsterInconsistency}(\tau_1, \tau_2, 0.3)$$

not holding, then the fusion rule would fail. In other words, if PotentialDempsterInconsistency$(\tau_1, \tau_2, 0.8)$ and DempsterInconsistency$(\tau_1, \tau_2, 0.3)$ hold, then these would cause the fusion rule to fail, and hence merging based on Dempster-Shafer theory in this fusion rule would be blocked.

4.4 Merge possibility distributions in possibility theory

4.4.1 From necessity measures to possibility distributions

A possibility-valid uncertainty component usually specifies a partial necessity measure. Below we first recover the possibility distribution associated with this necessity measure using the minimum specificity principle.

Let a possibility-valid uncertainty component be

$$\langle \mathtt{possibility} \rangle \sigma_1, .., \sigma_p \langle /\mathtt{possibility} \rangle$$

where $\sigma_i \in \{\sigma_1, .., \sigma_p\}$ is of the form

$$\langle \mathtt{ness\ value} = \kappa_i \rangle \psi_i \langle /\mathtt{ness} \rangle$$

and ψ_i is of the form

$$\langle \mathtt{nessitem} \rangle \phi_{i_1} \langle /\mathtt{nessitem} \rangle \cdots \langle \mathtt{nessitem} \rangle \phi_{i_x} \langle /\mathtt{nessitem} \rangle$$

We denote the frame associated with a possibility-valid uncertainty component as $\Omega = \{\phi_1, .., \phi_n\}$, and also let $\psi_i = \{\phi_{i_1}, .., \phi_{i_x}\}$ in order to make the subsequent description simpler. In this way, a possibility-valid uncertainty component can be viewed as consisting of a finite set of weighted subsets of Ω, $\{(\psi_i, \kappa_i), i = 1, .., p\}$, where κ_i is interpreted as a lower bound on the degree of necessity $N(\psi_i)$. This representation is consistent with notations in [DP87] and analogous with notations in possibilistic knowledge bases using possibilistic logic (e.g., [BDP97, BDKP00]), where uncertain knowledge is represented as a set of weighted formulae, $\{(p_i, a_i), i = 1, .., n\}$. A subset ψ_i and formula p_i are thought to be equivalent if p_i if defined as

$$p_i = \vee q_j, \ \ \texttt{where } q_j \texttt{ stands for } ``\phi_j \texttt{ is true}", \phi_j \in \psi_i$$

Therefore, when one of the elements in ψ_i is definitely true, formula p_i is definitely true as well.

There is normally a family of possibility distributions associated with a given possibility-valid uncertainty component, with each of the distributions satisfying the condition

$$1 - \mathsf{max}\{\pi(\phi) | \phi \in \bar{\psi_i}\} \geq \kappa_i$$

That is, a necessity measure N derived from such a compatible possibility distribution π guarantes that $N(\psi_i) \geq \kappa_i$.

A common method to select one of the compatible possibility distributions is to use the **minimum specificity principle** [DP87]. Let $\{\pi_j, j = 1, .., m\}$ be all the possibility distributions that are compatible with $\{(\psi_i, \kappa_i), i = 1, .., p\}$. A possibility distribution $\pi_l \in \{\pi_j, j = 1, .., m\}$ is said to be the least specific possibility distribution among $\{\pi_j, j = 1, .., m\}$ if $\not\exists \pi_t \in \{\pi_j, j = 1, .., m\}, \pi_t \neq \pi_l$ such that $\forall \phi, \pi_t(\phi) \geq \pi_l(\phi)$.

The minimum specificity principle allocates the greatest possibility degrees in agreement with the constraints $N(\psi_i) \geq \kappa_i$. This possibility distribution always exists and is defined as ([DP87, BDP97])

$$\forall \phi \in \Omega, \pi(\phi) = \mathsf{min}\{1 - \kappa_i | \phi \in \bar{\psi_i}\} \tag{1}$$

Each value, $1 - \kappa_i$ for $\phi \in \bar{\psi_i}$, specifies a maximum degree of possibility that ϕ can take given that at least κ_i amount of necessity has been allocated to its opponent, e.g., $\Pi(\psi_i) \leq 1 - \kappa_i$. Among several possible maximum degrees of possibility for ϕ, the operator min selects the lowerest degree of possibility.

When ϕ appears in all specified subsets (each with a necessity degree), ϕ is absolutely possible and shall have the maximum degree of possiblity 1.

$$\forall \phi \in \Omega, \pi(\phi) = \begin{cases} \mathsf{min}\{1 - \kappa_i | \phi \notin \psi_i\} & \mathsf{when} \exists \psi_i \mathsf{such\ that} \phi \notin \psi_i \\ 1 & \mathsf{otherwise} \end{cases} \tag{2}$$

and it can further be rewritten as

$$\forall \phi \in \Omega, \pi(\phi) = \begin{cases} 1 - \mathsf{max}\{\kappa_i | \phi \notin \psi_i\} & \mathsf{when} \exists \psi_i \mathsf{such\ that} \phi \notin \psi_i \\ 1 & \mathsf{otherwise} \end{cases} \tag{3}$$

Definition 32 *Let a possibility-valid uncertainty component be*

$$\langle \texttt{possibility} \rangle \sigma_1, .., \sigma_p \langle / \texttt{possibility} \rangle$$

where

1. $\sigma_i \in \{\sigma_1, .., \sigma_p\}$ *is in the form* $\langle \texttt{poss value} = \kappa_i \rangle \psi_i \langle / \texttt{poss} \rangle$

2. ψ_i is of the form $\langle\texttt{nessitem}\rangle\phi_{i_1}\langle/\texttt{nessitem}\rangle\cdots\langle\texttt{nessitem}\rangle\phi_{i_x}\langle/\texttt{nessitem}\rangle$

and the set of weighted subsets is $\{(\psi_i,\kappa_i), i=1,..,p\}$.

Let the possibility distribution obtained from using the minimum specificity principle be $\pi:\Omega\rightarrow[0,1]$, where

for each $\phi\in\Omega$, $\pi(\phi)=1-\nu$

where

$$\nu=\begin{cases}\mathsf{max}\{\kappa_1,\kappa_2,..,\kappa_t\} & \phi\notin\psi_j, j=1,2,..,t\ \ (\texttt{where } p\geq t>0)\\ 0 & \texttt{otherwise}\end{cases}$$

Example 20 *Consider the following two possibility-valid components with associated frame of discernment $\Omega=\{\phi_1,\phi_2,\phi_3,\phi_4\}$.*

$\langle\texttt{possibility}\rangle$
$\quad\langle\texttt{ness value} = \text{“0.2”}\rangle$
$\qquad\langle\texttt{nessitem}\rangle\phi_1\langle/\texttt{nessitem}\rangle$
$\qquad\langle\texttt{nessitem}\rangle\phi_2\langle/\texttt{nessitem}\rangle$
$\quad\langle/\texttt{ness}\rangle$
$\quad\langle\texttt{ness value} = \text{“0.3”}\rangle$
$\qquad\langle\texttt{nessitem}\rangle\phi_3\langle/\texttt{nessitem}\rangle$
$\quad\langle/\texttt{ness}\rangle$
$\langle/\texttt{possibility}\rangle$

$\langle\texttt{possibility}\rangle$
$\quad\langle\texttt{ness value} = \text{“0.2”}\rangle$
$\qquad\langle\texttt{nessitem}\rangle\phi_1\langle/\texttt{nessitem}\rangle$
$\qquad\langle\texttt{nessitem}\rangle\phi_2\langle/\texttt{nessitem}\rangle$
$\quad\langle/\texttt{ness}\rangle$
$\quad\langle\texttt{ness value} = \text{“0.3”}\rangle$
$\qquad\langle\texttt{nessitem}\rangle\phi_2\langle/\texttt{nessitem}\rangle$
$\qquad\langle\texttt{nessitem}\rangle\phi_3\langle/\texttt{nessitem}\rangle$
$\quad\langle/\texttt{ness}\rangle$
$\langle/\texttt{possibility}\rangle$

π_1 and π_2 below are obtained from the left and right possibility-valid components respectively.

$$\pi_1(\phi_1)=0.7, \pi_1(\phi_2)=0.7, \pi_1(\phi_3)=0.8, \pi_1(\phi_4)=0.7$$

$$\pi_2(\phi_1)=0.7, \pi_2(\phi_2)=1, \pi_2(\phi_3)=0.8, \pi_2(\phi_4)=0.7$$

4.4.2 Inconsistency within a possibility-valid component

A possibility distribution is not normal if $\forall\phi, \pi(\phi)<1$. The value $1-\mathsf{max}_{\phi\in\Omega}\pi(\phi)$ is called **the degree of inconsistency** of the original possibility-valid component. In this situation, condition if "$N(A)>0$ then $\Pi(A)=1$", is no longer valid.

For instance, in Example 20, the possibility-valid component on the left is inconsistent since $\forall\phi, \pi(\phi)<1$, whilst the right one is consistent, because $1-\mathsf{max}_{\phi\in\Omega}(\pi_2(\phi))=0$.

Proposition 1 *Let $\{(\psi_i,a_i),\ i=1,..,p\}$ be weighted subsets of Ω and specified in a possibility-valid component with respect to frame of discernment Ω. This possibility-valid component is* **consistent** *iff $\cap_i\psi_i\neq\emptyset$.*

The degree of inconsistency of a possibility-valid component may not be the only merit to judge the quality of a source, as illustrated by the following example.

Example 21 *Consider the following two possibility-valid components with associated frame of discernment* $\Omega = \{\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6\}$.

```
⟨possibility⟩
   ⟨ness value = "0.2"⟩
      ⟨nessitem⟩φ1⟨/nessitem⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.3"⟩
      ⟨nessitem⟩φ3⟨/nessitem⟩
      ⟨nessitem⟩φ4⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.2"⟩
      ⟨nessitem⟩φ5⟨/nessitem⟩

   ⟨/ness⟩
   ⟨ness value = "0.3"⟩
      ⟨nessitem⟩φ6⟨/nessitem⟩

   ⟨/ness⟩
⟨/possibility⟩
```

```
⟨possibility⟩
   ⟨ness value = "0.2"⟩
      ⟨nessitem⟩φ1⟨/nessitem⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.3"⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
      ⟨nessitem⟩φ3⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.3"⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
      ⟨nessitem⟩φ4⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.3"⟩
      ⟨nessitem⟩φ4⟨/nessitem⟩
      ⟨nessitem⟩φ5⟨/nessitem⟩
   ⟨/ness⟩
⟨/possibility⟩
```

The two possibility distributions from these two possibility-valid components using Equations (1) are

$$\pi_1(\phi_1) = 0.7, \pi_1(\phi_2) = 0.7, \pi_1(\phi_3) = 0.7, \pi_1(\phi_4) = 0.7, \pi_1(\phi_5) = 0.7, \pi_1(\phi_6) = 0.7$$

and

$$\pi_2(\phi_1) = 0.7, \pi_2(\phi_2) = 0.7, \pi_2(\phi_3) = 0.7, \pi_2(\phi_4) = 0.7, \pi_2(\phi_5) = 0.7, \pi_2(\phi_6) = 0.7$$

The degrees of inconsistencies of the two possibility-valid components are the same, $1 - \max_{\phi\in\Omega}(\pi_1(\phi)) = 0.3$ and $1 - \max_{\phi\in\Omega}(\pi_2(\phi)) = 0.3$. However, if we examine the structure of the weighted subsets ψ_i^1, ψ_j^2 in detail, we will find that the right possibility-valid component is more coherent than the left one, since there is a significant overlap among subsets ψ_j^2 in this component. While any two subsets in the first component have no common elements. This observation leads to the following two definitions.

Definition 33 *Let a possibility-valid uncertainty component be*

$$\langle \texttt{possibility} \rangle \sigma_1, .., \sigma_p \langle /\texttt{possibility} \rangle$$

where

1. $\sigma_i \in \{\sigma_1, .., \sigma_p\}$ *is in the form* $\langle \texttt{ness value} = \kappa_i \rangle \psi_i \langle /\texttt{ness} \rangle$

2. ψ_i *is of the form* $\langle \texttt{nessitem} \rangle \phi_{i_1} \langle /\texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_{i_x} \langle /\texttt{nessitem} \rangle$

and the corresponding set of weighted subsets be $\{(\psi_i, \kappa_i), i = i, .., p\}$.

This component is said to be **inconsistent but with good quality**, *if there exists at least one* ψ_j, *called a* **separable element**, *such that*

$$(\bigcap_{i=1, i\neq j}^{p} \psi_i) \neq \emptyset \text{ and } \bigcap_{i=1}^{p} \psi_i = \emptyset$$

There can be several separable elements ψ_j satisfying this definition given a possibility-valid component.

Definition 34 *Let a possibility-valid uncertainty component be*

$$\langle \texttt{possibility} \rangle \sigma_1, .., \sigma_p \langle / \texttt{possibility} \rangle$$

where

1. $\sigma_i \in \{\sigma_1, .., \sigma_p\}$ *is in the form* $\langle \texttt{ness value} = \kappa_i \rangle \psi_i \langle / \texttt{ness} \rangle$

2. ψ_i *is of the form* $\langle \texttt{nessitem} \rangle \phi_{i_1} \langle / \texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_{i_x} \langle / \texttt{nessitem} \rangle$

and the corresponding set of weighted subsets be $\{(\psi_i, \kappa_i), i = 1, .., p\}$.

This component is said to be **inconsistent with lower quality**, *if for every pair* (ψ_i, ψ_j), $\psi_i \cap \psi_j = \emptyset$, *when* $\psi_i \neq \psi_j$.

The possibility distribution π associated with an inconsistent possibility-valid component can be normalized by one of the three normalization rules as reviewed in [BDP97]:

$$\pi_{n_1}(\phi) = \frac{\pi(\phi)}{\mathsf{max}\{\pi(\phi_i)\}} \tag{4}$$

$$\pi_{n_2}(\phi) = \begin{cases} 1 & \texttt{if} \pi(\phi) = \mathsf{max}\{\pi(\phi_i)\} \\ \pi(\phi) & \texttt{otherwise} \end{cases} \tag{5}$$

$$\pi_{n_3}(\phi) = \pi(\phi) + (1 - \mathsf{max}\{\pi(\phi_i)\}) \tag{6}$$

All these rules satisfy the minimal requirements for a normalized possibility distribution:

- $\exists \phi, \pi_{n_i}(\phi) = 1$
- if π is normal, then $\pi_{n_i} = \pi$
- $\forall \phi, \phi'$, if $\pi_{n_i}(\phi) < \pi_{n_i}(\phi')$ iff $\pi(\phi) < \pi(\phi')$

The first normalization rule is most common and is consistent with the normalization procedure in probability theory.

The degree of possibility of every element in Ω, a pre-defined set, is increased both in the first and the third rules. While in the second rule, only the current most possible elements are assigned with the maximum possibility. We harness the second rule with Definition (33) and assign the maximum possibility to the elements that have appeared in all but one subset in a possibility-valid component.

$$\pi_{n_4}(\phi) = \begin{cases} 1 & \phi \in (\bigcap_{i=1}^{p} \psi_i), \psi_i \neq \psi_j, \psi_j \text{ is a separable element} \\ & \text{s.t. if} \exists \phi_l \in (\bigcap_{i=1}^{p} \psi_i), \psi_i \neq \psi_l, \psi_l \text{ is a separable element} \\ & \text{ then } \pi(\phi) > \pi(\phi_l) \\ \pi(\phi) & \texttt{otherwise} \end{cases} \tag{7}$$

When there are several elements $\phi_i, .., \phi_j$ satisfying Equation (7) and they all have the same degree of possibility distribution, e.g., $\pi(\phi_i) = \pi(\phi_j)$, then we arbitrarily chose one of them to normalize.

In this paper, we assume that all the possibility-valid components are either *consistent* or *inconsistent but with good quality*, and leave the *inconsistent but with lower-quality* possibility-valid components to future papers where we will examine how to normalize such a distribution and how to discount it ([BDP97]) when there are more (reliable) sources available.

4.4.3 Merging possibility distributions

With two possibility-valid uncertainty components, there will be corresponding possibility distributions associated with them respectively. In this subsection, we discuss how to merge two (or more) possibility distributions.

The two basic combination rules in possibility theory are the **conjunction** and the **disjunction** of possibility distributions ([BDP97]) when n possibility distributions are given on the same frame of discernment.

$$\forall \phi, \pi_{cm}(\phi) = \mathsf{min}_{i=1}^{n}(\pi_i(\phi)) \tag{8}$$

$$\forall \phi, \pi_{dm}(\phi) = \mathsf{max}_{i=1}^{n}(\pi_i(\phi)) \tag{9}$$

The conjunction rule uses the minimum operation, so that the most specific source will determine the merged possibility distribution. The conjunction rule is useful and produces sensible results only when all the sources are taken as equally and fully reliable and these sources are highly in agreement with each other. Since otherwise, any information prefered by some sources but not by the others will be rejected. Therefore, the conjunction rule can lead to a new possibility distribution that is not normal, even though all the original possibility distributions are normal. When this happens, the merged possibility distribution expresses an inconsistency among the sources. Furthermore, it suggests that some sources (at least one) is wrong when all the degrees of possibility $\pi_{cm}(\phi)$ are significantly smaller than 1.

On the other hand, the disjunction rule uses the maximum operation, so the least specific source will determine π_{dm}. This operation takes the most optimistic view from all the sources and does not reject any suggestion from every source. A set of sources that should not be merged by the conjunction rule can be merged with the disjunction rule to generate the maximum coverage of the information provided by all the sources.

Here, any possibility-valid component that is *inconsistent but with good quality* is normalized using Equation (7) before being merged with other possibility-valid components. Therefore, all the possibility-valid components are normal when applying conjunctive or disjunctive operations.

Definition 35 *Let the following be two possibility-valid uncertainty components*

$$\langle \texttt{possibility} \rangle \sigma_1^1, .., \sigma_p^1 \langle / \texttt{possibility} \rangle$$
$$\langle \texttt{possibility} \rangle \sigma_1^2, .., \sigma_q^2 \langle / \texttt{possibility} \rangle$$

where

1. $\sigma_i^1 \in \{\sigma_1^1, .., \sigma_p^1\}$ *is in the form* $\langle \texttt{ness value} = \kappa_i^1 \rangle \psi_i^1 \langle / \texttt{ness} \rangle$
2. ψ_i^1 *is of the form* $\langle \texttt{nessitem} \rangle \phi_{i_1}^1 \langle / \texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_{i_x}^1 \langle / \texttt{nessitem} \rangle$
3. $\{(\psi_i^1, \kappa_i^1), i = 1, .., p\}$ *is the set of weighted subsets with respect to* $\{\sigma_1^1, .., \sigma_p^1\}$
4. $\sigma_j^2 \in \{\sigma_1^2, .., \sigma_q^2\}$ *is in the form* $\langle \texttt{ness value} = \kappa_j^2 \rangle \psi_j^2 \langle / \texttt{ness} \rangle$
5. ψ_j^2 *is of the form* $\langle \texttt{nessitem} \rangle \phi_{j_1}^2 \langle / \texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_{j_y}^2 \langle / \texttt{nessitem} \rangle$
6. $\{(\psi_j^2, \kappa_j^2), j = 1, .., q\}$ *is the set of weighted subsets with respect to* $\{\sigma_1^2, .., \sigma_q^2\}$
7. π_1 *and* π_2 *are the corresponding possibility distributions of the components respectively*

Let the **conjunctively** *merged uncertainty component be as follows.*

$$\langle\texttt{possibility}\rangle\sigma_1, .., \sigma_s\langle/\texttt{possibility}\rangle$$

where each $\sigma_k \in \{\sigma_1, .., \sigma_s\}$ *is of the form*

$$\langle\texttt{ness value} = \kappa_k\rangle\psi\langle/\texttt{ness}\rangle$$

and

$$\kappa_k = 1 - \mathsf{max}\{\pi(\phi)|\phi \in \bar{\psi}\}$$

where π *is the merged possibility distribution with the conjunctive operation*

$$\pi(\phi) = \mathsf{min}(\pi_1(\phi), \pi_2(\phi))$$

and ψ *is of the form*

$$\langle\texttt{nessitem}\rangle\phi_1\langle/\texttt{nessitem}\rangle \cdots \langle\texttt{nessitem}\rangle\phi_z\langle/\texttt{nessitem}\rangle$$

and

$$\psi = \{\phi_1, .., \phi_z\} \in \{\psi_1^1, \psi_2^1, .., \psi_p^1\} \cup \{\psi_1^2, \psi_2^2, .., \psi_q^1\}$$

Although it is possible to generate the necessity measure for all subsets once the merged possibility distribution is known, we only consider the degree of necessity of those subsets which have been specified in the original possibility-valid components. This will provide a consistent structure with the input news reports.

Definition 36 *Let the news terms* τ_1 *and* τ_2 *each denote a possibility-valid component and let* `X` *be a logical variable. The condition literal* `ConjunctivePossibility`$(\tau_1, \tau_2, \texttt{X})$ *is such that* `X` *is evaluated to* τ_3 *where* τ_3 *is the news term denoting the conjunctively merged possibility component obtained by Definition 35.*

The `ConjunctivePossibility` condition literal can be used as a condition in a fusion rule for merging pairs of structured news reports that each incorporate a possibility-valid uncertainty component.

Example 22 *Consider the following two possibility-valid components with related frame* $\Omega = \{\phi_1, \phi_2, \phi_3, \phi_4, \phi_5\}$.

```
⟨possibility⟩
   ⟨ness value = "0.4"⟩
      ⟨nessitem⟩φ1⟨/nessitem⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.5"⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
      ⟨nessitem⟩φ3⟨/nessitem⟩
      ⟨nessitem⟩φ4⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.4"⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
   ⟨/ness⟩
⟨/possibility⟩
```

```
⟨possibility⟩
   ⟨ness value = "0.3"⟩
      ⟨nessitem⟩φ1⟨/nessitem⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.5"⟩
      ⟨nessitem⟩φ1⟨/nessitem⟩
      ⟨nessitem⟩φ2⟨/nessitem⟩
      ⟨nessitem⟩φ3⟨/nessitem⟩
   ⟨/ness⟩
   ⟨ness value = "0.4"⟩
      ⟨nessitem⟩φ1⟨/nessitem⟩
      ⟨nessitem⟩φ4⟨/nessitem⟩
   ⟨/ness⟩
⟨/possibility⟩
```

For `ConjunctivePossibility`$(\tau_1, \tau_2, \texttt{X})$, *the logical variable* `X` *is evaluated to* τ_3 *where* τ_3 *is the news term denoting the following uncertainty component in the merged news report.*

⟨possibility⟩
⟨ness value = “0.4”⟩
⟨nessitem⟩ϕ_1⟨/nessitem⟩
⟨nessitem⟩ϕ_2⟨/nessitem⟩
⟨/ness⟩
⟨ness value = “0.5”⟩
⟨nessitem⟩ϕ_2⟨/nessitem⟩
⟨nessitem⟩ϕ_3⟨/nessitem⟩
⟨nessitem⟩ϕ_4⟨/nessitem⟩
⟨/ness⟩
⟨ness value = “0.4”⟩
⟨nessitem⟩ϕ_2⟨/nessitem⟩
⟨/ness⟩
⟨ness value = “0.5”⟩
⟨nessitem⟩ϕ_1⟨/nessitem⟩
⟨nessitem⟩ϕ_2⟨/nessitem⟩
⟨nessitem⟩ϕ_3⟨/nessitem⟩
⟨/ness⟩
⟨ness value = “0.4”⟩
⟨nessitem⟩ϕ_1⟨/nessitem⟩
⟨nessitem⟩ϕ_4⟨/nessitem⟩
⟨/ness⟩
⟨/possibility⟩

In this combination procedure, the merged possibility distribution, π, is not explicitly calculated. However, in practice, this distribution will be calculated and stored in the background knowledgebase. This message serves two purposes, it will be used to check any inconsistencies among multiple sources and it will be used to merge with other types of uncertain information, e.g., belfunction components, when heterogeneous uncertainty components are involved.

Here, we consider the first role of the merged possibility distribution and leave the second to the next section.

In the above example, the merged possibility distribution π gives

$$\pi(\phi_1) = 0.5, \pi(\phi_2) = 0.6, \pi(\phi_3) = 0.6, \pi(\phi_4) = 0.5, \pi(\phi_5) = 0.5$$

which is not normal, since all the values are strictly less than 1.0, although both of the original possibility distribution are normal.

The definition below aims to detect this inconsistency.

Definition 37 *Let the news terms τ_1 and τ_2 each denote a possibility-valid component and let* `X` *be a logical variable. If the merged possibility distribution π, by* `ConjunctivePossibility`$(\tau_1, \tau_2, \mathtt{X})$*, satisfies $\forall \phi, \pi(\phi) < 1$ then τ_1 and τ_2 are said to be* **potentially inconsistent***. If τ_1 and τ_2 are potentially inconsistent, then the condition literal* `PotentialPossibilityInconsistency`(τ_1, τ_2, X) *holds.*

Instead of normalizing any merge results, the conflict information presented in the merged results will be used to detect potential inconsistencies among sources. The `PotentialPossibilityInconsistency` predicate is used as an extra condition to restrict the application of the `ConjunctivePossibility` predicate defined in Definition 36 in fusion rules. When it is true, alternative merging operations, especially the disjunctive operation, will be applied to merge the input news reports.

Definition 38 *Let the following be two possibility-valid uncertainty components*

$$\langle \texttt{possibility} \rangle \sigma_1^1, .., \sigma_p^1 \langle /\texttt{possibility} \rangle$$
$$\langle \texttt{possibility} \rangle \sigma_1^2, .., \sigma_q^2 \langle /\texttt{possibility} \rangle$$

where

1. $\sigma_i^1 \in \{\sigma_1^1, .., \sigma_p^1\}$ *is in the form* $\langle \texttt{ness value} = \kappa_i^1 \rangle \psi_i^1 \langle /\texttt{ness} \rangle$
2. ψ_i^1 *is of the form* $\langle \texttt{nessitem} \rangle \phi_{i_1}^1 \langle /\texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_{i_x}^1 \langle /\texttt{nessitem} \rangle$
3. $\{(\psi_i^1, \kappa_i^1), i = 1, .., p\}$ *is the set of weighted subsets with respect to* $\{\sigma_1^1, .., \sigma_p^1\}$
4. $\sigma_j^2 \in \{\sigma_1^2, .., \sigma_q^2\}$ *is in the form* $\langle \texttt{ness value} = \kappa_j^2 \rangle \psi_j^2 \langle /\texttt{ness} \rangle$
5. ψ_j^2 *is of the form* $\langle \texttt{nessitem} \rangle \phi_{j_1}^2 \langle /\texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_{j_y}^2 \langle /\texttt{nessitem} \rangle$
6. $\{(\psi_j^2, \kappa_j^2), j = 1, .., q\}$ *is the set of weighted subsets with respect to* $\{\sigma_1^2, .., \sigma_q^2\}$
7. π_1 *and* π_2 *are the corresponding possibility distributions of the components respectively*

Let the **disjunctively** *merged uncertainty component be as follows.*

$$\langle \texttt{possibility} \rangle \sigma_1, .., \sigma_s \langle /\texttt{possibility} \rangle$$

where each $\sigma_k \in \{\sigma_1, .., \sigma_s\}$ *is of the form*

$$\langle \texttt{ness value} = \kappa_k \rangle \psi \langle /\texttt{ness} \rangle$$

and

$$\kappa_k = 1 - \mathsf{max}\{\pi(\phi) | \phi \in \bar{\psi}\}$$

where π *is the merged possibility distribution with the disjunctive operation*

$$\pi(\phi) = \mathsf{max}(\pi_1(\phi), \pi_2(\phi))$$

and ψ *is of the form*

$$\langle \texttt{nessitem} \rangle \phi_1 \langle /\texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_z \langle /\texttt{nessitem} \rangle$$

where

$$\psi = \{\phi_1, .., \phi_z\} \in \{\psi_1^1, \psi_2^1, .., \psi_p^1\} \cup \{\psi_1^2, \psi_2^2, .., \psi_q^2\} \cup \{\psi_i^1 \cup \psi_j^2 | i = 1, .., p, j = 1, .., q\}$$

Definition 39 *Let the news terms* τ_1 *and* τ_2 *each denote a possibility-valid component and let* $\texttt{X}$ *be a logical variable. The condition literal* $\texttt{DisjunctivePossibility}(\tau_1, \tau_2, \texttt{X})$ *is such that* $\texttt{X}$ *is evaluated to* τ_3 *where* τ_3 *is the news term denoting the disjunctively merged possibility component obtained by Definition 38.*

When applying `DisjunctivePossibility`(τ_1, τ_2, X) to the two possibility-valid components in Example 22, the merged possibility-valid component is:

⟨possibility⟩
  ⟨ness value = “0.4”⟩
    ⟨nessitem⟩ϕ_1⟨/nessitem⟩
    ⟨nessitem⟩ϕ_2⟨/nessitem⟩
  ⟨/ness⟩
  ⟨ness value = “0.0”⟩
    ⟨nessitem⟩ϕ_2⟨/nessitem⟩
    ⟨nessitem⟩ϕ_3⟨/nessitem⟩
    ⟨nessitem⟩ϕ_4⟨/nessitem⟩
  ⟨/ness⟩
  ⟨ness value = “0.0”⟩
    ⟨nessitem⟩ϕ_2⟨/nessitem⟩
  ⟨/ness⟩
  ⟨ness value = “0.4”⟩
    ⟨nessitem⟩ϕ_1⟨/nessitem⟩
    ⟨nessitem⟩ϕ_2⟨/nessitem⟩
    ⟨nessitem⟩ϕ_3⟨/nessitem⟩
  ⟨/ness⟩
  ⟨ness value = “0.0”⟩
    ⟨nessitem⟩ϕ_1⟨/nessitem⟩
    ⟨nessitem⟩ϕ_4⟨/nessitem⟩
  ⟨/ness⟩
  ⟨ness value = “0.4”⟩
    ⟨nessitem⟩ϕ_1⟨/nessitem⟩
    ⟨nessitem⟩ϕ_2⟨/nessitem⟩
    ⟨nessitem⟩ϕ_4⟨/nessitem⟩
  ⟨/ness⟩
  ⟨ness value = “0.5”⟩
    ⟨nessitem⟩ϕ_1⟨/nessitem⟩
    ⟨nessitem⟩ϕ_2⟨/nessitem⟩
    ⟨nessitem⟩ϕ_3⟨/nessitem⟩
    ⟨nessitem⟩ϕ_4⟨/nessitem⟩
  ⟨/ness⟩
⟨/possibility⟩

and the merged possibility distribution is:

$$\pi(\phi_1) = 1.0, \pi(\phi_2) = 1.0, \pi(\phi_3) = 0.6, \pi(\phi_4) = 0.6, \pi(\phi_5) = 0.5$$

When using the disjunctive operator, such as max, there can be many subsets that have the degree of necessity measure with 0.0 in the merged XML document, as demonstrated above. These segments will unnecessarily expand the XML code. If we revise the calculation of κ_k in Definition 38 as

$$\kappa_k = 1 - \mathsf{max}\{\pi(\phi) | \phi \in \bar{\psi}\} > 0$$

then the merged XML document will overcome this problem.

Both the conjunctive and disjunctive operations are associative. When there are more than two possibility-valid components available, we will merge two of them first and then merge the merged result with the third one and so on, until all the components have been considered.

Since inconsistency may appear at any stage when using the conjunctive operation, especially if two conflicting sources are merged before consistent sources are combined, we will leave the study of how to prioritize a list of sources for combination, e.g., based on the reliability of a source in the past, or the degree of consistency of two sources, to a future paper. A starting point for this would be measuring the degree of inconsistency ([Hun02c, Hun03, KLM03]) and further measuring the quality of uncertainty components when they have the same degree of inconsistency [HL04c].

5 Merging heterogeneous uncertainty components

In the previous section, we have provided predicates for use in conditions of fusion rules for merging information in homogeneous news reports with either probability distributions, or mass functions, or possibility measures respectively. Since structured news reports can come from various sources for various purposes, individual structured news reports may use different mechanisms to represent uncertainty information. In this section, we discuss how to merge two structured news reports with the uncertainty components in two different forms.

5.1 Merging possibility with beliefs

This subsection particularly focuses on how to merge belfunction-valid components with possibility-valid components. Since probability distributions can be regarded as special cases of mass functions, fusion rules for merging possibility-valid components and belfunction-valid components will equally apply to merging probability-valid and possibility-valid news reports.

5.1.1 Relationship between DS theory and possibility theory

In [Sha76] Chapter 10, a belief function is called a **consonant function** if its focal elements are nested. That is, if S_1, S_2 ,.., S_n are the focal elements with S_{i+1} containing more elements than S_i, then $S_1 \subset S_2 \subset .. \subset S_n$. Let Bel be a consonant function, and Pl be its corresponding plausibility function, Bel and Pl have the following properties:

$$Bel(A \cap B) = \mathsf{min}(Bel(A), Bel(B)) \texttt{ for all } A, B \subseteq \wp(\Omega).$$

$$Pl(A \cup B) = \mathsf{max}(Pl(A), Pl(B)) \texttt{ for all } A, B \subseteq \wp(\Omega).$$

These two properties are exactly the requirements of necessity and possibility measures in possibility theory. Therefore, necessity and possibility measures are special cases of belief and plausibility functions.

Shafer further defined a function, called **contour function** $f : \Omega \to [0, 1]$, for a consonant function through equation

$$f(\phi) = Pl(\{\phi\})$$

For a subset $A \subseteq \Omega$,

$$Pl(A) = \mathsf{max}_{\phi \in A} f(\phi) \tag{10}$$

Equation (10) matches the definition of possibility measures from a possibility distribution. So a contour function is a possibility distribution.

Now, we look at the properties of the corresponding plausibility function of a consonant function.

Proposition 2 *Let Bel be a consonant function. Let m and Pl be its corresponding mass and plausibility functions respectively. Let the set of focal elements be A_1, A_2,.., A_p, with $A_i \subset A_{i+1}$. Then the following properties hold:*

1. $Pl(\{\phi\}) = 1$, *for all* $\phi \in A_1$
2. $Pl(\{\phi_{i_l}\}) = Pl(\{\phi_{i_t}\})$ *for any* $\phi_{i_l}, \phi_{i_t} \in A_i \setminus A_{i-1}$ *where* $i > 1$
3. $Pl(\{\phi_l\}) > Pl(\{\phi_j\})$, *for* $\phi_l \in A_i \setminus A_{i-1}$ *and* $\phi_j \in A_{i+1} \setminus A_i$ *where* $i > 1$
4. $Pl(\{\phi\}) = 0$ *for any* $\phi \notin A_p$

Proof

Let us first prove that $Pl(\{\phi\}) = 1$, for all $\phi \in A_1$.

Since the set of focal elements are nested and A_1 is the smallest among all the focal elements, $A_1 \cap A_i = A_1$ for all the other focal elements A_i. This leads to

$$\begin{aligned} Pl(\{\phi\}) &= \Sigma\{m(A_i) | \{\phi\} \cap A_i \neq \emptyset, \phi \in A_1\} \\ &= \Sigma\{m(A_i) | i = 1, 2, .., p\} \\ &= 1 \end{aligned}$$

Next, we prove $Pl(\{\phi_{i_l}\}) = Pl(\{\phi_{i_t}\})$ for any $\phi_{i_l}, \phi_{i_t} \in A_i \setminus A_{i-1}$ where $i > 1$.

In fact, for all $\phi_j \in A_i \setminus A_{i-1}$, we have

$$\begin{aligned} Pl(\{\phi_j\}) &= \Sigma\{m(A_s) | \{\phi_j\} \cap A_s \neq \emptyset\} \\ &= \Sigma\{m(A_s) | s >= i, \texttt{since} \{\phi_j\} \cap A_t = \emptyset, \texttt{for } t < i\} \\ &= \Sigma\{m(A_s) | s = i, i+1, .., p\} \end{aligned}$$

Therefore, for any $\phi_{i_l}, \phi_{i_t} \in A_i \setminus A_{i-1}$, we have $Pl(\{\phi_{i_l}\}) = Pl(\{\phi_{i_t}\}) = \Sigma\{m(A_s) | s = i, i+1, .., p\}$.

Now, we prove the third property.

Let $\phi_l \in A_i \setminus A_{i-1}$, similar to the above calculation, we have

$$\begin{aligned} Pl(\{\phi_l\}) &= \Sigma\{m(A_s) | \{\phi_l\} \cap A_s \neq \emptyset\} \\ &= \Sigma\{m(A_s) | s >= i, \texttt{since} \{\phi_l\} \cap A_t = \emptyset, \texttt{for } t < i\} \\ &= \Sigma\{m(A_s) | s = i, i+1, .., p\} \end{aligned}$$

and for any $\phi_j \in A_{i+1} \setminus A_i$, we equally have

$$\begin{aligned} Pl(\{\phi_j\}) &= \Sigma\{m(A_s) | \{\phi_j\} \cap A_s \neq \emptyset\} \\ &= \Sigma\{m(A_s) | s >= i+1, \texttt{since} \{\phi_j\} \cap A_t = \emptyset, \texttt{for } t \leq i\} \\ &= \Sigma\{m(A_s) | s = i+1, i+2, .., p\} \end{aligned}$$

These two equations lead to

$$Pl(\{\phi_l\}) = m(A_i) + Pl(\{\phi_j\}) \texttt{ for } \phi_l \in A_i \setminus A_{i-1} \texttt{ and } \phi_j \in A_{i+1} \setminus A_i$$

which says $Pl(\{\phi_l\}) > Pl(\{\phi_j\})$ since $m(A_i) > 0$.

Finally, for any $\phi \notin A_p$, we have

$$\begin{aligned} Pl(\{\phi\}) &= \Sigma\{m(A_s) | \{\phi\} \cap A_s \neq \emptyset\} \\ &= 0 \ \ (\texttt{since } \{\phi\} \cap A_t = \emptyset, \ \texttt{for } t = 1, .., p) \end{aligned}$$

End

Example 23 *Let Bel be a consonant function and m be its mass function with the following assignment.*

$\langle\texttt{belfunction}\rangle$
$\langle\texttt{mass value} = \text{“0.3”}\rangle$
$\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle$
$\langle/\texttt{mass}\rangle$
$\langle\texttt{mass value} = \text{“0.2”}\rangle$
$\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle$
$\langle/\texttt{mass}\rangle$
$\langle\texttt{mass value} = \text{“0.1”}\rangle$
$\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_4\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_6\langle/\texttt{massitem}\rangle$
$\langle/\texttt{mass}\rangle$
$\langle\texttt{mass value} = \text{“0.3”}\rangle$
$\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_4\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_6\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_8\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_9\langle/\texttt{massitem}\rangle$
$\langle/\texttt{mass}\rangle$
$\langle\texttt{mass value} = \text{“0.1”}\rangle$
$\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_4\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_6\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_8\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_9\langle/\texttt{massitem}\rangle$
$\langle\texttt{massitem}\rangle\phi_{11}\langle/\texttt{massitem}\rangle$
$\langle/\texttt{mass}\rangle$
$\langle/\texttt{belfunction}\rangle$

The focal elements are $A_1 = \{\phi_1, \phi_2\}$, $A_2 = \{\phi_1, \phi_2, \phi_3\}$, $A_3 = \{\phi_1, \phi_2, \phi_3, \phi_4, \phi_6\}$, $A_4 = \{\phi_1, \phi_2, \phi_3, \phi_4, \phi_6, \phi_8, \phi_9\}$, and $A_5 = \{\phi_1, \phi_2, \phi_3, \phi_4, \phi_6, \phi_8, \phi_9, \phi_{11}\}$.

From this mass assignment, the degree of plausibility on each singleton set can be calculated as follows.

$$Pl(\{\phi_1\}) = Pl(\{\phi_2\}) = 1$$

$$Pl(\{\phi_3\}) = 0.7$$

$$Pl(\{\phi_4\}) = Pl(\{\phi_6\}) = 0.5$$

$$Pl(\{\phi_8\}) = Pl(\{\phi_9\}) = 0.4$$

$$Pl(\{\phi_{11}\}) = 0.1$$

And for any other element in the frame of discernment, its degree of plausibility is 0.

This proposition reveals several interesting points in relation to the values of plausibility function on singleton sets. First of all, it says that all the elements in the smallest focal element set must have the maximum degree of plausibility value. So any element, ϕ with value $Pl(\{\phi\}) = 1$, declares that ϕ must be in the smallest focal element. Secondly, for any two elements ϕ_i and ϕ_j if they have the same degree of plausibility, that is $Pl(\{\phi_i\}) = Pl(\{\phi_j\})$, then these two elements always appear together in any focal element set. Thirdly, all those elements with positive Pl values are partitioned into subsets based on their degrees of plausibility, with the elements having the maximum degree in the innermost layer, and the elements having the lowest degree in the outmost layer.

We now investigate how to recover a mass function from a consonant function without exhausting all the subsets of a frame of discernment.

Proposition 3 [2] *Let Bel be a consonant function and Pl be its plausibility function where $Pl(\{\phi\})$ is known for each ϕ when $Pl(\{\phi\}) > 0$ for frame of discernment Ω. Let B_1, B_2,.., B_p be disjoint subsets of Ω such that $Pl(\{\phi_i\}) = Pl(\{\phi_j\})$ when both $\phi_i, \phi_j \in B_i$, and $Pl(\{\phi_i\}) > Pl(\{\phi_j\})$ if $\phi_i \in B_i$ and $\phi_j \in B_{i+1}$, then the following properties hold:*

1. *Let $A_i = \cup\{B_j | j = 1, .., i\}$ for $i = 1, 2, .., p$, then subsets $A_1, A_2, .., A_n$ are nested and are focal elements of Bel*
2. *Let $m(A_i) = Pl(\{\phi_i\}) - Pl(\{\phi_j\})$ where $\phi_i \in B_i$ and $\phi_j \in B_{i+1}$ for $i = 1, .., p-1$. Let $m(A_p) = Pl(\{\phi\})$ where $\phi \in B_p$. Then m is the corresponding mass function on focal elements A_i*

Proof

It is easy to prove that subsets A_i are nested, since $A_i \subset A_{i+1}$ is true based on the construction of each A_i.

Also,

$$Pl(\{\phi_i\}) = Pl(\{\phi_j\}) \texttt{ when } \phi_i, \phi_j \in (A_{i+1} \setminus A_i) = B_i$$

and

$$Pl(\{\phi_i\}) > Pl(\{\phi_j\}) \texttt{ when } \phi_i \in B_i = A_i \setminus A_{i-1} \texttt{ and } \phi_j \in B_{i+1} = A_{i+1} \setminus A_i$$

Furthermore, $Pl(\{\phi\}) = 0$ for any other $\phi \notin A_p$, and $Pl(\{\phi_i\}) = 1$ for $\phi_i \in A_1 = B_1$, because (1) $Pl(\{\phi_i\})$ must have the highest degree of plausibility, and (2) if $Pl(\{\phi_i\}) < 1$ then there exists a focal element A_l such that $A_1 \cap A_l = \emptyset$ which contradicts with the assumption that Bel is consonant.

Therefore, all the properties in Proposition 2 are true for Pl on A_i, so A_i are the focal elements.

Next, to prove that m is the corresponding mass function, we need to prove that $\Sigma_i m(A_i) = 1$.

Based on the definition of m, we have

$$m(A_p) = Pl(\{\phi_j\}) \texttt{ when } \phi_j \in B_p$$

and

$$m(A_{p-1}) = Pl(\{\phi_i\}) - Pl(\{\phi_j\}) = Pl(\{\phi_i\}) - m(A_p) \texttt{ for } \phi_i \in B_{p-1} \texttt{ and } \phi_j \in B_p$$

So

$$Pl(\{\phi_i\}) = m(A_p) + m(A_{p-1}) \texttt{ for } \phi_i \in B_{p-1}$$

which leads to

$$m(A_{p-2}) = Pl(\{\phi_i\}) - Pl(\{\phi_j\}) = Pl(\{\phi_i\}) - (m(A_p) + m(A_{p-1})) \texttt{ for } \phi_i \in B_{p-2} \texttt{ and } \phi_j \in B_{p-1}$$

[2] Whilst Propositions 2 and 3 may be regarded as established facts, we have added proofs to make the paper self-contained.

This procedure can be repeated for all A_i, until

$$m(A_1) = Pl(\{\phi_i\}) - Pl(\{\phi_j\}) = Pl(\{\phi_i\}) - (\Sigma_j m(A_j)) \texttt{ for } \phi_i \in B_1 \texttt{ and } j = 2, .., p$$

which gives

$$Pl(\{\phi_i\}) = m(A_1) + \{(\Sigma_j m(A_j)) | j = 2, .., p\} = \Sigma_{i=1}^{p} m(A_i)$$

Since $Pl(\{\phi_i\}) = 1$ for $\phi_i \in B_1$, we eventually have $1 = \Sigma_{i=1}^{p} m(A_i)$.

End

Looking at Example 23 again, it is not difficult to see that $B_1 = \{\phi_1, \phi_2\}$, $B_2 = \{\phi_3\}$, $B_3 = \{\phi_4, \phi_6\}$, $B_4 = \{\phi_8, \phi_9\}$ and $B_5 = \{\phi_{11}\}$. Also, $A_i = \cup_j B_j$, and $m(A_i) = Pl(\{\phi_i\}) - Pl(\{\phi_j\})$ where $\phi_i \in B_i$ and $\phi_j \in B_{i+1}$ for $1 \leq i \leq n$. For example,

$$m(A_3) = Pl(\{\phi_4\}) - Pl(\{\phi_8\}) = 0.5 - 0.4 = 0.1$$

Lemma 1 *Let π be a possibility distribution on frame of discernment Ω and is normal. Let B_1, B_2,.., B_p be disjoint subsets of Ω such that $\pi(\phi_i) = \pi(\phi_j)$ when both $\phi_i, \phi_j \in B_i$, and $\pi(\phi_i) > \pi(\phi_j)$ if $\phi_i \in B_i$ and $\phi_j \in B_{i+1}$, then the following properties hold:*

1. *Let $A_i = \cup\{B_j | j = 1, .., i\}$ for $i = 1, 2, .., p$, then subsets $A_1, A_2, .., A_p$ are nested*
2. *Let $m(A_i) = \pi(\phi_i) - \pi(\phi_j)$ where $\phi \in B_i$ and $\phi_j \in B_{i+1}$ for $i = 1, .., p-1$. Let $m(A_p) = \pi(\phi)$ where $\phi \in B_p$. Then m is a mass function on focal elements A_i*
3. *Let Bel be the belief function corresponding to m defined above, then Bel is a consonant function*

This Lemma can be easily proved based on Proposition 3 and the fact that a normal possibility distribution is a contour function $f(\phi) = Pl(\{\phi\})$. This Lemma says that from any possibility distribution π which is normal, it is possible to recover its corresponding mass function.

The nature of Lemma 1 was first observed in [DP82] where the relationship between the possibility theory and belief functions was discussed. This relationship was further referred to in several papers subsequently ([DP88b, DP98b, DNP00]). Here we state the Lemma again to make the paper self-contained.

Definition 40 *Let the following be a possibility-valid component*

$$\langle \texttt{possibility} \rangle \sigma_1, .., \sigma_p \langle / \texttt{possibility} \rangle$$

where

1. $\sigma_i \in \{\sigma_1, .., \sigma_p\}$ *is in the form* $\langle \texttt{ness value} = \kappa_i \rangle \psi_i \langle / \texttt{ness} \rangle$
2. ψ_i *is of the form* $\langle \texttt{nessitem} \rangle \phi_{i_1} \langle / \texttt{nessitem} \rangle \cdots \langle \texttt{nessitem} \rangle \phi_{i_x} \langle / \texttt{nessitem} \rangle$

Let the corresponding possibility distribution obtained by Definition 32 be π which is normal. For all $\phi \in \Omega$ when $\pi(\phi) > 0$, let B_1, B_2,.., B_n be disjoint subsets such that $\pi(\phi_i) = \pi(\phi_j)$ when both $\phi_i, \phi_j \in B_i$, and $\pi(\phi_i) > \pi(\phi_j)$ if $\phi_i \in B_i$ and $\phi_j \in B_{i+1}$.

Let the converted belfunction-valid component be

$$\langle \texttt{belfunction} \rangle \sigma'_1, .., \sigma'_n \langle / \texttt{belfunction} \rangle$$

where each $\sigma'_k \in \{\sigma'_1, .., \sigma'_n\}$ *is of the form*

$$\langle\texttt{mass value} = \kappa'_k\rangle\psi'_k\langle/\texttt{mass}\rangle$$

and ψ'_k *is of the form*

$$\langle\texttt{messitem}\rangle\phi_{k_1}\langle/\texttt{messitem}\rangle \cdots \langle\texttt{messitem}\rangle\phi_{k_y}\langle/\texttt{messitem}\rangle$$

where

$$\{\phi_{k_1}, .., \phi_{k_y}\} = \bigcup B_j \texttt{ for } j = 1, 2, .., k$$

and

$$\kappa'_k = \begin{cases} \pi(\phi_i) - \pi(\phi_j) & \texttt{where } \phi_i \in B_k \texttt{ and } \phi_j \in B_{k+1} \texttt{ when } k < n \\ \pi(\phi_i) & \texttt{where } \phi_i \in B_n \texttt{ when } k = n \end{cases}$$

Definition 41 *Let the news term* τ *be a possibility-valid component and* π *be its corresponding possibility distribution which is normal. Let* X *be a logical variable. The condition literal* $\texttt{PossibilityBelfunction}(\tau, X)$ *is such that* X *is evaluated to* τ' *where* τ' *is the news term denoting a belfunction-valid component obtained by Definition 40.*

Example 24 *Let* τ *be the news term denoting the possibility-valid component on the right-hand side of Example 20 and let* π_2 *be its corresponding possibility distribution with details*

$$\pi_2(\phi_1) = 0.7, \pi_2(\phi_2) = 1.0, \pi_2(\phi_3) = 0.8, \pi_2(\phi_4) = 0.7$$

then we have

$$Pl(\{\phi_1\}) = 0.7, Pl(\{\phi_2\}) = 1.0, Pl(\{\phi_3\}) = 0.8, Pl(\{\phi_4\}) = 0.7$$

since a possibility distribution is a contour function (of a consonant function) which defines plausibility values on singletons. Therefore, the disjoint subsets for this consonant function are

$$B_1 = \{\phi_2\}, B_2 = \{\phi_3\}, B_3 = \{\phi_1, \phi_4\}$$

and the corresponding focal elements are

$$A_1 = B_1, A_2 = B_1 \cup B_2, A_3 = B_1 \cup B_2 \cup B_3$$

The condition literal $\texttt{PossibilityBelfunction}(\tau, X)$ *generates a belfunction-valid component denoted by* X *as below which defines a mass function* m *with* $m(A_1) = 0.2, m(A_2) = 0.1, m(A_3) = 0.7$*:*

$$\begin{array}{l}
\langle\texttt{belfunction}\rangle \\
\quad\langle\texttt{mass value} = \text{“0.2”}\rangle \\
\qquad\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle \\
\quad\langle/\texttt{mass}\rangle \\
\quad\langle\texttt{mass value} = \text{“0.1”}\rangle \\
\qquad\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle \\
\qquad\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle \\
\quad\langle/\texttt{mass}\rangle \\
\quad\langle\texttt{mass value} = \text{“0.7”}\rangle \\
\qquad\langle\texttt{massitem}\rangle\phi_1\langle/\texttt{massitem}\rangle \\
\qquad\langle\texttt{massitem}\rangle\phi_2\langle/\texttt{massitem}\rangle \\
\qquad\langle\texttt{massitem}\rangle\phi_3\langle/\texttt{massitem}\rangle \\
\qquad\langle\texttt{massitem}\rangle\phi_4\langle/\texttt{massitem}\rangle \\
\quad\langle/\texttt{mass}\rangle \\
\langle/\texttt{belfunction}\rangle
\end{array}$$

5.2 Merging multiple heterogeneous uncertainty components

Following the discussion above regarding belief functions and possibility/necessity measures and the conversion of a possibility distribution into a mass function, as well as the view we take in this paper that a probability distribution is a special case of mass function, we are ready to demonstrate the merging of multiple structured news reports with heterogeneous forms of uncertainty representations. The umbrella for this unified approach to merging is the Dempster's combination rule, used after non-mass function uncertainty components are converted into mass function components. We will also show how this merge process can be efficiently performed when one of the news reports to be merged has a probability-valid component.

Example 25 *Let the three structured news reports with uncertainty components be as shown below and be denoted by new terms τ_1, τ_2 and τ_3 respectively.*

⟨probability⟩
⟨prob value = "0.4"⟩ϕ_1⟨/prob⟩
⟨prob value = "0.5"⟩ϕ_2⟨/prob⟩
⟨prob value = "0.1"⟩ϕ_3⟨/prob⟩
⟨/probability⟩

⟨belfunction⟩
⟨mass value = "0.4"⟩
⟨massitem⟩ϕ_1⟨/massitem⟩
⟨massitem⟩ϕ_2⟨/massitem⟩
⟨/mass⟩
⟨mass value = "0.6"⟩
⟨massitem⟩ϕ_3⟨/massitem⟩
⟨/mass⟩
⟨/belfunction⟩

⟨possibility⟩
⟨ness value = "0.6"⟩
⟨nessitem⟩ϕ_1⟨/nessitem⟩
⟨nessitem⟩ϕ_2⟨/nessitem⟩
⟨/ness⟩
⟨ness value = "0.6"⟩
⟨nessitem⟩ϕ_1⟨/nessitem⟩
⟨nessitem⟩ϕ_4⟨/nessitem⟩
⟨/ness⟩
⟨/ness value = "1.0"⟩
⟨nessitem⟩ϕ_1⟨/nessitem⟩
⟨nessitem⟩ϕ_3⟨/nessitem⟩
⟨/ness⟩
⟨/possibility⟩

We first convert news term τ_1 into a belfunction-valid component, denoted by τ_1'. We then merge it with belfunction-valid component τ_2 using $\texttt{Dempster}(\tau_1', \tau_2, \texttt{X})$ *resulting in an uncertainty component as follows, denoted as τ_2'.*

⟨belfunction⟩
⟨mass value = "0.381"⟩
⟨massitem⟩ϕ_1⟨/massitem⟩
⟨/mass⟩
⟨mass value = "0.476"⟩
⟨massitem⟩ϕ_2⟨/massitem⟩
⟨/mass⟩
⟨mass value = "0.143"⟩
⟨massitem⟩ϕ_3⟨/massitem⟩
⟨/mass⟩
⟨/belfunction⟩

Based on Definition 40, the belfunction-valid component converted from the possibility-valid component τ_3, denoted as τ_3', is derived as

```
⟨belfunction⟩
  ⟨mass value = "0.6"⟩
    ⟨massitem⟩φ1⟨/massitem⟩
  ⟨/mass⟩
  ⟨mass value = "0.4"⟩
    ⟨massitem⟩φ1⟨/massitem⟩
    ⟨massitem⟩φ3⟨/massitem⟩
  ⟨/mass⟩
⟨/belfunction⟩
```

Finally, merging τ_2' with τ_3' by $\texttt{Dempster}(\tau_2', \tau_3', X)$*, we obtain the merged belfunction-valid component that carries the impact of all the uncertain information.*

```
⟨belfunction⟩
  ⟨mass value = "0.8695"⟩
    ⟨massitem⟩φ1⟨/massitem⟩
  ⟨/mass⟩
  ⟨mass value = "0.1305"⟩
    ⟨massitem⟩φ3⟨/massitem⟩
  ⟨/mass⟩
⟨/belfunction⟩
```

This final output strongly suggests that ϕ_1 is the most likely outcome after considering all three news reports. Since Dempster's rule is associative, any sequence of merging would produce the same result as we have obtained above.

It should be noted that if there is a probability-valid component among multiple heterogeneous uncertainty components, the final outcome will have degrees of uncertainty on singleton sets rather than on subsets, as proved in [Sha76].

Let Bel_1 be a belfunction-valid component where every focal element of Bel contains only a single element, it is in fact a probability-valid component. Let Bel_2 be a belief function and Pl_2 be its plausibility function. The combination of these two belief functions results in a belfunction-valid component that is a probability-valid component and it can be obtained by the following simple equation [Sha76]:

$$Bel(\{\phi\}) = kBel_1(\{\phi\})Pl_2(\{\phi\}) \tag{11}$$

where

$$k = (\Sigma_{\phi \in \Omega} Bel_1(\{\phi\})Pl_2(\{\phi\}))^{-1}$$

Equation (11) can be rewritten as

$$Bel(\{\phi\}) = km_1(\{\phi\})Pl_2(\{\phi\}) \tag{12}$$

since $m_1(\{\phi\}) = Bel_1(\{\phi\})$ when Bel_1 has focal elements with only singleton sets.

When a probability-valid component is to be merged with a belfunction-valid component (or a converted one from a possibility-valid component), we use the following rule, instead of condition literal $\texttt{Dempster}(\tau_1, \tau_2, X)$, to merge them to simplify the calculation.

Definition 42 *Let the following be a probability-valid and a belfunction-valid uncertainty components respectively.*

$$\langle\texttt{probability}\rangle\sigma_1^1, .., \sigma_p^1\langle/\texttt{probability}\rangle$$

$$\langle\texttt{belfunction}\rangle\sigma_1^2,..,\sigma_q^2\langle/\texttt{belfunction}\rangle$$

where

1. $\sigma_i^1 \in \{\sigma_1^1,..,\sigma_p^1\}$ *is of the form* $\langle\texttt{prob value} = \kappa_i^1\rangle\phi_i^1\langle/\texttt{prob}\rangle$*, and* $\phi_i^1 \in \Omega$.
2. $\sigma_j^2 \in \{\sigma_1^2,..,\sigma_q^2\}$ *is of the form* $\langle\texttt{mass value} = \kappa_j^2\rangle\psi_i^2\langle/\texttt{mass}\rangle$
3. ψ_j^2 *is of the form* $\langle\texttt{massitem}\rangle\phi_{j_1}^2\langle/\texttt{massitem}\rangle\cdots\langle\texttt{massitem}\rangle\phi_{j_y}^2\langle/\texttt{massitem}\rangle$

Let the **combined belfunction-valid component** *be*

$$\langle\texttt{belfunction}\rangle\sigma_1,..,\sigma_s\langle/\texttt{belfunction}\rangle$$

where each $\sigma_k \in \{\sigma_1,..,\sigma_s\}$ *is of the form* $\langle\texttt{mass value} = \kappa_k\rangle\phi_k\langle/\texttt{mass}\rangle$ *and*

$$\exists\phi_i^1 \in \{\phi_1^1,..,\phi_p^1\} \texttt{ such that } \phi_i^1 = \phi_k$$

and

$$\kappa_k = k\kappa_i^1 Pl(\{\phi_k\})$$

where Pl *is the plausibility function for the belfunction-valid component with*

$$Pl(\{\phi\}) = \Sigma\{\kappa_j^2 | \phi \in \{\phi_{j_1}^2,..,\phi_{j_y}^2\}\} \texttt{ for each } \phi \in \Omega,$$

and

$$k = (\Sigma\{\kappa_i^1 Pl(\{\phi_i\}) | \phi_i \in \{\phi_1^1,..,\phi_p^1\})^{-1}$$

It is meaningful to calculate those $Bel(\{\phi\})$ only when $\{\phi\}$ is a focal element in m_1. Otherwise, $\{\phi\}$ is not a focal element in the combined belief function. We have incorporated this fact in the above definition to further reduce the calculation.

To differentiate this combination procedure from that in Definition 26, we name the following condition literal `BayesianMerge`.

Definition 43 *Let the news term* τ_1 *denote a probability-valid component and news term* τ_2 *denote a belfunction-valid component. Let* X *be a logical variable. The condition literal* $\texttt{BayesianMerge}(\tau_1, \tau_2, X)$ *is such that* X *is evaluated to* τ *where* τ *is the news term denoting a belfunction-valid component obtained by Definition 42.*

If a number of news terms have been selected for answering a user's query each with an uncertainty valid component as we have defined in previous sections, news terms with homogeneous uncertainty valid components are merged using appropriate fusion rules first before heterogeneous uncertainty components are merged. This is especially so for possibility-valid components, since a converted belfunction-valid component from a merged (either conjunctive or disjunctive) possibility distribution is not guaranted to be the same as the combined belfunction-valid component from two converted belfunction-valid components. That is, in general

$$\texttt{ConjunctiveMerge}(\tau_1, \tau_2, X) \wedge \texttt{PossibilityBelfunction}(X, X')$$

is not equivalent to

$$\texttt{PossibilityBelfunction}(\tau_1, X_1) \wedge \texttt{PossibilityBelfunction}(\tau_2, X_2) \wedge \texttt{Dempster}(X_1, X_2, X')$$

Therefore, it is highly possible that three heterogeneous news terms would be the last to be merged, as illustrated in Example 25.

During a multiple-step merge process, a higher degree of conflict among homogeneous news terms will be resolved using other techniques. For example, if the numerical values are highly conflicting, then we could use priorities over sources or to eliminate or discount on less reliable ones, or we could use some other background knowledge to resolve the conflict before proceeding. However, these are not the topics being addressed in this paper and we shall leave them to later papers following proposals for aggregation in [HS04].

6 Discussion

The approach of fusion rules suggests an implementation based on existing automated reasoning technology and on XML programming technology. Once information is in the form of XML documents, a number of technologies for managing and manipulating information in XML are available. We have developed a prototype implementation in Java for executing fusion rules that are marked up in FusionRuleML and constructing the merged reports [HS03, HS04]. Background knowledge is handled in a Prolog system and is queried by the Java implementation.

The extension of the fusion rules approach to modelling and reasoning with uncertain information on textentries in this paper provides a general framework to represent evidence and knowledge with uncertainty. With the increasing use of XML for modelling variety types of data for e-science, there is potentially a wide range of applications of the framework beyond news reports. For instance, the framework has been used to represent a large number of pieces of evidence as input for an efficient algorithm on evidence combination [LHA03].

To summarize, we have made the following novel contributions in the paper:

- A unified modelling method is presented in this paper for representing uncertain information in three types of uncertainty formalism, together with conversions among them. This method enables users to encode these types of uncertainty conveniently.
- Merging procedures (or predicates) are defined to merge multiple news reports involving either the same type of uncertainty formalisms or heterogenous ones. Although the relationship between belief functions, especially consonant belief functions and possibility theory has been discussed in many research papers, the combination (or merging process) of these two types of uncertain information in practice does not seem to have been reported. This paper presents a tool to model them and convert one type to another to facilitate the merging process. From this point of view, this is a new contribution of the paper.
- Inconsistency detection methods are proposed to detect potential contradictions of merging. The detection of inconsistencies among mass functions is a derivation of the well known perception in DS theory that the mass value assigned to the empty set before normalization represents the degree of conflict (inconsistency). We adopt this perception in the paper and use an extra predicate to detect potential conflicts when $m(\emptyset)$ is large. Then an additional predicate is developed to veto any wrong conclusions about a potential inconsistency. These two additional predicates are a new contribution to conflict detection in DS theory.

The merging of uncertain information on subtrees as well as on different levels of granularities of concepts are both discussed in [HL04a, HL04b]. The combination of dependent sources of evidence and the discounting of less reliable sources when inconsistencies exist will be studied in future papers. At present, we assume that evidence has always been given on a set of mutually exclusive and exhaustive values. Assigning and reasoning with evidence on elements that are not mutually exclusive will also be examined in the future.

We also need to investigate confluence issues for merging uncertain information using fusion rules. Merging operators for uncertainty measures are not always associative and sources of information can be conflicting. This lack of confluence is in some sense incompatible with declarative programming (while a strong point of rule programming is that, ideally, rules should be considered independent of one another). One approach we are exploring to address this is to analyse the nature of the inconsistency arising in the different sources, and using this analysis to prioritise the sources to be merged.

Another issue for further investigation is our approach to merging several uncertainty components with different uncertainty measures. For this, we have suggested merging separately the possibility-valid components and belfunction-valid compnents, and then turning the resulting possibility-valid component into a belfunction-valid one so as to be able to perform a last merging step. The rational here is to apply the most suitable merging operators, given multiple uncertainty components, in order to preserve the consistency within each theory. It is also because that although a possibility-valid component can be converted into a consonant function, and two such converted belief functions can be combined with Dempster's combination rule, the combined result is no longer a consonant function [DP82]. On the other hand, if we merge the original possibility-valid components as they are, the merged result (after normalization) is a consonant function.

However it may be possible to compare the resulting uncertainty measure above with the one obtained by turning every possibility-valid component into a belfunction-valid one during a preliminary step, then merging all the belfunction-valid components using Dempster's rule. One issue is whether in practice, these two sequences of merging would be much different.

In [NJ02], a probabilistic XML model was presented to deal with information with uncertainty that was in the form of probabilities. Using this model, we can construct an XML report as in Figure 1. Two types of probability assignments are distinguished, mutually exclusive or not mutually exclusive. For the first type, probabilities are assigned to single atoms where only one of these atoms can be true, and the total sum of probability values is less than or equal to 1 (as for $\langle\texttt{precipitation}\rangle$). For the second type, two single atoms can be compatible, so the total sum of probabilities can be greater than 1 (as for $\langle\texttt{cities}\rangle$).

This model allows probabilities to be assigned to multiple granularities. When this occurs, the probability of an element is true is conditioned upon the existence of its parent (with probability), and so on until up to the root of the tree. For example, if we would like to know the probability of `sunny` in `London`, we have

$$
\begin{aligned}
&\texttt{Prob}(\texttt{precipitation} = \texttt{sunny} \wedge \texttt{cityName} = \texttt{London})\\
&= \texttt{Prob}(\texttt{precipitation} = \texttt{sunny}) * \texttt{Prob}(\texttt{cityName} = \texttt{London})\\
&\qquad\qquad * \texttt{Prob}(\texttt{precipitation} = \texttt{sunny} \wedge \texttt{cityName} = \texttt{London} \mid \texttt{city}) * \texttt{Prob}(\texttt{city} \mid \texttt{cities})\\
&\qquad\qquad * \texttt{Prob}(\texttt{cities} \mid \texttt{report}) * \texttt{Prob}(\texttt{report})\\
&= 0.1 * 1.0 * 0.7 * 1.0 * 1.0 * 1.0 = 0.07
\end{aligned}
$$

Therefore, the probability associated with a textentry (at any level) is treated as the conditional probability under its parent. A query is answered by tracing the relevant branches with the textentries specified by the query, and calculating probabilities using the conditional probabilities along these branches. These derived probabilties are then either multiplied or added depending on whether the "and" or the "or" operation are used in the original query. For instance, the query "London is either sunny or rain on 19/3/02" is evaluated as:

```
⟨report⟩
 ⟨source⟩TV1⟨/source⟩
 ⟨date⟩19/3/02⟨/date⟩
 ⟨cities⟩
  ⟨city Prob = “0.7”⟩
     ⟨cityName⟩London⟨/cityName⟩
     ⟨precipitation⟩
      ⟨Dist type = “mutually − exclusive”⟩
       ⟨Val Prob = “0.1”⟩sunny⟨/Val⟩
       ⟨Val Prob = “0.7”⟩rain⟨/Val⟩
      ⟨/Dist⟩
     ⟨/precipitation⟩
  ⟨/city⟩
  ⟨city Prob = “0.4”⟩
     ⟨cityName⟩GreaterLondon⟨/cityName⟩
     ⟨precipitation⟩
      ⟨Dist type = “mutually − exclusive”⟩
       ⟨Val Prob = “0.2”⟩sunny⟨/Val⟩
       ⟨Val Prob = “0.6”⟩rain⟨/Val⟩
      ⟨/Dist⟩
     ⟨/precipitation⟩
  ⟨/city⟩
 ⟨/cities⟩
⟨/report⟩
```

Figure 1: An XML report using the framework in ProTDB [NJ02].

$$
\begin{aligned}
&\texttt{Prob}(\texttt{cityName} = \texttt{London} \wedge ((\texttt{precipitation} = \texttt{sunny}) \vee (\texttt{precipitation} = \texttt{rain})))\\
&= \texttt{Prob}(\texttt{cityName} = \texttt{London}) * \texttt{Prob}(\texttt{precipitation} = \texttt{sunny})\\
&\qquad *\texttt{Prob}(\texttt{cityName} = \texttt{London} \wedge \texttt{precipitation} = \texttt{sunny} \mid \texttt{city}) * \texttt{Prob}(\texttt{city} \mid \texttt{cities})\\
&\qquad *\texttt{Prob}(\texttt{cities} \mid \texttt{report}) * \texttt{Prob}(\texttt{report})\\
&\quad +\texttt{Prob}(\texttt{cityName} = \texttt{London}) * \texttt{Prob}(\texttt{precipitation} = \texttt{rain})\\
&\qquad *\texttt{Prob}(\texttt{cityName} = \texttt{London} \wedge \texttt{precipitation} = \texttt{rain} \mid \texttt{city}) * \texttt{Prob}(\texttt{city} \mid \texttt{cities})\\
&\qquad *\texttt{Prob}(\texttt{cities} \mid \texttt{report}) * \texttt{Prob}(\texttt{report})\\
&= (1.0 * 0.1 * 0.7 * *1.0 * 1.0 * 1.0) + (1.0 * 0.7 * 0.7 * 1.0 * 1.0 * 1.0) = 0.07 + 0.49 = 0.56.
\end{aligned}
$$

The main advantage of this model is that it allows probabilities to be assigned to multiple granularities and provides a means to calculate the joint probability from them. However, it does not merge multiple probabilistic XML on the same issue. On the contrary, our uncertainty XML model focuses on multiple XML datasets and provides a set of means to merge opinions with uncertainty from different sources. We have not yet in this paper considered uncertainties at non-leaf level. In this sense, the research in [NJ02] and ours is complementary to each other. However, since we allow uncertain information to be specified in a variety of forms other than just probabilities, dealing with multiple granular uncertainty information will not be as straightforward as in the case for probabilities only. We will focus on this issue in a forthcoming paper.

Our logic-based approach differs from other logic-based approaches for handling inconsistent information such as belief revision theory (e.g. [Gar88, DP98a, KM91, LS98]) and knowledgebase merging (e.g. [KP98, BKMS92]). These proposals are too simplistic in certain respects for handling news reports. Each of them has one or more of the following weaknesses: (1) One-dimensional preference ordering over sources of information — for news reports we require finer-grained preference orderings; (2) Primacy of updates in belief revision — for news reports, the newest reports are not necessarily the best reports; and (3) Weak merging based on a meet operator — this causes unnecessary loss of information. Furthermore, none of these proposals incorporate actions on inconsistency or context-dependent rules specifying the information that is to be incorporated in the merged information, nor do they offer a route for specifying how merged reports should be composed.

Other logic-based approaches to fusion of knowledge include the KRAFT system and the use of Belnap's four-valued logic. The KRAFT system uses constraints to check whether information from heterogeneous sources can be merged [PHG$^+$99, HG00]. If knowledge satisfies the constraints, then the knowledge can be used. Failure to satisfy a constraint can be viewed as an inconsistency, but there are no actions on inconsistency. In contrast, Belnap's four-valued logic uses the values "true", "false", "unknown" and "inconsistent" to label logical combinations of information (see for example [LSS00]). However, this approach does not provide actions in case of inconsistency.

Merging information is also an important topic in database systems. A number of proposals have been made for approaches based in schema integration (e.g. [PM98]), the use of global schema (e.g. [GM99]), and conceptual modelling for information integration based on description logics [CGL$^+$98b, CGL$^+$98a, FS99, PSB$^+$99, BCVB01]. These differ from our approach in that they do not seek an automated approach that uses domain knowledge for identifying and acting on inconsistencies. Heterogeneous and federated database systems also could be relevant in merging multiple news reports, but they do not identify and act on inconsistency in a context-sensitive way [SL90, Mot96, CM01], though there is increasing interest in bringing domain knowledge into the process (e.g. [Cho98, SO99]).

Our approach also goes beyond other technologies for handling news reports. The approach of wrappers offers a practical way of defining how heterogeneous information can be merged (see for example [HGNY97, Coh98, SA99]). However, there is little consideration of problems of conflicts arising between sources. Our approach therefore goes beyond these in terms of formalizing reasoning with inconsistent information and using this to analyse the nature of the news report and for formalizing how we can act on inconsistency.

Acknowledgement The authors are grateful to Philippe Smets, Didier Dubois, Henri Prade, and Salem Benferhat for helpful comments and discussions when developing the ideas in the paper. The authors are also grateful to the anonymous referees who have provided numerous suggestions to improve the paper.

References

[AS01] A Ayoun and Ph Smets. Data association in muti-target detection using the transferable belief model. *Int. J. of Intelligent Systems*, Vol 16 (10):1167-1182, 2001.

[BCVB01] S Bergamaschi, S Castano, M Vincini, and D Beneventano. Semantic integration of heterogeneous information sources. *Data and Knowledge Engineering*, 36:215–249, 2001.

[BDP97] S Benferhat, D Dubois, and H Prade. From semantic to syntactic appraoch to information combination in possibilitic logic. In *Aggregation and Fusion of Imperfect Information*, 141-151, Bernadette Bouchon-Meunier (Ed.). Physica Verlag, 1997.

[BDKP00] S Benferhat, D Dubois, S Kaci, and H Prade. Encoding classical fusion in ordered knowledge bases framework. In *Linkping Electronic Articles in Computer and Information Science*, Vol. 5, No. 027, 2000.

[BGP92] D Barbara, H Garcia-Molina, and D Porter. The management of probabilistic data. *IEEE Transactions on Knowledge and Data Engineering*, 4:487–502, 1992.

[BKMS92] C Baral, S Kraus, J Minker, and V Subrahmanian. Combining knowledgebases of consisting of first-order theories. *Computational Intelligence*, 8:45–71, 1992.

[BP91] P Bosc and O Pivert. Some algorithms for evalutating fuzzy relational queries. In *Uncertainty in Knowledge Bases (IPMU'90)*, LNCS 521, 431–442. Springer, 1991.

[CGL+98a] D Calvanese, G De Giacomo, M Lenzerini, D Nardi, and R Rosati. Description logic framework for information integration. In *Proceedings of the 6th Conference on the Principles of Knowledge Representation and Reasoning (KR'98)*, 2–13. Morgan Kaufmann, 1998.

[CGL+98b] D Calvanese, G De Giacomo, M Lenzerini, D Nardi, and R Rosati. Source integration in data warehousing. In *Proceedings of the 9th International Workshop on Database and Expert Systems (DEXA'98)*, 192–197. IEEE Press, 1998.

[Cho98] L Cholvy. Reasoning with data provided by federated databases. *Journal of Intelligent Information Systems*, 10:49–80, 1998.

[CL96] J Cowie and W Lehnert. Information extraction. *Communications of the ACM*, 39:81–91, 1996.

[CM01] L Cholvy and S Moral. Merging databases: Problems and examples. *International Journal of Intelligent Systems*, 10:1193–1221, 2001.

[Coh98] W Cohen. A web-based information system that reasons with structured collections of text. In *Proceedings of Autonomous Agents'98*,400-407, 1998.

[CP87] R. Cavallo and M. Pittarelli. The theory of probabilistic databases. In *Proceedings of VLDB'87*, 71–81, 1987.

[Dem67] A P Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annuals of Mathematics and Statistics*, 38:325–339, 1967.

[DP82] D Dubois and H Prade. *On several representations of an uncertain body of evidence*. *Fuzzy Information and Decision Processes*, 167-181, Gupta and Sanchez (Eds.). North-Holland Publishing Company, 1982.

[DP87] D Dubois and H Prade. *The principle of minimum specificity as a basis for evidential reasoning*. *Uncertainty in Knowledge-Based Systems*, 75-84, Bouchon and Yager (Eds.). Springer-Verlag, 1987.

[DP88a] D Dubois and H Prade. *Possibility theory: An approach to the computerized processing of uncertainty*. Plenum Press, 1988.

[DP88b] D Dubois and H Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4:244-264. 1988

[DP98a] D Dubois and H Prade, editors. *Handbook of Defeasible Resoning and Uncertainty Management Systems*, volume 3. Kluwer, 1998.

[DP98b] D Dubois and H Prade Possibility theory: Qualitative and quantitative aspects. *Handbook of Defeasible Reasoning and Uncertainty Management Systems* , Vol 1:169-226, Gabbay and Smets (eds.). Kluwer Academic Publ., Dordrecht, 1998.

[DNP00] D Dubois, H Nguyen and H Prade Possibility theory, probability and fuzzy sets, Misunderstandings, bridges and gaps. *Fundamentals of Fussy Sets*, Chapter 7, 343-438, Dubois and Prade (eds.), *The Handbooks of Fuzzy Sets Series*.

[FS99] E Franconi and U Sattler. A data warehouse conceptual data model for multidimensional aggregation. In S Gatziu, M Jeusfeld, M Staudt, and Y Vassiliou, editors, *Proceedings of the Workshop in Design and Management of Data Warehouses*, Vol 19 of *CEUR Workshop Proceedings*, 13-1 to 13-10, 1999.

[Gar88] P Gardenfors. *Knowledge in Flux*. MIT Press, 1988.

[GM99] G Grahne and A Mendelzon. Tableau techniques for querying information sources through global schemas. In *Proceedings of the 7th International Conference on Database Theory (ICDT'99)*, LNCS 1540:332-347. Springer, 1999.

[HG00] K Hui and P Gray. Developing finite domain constraints – a data model approach. In *Proceedings of Computational Logic (CL'2000)*, LNAI 1861:448–462. Springer, 2000.

[HGNY97] J Hammer, H Garcia-Molina, S Nestorov, and R Yerneni. Template-based wrappers in the TSIMMIS system. In *Proceedings of ACM SIGMOD'97*, 532–535. ACM Press, 1997.

[HL04a] A Hunter and W Liu. Logical reasoning with multiple granularities of uncertainty in semi-structured information. *Proceedings of the International Conference on Information Processing and Management of Uncertainty*, 1009-1016, 2004.

[HL04b] A Hunter and W Liu. Merging uncertain information with semantic heterogeneity in XML. (journal submission) 2004.

[HL04c] A Hunter and W Liu. Measuring the quality of inconsistent uncertain information and quality of fusion. Technical Report, Department of Computer Science, UCL. 2004.

[HS03] A Hunter and R Summerton. Propositional fusion rules. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, LNCS 2711:502-514. Springer, 2003.

[HS04] A Hunter and R Summerton. Fusion rules for context-dependent aggregation of structured news reports. *Journal of Applied Non-Classical Logic*, 14(3):329-366, 2004.

[Hun00] A Hunter. Merging potentially inconsistent items of structured text. *Data and Knowledge Engineering*, 34:305–332, 2000.

[Hun02a] A Hunter. Logical fusion rules for merging structured news reports. *Data and Knowledge Engineering*, 42:23–56, 2002.

[Hun02b] A Hunter. Merging structured text using temporal knowledge. *Data and Knowledge Engineering*, 41:29–66, 2002.

[Hun02c] A Hunter. Measuring inconsistency in knowledge via quasi-classical models. *Proceedings of the 18th National Conference on Artificial Intelligence (AAAI'2002)*, 68–73, 2002.

[Hun03] A Hunter. Evaluating the significance of inconsistency. *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*, 468–473, 2003.

[KLM03] S Konieczny, J Lang, and P Marquis. Quantifying information and contradiction in propositional logic through epistemic actions. *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*, 106-111, 2003.

[KM91] H Katsuno and A Mendelzon. On the difference between updating a knowledgebase and revising it. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning(KR'91)*, 387–394. Morgan Kaufmann, 1991.

[KP98] S Konieczny and R Pino Perez. On the logic of merging. In *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*, 488–498. Morgan Kaufmann, 1998.

[Lee92] S K Lee. Imprecise and uncertain information in databases: An evidential approach. *Proceedings of Sixth International Conference on Data Engineering (ICDE)*, 614–621, 1992.

[LHA03] W Liu, X Hong and K Adamson. Efficient implementation of evidence combination using parallel techniques. Technical Report, Faculty of Engineering, University of Ulster, 2003.

[LS98] P Liberatore and M Schaerf. Arbitration (or how to merge knowledgebases). *IEEE Transactions on Knowledge and Data Engineering*, 10:76–90, 1998.

[LSS00] Y Loyer, N Spyratos, and D Stamate. Integration of information in four-valued logics under non-uniform assumptions. In *Proceedings of 30th IEEE International Symposium on Multiple-Valued Logic (ISMVL2000)*, 185-192. IEEE Press, 2000.

[Mot96] A Motro. Cooperative database systems. *International Journal of Intelligent Systems*, 11:717–732, 1996.

[NJ02] A Nierman and H Jagadish. ProTDB: Probabilistic data in XML. In *Proceedings of the International Conference on Very Large Databases (VLDB'02)*, LNCS 2590: 646–657. Springer, 2002.

[NS94] R Ng and V S Subrahmanian. Stable semantics for probabilistic deductive databases. *Information and Computation*, 110(1):42–83, 1994.

[Pet95] F E Petry. *Fuzzy Databases: Principles and Applications*. Kluwer Academic Publishers, 1995.

[PHG+99] A Preece, K Hui, A Gray, P Marti, T Bench-Capon, D Jeans, and Z Cui. The KRAFT architecture for knowledge fusion and transformation. In *Research and Development in Intelligence Systems XVI*, 23–38. Springer, 1999.

[PM98] A Poulovassilis and P McBrien. A general formal framework for schema transformation. *Data and Knowledge Engineering*, 28:47–71, 1998.

[PSB+99] N Paton, R Stevens, P Baker, C Goble, S Bechhofer, and A Brass. Query processing in the TAMBIS bioinformatics source integration system. In *Proceedings of the 11th International Conference on Scientific and Statistical Databases*, 138–147. IEEE Press, 1999.

[SA99] A Sahuguet and F Azavant. Building light-weight wrappers for legacy web data-sources using W4F. In *Proceedings of the International Conference on Very Large Databases (VLDB'99)*, 738-741, 1999.

[SDK95] S Sandri, D Dubois, and H Kalfsbeek. Elicitation, assessment and polling of expert judgements using possibility theory. *IEEE Transactions on Fuzzy Systems*, 3:313–335, 1995.

[Sha76] G Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[SL90] A Sheth and J Larson. Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Surveys*, 22:183–236, 1990.

[Sme88] Ph Smets. Belief functions. In Ph Smets, A Mamdani, D Dubois, and H Prade, editors, *Non-Standard Logics for Automated Reasoning*, 253–286. Academic Press, 1988.

[SO99] K Smith and L Obrst. Unpacking the semantics of source and usage to perform semantic reconciliation in large-scale information systems. In *ACM SIGMOD RECORD*, Vol 28:26–31, 1999.