Transactions of the Korean Nuclear Society Autumn Meeting
PyeongChang, Korea, October 30-31, 2008

Some Techniques for Aggregation of Uncertain Cardinal Information in Risk-Informed Applications

Seong Ho KIM[*a, c], Sang Hoon HAN[a], Joon Eon YANG[a], NamDuk SUH[b]
[a]*Korea Atomic Energy Research, Yuseong POBox150, Daejeon*
[b]*Korea Institute of Nuclear Safety, Yuseong, Daejeon*
[c]*Systemia G&E Inc.,399-8 Doryeong Dong, Yuseong Gu, Daejeon*
[*]*Corresponding author: well48@hanmir.com*

1. Introduction

Information aggregation (also known as information fusion) is viewed as the process that combines information obtained from different multiple sources such as experts, models, and sensors. The aggregated information is supposed to be more efficient and potentially more accurate than individual information achieved by means of a single source. Its application area is used in computer vision, medical image processing, pattern recognition, multi-sensor data fusion, decision making, *etc*.

There are different techniques for an information aggregation that have been successfully used in the science and engineering fields. For example, we find it much more powerful to devise aggregators in terms of the following principles: Bayesian network, Dempster-Shafer theory, and/or fuzzy integral.

Typically, risk-informed applications [Zio 2008; USNRC 2008] for a management of complex systems like nuclear power plants deal with information from rare events, subjective expert judgments, and similar data of other plants. In particular, various techniques for the probabilistic risk assessment (PRA) of a nuclear safety (*e.g.*, human reliability analysis, fault tree analysis, event tree analysis) are based on the theory of a probability measure, which is a subset of fuzzy measures such as a belief measure and a plausibility measure. Fig.1 shows a relation among fuzzy measures.

The main purpose of this work is to develop a non-additive Bayesian network methodology as an aggregation tool.

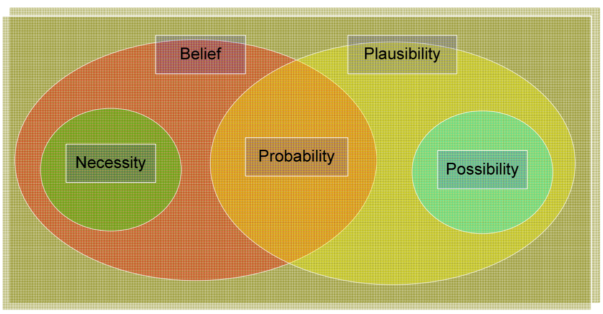

Figure 1. Venn diagram of some fuzzy measures

2. Methods and Data

Among the various approaches for an aggregation of multiple information sources, a **Bayesian network** (BN) approach is used for aggregating information, above all, at an evaluation level rather than at a higher level. A **Dempster-Shafer theory** (DST) approach is involved at a lower level as well as at a high level like a decision level.

2.1 Bayesian network method

BN is one of the probabilistic graphical approaches. A Bayesian network is often referred to as a Bayes net, a belief network, or a Bayesian belief network. It consists of a directed acyclic graph (DAG) for a representation and conditional probability tables (CPTs) for an evaluation. DAG has nodes and arrows. Arrows are referred to as edges, directed arcs from an influencing node (*i.e.*, parent) to an influenced node (*i.e.*, child). Here, nodes stand for variables (proposal, hypothesis, or evidence) and arrows for causal or influential relationships between nodes. CPTs, also known as node probability tables (NPTs), define the probabilistic relationship of each node given its respective parents occurring. Root nodes (*i.e.*, nodes without parents) are expressed in terms of their marginal probability distributions.

BN method has some advantages over traditional methods as follows: 1) It enables us to treat various phenomena with a degree of dependencies among proposals such as a common cause failure [Torres-Toledano and Sucar 1998]; 2) It facilitates in aggregating information from a variety of sources, including experimental data, historical data, and prior expert opinions [Marquez *et al*. 2007]. On the contrary, there are some limitations to the BN method. We summarize a severe limitation as follows: It can only handle singleton proposals due to a strong constraint on the probability measure, that is, additivity.

To deal with the limitation of the BN method, a hybrid method such as an evidential BN method is proposed [Simon *et al*. 2008]. Evidential BN enables us to treat an epistemic uncertainty at a system evaluation level. Here, a type of epistemic uncertainty is due to the lack of knowledge about the system of interest and is reducible, while a type of aleatory uncertainty results from the randomness of the system behavior and is irreducible.

System **S** under consideration consists of two components **C1** and **C2**. Each component has 3-states as {{u}, {d}, {u, d}}. Here, {u}, {d}, and {u, d} denote the up (or success) state, the down (or failure) state, and the unknown state, respectively. Fig 2 shows a BN with 3 nodes for representing system **S**. Here, the symbol $\mathbf{\Omega}$ means the frame of discernment, that is, all the possible states of the problem.

As a mathematical model, the relation between two nodes can be formulated by Bayes' rule, as Eq. (1). The rule requires us to estimate the entire prior and conditional probabilities. The probability distribution for system **S** can be written as Eq. (2) [Stamelos *et al.* 2003]:

P(H | I) = P (I | H) P(H)/P(I) . (1)

$$P(S) = NPT(S) \cdot vector(\ P(C1, C2)\ ) \quad (2)$$

, where the marginal distribution for **C1**, **C2** is

$$P(C1, C2) = NPT(C2) \cdot diag(\ NPT(C1)\ ). \quad (3)$$

Here, NPT(S) is a matrix representing NPT for node **S**. Similarly, NPT(C2) and NPT(C1) are matrices for nodes **C2** and **C1**, respectively.

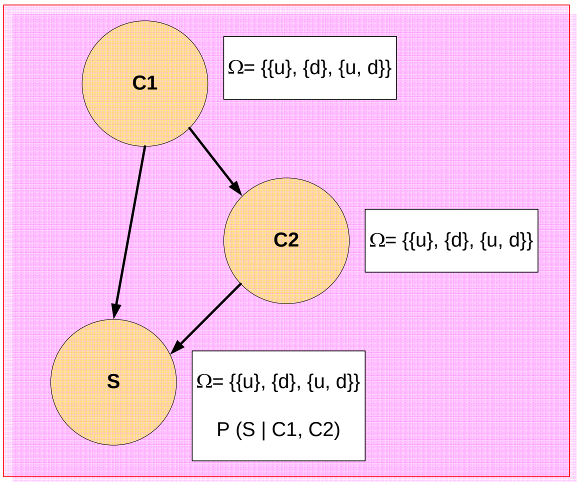

Figure 2. Bayesian network for two-component system

To introduce a belief measure Bel(·), the degree of belief for a proposal, a matrix NPT(Bel) for node **S** is defined as [Simon *et al*. 2008]:

$$NPT(Bel) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad (4)$$

Similarly, concerning a plausibility measure, the degree of plausibility for a proposal, a matrix NPT(Pl) for node **S** is defined

$$NPT(Pl) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (5)$$

2.2 Dempster-Shafer theory

To aggregate information including data under an uncertainty, a number of rules for a combination of evidence are studied, instead of using a traditional probabilistic theory [Sentz and Ferson 2002].

In the present work, we introduce the Dempster rule for a combination, as described in Eq. (6). Here, the operator $\oplus$ for two mass functions means an orthogonal sum, and is both commutative and associative.

The algorithm for the Dempster rule is as follows: 1) Define the frame of discernment; 2) Configure a combination matrix for the mass functions; 3) Establish a measure of conflict; and 4) Calculate an allocation of the combined evidence.

$$m(S) = m_1 \oplus m_2(S) = \begin{cases} \dfrac{\sum_{C1 \cap C2 = S} m_1(C1) m_2(C2)}{1 - \sum_{C1 \cap C2 = \phi} m_1(C1) m_2(C2)} & S \neq \phi \\ 0 & Other \end{cases} \quad (6)$$

3. Numerical Examples

In this section, we introduced the results obtained from the aggregators using the BN method and the Dempster rule method, respectively. For **C1**, (P(C1={u}), P(C1={d}), P(C1={u, d}))=(76.87, 13.13, 10.00) in %. For **C2**, we assign (81.87, 18.13, 0.00) in %.

For the BN aggregator for the 2-component system, we assign an AND logic to node **S** and then the logics for a belief measure and a plausibility measure. The results for S are obtained as follows: P(S)= (62.94, 28.87, 8.19) for the states ({u}, {d}, {u, d}). In addition, Bel(S={u})=62.94% and Pl(S={u})=71.13% for the states ({u}, {d}, {u, d}). Hence, for **S**, we can represent reliability with an uncertainty as [62.94%, 71.13%] due to an epistemic uncertainty for **C1**, 10.00%.

Using the Dempster rule aggregator, we obtained the aggregated reliability values as follows: m(S={u})= 94.43%, m(S={d})= 5.57%, and m(S={u, d})= 0.00%.

4. Conclusive Remarks

In the present work, we proposed two aggregators using the Bayesian network method and the Dempster rule method with numerical examples. As for handling an epistemic uncertainty, we found that the Dempster-Shafer theory was very influential.

For a future work, various other aggregation methods will be modeled.

Acknowledgement

This work is partially supported by the MEST as the Nuclear R&D Program.

REFERENCES

[1] C. Simon et al. (2008), Reliability Engineering and System Safety 93, 950-963.

[2] I. Stamelos, L. et al. (Jan 2003), Information and Software Technology 45 (1), 51-60..

[3] USNRC (2008): Tutorial on probabilistic risk assessment (PRA), Risk-Informed Regulation for Technical Staff, 38pp.

[4] E. Zio (2008), Nuclear Engineering and Technology 40 (5), 327-348.

[5] D. Marquez et al. (2007): A new Bayesian network approach to reliability modelling, 5th International Mathematical Methods in Reliability Conference (MMR 07), Glasgow 1-4 July 2007.

[6] K. Sentz and S. Ferson (2002): Combination of evidence in Dempster-Shafer theory, SAND 2002-0835, SAND Technical Report.