Delft University of Technology

Delft Center for Systems and Control

Technical report 15-006

Reasoning under uncertainty for knowledge-based fault diagnosis: A comparative study*

K. Verbert, B. De Schutter, and R. Babuška

If you want to cite this report, please use the following reference instead:
K. Verbert, B. De Schutter, and R. Babuška, "Reasoning under uncertainty for knowledge-based fault diagnosis: A comparative study," *Proceedings of the 9th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SafeProcess 2015)*, Paris, France, pp. 422–427, Sept. 2015. doi:10.1016/j.ifacol.2015.09.563

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

* This report can also be downloaded via https://pub.bartdeschutter.org/abs/15_006.html

Reasoning under Uncertainty for Knowledge-Based Fault Diagnosis: A Comparative Study

K. Verbert, B. De Schutter, R. Babuška

Delft Center for Systems and Control (DCSC), Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands, (e-mail: {k.a.j.verbert, b.deschutter, r.babuska}@tudelft.nl)

Abstract: This paper addresses reasoning under uncertainty for knowledge-based fault diagnosis. We illustrate how the fault diagnosis task is influenced by uncertainty. Furthermore, we compare how the diagnosis task is solved in the Bayesian and the Dempster-Shafer reasoning framework, in terms of both diagnostic performance and additional objectives, like transparency, adaptability, and computational efficiency. Since the diagnosis problem is influenced by different kinds of uncertainty, it is not straightforward to determine the optimal reasoning method. First, the different uncertain influences all have their own characteristics, asking for different reasoning approaches. So, to solve the whole problem in one reasoning framework, approximations and trade-offs need to be made. Second, which types of uncertainty are present and to what extent, is highly application-specific. Therefore, the best framework can only be assigned after the problem, the uncertainty characteristics, and the user requirements are known.

Keywords: Fault diagnosis, Bayesian inference, Dempster-Shafer inference.

1. INTRODUCTION

Condition-based maintenance aims to optimize maintenance planning based on actual system data. A maintenance scheme is optimal when it minimizes a function of downtime and costs. Generally, this implies that we want to perform adequate maintenance at a convenient time "just" before a system failure occurs. To be successful, one needs to estimate the time of occurrence and the type of an upcoming failure from the condition monitoring data. For the latter task, techniques from fault diagnosis are generally considered. Especially for safety-critical systems, like medical devices, nuclear reactors, and railway systems, this task is challenging due to the presence of uncertainty.

The diagnosis task is influenced by uncertainty in various ways. First, the available monitoring data may be incomplete, incorrect, and imprecise, e.g. due to sensors with a finite accuracy and suffering from drift and off-sets. Second, the available knowledge relating monitoring data to system health is usually uncertain, i.e. incomplete, subjective, and partly incorrect. Finally, the relations between monitoring data and system health are not deterministic, i.e. the system health depends on factors we do not know.

We focus on reasoning under uncertainty for the purpose of knowledge-based fault diagnosis. We investigate:

(1) How the diagnosis task is influenced by uncertainty;
(2) How the diagnosis task fits within the Bayesian and Dempster-Shafer (D-S) reasoning frameworks;
(3) Which additional objectives (e.g. transparency) are of relevance to assess a reasoning method.

Although reasoning under uncertainty has already received a lot of attention in the literature, most of the works focus on advocating one particular reasoning method (e.g. Bayesian or D-S inference) in general, i.e., without referring to a specific application (Cobb and Shenoy, 2003b; Dubois et al., 1996; Dubois and Prade, 2001; Ferson and Ginzburg, 1996; Lindley, 1987; Oukhellou et al., 2008, 2010; Smets, 1992) or apply one particular reasoning method to a specific problem (Basir and Yuan, 2007; Oukhellou et al., 2008, 2010; Sallak et al., 2013). In this work, we specifically consider reasoning for the purpose of fault diagnosis. By analyzing the different uncertainty sources that affect the diagnosis task, we investigate to what extent Bayesian and D-S reasoning suit the problem under consideration. In addition, we take objectives, like computational efficiency and transparency into account.

This paper is organized as follows: In Section 2, the diagnosis problem is introduced. Section 3 briefly reviews the Bayesian and D-S reasoning frameworks. Next, in Section 4, a diagnosis problem is solved in both the Bayesian and the D-S framework. In Section 5, the two reasoning methods are compared and additional performance criteria and trade-offs are discussed.

2. THE DIAGNOSIS TASK

A knowledge-based approach is considered for fault diagnosis, which comprises the determination of the cause(s) of any abnormal system behavior. As input for the diagnosis, we have ℓ raw monitoring signals x_1 till x_ℓ and as output we aim to determine the system health H, i.e. whether or not the system is healthy and if not what is the actual cause of the unhealthy behavior. To arrive from the raw monitoring signals x_1 till x_ℓ at a prediction of the current system health H, the following steps need to be taken:

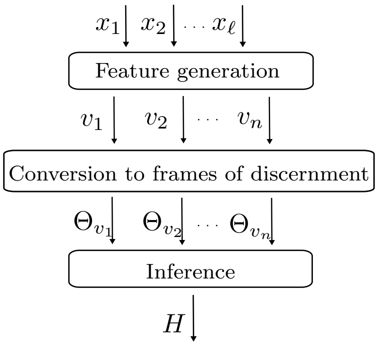

Fig. 1. Different steps in a knowledge-based diagnosis task.

(1) Feature generation, i.e. extracting features from the monitoring signals that are useful for diagnosis.
(2) Conversion from the feature space to a finite set of values, called the frame of discernment.
(3) Inference, i.e. deriving the current system health from the values of the features.

These steps are illustrated in Figure 1. For the purpose of diagnosis, n independent features v_1 till v_n are derived from x_1 till x_ℓ. For each feature v_i, its actual "value" is determined and represented by a distribution function over the frame of discernment Θ_{v_i}. The type of distribution function used, depends on the framework considered for the diagnosis, e.g. when a Bayesian approach is considered, the distribution function has the form of a probability distribution and when a D-S approach is considered, it has the form of a mass function. Based on the distribution functions of all features, system health is inferred.

The diagnosis task outlined above, may be influenced by uncertainty at the different stages of the diagnosis process:

(1) The monitoring signals x are not a perfect representation of the quantity we aim to measure.
(2) In general, it is not straightforward to derive a distribution function over the frame of the discernment Θ_{v_i} based on the behavior of feature v_i. This means that approximations (e.g. based on subjective human judgment) need to be made.
(3) The knowledge base, i.e. the relations assumed between features and system health, is not completely complete and correct.

The exact nature of these different uncertain influences is highly application-specific. For reasoning purposes, uncertainty is often divided into:

(1) Statistical (i.e. aleatory) uncertainty
(2) Systematic (i.e. epistemic) uncertainty

Statistical uncertainty represents intrinsic variability, whereas systematic uncertainty arises due to a lack of knowledge. The two are often distinguished using the fact that systematic uncertainty can be reduced by gathering more data or knowledge, whereas statistical uncertainty cannot be reduced (Ferson and Ginzburg, 1996; Kiureghian and Ditlevsen, 2009). So, ideally we would like to eliminate all systematic uncertainty by improving our diagnosis setup (e.g. by placing additional sensors or by improving our knowledge base), so that only statistical uncertainty remains. In practice, this is not always possible e.g. because our knowledge is (still) not sufficient or because placing additional sensors is not possible or too expensive. Therefore, in the remainder we consider reasoning in the presence of both statistical and systematic uncertainty. We assume that systematic uncertainty is present in the form of subjective and incomplete information and we restrict ourselves to two well-known reasoning frameworks, Bayesian and D-S inference. Bayesian theory is a framework, especially designed to handle subjective probabilities, but claimed to be good in handling all kinds of uncertainty (Lindley, 1987). The D-S framework is especially suitable to handle incomplete information (Dempster, 1967; Shafer, 1976, 1990; Smets, 1994).

3. BAYESIAN AND DEMPSTER-SHAFER REASONING

This section contains background information about the two reasoning frameworks considered. Section 3.1 on Bayesian reasoning is based on (Darwiche, 2009; Pearl, 1988; Pearl and Russel, 2001; Wiegerinck et al., 2010) and Section 3.2 on D-S reasoning is based on (Cobb and Shenoy, 2003a; Smets, 1978, 1990, 1994; Yager, 1987).

3.1 Bayesian framework

The Bayesian framework is based on probabilities. At each reasoning step, a probability between zero and one inclusive is assigned to *each individual element* a in the frame of discernment Θ_Y of variable Y such that

$$\sum_{a \in \Theta_Y} \Pr(a) = 1 \tag{1}$$

When a new evidence $b \in \Theta_X$ regarding a variable X that is related to variable Y becomes available, the probability distribution of Y is updated using *Bayes' rule*:

$$\Pr(a|b) = \frac{\Pr(b|a)\Pr(a)}{\sum_{a' \in \Theta_Y} \Pr(b|a')\Pr(a')} \tag{2}$$

with $\Pr(a)$ the prior probability of a, $\Pr(a|b)$ the posterior probability of a, i.e. the probability of a after observing b, and $\Pr(b|a)$ the likelihood function, i.e. the probability of observing b given a.

Because in practice the available information is not always exactly in the desired format, i.e. represented by probability distributions, rules have been defined to approximate probability functions from non-probabilistic information. One important rule is the *principle of insufficient reasoning*, which states that in the absence of knowledge, all possible outcomes should be assigned equal probabilities. Another commonly used rule is the *additivity axiom*, which directly follows from (1) and states that:

$$\Pr(a) + \Pr(\sim a) = 1 \tag{3}$$

3.2 Dempster-Shafer (D-S) framework

The theory of belief functions was developed to handle incomplete information. This is realized by allowing the assignment of belief to *sets of elements of* Θ_Y instead of assigning belief only to individual elements, like in the Bayesian framework. A belief function $m^{\Theta_Y} : 2^{\Theta_Y} \to [0,1]$ is a function that assigns a "mass of belief" to each subset A of Θ_Y such that:

$$\sum_{A \subseteq \Theta_Y} m^{\Theta_Y}(A) = 1 \tag{4}$$

When a new evidence about Y, in the form of a mass function $m_{\mathrm{e}}^{\Theta_Y}$, becomes available, the mass function m^{Θ_Y} is updated using Dempster's combination rule:

$$m_{\mathrm{u}}^{\Theta_Y}(A) = \begin{cases} 0 & \text{if } A = \emptyset, \\ K \displaystyle\sum_{\substack{A' \cap A'' = A \\ A', A'' \subseteq \Theta_Y}} m^{\Theta_Y}(A') m_{\mathrm{e}}^{\Theta_Y}(A'') & \text{otherwise} \end{cases} \quad (5)$$

with m^{Θ_Y}, $m_{\mathrm{e}}^{\Theta_Y}$, and $m_{\mathrm{u}}^{\Theta_Y}$ mass functions on the same space Θ_Y, K a normalization constant, and $m_{\mathrm{u}}^{\Theta_Y}$ the updated mass function.

As the available information is not always in the desired format e.g. we may have conditional information or pieces of information on different spaces, operations have been defined to convert this information to the required format. The most common ones are:

(1) Ballooning extension (Smets, 1978) to convert conditional information to a mass function on the joint space;
(2) Cylindrical extension (Cobb and Shenoy, 2003a) to convert a mass function to a mass function on a larger space;
(3) Marginalization (Cobb and Shenoy, 2003a) to convert a mass function to a mass function on a smaller space.

4. UNCERTAINTY REASONING FOR DIAGNOSIS

To illustrate the different reasoning steps and the differences between Bayesian and D-S reasoning, in this section, we work out an uncertain diagnosis problem in both frameworks. Before analyzing the diagnosis problem in the two frameworks, we specify the diagnosis task considered.

4.1 Problem specification

We consider a simple reasoning task with the aim to determine system health H based on condition monitoring data. As we focus on reasoning, we assume that the monitoring signals are already transformed to the feature space and that information regarding the features v_1 till v_{n} becomes available. So, based on evidence regarding the feature values, we aim to determine system health H. This way, the diagnosis task comprises the following subtasks:

T_1 : Transforming the uncertain knowledge base to the desired format, i.e. conditional probabilities for Bayesian reasoning and mass functions for D-S reasoning;
T_2 : Transforming the uncertain information regarding the features to the desired format, i.e. a specific value for Bayesian reasoning and a mass distribution function for D-S reasoning;
T_3 : Inferring system health.

For this particular example, there are only two relevant features, v_1 and v_2, and the frames of discernment of system health H and features v_1 and v_2 are defined as:

$$\Theta_H = \{\mathrm{f}_1, \mathrm{f}_2\}$$
$$\Theta_{v_1} = \{\mathrm{a,b,c,d}\}$$
$$\Theta_{v_2} = \{\mathrm{y,z}\}$$

The following knowledge base is available for inference:

s_1 : **If** $H = \mathrm{f}_1$ **then** $\Pr(v_1 = \mathrm{b}) = 0.85$
s_2 : **If** $H = \mathrm{f}_2$ **then** $\Pr(v_1 = \mathrm{c}) = 0.95 \wedge \Pr(v_1 = \mathrm{d}) = 0.05$
s_3 : **If** $H = \mathrm{f}_1$ **then** $\Pr(v_2 = \mathrm{y}) = 1$
s_4 : **If** $H = \mathrm{f}_2$ **then** $\Pr(v_2 = \mathrm{z}) = 0.7$

Furthermore, we assume that no prior knowledge regarding the relative probabilities of the two health states in Θ_H is available. The evidences regarding v_1 and v_2 are available in the form of incomplete distribution functions over the feature values:

e_1 : $\Pr(v_1 = \mathrm{d}) = 0.3$, $\Pr(v_1 \neq \mathrm{d}) = 0.7$
e_2 : $\Pr(v_1 = \mathrm{c} \vee v_1 = \mathrm{d}) = 1$
e_3 : $\Pr(v_2 = \mathrm{z}) = 0.8$

Note that these distribution functions are not yet in the format required by the Bayesian and D-S reasoning framework. That is why step T_2 is required.

4.2 Bayesian reasoning

In the Bayesian framework, the diagnosis task is graphically represented by the Bayesian network as is shown in Figure 2. Here, the variables of interest are system health H and the features v_1 and v_2. The links in the network indicate that system health H affects both v_1 and v_2, but features v_1 and v_2 do not influence each other.

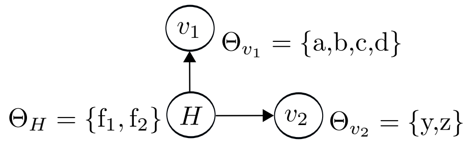

Fig. 2. Bayesian network of the diagnosis problem.

In our example, we have no information regarding the prior probability function of H and we have only partial knowledge regarding the conditional probability tables of v_1 and v_2. Therefore, we use the principle of insufficient reasoning to obtain a probability function on H:

$$\Pr(\mathrm{f}_1) = \Pr(\mathrm{f}_2) = 0.5 \quad (6)$$

Furthermore, given knowledge rules s_1 till s_4 and considering the principle of insufficient reasoning and the additivity axiom (3), the conditional probability tables of v_1 and v_2, as given in Table 1 are obtained.

Table 1. Conditional probability table of v_1 and v_2.

	v_1				v_2	
H	a	b	c	d	y	x
f_1	0.05	0.85	0.05	0.05	1	0
f_2	0	0	0.95	0.05	0.3	0.7

Now that the network is defined, we can update the network based on evidences e_1, e_2, e_3. First, a decision regarding the value of v_1 and v_2 needs to be made. To determine the value of v_1, we combine the information in e_1 and e_2. First, we approximate e_1 by a probability distribution function (principle of insufficient reasoning):

$$\Pr(\mathrm{d}) = 0.3$$
$$\Pr(\mathrm{a}) = \Pr(\mathrm{b}) = \Pr(\mathrm{c}) = 0.2333 \quad (7)$$

Next, this information is conditioned based on e_2 using (2), resulting in the following probability distribution:

$$\begin{aligned}
\Pr(\mathrm{d}|\mathrm{c} \vee \mathrm{d}) &= \frac{1 \cdot 0.3}{0.3 + \frac{0.7}{3}} = 0.5625 \\
\Pr(\mathrm{c}|\mathrm{c} \vee \mathrm{d}) &= \frac{1 \cdot \frac{0.7}{3}}{0.3 + \frac{0.7}{3}} = 0.4375 \\
\Pr(\mathrm{a}|\mathrm{c} \vee \mathrm{d}) &= 0 \\
\Pr(\mathrm{b}|\mathrm{c} \vee \mathrm{d}) &= 0
\end{aligned} \tag{8}$$

As Bayes' rule requires a hard decision regarding the value of v_1, we decide that:

$$v_1 = \mathrm{d} \tag{9}$$

Similarly from e_3 we derive:

$$\begin{aligned}
\Pr(\mathrm{z}) &= 0.8 \\
\Pr(\mathrm{y}) &= 0.2
\end{aligned} \tag{10}$$

from which we decide that:

$$v_2 = \mathrm{z} \tag{11}$$

Now we can update the network using Bayes' rule (2). First, conditioning on v_1 results in:

$$\begin{aligned}
\Pr(\mathrm{f}_1|\mathrm{d}) &= \frac{\Pr(\mathrm{d}|\mathrm{f}_1)\Pr(\mathrm{f}_1)}{\Pr(\mathrm{d}|\mathrm{f}_1)\Pr(\mathrm{f}_1) + \Pr(\mathrm{d}|\mathrm{f}_2)\Pr(\mathrm{f}_2)} = 0.5 \\
\Pr(\mathrm{f}_2|\mathrm{d}) &= \frac{\Pr(\mathrm{d}|\mathrm{f}_2)\Pr(\mathrm{f}_2)}{\Pr(\mathrm{d}|\mathrm{f}_2)\Pr(\mathrm{f}_2) + \Pr(\mathrm{d}|\mathrm{f}_2)\Pr(\mathrm{f}_2)} = 0.5
\end{aligned} \tag{12}$$

From this probability distribution, it follows that the evidences regarding v_1 do not provide much information regarding H.

Next, the obtained posterior probability distribution of H is used to condition the evidence regarding v_2 on:

$$\begin{aligned}
\Pr(\mathrm{f}_1|\mathrm{d} \wedge \mathrm{z}) &= \frac{\Pr(\mathrm{z}|\mathrm{f}_1)\Pr(\mathrm{f}_1|\mathrm{d})}{\Pr(\mathrm{z}|\mathrm{f}_1)\Pr(\mathrm{f}_1|\mathrm{d}) + \Pr(\mathrm{z}|\mathrm{f}_2)\Pr(\mathrm{f}_2|\mathrm{d})} = 0 \\
\Pr(\mathrm{f}_2|\mathrm{d} \wedge \mathrm{z}) &= \frac{\Pr(\mathrm{z}|\mathrm{f}_2)\Pr(\mathrm{f}_2|\mathrm{d})}{\Pr(\mathrm{z}|\mathrm{f}_2)\Pr(\mathrm{f}_2|\mathrm{d}) + \Pr(\mathrm{z}|\mathrm{f}_2)\Pr(\mathrm{f}_2|\mathrm{d})} = 1
\end{aligned} \tag{13}$$

So, given v_1 and v_2, we conclude that $H = \mathrm{f}_2$.

4.3 Dempster-Shafer belief networks

In the D-S framework, the diagnosis problem is graphically represented by the valuation network as is shown in Figure 3. The rounded rectangles represent the variables of interest, i.e. system health and features v_1 and v_2, and the hexagons are valuations, representing the knowledge about the relationships between the connected variables (Shenoy, 1992). To define a D-S valuation network, each valuation must be represented by a mass function, e.g. for the relation between v_1 and H a mass function $m^{\Theta_{v_1} \times H}$ on the space $\Theta_{v_1} \times \Theta_H$ is required. As the available information is conditional, the ballooning extension is used to derive the mass functions on the joint spaces of the related variables. First, the ballooning extension is used to transform rule s_1 to a mass function on $\Theta_{v_1} \times \Theta_H$:

$$\begin{aligned}
m^{\Theta_{v_1}}(\cdot|\mathrm{f}_1)^{\Uparrow \Theta_{v_1} \times \Theta_H}(\{(\mathrm{b}, \mathrm{f}_1), (\cdot, \mathrm{f}_2)\}) &= 0.85 \\
m^{\Theta_{v_1}}(\cdot|\mathrm{f}_1)^{\Uparrow \Theta_{v_1} \times \Theta_H}(\Theta_{v_1} \times \Theta_H) &= 0.15
\end{aligned} \tag{14}$$

Similarly, rule s_2 is transformed to a mass function on $\Theta_{v_1} \times \Theta_H$:

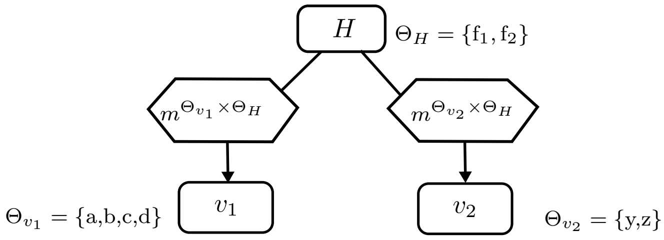

Fig. 3. D-S representation of the diagnosis problem.

$$\begin{aligned}
m^{\Theta_{v_1}}(\cdot|\mathrm{f}_2)^{\Uparrow \Theta_{v_1} \times \Theta_H}(\{(\mathrm{c}, \mathrm{f}_2), (\cdot, \mathrm{f}_1)\}) &= 0.95 \\
m^{\Theta_{v_1}}(\cdot|\mathrm{f}_2)^{\Uparrow \Theta_{v_1} \times \Theta_H}(\{(\mathrm{d}, \mathrm{f}_2), (\cdot, \mathrm{f}_1)\}) &= 0.05
\end{aligned} \tag{15}$$

with $m^{\Theta_{v_1}}(\cdot|\mathrm{f}_1)^{\Uparrow \Theta_{v_1} \times \Theta_H}(\cdot)$ and $m^{\Theta_{v_1}}(\cdot|\mathrm{f}_2)^{\Uparrow \Theta_{v_1} \times \Theta_H}(\cdot)$ mass functions on the space $\Theta_{v_1} \times \Theta_H$ that result from applying the ballooning extension on the conditional information about Θ_{v_1}. Combining (14) and (15) using (5) results in:

$$\begin{aligned}
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{b}, \mathrm{f}_1), (\mathrm{c}, \mathrm{f}_2)\}) &= 0.85 \cdot 0.95 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{b}, \mathrm{f}_1), (\mathrm{d}, \mathrm{f}_2)\}) &= 0.85 \cdot 0.05 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{c}, \mathrm{f}_2), (\cdot, \mathrm{f}_1)\}) &= 0.15 \cdot 0.95 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{d}, \mathrm{f}_2), (\cdot, \mathrm{f}_1)\}) &= 0.15 \cdot 0.05
\end{aligned} \tag{16}$$

Next, the ballooning extension is used to extend rules s_3 and s_4 to a mass function on the space $\Theta_{v_2} \times \Theta_H$.

$$m^{\Theta_{v_2}}(\cdot|\mathrm{f}_1)^{\Uparrow \Theta_{v_2} \times \Theta_H}(\{(\mathrm{y}, \mathrm{f}_1), (\cdot, \mathrm{f}_2)\}) = 1 \tag{17}$$

and applying the ballooning extension on s_4 gives:

$$\begin{aligned}
m^{\Theta_{v_2}}(\cdot|\mathrm{f}_2)^{\Uparrow \Theta_{v_2} \times \Theta_H}(\{(\mathrm{z}, \mathrm{f}_2), (\cdot, \mathrm{f}_1)\}) &= 0.7 \\
m^{\Theta_{v_2}}(\cdot|\mathrm{f}_2)^{\Uparrow \Theta_{v_2} \times \Theta_H}(\Theta_{v_2} \times \Theta_H) &= 0.3
\end{aligned} \tag{18}$$

Combining (17) and (18) using (5) yields:

$$\begin{aligned}
m^{\Theta_{v_2} \times \Theta_H}(\{(\mathrm{z}, \mathrm{f}_2), (\mathrm{y}, \mathrm{f}_1)\}) &= 0.7 \\
m^{\Theta_{v_2} \times \Theta_H}(\{(\mathrm{y}, \mathrm{f}_1), (\cdot, \mathrm{f}_2)\}) &= 0.3
\end{aligned} \tag{19}$$

Now that the network is defined, we can update the network based on evidences regarding the features v_1 and v_2. We first convert the evidences e_1, e_2, e_3 to the appropriate format. To obtain a mass function regarding v_1, we combine the information in e_1 and e_2. To this aim, both evidence e_1 and e_2 are represented by a mass function, after which the two functions are combined. Evidence e_1 is represented by mass function $m_1^{\Theta_{v_1}}$:

$$\begin{aligned}
m_1^{\Theta_{v_1}}(\{\mathrm{d}\}) &= 0.3 \\
m_1^{\Theta_{v_1}}(\{\mathrm{a}, \mathrm{b}, \mathrm{c}\}) &= 0.7
\end{aligned} \tag{20}$$

Evidence e_2 is represented by mass function $m_2^{\Theta_{v_1}}$:

$$m_2^{\Theta_{v_1}}(\{\mathrm{c}, \mathrm{d}\}) = 1 \tag{21}$$

Combining $m_1^{\Theta_{v_1}}$ and $m_2^{\Theta_{v_1}}$: using (5), yields:

$$\begin{aligned}
m^{\Theta_{v_1}}(\{\mathrm{d}\}) &= 0.3 \\
m^{\Theta_{v_1}}(\{\mathrm{c}\}) &= 0.7
\end{aligned} \tag{22}$$

A mass function regarding v_2 follows from evidence e_3:

$$\begin{aligned}
m^{\Theta_{v_2}}(\{\mathrm{z}\}) &= 0.8 \\
m^{\Theta_{v_2}}(\Theta_{v_2}) &= 0.2
\end{aligned} \tag{23}$$

First, we update the network based on the evidences regarding v_1. Therefore, we combine $m^{\Theta_{v_1}}$ with the corresponding valuation function $m^{\Theta_{v_1} \times \Theta_H}$. As the two mass functions are defined on different space, we first vacuously extend $m^{\Theta_{v_1}}$ to the space $\Theta_{v_1} \times \Theta_H$ using the cylindrical extension, so that we end up with two mass functions on the same space:

$$m^{\Theta_{v_1} \uparrow \Theta_{v_1} \times \Theta_H}\big(\{(\mathrm{d}, \mathrm{f}_1), (\mathrm{d}, \mathrm{f}_2)\}\big) = 0.3;$$
$$m^{\Theta_{v_1} \uparrow \Theta_{v_1} \times \Theta_H}\big(\{(\mathrm{c}, \mathrm{f}_1), (\mathrm{c}, \mathrm{f}_2)\}\big) = 0.7. \quad (24)$$

with $m^{\Theta_{v_1} \uparrow \Theta_{v_1} \times \Theta_H}(\cdot)$ the cylindrical extension of a mass function on Θ_{v_1} on the space $\Theta_{v_1} \times \Theta_H$. Combining this mass function with the valuation function $m^{\Theta_{v_1} \times \Theta_H}$ (16) according to (5) results in:

$$\begin{aligned}
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{d}, \mathrm{f}_2)\}) &= 0.3 \cdot 0.85 \cdot 0.05 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{c}, \mathrm{f}_2)\}) &= 0.7 \cdot 0.85 \cdot 0.95 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{d}, \mathrm{f}_1)\}) &= 0.3 \cdot 0.15 \cdot 0.95 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{c}, \mathrm{f}_1), (\mathrm{c}, \mathrm{f}_2)\}) &= 0.7 \cdot 0.15 \cdot 0.95 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{d}, \mathrm{f}_1), (\mathrm{d}, \mathrm{f}_2)\}) &= 0.3 \cdot 0.15 \cdot 0.05 \\
m^{\Theta_{v_1} \times \Theta_H}(\{(\mathrm{c}, \mathrm{f}_1)\}) &= 0.7 \cdot 0.15 \cdot 0.05 \quad (25)
\end{aligned}$$

Marginalization of $m^{\Theta_{v_1} \times \Theta_H}$ on Θ_H gives:

$$\begin{aligned}
m^{\Theta_{v_1} \times \Theta_F \downarrow \Theta_H}(\{\mathrm{f}_1\}) &= \frac{0.3 \cdot 0.15 \cdot 0.95 + 0.7 \cdot 0.15 \cdot 0.05}{0.728} = 0.07 \\
m^{\Theta_{v_1} \times \Theta_F \downarrow \Theta_H}(\{\mathrm{f}_2\}) &= \frac{0.3 \cdot 0.85 \cdot 0.05 + 0.7 \cdot 0.85 \cdot 0.95}{0.728} = 0.79 \\
m^{\Theta_{v_1} \times \Theta_F \downarrow \Theta_H}(\{\mathrm{f}_1, \mathrm{f}_2\}) &= \frac{0.7 \cdot 0.15 \cdot 0.95 + 0.3 \cdot 0.15 \cdot 0.05}{0.728} = 0.14 \quad (26)
\end{aligned}$$

with $m^{\Theta_{v_1} \times \Theta_F \downarrow \Theta_H}(\cdot)$ the marginalization of a mass function on the space $\Theta_{v_1} \times \Theta_H$ on the space Θ_H. So, based on only the evidences regarding v_1, we conclude that most probably $H = \mathrm{f}_2$.

Next, we update the network based on the evidence regarding v_2. Extending $m^{\Theta_{v_2}}$ to the space $\Theta_{v_2} \times \Theta_H$ using the cylindrical extension yields:

$$m^{\Theta_{v_2} \uparrow \Theta_{v_2} \times \Theta_H}\big(\{(\mathrm{z}, \mathrm{f}_1), (\mathrm{z}, \mathrm{f}_2)\}\big) = 0.8$$
$$m^{\Theta_{v_2} \uparrow \Theta_{v_2} \times \Theta_H}\big(\Theta_{v_2} \times \Theta_H\big) = 0.2 \quad (27)$$

Combining (27) with the valuation function $m^{\Theta_{v_2} \times \Theta_F}$ (19) according to (5) gives:

$$\begin{aligned}
m^{\Theta_{v_2} \times \Theta_H}(\{(\mathrm{z}, \mathrm{f}_2)\}) &= 0.7 \cdot 0.8 + 0.3 \cdot 0.8 \\
m^{\Theta_{v_2} \times \Theta_H}(\{(\mathrm{z}, \mathrm{f}_2), (\mathrm{y}, \mathrm{f}_1)\}) &= 0.7 \cdot 0.2 \\
m^{\Theta_{v_2} \times \Theta_H}(\{(\mathrm{y}, \mathrm{f}_1), (\cdot, \mathrm{f}_2)) &= 0.3 \cdot 0.2 \quad (28)
\end{aligned}$$

Marginalization of $m^{\Theta_{v_2} \times \Theta_H}$ on Θ_H gives:

$$\begin{aligned}
m^{\Theta_{v_2} \times \Theta_H \downarrow \Theta_H}(\{\mathrm{f}_2\}) &= 0.7 \cdot 0.8 + 0.3 \cdot 0.8 \\
m^{\Theta_{v_2} \times \Theta_H \downarrow \Theta_H}(\{\mathrm{f}_1, \mathrm{f}_2\}) &= 0.7 \cdot 0.2 + 0.3 \cdot 0.2 \quad (29)
\end{aligned}$$

Combining (26) and (29) accordingly (5) results in the final mass distribution:

$$\begin{aligned}
m^{\Theta_H}(\{\mathrm{f}_1\}) &= 0.01 \\
m^{\Theta_H}(\{\mathrm{f}_2\}) &= 0.96 \\
m^{\Theta_H}(\{\mathrm{f}_1, \mathrm{f}_2\}) &= 0.03 \quad (30)
\end{aligned}$$

from which we conclude that probably $H = \mathrm{f}_2$.

5. PERFORMANCE REQUIREMENTS AND CONSIDERATIONS

In the previous section, we have seen how a diagnosis problem can be solved in both the Bayesian and the D-S framework. In this section, we compare the results where we take, next to diagnostic performance, additional performance criteria, like computational efficiency and transparency, into account.

As fault diagnosis comprises the determination of the cause(s) of the abnormal behavior of the monitoring signal(s), which is a causal reasoning task, Bayesian reasoning seems particularly suited for this task. However, in practice, the available knowledge and data are not in Bayesian format (i.e. are not represented by a probability distribution) and approximations need to be made. The D-S framework is perfectly suited to handle knowledge that is not purely probabilistic, but is less suited for causal reasoning. This means that when choosing a method it needs to be considered to what extent the actual data and knowledge are probabilistic and what the consequences are of approximating the non-probabilistic information by probability distributions. Unfortunately, a good insight into the characteristics of all uncertain influences is often not available and approximations have to be used.

Other objectives that need to be taken into account when selecting an appropriate reasoning approach are e.g., computational efficiency, clarity of inference, adaptability, and flexibility of reasoning. *Computational efficiency* may be of importance for on-line diagnosis tasks. Computationally, D-S networks are more expensive to evaluate than Bayesian networks (Cobb and Shenoy, 2003a; Haenni and Lehmann, 2003). So, even when, for a particular application, D-S networks may regarded as theoretically superior to Bayesian networks, the Bayesian solution may outperform the D-S solution in practice because the diagnosis can be carried out with a smaller delay.

Clarity of inference is an objective that is relevant in many practical applications as the implementation of a decision support system within a company is much easier when the system is intuitive and understandable. Considering "clarity of inference" for the user, Bayesian networks outperform D-S networks. In contrast, for the experiment designer, the output of a D-S framework is often considered clearer, as the D-S framework makes a distinction between probabilistic information and ignorance. Whereas in the D-S framework two distinct outcomes are obtained in the situation that no information regarding system health is available, i.e. $m(\Theta_H) = 1$, and the situation in which we have the information that all faults are equally likely, i.e. $m(\mathrm{f}_1) = m(\mathrm{f}_2) = ... = m(\mathrm{f}_n) = 1/n$, in the Bayesian framework, both situations result in the same probability function, $\Pr(\mathrm{f}_1) = \Pr(\mathrm{f}_2) = ... = \Pr(\mathrm{f}_n) = 1/n$. The additional information provided by the D-S outcome can be used to reconsider the diagnosis setup (e.g. an incomplete outcome gives rise to extend the knowledge base) or to assist decision making, e.g. by choosing a conservative decision approach when the diagnosis outcome is ignorant.

Finally, the relevance of the objectives *adaptability* and *flexibility of reasoning* is rather application-specific. Do we expect that our data and knowledge will change over time? So, do we really need an easily adaptable system? Or, do we really need complex types of reasoning? Bayesian reasoning is promoted for its ability to easily handle complex types of reasoning, like explaining away and bi-directional (both predictive and diagnostic) reasoning (Pearl, 1988). When we consider the diagnosis problem as introduced in Section 2, there is no need for complex types of reasoning as only evidence of the consequence variables is expected and there are no variables that are influenced by more than one variable. However, the situation changes when we take besides system data, external influences, e.g. weather, into account as inputs for the diagnosis. In this situation, deviations in the monitoring data m can be explained by system health and/or by the external influences, in which case there is a need for bi-directional inference and explaining away.

6. CONCLUSIONS

We have discussed how a knowledge-based diagnosis approach is influenced by uncertainty and how the uncertain problem can be solved in the Bayesian and Dempster-Shafer framework. We conclude that the final choice for a reasoning method is highly application-specific, both because each problem has different uncertain influences, requiring other reasoning strategies and because each problem has another weighting of additional objectives, like computational efficiency. Therefore, the optimal reasoning strategy can only be assigned after the problem (including uncertainty characteristics) and the user requirements are known.

ACKNOWLEDGEMENTS

This research is supported by ProRail and the Dutch Technology Foundation STW, which is part of the Netherlands Organization for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement nr 324432.

REFERENCES

Basir, O. and Yuan, X. (2007). Engine fault diagnosis based on multi-sensor information fusion using Dempster-Shafer evidence theory. *Information Fusion*, 8(4), 379–386.

Cobb, B.R. and Shenoy, P.P. (2003a). A comparison of Bayesian and belief function reasoning. *Information Systems Frontiers*, 5(4), 345–358.

Cobb, B. and Shenoy, P. (2003b). A comparison of methods for transforming belief function models to probability models. In T. Nielsen and N. Zhang (eds.), *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. Springer-Verlag, Berlin.

Darwiche, A. (2009). *Modeling and Reasoning with Bayesian Networks*. Cambridge University Press.

Dempster, A. (1967). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2), 325–339.

Dubois, D. and Prade, H. (2001). Possibility theory, probability theory and multiple-valued logics: A clarification. *Annuals of Mathematics and Artificial Intelligence*, 32(1-4), 35–66.

Dubois, D., Prade, H., and Smets, P. (1996). Representing partial ignorance. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 26(3), 361–377.

Ferson, S. and Ginzburg, L. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2), 133–144.

Haenni, R. and Lehmann, N. (2003). Implementing belief function computations. *International Journal of Intelligent Systems*, 18(1), 31–49.

Kiureghian, A.D. and Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, 31(2), 105–112.

Lindley, D. (1987). The probability approach to the treatment of uncertainty in artificial intelligence and expert systems. *Statistical Sciences*, 2(1), 17–24.

Oukhellou, L., Côme, E., Bouillaut, L., and Aknin, P. (2008). Combined use of sensor data and structural knowledge processed by Bayesian network: Application to a railway diagnosis aid scheme. *Transportation Research Part C: Emerging Technologies*, 16(6), 755–767.

Oukhellou, L., Debiolles, A., Denoeux, T., and Aknin, P. (2010). Fault diagnosis in railway track circuits using Dempster-Shafer classifier fusion. *Engineering Applications of Artificial Intelligence*, 23(1), 117–128.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc.

Pearl, J. and Russel, S. (2001). Bayesian networks. In M. Arbib (ed.), *Handbook of Brain Theory and Neural Networks*. MIT press.

Sallak, M., Schön, W., and Aguirre, F. (2013). Reliability assessment for multi-state systems under uncertainties based on the dempster–shafer theory. *IIE Transactions*, 45(9), 995–1007.

Shafer, G. (1976). *A Mathematical Theory of Evidence*, volume 1. Princeton University Press, Princeton.

Shafer, G. (1990). Belief functions. In G. Shafer and J. Pearl (eds.), *Readings in Uncertain Reasoning*. Morgan Kaufman, San Francisco.

Shenoy, P. (1992). Valuation-based systems for Bayesian decision analysis. *Operations research*, 40(3), 463–484.

Smets, P. (1978). *Un Modéle Mathématico-Statistique Stimulant le Processus du Diagnostic Médical*. Ph.D. thesis, Université de Bruxelles.

Smets, P. (1990). The transferable belief model and other interpretations of Dempster-Shafer's model. In *Proceedings of the 6th Annual Conference on Uncertainty in Artificial Intelligence*, 375–383. Amsterdam, the Netherlands.

Smets, P. (1992). Resolving misunderstandings about belief functions. *International Journal of Approximate Reasoning*, 6(3), 321–344.

Smets, P. (1994). What is Dempster-Shafer's model? In R. Yager, J. Kacprzyk, and M. Fedrizzi (eds.), *Advances in the Dempster-Shafer Theory of Evidence*, 5–34. John Wiley & Sons, Inc.

Wiegerinck, W., Kappen, H., and Burgers, W. (2010). Bayesian networks for expert systems: Theory and practical applications. In R. Babuška and F. Groen (eds.), *Interactive Collaborative Information Systems*, volume 281 of *Studies in Computational Intelligence*, 547–578. Springer.

Yager, R. (1987). On the Dempster-Shafer framework and new combination rules. *Information Science*, 41(2), 93–137.