WISA
A Modular and Hybrid Expert System for Machine and Plant Diagnosis

Mario Thron, Thomas Bangemann and Nico Suchold
Institut für Automation und Kommunikation e.V. Magdeburg, Steinfeldstr 3, D-39179 Barleben, Germany

Keywords: Industrial diagnosis, expert system, Fuzzy Logic, Bayesian Network, logic description, modularity.

Abstract: Expert systems are well known tools for diagnosis purposes in medicine and industry. One problem is the hard efford, to create the knowledge base. This article describes an expert system for industrial diagnosis and shows an efficient approach for the creation of the rule base, which is based on the reusage of knowledge modules. These knowledge modules are representants for assets like devices, machines and plants. The article encourages manufacturers of such assets to provide diagnosis knowledge bases by using a proposed multi-paradigm rule definition language called HLD (Hybrid Logic Description).
Rule based knowledge may be expressed by using various methodologies, which differ in expressiveness but also in runtime performance. The HLD allowes rules to be defined as propositional logic with or without the use of certainty factors, as Fuzzy Logic or as probabilistic rules as in Bayesian Networks. The most effective rule type may be chosen to describe causal dependencies between symptoms and failures. An evaluation prototype implementation has been developed in the research project WISA, which includes a sotftware tool chain for handling HLD files.

1 INTRODUCTION

A fundamental goal of industrial companies is to optimize the use of their production equipment, which consists of simple assets like devices (sensors, actuators, etc.), assembled machines and even more complex assets like production lines and plants. The term ‘optimization’ relates here to maximization of the asset performance and to minimization of production down-times as effects of failures.

This article is dedicated to the maximization of the equipment availability. Industrial companies prefer the use of reliable equipment over managing failures. *Failure Mode and Effects Analysis* (*FMEA*) or *Failure Mode, Effects and Criticality Analysis* (*FMECA*) may be used for engineering of robust production systems. But the production equipment may fail even in the case of a careful design. Thus different maintenance strategies have been developed in order to increase the availability of the assets:

- *Predictive Maintenance.* Assets are continuously monitored to get information on the abrasion state. The remaining asset life time is predicted in order to schedule necessary maintenance activities like repairing or asset exchanges.
- *Preventive Maintenance.* Maintenance activities are scheduled according to the machine run time and absolute machine life time. Time based abrasion models are used.
- *Reactive Maintenance.* The equipment is repaired in cases of failures according.

Traditionally a mixture of preventive and reactive strategies is used. Predictive and preventive maintenance may only be complementary elements to the reactive maintenance, since they left a remaining risk. This article emphases the optimization of reactive maintenance by

- decreasing the diagnosis time to find the failure causes and the repair information,
- integrate diagnosis results into the IT supported planning workflows of the companies.

An approved approach to increase diagnosis efficiency is the use of expert systems as shown in figure1. It contains a knowledge acquisition component, which is the interface for the expert. He may input new information about the respective knowledge domain into the knowledge base. Expert systems guide the user with questions to input only most

Thron M., Bangemann T. and Suchold N. (2008).
WISA - A Modular and Hybrid Expert System for Machine and Plant Diagnosis.
In *Proceedings of the Fifth International Conference on Informatics in Control, Automation and Robotics - ICSO*, pages 338-341
DOI: 10.5220/0001497003380341

relevant symptom information for determine failure causes.

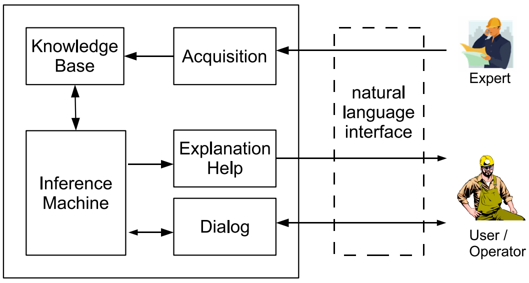

Figure 1: Structure of an expert system.

According to Moore's Law and its derivatives computer performance and memory availability doubles roughly every two years. Thus some handicaps of expert systems, regarding space and time complexity of applied algorithms, are reduced by time.

The WISA project focuses one of the remaining problems, which is the reduction of necessary effort of the experts to establish and maintain the knowledge base. This problem avoids the broad assignment of knowledge based systems in the industry.

The WISA project is funded by the German Ministry of Economy and Employment. It is registered under reg.-no. IW060215.

2 WISA SYSTEM DESCRIPTION

2.1 Standards based Knowledge Integration

Key factors of decreasing the effort for the creation and maintenance of knowledge bases are to reuse formalized knowledge and to spread the work among different experts. Another one is the formalization of the knowledge by the creators of the equipment, since they already own the necessary expertise.

Thus there is a need for a language, which makes it possible to describe knowledge bases. The complexity of this language should be small and reflect the needs of the application context (diagnosis) only. In any case it should be a standardized language to become relevant for the industry.

There are different approaches to establish a standard. The WISA project aims to be part of the following approach: *problem definition* → *finding solution ideas* → *proof of concept* → *publication* → foundation of a standardization consortium → redefinition of the proposals → reference implementation → finalization of the standard. The project leads up to the publications, which are necessary to involve candidates for a standardization consortium.

The subject for standardization is a language for description of industrial equipment. It is called Hybrid Logic Description (HLD). Design goals coming from the application domain of intelligent diagnosis where:

- handling of logic and causal dependencies between failure symptoms and their causes,
- handling of uncertain assertions about these dependencies,
- handling of vague information about the symptoms and their causes.

The HLD design includes propositional logic in combination with certainty factors, Fuzzy Logic and Bayesian networks.

2.2 Language Features of the HLD

2.2.1 Logical Description

A natural kind of causal reasoning is propositional logic. A simple sentence for diagnosis purposes could be

IF symptom1 AND symptom2 THEN cause1

The knowledge base of a propositional logic system would consist of a bunch of such rules. Variables contain truth values therein. Operators like NOT, AND, OR, IMPLICATION, BICONDITIONAL may be used in the rule sets.

Diagnosis for a broad variety of devices may be done by using propositional logic. Even good cars or TVs ship with a manual containing chapter for first aid diagnosis in a kind of propositional logic descriptions. A look into these documents shows that there is seldom a need for operators of first-order or higher-order logic like ALL- or EXISTENCE operators. Thus the WISA project focused on propositional logic for handling binary variables.

Certainty Factors are extensions to propositional logic. They were developed in the 1970s for medical diagnosis expert systems, see for example (Russell and Norvig, 2003) or (Norvig, 2004). The principal idea is that the same observable symptomatic may have different causes but with different certainty. Thus the conclusion of a rule is weighted by using a *certainty factor* (*CF*).

Certainty factors are born by empirical development. The value range of a *CF* is part of the interval $[-1,+1]$, where selected meanings for the variable value are:

```
CF = +1     => true
CF =  0     => unknown
CF = -1     => false
```

Disjunctive and conjunctive combinations of conditions are allowed in logic systems with certainty factors. A variable may be the conclusion in different rules. Thus there exists a combination calculus for merging different conclusion results.

One problem with certainty factors is that the memory space, which is necessary for the computation of the conclusion, grows with complexity $O(c^n)$ because of a necessary breadth-first search. Therein c is an approximation of the average count of conditions in the rules, whereas n is the number of rules, which are involved in the recursive computation of the conclusion. Another problem is the sloppy handling of certainties for maximization of efficiency. Bayesian Networks are additionally included into the HLD, which avoid these problems but expect input of probabilities from the expert, which are difficult to estimate.

2.2.2 Fuzzy Logic

Propositional logic deals perfectly with truth values. But in industrial plants there are symptoms and failures with continuous value domains, like temperatures, pressures, speed and other physical and chemical phenomenon.

Fuzzy Logic is a better way to deal with this kind of information than handling truth variables, which represent special value ranges. Continuous value variables are handled as linguistic variables in Fuzzy Logic. Attributes (with a representation as fuzzy set) are spread over ranges of the full variable value space. Attributes are characterized by membership functions, which map a continuous space $[0..1]$ over the related value range. This allows a more accurate reasoning in Fuzzy Logic than in propositional logic with discretized values.

An interesting point is that the notations of rules in Fuzzy Logic are very similar to the notations in propositional logic. Thus experts have to rearrange their thinking not too much when using both.

But the inference processes in Fuzzy Logic is expensive. It is done in a *fuzzyfication - inference - defuzzyfication* approach. While fuzzyfication is a straight forward computation of the membership value, inference and defuzzyfication are time consuming operations. Thus Fuzzy Logic should only be used in diagnosis scenarios where really necessary.

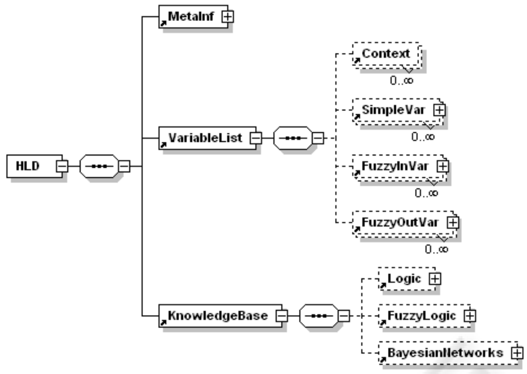

Figure 2: High level elements of the HLD specification.

2.2.3 Bayesian Networks

Bayesian Networks deal with event probabilities. They consist of directed acyclic graphs. The nodes consist of random variables arcs notify dependencies. An event is defined in a manner, that a random variable fits a specific state. According to (Jensen, 1996) there is then the possibility to declare rules in the form:

IF event A has taken place, THEN B has taken place with certainty x.

In a diagnosis applications event B may be an observed symptom, while A may be a cause. The probability of B is computable by using the Bayes' rule:

$$P(B|A,C) = \frac{P(A|B,C)P(B|C)}{P(A|C)}. \tag{1}$$

The variable C is a context variable therein. Noteworthy is that the Bayesian Network theory stands completely on a solid mathematical fundament, see (Pearl, 2000) for example. There is no approximation needed like in the Certainty Factor theory. Pearl showed, that Bayesian Networks admit d-separation. The usage of this circumstance leads to very efficient computations of the conditional probabilities in Bayesian Networks.

2.3 Language Formalization

XML has been chosen as base format for the HLD. Some requirements where:

- integrated UNICODE support,
- automatic support for a validation of instance documents,
- broad acceptance in industry.

XML is used for example as message transport format for engineering and commissioning tools (e.g. by the FDT Group and OPC Foundation), for maintenance and application integration purposes (e.g. by the MIMOSA Alliance) and for formal device descriptions (e.g. by PROFIBUS International).

Figure 2 shows the base levels of the XML-Schema based definition of the HLD. Meta information (MetaInf) are used to describe an equipment type verbally. The other elements contain the formal description of the algorithmic HLD elements as introduced in the previous sections.

2.4 Integration Concept

A machine or plant is mostly an assembly of equipment of different vendors. Thus the diagnosis descriptions have to integrate and extend descriptions of aggregated equipments. Figure 3 gives an abstract overview of assembled equipment element as considered in HLD descriptions.

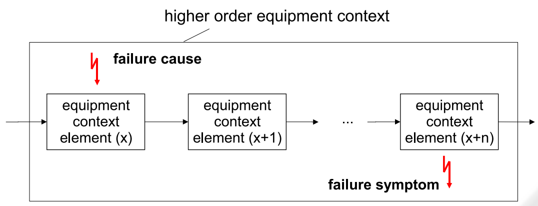

Figure 3: High level elements of the HLD specification.

An equipment is described by a single HLD file. If it contains other types of equipment then these are referenced as contexts (see also figure 2). It is possible to define rules or Bayesian Networks, which contain variables of the aggregated equipment contexts since they are referenced by context paths to variables.

2.5 Wisas - A HLD Tool Chain

It is necessary to show, that the approach leads to an efficient diagnosis, in order to gain attention and acceptance by industrial companies. But it is also necessary to show how various tools may downsize the effort for creation and maintenance of the knowledge bases. A tool chain is built in the WISA project, which is called WISAS and which consists of following components:

- *Editor*. Involved industrial partners showed that it is necessary to leave the vocabulary of AI technologies but to use the vocabulary of maintainers. This language transformation is done by a graphical editor component.
- *Validator*. This component is necessary to support creation of correct HLD files.
- *Interpreter*. This is the reasoning component for interpretation of HLD files.
- *Documentation generator*. It is assumed that industrial companies want to provide written documents for diagnosis. This component is dedicated to reach that goal.
- *Package management*. There is a need to deliver a bunch of HLD files for description of higher order equipment. There may be complex dependencies, which are handled by a package build tool and a HLD repository maintenance tool, which may use Internet connections.
- *Planning system integration*. The last but yet unfinished part of the project is to define adapters to ERP and CMMS systems.

3 CONCLUSIONS

The initial problem to reduce the effort for creation of knowledge bases for intelligent industrial diagnosis systems has been solved by a modularization of the knowledge bases. It has shown that the industrial relevance will only be reached if there will be a standardization process. Therefore the project WISA proposes a knowledge description language for the special purpose of industrial diagnosis. A prototypical tool chain has additionally been developed in order to find more acceptance at industrial companies.

There have been critical discussions about the approach especially if XML is the right base technology. Some advantages as the possibilities to validate and transform XML files has been higher rated than the disadvantage of poor parsing performance compared to special language parsers.

There are some open questions regarding time and space complexity. These questions will be answered until the end of the project in August 2008. Implementation of future HLD interpreters and editors will benefit from using on-line data, which could be used for automatic triggering a diagnosis process but also for learning procedures for the Bayesian Networks.

REFERENCES

Jensen, F. V. (1996). *An introduction to Bayesian networks*. UCL Press Limited, London, 1st edition edition.

Norvig, P. (2004). *Paradigms of artificial intelligence programming: case studies in Common Lisp*.

Pearl, J. (2000). *CAUSALITY Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, 1st edition edition.

Russell, S. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, 2nd edition edition.