INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'09
24 - 27 AUGUST 2009, STANFORD UNIVERSITY, STANFORD, CA, USA

KNOWLEDGE-BASED FEEDBACK OF PRODUCT USE INFORMATION INTO PRODUCT DEVELOPMENT

Michael Abramovici[1], Manuel Neubach[1], Madjid Fathi[2] and Alexander Holland[2]
(1) Ruhr-University Bochum, Germany (2) University of Siegen, Germany

ABSTRACT

Since real product use information is not available, Design Simulation and Design for X methods in many ways rely on assumptions regarding the product use today. These assumptions generally differ from the real conditions of product use. There are various reasons for not feeding back product use information. First, current business models lead to a loss of access to the product after sale. Second, due to their price and size appropriate sensors are only rarely embedded in the product. Third, there is a lack of an integrated framework for feeding back product use information into product development.
This paper presents a new solution approach for the integration of product use information into product development. The first part of the paper provides a summary of the developed solution. While aspects like data management and knowledge discovery have been covered in previous work, this paper focuses on the representation of empirical product use information and the use of knowledge-based inference methods in order to carry out "What-If" analyses. These can help the product developer to improve the design of next generation products.

Keywords: Product Use Information (PUI), Product Lifecycle Management (PLM), Knowledge Representation, Artificial Intelligence

1 INTRODUCTION

Design Simulation and Design for X methods and tools support the anticipation of product behavior during its use [1,2]. Unfortunately, the real conditions of the product use and environmental use conditions differ from the design assumptions. Today, the acquisition, aggregation and analysis of product field data for design purposes are fairly difficult due to different reasons.
First, within the current producer-customer business models the product suppliers do not have any access to the customer's product environment. Second, sensors embedded within products for monitoring product use parameters like loads and environmental data are rarely used due to their high price and large size. Third, an integrated conceptual framework for filtering, aggregating and analyzing field data for design feedback as well as appropriate IT infrastructures are not available.
The emerging shift within manufacturing companies from selling products to offering customer-specific product service systems (IPS²) will expand the responsibility of producers to the whole product lifecycle [3] and will facilitate an easier access to product use information. Furthermore, the progress in the miniaturization of embedded micro sensors, their price reduction as well as advances in the information technology will allow an easier capturing and processing of product use information as feedback for the development of improved products.
In this changed industrial and technological environment the project described in this paper aims at the development of a new solution for the acquisition, aggregation and analysis of product use information. This solution is based upon knowledge-based methods like Bayesian network inference and is integrated in an extended Product Lifecycle Management (PLM [4]) solution. Within the scope of a feedback cycle, product use information and deduced knowledge from previous product generations can be incorporated into the development of subsequent product generations in a target-oriented fashion and can thus provide faster product improvements, lower development costs, increased product quality and lower maintenance expenses for the use phase. The IT prototype of the proposed approach has been realized as an extension of the commercial PLM software Teamcenter Engineering (by Siemens PLM software) and validated by a use case of a rotary spindle used in a wire electrical discharge grinding (WEDG) machine.

2 OVERVIEW OF THE DEVELOPED SOLUTION APPROACH FOR INTEGRATING PRODUCT USE INFORMATION

Figure 1 provides an overview of the developed overall solution for the feedback of product use information into product development. The upper half of the figure shows the situation during the product use phase for various customers. Every customer uses another instance of the product *i* within individual environmental conditions and load scenarios (1). Many customers maintain data bases with product use information for condition monitoring purposes, which include sensor data, environmental parameters, failures and incidences of maintenance, locally and isolated from each other (2).

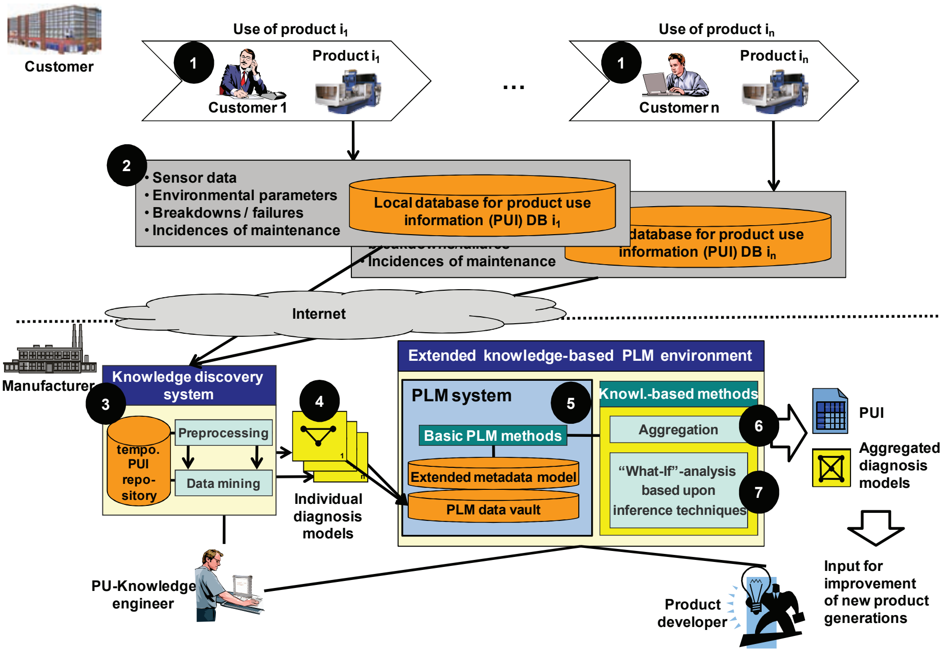


Figure 1: Overall solution for feeding back product use information into product development

So far, if not totally neglected, product use information has been used for process optimization or for the prognosis of incidences of breakdown or failure. The approach described in this paper intends that the product use information is led back to the product development in the course of a feedback cycle. This serves as a basis for deducing room for improvement and optimization for new product generations.

For this, the data generated at various customer locations first has to flow back to the manufacturer (lower half of the figure). The knowledge engineer edits the raw product use information with the help of knowledge discovery methods (3), deduces interrelationships between sensor data, environmental parameters, breakdowns and incidences of maintenance both qualitatively and quantitatively, and finally manages the resulting individual diagnosis models (4) in the PLM data archive. In this context, the metadata model forming the core of an PLM system (5) has been extended to manage not only traditional product type but also product item data. (This aspect is taken up in a further paper by the authors. It describes in detail the extension of a commercial PLM system for an integrated management of product item and product type data [5].)

By using knowledge-based methods within a PLM system it is possible to aggregate knowledge (6) from the data collected and edited with the help of knowledge discovery methods. The knowledge engineer can enter the product use information acquired from various customers in the PLM data model, generate individual diagnosis models and finally aggregate them. Thereby he can deduce a representative diagnosis model which takes differing environmental and load scenarios and their impacts on the machine condition, generally varying for individual customers, into consideration. The

mentioned aspects of learning and aggregating individual diagnosis models based on machine learning and fusion algorithms were discussed on a technical level in [6] and will not be repeated in this paper.
On the basis of inference methods the product developer in collaboration with the knowledge engineer can interactively apply the aggregated diagnosis models in order to carry out simulations and "What-If" analyses (7). For instance, definite load scenarios and environmental parameters can be set and depending on this probabilities for certain machine conditions, instances of breakdowns and recommended maintenance intervals can be calculated. All these simulations are based on data empirically gained during the product use phase. Such analyses form the basis for product developers for identifying critical components and deduce room for improvement of new product generations. The topic of representing product use information and deducing knowledge on the basis of knowledge-based methods is central to the following chapters.
The solution outlined in this section can be understood as an assistant for product developers. Deducing and realizing room for product improvement constitutes a creative process which cannot be automated with today's technologies. The aggregated diagnosis models along with analyzing techniques based on artificial intelligence methods should therefore rather support product developers in carrying out this creative process efficiently.

3 REPRESENTATION OF PRODUCT USE KNOWLEDGE

3.1 Requirements for the knowledge representation

The term product use information has already been distinguished from the purely subjective customer feedback in the introduction of this paper. Thus, no suggestions for improvement or positive/ negative reviews from customers as well as demands on future product generations expressed by users should be acquired and represented (These topics are already addressed in [7]). The focus of the represented data from the product use phase, which is to be considered within the scope of the present paper, is rather on objectively measurable information which accumulates during the use of a product.
This data strongly depends on the product examined. However, by concentrating on complex production machines and their components recurrent classes of data could be found. A structuring of the product use information to be represented into the following classes was conducted on the basis of the analysis of several case studies (stepper for the use of wafers in chip production, Wire Electrical Discharge Grinding (WEDG) machines for the electrical discharge machining of work pieces and other production machines).

- **Sensor data of the machine**: the machine parameters captured in the course of a condition monitoring are part of this class. Examples are engine speed, consumption of operating materials, machine running times, voltages etc.
- **Environmental parameters**: All objectively measurable ambient factors which have an influence on the operation of the examined machine fall into this category. Depending on the machine this can be, for instance, temperature, pressure, humidity of the ambient air.
- **Quality parameters of produced items**: complementary to the already discussed sensor data of the examined machine monitored quality parameters of manufactured items can (also via sensors) provide information about the condition of the machine. A concrete example is the proportion of functioning chips on a wafer (yield).
- **Failures/ breakdowns:** failures of components and other reduction of functioning are subsumed under this class.
- **Incidences of maintenance:** maintenances and repair measures as well as the exchange of components are part of this group.

Knowledge representation and reasoning techniques are part of the field of artificial intelligence which is concerned with how knowledge can be represented symbolically and manipulated in an automated way through reasoning. The knowledge engineer dealing with the problem context at hand has to represent this knowledge in a formalized way based on an acquisition process for the acquisition and structuring of explicit and implicit knowledge from the product use phase. In which form can the gained sensor data, environmental parameters, breakdown data and incidences of maintenance be represented in order to use them in the processing stage for deducing coherences? Which demands have to be satisfied by the methodology for knowledge representation?

Hereafter, requirement potentials are discussed which are made on the representation and the management of product use information in order to be able to deduce knowledge-based improvement potentials from individual product entities and more general on the product type level.

- **Automatic transformation of PUI in PUK**: In order to conclude improvement potentials from a knowledge-based feedback into the product development it is insufficient to provide product developers with the raw product use information (PUI) gained in the product use phase (e.g. condition monitoring data of a WEDG machine). Rather, these have to be edited intelligently and the product use knowledge (PUK) existing implicitly in the data has to be extracted. Therefore, one of the most important requirements on the methodology for knowledge representation is to enable deducing PUK as completely automated as possible from case data gained empirically during product use.
- **Integration of expert knowledge**: The integration of a priori knowledge enables an integration of basic conditions or known dependencies in the preliminary stages of the model. For instance, by this, qualitative dependencies of load scenarios, environmental parameters or component breakdowns can be incorporated into the model as given expert knowledge.
- **Representation of uncertain knowledge**: The representation of uncertain knowledge (as the result of incomplete, imprecise, inconsistent and defective data) and the handling of this type of knowledge is another important requirement especially with regard to the application domain 'product use information' since sensor data may be imprecise or missing and the interrelationships between sensor data, failures and incidences of maintenance contain a large degree of uncertainty, particularly in the case of new technologies.
- **Aggregation of PUK**: Mechanisms of aggregation and fusion are necessary for the consolidation of individual models in order to ensure a higher representativeness and relevance of the resulting knowledge representation model.
- **Suitability for inference ("What-If" Analysis)**: The support of intelligent inference techniques is especially decisive with regard to the application of the methodology for knowledge representation so that simulations and analyses (Which influence do certain environmental and load scenarios have on a particular component failure?) can be conducted on this basis.
- **Intuitive graphical visualization:** Intuitive graphic means for the visualization of the derived product use knowledge are necessary in order to support product developers.
- **Model interpretability**: The interpretability of the model and the conclusions assessed on the basis of inference techniques constitute a mental factor which should not be underestimated.

3.2 State of the art of methods and techniques for knowledge representation

In the further course of this chapter models for knowledge representation are compactly presented and compared with each other. The models are structurally subdivided into various categories and the most important models of each category are presented. Afterwards, the most relevant approaches for the structured representation of product use knowledge are evaluated. In this context, it is advisable to give a definition of knowledge representation before the individual knowledge representation models are introduced.

Knowledge representation is to be defined as a set of syntactic and semantic conventions for describing things and circumstances. The syntax specifies a set of rules which can be used for combining and grouping the symbols on which the knowledge handling is based. Thereby, expressions of the representation language can be formulated. The semantic describes the meaning of these expressions [8].

A way of structuring the most established knowledge representation forms is shown in Figure 2. In principle, it can be distinguished between procedural and non-procedural knowledge representation forms. Procedural knowledge representation forms are characterized by focusing on the description of procedures [9], while non-procedural knowledge representation forms rather concentrate on the individual knowledge elements and the relations between these elements [10].

A disadvantage of procedural knowledge representation forms often cited is the mixing of application-specific knowledge and general problem solution knowledge. As a consequence, the flexibility and the maintainability of such systems are strongly restricted, because a direct intervention in the program code becomes necessary with regard to an extension of the knowledge basis.

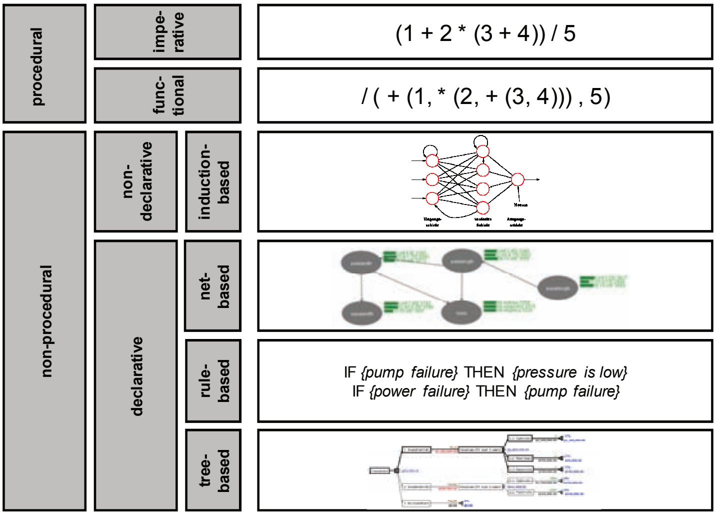


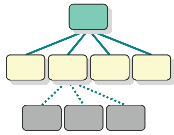
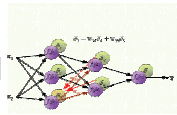
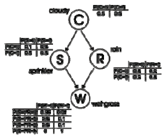
Figure 2: Structuring of knowledge representation forms and examples

Non-procedural knowledge representation forms stand out due to a clear separation of general problem solution knowledge and application-specific knowledge. Rule-based systems constitute a good example. The domain-independent inference mechanisms (algorithms for the deduction of new facts through conclusions) are contained in the problem solution component while the actual domain knowledge is kept segregated in a rule data base. This separation allows the explicit mapping of knowledge in a program logic. However, the resulting disadvantage is a higher initial effort for creating such systems. On the other hand, it leads to a considerably improved maintainability. By depositing the case-specific expert knowledge within a closed knowledge basis it becomes exchangeable, upgradeable and modifiable without having to change the program code.
On the next granulation level non-procedural knowledge representation forms can be subdivided into declarative and non-declarative knowledge representation forms. An example for declarative knowledge representation forms has already been given above: rule-based systems.
Declarative knowledge representation forms like Bayesian networks are characterized by a presentation of facts and the relations among them in order to gain new knowledge on this basis [11]. With regard to the aspect of interpretability conclusions can be transparently understood by the user. This aspect is not given for non-declarative approaches and should not be underestimated.
Artificial neural networks are a typical example for a non-declarative knowledge representation form. Artificial neural networks are capable of approximating functional coherences on the basis of case knowledge in form of training data sets [12]. Hereby, the interpretability by the user gets lost because it cannot explicitly be explained how neural networks arrive at certain conclusions. This can also complicate the acceptance of systems based on such representation forms.
Procedural knowledge representation forms are inappropriate for modeling product use knowledge deduced from PUI because they concentrate rather on sequences than on the elements themselves and the relations between them. Therefore, the following evaluation focuses on rule-based systems (RBSs), Tree-Based Models (TBMs), Artificial Neural Networks (ANNs) and Bayesian Networks (BNs) as the most relevant non-procedural knowledge representation methods.

3.3 Evaluation of methods for knowledge representation

The knowledge representation forms most relevant for modeling product use knowledge with regard to the requirement criteria learning, integration of a priori knowledge, interpretability, inference, representation of uncertain and vague knowledge, visualization and aggregation techniques presented in 3.1 are compared and critically evaluated below (see table 1). The methodology which fulfills the criteria for the knowledge-based processing of feedback information from the product use phase in the most target-oriented way will subsequently be applied in a practical scenario.

Table 1. Evaluation of various models for knowledge representation regarding the requirements made in chapter 3.1

Representation Requirements	Rule-Based Systems	Tree-Based Models	Artificial Neural Networks	Bayesian Networks
	IF {pump failure = true} THEN {pressure = low} IF {power failure = true} THEN {pump failure = true}			
Automatic transformation of PUI in PUK	◑	◔	●	●
Integration of expert knowledge	●	●	○	●
Representation of uncertain knowledge	◑	◑	◑	●
Aggregation of PUK	◑	◔	◕	●
Suitability for inference (“What-If” Analysis)	●	◑	●	●
Intuitive PUK graphical visualization	◑	◕	○	◕
Model interpretability	◕	●	○	●

Fulfillment of the requirements: ○-not, ◔-only to a small proportion, ◑-partly, ◕-mostly, ●-completely

With regard to automatic learning of knowledge on the basis of case data ANNs and BNs can play their advantages. Training data and appropriate learning algorithms (e.g. the backpropagation algorithm for ANNs [10]) enable learning general coherences from exemplary data sets as they occur in the field of product use information.

The integration of a priori expert knowledge turns out to be difficult in case of an ANN as the entire knowledge has to be learned on the basis of case studies. Serious disadvantages also arise in terms of options for the interpretation and visualization of knowledge. ANNs are not suitable for explaining an output semantically as the (artificial) neurons of the inner layers of an ANN in themselves do not possess a semantic interpretation and, hence, the ANN has to be regarded as a black box.

On the other hand, BNs offer the possibility to incorporate a priori expert knowledge into the model in addition to experimentally gained data. On a qualitative level dependencies between random variables can be modeled through manually integrated directed edges while on a quantitative level conditional probabilities can be determined by experts on the basis of theoretical insights, empirical studies and subjective estimates.

Further advantages of Bayesian networks are the representation and processing possibilities of uncertain knowledge. In contrast to other examined knowledge representation forms it is not only possible to deduce the most probable diagnosis for a given set of symptoms, but also to determine the degree of uncertainty for the deduced conclusion. Furthermore, other possible diagnoses can be calculated according to descending probability. Inconsistencies, which for example occur for RBSs due to the local treatment of the factor uncertainty, can be avoided by a holistic examination.

RBSs, KNNs as well as BNs offer the possibility of inference and hereupon based “What-If” analyses. The requirements on visualization possibilities for product use information, however, are best fulfilled by BNs because of their clearly arranged representation of qualitative coherences in terms of a directed graph.

Especially relevant are possibilities for the aggregation of several knowledge representation models in order to be able to conclude general conclusions for the entire product class from the merged model. For the aggregation of Bayesian networks several concepts and algorithms are available [6]. Postulating that the same nodes in different networks also represent the same domains, an automated aggregation is made possible. Approaches for the individual weighting of several networks during the aggregation exist as well and contribute to the creation of a representative aggregated network. Concerning neural networks algorithms for aggregation exist as well (see e.g. [13]). Here, the same

structure is assumed for all networks to be combined. The aggregation of several TBMs proves to be difficult unless there is not exactly the same topology for all trees and only associated probabilities are to be aggregated. The aggregation of several rule bases turns out to be complex as it has to be tested in a holistic approach whether additional rules lead to inconsistencies. In case of similar rules it then has to be individually determined which rules should be incorporated into the aggregated rule basis.
As measured by the requirements put forward in chapter 3.1 Bayesian networks overall appear as the most promising knowledge representation form for modeling product use knowledge especially with regard to aggregation, interpretation, inference and visualization capabilities.

4 PROPOSED KNOWLEDGE-BASED APPROACH

4.1 Use case scenario

In engineering a vast number of manufacturing methods exists which are applied by several machine tools to handle a work piece. In general, parts are formed by archetyping, transforming, disconnecting, assembling, coating or changing certain material properties.
The procedure of erosion belongs to the disconnecting procedures since, during the process of manufacturing, the cohesion of the work piece gets changed. Within the disconnecting procedures a further breakdown is carried out so that a certain type of cohesion modification is reflected. Thereby the spark erosion is assigned to the erosive procedures. Erosive procedures are characterized by not performing any mechanical action during the adaptation of the work piece. This enables a handling which takes place completely independently of any attributes of the work piece, for example without consideration of a work piece's hardness.
The principle of spark erosion is based, as the name suggests, on sparks and their thermal impacts on a work piece. Particles are separated by the sparks and afterwards removed by mechanical and/or electro-magnetic power. This process is also called Electrical Discharge Machining (EDM) [14].
Thereby the sparks emerge via electrical discharges between certain tools and the work piece and, furthermore, create a high temperature at the point of working. Besides the material removal on a work piece, additionally, there is a noticeable metal removal on the working tool. For cooling purposes and for the removal of segregated material there is a dielectric fluid. It is characterized by an especially poor conductibility and so isolates the electric wire and the work piece.

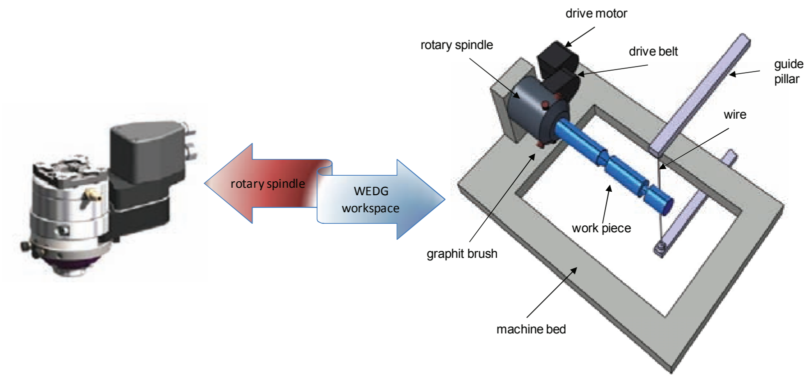


Figure 3: Rotary spindle and schematic representation of the WEDG workspace

Next, the general structure of a wire-electro discharge machine is presented. In these machines an implementing electrode in form of a wire is used, thus contact-free consigning its image on a work piece. Since the wire itself thereby incurs a certain material removal, it is continuously replenished by an engine in order to provide a constant material removal on the work piece. A generator supplies the working tool and the work piece with required voltage, so that a discharge and the associated appearance of erosive sparks become at all possible at the working point. The relative movement of wire guide and work piece is taken on by a separate control.
EDM machines working with rotary spindles are commonly called Wire Electrical Discharge Grinding (WEDG). These machines are mostly used in the manufacturing of micro-structured work pieces. Figure 3 illustrates the structure of a rotary spindle. The rotary spindle is assembled on the machine

bed and put into motion by a drive motor with a drive belt. The electrode (work piece) is fixed to the spindle with some kind of chuck and therefore rotates with exactly the same angular speed. The speed is specified by the drive motor, which, by its own, is regulated by a control instance. Simultaneously, the erosion wire is continually run by a wire guide. This represents the antipole for the electrode. In order to enable the accruement of erosion sparks the working point is continually supplied with dielectrics. The electrode on the other side gets its power from graphite brushes which force the current conduction inside the ball bearing of the rotary spindle to decrease. The drive motor's electric circuit is strictly disconnected from the electric circuit used for the erosion by an insulating plate. The rotary spindle and drive motor are continually provided with compressed air, so that any intrusion of liquids or removed material is avoided.

4.2 Bayesian networks for modeling product use knowledge

Bayesian networks can model dependencies between incidences like e.g. breakdowns or maintenance adequately on the basis of probabilistic constructs. Here, a Bayesian network represents a causal or probabilistic net which is appropriate for representing uncertain knowledge and resulting possible conclusions. It consists of a directed acyclic graph (DAG) in which nodes represent incidences as random variables and directed edges represent conditional dependencies. Every node is given a conditional probability distribution of the random variable it represents. If new critical values appear, updated probability distributions of other random variables can be calculated by means of dedicated nodes in the Bayesian network.

A Bayesian network consists of a qualitative structural and quantitative numeric component. The qualitative component represents the coherences between the random variables of the problem scenario as well as the dependencies between product use information (like conditional dependent, independent) expressed through the graph-based structure. A Bayesian network can compactly describe the common probability distribution of all involved random variables by using known conditional independencies. Qualitatively, relations of dependencies and independencies can be depicted. Every random variable X, which possesses finitely many conditions $x_1, x_2, \ldots, x_n$, is allocated a table of conditional probability distributions for every possible combination of conditions $a, \ldots, z$ of the parent nodes $A, \ldots, Z$ of X as follows:

$$P(X = x_i | a, \ldots, z) \;\forall i = 1, \ldots n. \tag{1}$$

Regarding especially the root nodes one is concerned with unconditional probability distributions as a priori distributions. If a Bayesian network consists of n random variables $X_1, X_2, \ldots, X_n$, the common probability distribution of all nodes can be expressed as follows:

$$P(X_1, X_2, \ldots, X_n) = \prod_i P(X_i | pa(X_i)). \tag{2}$$

The direct parent nodes of the random variables X_i will be expressed in this context as $pa(X_i)$. The example of the abovementioned rotary spindle presented in Figure 4 serves as illustration. Such a network structure can be achieved on the basis of empirically gained data with the help of a qualitative algorithm for learning structures of Bayesian networks [6]. For motivation and clearly arranged graphic visualization the rotary spindle, first consisting of the three random variables

R: rotation speed
M: last maintenance
B: crack of drive belt,

will be indicated with 2 conditions for each random variable.

The node *crack of drive belt (B)* has the two parent nodes *rotation speed (R)* and *last maintenance (M)*. In this rotary spindle net the node *crack of drive belt* has two conditions: *true (t)* and *false (f)*. Node *rotation speed* has the two conditions *high (h)* and *low (l)* and node *last maintenance* has the two conditions *less20d (l20)* und *greaterequal20d (ge20)*. The qualitative component of the Bayesian network is already given in Figure 4.

The quantitative component in this scenario consists of the a priori probabilities $P(R=h)$, $P(R=l)$, $P(M=l20)$ and $P(M=ge20)$ in the root nodes as well as the conditional probability tables (CPTs) which are exemplarily represented for the node *crack of drive belt* as follows:

$$\begin{pmatrix} P(B=t|R=h, M=l20) & P(B=t|R=h, M=ge20) & P(B=t|R=l, M=l20) & P(B=t|R=l, M=ge20) \\ P(B=f|R=h, M=l20) & P(B=f|R=h, M=ge20) & P(B=f|R=l, M=l20) & P(B=f|R=l, M=ge20) \end{pmatrix}. \tag{3}$$

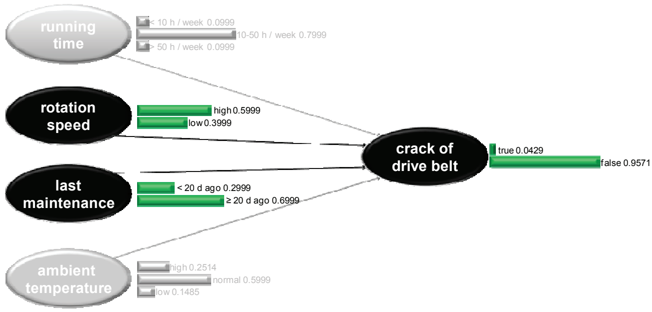


Figure 4: Influencing random variables causing the failure "crack of drive belt"

Belief values represent the confidence that a given node is in a certain condition. The initial belief value *Bel(R)* is given for the root node *R* for example through the a priori probability *P(R)*. For arbitrary nodes *X* the belief value *Bel(X)* can be declared as $P(X|O_X)$ in which O_X describes all nodes except *X*. A Bayesian network is initialized as soon as all belief values are calculated. If new information is available, the belief values have to be updated for all nodes. Efficient algorithms exist for the propagation of information (compare [15]).

To calculate the overall probability that node *crack of drive belt* is in state *true* the CPT of the node *crack of drive belt* and also the two parent nodes, *rotation speed (high/low)* and *last maintenance (l20/ge20),* are required. Thereby $P(B=t)$ can be calculated as follows:

$$\begin{aligned} P(B=t) &= P(B=t \mid R=h, M=l20)\cdot P(R=h)\cdot P(M=l20) \\ &+ P(B=t \mid R=h, M=ge20)\cdot P(R=h)\cdot P(M=ge20) \\ &+ P(B=t \mid R=l, M=l20)\cdot P(R=l)\cdot P(M=l20) \\ &+ P(B=t \mid R=l, M=ge20)\cdot P(R=l)\cdot P(M=ge20). \end{aligned} \tag{4}$$

The introduced spindle scenario can be extended to 5 random variables by adding further product use information like the environmental parameter temperature and the sensor parameter machine running time. Through information granulation the characteristics of the conditions of the random variables can be fine or coarse granularly be realized (e.g. extension of the conditions of the random variable rotation speed from two (low/ high) to five (0 min^{-1}, 30 min^{-1}, 100 min^{-1}, 300 min^{-1} and 500 min^{-1}) values). Here, it is essential to determine a balanced standard between inference with regard to interpretation time and the complexity concerning the CPT assignments of the represented Bayesian network. The rotary spindle network extended to 5 random variables cannot only be used for deducing and comparing quantity measures (for instance for determining which breakdowns happen most frequently) but also for investigating interrelationships between identified critical components and load, maintenance and environmental scenarios on the basis of "What-If" analyses more closely.

4.3 Use of "What-If" Analysis in order to deduce room for product improvements

Besides averaged distributions on how susceptible individual components are and under which load scenarios and environmental parameters the spindle operates at various customer locations, the model can also be used for simulation purposes. After identifying the failures or breakdowns which occur most frequently it is possible to conduct a detailed analysis on which factors influence a certain breakdown.

Coherences between sensor data of the rotary spindle (*rotation speed*), incidences of maintenance over time (*last maintenance*) and breakdowns of individual rotary spindle components (*crack of drive belt*) can be deduced on the qualitative as well as the quantitative level. On which parameters do the different breakdowns depend qualitatively? How does the probability for the cracked drive belt quantitatively change in the extended example from chapter 4.1, if the spindle is used more than 50 hours a week?

These and similar questions can be answered on the basis of an underlying inference engine by applying evidences like "running time of the spindle more than 50 hours a week" (compare red bar in Figure 5) and propagating them in the network. The outcome of this is a doubling of error probability with regard to the crack of the drive belt on the basis of acquired product use information.

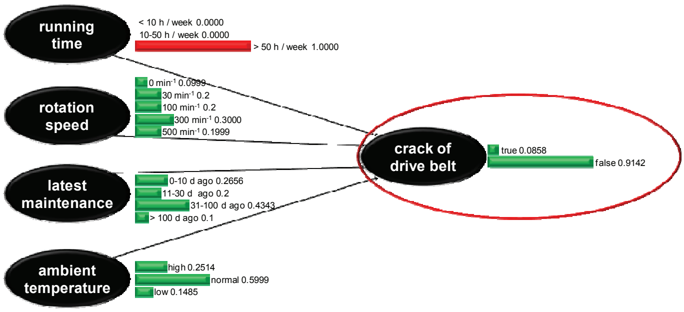


Figure 5: Influence of the weekly running time on a crack of the drive belt

However, if it is additionally known that the last maintenance has taken place 0-10 days ago, the probability for a crack of the drive belt decreases to 3,14% (see Figure 6).

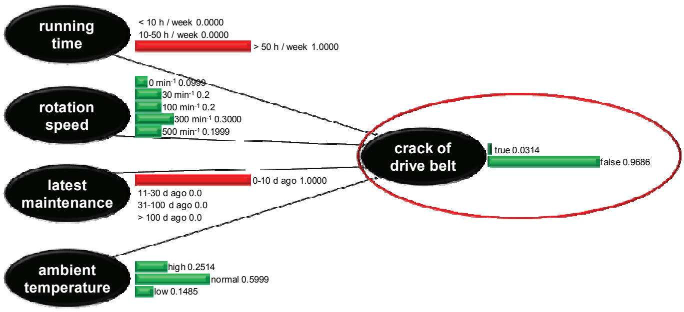


Figure 6: Influence of a running time of more than 50 hours a week on a crack of the drive belt in case the last maintenance has taken place 0-10 days ago

In case the drive belt cracks nevertheless, on the basis of the acquired data material one can assume a high probability that the rotation speed and/ or the ambient temperature have been high (compare Figure 7).

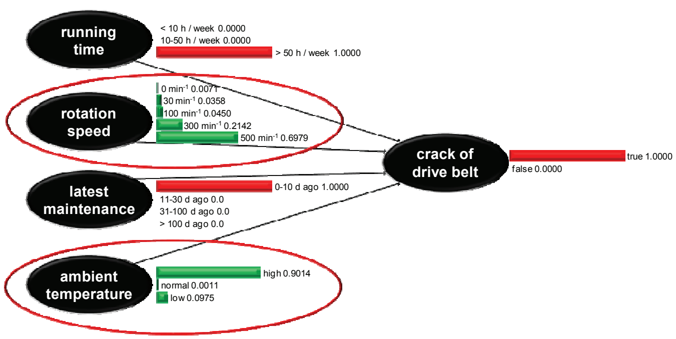


Figure 7: Most probable causes for a crack of the drive belt with regard to the set assignments for running time and date of last maintenance

While the given example based upon fictitious data sets only serves the purpose of clarifying the potential of the introduced framework by means of an easily understandable example, in practice a significantly wider range of product use information accumulates. In contrast to the given example, many of the relations between machine sensor data, environmental parameters, breakdowns and incidences of maintenance can neither be known on a qualitative nor on a quantitative level.

On the basis of machine learning algorithms and appropriate aggregation methods as presented in [6] coherences between product use information can be revealed and represented as a Bayesian network. With the help of the presented approach scenarios can also be simulated in case of complex networks

in the course of a “What-If” analysis and their impacts on relevant variables can be examined in order to support the product developer in deducing improvement potentials.
To sum up, knowledge representation models automatically established on the basis of empirically acquired product use information (below also diagnosis models) offer the following possibilities:

- Identification of the components breaking down most frequently
- Clearly arranged graphic visualization of qualitative interrelationships
- Discovering factors (scenarios of ambience, load and maintenance) that have an influence on certain breakdowns
- Deduction of quantitative dependencies on the basis of empiric product use information

Especially in case of complex machines being based on technologies which yet cannot be entirely “understood” the processed diagnosis models can serve as a basis for the deduction of design, implementation and use guidelines.

5 CONCLUSION AND OUTLOOK

The concept for a knowledge-based feedback of product use information into product development, as described in the first part of the paper, has been prototypically implemented. In this context, the PLM solution Teamcenter Engineering of Siemens PLM Software was chosen as a testbed. The functionality was enhanced with regard to creating and modifying product items, their association to the appropriate product type, linking maintenance events, condition monitoring data and diagnosis models, including support for visualization and “What-If” analyses [5].
For the appropriate representation of domain knowledge, which should also comprise breakdowns, incidences of maintenance and the dependencies between all involved elements besides sensor data and environmental parameters, in the second part of the paper various knowledge representation forms have been considered and critically evaluated with regard to the proposed requirements. Bayesian networks have proven to be the most promising model, especially in view of aggregation, interpretation, inference and visualization capabilities.
In order to evaluate the fundamental practicability of the concept an exemplary scenario has been chosen, which describes the essential steps for the knowledge-based feedback of product use information.
However, the need for an industrial testing and evaluation still exists. As the success of the concept strongly depends on the willingness of the customers to provide individual product use information, additionally, there is a high demand for a motivation concept for customers today. Nevertheless, the emerging trend at manufacturing companies as well as service providers to break with the traditional product and service understanding and to address the integrated consideration of products and services as customer-oriented overall solutions (Industrial Product-Service Systems (IPS^2) [16]) will enable IPS^2 providers to get easy access to information generated in the product use phase at various customer locations. In effect this will facilitate the industrial implementation of the presented solution approach.
The paper at hand focuses explicitly on the product area and excludes the service sector from consideration. However, the observable trend towards IPS^2 will demand an integration of additional types of feedback in the PLM concept. Basically, three different types of feedback can be acquired in addition to an IPS^2: First, product-related feedback, second, service-related feedback and third, IPS^2-related feedback.
Within the realms of product-related feedback active feedback (subjective requirements, reviews, customer satisfactions) [7] and passive feedback (objectively measurable product use information) dealt within the present paper can be distinguished. Principally, this structure can also be transferred to service-related feedback. However, it is insufficient to cover feedback which can neither be directly allocated to products nor services.
Such an IPS^2-related feedback could be, for instance, the request for a greater availability of a machine in the course of an availability-oriented business model [3], which adequately induces certain product and service adaptations. The consideration of such mutual relations between product and service shares at the moment is still an object of basic research. In this context the concept for the knowledge-based feedback of product use information into product development presented in the paper at hand is one of the puzzle pieces necessary for an IPS^2 Feedback Management.

REFERENCES

[1] ElMaraghy, H. and ElMaraghy, W. *Advances in Design, Advanced Series in Manufacturing*, 2006 (Springer-Verlag, London).

[2] Kota, S.; Chakrabarti, A. Development of a Method for Estimating Uncertainty in Evaluation of environmental Impacts during Design. In *16th International Conference on Engineering Design (ICED'07),* Paris, France, August 2007.

[3] Meier, H.; Kortmann, D. Leadership - From Technology to Use; Operation Fields and Solution Approaches for the Automation of Service Processes of Industrial Product-Service-Systems. In *14th CIRP Conference on Life Cycle Engineering, LCE 2007*, Tokyo, Japan, June 2007, pp.159-163.

[4] Abramovici, M. Future Trends in Product Lifecycle Management. In *17th CIRP Design Conference*, Berlin, Germany, March 2007.

[5] Abramovici, M.; Neubach, M.; Fathi, M.; Holland, A. Enhancing a PLM System in Regard to the Integrated Management of Product Item and Product Type Data. In *IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2008)*, Singapore, October 2008.

[6] Abramovici, M.; Neubach, M.; Fathi, M.; Holland, A. Competing Fusion for Bayesian Applications. In *12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2008)*, Málaga, Spain, June 2008, pp.378-385.

[7] Abramovici, M.; Schulte, S. Optimising Customer Satisfaction by Integrating the Customer's Voice into Product Development. In *16th International Conference on Engineering Design (ICED'07),* Paris, France, August 2007.

[8] Spur G. and Krause F.-L. *Das virtuelle Produkt – Management der CAD-Technik*, 1997 (Carl Hanser Verlag, Munich).

[9] Bullinger H.-J.; Warschat J.; Lay, K. *Künstliche Intelligenz in Konstruktion und Arbeitsplanung*, 1989 (Verlag Moderne Industrie, Landsberg/Lech).

[10] Russel S. and Norvig P. *Artificial Intelligence. A Modern Approach,* 2003 (Prentice Hall, New Jersey).

[11] Mertens P. *Expertensysteme in der Produktion*, 1990 (Oldenbourg, Munich).

[12] Haykin, S. *Neural Networks and Learning Machines*, 2008 (Prentice Hall, New Jersey).

[13] Granitto P.M.; Verdes P.F.; Ceccatto, H.A. Neural Network Ensembles: Evaluation of Aggregation Algorithms. *Artificial Intelligence*, 2005, Volume 163, Number 2, 139-162.

[14] Weck M. and Brecher C. *Werkzeugmaschinen - Maschinenarten und Anwendungsbereiche,* 2005 (Springer-Verlag, Berlin Heidelberg New York).

[15] Spiegelhalter, D.J.; Lauritzen, S.L. Probabilistic Networks and Expert Systems. Springer-Verlag Berlin Heidelberg New York, 1999.

[16] Meier, H.; Uhlmann, E.; Kortmann, D. Hybride Leistungsbündel – Nutzenorientiertes Produktverständnis durch interferierende Sach- und Dienstleistungen. In *wt Werkstattstechnik online*, 7/2005, (Springer-VDI-Verlag, Düsseldorf).

Contact: Manuel Neubach
Ruhr-University Bochum
IT in Mechanical Engineering (ITM)
Universitätsstr. 150, IB 02/140
44801, Bochum
Germany
Phone: +49 (0)234 / 32 22115
Fax: +49 (0)234 / 32 14443
Email: manuel.neubach@itm.rub.de

Manuel Neubach studied computer science at the University of Dortmund. Since 2005 he is research assistant at the Chair of Information Technology in Mechanical Engineering (ITM) of the Ruhr-University Bochum. He has worked on different research and industrial projects in the field of product information management and configuration management; in particular he is interested in transferring methods in the field of artificial intelligence to applications in engineering.