From Theory to Practice: Model-Based Diagnosis in Industrial Applications

Roxane Koitz[1] and Franz Wotawa[2]

[1,2] *Institute for Software Technology, Graz, Styria, 8010, Austria*
rkoitz@ist.tugraz.at
wotawa@ist.tugraz.at

ABSTRACT

Due to the increasing complexity of technical systems, accurate fault identification is crucial in order to reduce maintenance costs and system downtime. Model-based diagnosis has been proposed as an approach to improve fault localization. By utilizing a system model, possible causes, i.e. defects, for observable anomalies can be computed. Even though model-based diagnosis rests on solid theoretical background, it has not been widely adopted in practice. The reasons are twofold: on the one hand it requires an initial modeling effort and on the other hand a high computational complexity is associated with the diagnosis task in general. In this paper we address these issues by proposing a process for abductive model-based diagnosis in an industrial setting. Suitable models are created automatically from failure assessments available. Further, the compiled system descriptions reside within a tractable space of abductive diagnosis. In order to convey the feasibility of the approach we present results of an empirical evaluation based on several failure assessments.

1. INTRODUCTION

Fault diagnosis of technical systems has gained attention on account of safety and economic considerations in various fields such as artificial intelligence or fault detection and isolation (FDI). In order to improve diagnostic reasoning, the notion and foundations of model-based diagnosis have been investigated (Reiter, 1987; de Kleer & Williams, 1987). Model-based diagnosis as part of artificial intelligence rests on a formal description of the system to be diagnosed and derives root causes from observable anomalies. Within the last decades a solid theoretical background has been established with two approaches emerging: consistency-based and abductive diagnosis. The former depends on knowledge of the correct system behavior and infers diagnoses via inconsistencies (Reiter, 1987; de Kleer & Williams, 1987). In contrast, the abductive technique is based on models representing faults and their manifestations. It exploits the concept of entailment to compute abductive explanations for given observations (Console, Dupré, & Torasso, 1989). Although building upon different ideas, the close relationship between the two approaches has been proven (Console, Dupre, & Torasso, 1991). Cordier et al. (2004) have bridged the gap between the FDI and the consistency-based approach by investigating their relations and developing a unified framework.

Over the years there have been applications in various domains, such as space probes (Williams & Nayak, 1996), the automotive industry (Struss, Malik, & Sachenbacher, 1996), or environmental decision support systems (Wotawa, Rodriguez-Roda, & Comas, 2010). Several projects have been engaged in developing methods to integrate model-based diagnosis into industrial processes (Milde, Guckenbiehl, Malik, Neumann, & Struss, 2000; Fleischanderl, Havelka, Schreiner, Stumptner, & Wotawa, 2001). These efforts, however, mostly focus on the consistency-based method. In general, the model-based approach has not been accepted in practice, mainly due to the initial modeling and the computational complexity (Console & Dressler, 1999; Zoeteweij, Pietersma, Abreu, Feldman, & Van Gemund, 2008).

In this context we propose a process that relies first on Failure Mode Effect Analysis (FMEA) in order to develop system models while keeping knowledge acquisition affordable and second uses restricted logical formalisms where abduction is still tractable (Eiter & Gottlob, 1995; Nordh & Zanuttini, 2008). FMEA as a reliability analysis tool is growing in importance as it has been established as a mandatory task in certain industries, especially for systems that require a detailed safety assessment (Catelani, Ciani, & Luongo, 2010). An FMEA is a systematic analysis of possible component faults and the consequences said faults have on the system behavior and function (Hawkins & Woollons, 1998). Since it represents a clear causal dependency from specific fault modes

to symptoms, an FMEA provides information needed for abductive reasoning (Wotawa, 2014). Model-based diagnosis research has been concerned with the automatic creation of FMEAs based on models (Price & Taylor, 2002; Struss & Fraracci, 2012), however, we are considering the reverse process, utilizing FMEAs already available to develop suitable system descriptions.

This paper describes a process for applying abductive model-based diagnosis to an industrial setting. The remainder is structured as follows. After defining the abduction problem and presenting one algorithm capable of computing abductive explanations, we outline our suggested process in Section 3. In particular, we focus on the modeling methodology from FMEAs, show that the generated models exhibit a certain topology resulting in a manageable computational complexity, and discuss possibilities to improve the initial diagnosis results. Section 4 covers an empirical evaluation of the abductive diagnosis algorithm for models derived from multiple FMEAs. Section 5 completes the paper and argues in favor of the process' feasibility.

2. PRELIMINARIES

We assume standard definitions for propositional logic throughout this section (Chang & Lee, 2014). Abductive inference generates plausible explanations for a given set of observations by relying on the notion of entailment (Poole, Goebel, & Aleliunas, 1987). A set of premises ψ logically entails a conclusion ϕ if and only if for any interpretation in which ψ is true ϕ is also true. We write this relation as $\psi \models \phi$ and call ϕ a logical consequence of ψ. To utilize this type of reasoning, the abductive model-based diagnosis approach depends on a formalization of the relationship between faults and discoverable manifestations to derive causes for observed symptoms.

In general, abduction is an intractable problem, i.e. it cannot be solved by a polynomial-time algorithm. However, there are tractable subsets of logic, such as propositional definite Horn theory (Nordh & Zanuttini, 2008). We draw upon these findings and consider the propositional Horn clause abduction problem (PHCAP) as defined by Friedrich et al. (1990). A PHCAP links causes to effects via propositional Horn sentences. Let HC be the set of Horn clauses. Along similar lines as Friedrich et al. (1990), we define a knowledge base as a set of Horn clauses from HC over a finite set of propositional variables.

Definition 1 *A knowledge base (KB) is a tuple (A,Hyp,Th) where A denotes the set of propositional variables, Hyp $\subseteq$ A the set of hypotheses, and Th $\subseteq$ HC the set of Horn clause sentences over A.*

The set of hypotheses denotes the propositional variables which are possible causes and that we can assume to either be true or false. Later in our modeling methodology these hypotheses refer to component-based fault modes. Th represents the theory which contains rules describing the connections between hypotheses and their effects. In order to form an abduction problem, a set of observations, i.e. discovered effects, has to be considered for which explanations are to be computed.

Definition 2 *Given a knowledge base (A,Hyp,Th) and a set of observations Obs $\subseteq$ A then the tuple (A,Hyp,Th,Obs) forms a propositional Horn clause abduction problem (PHCAP).*

Definition 3 *Given a PHCAP (A,Hyp,Th,Obs). A set $\Delta \subseteq$ Hyp is a solution if and only if $\Delta \cup Th \models Obs$ and $\Delta \cup Th \not\models \bot$. A solution Δ is parsimonious or minimal if and only if no set $\Delta' \subset \Delta$ is a solution.*

A solution to a PHCAP is a set of hypotheses which logically entails the observations together with the background theory, i.e. $\Delta \cup Th \models Obs$. In addition, we require $\Delta \cup Th$ to be consistent, as from inconsistencies anything can be inferred. Considering that a solution comprises a set of hypotheses explaining the observations it is equivalent to an abductive diagnosis. While Definition 3 does not impose the limitation on the diagnosis to be minimal, in most practical applications only parsimonious solutions are of interest. Therefore, if not specified otherwise, we refer to minimal diagnoses simply as diagnoses. Notice that finding solutions to a given PHCAP is an NP-complete problem. We refer the interested reader to Friedrich, Gottlob, and Nejdl (1990) for a proof.

While there are several abductive reasoning systems, such as Theorist (Poole et al., 1987), it is well known that Assumption-Based Truth Maintenance Systems (ATMS) (de Kleer, 1986a, 1986b) are capable of deriving abductive explanations as well. The ATMS employs a graph structure where hypotheses, observations, and contradiction are represented as nodes. Implications determine the directed edges in the graph. Each node has a label assigned which contains the set of hypotheses said node can be inferred from. By updating the labels, the ATMS retains consistency. In case a single effect is observed, the label of the corresponding proposition already contains the abductive explanations. To handle multiple observations, a single rule is added, comprising a conjunction of the observations on the left hand side and a new proposition on the right hand side, i.e. $o_1 \wedge o_2 \ldots \wedge o_n \rightarrow obs$. Every set contained in the label of obs constitutes a solution to the particular PHCAP. Wotawa, Rodriguez-Roda, and Comas (2009) propose Algorithm 1 employing an ATMS and returning consistent abductive explanations.

3. PROCESS

Intercalating abductive diagnosis into real-world applications faces two major issues: constructing the domain model and the complexity of diagnosis. In this regard we define a pro-

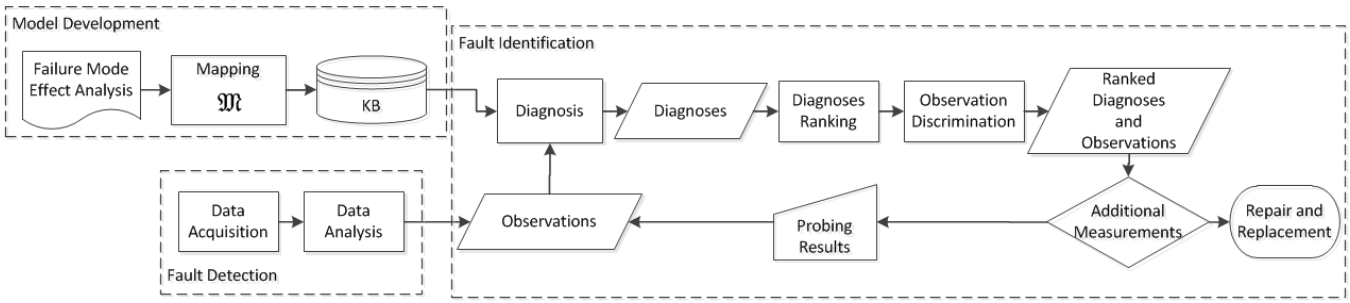


Figure 1. Process for incorporating abductive model-based diagnosis in an industrial setting.

Algorithm 1 abductiveExplanations

procedure ABDUCTIVEEXPLANATIONS (A, Hyp, Th, Obs)
  Add Th to $ATMS$
  Add $\bigwedge_{o \in Obs} o \rightarrow obs$ to $ATMS$ ▷ $obs \notin A$
  return the label of obs.
end procedure

cess to address these problems that takes advantage of information available and the structure of the resulting system descriptions. We divide the process into three main steps, as can be seen in Figure 1:

1. Model Development
2. Fault Detection
3. Fault Identification

We give a short overview of the stages and subsequently elaborate on certain parts in the following sections.

1. Model Development. As mentioned earlier abductive model-based diagnosis relies on an explicit description of the system behavior in presence of a fault. Our modeling methodology utilizes FMEAs available. As these assessments capture knowledge on failures and their symptoms, the mapping to a corresponding abductive knowledge base (KB), as defined in the previous section, is straightforward.
   Since abductive diagnosis depends on the premise of model completeness, we assume that all significant fault modes for each contributing part of the system are being considered in the analysis. Furthermore, our mapping approach expects consistent effect descriptions, i.e. a symptom is described in a uniform way throughout the FMEA. Since FMEAs usually consider single faults the resulting diagnostic system holds the single fault assumption. Note that the model can be compiled automatically offline.
2. Fault Detection. Abductive diagnosis derives possible explanations for observed anomalies, hence to initiate the diagnosis process, the presence of a fault has to be detected. Within our process, we assume the manifestation of a fault is discovered by a monitoring system and therefore do not consider the data acquisition or analysis in detail.
3. Fault Identification. Once the presence of a disturbance has been established, the possible causes associated with the observations are to be computed. Due to the knowledge represented in FMEAs, abductive diagnosis poses an intuitive approach for fault identification. We already discussed one possible algorithm capable of computing abductive diagnoses in the previous section. The process, however, is not limited to the use of this exact procedure (Koitz & Wotawa, 2015a).

In the course of this paper we explain further improvements to the initial set of solutions. In particular, we show a simple diagnoses ranking according to probability theory and how to determine the next probing point in order to diminish the number of solutions.

3.1. Model Development

The initial construction of the system description related to model-based diagnosis hinders a widespread industrial adoption. To automate this task, we propose a mapping function associating entries from an FMEA with propositional Horn clauses. Even though there are several possible representation languages suitable for diagnosis, logics provide a precise semantic of entailment necessary for abductive diagnosis.

Formally, we create a knowledge base KB. To avoid some of the inefficiencies due to complexity, we focus on a subset of logics, namely definite propositional Horn clauses. This does not impose a restriction in our case, as this representation is specific enough to capture the information contained in the FMEAs.

3.1.1. Running Example

The converter of an industrial wind turbine will act as our running example to illustrate the modeling process. Since the converter allows operation at variable speed whilst con-

Table 1. Excerpt of the FMEA of the converter.

Component	Fault Mode	Effect
Fan - Pin	Corrosion	T_cabinet, P_turbine
Fan - Bearing Running Surface	Thermo-mechanical fatigue (TMF)	T_cabinet, P_turbine
Buck Boost - Electrolyte Capacitor	Electrical chemical aging	T_power_cabinet, P_turbine, Equivalent series resistance (>)
Buck Boost - Electrolyte Capacitor	Electrical chemical aging	T_power_cabinet, P_turbine, Alarm code (over voltage, link), Equivalent series resistance (<), Electrolyte trace
IGBT - Wire Bonding	High-cycle fatigue (HCF)	T_inverter_cabinet, T_nacelle, P_turbine

necting the turbine to a constant frequency grid, it is a fundamental part of a modern industrial wind turbine. Table 1 depicts a portion of the corresponding FMEA omitting all parts concerned with reliability analysis, e.g. severity ratings. Each record contains information on a component's possible fault mode and the failure's effects. For example, $P\_turbine$ refers to a deviation between expected and measured turbine power output and $T\_cabinet$ indicates a higher than predicted temperature in the inverter cabinet (Gray, Koitz, Psutka, & Wotawa, 2015). Notice that the effect descriptions in the third column are consistent throughout the example.

We assume an FMEA comprises a set of components $COMP$, their potential fault modes $MODES$, and the set of effects which we define as a subset of the set of propositional variables $PROPS$.

Definition 4 *An FMEA is a set of tuples (C, M, E) where $C \in COMP$ is a component, $M \in MODES$ is a fault mode, and $E \subseteq PROPS$ is a set of effects.*

Since the FMEA already represents the relation between defects and their manifestations the conversion to a suitable abductive model is straightforward. The mapping function $\mathfrak{M} : 2^{FMEA} \mapsto HC$ generates corresponding propositional Horn clauses for each entry of the FMEA, i.e. rules describing the connections between a component-based fault mode and its effects.

Definition 5 *Given an FMEA, the function $\mathfrak{M}$ is defined as follows:*

$$\mathfrak{M}(FMEA) =_{def} \bigcup_{t \in FMEA} \mathfrak{M}(t) \tag{1}$$

where

$$\mathfrak{M}(C, M, E) =_{def} \{mode(C, M) \rightarrow e \,|\, e \in E\} \tag{2}$$

Hypotheses, hence all propositional variables allowed to contribute as a cause, are represented as the proposition $mode(C, M)$, where C and M relate to the corresponding component and fault mode, respectively. Equation. (3) defines Hyp in this modeling context.

$$Hyp =_{def} \bigcup_{(C,M,E) \in FMEA} \{mode(C, M)\} \tag{3}$$

Considering the running example. We would obtain the following elements for the set of hypotheses:

$$Hyp = \left\{ \begin{array}{c} mode(Fan\_Pin, Corrosion), \\ mode(Fan\_Bearing\_Running\_Surface, \\ Thermo\_mechanical\_fatigue\_(TMF)), \\ \ldots \end{array} \right\}$$

Equation (4) defines the set of propositional variables as the union of all effects and hypotheses stored in the FMEA.

$$A =_{def} \bigcup_{(C,M,E) \in FMEA} E \cup \{mode(C, M)\} \tag{4}$$

Continuing our converter example:

$$A = \left\{ \begin{array}{c} mode(Fan\_Pin, Corrosion), \\ T\_cabinet, P\_turbine, \ldots \end{array} \right\}$$

Applying $\mathfrak{M}$ results in the following theory Th completing the $KB_{Converter}$:

$$Th = \left\{ \begin{array}{c} mode(Fan\_Pin, Corrosion) \rightarrow T\_cabinet, \\ mode(Fan\_Pin, Corrosion) \rightarrow P\_turbine, \\ mode(Fan\_Bearing\_Running\_Surface, \\ Thermo\_mechanical\_fatigue\_(TMF)) \\ \rightarrow T\_cabinet, \\ \ldots \end{array} \right\}$$

It is worth noticing that Th constructed from an FMEA via $\mathfrak{M}$ features bijunctive definite Horn clauses. To ensure that contradicting observations are omitted during diagnosis, additional Horn clauses are created in Th, stating that an effect and its complement cannot occur at the same time, i.e. $e \wedge \neg e \rightarrow \bot$. For example, we would include the rule $Equivalent\_series\_resistance\_(<) \wedge Equivalent\_series\_resistance\_(>) \rightarrow \bot$ in the theory.

3.1.2. One Single Fault Diagnosis Property

The appropriateness of the models obtained from the FMEA is yet to be examined. Due to the fact that abductive explanations are consistent by definition and complete given an exhaustive search, suitability refers to the characteristic of the model that given all necessary information a single diagnosis can be computed. We refer to this feature as the One Single Fault Diagnosis Property (OSFDP).

Definition 6 *Given a KB (A, Hyp, Th). KB fulfills the*

OSFDP if the following hold:

$\forall m \in Hyp : \exists Obs \subseteq A : \{m\}$ *is a diagnosis of* (A, Hyp, Th, Obs) *and* $\neg\exists m' \in Hyp : m' \neq m$ *such that*$\{m'\}$ *is a diagnosis for the same PHCAP.*

The property can be checked by computing for each $h \in Hyp$ the set of propositions $\delta(h)$, such that $\{h\} \cup Th \models \delta(h)$. In case $\{h\} \cup Th$ leads to a contradiction, $\delta(h)$ equals $\emptyset$. If for any two hypotheses the derived propositions are the same, the OSFDP is not satisfiable. Besides determining whether single fault diagnoses can be computed, the absence of the property indicates that KB is not complete, i.e. information is missing. In the case of FMEAs this can signal that internal variables or observations have not been contemplated during the analysis. A polynomial time algorithm for testing whether the property is satisfied can be found in Wotawa (2014).

A simple procedure to enforce the OSFDP treats indistinguishable faults as a unit. Hence, each set of indistinguishable hypotheses $\{h_1, h_2, \ldots h_n\}$ is replaced by a new hypothesis h'. We proceed with these substitutions until the OSFDP is fulfilled. Algorithm 2 ensures that after termination the given KB satisfies the property. It assumes that for each hypothesis in Hyp the set $\delta(h)$ has already been computed. Due to the finite number of hypotheses as well as possible effects contained in $\delta(h)$, the procedure must halt. Further, the complexity of the algorithm is determined by the three nested loops, hence $O(|Hyp|^2 + |A - Hyp|)$.

Enforcing the OSFDP has a practical rational: Since the indistinguishable faults cannot be differentiated, all components have to be repaired or replaced in case they are part of the diagnosis. Thus, treating them as a single unit during diagnosis does not influence the result; however, it does have an effect on the computational effort because it reduces the number of possible hypotheses to consider.

Our running example of the converter does not fulfill the OSFDP, since $mode(Fan\_Pin, Corrosion)$ and $mode(Fan\_Bearing\_Running\_Surface, Thermo\_mechanical\_fatigue\_(TMF))$ are not distinguishable. By removing both hypotheses and introducing $h' = mode((Fan\_Pin, Fan\_Bearing\_Running\_Surface), (Corrosion, Thermo\_mechanical\_fatigue\_(TMF)))$ the property is fulfilled.

3.2. Fault Identification

Since the modeling methodology generates definite Horn theories, abductive reasoning is tractable (Nordh & Zanuttini, 2008). Due to the structure of the FMEAs, the resulting system descriptions are acyclic and contain solely bijunctive clauses. There are two exceptions: the formulae generated to account for contradicting observations and the implication mapping all observed effects to the proposition obs. Further

Algorithm 2 distinguishHypotheses

```
procedure DISTINGUISHHYPOTHESES (KB(A, Hyp, Th))
    Ψ[|Hyp|] ← Hyp
    for all h1 ∈ Ψ do
        for all h2 ∈ Ψ do
            if h1 ≠ h2 then
                if δ(h1) = δ(h2) and δ(h1) ≠ ∅ then
                    Create new hypothesis h'
                                                ▷ h' ∉ Hyp
                    Add h' to Ψ
                    Add h' to A
                    for all e ∈ δ(h1) do
                        Add (h' → e) to Th
                        Remove (h1 → e) from Th
                        Remove (h2 → e) from Th
                    end for
                    Remove h1 ∧ h2 from Ψ
                    Remove h1 ∧ h2 from A
                end if
            end if
        end for
    end for
    return KB(A, Ψ, Th)
end procedure
```

the intersection of the set of hypotheses and effects is empty. These features of the model all reduce the computation complexity in regard to the abduction problem. In particular, abductive diagnosis requires polynomial time in our case. For a more detailed discussion we refer the interested reader to Koitz and Wotawa (2015b).

3.2.1. Observation Discrimination

Generally, there might be an exponential number of diagnoses. In a real world context, however, a single solution is preferred. Probe selection has been proposed as a way to minimize the number of results. While other approaches assume an interleaved process between diagnosis and repair (Friedrich et al., 1990), Wotawa (2011) suggests computing all solutions and subsequently adding new symptoms which allow to discriminate diagnoses. A discriminating observation is a measurement not yet considered, which decreases the number of possible faults.

Definition 7 *Given a PHCAP (A,Hyp,Th,Obs) and two diagnoses* Δ_1 *and* Δ_2*. A new observation* $o \in A \setminus Obs$ *discriminates two diagnoses if and only if* Δ_1 *is a diagnosis for (A,Hyp,Th,Obs* $\cup \{o\}$*) but* Δ_2 *is not.*

According to information theory, the observation with the highest entropy $H(o)$ (Eq. (5)) provides the best probing point (de Kleer & Williams, 1987).

$$H(o) = -p(o) \cdot \log_2 p(o) - (1 - p(o)) \cdot \log_2(1 - p(o)) \quad (5)$$

$p(o)$ denotes the probability of observation o and is defined in Eq. (6).

Table 2. Features of the FMEAs and experimental results. For each component we conducted the experiment using the original model as well as a model fulfilling the OSFDP. The last three columns display the maximum number of single faults, double faults, and triple faults, respectively.

		Model Structure			Runtime [in ms]				#Diagnoses				
Component		**#Hyp**	**#Effects**	**#Rules**	**MIN**	**MAX**	**AVG**	**MED**	**MAX**	**AVG**	**SF**	**DF**	**TF**
Electrical circuit	Original	32	17	52	< 1	994	48.04	2	792	191.61	11	22	44
	OSFDP	15	17	35	< 1	40	0.99	1	1	1	1	1	1
Ford connector system	Original	17	17	56	< 1	204	2.08	1	18	3.14	6	18	18
	OSFDP	15	17	49	< 1	172	1.37	1	18	2.85	6	12	18
HIFI - FPU	Original	17	11	35	< 1	214	5.17	1	189	8.21	7	21	27
	OSFDP	9	11	27	< 1	18	0.83	1	12	1.59	3	6	6
MiTS1	Original	17	21	47	< 1	307	7.59	1	12	5.40	3	3	6
	OSFDP	13	21	43	< 1	312	5.38	1	1	1.00	1	1	1
MiTS 2	Original	22	15	48	< 1	191	6.60	2	288	37.44	8	24	24
	OSFDP	14	15	37	< 1	23	1.04	1	10	1.96	5	10	10
PCB	Original	10	11	24	< 1	140	1.29	0	2	1.55	2	2	2
	OSFDP	9	11	23	< 1	140	0.87	0	1	1.00	1	1	1
ACD	Original	13	16	52	< 1	210	4.47	1	15	2.44	5	8	15
	OSFDP	12	16	39	< 1	199	3.51	1	10	1.77	5	5	10
Inverter	Original	29	38	165	< 1	4830	34.80	10	1280	33.00	19	57	76
	OSFDP	23	38	124	< 1	331	9.91	4	144	6.04	13	39	26
Rectifier	Original	20	17	93	< 1	53	3.80	3	160	10.76	15	40	64
	OSFDP	14	17	66	< 1	176	4.11	2	30	3.50	9	18	24
Transformer	Original	4	8	22	< 1	70	0.73	0	4	1.17	4	2	2
	OSFDP	4	8	22	< 1	153	1.05	0	4	1.17	4	2	2
Backup components	Original	25	30	114	< 1	856	14.77	5	864	25.08	9	42	210
	OSFDP	19	30	95	< 1	172	3.67	3	90	3.53	7	30	72
Main bearing	Original	3	5	20	< 1	191	1.68	0	3	2.41	3	0	0
	OSFDP	2	5	15	< 1	184	0.84	0	2	1.41	2	0	0

$$p(o) = \frac{|\{\Delta | \Delta \in \Delta\text{-Set}, \Delta \cup Th \models \{o\}\}|}{|\Delta\text{-Set}|} \tag{6}$$

Δ-Set is the set of diagnoses obtained as a solution to the PHCAP. Once the next best probing point has been selected and the additional measurements have been taken, the probing results are passed on to the diagnosis engine as observations and the fault identification process is restarted.

3.2.2. Fault Ranking

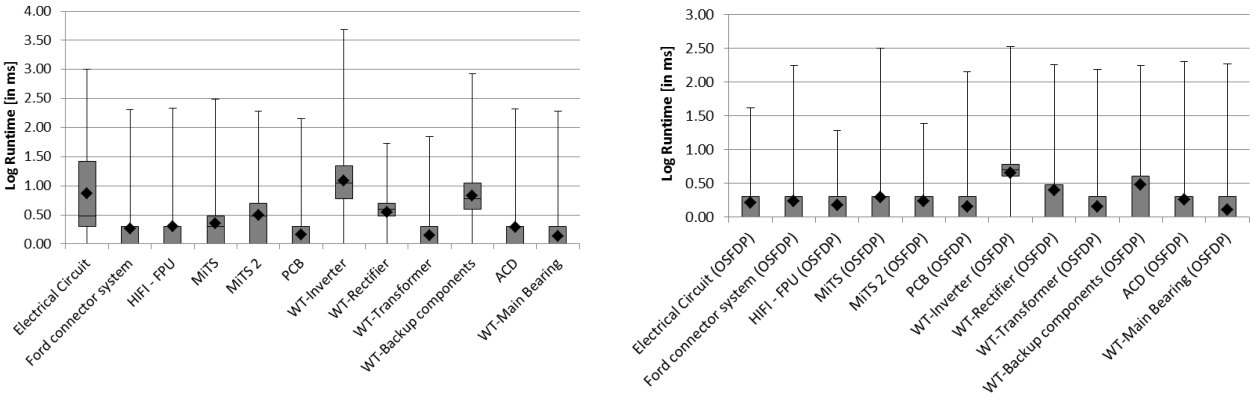
We assume independence amongst faults. Hence, the probability of a diagnosis Δ, derived from the knowledge base KB and given observations Obs, can be computed by Eq. (7).

$$p(\Delta) = \prod_{h \in \Delta} p(h) \prod_{h \notin \Delta} (1 - p(h)) \tag{7}$$

$p(h)$ represents the a-prior probability of the fault h. We presume that the fault probabilities are known, e.g. from the manufacturer or fault history analysis. Given a $PHCAP$ we compute $p(\Delta)$ for all diagnoses in Δ-Set and subsequently assign ranks correspondingly.

4. EMPIRICAL EVALUATION

In this section we report on our test scenarios and results. We obtained several publicly available as well as project internal FMEAs considering diverse technical systems and subsystems. Subsequently, we created corresponding abductive knowledge bases KB from the analyses via the mapping function $\mathfrak{M}$. The FMEAs cover electrical circuits, a connector system by Ford, the Focal Plane Unit (FPU) of the Heterodyne Instrument for the Far Infrared (HIFI) built for the Herschel Space Observatory, printed circuit boards (PCB), the Anticoincidence Detector (ACD) mounted on the Large Area Telescope of the Fermi Gamma-ray Space Telescope, the Maritim ITStandard (MiTS), as well as rectifier, inverter, transformer, main bearing, and backup components of an industrial wind turbine. As can be seen from Table 2 these FMEAs vary in the number of components and faults (i.e. hypotheses), observations (i.e. effects), as well as connections between faults and effects (i.e. rules). Note that the numbers referred to in the table correspond to the underlying FMEAs and not to the abductive model. We tested each FMEA for the OSFDP and as Table 2 reveals none, except of the model resulting from the transformer's failure assessment, of the original models satisfies the property. To generate models which fulfill the OSFDP, each set of indistinguishable hypotheses was exchanged with a new single hypothesis representing said set. Thus, the number of hypotheses and



(a) Original models

(b) Adapted models fulfilling the OSFDP

Figure 2. Box-and-whisker plots of the underlying statistical distributions of the log runtimes.

rules diminishes for the adapted models. We do not report on the computation time of the mapping, as model generation is executed offline and the conversions we have computed so far took less than a second.

After model compilation, we examined the performance of a Java implementation of Algorithm 1 on the generated KBs. The evaluation was performed on an Intel Core i7-4700MQ processor (2.60 GHz) with 8 GB RAM on Windows 7 Enterprise (64-bit). Note that our implementation utilizes an unfocussed ATMS (Forbus & de Kleer, 1988). For each FMEA we ran the algorithm for $|obs|$ from 1 to maximum number of effects possible. The observation set was generated randomly, however, we utilized the same observations for the original as well as for the adapted model. The results reported in Table 2 have been obtained from 100 trials. Unsurprisingly, the runtime increases with the number of rules to consider. As the results show while there are maximum computation times of around five seconds, the median of the distributions is located around and below ten milliseconds. Comparing the original model to the OSFDP variant, we see a performance advantage of the latter for the majority of examples. It is worth noticing that even though the transformer example already satisfied the OSFDP, the runtimes deviate. These discrepancies can be attributed to the small unit of measurement in the millisecond range.

Figure 2 depicts the underlying statistical distribution of the performance for the original and the adapted models. In order to determine whether the adapted models are superior in regard to the diagnosis performance, we used an adaptation of the sign test as described by Stumptner and Wotawa (2001). Suppose paired runtime data $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ from the original and adapted models, respectively. We propose the hypothesis $H_0 : mX = mY$, stating a median difference of zero. $H_1 : m_d > 0$ is our alternative hypothesis, where m_d denotes the median of $X - Y$. Let Z be the sum of pairs, where $x_i > y_i$. Given H_0 is true, the test statistic $Z \sim B_{0.5,n}$ has to be binomial distributed. We refute H_0 and accept H_1 if the critical value z_α is smaller than the z value from the sample. Since there is a large number of samples in our evaluation, the critical values for the sign test are not based directly on the binomial distribution, but rather on a normal approximation. For $\alpha = 0.05$ we accepted H_1, i.e. the runtime performance for the adapted models is superior to the original ones.

5. CONCLUSION

In the course of the presented research, we proposed a process to facilitate the adoption of abductive model-based diagnosis in industrial practice. FMEAs contain information suitable for abductive system descriptions and allow us to automatically generate models offline. Exploiting failure assessments is a feasible approach, as on the one hand this sort of analyses is becoming increasingly important and on the other hand the abduction problem corresponding to the contents of these documents is computationally feasible. We evaluated the resulting models on an implementation of an abductive diagnosis algorithm to identify corresponding performance trends. The results indicate that the computation times are on average under half a second. We argue that the automated modeling based on FMEAs allows for immediate reuse of information and thus provides a convenient way to employ abductive model-based diagnosis without the associated modeling effort.

ACKNOWLEDGMENT

The work presented in this paper has been supported by the FFG project Applied Model Based Reasoning (AMOR) under grant 842407. We would further like to express our gratitude to our industrial partner, Uptime Engineering GmbH.

NOMENCLATURE

α	significance level
A	set of propositional variables
ACD	Anticoincidence Detector
$ATMS$	Assumption-Based Truth Maintenance System
C	component
$COMP$	set of components
$\delta(h)$	propositions entailed by hypothesis h
Δ	diagnosis
Δ-Set	set containing all diagnoses
DF	double fault
E	set of effects
FDI	fault detection and isolation
$FMEA$	Failure Mode Effect Analysis
H	entropy
HC	set of Horn clauses
$HIFI\text{-}FPU$	Focal Plane Unit of the Heterodyne Instrument for the Far Infrared built for the Herschel Space Observatory
Hyp	set of hypotheses
KB	Knowledge Base
m	median value
M	fault mode
$\mathfrak{M}$	mapping function
$MiTS$	Maritim ITStandard
$MODES$	set of fault modes
obs	set of observations
Obs	set of all possible observations
$OSFDP$	One Single Fault Diagnosis Property
p	probability
PCB	Printed Circuit Boards
$PHCAP$	Propositional Horn Clause Abduction Problem
$PROPS$	set of propositional variables
SF	single fault
Th	theory
TF	triple fault
X, Y	random variables
Z	test statistic
z_α	critical value

REFERENCES

Catelani, M., Ciani, L., & Luongo, V. (2010). The FMEDA approach to improve the safety assessment according to the IEC61508. *Microelectronics Reliability*, *50*, 1230–1235.

Chang, C.-L., & Lee, R. C.-T. (2014). *Symbolic logic and mechanical theorem proving*. Academic press.

Console, L., & Dressler, O. (1999). Model-based diagnosis in the real world: lessons learned and challenges remaining. In *Ijcai* (pp. 1393–1400).

Console, L., Dupré, D. T., & Torasso, P. (1989). A theory of diagnosis for incomplete causal models. In *Proceedings* 11^{th} *international joint conf. on artificial intelligence* (pp. 1311–1317).

Console, L., Dupre, D. T., & Torasso, P. (1991). On the Relationship Between Abduction and Deduction. *Journal of Logic and Computation*, *1*(5), 661–690. doi: 10.1093/logcom/1.5.661

Cordier, M.-O., Dague, P., Lévy, F., Montmain, J., Staroswiecki, M., & Travé-Massuyès, L. (2004). Conflicts versus analytical redundancy relations: a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspectives. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *34*(5), 2163–2177.

de Kleer, J. (1986a). An assumption-based TMS. *Artificial Intelligence*, *28*, 127–162.

de Kleer, J. (1986b). Problem solving with the ATMS. *Artificial Intelligence*, *28*(2), 197–224.

de Kleer, J., & Williams, B. C. (1987). Diagnosing multiple faults. *Artificial Intelligence*, *32*(1), 97–130.

Eiter, T., & Gottlob, G. (1995). The complexity of logic-based abduction. *Journal of the ACM*, *42*(1), 3–42.

Fleischanderl, G., Havelka, T., Schreiner, H., Stumptner, M., & Wotawa, F. (2001). *Dike-a model-based diagnosis kernel and its application*. Springer.

Forbus, K. D., & de Kleer, J. (1988, August). Focusing the ATMS. In *Proceedings aaai* (pp. 193–198). Saint Paul, Minnesota: Morgan Kaufmann.

Friedrich, G., Gottlob, G., & Nejdl, W. (1990, September). Hypothesis classification, abductive diagnosis and therapy. In *Proceedings of the international workshop on expert systems in engineering.* Vienna: Springer Verlag, Lecture Notes in Artificial Intelligence, Vo. 462.

Gray, C. S., Koitz, R., Psutka, S., & Wotawa, F. (2015). An abductive diagnosis and modeling concept for wind power plants. In *9th IFAC symposium on fault detection, supervision and safety of technical processes.*

Hawkins, P. G., & Woollons, D. J. (1998). Failure modes and effects analysis of complex engineering systems using functional models. *Artificial Intelligence in Engineering*, *12*, 375–397.

Koitz, R., & Wotawa, F. (2015a). Finding explanations: an empirical evaluation of abductive diagnosis algorithms. In *Proceedings of the international workshop on defeasible and ampliative reasoning (DARe-15).*

Koitz, R., & Wotawa, F. (2015b). On the feasibility of abductive diagnosis for practical applications. In *9th ifac symposium on fault detection, supervision and safety of technical processes.*

Milde, H., Guckenbiehl, T., Malik, A., Neumann, B., & Struss, P. (2000). Integrating model-based diagnosis techniques into current work processes–three case

studies from the india project. *AI Communications*, *13*(2), 99–123.

Nordh, G., & Zanuttini, B. (2008). What makes propositional abduction tractable. *Artificial Intelligence*, *172*, 1245–1284.

Poole, D., Goebel, R., & Aleliunas, R. (1987). *Theorist: A logical reasoning system for defaults and diagnosis*. Springer.

Price, C. J., & Taylor, N. S. (2002). Automated multiple failure FMEA. *Reliability Engineering & System Safety*, *76*, 1–10.

Reiter, R. (1987). A theory of diagnosis from first principles. *Artificial Intelligence*, *32*(1), 57–95.

Struss, P., & Fraracci, A. (2012). Modeling hydraulic and software components for automated fmea of a braking system. In *Proceedings of the 23rd workshop on the principles of diagnosis. great malvern, uk.*

Struss, P., Malik, A., & Sachenbacher, M. (1996). Case studies in model-based diagnosis and fault analysis of car-subsystems. In *Proceedings 1st int'l workshop model-based systems and qualitative reasoning* (pp. 17–25).

Stumptner, M., & Wotawa, F. (2001). Diagnosing tree-structured systems. *Artificial Intelligence*, *127*(1), 1–29.

Williams, B. C., & Nayak, P. P. (1996). A model-based approach to reactive self-configuring systems. In *Proceedings of the national conference on artificial intelligence* (pp. 971–978).

Wotawa, F. (2011). On the use of abduction as an alternative to decision trees in environmental decision support systems. *International journal of agricultural and environmental information systems*, *2*(1), 63–82.

Wotawa, F. (2014). Failure mode and effect analysis for abductive diagnosis. In *Proceedings of the international workshop on defeasible and ampliative reasoning (DARe-14)* (Vol. 1212). CEUR Workshop Proceedings, ISSN 1613-0073. (http://ceur-ws.org/Vol-1212/)

Wotawa, F., Rodriguez-Roda, I., & Comas, J. (2009). Abductive Reasoning in Environmental Decision Support Systems. In *AIAI workshops* (pp. 270–279).

Wotawa, F., Rodriguez-Roda, I., & Comas, J. (2010). Environmental decision support systems based on models and model-based reasoning. *Environmental Engineering and Management Journal*, *9*(2), 189–195.

Zoeteweij, P., Pietersma, J., Abreu, R., Feldman, A., & Van Gemund, A. J. (2008). Automated fault diagnosis in embedded systems. In *Secure system integration and reliability improvement, 2008. ssiri'08. second international conference on* (pp. 103–110).