On Structural Properties to Improve FMEA-Based Abductive Diagnosis

Roxane Koitz* and **Franz Wotawa**
Institute for Software Technology
Graz University of Technology, Graz, Austria
e-mail: {rkoitz, wotawa}@ist.tugraz.at

Abstract

Abductive Model-Based Diagnosis (MBD) provides an intuitive approach to fault identification by reasoning on a description of the system to be diagnosed. Nevertheless, its computational complexity hinders a vast adoption and thus motivates further evaluation of efficient methods. In this paper, we investigate the structural metrics inherent to models and diagnosis problems generated on the basis of Failure Mode Effect Analysis (FMEA). Proceeding on the metrics developed, we investigate their potential as classification features to identify the most suitable diagnosis algorithm for a particular diagnosis problem. Evaluated on artificial and practical samples, our approach shows that the classifier trained on the described metrics is able to indicate the most efficient method in case of a specific diagnosis scenario.

1 Introduction

Growing complexity of technical systems complicates an effective as well as efficient fault identification, causing automated diagnosis to be of increasing interest from a theoretical as well as applied point of view. An extensive body of research has concerned itself with model-based diagnosis (MBD) [Reiter, 1987] which reasons on causes for observed anomalies using a description of the system. The abductive approach within this framework exploits a model of how failures manifest themselves within the system [Console *et al.*, 1991]. By relying on the concept of entailment, abductive reasoning provides consistent root causes for failure indicators in an intuitive way.

Abduction is not merely relevant in the field of diagnostics, but has been applied to diverse fields such as planning [Poole and Kanazawa, 1994] or ontology debugging [Lambrix *et al.*, 2013]. Various approaches to compute abductive explanations have been developed over the last decades, such as set covering [Patil *et al.*, 1982; Guan and Jiang, 2013], abductive logic programming [Kakas *et al.*, 1992; Denecker and De Schreye, 1998], proof-tree completion [McIlraith, 1998], or consequence finding [Marquis, 2000].

Within this paper we address two methods: abductive MBD and the parsimonious set covering theory. In the latter a simple diagnosis problem comprises a set of causes, manifestations, and a causal associative network connecting these disjoint sets. A diagnosis is defined as the disorders which cover, i.e. explain, a given set of symptoms. The parsimonious set covering approach has been formalized and later extended to include Bayesian probabilities [Peng and Reggia, 1990]. Several refinements to the basic theory have been proposed such as the improvement of models with additional knowledge or the inclusion of more complex covering relations [Baumeister and Seipel, 2002].

Abduction is a hard problem with an exponential number of solutions in the worst case. Even though certain model representations are tractable [Nordh and Zanuttini, 2008], computing solutions for instances of reasonable size and complexity remains a challenge. Therefore, in this paper we investigate the algorithm selection problem [Rice, 1975] for abductive diagnosis. Algorithm selection addresses the issue of choosing the best performing method for a particular problem instance and advocates for the importance of structural properties of the problem space to determine the preferred approach [Smith-Miles, 2009]. It has been applied for example in SAT solving [Xu *et al.*, 2008], graph coloring problems [Musliu and Schwengerer, 2013], or tree-decomposition [Morak *et al.*, 2012]. Our problem space is restricted to propositional Horn clause models generated from Failure Mode Effect Analysis (FMEA), which is a failure assessment available in practice. The formal descriptions of these models present certain structural traits, which are used as features in the algorithm selection process. Based on these model attributes and a set of experiments, a machine learning classifier was trained to decide on the most run time efficient abductive reasoning algorithm for a distinct diagnosis problem. We embedded the selection process within a meta-algorithm, which generates the structural metrics for a given diagnosis scenario, categorizes it on the previously trained classifier and computes the diagnoses using the "best" algorithm according to the prediction.

2 Abductive Diagnosis

Within this section we present two abductive diagnosis methods, namely abductive MBD building on propositional Horn clauses and the simple set covering approach. We show that

*Authors are listed in alphabetical order.

these two formalizations in their restricted form can be interchanged.

2.1 Model-Based Diagnosis

In abductive MBD a system description holds information on how failures affect system variables. On basis of the knowledge and given an observable anomaly, the task is to search for a set of causes, which together with the model logically entail the observation. Furthermore, the explanations have to be consistent with the underlying theory. Since abductive inference is an intractable problem, research has focused on logical subsets which allow to compute explanations in polynomial time [Eiter and Gottlob, 1995]. Based on these findings we focus on a propositional Horn clause formalization [Friedrich *et al.*, 1990].

Definition 1 (Knowledge base (KB)) *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 the set of Horn clause sentences over A.*

A knowledge base provides the underlying model for our reasoning, where the set of hypotheses comprises the propositional variables constituting a cause or explanation. The remaining propositional variables. i.e. $\{A \setminus Hyp\}$, are effects or symptoms and the theory determines the relations between hypotheses and effects.

Example 1: $Hyp = \{h_1, h_2, h_3\}$, $A = \{h_1, h_2, h_3, o_1, o_2, o_3\}$, $Th = \{\ h_1 \rightarrow o_1, h_2 \rightarrow o_1, h_2 \rightarrow o_2, h_3 \rightarrow o_2, h_3 \rightarrow o_3\ \}$

An abduction problem considers a knowledge base KB and a set of observations, i.e. current symptoms, for which explanations should be computed.

Definition 2 ***(Propositional Horn Clause Abduction Problem (PHCAP))*** *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 (Diagnosis; Solution of a PHCAP) *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 $\not\models \perp$. A solution Δ is parsimonious or minimal if and only if no set $\Delta' \subset \Delta$ is a solution.*

A solution to the PHCAP is an abductive diagnosis, as it provides hypotheses consistently explaining the occurrence of an observation.

Example 1 (cont): Let us assume we can observe o_1 and o_3, i.e. $Obs = \{o_1, o_3\}$. The solution to the PHCAP, i.e. the abductive diagnoses, are $\Delta_1 = \{h_1, h_3\}$ and $\Delta_2 = \{h_2, h_3\}$.

An Assumption-Based Truth Maintenance Systems (ATMS) [de Kleer, 1986] is able to compute abductive explanations. Based on a directed graph, where propositional variables are nodes and the edges are determined by the theory, each node owns a label which stores the hypotheses implying said node. Whenever a new clause is added to the theory, the ATMS updates its nodes' labels and further maintains a consistent state. By adding a rule of the type $o_1 \wedge \ldots \wedge o_n \rightarrow explain$ containing all elements of Obs on the left hand side and a new variable on the right hand side, the ATMS computes the abductive solution as the label of $explain$.

2.2 Set Covering

Peng and Reggia [1990] developed the parsimonious set covering theory as a formal approach to abductive diagnosis relying on an associative network of causal connections between disorders and manifestations. A simple diagnosis problem is defined as a 4-tuple $P = < D, M, C, M^+ >$, where D is the set of disorders, M denotes the set of manifestations, C describes the relations in the causation network and M^+ comprises the current set of observations. Considering the definitions of the previous subsection D refers to Hyp. Further, the manifestations M, describe the remaining propositions not included within the hypotheses which are in fact the effects. The causal relations C are given by the theory, i.e. there exists a relation between a disorder d_i and a manifestation m_j whenever there is a clause $d_i \rightarrow m_j$ contained within the theory. Since M^+ provides the distinguished subset of symptoms observed it corresponds to Obs. As the mapping between a PHCAP and a set covering problem is straightforward, we will use the wording as defined in the previous subsection, e.g. Hyp refers to the set of hypotheses, causes, disorders, etc. For clarity we add one additional set missing from the logic-based framework, i.e. M, which we define as $\{A \setminus Hyp\}$. The simple set covering model is equivalent to logic-based abduction with a theory restricted to definite Horn clauses [McIlraith, 1998].

Example 1: (cont) Considering our example from before, the diagnosis problem can be reduced to set covering: $Hyp = \{h_1, h_2, h_3\}$, $M = \{o_1, o_2, o_3\}$, $Obs = \{o_1, o_3\}$ and $Th = \{< h_1, o_1 >, < h_2, o_1 >, < h_2, o_2 >, < h_3, o_2 >, < h_3, o_3 >\}$.

In order to define a solution to a diagnosis problem within this framework, we define for every hypothesis the set *effects*$(h_i) = \{m_j | < h_i, m_j > \in Th\}$, i.e. the set of objects directly caused by h_i, and respectively for each effect, i.e. m_j, the set *causes*$(m_j) = \{h_i | < h_i, m_j > \in Th\}$, i.e. the set of objects which can directly cause m_j [Peng and Reggia, 1990]. Thus, for any subset of disorders Hyp_I, we can determine the objects directly caused by it as

$$\textit{effects}(Hyp_I) = \bigcup_{h_i \in Hyp_I} \textit{effects}(h_i)$$

Along similar lines, we can observe that

$$\textit{causes}(M_J) = \bigcup_{m_j \in M_J} \textit{causes}(m_j)$$

Example 1 (Cont): For example *causes*$(o_1) = \{h_1, h_2\}$ and *effects*$(\{h_1, h_2\}) = \{o_1, o_2\}$.

Definition 4 (Cover) *A set $Hyp_I \subseteq Hyp$ is said to cover $M_J \subseteq M$ if $M_J \subseteq$ effects(Hyp_I) and there exists no $Hyp'_I \subset Hyp_I$ with $M_J \subseteq$ effects(Hyp'_I).*

A cover relation exists between a disorder and a manifestation whenever the latter is causally inferred from the former and is subset minimal. While minimality is not a necessary condition for a cover in the original definition, we enforce this parsimonious criteria as we are only interested in minimal diagnoses [Peng and Reggia, 1990].

Definition 5 (Set Cover Diagnosis) *Given a diagnosis problem P. A set $\Delta \subseteq Hyp$ is said to be a diagnosis iff Δ covers Obs.*

Example 1 (Cont): In case we have the same observations as before, i.e. o_1 and o_3, we can obtain the set covering diagnoses by determining the disorder sets Hyp_I where *effects*(Hyp_I) cover Obs, which are *effects*$(\{h_1, h_3\}) = \{o_1, o_2, o_3\}$ and *effects*$(\{h_2, h_3\}) = \{o_1, o_2, o_3\}$. Hence, the diagnoses are $\Delta_1 = \{h_1, h_3\}$ and $\Delta_2 = \{h_2, h_3\}$.

As it has been shown previously, set covering is equivalent to the hitting set (HS) problem [Karp, 1972]. Let S be a collection of sets, then a set $h \subseteq \bigcup_{s_i \in S} s_i$ is a hitting set for S such that $\forall s_i \in S : h \cap s_i \neq \emptyset$ [Greiner *et al.*, 1989]. Since a cover states that a certain disorder causally infers a manifestation, we can utilize the set *causes*(m_j) as previously defined as a similar cover indicator. For each manifestation the set *causes*(m_j) contains the information on all disorder causing m_j. By computing the hitting set of *causes*(m_j) we derive a disjunction of all disorders included, i.e. each disorder constitutes a possible solution. In case we obtain a set of observable manifestations $m_1, \ldots, m_n \in Obs$, the hitting sets of all *causes*$(m_j) \in Obs$ comprises the diagnoses. This is apparent, as to account for all current manifestations one disorder causing each manifestation has to be present within a single solution. Again we focus on parsimonious solutions, therefore we are solely interested in subset minimal hitting sets (MHS) [Peng and Reggia, 1990].

Definition 6 (Abductive Hitting Set Diagnosis) *Given a diagnosis problem P. A set $\Delta \subseteq Hyp$ is said to be a minimal diagnosis iff Δ is a MHS of S, where $\forall m_j \in Obs$: causes$(m_j) \in S$.*

Example 1 (Cont): The *causes* sets for the current manifestations are *causes*$(o_1) = \{h_1, h_2\}$ and *causes*$(o_3) = \{h_3\}$, thus *causes*$(o_1) \in S$ and *causes*$(o_3) \in S$. The MHS of S correspond to Δ_1 and Δ_2.

3 Structural Analysis

Since developing suitable system models for MBD is a tedious task, there are approaches taking advantage of knowledge available to automatically generate system descriptions [Sterling *et al.*, 2014]. A recent technique utilizes the information captured within FMEAs as the basis of their abduction models [Wotawa, 2014]. FMEA is of increasing interest as a systematic assessment of reliability on a component level [Catelani *et al.*, 2010]. It encapsulates possible faults as well as the way they reveal themselves in different system variables [Hawkins and Woollons, 1998]. Therefore, it provides information on causal relations between failures and their symptoms, which in turn can be used as a knowledge base in an abductive diagnosis context [Wotawa, 2014].

3.1 FMEA-Based Model Development

On basis of the relation between causes and effects recorded in the FMEA, a mapping function creates a propositional KB as described in the previous section, comprising a set of variables, hypotheses, and a propositional Horn clause theory [Wotawa, 2014]. Example 2 shows an FMEA with three component-based faults and their effects. In the KB each component-fault mode pair is represented by a propositional variable $mode(C, M)$, where C is the component and M is the fault mode. These propositional variables form the set Hyp of the KB, while the set of variables A encompasses all of these hypotheses as well as propositional variables representing the recorded effects. In order to create a Horn theory Th, the mapping function generates for each row in the FMEA a set of clauses. Each clause is an implications from the hypothesis, i.e. the component-fault mode pair, to one single effect variable. Thus, for example, for the second row of Table 1 two clauses are generated. The theory then is the union over all developed Horn clauses.

Component	Fault Mode	Effect
Fan	Corrosion	P_turbine
Fan	TMF	T_cabinet, P_turbine
IGBT	HCF	T_cabinet, T_nacelle

Table 1: Simplified FMEA

Example 2: For the example given in Table 1 we derive the following set of hypotheses, variables and Horn clauses:

$$Hyp = \left\{ \begin{array}{c} mode(Fan, Corrosion), \\ mode(Fan, TMF), mode(IGBT, HCF) \end{array} \right\}$$

$$A = \left\{ \; mode(Fan, Corrosion), T\_cabinet, P\_turbine, \ldots \; \right\}$$

$$Th = \left\{ \begin{array}{c} mode(Fan, Corrosion) \rightarrow P\_turbine, \\ mode(Fan, TMF) \rightarrow P\_turbine, \\ mode(Fan, TMF) \rightarrow T\_cabinet, \\ mode(IGBT, HCF) \rightarrow T\_cabinet, \\ mode(IGBT, HCF) \rightarrow T\_nacelle \end{array} \right\}$$

For a detailed discussion on how to obtain the KB from FMEAs we refer to the paper by Wotawa [2014].

3.2 Structural Metrics

In this section we present the structural metrics extracted from the models and utilized as attributes in classification. Analyzing the structure of the Horn theory, we can observe that the models constructed on basis of FMEAs with the mapping function presented by Wotawa [2014] are bijunctive definite Horn clauses, i.e. each clause is an implication from a single proposition representing a hypothesis to a single effect variable [Wotawa, 2014]. Hence, we can easily represent the theory as an acyclic directed graph (DAG) with a forward structure from causes to effects. Which in fact is the same structure as the problems in the simple set covering theory where hypotheses and manifestations are disjoint sets.

Based on the theory, we can easily represent the model as a hypergraph $H = (V, E)$, where V is the set of vertices and E denotes the set of hyperedges with $\forall e \in E : e \subseteq V$. Concerning our models the node set comprises all propositional variables, while hyperedges are determined by the theory; for each clause there exists a hyperedge containing the propositional variables of the clause, i.e. $\forall a \in A : a \in V$ and $\forall c \in Th : \bigcup_{l \in c} |l| \in E$ where $||$ is a function mapping literals to the underlying propositions ignoring negations, i.e., $|\neg p| = p$ and $|p| = p$ for all $p \in A$. Furthermore, we can assign a label to each vertex within a hyperedge e, as follows:

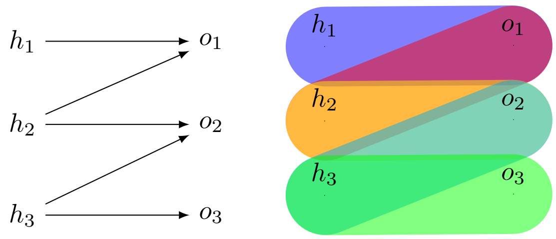


Figure 1: DAG (left) and hypergraph (right) representation of *Example 1*.

$$label(v) = \begin{cases} \{v\} & \text{if } v \in Hyp \\ \bigcup\limits_{x \in E \wedge x \neq v} label(x) & \text{otherwise} \end{cases}$$

Obviously, in case we examine the vertices representing manifestations, the labels correspond to the $causes$-sets, since both contain the hypotheses inducing the corresponding manifestation. Hence, we can utilize the label of the current observations in computing the abductive diagnoses by means of hitting sets. Note, that we choose this representation to be able to handle intermediate effects, i.e. manifestations leading to additional symptoms, which are not represented in the simple set covering theory. In this paper, however, we focus on the simple structure as inherent to models created from FMEAs.

Figure 1 shows the hypergraph built for *Example 1*. On top of the DAG and hypergraph representation of the models, we can extract certain metrics relating to the underlying structure. An intuitive measure is the number of causes in the model, since abductive reasoning is exponential in the size of Hyp. Furthermore, basic quantities include the number of effects and relations occurring in the model.

Outdegree and Indegree Considering the directed graph determined by the theory we can compute the outdegree of each hypothesis node, i.e. the number of effects caused by said node, as well as the indegree of each effect, i.e. the number of hypotheses implying said manifestation. Using the set covering formalization we can define $\mathit{outdegree}(h_i) = |\mathit{effects}(h_i)|$ and similarly $\mathit{indegree}(m_j) = |\mathit{causes}(m_j)|$. Collected over the entire model, these measures provide an intuitive metric of the basic magnitude of the theory and the connectedness of the graph.

Covering and Overlap We can easily determine a covering metric for any given pair of hypotheses by building the ratio between common manifestations and the total number of manifestations caused by the hypotheses:

$$\mathit{covering}(h_i, h_j) = \frac{|\mathit{effects}(h_i) \cap \mathit{effects}(h_j)|}{|\mathit{effects}(h_i) \cup \mathit{effects}(h_j)|}$$

In a similar manner, we define the overlap of two effects as their shared causes in relation to all their causes:

$$\mathit{overlap}(o_i, o_j) = \frac{|\mathit{causes}(o_i) \cap \mathit{causes}(o_j)|}{|\mathit{causes}(o_i) \cup \mathit{causes}(o_j)|}$$

In case there is a unique hypothesis explaining an observation, this is referred to as a pathognomonic effect [Peng and Reggia, 1990]. Whenever a pathognomonic symptom is involved, we cannot compute an overlap relation, as there cannot be any shared hypotheses.

Independent Diagnosis Subproblem Independent diagnosis subproblems occur whenever the directed graph or hypergraph are not connected, i.e. there exist subproblems within the model which have disjoint hypotheses and effect sets. Note that if all effects are pathognomonic, then each cause-effect relation represents its own diagnosis subproblem and thus the diagnosis model is orthogonal. Imagine the clause $h_2 \rightarrow o_2$ missing from the theory of *Example 1*. In this case we would have two independent diagnosis subproblems, namely one including h_1, h_2 and o_1 and the other one is h_3, o_2 and o_3.

Path Length Another measure of connectedness within the model is the path length on the hypergraph. In particular, we measure the length of paths between nodes representing hypotheses. Note, that for a model there are possibly several hypergraphs depending on the number of independent diagnosis subproblems, thus we disregard paths between nodes belonging to different diagnosis subproblems.

Observation Dependent Metrics Since not only the topology of the model is of interest, but also the structure of the current diagnosis problem, we measure the effect covering among the elements of Obs and determine the number of diagnosis subproblems involved, in case several exist.

4 Algorithm Selection and Meta-Approach

Since abductive reasoning belongs to the NP-hard problems, a research area of interest is to discover efficient methods to compute explanations. Thus, we investigate the feasibility to utilize the metrics defined in the previous section to train a classifier able to select the most time efficient approach for a particular diagnosis scenario. The general idea of the meta-algorithm based on algorithm selection is straightforward; we train a classifier on a distinct training set and whenever the diagnosis process is triggered by a detected anomaly, we retrieve the trained classifier, collect the metrics of the particular PHCAP and create the attributes for classification based on the measures. By providing these features to the machine learning algorithm, we in turn retrieve a predicted best algorithm for this scenario. Subsequently, we can instantiate the diagnosis engine with the corresponding abduction method as well as diagnosis problem and compute the abductive explanations. Algorithm 1 describes our meta-approach.

Algorithm 1 MetAB

```
procedure METAB (A, Hyp, Th, Obs)
    model_weka ← retrieveModel()
    f[] ← retrieveMetrics(A, Hyp, Th, Obs)
    algorithm ← classify(f, model_weka)
    Δ − Set ← diagnose(algorithm, A, Hyp, Th, Obs)
    return Δ − Set
end procedure
```

To evaluate the feasibility of the meta-technique we conducted experiments on two different data sets which we will

Actual \ Predicted	Boolean	ATMS	BHSTree	HST	HSDAG	Total
Boolean	186/20	22/0	25/4	0/1	0/1	233/26
ATMS	35/0	32/0	12/0	0/0	0/0	79/0
BHSTree	16/2	6/0	90/6	0/0	0/0	112/8
HST	0/0	0/0	0/1	0/0	0/0	0/1
HSDAG	0/3	0/0	0/0	0/0	0/4	0/7
Total	237/25	60/0	127/11	0/1	0/5	424/42

Table 2: Confusion Matrix for the test sets (FMEA / artificial)

explain in detail in the upcoming subsection. For the machine learning part of our meta-algorithm we employed the Waikato Environment for Knowledge Analysis (WEKA) library [Hall *et al.*, 2009], which provides a vast number of classification methods. The abductive reasoning algorithms forming our prediction categories are an ATMS as well as several HS algorithms, namely the Binary Hitting Set Tree (BHSTree) [Lin and Jiang, 2003], HSDAG [Reiter, 1987], HST [Wotawa, 2001] and the Boolean approach [Lin and Jiang, 2003]. To collect runtime data for the training and test set for classification, we exploited a Java implementation of an ATMS as well as a Java implementation[1] of BHSTree and Python implementations [Quaritsch and Pill, 2014][2] of the remaining HS methods.

4.1 Data

Our data for classification originate from two separate sources; on the one hand a small corpus of FMEAs and on the other hand generated artificial examples. The former comprises publicly available as well as internally used FMEAs recording fault knowledge from diverse domains. We automatically mapped these failure assessments to abductive knowledge bases, which we can use for logic-based as well as set covering abduction. All in all we conducted experiments on twelve FMEAs with various numbers of hypotheses ($4 \leq |Hyp| \leq 32$), effects ($5 \leq |M| \leq 30$) and clauses ($12 \leq |Th| \leq 105$). For each experiment we randomly chose the number of observations as well as the manifestations themselves from the effects available within the model. A total of 2120 experiments where conducted on the FMEA sample set.

In case of the artificial portion of our classification data, we produced examples with a varying number of hypotheses ($10 \leq |Hyp| \leq 500$), effects ($4 \leq |M| \leq 13001$) and clauses ($100 \leq |Th| \leq 13500$). Furthermore, we chose the effect overlap randomly as well as the out degree of the disorders. Due to the implementation of the artificial example generator, we do not observe several independent diagnosis subproblems within these models. We collected the data on 252 experiments with $|Obs|$ ranging from 1 to 25. Clearly, the majority of these models is larger in size than the FMEA examples and thus computationally more expensive.

With each experiment run we collected the following 27 metrics in accordance to the previous section: the number of hypotheses, number of effects, number of clauses, outdegree of hypotheses (average, maximum, standard deviation), indegree of effects (average, maximum, standard deviation), hypothesis covering for each pair of hypotheses of the entire model (average, maximum, standard deviation), effect overlap for each pair of manifestations (average, maximum, standard deviation), path length between hypotheses on basis of the hypergraph (average, maximum, standard deviation), number of independent diagnosis subproblems and based on the current observations the size of Obs, the indegree of the current observation nodes (average, maximum, standard deviation), the effect overlap (average, maximum, standard deviation) and number of independent diagnosis problems. All these metrics build our feature vector for the classification. The variable to be predicted is the algorithm, i.e. ATMS, BHSTree, HSDAG, HST, or Boolean, which would be the most efficient on the current diagnosis problem.

Each experiment data series was split into a training set comprising 80% of the data and a test set of 20%. Dividing the data, we made sure to split it in such a way that the test set comprises models of various sizes. Before selecting the classification method, we performed cross validation on several classification algorithms available in WEKA on the training data. Based on the accuracy obtained we decided to use a multilayer perceptron as the classifier for the FMEA-based models and multinomial logistic regression for the artificial examples.

	FMEA	**AI**
Classification Method	Multilayer Perceptron	Multinomial Logistic Regression
|Training Set|	1696	210
|Test Set |	424	42
Total Test Time	5ms	< 1 ms
Correctly Classified Instances	308 (72.64 %)	30 (71.43 %)
Incorrectly Classified Instances	116 (27.36 %)	12 (28.57 %)
Mean absolute error	0.17	0.17

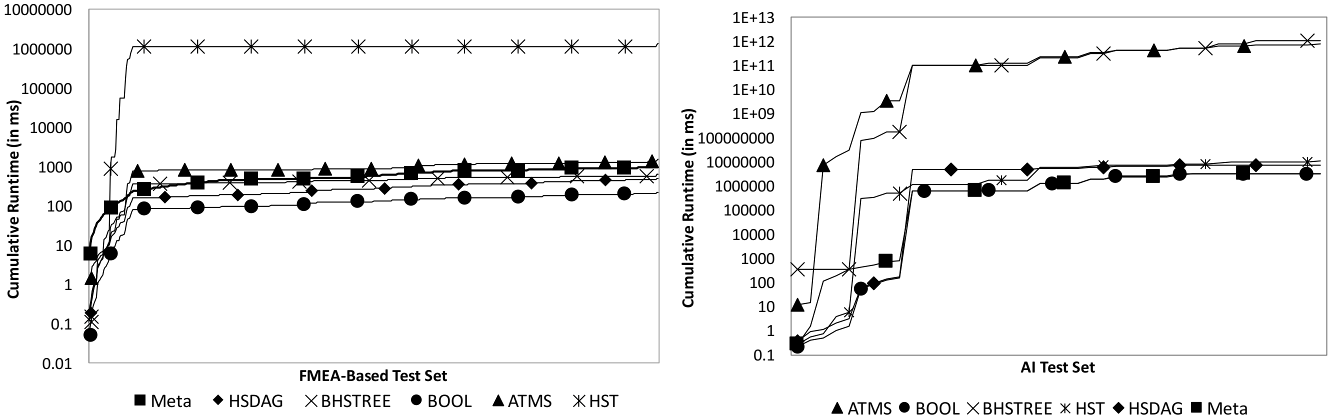
Table 3: Classification Statistics

4.2 Evaluation Results

As can be seen in Table 3 the classification based on the metrics reaches a satisfactory success rate on the FMEA-based as well as artificial examples. The confusion matrix in Table 2 shows the number of correctly and wrongly classified instances. The rows represent the actual number of instances, i.e. the number of samples the corresponding algorithm was the most efficient, while the columns show the predicted outcome. For example, the cell in the first row in column three states "25/4", meaning that for the FMEA-based samples 25 times the classifier predicted BHSTree to be the most effi-

[1]http://www.ist.tugraz.at/modremas/index.html

[2]http://modiaforted.ist.tugraz.at/downloads/pymbd.zip



(a) FMEA-based samples

(b) Artificial samples

Figure 2: Cumulative runtime for the test sets

cient approach when in fact the Boolean algorithm was the fastest. For the artificial samples 4 examples where wrongly classified as BHSTree instead of the Boolean approach.

Note that for the FMEA samples HST and HSDAG were never predicted or actually measured to be the best performing algorithm, the same holds for the ATMS in the artificial examples. Further, we can see that the Boolean approach was the most efficient, followed by BHSTree on both test sets. The confusion matrix further shows that the neural network had difficulties in classifying instances where the ATMS succeeded, as it categorizes it in nearly 41% incorrectly and the same holds for the HSDAG in the artificial examples. Yet, whenever the classifier categorizes the problem incorrectly, the suggested algorithm is the second or third most efficient.

To discover whether our meta-approach provides an efficiency improvement, we compared computation time on both test sets for all methods, i.e. our meta-algorithm and each abductive reasoning technique. The runtime for the meta-algorithm is determined by first, the computation of the metrics, second the time it takes to create the feature vector, supply it to the classifier and predicting the best algorithm and third the diagnosis time of the suggested procedure. Figure 2 shows the cumulative runtime for the test data with the x-axis representing samples from the test set. As can be seen from the figure, our meta-algorithm is not able to outperform on average (2.22 ms) all direct diagnosis methods (Boolean approach 0.48 ms, BHSTree 1.36 ms, HSDAG 1.09 ms, ATMS 3.14 ms, HST 2643.17ms) for the FMEA-based examples. The reason being that the mean time to collect the metrics of the PHCAP requires 1.71 ms on the tested examples, which is close the actual diagnosis time of these problems. In case of the artificial examples our approach performs well, i.e. on average the meta-algorithm is the most efficient. Due to the size of these samples, the computation of the properties only demands a fraction of the actual diagnosis run time. On average our algorithm requires 72674.36 ms and is 99.9% faster than the ATMS and BHSTree, 71.53 % faster then HST, 56.58% faster than the HSDAG and 2.77% faster than the Boolean approach.

5 Conclusion and Future Work

Computation time of abductive diagnosis depends primarily on the underlying model. Thus, we investigated algorithm selection, as a form to predict the best performing method based on the structural properties of problem instances. We explore metrics inherent to the structure of FMEA-based models, which form the feature vector for a classifier as part of a meta-approach. Evaluated on a test set, the metrics led to a satisfactory selection of the best algorithm for a particular diagnosis problem. In case of the FMEA-based samples our meta-algorithm was restrained by the time necessary to construct the feature vector and thus could not outperform all abduction methods. Our approach shows its value when operating on larger problem instances, where it performs well and in fact is the most efficient in comparison.

Despite the satisfactory classification results, there are certain metrics worth investigating such as treewidth based on the decomposition of the hypergraph as well as different attribute combinations to determine the metrics most suited. Regarding the restricted representation class, we plan on expanding our approach to more expressive models.

Acknowledgments

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

References

[Baumeister and Seipel, 2002] J. Baumeister and D. Seipel. Diagnostic reasoning with multilevel set-covering models. Technical report, DTIC Document, 2002.

[Catelani *et al.*, 2010] M. Catelani, L. Ciani, and V. Luongo. The FMEDA approach to improve the safety assessment according to the iec61508. *Microelectronics Reliability*, 50:1230–1235, 2010.

[Console *et al.*, 1991] Luca Console, Daniele Theseider Dupre, and Pietro Torasso. On the Relationship Between Abduction and Deduction. *Journal of Logic and Computation*, 1(5):661–690, 1991.

[de Kleer, 1986] Johan de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28:127–162, 1986.

[Denecker and De Schreye, 1998] Marc Denecker and Danny De Schreye. SLDNFA: an abductive procedure for abductive logic programs. *The journal of logic programming*, 34(2):111–167, 1998.

[Eiter and Gottlob, 1995] Thomas Eiter and Georg Gottlob. The complexity of logic-based abduction. *Journal of the ACM (JACM)*, 42(1):3–42, 1995.

[Friedrich *et al.*, 1990] Gerhard Friedrich, Georg Gottlob, and Wolfgang Nejdl. Hypothesis classification, abductive diagnosis and therapy. In *Expert Systems in Engineering Principles and Applications*, pages 69–78. Springer, 1990.

[Greiner *et al.*, 1989] Russell Greiner, Barbara A Smith, and Ralph W Wilkerson. A correction to the algorithm in reiter's theory of diagnosis. *Artificial Intelligence*, 41(1):79–88, 1989.

[Guan and Jiang, 2013] Qixue Guan and Yueqiu Jiang. The parsimonious covering theory applied for communication equipment malfunction. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2013 10th International Conference on*, pages 637–641. IEEE, 2013.

[Hall *et al.*, 2009] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I.H. Witten. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.

[Hawkins and Woollons, 1998] Peter G. Hawkins and Davis J. Woollons. Failure modes and effects analysis of complex engineering systems using functional models. *Artificial Intelligence in Engineering*, 12:375–397, 1998.

[Kakas *et al.*, 1992] Antonis C. Kakas, Robert A. Kowalski, and Francesca Toni. Abductive logic programming. *Journal of logic and computation*, 2(6):719–770, 1992.

[Karp, 1972] Richard M Karp. *Reducibility among combinatorial problems*. Springer, 1972.

[Lambrix *et al.*, 2013] Patrick Lambrix, Fang Wei-Kleiner, Zlatan Dragisic, and Valentina Ivanova. Repairing missing is-a structure in ontologies is an abductive reasoning problem. In *WoDOOM*, pages 33–44, 2013.

[Lin and Jiang, 2003] Li Lin and Yunfei Jiang. The computation of hitting sets: Review and new algorithms. *Information Processing Letters*, 86(4):177–184, 2003.

[Marquis, 2000] Pierre Marquis. Consequence finding algorithms. In *Handbook of Defeasible Reasoning and Uncertainty Management Systems*. Springer, 2000.

[McIlraith, 1998] Sheila A McIlraith. Logic-based abductive inference. *Knowledge Systems Laboratory, Technical Report KSL-98-19*, 1998.

[Morak *et al.*, 2012] Michael Morak, Nysret Musliu, Reinhard Pichler, Stefan Rümmele, and Stefan Woltran. Evaluating tree-decomposition based algorithms for answer set programming. In *Learning and Intelligent Optimization*, pages 130–144. Springer, 2012.

[Musliu and Schwengerer, 2013] Nysret Musliu and Martin Schwengerer. Algorithm selection for the graph coloring problem. In *Learning and Intelligent Optimization*, pages 389–403. Springer, 2013.

[Nordh and Zanuttini, 2008] Gustav Nordh and Bruno Zanuttini. What makes propositional abduction tractable. *Artificial Intelligence*, 172(10):1245–1284, 2008.

[Patil *et al.*, 1982] Ramesh S Patil, Peter Szolovits, and William B Schwartz. Modeling knowledge of the patient in acid-base and electrolyte disorders. *Artificial Intelligence in Medicine*, 345348, 1982.

[Peng and Reggia, 1990] Yun Peng and James A Reggia. *Abductive inference models for diagnostic problem-solving*. Springer, 1990.

[Poole and Kanazawa, 1994] David Poole and Keiji Kanazawa. A decision-theoretic abductive basis for planning. In *AAAI Spr. Symp. on Decision-Theoretic Planning*, 1994.

[Quaritsch and Pill, 2014] Thomas Quaritsch and Ingo Pill. Pymbd: A library of mbd algorithms and a light-weight evaluation platform. *Proceedings of Dx-2014*, 2014.

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

[Rice, 1975] John R Rice. The algorithm selection problem. *Advances in Computers*, 15, 1975.

[Smith-Miles, 2009] Kate A Smith-Miles. Cross-disciplinary perspectives on meta-learning for algorithm selection. *ACM Computing Surveys*, 41(1):6, 2009.

[Sterling *et al.*, 2014] Raymond Sterling, Peter Struss, Jesús Febres, Umbreen Sabir, and Marcus M Keane. From modelica models to fault diagnosis in air handling units. In *10th International Modelica Conference*, 2014.

[Wotawa, 2001] Franz Wotawa. A variant of reiter's hitting-set algorithm. *Information Processing Letters*, 79(1):45–51, 2001.

[Wotawa, 2014] Franz Wotawa. Failure mode and effect analysis for abductive diagnosis. In *Proc. Intl. Workshop on Defeasible and Ampliative Reasoning (DARe-14)*, 2014.

[Xu *et al.*, 2008] Lin Xu, Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Satzilla: portfolio-based algorithm selection for sat. *Journal of Artificial Intelligence Research*, pages 565–606, 2008.