Reducing Model-Based Diagnosis to Knowledge Base Debugging

Patrick Rodler, Konstantin Schekotihin

Alpen-Adria Universität Klagenfurt

Contents

Overview

- Model-based Diagnosis (MBD) is a principled approach to fault localization in any type of system that can be described in a formal structured way
- Knowledge Base Debugging (KBD) draws on concepts from MBD to find faults in a monotonic knowledge base
- We show that KBD is a generalization of MBD in that
  - any MBD problem can be reduced to a KBD problem
  - solutions of the MBD problem can be directly extracted from solutions of the KBD problem
- The sequential MBD problem is a special case of the sequential KBD problem in that the latter allows a user to provide more types of measurements and specify additional requirements (beyond consistency)
- Consequently: KBD approaches can be applied to all systems amenable to MBD

MBD I

Definition 1 (System)

A *system* is a tuple $(\text{SD}, \text{COMPS})$ where SD, the system description, is a set of first-order sentences, and COMPS, the system components, is a finite set of constants $c_1, \ldots, c_n$.

Let $\text{SD}_{beh} := \{\neg \text{AB}(c) \rightarrow beh(c) \mid c \in \text{COMPS}\}$ where $beh(c)$ denotes the first-order sentence describing the expected behavior of $c \in \text{COMPS}$. General axioms describing the system domain or descriptions of the interplay between the system components are comprised by SD_{gen}. So, $\text{SD} = \text{SD}_{beh} \cup \text{SD}_{gen}$.

Definition 2 (MBD-DPI)

Let OBS (observations) be a finite set of first-order sentences, MEAS (measurements) be a finite set including finite sets m_i of first-order sentences, and $(\text{SD}, \text{COMPS})$ be a system. Then $(\text{SD}, \text{COMPS}, \text{OBS}, \text{MEAS})$ is an *MBD diagnosis problem instance (MBD-DPI)*.

MBD II

Definition 3 ($\mathrm{SD}^*[\Delta]$)

Let *DPI* $:= (\mathrm{SD}, \mathrm{COMPS}, \mathrm{OBS}, \mathrm{MEAS})$ be an MBD-DPI and U_{MEAS} denote the union of all $m \in \mathrm{MEAS}$. Then $\mathrm{SD}^*[\Delta] :=$ $\mathrm{SD} \cup \{\mathrm{AB}(c) \mid c \in \Delta\} \cup \{\neg\mathrm{AB}(c) \mid c \in \mathrm{COMPS} \setminus \Delta\} \cup \mathrm{OBS} \cup U_{\mathrm{MEAS}}$ for $\Delta \subseteq \mathrm{COMPS}$ denotes the behavior description of the system $(\mathrm{SD}, \mathrm{COMPS})$

- under the current state of knowledge given by the *DPI* in terms of OBS and MEAS, and
- under the assumption that all components in $\Delta \subseteq \mathrm{COMPS}$ are faulty and all components in $\mathrm{COMPS} \setminus \Delta$ are healthy.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT I WIEN GRAZ

MBD III

Definition 4 (MBD-Diagnosis)

Let *DPI* $:= (\textsc{sd}, \textsc{comps}, \textsc{obs}, \textsc{meas})$ be an MBD-DPI. Then $\Delta \subseteq \textsc{comps}$ is an *MBD-diagnosis for DPI* iff $\textsc{sd}^*[\Delta]$ is consistent (i.e. Δ explains $\textsc{obs}$ and $\textsc{meas}$). An MBD-diagnosis Δ for *DPI* is called *minimal* iff there is no MBD-diagnosis Δ' for *DPI* such that $\Delta' \subset \Delta$.

Problem 1 (Sequential MBD)

Given: An MBD-DPI *DPI* $:= (\textsc{sd}, \textsc{comps}, \textsc{obs}, \textsc{meas})$ and a diagnostic goal G. **Find:** $\textsc{meas}_{new} \supseteq \emptyset$ and Δ, where $\textsc{meas}_{new}$ is a set of new measurements such that Δ is a minimal MBD-diagnosis for the MBD-DPI $DPI_{new} := (\textsc{sd}, \textsc{comps}, \textsc{obs}, \textsc{meas} \cup \textsc{meas}_{new})$ and Δ satisfies G.

Remark: Examples for diagnostic goals G are the presence of one highly probable or just a single remaining (minimal) diagnosis.

KBD – Diagnosis Problem Instance

Definition 5 (KBD-DPI)

Let

- $\mathcal{K}$ be a KB,
- P, N be sets including sets of sentences,
- $R \supseteq \{\text{consistency}\}$ be a set of (logical) requirements,
- $\mathcal{B}$ be a KB such that $\mathcal{K} \cap \mathcal{B} = \emptyset$ and $\mathcal{B}$ satisfies all requirements $r \in R$,
- the cardinality of all sets $\mathcal{K}, \mathcal{B}, P, N$ be finite, and
- all sets $\mathcal{K}, \mathcal{B}, P, N$ be formulated over some monotonic logic.

Then we call the tuple $\langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$ a *KBD diagnosis problem instance (KBD-DPI)*.

Solution KB

Definition 6 (Solution KB)

Let $DPI := \langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$ be a KBD-DPI. Then a KB $\mathcal{K}^*$ is called *solution KB w.r.t. DPI* iff all the following conditions hold:

$$\forall r \in R \quad : \quad \mathcal{K}^* \cup \mathcal{B} \text{ fulfills } r \tag{1}$$

$$\forall p \in P \quad : \quad \mathcal{K}^* \cup \mathcal{B} \models p \tag{2}$$

$$\forall n \in N \quad : \quad \mathcal{K}^* \cup \mathcal{B} \not\models n. \tag{3}$$

A solution KB $\mathcal{K}^*$ w.r.t. *DPI* is called *maximal* iff there is no solution KB $\mathcal{K}'$ w.r.t. *DPI* such that $\mathcal{K}' \cap \mathcal{K} \supset \mathcal{K}^* \cap \mathcal{K}$ (i.e. $\mathcal{K}^*$ has a set-maximal intersection with $\mathcal{K}$ among all solution KBs).

KBD-Diagnosis and KBD-Conflict

Definition 7 (KBD-Diagnosis)

Let $DPI := \langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$ be a KBD-DPI. A set of sentences $\mathcal{D} \subseteq \mathcal{K}$ is called a *KBD-diagnosis w.r.t. DPI* iff $(\mathcal{K} \setminus \mathcal{D}) \cup U_P$ is a solution KB w.r.t. *DPI* (i.e. $\mathcal{K}^* := (\mathcal{K} \setminus \mathcal{D}) \cup U_P$ satisfies (1) – (3)). A KBD-diagnosis $\mathcal{D}$ w.r.t. *DPI* is *minimal* iff there is no $\mathcal{D}' \subset \mathcal{D}$ such that $\mathcal{D}'$ is a KBD-diagnosis w.r.t. *DPI*.

Definition 8 (KBD-Conflict)

Let $DPI := \langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$ be a KBD-DPI. A set of formulas $\mathcal{C} \subseteq \mathcal{K}$ is called a *KBD-conflict w.r.t. DPI* iff $\mathcal{C} \cup U_P$ is not a solution KB w.r.t. *DPI* (i.e. $\mathcal{K}^* := \mathcal{C} \cup U_P$ violates at least one of (1) – (3)). A KBD-conflict $\mathcal{C}$ w.r.t. *DPI* is *minimal* iff there is no $\mathcal{C}' \subset \mathcal{C}$ such that $\mathcal{C}'$ is a KBD-conflict w.r.t. *DPI*.

Canonical Solution KB

In general, the (maximal) solution KB resulting from the deletion of one and the same set $\mathcal{D}$ from $\mathcal{K}$ is not unique. Since

- P does not justify the inclusion of sentences (semantically) different from U_P, and
- only one solution KB is sought

we define:

Definition 9 (Canonical Solution KB)

$(\mathcal{K} \setminus \mathcal{D}) \cup U_P$ is the *canonical solution KB for* $\mathcal{D}$ *w.r.t. DPI* iff $(\mathcal{K} \setminus \mathcal{D}) \cup U_P$ is a solution KB w.r.t. *DPI*.

Canonical Solution KB vs. KBD-Diagnosis

The relationship between maximal canonical solution KBs and minimal KBD-diagnoses w.r.t. a DPI is as follows (cf. [Rodler, 2015]):

Property 1

Let *DPI* be a KBD-DPI. Then the set of all maximal canonical solution KBs w.r.t. *DPI* is given by

$$\{(\mathcal{K} \setminus \mathcal{D}) \cup U_P \mid \mathcal{D} \text{ is a minimal KBD-diagnosis w.r.t. } DPI\}$$

Therefore, KBD methods focus on the computation of minimal KBD-diagnoses in order to find all maximal canonical solution KBs.

KBD-Diagnoses, Solution KBs and KBD-Conflicts

The relationship between the notions *KBD-diagnosis*, *solution KB* and *KBD-conflict* is as follows:

Property 2

Let $\mathcal{D} \subseteq \mathcal{K}$. Then the following statements are equivalent:

1. $\mathcal{D}$ is a KBD-diagnosis w.r.t. $\langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$.
2. $(\mathcal{K} \setminus \mathcal{D}) \cup U_P$ is a solution KB w.r.t. $\langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$.
3. $(\mathcal{K} \setminus \mathcal{D})$ is not a KBD-conflict w.r.t. $\langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$.

Sequential KB Debugging

The sequential KBD problem which seeks a set of test cases in order to achieve a diagnostic goal G is defined as follows:

Problem 2 (Sequential KBD)

Given: A KBD-DPI $DPI := \langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$ and a diagnostic goal G.
Find: $P_{new}, N_{new} \supseteq \emptyset$ and $\mathcal{D}$, where P_{new}, N_{new} are sets of positive and negative test cases, respectively, such that $\mathcal{D}$ is a minimal KBD-diagnosis w.r.t. $DPI_{new} := \langle \mathcal{K}, \mathcal{B}, P \cup P_{new}, N \cup N_{new} \rangle_R$ and $\mathcal{D}$ satisfies G.

Reducing MBD to KBD

Any MBD-DPI can be modeled as a KBD-DPI, and the solutions of the latter directly yield the solutions of the former:

Theorem 1 (Reduction of MBD to KBD)

Let *mDPI* $:= (\mathrm{SD}, \mathrm{COMPS}, \mathrm{OBS}, \mathrm{MEAS})$ be an MBD-DPI where $\mathrm{COMPS} = \{c_1, \ldots, c_n\}$. Then:

- *mDPI* can be formulated as a KBD-DPI *kDPI* such that there is a bijective correspondence between KBD-diagnoses for *kDPI* and MBD-diagnoses for *mDPI*.
- All MBD-diagnoses for *mDPI* can be computed from the KBD-diagnoses for *kDPI*.

Reducing MBD to KBD

Proof:

We first show how *mDPI* can be formulated as a KBD-DPI *kDPI*. To this end, we specify how $kDPI = \langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$ can be written in terms of the components of $mDPI = (\mathrm{SD}_{beh} \cup \mathrm{SD}_{gen}, \mathrm{COMPS}, \mathrm{OBS}, \mathrm{MEAS})$:

$$\mathcal{K} = \{\alpha_i \mid \alpha_i := beh(c_i), c_i \in \mathrm{COMPS}\} \quad (4)$$

$$\mathcal{B} = \mathrm{OBS} \cup \mathrm{SD}_{gen} \quad (5)$$

$$P = \mathrm{MEAS} \quad (6)$$

$$N = \emptyset \quad (7)$$

$$R = \{consistency\} \quad (8)$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT I WIEN GRAZ

Reducing MBD to KBD

Proof (cont'd):

That is, $\mathcal{K}$ captures $\mathrm{SD}_{beh} \cup \{\neg\mathrm{AB}(c_i) \mid c_i \in \mathrm{COMPS}\}$, i.e. the nominal behavioral descriptions of all system components. By Def. 7, $\mathcal{D} \subseteq \mathcal{K}$ is a KBD-diagnosis for *kDPI* iff both

$$(\mathcal{K} \setminus \mathcal{D}) \cup \mathcal{B} \cup U_P \text{ meets all } r \in R \text{ (i.e. is consistent)} \tag{9}$$

and

$$(\mathcal{K} \setminus \mathcal{D}) \cup \mathcal{B} \cup U_P \not\models n \text{ for all } n \in N \tag{10}$$

hold.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Reducing MBD to KBD

Proof (cont'd):

Let now $\mathcal{D}$ be an arbitrary KBD-diagnosis for *kDPI* such that $\mathcal{D} = \{\alpha_i \mid i \in I\}$ for the index set $I \subseteq \{1, \dots, n\}$.
Using (4) – (8) above, condition (9) for $\mathcal{D}$ is equivalent to the consistency of

$$\mathsf{SD}_{beh} \cup \{\mathsf{AB}(c_i) \mid i \in I\} \cup \{\neg\mathsf{AB}(c_i) \mid i \in \{1, \dots, n\} \setminus I\} \cup \mathsf{OBS} \cup \mathsf{SD}_{gen} \cup U_{\mathsf{MEAS}}$$

which in turn yields that

$$\begin{aligned} \mathsf{SD} &\cup \{\mathsf{AB}(c_i) \mid c_i \in \Delta\} \\ &\cup \{\neg\mathsf{AB}(c_i) \mid c_i \in \mathsf{COMPS} \setminus \Delta\} \\ &\cup \mathsf{OBS} \cup U_{\mathsf{MEAS}} \text{ is consistent} \end{aligned} \tag{11}$$

for $\Delta := \{c_i \mid i \in I\}$.

Reducing MBD to KBD

Proof (cont'd).

But, (11) is exactly the condition defining an MBD-diagnosis (see Def. 4). Note, since $N = \emptyset$ by (7), condition (10) is met for any $\mathcal{D}$ satisfying (9) and can thus be neglected. Hence, $\mathcal{D} = \{\alpha_i \mid i \in I\} \subseteq \mathcal{K}$ is a KBD-diagnosis w.r.t. *kDPI* iff $\Delta = \{c_i \mid i \in I\} \subseteq$ COMPS is an MBD-diagnosis for *mDPI*. □

Reducing MBD to KBD

Also, there is a bijective correspondence between KBD-conflicts and MBD-conflicts:

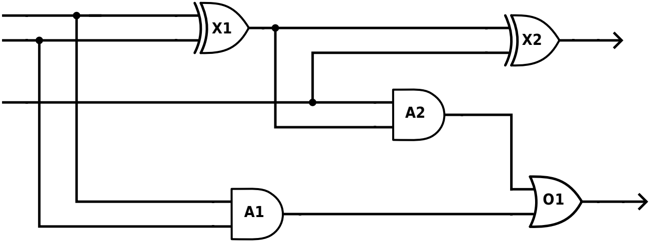
Theorem 2

Let $mDPI = (\textsc{sd}, \textsc{comps}, \textsc{obs}, \textsc{meas})$ be an MBD-DPI and $kDPI = \langle \mathcal{K}, \mathcal{B}, P, N \rangle_R$ a KBD-DPI modeling $mDPI$ as per (4) – (8). Further, let $\textsc{comps} = \{c_1, \dots, c_n\}$ and $I \subseteq \{1, \dots, n\}$. Then, $C = \{c_i \mid i \in I\} \subseteq \textsc{comps}$ is an MBD-conflict for $mDPI$ iff $\mathcal{C} = \{\alpha_i \mid i \in I\} \subseteq \mathcal{K}$ is a KBD-conflict w.r.t. $kDPI$.

Proof.

$\mathcal{C}$ is a KBD-conflict w.r.t. $kDPI$ iff $\mathcal{K} \setminus \mathcal{C} = \{\alpha_i \mid i \in \{1, \dots, n\} \setminus I\}$ is not a KBD-diagnosis w.r.t. $kDPI$ (Property 2) iff $\{c_i \mid i \in \{1, \dots, n\} \setminus I\}$ is not an MBD-diagnosis for $mDPI$ (Theorem 1) iff $\{c_i \mid i \in I\} = C$ is an MBD-conflict for $mDPI$ ([Reiter, 1987, Prop. 4.2]). $\square$

Full-Adder MBD Example (Genesereth, 1984)



circuit inputs (from top to bottom)
1
0
1
circuit outputs (from top to bottom)
1
0

Example (cont'd)

i	α_i	SD_{beh}	SD_{gen}	OBS
1	$\neg\text{AB}(X_1) \to beh(X_1)$	•		
2	$\neg\text{AB}(X_2) \to beh(X_2)$	•		
3	$\neg\text{AB}(A_1) \to beh(A_1)$	•		
4	$\neg\text{AB}(A_2) \to beh(A_2)$	•		
5	$\neg\text{AB}(O_1) \to beh(O_1)$	•		
6	$out(X_1) = in2(A_2)$		•	
7	$out(X_1) = in1(X_2)$		•	
8	$out(A_2) = in1(O_1)$		•	
9	$in1(A_2) = in2(X_2)$		•	
10	$in1(X_1) = in1(A_1)$		•	
11	$in2(X_1) = in2(A_1)$		•	
12	$out(A_1) = in2(O_1)$		•	
13	$in1(X_1) = 1$			•
14	$in2(X_1) = 0$			•
15	$in1(A_2) = 1$			•
16	$out(X_2) = 1$			•
17	$out(O_1) = 0$			•
	COMPS			
	$\{X_1, X_2, A_1, A_2, O_1\}$			
c	$beh(c)$ for $c \in$ COMPS			
X_1	$out(X_1) = xor(in1(X_1), in2(X_1))$			
X_2	$out(X_2) = xor(in1(X_2), in2(X_2))$			
A_1	$out(A_1) = xor(in1(A_1), in2(A_1))$			
A_2	$out(A_2) = xor(in1(A_2), in2(A_2))$			
O_1	$out(O_1) = xor(in1(O_1), in2(O_1))$			
i	MEAS			
×	×			

MBD-DPI

$\Rightarrow$

i	α_i	$\mathcal{K}$	$\mathcal{B}$
1	$out(X_1) = xor(in1(X_1), in2(X_1))$	•	
2	$out(X_2) = xor(in1(X_2), in2(X_2))$	•	
3	$out(A_1) = and(in1(A_1), in2(A_1))$	•	
4	$out(A_2) = and(in1(A_2), in2(A_2))$	•	
5	$out(O_1) = or(in1(O_1), in2(O_1))$	•	
6	$out(X_1) = in2(A_2)$		•
7	$out(X_1) = in1(X_2)$		•
8	$out(A_2) = in1(O_1)$		•
9	$in1(A_2) = in2(X_2)$		•
10	$in1(X_1) = in1(A_1)$		•
11	$in2(X_1) = in2(A_1)$		•
12	$out(A_1) = in2(O_1)$		•
13	$in1(X_1) = 1$		•
14	$in2(X_1) = 0$		•
15	$in1(A_2) = 1$		•
16	$out(X_2) = 1$		•
17	$out(O_1) = 0$		•
i	$p_i \in P$		
×	×		
i	$n_i \in N$		
×	×		
i	$r_i \in R$		
1	consistency		

KBD-DPI

Example (cont'd)

$\Downarrow$

min MBD-conflicts
$\{X_1, X_2\}, \{X_1, A_2, O_1\}$
min MBD-diagnoses
$\{X_1\}, \{X_2, A_2\}, \{X_2, O_1\}$

Minimal MBD-conflicts and MBD-diagnoses

$\Leftarrow$

min KBD-conflicts
$\{\alpha_1, \alpha_2\}, \{\alpha_1, \alpha_4, \alpha_5\}$
min KBD-diagnoses
$\{\alpha_1\}, \{\alpha_2, \alpha_4\}, \{\alpha_2, \alpha_5\}$

Minimal KBD-conflicts and KBD-diagnoses

Lessons learned...

- Can find all MBD-diagnoses and MBD-conflicts for any MBD problem by representing it as a KBD problem and solving the latter for KBD-diagnoses and KBD-conflicts, respectively
- Sequential MBD problem is a special case of Sequential KBD problem: Former can be solved for a given MBD-DPI *MP* by reducing *MP* to a KBD-DPI *KP* and solving latter for *KP* under the restriction that $N_{new} = \emptyset$
- Methods targeting KBD-problem are more general than those addressing MBD-problem as they allow the specification of negative information N_{new} in addition to positive one (P_{new} and $\textsc{meas}_{new}$, respectively)
- Existing KBD methods such as [Felfernig et al., 2004], [Shchekotykhin et al., 2012] and [Rodler, 2015] are suitable to be used for solving arbitrary MBD problems as per [Reiter, 1987] and [de Kleer & Williams, 1987]

References

[de Kleer & Williams, 1987] DE KLEER, J., AND WILLIAMS, B. C.
Diagnosing multiple faults.
Artificial Intelligence 32, 1 (1987), 97–130.

[Felfernig et al., 2004] FELFERNIG, A., FRIEDRICH, G., JANNACH, D., AND STUMPTNER, M.
Consistency-based diagnosis of configuration knowledge bases.
Artificial Intelligence 152, 2 (2004), 213 – 234.

[Reiter, 1987] REITER, R.
A Theory of Diagnosis from First Principles.
Artificial Intelligence 32, 1 (1987), 57–95.

[Rodler, 2015] RODLER, P.
Interactive Debugging of Knowledge Bases.
PhD thesis, Alpen-Adria Universität Klagenfurt, 2015. *goo.gl/NTszUY*.

[Shchekotykhin et al., 2012] SHCHEKOTYKHIN, K., FRIEDRICH, G., FLEISS, P., AND RODLER, P.
Interactive Ontology Debugging: Two Query Strategies for Efficient Fault Localization.
Web Semantics: Science, Services and Agents on the World Wide Web 12-13 (2012), 88–103.