Title | A general characterization of Model-Based Diagnosis  
Author(s) | Provan, Gregory  
Publication date | 2016-01  
Type of publication | Conference item  
| [http://dx.doi.org/10.3233/978-1-61499-672-9-1565](http://dx.doi.org/10.3233/978-1-61499-672-9-1565)  
Access to the full text of the published version may require a subscription.  
Rights | © 2016, the Authors and IOS Press. This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0)  
| [https://creativecommons.org/licenses/by-nc/4.0/](https://creativecommons.org/licenses/by-nc/4.0/)  
Item downloaded from | [http://hdl.handle.net/10468/3770](http://hdl.handle.net/10468/3770)  
Downloaded on 2020-08-31T17:15:10Z
A General Characterization of Model-Based Diagnosis

Gregory Provan

Abstract. The Model-Based Diagnosis (MBD) framework developed by Reiter has been a strong theoretical foundation for MBD, yet is limited to models that are described in terms of logical sentences. We propose a more general framework that covers a wide range of modelling languages, ranging from AI-based languages (e.g., logic and Bayesian networks) to FDI-based languages (e.g., linear Gaussian models). We show that a graph-theoretic basis for decomposable system models can be augmented with several languages and corresponding inference algorithms based on valuation algebras.

1 A GENERAL MBD FRAMEWORK

We propose a framework for extending the Reiter MBD approach [7] by integrating several MBD and FDI approaches within a decomposable graphical framework in which the modeling language and inference are specified by a valuation algebra [6].

More formally, we define an MBD framework using a tuple \((G, T, \Gamma)\), where \(G\) is a factor graph [5], \(T\) is the diagnosis task, and \(\Gamma\) is a valuation algebra [6]. The factor graph \(G\) specifies a system topology in terms of a decomposable relation \(\Psi\), defined over a set \(V\) of variables, such that \(\Psi = \bigotimes_i \psi_i(V_i)\), where \(\psi_i\) is a sub-relation, \(\otimes\) is the composition operator and \(V_i \subseteq V\). This decomposition can be mapped to a graph, e.g., a DAG for a Bayesian network (BN) or an undirected graph for a Markov network. The diagnosis task is given by the tuple \(T = (D, y, R)\), where \(D\) is the task specification; \(y\) is the required system measurement; and residual \(R(\Psi, y)\) indicates a discrepancy between observation \(y\) and model-derived prediction \(y\) using some distance measure. The valuation algebra \(\Gamma\) specifies (1) the underlying language for the diagnosis system, and (2) the inference necessary to compute the diagnosis for the task \(T\), e.g., multiple-fault subset-minimal diagnosis or Most-Probable Explanation.

This decomposable representation can encode a wide range of diagnosis models, including propositional logic models, FDI dynamical systems models, as well as stochastic models (Bayesian networks, HMMs, and linear Gaussian models). Previous work, e.g., [4] has shown AI-based approaches to diagnosis [7, 2] can be described by valuation algebras. Here, we extend this to include FDI approaches based on ordinary differential equations (ODEs), and show the importance of system structure and diagnosis task in specifying the full diagnosis representation.

This framework has several outcomes. First, it enables a clear separation between models and inference (although the two are linked). Specifically, the model structure encoded as a factor graph that governs inference complexity. For example, tree-structured factor graphs are all poly-time computable. Second, the factor graph encoding of models clarifies the structural difference between AI-based approaches and FDI-based approaches.

2 REPRESENTING MULTIPLE MODEL TYPES

We can represent a system (or plant) model \(\Psi\) using a factor graph, which represents the physical connectivity of \(\Psi\) in terms of a structured decomposition of \(\Psi\) into sub-relations. Consider Figure 1(a), which shows an example of a tank with a valve, where we control the level \(x\) in the tank by controlling the inflow \(f_1\) and the valve state \(V\). There are two possible failures in the system: (1) the tank can leak, with failure mode \(\phi_T\), and (2) the valve can malfunction, with failure mode \(\phi_V\). This example has variables \(\{f_1, f_2, f_3, x, \phi_T, \phi_V\}\) and three relations: \(\psi_1(f_1, \phi_T, x), \psi_2(x, f_2), \) and \(\psi_3(f_2, f_3, \phi_V)\). \(\psi_1\) represents how the tank’s fluid level \(x\) is governed by inflow \(f_1\) and fault (leak) \(\phi_T\). \(\psi_2\) represents how outflow is governed by fluid height \(x\), and \(\psi_3\) represents the valve’s impact on flows \(f_2, f_3\).

Given such a decomposition, we can represent the modelling language as a valuation over \(\psi_1 \otimes \psi_2 \otimes \psi_3\). For example, if we choose a probabilistic algebra then we obtain a diagnostic BN model, for which Figure 1(b) shows the structure and valuation \(P(V) = P(f_1)P(\phi_T)P(\phi_V)P(f_1|\phi_T, x)P(x|f_2)P(f_1|f_2, \phi_V)\). \(P(x|f_1, \phi_T)\) defines the conditional dependence of tank level \(x\) on the inflow \(f_1\) and the tank fault-state \(\phi_T\), and \(P(f_3|f_2, \phi_V)\) the conditional dependence of flow from this system, \(f_3\), on the tank outflow \(f_2\) and the valve fault-state \(\phi_V\).
We perform inference in $\Psi$ by message-passing and valuation updating [6]. Figure 1(c) shows how we can compute a diagnosis (i.e., evaluate $\phi_T$, $\phi_V$) by passing messages among the nodes starting with the control and observation settings $S$. If all assignments in $S$ are nominal (nom), the “diagnosis” is $P(\phi_T = fault) = .004$ and $P(\phi_V = fault) = .004$. If scenario $S$ has control assignment $f_1$=nom, and observation of $f_2$, $f_3$ both as low, we obtain a diagnosis given by $P(\phi_T = fault) = .067$ and $P(\phi_V = fault) = .009$, i.e., a faulty tank is the most likely diagnosis.

We can represent several different tank models by keeping the same decomposable structure and changing the valuation. For example, we obtain a qualitative model by replacing (1) the conditional probability tables with qualitative relations, and (2) passing qualitative messages (e.g., $\{+,-,0\}$ rather than discrete-valued probabilities), and using qualitative inference rather than Bayesian updating.

Table 1 summarizes the properties of several models characterized by our framework, defining the language, model structure, the underlying semi-ring, and the inference complexity. The language and task can be characterized by the valuation semi-ring $(Z, O_1, O_2)$, which consists of a set $Z$ and two operations $(O_1, O_2)$ [3]. The last column of Table 1 shows the inference complexity for which the primary determinant is the topology [1]: any non-tree topology is likely to be NP-hard for computing a task requiring at least one multiple-fault diagnosis, whereas tree topologies are poly-time solvable for the majority of languages and tasks.

A valuation is a measure over the possible values of a set $V$ of variables [3]. Each valuation $\psi$ refers to a finite set of variables $d(\psi) \subseteq V$, called its domain. Given the power set $\mathcal{P}$ of $V$ and a set $\psi$ of valuations with their domains in $\mathcal{P}$, we can define 3 key operations: (1) Labeling: $\psi \mapsto d(\psi)$, which returns the domain of each valuation; (2) Combination: $(\psi_1, \psi_2) \mapsto \psi_1 \otimes \psi_2$, which specifies functional composition, i.e., the aggregation of data from multiple functions; (3) Projection: $(\psi, V) \mapsto \psi^{V'}$ for $V' \subseteq d(\psi)$, which specifies the computation of a query (set of variables) of interest.

Given an observation $y$, we specify diagnosis within a valuation algebra as a two-step process: (1) residual analysis (RA); and (2) fault isolation (FI).

### Residual analysis
This inference depends on the type of residual. AI logic-based approaches compute RA using a consistency check, denoted $\Psi^{10}$. FDI continuous-valued systems compute RA as $R = |\hat{y} - y|$, where $\hat{y}$ is the model’s prediction. Residual-specific FG structure may be necessary to enable us to compute $\Psi^{1\Gamma}$.

### Fault isolation
Isolating a diagnosis is equivalent to projecting the marginal over the fault-mode variables $\phi$, denoted $\Psi^{1\phi} = (\psi_1 \otimes \cdots \otimes \psi_n)^{1\phi}$. Diagnostic inference requires all 3 valuation operations, in particular combination and projection. The task also may change the FG structure and operations required. For example, different operations are required for computing a posterior distribution $P(\phi|y)$ as opposed to the Most Probable Explanation (MPE).

Given an observation $y$ and prediction $\hat{y}$, the typical objective of a diagnosis process is to identify the system fault-state that minimises the residual vector: $\psi^* = \arg\min_{\psi \in \Psi} R(\Psi, y)$. The full paper generalizes the inference metric to define our diagnosis task as jointly minimizing the accuracy (based on $R$) and the inference complexity.

### 3 CONCLUSION
This article has presented a general framework for MBD that integrates several approaches developed within different communities, most notably the AI and FDI communities. By characterizing MBD using the triple $(\mathcal{G}, T, \Gamma)$, we show structural similarities in MBD techniques using the underlying graph $\mathcal{G}$. The valuation algebra $\Gamma$ enables us to demonstrate the operations and message-passing techniques underlying the MBD approaches. As a consequence, we are able to identify similarities among MBD approaches, thereby paving the way for a more holistic approach to MBD and potential cross-pollination of MBD inference techniques.

### REFERENCES

<table>
<thead>
<tr>
<th>Approach</th>
<th>Language</th>
<th>Structure</th>
<th>Task</th>
<th>Semi-Ring</th>
<th>RA</th>
<th>FI</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atemporal</td>
<td>Reiter</td>
<td>Prop. Logic</td>
<td>DAG</td>
<td>Dx</td>
<td>${0,1,\land,\lor}$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
</tr>
<tr>
<td></td>
<td>ATMS</td>
<td>M-V Logic</td>
<td>DAG</td>
<td>all Dx $\langle 2^n,\land,\lor\rangle$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
<td>NP-hard</td>
</tr>
<tr>
<td>Qualitative</td>
<td>Constraint</td>
<td>Constraint</td>
<td>arbitrary</td>
<td>all Dx $\langle 2^n,\land,\lor\rangle$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
<td>NP-hard</td>
</tr>
<tr>
<td>Constraint Network</td>
<td>BN-Posterior</td>
<td>PGM</td>
<td>DAG</td>
<td>$P(\phi</td>
<td>y)$</td>
<td>${0,1,\land,\lor}$</td>
<td>$\Psi^{10}$</td>
</tr>
<tr>
<td></td>
<td>BN-MPE</td>
<td>PGM</td>
<td>DAG</td>
<td>MPE</td>
<td>${0,1,\max,\times}$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
</tr>
<tr>
<td>Temporal</td>
<td>DES</td>
<td>M-V Logic</td>
<td>arbitrary</td>
<td>all Dx $\langle 2^n,\land,\lor\rangle$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
<td>NP-hard</td>
</tr>
<tr>
<td></td>
<td>DBN</td>
<td>PGM</td>
<td>DAG</td>
<td>$P(\phi)$</td>
<td>${0,1,\land,\lor}$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
</tr>
<tr>
<td></td>
<td>HMM</td>
<td>PGM</td>
<td>tree</td>
<td>$P(\phi)$</td>
<td>${0,1,\land,\lor}$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
</tr>
<tr>
<td></td>
<td>FDI</td>
<td>ODE</td>
<td>bipartite</td>
<td>Dx</td>
<td>${0,1,\land,\lor}$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
</tr>
<tr>
<td></td>
<td>Kalman filter</td>
<td>PGM</td>
<td>tree</td>
<td>MPE</td>
<td>${0,1,\land,\lor}$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
</tr>
<tr>
<td></td>
<td>Particle filter</td>
<td>PGM</td>
<td>arbitrary</td>
<td>MPE</td>
<td>${0,1,\land,\lor}$</td>
<td>$\Psi^{10}$</td>
<td>$\Psi^{10}$</td>
</tr>
</tbody>
</table>

Table 1. Classification of Diagnosis Approaches using Generalized Approach. RA and FI denote Residual Analysis and Fault Isolation, respectively. Shaded rows denote stochastic methods, and unshaded rows denote deterministic methods. BN denotes Bayesian-network. For the language, Prop and M-V denote propositional and multivalued respectively; Q-Relation denotes Qualitative Relation; PGM refers to probabilistic graphical model, and ODE to Ordinary Differential Equation. For the task, Dx corresponds to a single diagnosis, $P(\phi)$ to the probability distribution over $\phi$, and MPE to Most-Probable Explanation.