Model-Based Diagnosis

Alban Grastien
Summary

1. What is Diagnosis

2. Model-Based Diagnosis Techniques
   - Pattern-Based Diagnosis
   - Explanation-Based Diagnosis
   - Hypothesis-Based Diagnosis

3. Diagnosability
Diagnosis

Observation: the car does not start

What is wrong?
- Flat battery
- Broken starter
- Out-of-petrol tank
Applications

Examples

- Transport (car, truck, plane, helicopter, rocket, submarine, extraplanetary rover)
- Infrastructure (communication, electricity, water networks)
- Complex systems (assembly line, circuit, robot, HVAC)
- Software (testing/debbuging, web-services)
Model-Based Diagnosis

Principle

- **Model**: description of a system
  - Knowledge about “how the world works”
  - [Russel and Norvig, 2003]

- **Diagnoser**: generic algorithm
Model-Based Diagnosis

Principle

- Model: description of a system
  Knowledge about “how the world works”
  [Russel and Norvig, 2003]
- Diagnoser: generic algorithm

Advantages

- If the system is (slightly) changed, only the model needs to be (slightly) changed
- It is easier to distinguish the model from the solver
- Complexity results apply
- Additional analyses can be performed (e.g. diagnosability)
Examples

- Propositional logic
- First-order logic
- Causal (Bayesian) networks
- Continuous systems
- Discrete event systems (DES)
- Hybrid systems
Summary

1. What is Diagnosis

2. Model-Based Diagnosis Techniques
   - Pattern-Based Diagnosis
   - Explanation-Based Diagnosis
   - Hypothesis-Based Diagnosis

3. Diagnosability
Principle

- Associate certain “patterns” on the observations with diagnostic information
Pattern Example: Residual

Principle

Off-line:

- Input: a set of equations on observable and unobservable variables modeling the nominal behaviour of the components in the system
- Generate equations defined purely on the observable variables such that the result must equal 0
- Identify the components involved in each equation

On-line:

- Test that the equations indeed evaluate to 0
- For each equation that does not evaluate to 0, one of the component involved in the equation must be faulty
Example

Model:
- Pump: \( x = \rho \times K_x \) (1)
- Pipe: \( y + z = x \) (2)
- Tank 1: \( y = \dot{l}_1 \times K_y \) (3)
- Tank 2: \( z = \dot{l}_2 \times K_z \) (4)

Constants: \( K_x, K_y, K_z \)

Observable variables:
- \( \rho \)
- \( l_1 \) (hence \( \dot{l}_1 \))
- \( l_2 \) (hence \( \dot{l}_2 \))

\[
(\rho \times K_x) - (\dot{l}_1 \times K_y) - (\dot{l}_2 \times K_z) = 0
\]
Pattern Example: Chronicle

Principle

- A chronicle is a collection of alarms/logs with time constraints
- A chronicle is associated with some diagnostic information

Example

\[ \text{CB}_1 \text{ close} \leftrightarrow \text{CB}_1 \text{ open} \quad \text{[2,3]} \quad \text{CB}_2 \text{ open} \quad \text{[2,3]} \quad \text{CB}_2 \text{ close} \]

\[ \text{CB}_3 \text{ close} \leftrightarrow \text{CB}_3 \text{ open} \quad \text{[2,3]} \quad \text{[2,3]} \quad \text{CB}_4 \text{ open} \quad \text{[2,3]} \quad \text{CB}_4 \text{ close} \]

→ transient fault isolated by circuit breakers (equipped with automatic reclosers)
Pattern Example: Diagnoser

Principle

Project the model on the observations so that the diagnoser is a function that associates the observations to the diagnosis
Pattern: Advantages / Issues

Advantages
- Intuitive
- Easy to implement / test
- Generally quick

Issues
- Completeness
- Size / number of patterns for complex systems / faults
Summary

1. What is Diagnosis
2. Model-Based Diagnosis Techniques
   - Pattern-Based Diagnosis
   - Explanation-Based Diagnosis
   - Hypothesis-Based Diagnosis
3. Diagnosability
Explanations

Principle

- Compute all the system behaviours that are compatible with the observation
- Extract the relevant information from these behaviours
Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Circuit

- The model is a formula $\Phi_M$:
  \[
  (\neg Ab(Mul1) \rightarrow \text{output}(Mul1) = A \ast C) \land \ldots
  \]
- The observation is a formula $\Phi_O$: $(A = 2) \land (B = 2) \land \ldots$
- The explanations is the formula $\Phi_E = \Phi_M \land \Phi_O$
- The diagnosis is the projection of the explanations on the $Ab$ variables
  \[
  \Phi_\Delta = \text{Proj}_{Ab}(\Phi_E) = Ab(Mul1) \lor Ab(Add1) \lor
  (Ab(Mul2) \land Ab(Mul3)) \lor (Ab(Mul2) \land Ab(Add2))
  \]
The Compilation Map [Darwiche et al.]

- The formula $\Phi_{\Delta}$ should make it easy to answer questions
- The formulas must remain of reasonable size

$\Rightarrow$ theory of BDDs, DNNFs, DNF/CNF, SAT, etc.
Explanations Example: DES

Observations: $c, b, c, c$
Explanations Example: DES

Observations: $c, b, c, c$
**Issues**

- Complexity: the size of the model automaton is exponential in the number of components

  ⇒ Petri nets, decentralised/distributed diagnosis, abstraction, etc.
Advantages

- Supports the solution (explanations)
- Flexible wrt the diagnostic questions

Issue

- Complexity
Summary

1. What is Diagnosis

2. Model-Based Diagnosis Techniques
   - Pattern-Based Diagnosis
   - Explanation-Based Diagnosis
   - Hypothesis-Based Diagnosis

3. Diagnosability
Hypothesis Testing

Principle

- A hypothesis is a possible diagnosis: eg.
  - No fault
  - Components $Add_1$ and $Mul_2$ are faulty
  - A drift started on sensor $s_2$ before line BSW23 cut

- The compatibility (logical consistency) between certain hypotheses, the model, and the observations is tested.
Hypothesis Testing Example: Circuit

Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration
- Test $\emptyset \rightarrow$ conflict $Ab(\text{Add1}) \lor Ab(\text{Mul1}) \lor Ab(\text{Mul2})$
Hypothesis Testing Example: Circuit

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul2)$
- Test $\{Ab(Add1)\} \rightarrow$ possible diagnosis

Observations:
$A = B = E = 2$ ; $C = D = 3$ ; $F = 10$ ; $G = 12$
Hypothesis Testing Example: Circuit

Observations:
A = B = E = 2 ; C = D = 3 ;  
F = 10 ; G = 12

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul2)$
- Test $\{Ab(Add1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\} \rightarrow$ possible diagnosis
Hypothesis Testing Example: Circuit

Observations:
A = B = E = 2 ; C = D = 3 ; F = 10 ; G = 12

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul2)$
- Test $\{Ab(Add1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul2)\} \rightarrow$ conflict
  $Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul3) \lor Ab(Add2)$
Hypothesis Testing Example: Circuit

Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul2)$
- Test $\{Ab(Add1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul2)\} \rightarrow$ conflict
  \[Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul3) \lor Ab(Add2)\]
- Test $\{Ab(Mul2), Ab(Mul3)\} \rightarrow$ possible diagnosis
Hypothesis Testing Example: Circuit

Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration
- Test $\emptyset$ $\rightarrow$ conflict $Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul2)$
- Test $\{Ab(Add1)\}$ $\rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\}$ $\rightarrow$ possible diagnosis
- Test $\{Ab(Mul2)\}$ $\rightarrow$ conflict
  $Ab(Add1) \lor Ab(Mul1) \lor Ab(Mul3) \lor Ab(Add2)$
- Test $\{Ab(Mul2), Ab(Mul3)\}$ $\rightarrow$ possible diagnosis
- Test $\{Ab(Mul2), Ab(Add2)\}$ $\rightarrow$ possible diagnosis
Hypothesis Testing Example: Circuit (cont.)

Illustration

- Test \{Ab(Mul1), Ab(Mul2), Ab(Mul3), Ab(Add1), Ab(Add2)\} → possible diagnosis
- Test \{Ab(Mul1), Ab(Mul2), Ab(Mul3), Ab(Add1)\} → possible diagnosis
- Test \{Ab(Mul1), Ab(Mul2), Ab(Mul3)\} → possible diagnosis
- Test \{Ab(Mul1), Ab(Mul2)\} → possible diagnosis
- Test \{Ab(Mul1)\} → possible diagnosis
- Test ∅ → not a diagnosis

Observations:
A = B = E = 2 ; C = D = 3 ; F = 10 ; G = 12
Summary

1. What is Diagnosis

2. Model-Based Diagnosis Techniques
   - Pattern-Based Diagnosis
   - Explanation-Based Diagnosis
   - Hypothesis-Based Diagnosis

3. Diagnosability
Diagnosability

Principle

- If a fault occurs, will it be detected and identified?
- Which sensors can I drop, and still maintain (a decent level of) diagnosability?
Rephrasing the Problem

The model is not diagnosable if there exists a nominal behaviour $b$ and a faulty behaviour $b'$ that produce the same observations.

Twin Plant

- Make a copy $M'$ of the model $M$ and force a fault on the copy.
- Synchronise $M$ and $M'$ on the observations: $M \otimes M'$.
- Check $M \otimes M'$ for emptiness.

If $(b, b') \in M \otimes M'$, then the system is not diagnosable.
Testing Diagnosability Example: DES

\[ \begin{align*}
1 & \rightarrow 2 \\
3 & \rightarrow 4 \\
4 & \rightarrow 5 \\
5 & \rightarrow 6 \\
6 & \rightarrow 7 \\
7 & \rightarrow 8 \\
8 & \rightarrow 5 \\
5 & \rightarrow 6 \\
6 & \rightarrow 7 \\
7 & \rightarrow 8
\end{align*} \]
Testing Diagnosability Example: DES
Other Issues

- Modelling
- Dealing with uncertain diagnosis
- Incrementally computing the diagnosis
- Dealing with probabilities
- Increasing diagnosability