Model-Based Diagnosis

Alban Grastien

Australian Government
Department of Broadband, Communications and the Digital Economy
Australian Research Council

NICTA Funding and Supporting Members and Partners

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

NICTA

Examples

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

Model-Based Diagnosis

NICTA

Principle

- Model: description of a system
  Knowledge about “how the world works”

  [Russel and Norvig, 2003]
- Diagnoser: generic algorithm

Model-Based Diagnosis

NICTA

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)

Models

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

Patterns

NICTA

Principle

- Associate certain “patterns” on the observations with diagnostic information

Pattern Example: Residual

NICTA

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

Residual

Example

- Model:
  - Pump:
  - Pipe:
  - Tank 1:
  - Tank 2:

$$x = \rho \times K_x \quad (1)$$
$$y + z = x \quad (2)$$
$$y = \dot{l}_1 \times K_y \quad (3)$$
$$z = \dot{l}_2 \times K_z \quad (4)$$

- Constants: K_x, K_y, K_z
- Observable variables:
  - ρ
  - l_1 (hence $\dot{l}_1$)
  - l_2 (hence $\dot{l}_2$)

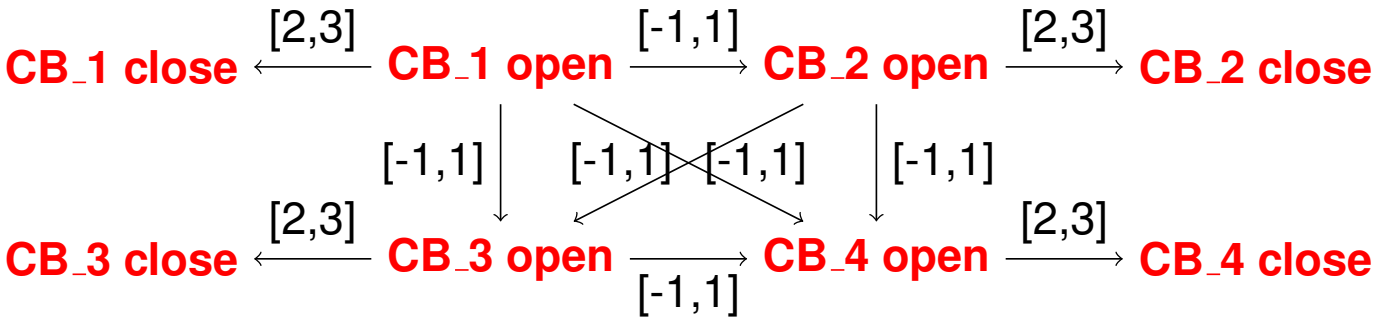
$$(\rho \times K_x) - \left(\dot{l}_1 \times K_y\right) - \left(\dot{l}_2 \times K_z\right) = 0$$

Pattern Example: Chronicle

Principle

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

Example



$\rightarrow$ transient fault isolated by circuit breakers (equipped with automatic reclosers)

Pattern Example: Diagnoser

NICTA

Principle

Project the model on the observations so that the diagnoser is a function that associates the observations to the diagnosis

Pattern: Advantages / Issues

NICTA

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

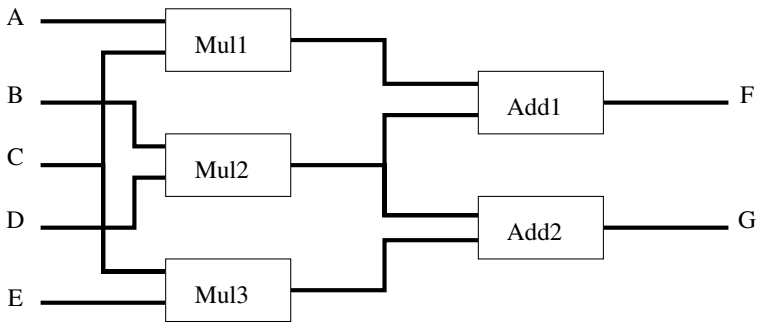
Explanations

NICTA

Principle

- Compute all the system behaviours that are compatible with the observation
- Extract the relevant information from these behaviours

Explanations Example: Circuit



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

Circuit

- The model is a formula Φ_M:

$$(\neg Ab(Mul1) \rightarrow \text{output}(Mul1) = A * C) \wedge \ldots$$

- The observation is a formula Φ_O: $(A = 2) \wedge (B = 2) \wedge \ldots$
- The explanations is the formula $\Phi_E = \Phi_M \wedge \Phi_O$
- The diagnosis is the projection of the explanations on the *Ab* variables

$$\Phi_\Delta = \text{Proj}_{Ab}(\Phi_E) = Ab(Mul1) \vee Ab(Add1) \vee (Ab(Mul2) \wedge Ab(Mul3)) \vee (Ab(Mul2) \wedge Ab(Add2))$$

Explanations Example: Circuit (cont.)

$$
\begin{aligned}
\Phi_\Delta \quad = \quad & Ab(Mul1) \\
\vee \quad & Ab(Add1) \\
\vee \quad & (Ab(Mul2) \wedge Ab(Mul3)) \\
\vee \quad & (Ab(Mul2) \wedge Ab(Add2))
\end{aligned}
$$

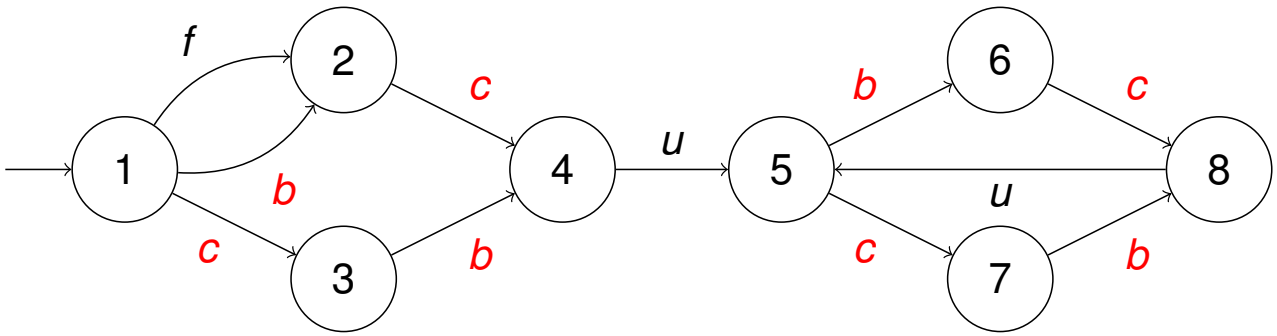
$$
\begin{aligned}
= \quad & (Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul2)) \\
\wedge \quad & (Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul3) \vee Ab(Add2))
\end{aligned}
$$

The Compilation Map [Darwiche et al.]

- The formula Φ_Δ should make it easy to answer questions
- The formulas must remain of reasonable size

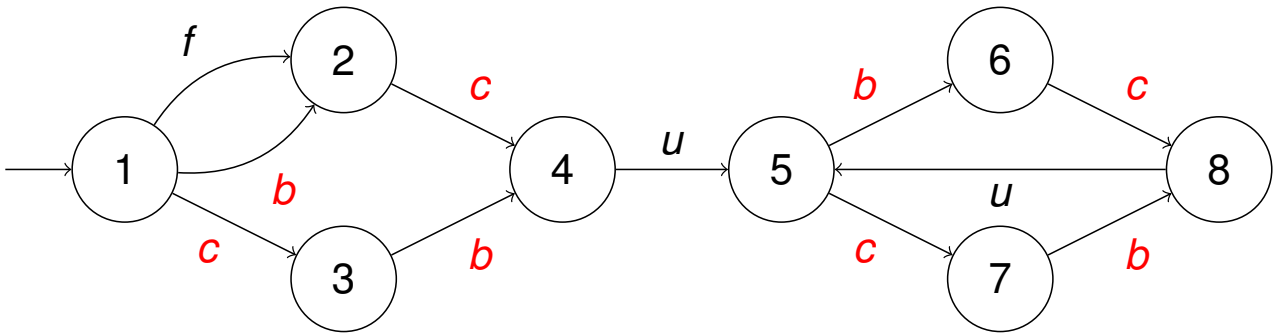
$\Rightarrow$ theory of BDDs, DNNFs, DNF/CNF, SAT, etc.

Explanations Example: DES

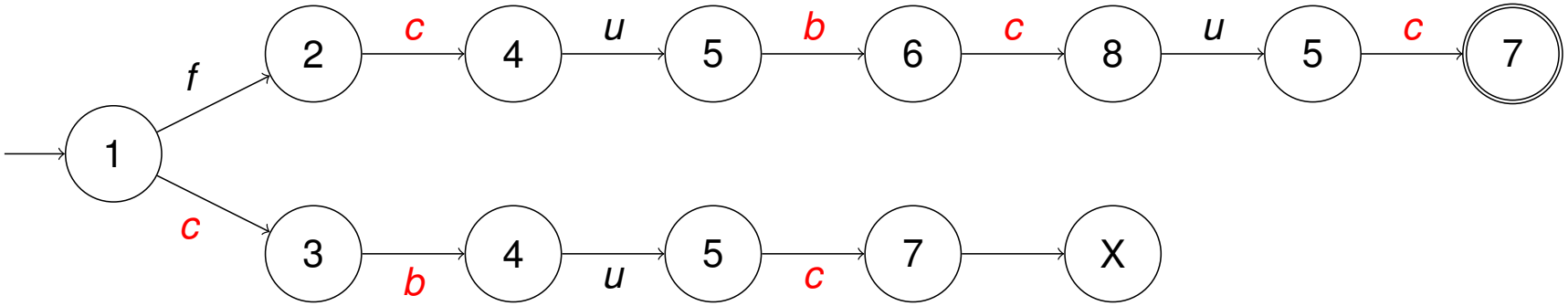


Observations: c, b, c, c

Explanations Example: DES



Observations: c, b, c, c



Explanations Example: DES (cont.)

NICTA

Issues

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

$\Rightarrow$ Petri nets, decentralised/distributed diagnosis, abstraction, etc.

Explanations: Advantages / Issues

NICTA

Advantages

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

Issue

- Complexity

Summary

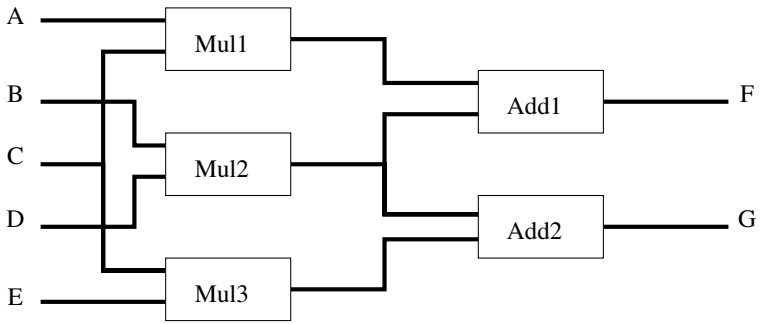
1. What is Diagnosis
2. Model-Based Diagnosis Techniques
   - Pattern-Based Diagnosis
   - Explanation-Based Diagnosis
   - Hypothesis-Based Diagnosis
3. Diagnosability

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

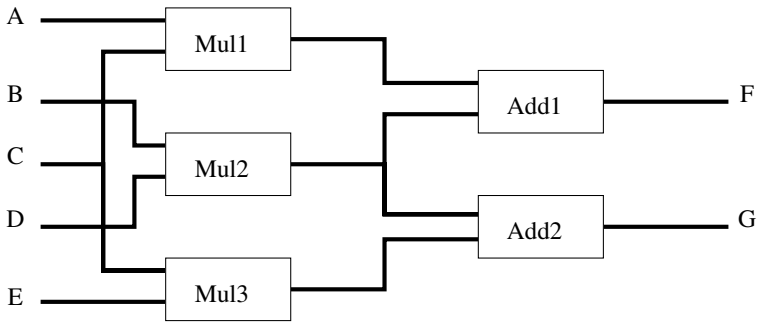


Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul2)$

Hypothesis Testing Example: Circuit

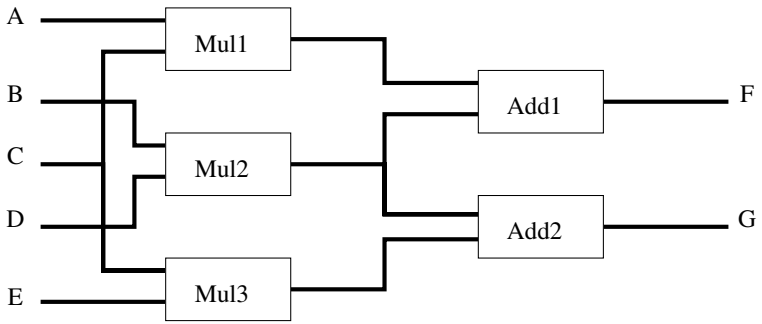


Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration

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

Hypothesis Testing Example: Circuit

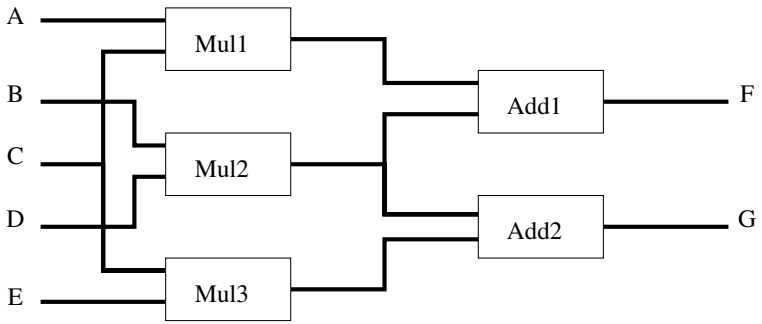


Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration

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

Hypothesis Testing Example: Circuit

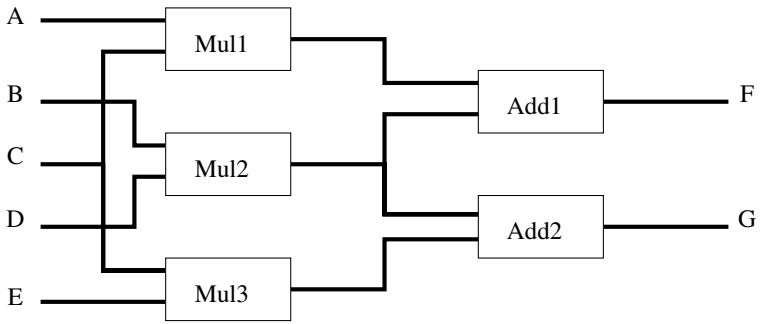


Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul2)$
- Test $\{Ab(Add1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul2)\} \rightarrow$ conflict
  $Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul3) \vee Ab(Add2)$

Hypothesis Testing Example: Circuit

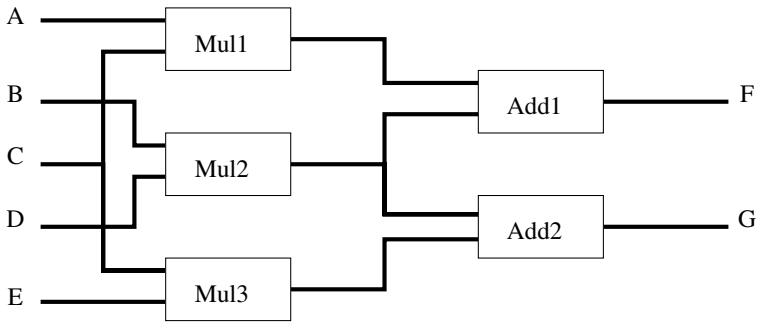


Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul2)$
- Test $\{Ab(Add1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul2)\} \rightarrow$ conflict
  $Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul3) \vee Ab(Add2)$
- Test $\{Ab(Mul2), Ab(Mul3)\} \rightarrow$ possible diagnosis

Hypothesis Testing Example: Circuit

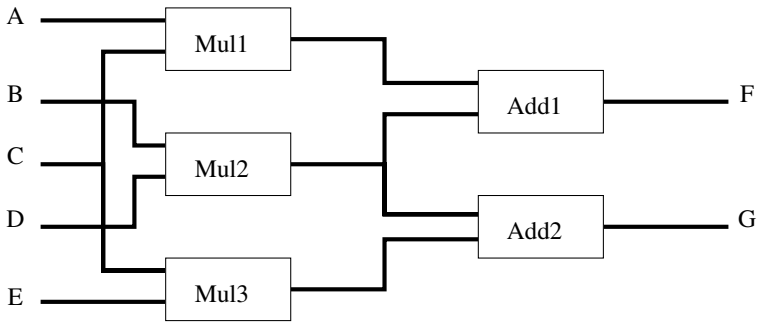


Observations:
A = B = E = 2 ; C = D = 3 ;
F = 10 ; G = 12

Illustration

- Test $\emptyset \rightarrow$ conflict $Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul2)$
- Test $\{Ab(Add1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul2)\} \rightarrow$ conflict
  $Ab(Add1) \vee Ab(Mul1) \vee Ab(Mul3) \vee Ab(Add2)$
- Test $\{Ab(Mul2), Ab(Mul3)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul2), Ab(Add2)\}) \rightarrow$ possible diagnosis

Hypothesis Testing Example: Circuit (cont.)



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

Illustration

- Test $\{Ab(Mul1), Ab(Mul2), Ab(Mul3), Ab(Add1), Ab(Add2)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1), Ab(Mul2), Ab(Mul3), Ab(Add1)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1), Ab(Mul2), Ab(Mul3)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1), Ab(Mul2)\} \rightarrow$ possible diagnosis
- Test $\{Ab(Mul1)\} \rightarrow$ possible diagnosis
- Test $\emptyset \rightarrow$ not a diagnosis

Summary

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

Diagnosability

NICTA

Principle

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

Testing Diagnosability

Rephrasing the Problem

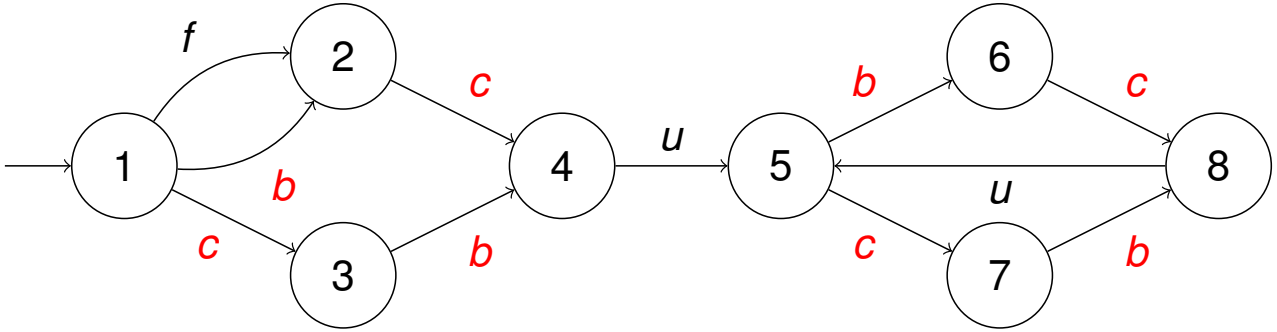
The model is not diagnosable if there exists a nominal behaviour b and a faulty behaviour b' that produce the same observations

Twin Plant

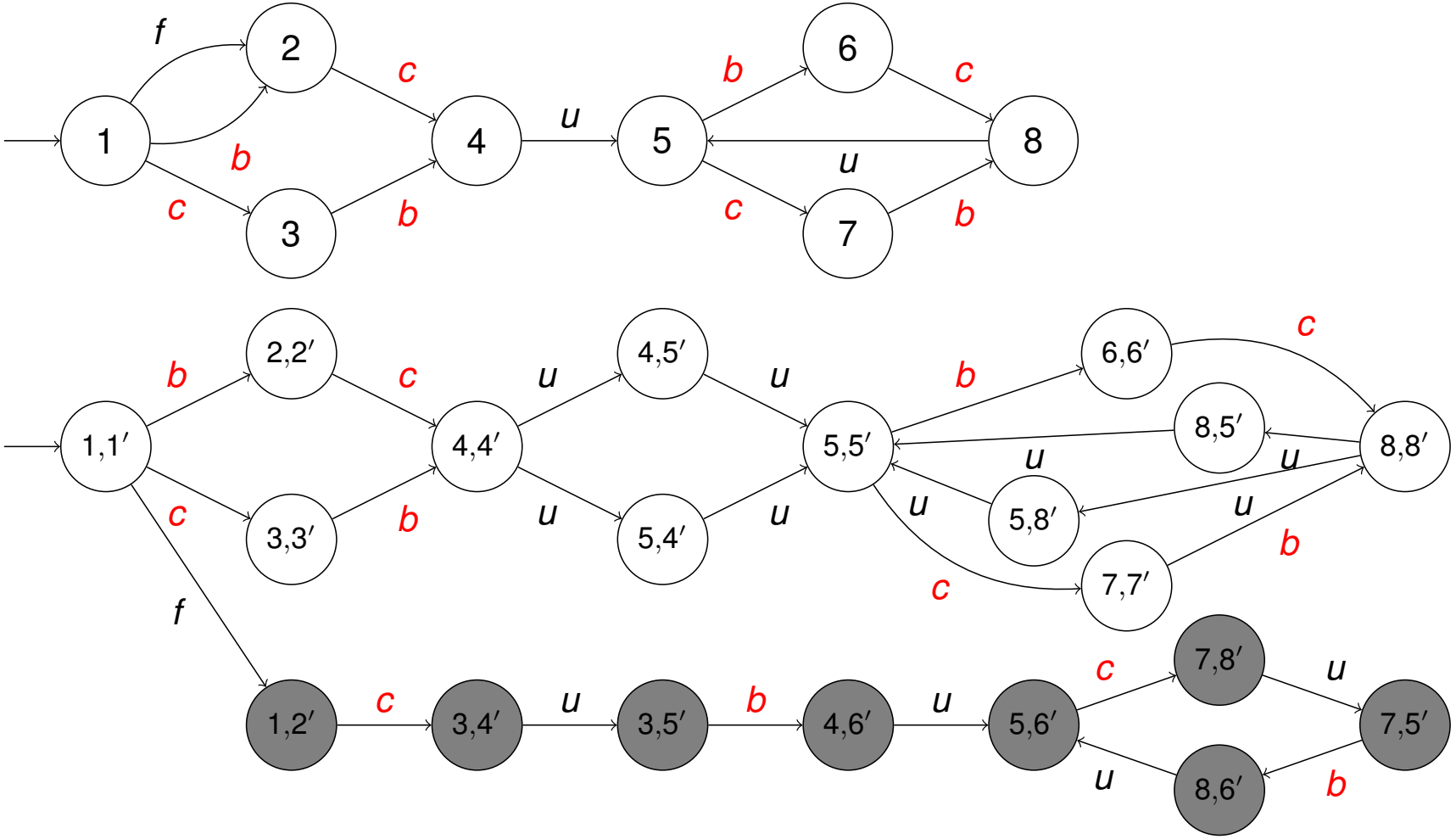
- Make a copy M' of the model M and force a fault on the copy
- Synchronise M and M' on the observations: $M \otimes M'$
- Check $M \otimes M'$ for emptiness

If $(b, b') \in M \otimes M'$, then the system is not diagnosable.

Testing Diagnosability Example: DES



Testing Diagnosability Example: DES



Other Issues

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