A Structural approach to Sensor Placement based on Symbolic Compilation of the Model

Gianluca Torta * **Pietro Torasso** *

* *Dipartimento di Informatica, Università di Torino*
Corso Svizzera 185, 10149 Torino (Italy)
e-mail: {torta,torasso}@di.unito.it.

Abstract: In the present paper we address the problem of computing the Minimal Additional Sensor Sets (MASS) that guarantee a desired level of diagnostic discrimination for a system.
Recently, techniques based on the symbolic compilation of qualitative system models have been proposed for the computation of MASS. The main contribution of this paper is the adaptation and application of such techniques to a structural approach suitable for the computation of MASS for component-oriented models consisting of sets of numerical equations. In this respect, the paper can be viewed as a bridge across the AI and FDI approaches to model-based sensor placement.
In particular, our approach derives the MASS for a given system starting from a Resolution Process Graph which associates a subset of equations to the set of unknown variables and defines the set of redundant relations. We show that the resulting method exploits the symbolic compilation techniques not only as a way to provide computational savings (including some theoretical guarantees on the computational complexity), but it also exhibits interesting features, such as handling of multiple faults.
As a test bed of the proposed approach, we have chosen the problem of the MASS computation for a gas turbine subsystem which has been previously adopted by other researchers to illustrate their proposal.

1. INTRODUCTION

In recent years the Model-Based Diagnosis community has devoted a significant attention to the problem of diagnosability and to the related problem of determining a set of sensors that guarantees the diagnosability of a given system (or at least a *desired* level of diagnosability since in some cases it is very hard to assure diagnosability in all possible conditions of the system).
In most cases one is interested in finding a sensor set that is *minimal* according to some criteria such as set inclusion, cardinality or total cost. The search space for the minimal sensor sets is usually specified by the system modeler as a set of potential measurement points, i.e. physical quantities that could be measured by placing a suitable sensor. Such a search space could be constrained either by positive information (e.g. we already have some sensors in place that "come for free") or negative information (e.g. the sensors placed in certain places are too unreliable or prone to failure or too costly). When such constraints are present, the problem of computing the Minimal Sensor Sets (MSS) is transformed into the problem of computing the Minimal Additional Sensor Sets (MASS).

The problem of computing the MSS (or MASS) has been widely studied in the FDI literature (e.g. Travé-Massuyès et al. [2001], Commault et al. [2006], Travé-Massuyès et al. [2006], Krysander and Frisk [2008]). It is worth mentioning that the problem is computationally hard and usually requires a significant amount of search since it requires that the solution is optimal. For this reason, in many cases the problem has been addressed by putting some additional constraints such as the single fault assumption, so that the requirement of diagnosability is simplified by requiring that, for each component, it is possible to discriminate whether it is ok or faulty under the assumption that at most one fault is present in the system.

Recently, in Torta and Torasso [2007], the computation of MASS has been approached in a quite different way by exploiting symbolic representation and compilation techniques. Such a proposal (and its extension Torta and Torasso [2008]) has been developed by taking a *symbolic AI* approach to the problem, and therefore addresses component based systems where the domain theory of the single components is given in terms of qualitative relations (in particular, this approach has been demonstrated on combinatorial digital circuits and qualitative models of hydraulic systems).
More specifically, in Torta and Torasso [2007], the discriminability relations for a given level of observability are parsimoniously encoded using Ordered Binary Decision Diagrams (OBDDs) and such discriminability relations are built from an extended model of the system to be diagnosed which includes a set of switches modeling the inclusion (or the exclusion), through a suitable placement of sensors, of potentially observable variables into the set of actual observations.
The most relevant result in Torta and Torasso [2007] concerns the fact that the minimization of the (additional) sensor sets can be done in polynomial time with respect to the size of the symbolic representation of the sets of sensors satisfying the discriminability requirements. Another important property of this approach is that it naturally handles the computation of MASS that guarantee diagnosability also when the single fault assumption is lifted.
The main goal of the present paper is to investigate whether (and to what extent) the symbolic approach proposed in Torta and Torasso [2007] can be extended and adapted in order to deal with system models given in terms of numeric equations.
As mentioned above, the problem of MASS computation for such class of systems has been previously investigated and many FDI approaches have been proposed. In this paper we concentrate specifically on the approach presented in Travé-

Massuyès et al. [2001]. Our focus is not only to reproduce, with a different method, the results of the existing approach, but also to investigate the potential benefits of our method in terms of computational costs and flexibility, i.e. release of some assumptions, such as the single fault assumption.

The paper is structured as follows. In section 2, we review some of the basic concepts developed in Torta and Torasso [2007] for qualitative relational models, tailoring the definitions to the purposes of this paper. In section 3, after reviewing the relevant parts of the approach of Travé-Massuyès et al. [2001], we propose a mapping from the Resolution Process Graph of a numeric equation model to a suitable qualitative relational model. In section 4 we show how MASS can be computed from the qualitative relational model, and some nice properties of such a computational process. The application of the proposed method is illustrated in section 5 using as a test bed a gas turbine subsystem taken from Travé-Massuyès et al. [2001]. Finally, in section 6, we review the contributions of this paper in the light of the existing literature and potential further developments.

2. DISCRIMINABILITY AND MASS FOR QUALITATIVE RELATIONAL MODELS

In this section we provide the formal setting for characterizing the notion of diagnostic discriminability and the one of Minimal Additional Sensor Set. We start from the definition of *System Description* according to which the model is given in terms of discrete variables and qualitative relations among them.

Definition 1. A *System Description* is a pair SD = ($\mathcal{SV}$, DT) where:

- $\mathcal{SV}$ is the set of discrete system variables partitioned in C (system components), X (exogenous variables) and E (endogenous variables). We will denote with $D(v)$ the *finite* domain of variable $v \in \mathcal{SV}$; in particular, for each $c \in C$, $D(c)$ consists of the values *ok* and *ab* for representing respectively the nominal and faulty behavioral modes [1]
- DT (Domain Theory) is a relation over the variables in $\mathcal{SV}$ [2]

We are now ready for introducing the notion of discriminability of a component. Such a notion clearly depends on the degree of *observability* of the system, represented as a subset $O \subseteq E$ of observable endogenous variables.

Definition 2. Let $c \in C$ be a system component. We say that c is *discriminable* w.r.t. observability $O \subseteq E$ iff for each instance $\mathcal{X}$ of X:

$$\Pi_O(\sigma_{c(ok)}(DT \bowtie \mathcal{X})) \cap \Pi_O(\sigma_{c(ab)}(DT \bowtie \mathcal{X})) = \emptyset$$

where Π, σ and $\bowtie$ are the *project*, *select* and *join* operations defined in the relational algebra.

According to the above definition, component c is considered discriminable w.r.t. a given observability O when, for each possible assignment $\mathcal{X}$ to the inputs, it is possible to tell whether c is ok or ab by looking just at the values of the O variables.

Note that the values of the observables O depend not only on the behavioral mode of c and the value of the inputs X, but also on the behavioral modes of the other components $C\backslash\{c\}$.

[1] Actually, the domain of component variables can contain more than one faulty behavioral mode. For the purpose of this paper, however, it is sufficient to consider the *ab* mode.

[2] Usually, in compositional systems, relation DT is obtained by joining a set of relations $DT_1, \ldots, DT_m$ modeling parts of the system behavior.

However, in the above definition we require that the values of O consistent with $c(ok)$ and the input $\mathcal{X}$ are disjoint from the values of O consistent with $c(ab)$ regardless of the assignments of behavioral modes to components $C\backslash\{c\}$.

This has also a significant consequence on the generality of the definition: in fact there is no assumption on the maximum number of simultaneous faults that may affect the system, i.e. any combination of *ok* and *ab* values for the components $C\backslash\{c\}$ is allowed.

Since we are interested in verifying whether a set of components is discriminable and what kind of observability guarantees such a discrimination, we introduce the notions of discriminability *requirement* and Minimal Sensor Set.

Definition 3. A discriminability requirement δ involving a component variable c is satisfied by an observability O iff c is discriminable w.r.t. O according to Definition 2.

We denote as $\Delta = \{c_{\Delta,1}, \ldots, c_{\Delta,k}\}$ a set of discriminability requirements; Δ is satisfied by O iff O satisfies the discriminability requirements for $c_{\Delta,1}, \ldots, c_{\Delta,k}$.

Definition 4. A *Minimal Sensor Set* w.r.t. the set of requirements Δ is an observability O^* which satisfies Δ such that no other observability O', $|O'| < |O^*|$ satisfies Δ.

The preference criterion for selecting Minimal Sensor Sets is based on the minimum cardinality, reflecting the assumption that all the sensors have an equal cost; we will comment on sensors with (qualitatively) different costs in section 6.

In many practical cases the availability of some sensors is known a priori, so one could be interested in solving the problem of minimizing the additional sensors.

Definition 5. Given a set of sensors (i.e. observability) $O_{AV} \subseteq E$, a *Minimal Additional Sensor Set* w.r.t. the set of requirements Δ is an observability O^* such that $O_{AV} \cup O^*$ satisfies Δ and for no other observability O', $|O'| < |O^*|$, $O_{AV} \cup O'$ satisfies Δ.

Our goal is to find a minimal observability O^* w.r.t. the requirements in Δ. As stated in the introduction, the adoption of switches for representing different degrees of observability has been exploited in previous works in order to make possible the computation of MASS for compiled qualitative models with some formal guarantees on the computational complexity (for details see Torta and Torasso [2007, 2008]).

Definition 6. Given an endogenous variable $e \in E$ its associated *observation switch* sw_e is a variable with $dom(sw_e) = \{yes, no\}$ which takes the value *yes* or *no* depending on the fact that endogenous variable e is currently observable or not . We denote as SW the set of switches.

A switch can be used to specify a relation between an endogenous variable e and what is observed about the value of e. In particular, when sw_e is set to no, variable e is not observed at all and therefore the observation has a special value abs (absent). Once the switches and their models are added to the Domain Theory DT we get an extended Domain Theory XDT.

In the following sections we will show how the idea of extending DT with switches can be exploited for computing MASS in a structural approach. In particular, we will show how the notion of indiscriminability given in Definition 2 can be reformulated in terms of assignments to switches and corresponding observed values; we will then discuss an algorithm which computes MASS based on this reformulated definition of discriminability.

name	component(s)	equation
r1	Injt	$q_3 = K_{inj} \times \sqrt{p_3 - cpd^*}$
r2		$q_4 - K_{li} \times q_3 = 0$
r3	GCVh	$q_2 = fsg \times \sqrt{fpg_2 - p_3}$
r4	GCVh	$q_3 - KI \times q_2 = 0$
r5	SRVh	$fqg = fsgr \times \sqrt{p_1^* - fpg_2}$
r6	SRVh	$q_2 - KI' \times fqg = 0$
r7	GCVm	$fsg = f(fag, 96hql^*)$
r8	SRVm	$fsgr = f(fagr, 96hql^*)$
r9	GCVm	$fsg = f(fsrout^*, 96hql^*)$
r10	SRVm	$fsgr = f(fprgout^*, fpg_2, 96hql^*)$
r11	SRVm & SRVh	$fpg_2 = f(fprgout^*)$

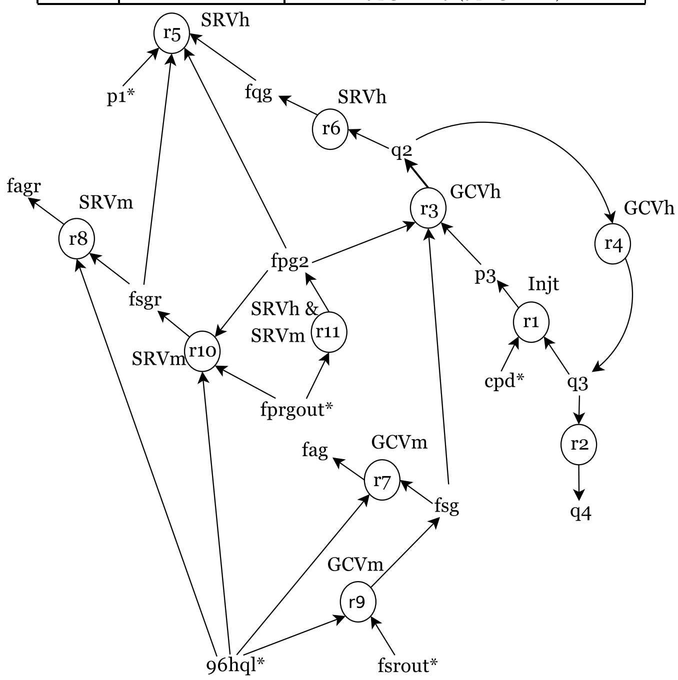

Fig. 1. Model and RPG for the GFS System.

3. A STRUCTURAL APPROACH BASED ON QUALITATIVE RELATIONAL MODELS

In this paper we consider the class of system models that was addressed in Travé-Massuyès et al. [2001]. In particular, each system model is characterized by a set of components and a set of numeric equations, where each equation is possibly associated with one or more components. The equations are defined in terms of endogenous variables (which are initially assumed to be all unknown) and exogenous variables (which are assumed to be known), and each component can be either in the *ok* or *ab* (abnormal) modes.

The upper part of Figure 1 shows the model of a Gas Fuel Subsystem (GFS) as presented in Travé-Massuyès et al. [2001], where the names of exogenous variables end with the star character.

3.1 Resolution Process Graph and Redundant Relations

As discussed in Travé-Massuyès et al. [2001], from this kind of numeric equation models it is possible to derive a Resolution Process Graph (RPG) which represents a causal ordering for the resolution of the equations in the system model, i.e. the RPG defines the dependency paths among variables which indicate the order in which every equation should be used to solve successively for the unknown variables [3].

In particular, for each unknown variable v, the RPG contains a node N_v with an incoming arc from a node N_{eq} which represents the equation eq matched with v in the RPG. If equation eq involves v plus other endogenous and exogenous variables $v_1, \ldots, v_k$, then nodes $N_{v_1}, \ldots, N_{v_k}$ in RPG are connected to node N_{eq} with outgoing arcs; the intended meaning is that values $v_1, \ldots, v_k$ will be used for determining the value of v via eq.

The number of equations in the model is, in general, greater or equal to the number of unknown variables; in the latter case some equations, called Redundant Relations (RR), appear as sink nodes in the RPG (i.e. they have only incoming arcs). The lower part of Figure 1 shows the RPG for the GFS model; note that equation $r5$ is a Redundant Relation. For example, variable p_3 has an incoming arc from its matched equation r_1 which, in turn, has incoming arcs from exogenous variable cpd^* and q_3.

Since RRs are not needed to determine the values of the unknown variables, they can be reformulated in terms of known variables (Analytical Redundant Relations); therefore, it is possible to check whether an RR r is satisfied or not in the situation defined by the known variables. This is reflected in the value of a variable called *residual* which assumes the value z (for *zero*) if r is satisfied and nz (for *non-zero*) is r is not satisfied.

Since the residual variables can take only two values, they can be modeled as discrete qualitative variables; moreover their value, which depends only on known values, is always known.

As in many other works (including Travé-Massuyès et al. [2001]), we assume that the exoneration working hypothesis holds: a faulty component c always implies that the residuals of all the (Analytical) RRs in which it is involved are non-zero.

In our example RPG of Figure 1, all of the system components are involved by the RR $r5$ since, in order to reformulate $r5$ in terms of the (known) exogenous variables $p_1^*, 96hql^*, fprgout^*, fsrout^*$ and cpd^*, we need equations $r1$, $r3$, $r4$, $r6$, $r9$, $r10$, $r11$ (and, of course $r5$ itself) which, all together, involve components $SRVh$, $Injt$, $GCVh$, $GCVm$, $SRVm$, i.e. all of the system components in the GFS.

3.2 Building a Qualitative Relational Model

In this section we show how, starting from an RPG, it is possible to automatically build a qualitative relational model which includes switch variables, along the lines of section 2. Each switch variable allows to represent the presence/absence of a sensor for measuring an endogenous variable e which, in the RPG, is matched with a non-RR r; the addition of the sensor makes relation r become a RR and, therefore, a new residual can be associated with r.

As we will see, the residuals of the RRs in the original RPG and the residuals of new RRs that arise as a consequence of observing endogenous system variables constitute the observable variables in the qualitative relational model.

We start by defining some of the qualitative system variables of the System Description which correspond to variables in the RPG:

- the set C of component variables contains a variable c with domain $D(c) = \{ok, ab\}$ for each component in the RPG
- the set X of exogenous variables contains a variable x with domain $D(x) = \{nom, abk\}$ for each exogenous variable in

[3] The RPG can be obtained straightforwardly from a *perfect matching* between the endogenous variables and a subset of the equations (Cassar and Staroswiecki [1997]).

the RPG; in particular, nom means that x has a nominal value, while abk means that x has an abnormal (but known) value

- the set E of endogenous variables contains a variable e with domain $D(e) = \{nom, abk, unk\}$ for each endogenous variable in the RPG; in particular, nom means that e has a nominal value, abk means that e has an abnormal (but known) value and unk means that e has an abnormal *and* unknown value

Note that, since we assume that exogenous variables are always known, they can only assume qualitative values nom and abk, while endogenous variables may also take an unknown abnormal value unk. We also define three additional sets of variables:

- the set RES of residual variables; for each equation r in the RPG, there is a variable $res_r \in RES$ with possible values abs (absent), z (zero) and nz (non-zero). As we will see, a residual variable res_r takes the abs value when the equation r is not a RR in the RPG and the endogenous variable e_r matched with r in the RPG is not sensorized
- the set $\hat{E}$ of *propagated* endogenous variables; for each endogenous variable e in the RPG, there is a variable $\hat{e} \in \hat{E}$ with possible values nom, abk and unk. As we will see, *after* an endogenous variable e has been observed with a sensor, the value that is propagated to solve the dependent equations may be different than the value of e itself (in particular, the value of $\hat{e}$ may be abk when the value of e is unk)
- the set SW of switch variables; for each non-RR r in the RPG, there is a variable $sw_{e_r} \in SW$ with possible values yes and no, whose meaning is that the endogenous variable e_r matched with r in the RPG *is* (resp. *is not*) sensorized

Note that we introduce a residual for each equation in the model since, by adding sensors, all of the equations may become RR and therefore have an associated residual. In the GFS system, the set of residuals RES, beside a residual res_{r5} for RR r_5, also contains residuals for the other equations $r_1 - r_4$ and $r_6 - r_{11}$; such residuals will take values different than abs when the switches of their matched variables are set to *yes*. For example, if sw_{p3} has value yes (i.e. p_3 is sensorized), relation r_1 becomes a RR and, therefore, the residual res_{r_1} will be allowed to take as value either z or nz, but not abs.

The next step consists in building a number of qualitative relations that specify the constraints among the C, X, E, RES, $\hat{E}$ and SW variables.

For each non-RR r, we build a qualitative relation QR_r which determines the value of the endogenous variable e_r matched with r in the RPG as a function of the values of the other (propagated) endogenous variables as well as the exogenous and component variables that appear in r (denoted respectively with $\hat{E}_r$, X_r and C_r). Figure 2 shows how QR_r is built. The need of expressing the value of e_r in terms of the values of the propagated endogenous variables $\hat{E}_r$ instead of the base endogenous variables E_r stems from the fact that each endogenous variable e that has an outgoing arc towards r in the RPG may have been observed with a sensor (see the description of switch relations below).

Relation QR_r contains tuples in which e_r has value nom, tuples where e_r has value abk and tuples where e_r has value unk (line 14).

The set QR_{nom} of tuples in which e_r has value nom is computed in lines 3-6 and contains all the assignments to $X_r \cup \hat{E}_r \cup C_r$ s.t. all the component variables are ok and all the exogenous and propagated endogenous variables have the nominal value nom.

Algorithm **BuildQR**

builds a qualitative relation QR_r corresponding to a non-RR r

e_r: variable matched with r in the RPG

$E_r = \{e_1, \ldots, e_l\}$: other endogenous variables in equation r

$X_r = \{x_1, \ldots, x_k\}$: exogenous variables in equation r

$C_r = \{c_1, \ldots, c_m\}$: component variables in equation r

1 $\hat{E}_r = \{\hat{e} \in \hat{E} : e \in E_r\}$
2 $D(r) = D(X_r) \times D(\hat{E}_r) \times D(C_r) \times D(\{e_r\})$
3 $\varphi_{nom} = (\forall \hat{e} \in \hat{E}_r : \hat{e}(nom))$
4 $\varphi_{nom} = \varphi_{nom} \wedge (\forall x \in X_r : x(nom))$
5 $\varphi_{nom} = \varphi_{nom} \wedge (\forall c \in C_r : c(ok))$
6 $QR_{nom} = \sigma_{\varphi_{nom} \wedge e_r(nom)} D(r)$
7 $\varphi_{abk} = (\exists \hat{e} \in \hat{E}_r : \hat{e}(abk)) \vee (\exists x \in X_r : x(abk))$
8 $\varphi_{abk} = \varphi_{abk} \wedge (\forall \hat{e} \in \hat{E}_r : \hat{e}(nom) \vee \hat{e}(abk))$
9 $\varphi_{abk} = \varphi_{abk} \wedge (\forall c \in C_r : c(ok))$
10 $QR_{abk} = \sigma_{\varphi_{abk} \wedge e_r(abk)} D(r)$
11 $\varphi_{unk} = (\exists \hat{e} \in \hat{E}_r : \hat{e}(unk))$
12 $\varphi_{unk} = \varphi_{unk} \vee (\exists c \in C_r : c(ab))$
13 $QR_{unk} = \sigma_{\varphi_{unk} \wedge e_r(unk)} D(r)$
14 $QR_r = QR_{nom} \cup QR_{abk} \cup QR_{unk}$

Fig. 2. Building Qualitative Relations for RRs.

Note that, in line 6, we make use of the set $D(r)$ which represents all the possible assignments to the variables that appear in QR_r, namely $X_r \cup \hat{E}_r \cup C_r \cup \{e_r\}$. Such a set is defined in line 2 as the Cartesian product of the domains (i.e. sets of possible assignments) of X_r, $\hat{E}_r$, C_r and e_r.

Technically, the set QR_{nom} is computed by building a logic condition φ_{nom} equivalent to what we have informally stated above, and by selecting (with the σ operator) from $D(r)$ those tuples that satisfy φ_{nom} and $e_r(nom)$.

Similarly, the set QR_{abk} of tuples in which e_r has value abk is computed in lines 7-10 and contains all the assignments to $X_r \cup \hat{E}_r \cup C_r$ s.t. all the component variables are ok and all the exogenous and propagated endogenous variables have either the nominal value nom or an abnormal but known value abk but at least one of them has value abk (otherwise e_r would have a nom value). Indeed, provided all of the components associated with equation r are ok, if one or more exogenous or propagated endogenous variables have abnormal (but known) values, the value of e_r that we predict by using r will be abnormal but it will be known.

Finally, the set of tuples QR_{unk} in which e_r has value unk is computed in lines 11-13 and contains all the assignments to $X_r \cup \hat{E}_r \cup C_r$ s.t. at least one component variable is ab or at least one propagated endogenous variable has an unknown abnormal value unk. Indeed, in such a case, the value of e_r that we predict by using r will be abnormal and unknown.

For the GFS system of Figure 1, we build a relation QR_r for all the equations $r_1, \ldots, r_{11}$, except for r_5 which is a RR.

For each non-RR r, we also build a qualitative switch relation SWR_r which expresses the relation between the endogenous variable e_r and its associated propagated variable $\hat{e_r}$ depending on the value of switch variable sw_{e_r}. Such a relation also determines the value of the residual res_r of r which is absent when the switch has value no (since, in such a case, r is a non-

Algorithm **BuildSWR**
builds a qualitative switch relation SWR_r corresponding to a non-RR r
Inputs:
e_r: variable matched with r in the RPG

1 $DSW(r) = D(sw_{e_r}) \times D(e_r) \times D(\hat{e}_r) \times D(res_r)$
2 $\varphi_{abs} = sw_{e_r}(no)$
3 $SWR_{abs} = \sigma_{\varphi_{abs} \wedge (\hat{e}_r = e_r) \wedge res_r(abs)} DSW(r)$
4 $\varphi_z = sw_{e_r}(yes) \wedge (e_r(nom) \vee e_r(abk))$
5 $SWR_z = \sigma_{\varphi_z \wedge (\hat{e}_r = e_r) \wedge res_r(z)} DSW(r)$
6 $\varphi_{nz} = sw_{e_r}(yes) \wedge e_r(unk)$
7 $SWR_{nz} = \sigma_{\varphi_{nz} \wedge \hat{e}_r(abk) \wedge res_r(nz)} DSW(r)$
8 $SWR_r = SWR_{abs} \cup SWR_z \cup SWR_{nz}$

Fig. 3. Building Switch Relations for non-RRs.

RR), but can take meaningful values when the switch has value yes (and r becomes a RR). Figure 3 shows how SWR_r is built.
Relation SWR_r contains tuples in which res_r has value abs, tuples where res_r has value z (zero) and tuples where res_r has value nz (non-zero) (line 8).
The set SWR_{abs} of tuples in which res_r has value abs (computed in lines 2-3) contains the assignments where the switch sw_{e_r} has value no, as explained above. In such a case, the value of the propagated variable $\hat{e}_r$ is set to be the same as the value of e_r (i.e. either nom, abk or unk).
The set SWR_z of tuples in which res_r has value z (computed in lines 4-5) contains the assignments where sw_{e_r} has value yes and the endogenous variable e_r (which is sensorized) has either the nominal value nom or an abnormal but known value abk. Also in such a case, the value of the propagated variable $\hat{e}_r$ is set to be the same as the value of e_r (i.e. either nom or abk).
Finally, the set of tuples SWR_{nz} in which res_r has value nz (computed in lines 6-7) contains only one assignment, where sw_{e_r} has value yes, the sensorized endogenous variable e_r has value unk and the propagated variable $\hat{e}_r$ has value abk. This is the case when, thanks to the presence of the sensor measuring e_r, the value of variable $\hat{e}_r$ (which is propagated to solve equations that are causally downstream in the RPG) becomes an abnormal but known value.

The third and last kind of relations we build are qualitative relations QRR_r which, for each RR r in the RPG, determine the value of the residual res_r as a function of the values of the exogenous, (propagated) endogenous and component variables that appear in r. Figure 4 shows how QRR_r is built.
Relation QRR_r contains tuples in which res_r has value z and tuples where res_r has value nz (line 9) (since r is a RR, its residual res_r is always known, so it never takes value abs).
The set QRR_z of tuples in which res_r has value z, is computed in lines 3-5 and contains all the assignments to $X_r \cup \hat{E}_r \cup C_r$ s.t. all the component variables are ok and all the propagated endogenous variables have either the nominal value nom or an abnormal but known value abk. In other words, provided all of the components associated with equation r are ok, the residual of r is zero even if one or more propagated endogenous variables have known abnormal values; indeed, when we evaluate r by substituting the known abnormal values and the predicted nominal values of the other propagated endogenous variables, the equation is satisfied.
Similarly, the set of tuples QRR_{nz} in which res_r has value nz, is computed in lines 6-8 and contains all the assignments to $X_r \cup \hat{E}_r \cup C_r$ s.t. at least one component variable is ab or

Algorithm **BuildQRR**
builds a qualitative relation QRR_r corresponding to a RR r
Inputs:
$E_r = \{e_1, \ldots, e_l\}$: endogenous variables in equation r
$X_r = \{x_1, \ldots, x_k\}$: exogenous variables in equation r
$C_r = \{c_1, \ldots, c_m\}$: component variables in equation r
res_r: residual variable for equation r

1 $\hat{E}_r = \{\hat{e} \in \hat{E} : e \in E_r\}$
2 $D(r) = D(X_r) \times D(\hat{E}_r) \times D(C_r) \times D(res_r)$
3 $\varphi_z = (\forall \hat{e} \in \hat{E}_r : \hat{e}(nom) \vee \hat{e}(abk))$
4 $\varphi_z = \varphi_z \wedge (\forall c \in C_r : c(ok))$
5 $QRR_z = \sigma_{\varphi_z \wedge res_r(z)} D(r)$
6 $\varphi_{nz} = (\exists \hat{e} \in \hat{E}_r : \hat{e}(unk))$
7 $\varphi_{nz} = \varphi_{nz} \vee (\exists c \in C_r : c(ab))$
8 $QRR_{nz} = \sigma_{\varphi_{nz} \wedge res_r(nz)} D(r)$
9 $QRR_r = QRR_z \cup QRR_{nz}$

Fig. 4. Building Qualitative Relations for RRs.

at least one propagated endogenous variable has an unknown abnormal value unk. Indeed, if one of the components is ab, the residual of r is non-zero by the exoneration hypothesis; and, even when all the components are ok, if we evaluate r by substituting the predicted nominal value for a variable that has an unknown abnormal value, the equation is *not* satisfied, i.e. its residual is non-zero.
For the example GFS system of Figure 1, we build a relation QRR_r only for equation r_5, which is the only RR in the RPG.

3.3 Compiled Extended Domain Theory

The method described in the previous section, starting from an RPG of a set of equations, creates a set of qualitative relations. In particular, as we have seen, the resulting set of relations contains a relation QR_r and a switch relation SWR_r for each non-RR, and a relation QRR_r for each RR.
It is then possible to build a qualitative Domain Theory XDT simply by joining all of these parts together (as mentioned in Definition 1). We use the notation XDT to stress the fact that the model built in this way is extended with switch variables and other related variables; indeed, as detailed in the previous section, beside the basic sets C, X and E of component, exogenous and endogenous variables, the set $\mathcal{XSV}$ of (extended) system variables also includes residuals RES, propagated endogenous variables $\hat{E}$ and switch variables SW, and the domain theory XDT is defined over all of these variables.

While the relations QR_r, SWR_r and QRR_r have limited size since they are just local theories involving a limited number of variables, the global Domain Theory XDT may be very large. For this reason it becomes of critical importance the ability of expressing XDT (and the other relations that we will needed for computing the MASS) in a compact form.
In particular, following Torta and Torasso [2007, 2008], in the present paper we have adopted Ordered Binary Decision Diagrams for encoding XDT and other relations involved in our algorithms. For space and clarity reasons, we will keep expressing the algorithms in the following sections in terms of relational algebra operations over extensional relations, instead of in terms of operations on OBDDs; the compilation of a relation with an OBDD and the mapping between relational algebra and OBDD operations is quite straightforward (for an OBDD-based implementation of diagnosis see Torasso and Torta [2006]).

Here we want just to point out that the compilation into OBDDs generally has a huge impact on reducing the size of the encodings of the relations, and will also provide an interesting computational complexity result on the minimization of MASS (see Property 1).
It is worth noting that the extended domain XDT has many similarities with the Extended HFS Matrix of Travé-Massuyès et al. [2001], since they both record the relations between component failures and values of the residuals, in a way conditioned by the set of sensors. However, there are relevant differences:
- XDT is parsimoniously computed and encoded in OBDD compiled form
- thanks to this fact, XDT is able to define the failures-residuals relations not only under the single-fault assumption, but for the combinatorially larger space of multiple-fault situations
- the sets of sensors which condition the failures-residuals relations are explicitly represented in XDT through the switch variables; this will be the basis for the computation of MASS described in the following section. In particular, we note that an instance $\mathcal{Y}_{SW}$ of the SW variables identifies a degree of observability (i.e. a Sensor Set) $O_{\mathcal{Y}_{SW}}$: a variable e is in $O_{\mathcal{Y}_{SW}}$ iff the switch sw_e is set to yes

Before describing the computation of MASS, we explicitly redefine the notion of discriminability into an equivalent one expressed in terms of the extended domain theory.

Definition 7. Let $c \in C$ be a system component. We say that c is *discriminable* w.r.t. observability $O \subseteq E$ iff for each instance $\mathcal{X}$ of X:
$\Pi_{RES}(\sigma_{c(ok)}(XDT \bowtie \mathcal{Y}_{SW} \bowtie \mathcal{X})\cap$
$\cap\Pi_{RES}(\sigma_{c(ab)}(XDT \bowtie \mathcal{Y}_{SW} \bowtie \mathcal{X}) = \emptyset$
where $\mathcal{Y}_{SW}$ is an assignment to SW variables s.t. $O_{\mathcal{Y}_{SW}} = O$

It is worth noting that, compared to Definition 2, a specific observability O is replaced by $\mathcal{Y}_{SW}$, while the observable traces are expressed in terms of the variables RES introduced for representing the residuals. This notion of indiscriminability will prove to be essential for computing MASS with the approach described in the next section.

4. COMPUTING MASS

Before describing in detail the computation of the MASS, we briefly summarize the main steps. The starting point is a set of discriminability requirements $\Delta = \{c_{\Delta,1}, \ldots, c_{\Delta,k}\}$ pointing out which components the user is interested to discriminate (for full diagnosability, Δ is equal to the whole set of components). Given a specific discriminability requirement $\delta_i = c_{\Delta,i}$, the system computes the set SS_{δ_i} which includes all the sensor sets that guarantee the discriminabilty of $c_{\Delta,i}$. We iterate the process for each discriminability requirement in Δ obtaining sets $SS_{\delta_1}, \ldots, SS_{\delta_k}$ and, by intersecting these sets, we get the set of all sensor sets SS_Δ which satisfy all the discriminability requirements in Δ.

At this point the user can go on in the analysis by requiring the system to compute the Minimal Sensor Sets or the Minimal Additional Sensor Sets. In particular, the user can specify a set of constraints Ω on the sensors: more specifically, it is possible to specify that an endogenous variable e is certainly observed (by adding a constraint in Ω that assigns the value yes to sw_e), or that e is not to be considered (in this case the constraint in Ω assigns value no to sw_e). Note that Ω can be conveniently expressed as a partial assignment to the SW variables.
If the user puts no constraint in Ω, the system computes the MSS, otherwise it computes the MASS, according to definitions 4 and 5.
Using SS_Δ and Ω, the minimization module is able to compute all the Minimum Additional Sensor Sets and therefore is able to provide the user with the MASS that satisfy his/her discrimination requirements.
It is worth noting that, in general, the globally optimal set of sensors (i.e. the MASS for $\Delta = \{c_{\Delta,1}, \ldots, c_{\Delta,k}\}$ and Ω) can *not* be obtained as the union of the locally optimal MASS for requirements $c_{\Delta,1}$, ..., $c_{\Delta,k}$, since such a union may be globally suboptimal. Therefore the optimization problem cannot be distributed but has to be solved at the global level and this is challenging in terms of computational cost; however, we will show that our compilation-based approach will allow us to have some formal guarantee on the computational complexity.

A final remark concerns the generality of the approach. In the approach sketched above there is no feature that depends on the particular kind of system model: in fact, the presented approach can be applied both to systems directly modeled as qualitative relation systems and to equation based systems once they have been reformulated in terms of qualitative systems.
In the following two subsections, we show how to compute SS_Δ and, then, how to compute the MASS by taking into consideration a set of constraints Ω on the sensors.

4.1 Computing SS_Δ

As stated above, we first compute SS_δ for each specific discriminability requirement δ involving component *c*. This computation is summarized in Figure 5 where the computational steps are expressed in terms of relational operations project, select and join. It is worth recalling that the actual implementation of the algorithm is based on OBDDs which encode the relations and the relational operations are expressed in terms of the operators working on the OBDDs.
Coming back to the algorithm, first of all we compute two relations H_{ok} and H_{ab} by restricting the extended domain theory XDT to the cases when the component c is ok and to the complementary cases when the component c is ab.
We have now to check that the two cases are discriminable in terms of observable traces. For this reason we project relations H_{ok} and H_{ab} on variables which are relevant for discriminability, and therefore we maintain in the two relations just the variables X (exogenous variables which are known), RES (residuals which are always known but take the value abs in case the corresponding endogenous variable has not an associated sensor) and SW (switches which allow to capture the different levels of observability). Note that the behavioral mode of the components are forgotten, since definition 7 requires that the two behavioral modes of c are discriminable regardless of the assignments of behavioral modes to the other components.

In line 5 we compute relation H_{com} by intersecting H_{ok} and H_{ab}. The tuples of relation H_{com} are assignments $\mathcal{X} \cup \mathcal{Y}_{SW} \cup \mathcal{Y}_{RES}$ to the $X \cup SW \cup RES$ variables; the presence of such a tuple in H_{com} means that, when exogenous variables have value $\mathcal{X}$ and we observe $\mathcal{Y}_{RES}$ as the values of residuals, there exists at least one assignment to the component variables consistent with $\mathcal{X}$ and $\mathcal{Y}_{RES}$ where $c(ok)$ holds and at least another such assignment where $c(ab)$ holds.
In other words, a tuple $\mathcal{X} \cup \mathcal{Y}_{SW} \cup \mathcal{Y}_{RES}$ in H_{com} indicates that, at least under input $\mathcal{X}$, $c(ok)$ and $c(ab)$ are not discriminable

ComputeSSdelta(XDT, c)

1 $H_{ok} = \sigma_{c(ok)}(XDT)$
2 $H_{ok} = \Pi_{X \cup SW \cup RES}(H_{ok})$
3 $H_{ab} = \sigma_{c(ab)}(XDT)$
4 $H_{ab} = \Pi_{X \cup SW \cup RES}(H_{ab})$
5 $H_{com} = H_{ok} \cap H_{ab}$
6 $SS_{com} = \Pi_{SW}(H_{com})$
7 $SS_{\delta} = D(SW) \backslash SS_{com}$

Fig. 5. Computation of SS_{δ}.

under the observability encoded by $\mathcal{Y}_{SW}$. Since Definition 7 requires that discriminability holds for all inputs $\mathcal{X}$, we project H_{com} on SW variables in order to isolate the observabilities that violate this requirement (i.e. in line 6 we compute SS_{com}). By complementing this set of observabilities, we obtain the set of observabilities SS_{δ} that satisfy the requirement of discrimination for component c. This is done in line 7 by subtracting SS_{com} from $D(SW)$ which represents the set of all possible assignment to SW.

Once the sets SS_{δ_i} have been computed for each discriminability requirement in Δ, the set of sensors which satisfy all the discriminability requirements can be easily computed as:

$$SS_{\Delta} = SS_{\delta_1} \cap \ldots \cap SS_{\delta_k}$$

4.2 Computing MASS from SS_{Δ}

The final step consists in computing the Minimum Additional Sensor Sets by exploiting SS_{Δ} and by taking into account the constraints provided by the user on the presence/absence of sensors (i.e Ω).

The computation of the MASS is performed by the function *CompMASS* (Figure 6). First of all (line 1) the algorithm takes into consideration the set of constraints Ω on the sensors provided by the user. Note that in case the user may not be willing or able to provide any constraint, $\Omega = \emptyset$.

The relation $SS_{\Delta,\Omega}$ contains now all the sensor sets satisfying the discrimination requirements in Δ compatible with the constraints in Ω.

The minimization (lines 2-6) is performed by exploiting a set of precomputed sets of sensors CSS_i, where a generic CSS_i contains all the possible combinations of switches sw_e with exactly i switches assuming the value yes.

In other words each CSS_i represents all the possible observabilities that involve exactly i observable variables. Therefore CSS_0 represents the case where nothing is observable (all the switches sw_e have value no) while $CSS_{|E|}$ represents the case when all the $|E|$ endogenous variables are actually observed (all the switches sw_e have value yes).

Due to lack of space, we do not report the details of the computation of sets CSS_i. The algorithm for such a computation (reported in Torta and Torasso [2007]) iteratively computes each CSS_i starting from CSS_0.

Given that we have at disposal the sets CSS_i, the minimization step can be implemented in a very simple way (lines 2-6): it is sufficient to verify whether the intersection of $SS_{\Delta,\Omega}$ with CSS_i is not empty starting from CSS_0. As soon as we find a non-empty intersection for index i, relation *MASS* represents the set of all the possible combinations of i sensors which satisfy both the discriminability requirements and the constraints on the sensors.

We have the guarantee that i is the minimum number of

CompMASS(SS_{Δ}, Ω)

1 $SS_{\Delta,\Omega} = SS_{\Delta} \bowtie \Omega$
2 i = 0
3 $MASS = SS_{\Delta,\Omega} \cap CSS_i$
4 **while** $MASS == \emptyset \wedge i < |E|$
5   i = i + 1
6     $MASS = SS_{\Delta,\Omega} \cap CSS_i$

Fig. 6. Computation of Set *MASS*.

switches since we have already verified that with $0, 1, \ldots, i-1$ sensors we fail in finding a solution.

Thanks to compilation of the relations with OBDDs and the fact that the OBDDs which encode sets CSS_i have a polynomial size in the number of switches SW (and consequently of the endogenous variables E, see Torasso and Torta [2006]), it turns out that the potentially very expensive task of computing the minimal (additional) sensor sets can be done in polynomial time with respect to the size of the OBDD $\mathcal{O}(SS_{\Delta})$ which encodes SS_{Δ}, as formalized by the following property.

Property 1. Let $\mathcal{O}(\mathcal{SS}_{\Delta})$ be an OBDD encoding set SS_{Δ}; then, OBDD $\mathcal{O}(MASS)$ encoding all the MASS can be computed by *CompMSS* in time $O(|E|^3 \cdot |\mathcal{O}(SS_{\Delta}|)$.

This property mirrors a similar result obtained for Minimum Cardinality Diagnoses whose proof is reported in Torasso and Torta [2006]. When the OBDD encoding SS_{Δ} is small, we have the guarantee that *CompMSS* can *always* be executed efficiently.

5. APPLICATION TO THE GFS

We have applied the approach described in the previous sections to the GFS system, whose RPG is reported in Figure 1. The experiments have been conducted with a Java implementation of the algorithms that uses the JBDD interface to the Buddy library for the OBDD operations; the test machine was equipped with an Intel Core Duo CPU at 2.4GHz and 2GB of RAM.

The OBDD representing the compiled XDT has a size of 776 nodes, and is computed in less than 10msec. Although the example system is quite small (51 variables including the switch, residual and propagated endogenous variables), 776 nodes are a very limited size for representing all of the tuples of the (global) extended domain theory XDT.

First, we compute the MASS under the same conditions as the ones of the example presented in Travé-Massuyès et al. [2001]:
- we make the single fault assumption; in order to enforce this assumption, we intersect the OBDD which represents XDT with an OBDD CD_1 representing all of the single fault diagnoses plus the assignment where all components are ok. The OBDD CD_1 is computed in a similar way as the Cardinality Sensor Sets, the main difference being that the role of yes/no sensors is played by ok/ab components - we let Ω be the set $\{sw_{fpg2}(yes), sw_{fqg}(yes), sw_{fsg}(yes), sw_{fsgr}(yes), sw_{q4}(no)\}$

The OBDD CD_1 has a size of 14 nodes, while the OBDD XDT_1 representing $XDT \cap CD_1$ has a size of 763 nodes, and is computed in a negligible amount of time.

We compute SS_{δ} for each requirement $\delta = c$, $c \in C$ and then intersect all of the SS_{δ} obtaining the set SS_{Δ} of all the sensor sets which satisfy all the requirements. The size of the OBDD representing SS_{Δ} is 63 nodes, while the maximum size of the sets SS_{δ} is 60 nodes. The time required for computing SS_{Δ}

starting from XDT_1 is around 20 msec.
Finally, we intersect SS_Δ with Ω obtaining the OBDD for $SS_{\Delta,\Omega}$ (20 nodes) and we compute the MASS from $SS_{\Delta,\Omega}$, obtaining a unique optimal solution $\{sw_{p3}(yes)\}$, which is the same as the one computed in Travé-Massuyès et al. [2001]. These computations take a negligible amount of time.
As an example variation that we can compute with a negligible amount of time, we also compute the MASS starting from the same SS_Δ, but with $\Omega = \emptyset$ (i.e. we compute the MSS under the single-fault assumption). It turns out that there is a unique MSS with cardinality 3, namely $\{sw_{fqg}(yes), sw_{q2}(yes), sw_{p3}(yes)\}$.
Another variation consists in releasing the single fault assumption and compute the sensor sets which guarantee diagnosability regardless of the number of faults affecting the system. In such a case, if Ω is set as in Travé-Massuyès et al. [2001], we find the single MASS $\{q3(yes), p3(yes)\}$.
This MASS, together with the 4 sensors required by Ω, results in the set of 6 sensors $\{sw_{fpg2}(yes)$, $sw_{fqg}(yes)$, $sw_{fsg}(yes)$, $sw_{fsgr}(yes)$, $sw_{p3}(yes)$, $sw_{q3}(yes)\}$. The total time needed for this computation is around 20 msec.
As a last variation, we let $\Omega = \emptyset$ in the multiple-fault case, finding 3 different solutions of cardinality 6, namely:

- $mss_1 = \{sw_{fag}(yes)$, $sw_{q3}(yes)$, $sw_{p3}(yes)$, $sw_{fpg2}(yes)$, $sw_{q2}(yes)$, $sw_{fsgr}(yes)\}$
- $mss_2 = \{sw_{fsg}(yes)$, $sw_{q3}(yes)$, $sw_{p3}(yes)$, $sw_{fpg2}(yes)$, $sw_{fqg}(yes)$, $sw_{fsgr}(yes)\}$
- $mss_3 = \{sw_{fsg}(yes)$, $sw_{q3}(yes)$, $sw_{p3}(yes)$, $sw_{fpg2}(yes)$, $sw_{q2}(yes)$, $sw_{fsgr}(yes)\}$

Note that mss_2 is the same sensor set as the one obtained above as the union of Ω and the MASS $\{q3(yes), p3(yes)\}$.

6. CONCLUSIONS

The problem of computing the MASS for numeric equations system models has been deeply investigated in the FDI literature (e.g. Travé-Massuyès et al. [2001], Commault et al. [2006], Travé-Massuyès et al. [2006], Krysander and Frisk [2008]).
In the present paper we have proposed and discussed a novel method for computing MASS by exploiting recent techniques based on the symbolic compilation of qualitative system models (Torta and Torasso [2007]) within a structural approach suitable for numerical equations models.
Our work addresses the problem starting from a Resolution Process Graph which is assumed to be computed with some existing techniques developed within the FDI community (e.g. Travé-Massuyès et al. [2001]); the RPG is then mapped to a qualitative relational model and symbolic AI techniques are applied in order to compute the MASS. In this respect, the paper can be viewed as a bridge work across the AI and FDI approaches to model-based sensor placement; it is worth mentioning that other works have aimed at bridging these two approaches for related problems such as diagnosability (Cordier et al. [2006]) but, as far as we know, the previous works do not directly address the computation of MASS.
The main advantage of the proposed approach consists in its flexibility that is made possible by the adoption of compilation techniques. In particular, we have shown that it is possible to compute and compactly represent a set SS_Δ compactly encoding all the sensor sets satisfying the given discriminability requirements. Once SS_Δ has been built, it is possible to perform a wide number of minimizations under different conditions with formal guarantees on the computational complexity. Moreover, the same method can be used for computing SS_Δ under the single-fault assumption or in the multiple faults case.
A crucial point of our approach is the possibility of compiling the extended domain theory XDT and the set SS_Δ into OBDDs of reasonable size. While this was certainly true for the small system model studied in this paper, Torta and Torasso [2007] reports experiments on systems of non-trivial size; in particular, the compiled XDT and SS_Δ are quite compact for a system (namely, the *c74182* digital circuit from ISCAS85) involving 70 components and 28 observation switches.
The approach described in the present paper could be extended in a number of ways. In particular, we have adopted the strong notion of indiscriminability of Travé-Massuyès et al. [2001], whereby a mode must be discriminable regardless of the current inputs. A weaker notion of indiscriminability may require that the modes are discriminable just under at least one combination of the input values; our approach can be applied without any major change to this case, as shown in Torta and Torasso [2007]. Similarly, the discriminability requirements addressed in our paper consist in the discrimination of $c(m)$ from $\neg c(m)$ under any combination of modes of the other components. The approach can be easily extended in order to deal with different kinds of requirements, e.g. discrimination of $c_i(m_i)$ from $c_j(m_j)$ (see again Torta and Torasso [2007]).
Finally, while in the present paper we have used minimum cardinality of sensors as a preference criterion, an obvious generalization would be to allow different costs for the sensors. Our approach can be generalized to cover the case of a limited number of (qualitative) possible costs for each sensor by extending the notion of CSS relation.

REFERENCES

J.P. Cassar and M. Staroswiecki. A structural approach for the design of failure detection and identification systems. In *Proc. IFAC Control of Industrial Systems*, 1997.

C. Commault, J. Dion, and S. Agha. Structural analysis for the sensor location problem in fault detection and isolation. In *Proc. IFAC World Congress*, pages 949–954, 2006.

M.O. Cordier, L. Travé-Massuyès, and X. Pucel. Comparing diagnosability in continuous and discrete-event systems. In *Proc. DX*, pages 55–60, 2006.

M. Krysander and E. Frisk. Sensor placement for fault diagnosis. *IEEE Transactions on Systems, Man and Cybernetics PART A*, 38(6):1398–1410, 2008.

P. Torasso and G. Torta. Model-based diagnosis through obdd compilation: a complexity analysis. *Lecture Notes in Computer Science*, 4155:287–305, 2006.

G. Torta and P. Torasso. Computation of minimal sensor sets from precompiled discriminability relations. In *Proc. DX*, pages 202–209, 2007.

G. Torta and P. Torasso. Computation of minimal sensor sets for conditional testability requirements. In *Proc. ECAI*, pages 805–806, 2008.

L. Travé-Massuyès, T. Escobet, and R. Milne. Model-based diagnosability and sensor placement. application to a frame 6 gas turbine sub-system. In *Proc. IJCAI*, pages 551–556, 2001.

L. Travé-Massuyès, T. Escobet, and X. Olive. Diagnosability analysis based on component-supported analytical redundancy relations. *IEEE Transactions on Systems, Man and Cybernetics PART A*, 36(6):1146–1160, 2006.