Conflicts versus analytical redundancy relations

A comparative analysis of the model based diagnosis approach from the Artificial Intelligence and Automatic Control perspectives

M.O. CORDIER, P. DAGUE, F. LÉVY, J. MONTMAIN, M. STAROSWIECKI,
L. TRAVÉ-MASSUYÈS

Abstract— **Two distinct and parallel research communities have been working along the lines of the Model-Based Diagnosis approach: the FDI community and the DX community that have evolved in the fields of Automatic Control and Artificial Intelligence, respectively. This paper clarifies and links the concepts and assumptions that underlie the FDI analytical redundancy approach and the DX consistency-based logical approach. A formal framework is proposed in order to compare the two approaches and the theoretical proof of their equivalence together with the necessary and sufficient conditions is provided.**

Index Terms— **Model Based Diagnosis, Fault Detection and Isolation, Potential Conflict, Analytical Redundancy Relation Support, Parity Space Approach vs Consistency-Based Logical Approach.**

I. Introduction

DIAGNOSIS is an increasingly active research domain, which can be approached from different perspectives according to the knowledge available. The so-called Model-Based Diagnosis (MBD) approach rests on the use of an explicit model of the system to be diagnosed. The occurrence of a fault is captured by discrepancies between the observed behavior and the behavior that is predicted by the model. Fault localization then rests on interlining the groups of components that are involved in each of the detected discrepancies. A definite advantage of this approach with respect to others, such as the relational approach [27] or the pattern recognition approach [16], is that it only requires knowledge about the normal operation of the system, following a consistency-based reasoning method.

Two distinct and parallel research communities have been using the MBD approach. The FDI community has evolved in the Automatic Control field from the seventies and uses techniques from control theory and statistical decision theory. It has now reached a mature state and a number of very good surveys exist in this field ([26], [18], [17], [22], [9]).

The DX community emerged more recently, with foundations in the fields of Computer Science and Artificial Intelligence ([39], [15], [20], [13]). Although the foundations are supported by the same principles, each community has developed its own concepts, tools and techniques, guided by their different modeling backgrounds. The modeling formalisms call indeed for very different technical fields; roughly speaking analytical models and linear algebra on the one hand and symbolic and qualitative models with logic on the other hand. The fact that each community has its own terminology and its own set of conferences and publications results in a poor understanding of the work in both sides.

The French IMALAIA group, supported by the French National Programs on Automatic Control *GDR-Automatique* and on Artificial Intelligence *GDR-I3* and AFIA (Association Française d'Intelligence Artificielle), has been working along these lines, benefiting from the work already performed by the ALARM group [7] and related work in France (e.g. [33]). The goals of this work are to agree upon a common DX/FDI terminology, to identify links in the concepts, similarities and complementarities in the DX and FDI methods, and to contribute to a unifying framework, thus taking advantage of the synergy of complementary techniques from the two communities

This paper, which considerably details and extends ([10], [11]), clarifies and links the concepts that underlie the FDI analytical redundancy approach and the DX consistency-based logical approach[1]. In particular, the link between *structured parity equations or analytical redundancy relations* (ARR for short) and *conflicts* (in the sense of Reiter) is clarified by introducing the notions of *potential conflict* or *ARR support* and interpreting a conflict as the support of a non satisfied ARR. This extension of ([10], [11]) also highlights the role of *completeness properties* on the set of ARRs and proves the *formal match* of the two approaches under completeness conditions which are clearly stated and discussed.

The FDI and DX model-based approaches used for fault isolation are analyzed from the two perspectives. It is shown that the first one, based on fault signatures, proceeds along a column interpretation of the fault signature matrix linking faults and ARRs whereas the later one, based on conflicts,

M.O. Cordier, IRISA, Université Rennes1, Campus de Beaulieu, 35042 Rennes Cedex (France).

P. Dague, F. Lévy, LIPN-UMR 7030, Université Paris 13, 99 avenue J.B. Clément, 93430 Villetaneuse (France).

J. Montmain, EMA-CEA, Site EERIE, Parc George Besse, 30035 Nîmes Cedex 1 (France).

M. Staroswiecki, LAIL-CNRS, EUDIL, Université Lille I, 59655 Villeneuve d'Ascq Cedex (France).

L. Travé-Massuyès, LAAS-CNRS, 7, avenue du Colonel Roche, 31077 Toulouse Cedex 4 (France).

Authors are listed in alphabetical order.

[1] These are not the only model-based approaches in their respective communities, but both are prototypic in the sense that most other approaches can be expressed in their formalism. The appellations FDI and DX approaches are abused in the following for these prototypic methods.

proceeds along a row interpretation.

The results provided by the two approaches are then shown to be identical under completeness conditions and the theoretical proof is included. This is proved in the no exoneration case under the single fault and the multiple fault assumptions, the exoneration case being left for further investigations. For the sake of clarity, the study is carried out in a pure consistency-based framework, i.e. without fault models.

The example that has been chosen to support the comparative analysis throughout the paper is the well-known system from [12] composed of three multipliers and two adders referred as the *polybox example* (figure 1). It refers for the sake of simplicity to a typical static system, but the comparison achieved in this paper applies as well to systems with a dynamic behavior. Indeed, a time variable may occur in behavioral equations, and thus in ARRs and signature matrix, and in observations. The only important limitation that is assumed is that the behavioral state (correct or faulty) of each component does not change during the diagnostic session, putting aside the problems of temporal diagnosis [5]. The question of incremental diagnosis and of the choice of the best next test point [14], [15] is also left aside : the set of observed variables is supposed to remain unchanged. In addition, the system is assumed to operate in an ideal non noisy and non disturbed environment.

.

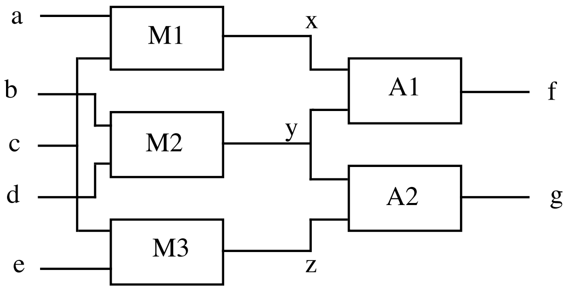


Fig. 1. The polybox system

The paper is organized as follows. Section II presents the FDI analytical redundancy approach and the DX logical approach, respectively. Section III proposes a unified framework for the two approaches. The assumptions and concepts adopted by the FDI and DX communities are outlined and the correspondence between conflicts and ARRs is exhibited. Section IV proves the equivalence of the two approaches in the no exoneration case. Finally, section V discusses the results and outlines several interesting directions for future investigation.

II. Presentation of The Two Approaches

II.1. The FDI analytical redundancy approach

II.1.1. The system model

A system is made of a set of components and a set of sensors, which provide a set of observations. The behavior model of the system expresses the constraints that link its descriptive variables. It is given by a set of relations, the formal expression of which depends on the type of knowledge (analytical, qualitative, production rules or numerical tables, etc.). It generally relies on a component-based description, which relates a set of constraints (or operators) to each component.

Example (polybox): Elementary components are the adders A1, A2 (operators +), the multipliers M1, M2, M3 (operators **.**) together with the set of sensors (identity operators adopted here for the sake of simplicity, and not represented on Figure 1).

Definition 2.1: The *system model SM* is defined as the behavioral model *BM*, i.e. the set of relations defining the system behavior, together with the observation model *OM*, i.e. the set of relations defining the observations that are performed on the system and the sensor models.

The set *V* of variables can be decomposed into the set of unknown variables *X* and the set of observed variables *O*.

Example (polybox continued):
$V = X \cup O$ where
$X = \{a, b, c, d, e, f, g, x, y, z\}$
$O = \{a_{obs}, b_{obs}, c_{obs}, d_{obs}, e_{obs}, f_{obs}, g_{obs}\}$

Behavioral Model (BM):
RM1: $x = a \cdot c$
RM2: $y = b \cdot d$
RM3: $z = c \cdot e$
RA1: $f = x + y$
RA2: $g = y + z$

Observation model (OM):
RSa: $a = a_{obs}$
RSb: $b = b_{obs}$
RSc: $c = c_{obs}$
RSd: $d = d_{obs}$
RSe: $e = e_{obs}$
RSf: $f = f_{obs}$
RSg: $g = g_{obs}$

II.1.2. The diagnosis problem

The diagnosis requirements define a set of identifiers $\{F_{op}\}$ as the set of faults *F* that may occur on an operator *op*. Without loss of generality, we assume that there is a one-to-one correspondence between components and operators (see discussion in III.3) and the set of faults is hence noted $\{F_c\}$ where *c* is a component.

Definition 2.2: The set of observations *OBS* is a set of relations of the form v_{obs} = val, where $v_{obs} \in O$ and val is in the domain of v_{obs}.

Example (polybox continued): $OBS = \{a_{obs} = 2,\ b_{obs} = 2,\ c_{obs} = 3,\ d_{obs} = 3,\ e_{obs} = 2,\ f_{obs} = 10,\ g_{obs} = 12\}$ is a set of observations.

Definition 2.3: A diagnosis problem is defined by the system model *SM*, a set of observations *OBS*, and a set of possible faults *F*.

II.1.3. Analytical redundancy relations

Definition 2.4: An *analytical redundancy relation* (ARR) is a constraint deduced from the system model which contains only observed variables, and which can therefore be evaluated from any *OBS*. It is noted $r = 0$, where r is called the *residual* of the ARR.

ARRs are used to check the consistency of the observations with respect to the system model *SM*. The ARRs are satisfied if the observed system behavior satisfies the model constraints.

ARRs can be obtained from the system model by eliminating the unknown variables[2].

Definition 2.5: For a given *OBS*, the instantiation of the residual r is noted val(r, *OBS*), abbreviated as val(r) when not ambiguous. Val(r, *OBS*) $= 0$ thus means that the observations satisfy the ARR $r = 0$.

Example (polybox continued):
Three redundancy relations are *ARR1*, *ARR2* and *ARR3*

ARR1: $r_1 = 0$ where $r_1 \equiv f_{obs} - a_{obs} \cdot c_{obs} - b_{obs} \cdot d_{obs}$
ARR2: $r_2 = 0$ where $r_2 \equiv g_{obs} - b_{obs} \cdot d_{obs} - c_{obs} \cdot e_{obs}$
ARR3: $r_3 = 0$ where $r_3 \equiv f_{obs} - g_{obs} - a_{obs} \cdot c_{obs} + c_{obs} \cdot e_{obs}$

ARR1, *ARR2* and *ARR3* are obtained from the models of M1, M2, A1; M2, M3, A2; and M1, M3, A1, A2, respectively. If we assume that the sensors are not faulty, then the ARRs can be rewritten as:

ARR1: $f - (a \cdot c + b \cdot d) = 0$
ARR2: $g - (b \cdot d + c \cdot e) = 0$
ARR3: $f - g - a \cdot c + c \cdot e = 0$

II.1.4. Signature matrix

Besides analytical redundancy relations, a fundamental concept in the FDI approach is that of *fault signature*. The theoretical signature of a fault can be viewed as the expected trace of the fault on the different ARRs, given the system model.

Definition 2.6: Given a set *SARR* of ARR_i: $r_i = 0$, with Card(*SARR*) $= n$, the (theoretical) signature of a fault F_j is given by the binary vector $FS_j = [s_{1j}, s_{2j}, \ldots, s_{nj}]^T$ in which s_{ij} is given by the following application

s: $SARR \times F \rightarrow \{0,1\}$
$(ARR_i, F_j) \rightarrow$ $s_{ij} = 1$ if the component affected by F_j is involved in ARR_i
$s_{ij} = 0$ otherwise.

The interpretation of some s_{ij} being 0 is that the occurrence of the fault F_j *does* not affect ARR_i, meaning that val(r_i) $= 0$. On the other hand, the interpretation of some s_{ij} being equal to 1 is that the occurrence of the fault F_j *is expected to* affect ARR_i, meaning that val(r_i) is now expected to be different from 0. This interpretation implicitly assumes that the occurrence of F_j is observable on the result of the ARR_i, or, equivalently, that if ARR_i is satisfied, then F_j is not present. As it will be stated later more formally, this is known as the *single fault exoneration (SF-exo) assumption*.

Definition 2.7: Given a set *SARR* of n ARRs, the signatures of a set of faults $F = \{F_1, F_2, \ldots, F_m\}$ all put together constitute the so-called *signature matrix* FS of dimension $n \times m$.

Example (polybox continued): the signature matrix for the set of single faults corresponding to components A1, A2, M1, M2 and M3, respectively, is given by:

TABLE I
POLYBOX SINGLE FAULTS SIGNATURE MATRIX

	F_{A1}	F_{A2}	F_{M1}	F_{M2}	F_{M3}
ARR1	1	0	1	1	0
ARR2	0	1	0	1	1
ARR3	1	1	1	0	1

II.1.5. Multiple faults

The case of multiple faults can be dealt with by expanding the number of columns of the signature matrix, leading to a total number of $2^m - 1$ columns if all the possible multiple faults are considered. The theoretical signature of a multiple fault is generally obtained from the signatures of single faults as explained below. Consider that F_j is a multiple fault corresponding to the simultaneous occurrence of k single faults $F_1, \ldots, F_k$, then the entries of the signature vector of F_j are given by:

$s_{ij} = 0$ if $s_{i1} = s_{i2} = \ldots = s_{ik} = 0$
$s_{ij} = 1$ otherwise, i.e. if $\exists\, l \in \{1,.., k\}$ such that $s_{il} = 1$

Example (polybox continued): the signature matrix above extended to double faults (all signatures of triple faults and above are identical to (1,1,1)) is given by:

TABLE II
POLYBOX DOUBLE FAULTS SIGNATURE MATRIX

	F_{A1}	F_{A2}	F_{M1}	F_{M2}	F_{M3}	$F_{A1 A2}$	$F_{A1 M1}$	$F_{A1 M2}$	$F_{A1 M3}$	$F_{A2 M1}$	$F_{A2 M2}$	$F_{A2 M3}$	$F_{M1 M2}$	$F_{M1 M3}$	$F_{M2 M3}$
ARR1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1
ARR2	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1
ARR3	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1

The interpretation of multiple fault signature entries is the same as for single faults. Given the way multiple fault signatures are derived from single fault signatures, this interpretation implies that the simultaneous occurrence of several faults is not expected to lead to situations in which the faults compensate, resulting in the non-observation of the

[2] The computation of a set of ARRs relies on elimination techniques which are left aside here. It is in general guided by structural analysis which can be formalized in a graph-theoretical framework (problem of finding a complete matching in a bi-partite graph).

multiple fault. As it will be stated later more formally, this is known as the *multiple fault exoneration (MF-exo) assumption*, which is a generalization of the exoneration assumption defined for single faults.

II.1.6. Diagnosis

The diagnosis sets in the FDI approach are given in terms of the faults accounted for in the signature matrix. The generation of diagnosis sets is based on a column interpretation of the signature matrix. The ARRs are instantiated with the observed values *OBS* and the associated residuals are determined, providing an *observed signature*, which can be compared with the fault theoretical signatures. This comparison is stated as a decision-making problem.

Definition 2.8: The signature of a given observation *OBS* is a binary vector $OS = [OS_1,\ldots,OS_n]^T$ where $OS_i = 0$ if and only if val(r_i, *OBS*) = *0* and $OS_i = 1$ otherwise.

The first step is to decide whether a residual value is zero or not, in the presence of noises and disturbances. This problem has been thoroughly investigated within the FDI community. It is generally stated as a statistical decision-making problem, making use of the available noise and disturbance models [4].

The second step is to actually decide which fault signatures the observed signature is consistent with. A solution to this decision-problem is to define the *consistency criterion* as follows.

Definition 2.9: An observed signature $OS = [OS_1,\ldots,OS_n]^T$ is consistent with a fault signature $FS_j = [s_{1j},\ldots,s_{nj}]^T$ if and only if $OS_i = s_{ij}$ for all *i*.

The consistency criterion adopted here has a clear semantics and is therefore appropriate for comparing the obtained diagnosis results with the ones obtained by the logical approach (cf. section 3). In practical situations (noisy environment for instance), this definition asking for a strict equality, is too demanding; it is why the FDI community generally accepts an approximate matching using a weaker *similarity-based consistency criterion* [6] (see V.3).

Definition 2.10: The *diagnosis sets* are given by the faults whose signatures are consistent with the observed signature.

Example (polybox continued): for different observed signatures, the results summarized in table III are obtained about single faults from the signature matrix of II.1.4, and about multiple faults from the extended signature matrix of II.1.5.

For the four first observed signatures, (0,0,0) and (1,1,0) have no multiple faults whereas for (0,1,1) and (1,0,1) the new double faults are supersets of single fault candidates; hence they do not need to be considered. Considering multiple faults does not bring thus more information for the four first observed signatures. This is not the case for the (1,1,1) signature for which double faults appear.

TABLE III
POLYBOX FDI DIAGNOSIS RESULTS FOR DIFFERENT OBSERVATION SIGNATURES

	OS				
ARR1	0	0	1	1	1
ARR2	0	1	0	1	1
ARR3	0	1	1	0	1
Single Fault Diagnoses	none	A2; M3	A1; M1	M2	none
Multiple Fault Diagnoses	none	(A2, M3)	(A1, M1)	none	All double faults but (A1, M1) and (A2, M3) + supersets

Another interesting point to note is that, in the polybox example, the same results are obtained for the three first observed signatures when the procedure is applied on *ARR1* and *ARR2* only:

$(OS_1,OS_2) = (0,0)$: no fault

$(OS_1,OS_2) = (0,1)$: A2 or M3 faulty

$(OS_1,OS_2) = (1,0)$: A1 or M1 faulty

In these examples, the use of ARR3, associated with r_3, does not provide any more localization power. This is obviously not the case for the two last observed signatures, for which r_3 is needed to disambiguate the signature (1,1). It can be noticed that *ARR3* was obtained from the combination of *ARR1* and *ARR2*. The contribution of this kind of additional redundancy relations and the existence of a minimal set of ARRs is discussed in V.1.

It is worth mentioning that the FDI community has developed a big amount of work for obtaining so-called *structured residuals*, which are designed so that every residual is sensitive to a subset of faults ([18], [32]). This provides a specific structure to the signature matrix. The localization power of a set of residuals can be derived from the properties of the signature matrix structure. Another approach is to design so-called *directional residuals*, which are designed so that the occurrence of a given fault gives a particular direction to the residual vector (observed signature). These methods make the choice of a set of ARRs whose signatures are more relevant than others.

II.2. The DX logical diagnosis approach

Reiter [39] proposed a logical theory of diagnosis. This theory is often referred to as diagnosis from first principles; i.e. given a description of a system together with observations of the system's behavior which conflict with the way the system is meant to behave, the problem is to determine those components of the system which, when not assumed to be operating normally, restore the consistency with the observed behavior.

This approach, also referred to as the consistency-based approach, was later extended and formalized in [13]. In the following we refer to the basic definition of [39] without considering posterior extensions and refinements.

II.2.1. The system model

The description of the behavior of the system is component-oriented and rests on first-order logic. The components are those elements subject to faults and that are part of the diagnosis of the system.

Definition 2.11: A system model is a pair (*SD*, *COMPS*) where:
1. *SD*, the *system description,* is a set of first order logic formulas with equality.
2. *COMPS*, the components of the system, is a finite set of constants.

The system description uses a distinguished predicate AB, interpreted to mean abnormal. $\neg\mathrm{AB}(c)$ with c belonging to COMPS hence describes the case where the component c is behaving correctly.

Example (polybox continued):
COMPS = {A1, A2, M1, M2, M3}
SD = {
$\mathrm{ADD}(x) \wedge \neg\mathrm{AB}(x) \Rightarrow \mathrm{Output}(x) = \mathrm{Input1}(x) + \mathrm{Input2}(x)$,
$\mathrm{MULT}(x) \wedge \neg\mathrm{AB}(x) \Rightarrow \mathrm{Output}(x) = \mathrm{Input1}(x) \,.\, \mathrm{Input2}(x)$,
ADD(A1), ADD(A2),
MULT(M1), MULT(M2), MULT(M3),
Output(M1) = Input1(A1), Output(M2) = Input2(A1),
Output(M2) = Input1(A2), Output(M3) = Input2(A2),
Input2(M1) = Input1(M3)
}

Let us note one aspect which differs somewhat from the description of the system in the FDI approach: with the distinguished predicate AB it is possible to link explicitly a physical component with the formulas describing its behavior and to make explicit the fact that the formulas describe the normal behavior of the component.
Formulas describing the behavior of the components are generally expressed by constraints and need a constraint solver to be processed. In the absence of such a constraint solver, they can be preprocessed by hand.

Example (polybox continued): the two first constraints above can be rewritten as:
$$\{\ \mathrm{ADD}(x) \wedge \neg\mathrm{AB}(x) \Rightarrow \mathrm{Output}(x) := \mathrm{Input1}(x) + \mathrm{Input2}(x),$$
$$\mathrm{ADD}(x) \wedge \neg\mathrm{AB}(x) \Rightarrow \mathrm{Input1}(x) := \mathrm{Output}(x) - \mathrm{Input2}(x),$$
$$\mathrm{ADD}(x) \wedge \neg\mathrm{AB}(x) \Rightarrow \mathrm{Input2}(x) := \mathrm{Output}(x) - \mathrm{Input1}(x),$$
$$\mathrm{MULT}(x) \wedge \neg\mathrm{AB}(x\ ) \Rightarrow \mathrm{Output}(x) := \mathrm{Input1}(x) \cdot \mathrm{Input2}(x),$$
$$\mathrm{MULT}(x) \wedge \neg\mathrm{AB}(x) \wedge \mathrm{Input2}(x) \neq 0 \Rightarrow \mathrm{Input1}(x) := \mathrm{Output}(x) / \mathrm{Input2}(x),$$
$$\mathrm{MULT}(x) \wedge \neg\mathrm{AB}(x) \wedge \mathrm{Input1}(x) \neq 0 \Rightarrow \mathrm{Input2}(x) := \mathrm{Output}(x) / \mathrm{Input1}(x)\ \}$$

II.2.2. The diagnosis problem

A diagnosis problem results from the discrepancy between the normal behavior of a system as described by the system model and a set of observations

Definition 2.12: A set of observations *OBS* is a set of first-order formulas.

Example (polybox continued): Suppose the polybox is given the inputs $a = 2$, $b = 2$, $c = 3$, $d = 3$, $e = 2$, and it outputs $f = 10$, $g = 12$ in response. The set of observations is represented by:
OBS = {Input1(M1) = 2, Input2(M1) = 3, Input1(M2) = 2, Input2(M2) = 3, Input2(M3) = 2, Output(A1) = 10, Output(A2) = 12}.

Definition 2.13: A diagnosis problem is a triple (*SD*, *COMPS*, *OBS*) where (*SD*, *COMPS*) is a system model and OBS a set of observations.
Note that this definition matches Definition 2.3 provided that each fault F corresponding to a set $\Delta \subseteq COMPS$ of components is described by:

$$\bigwedge_{c \in \Delta} \mathrm{AB}(c).$$

II.2.3. Diagnosis

A diagnosis is a conjecture that certain components of the system are behaving abnormally. This conjecture has to be consistent with what is known about the system and with the observations. Thus, a diagnosis is given by an assignment of a behavioral mode, AB or ¬AB, to each component of the system in a way consistent with the observations and the model.

Definition 2.14: A *diagnosis* for (*SD*, *COMPS*, *OBS*) is a set of components $\Delta \subseteq COMPS$ such that: $SD \cup OBS \cup \{\mathrm{AB}(c) \mid c \in \Delta\} \cup \{\neg\mathrm{AB}(c) \mid c \in COMPS - \Delta\}$ is consistent. A *minimal diagnosis* is a diagnosis Δ such that $\forall \Delta' \subset \Delta$, Δ' is not a diagnosis.

Following the principle of parsimony, minimal diagnoses are often the preferred ones.

Proposition 2.1: If every occurrence in the clausal form of *SD* ∪ *OBS* of an AB-literal is positive, the minimal diagnoses are sufficient to characterize all the diagnoses, i.e. the diagnoses are exactly the supersets of the minimal diagnoses.

The condition in proposition 2.1 is in particular satisfied when *SD* is limited to correct behavioral models expressed as necessary conditions, that is to the absence of explicit fault models, which is the case studied in this paper, and to the absence of exoneration models. Necessary conditions of correct behavior are of the form $\neg\mathrm{AB}(x) \Rightarrow \mathrm{CM}$, where CM is a formula describing the correct behavior of x. Explicit fault models are of the form $\mathrm{AB}_i(x) \Rightarrow \mathrm{FM}_i$, where FM_i is a formula describing a particular faulty behavior of x. Exoneration models express sufficient conditions of correctness of the form $\mathrm{CM} \Rightarrow \neg\mathrm{AB}(x)$, and can be generally seen as a very weak, non predictive, fault model, and are to some extent discussed in section IV.

By virtue of proposition 2.1, limiting ourselves to system descriptions made up of correct behavioral models expressed as necessary conditions means that diagnoses are characterized as supersets of minimal diagnoses. This limitation is assumed in the rest of the paper, except in IV.1.

II.2.3.1. R-conflicts

SEQA direct way of computing diagnoses based on definition 2.14 is a generate and test algorithm where subsets of components are selected, generating minimal ones first, and tested for consistency. The obvious problem is the inefficiency of this method. A method based upon the concept of conflict set has been proposed and is at the basis of most of implemented DX algorithms. This concept has been introduced by [39] and will be designated by R-conflict in this paper.

TABLE IV
POLYBOX DX DIAGNOSIS RESULTS FOR DIFFERENT OBSERVATION SIGNATURES

	OBS				
f =	12	12	10	10	10
g =	12	10	12	10	14
Minimal R-conflict	none	{A2, M2, M3}, {A1, A2, M1, M3}	{A1, M1, M2} {A1, A2, M1, M3}	{A1, M1, M2}, {A2, M2, M3}	{A1, M1, M2}, {A1, A2, M1, M3}, {A2, M2, M3}

Minimal diagnoses	{}	{A2};
		{M3};
		{A1, M2};
		{M1, M2}
			{A1};
			{M1};
			{A2, M2};
			{M2, M3}
				{M2}; {A1, A2};
				{A1, M3};
				{A2, M1};
				{M1, M3}
					{A1, A2}; {A1, M2};
					{A1, M3}; {A2, M1};
					{A2, M2}, {M1, M2};
					{M1, M3}; {M2, M3}.

Definition 2.15: An R-conflict for (SD, $COMPS$, OBS) is a set of components $C = \{c1, \ldots, ck\} \subseteq COMPS$ such that $SD \cup OBS \cup \{\neg \mathrm{AB}(c) \mid c \in C\}$ is inconsistent, i.e.: $SD \cup OBS \models \vee_{c \in C} \mathrm{AB}(c)$. A minimal R-conflict is an R-conflict, which does not strictly include (set inclusion) any R-conflict.

An R-conflict can be interpreted as follows: one at least of the components in the R-conflict is faulty in order to account for the observations; or equivalently it cannot be the case that all the components of the R-conflict behave normally. On the last expression of definition 2.15, it can be seen that an R-conflict identifies with a positive AB-clause which is an implicate of the system description and the observations.

Example (polybox continued): The polybox with the observations as seen above (f = 10, g = 12) has the following minimal R-conflicts: {A1, M1, M2} and {A1, A2, M1, M3} due to the abnormal value of 10 for f. Symmetrically, f = 12 and g = 10 yields {A2, M2, M3} and {A1, A2, M1, M3}. In the case f = 10 and g = 10, the two minimal R-conflicts are: {A1, M1, M2} and {A2, M2, M3}. In the case f = 10 and g = 14, the three minimal R-conflicts are: {A2, M2, M3}, {A1, M1, M2}, and {A1, A2, M1, M3}.

II.2.3.2. Computing minimal diagnosis using R-conflicts.

Using minimal R-conflicts, it is possible to give a characterization of minimal diagnoses, which provides a basis for computing them. By virtue of proposition 2.1 and following the hypothesis made at the end of II.3.2, minimal R-conflicts also provide a characterization of all diagnoses.

This characterization is based on the minimal hitting set definition which follows:

Definition 2.16: A *hitting set* for a collection $\mathscr{C}$ of sets is a set $H \subseteq \cup \{S \,/\, S \in \mathscr{C}\}$ such that $H \cap S \neq \{\}$ for each $S \in \mathscr{C}$. A hitting set is *minimal* if and only if no proper subset of it is a hitting set for $\mathscr{C}$.

A hitting set intersects each set of the collection. Obviously, in order to compute the minimal hitting sets of a collection $\mathscr{C}$ of sets, only those elements in $\mathscr{C}$ which are minimal have to be considered.

Proposition 2.2: Δ is a (minimal) diagnosis for (SD, $COMPS$, OBS) if and only if Δ is a (minimal) hitting set for the collection of (minimal) R-conflicts for (SD, $COMPS$, OBS).

Example (polybox continued): see Table IV.

A more general characterization of conflicts and diagnoses, available with exoneration models and with fault models, can be found in [13], allowing to get conflicts and diagnoses from prime implicates and prime implicants of the logical theory and giving then a way of computing diagnoses using a theorem prover. Our aim in this paper being to compare the basis of the FDI and DX approach in the absence of fault models, we do not consider these extensions of the theory and limit ourselves to the above definitions.

III. UNIFIED FRAMEWORK FOR THE DX AND FDI APPROACHES

This section first discusses the different ways DX and FDI formulate the diagnosis problem and links the different objects that underlie the concept of fault on each side. The notion of *potential conflict* or ARR *support* is introduced and the formal match of the two approaches is obtained, proving that a conflict can be interpreted as the support of a non satisfied ARR. The matrix framework is then proposed as suitable to strictly compare both approaches.

III.1. System model (SM) vs. system description (SD)

Both FDI and DX approaches are model-based. In FDI, the system model SM is composed of the behavior model BM and the observation model OM of the non faulty system. Behavioral laws are described in BM as constraints between variables (in general a set of ordinary differential and algebraic equations). Most works in the FDI community do not explicitly use the concept of component, and BM describes the system as a whole, using e.g. state space models. When component based models are used, topological knowledge is implicitly included as shared variables. The observation model describes which system variables are available from the sensors and the sensor models. In the simplest cases, the behavioral law of a non faulty sensor just equals some variable to the sensor output (an observed variable belonging to O): $a = a_{obs}$.

In DX, the system description SD includes explicit topological knowledge and behavioral models of components. The main difference with FDI is that the assumption of correct behavior of a component, which supports its model, is explicitly coded thanks to the AB predicate. So, if F is a formula[3] describing the correct behavior of a component c, *SM* just contains F (which implicitly means that the behavior of $\neg\text{AB}(c)$ is given by F) whereas SD explicitly contains the formula: $\neg\text{AB}(c) \Rightarrow$ F. Very often, the observation model OM is not present in DX. The equality $a = a_{obs}$ for each variable in O is thus implicitly assumed, and sensor faults are dealt with by considering sensors as components. To achieve a suitable comparison framework, further developments assume that the following property holds.

SRE Property (System Representation Equivalence): Let *SM* and *SD* respectively be a FDI and a DX model of the same system. The SRE property is true if each formula of *SM* representing (part of) a behavioral law of a component or sensor c appears in the right-hand side of an implication in *SD*, the left-hand side of which is $\neg\text{AB}(c)$ and conversely. *SM* is then simply obtained from *SD* by substituting False to all occurrences of the AB predicate.

In the following, by virtue of the SRE property, *SM* and *SD* are equally used. The restriction of SM (SD) to the behavioral law(s) of a set of components C is denoted by SM(C) (SD(C)).

III.2. FDI observations versus DX observations

In DX, the set of observations expresses as a set of first-order formulas. It is hence possible to express disjunctions of observations, which provides a powerful language. However, very often, only conjunctions of atomic formulas are used. In FDI, the observations are always conjunctions of equalities assigning a real value and/or possibly an interval value to an observed variable. In the following, to favor the comparative analysis, we do assume that we have the same observation language In both FDI and DX approaches, *OBS* is identical and made up of relations a_{obs} = v, which assign a value v to an observed variable.

III.3. FDI Faults vs. DX faults

DX adopts a component-centered modeling approach and defines a diagnosis as a set of (faulty) components. In FDI the concept of component is not in general the central one. Whereas DX abstracts the diagnosis process at the component level, FDI deepens the analysis down to variables and parameters. FDI faults hence rather correspond to the DX concept of *fault mode*. In general, several parameters can be associated with a given component, giving rise to different fault modes. The difference is that FDI faults are viewed as deviations with respect to the models of normal behavior whereas in DX's logical view the faulty behavior cannot be predicted from the normal model and the involved parameters. For deterministic models, two kinds of deviations are considered [19]:

- in the system parameters, which may take values different from the nominal ones. These are referred to as *multiplicative faults*[4].
- in known variables associated to the sensors and actuators. These are referred to as *additive faults*[4].

As a consequence, the columns of the signature matrix are generally associated with variables and parameters. The link between additive/multiplicative faults and components is hence easy to establish : sensor and actuator faults are generally modeled as additive faults whereas system component faults are modeled as multiplicative faults.

Note that, in FDI, system parameters may be physical parameters when the models are issued from physical first principles, or so called structural parameters when, typically, the model is the result of black-box identification. Structural parameters have no straightforward physical semantics. However, in some cases, it is possible to establish the (non necessarily one-to-one) correspondence with physical parameters [23]. In the two cases, the model developer must be able to make the link between parameters and physical components if the goal is fault isolation. On the other hand, linking variables to sensors and actuators is straightforward.

Conversely, the DX approach could easily account for FDI fault models by expressing the model at a finer granularity level. For instance, considering a single-input single-output (static) component c whose behavior depends on two parameters θ_1 and θ_2, the standard DX model given by:

$$\begin{array}{l} \text{COMPONENT}(x) \wedge \neg\text{AB}(x) \Rightarrow \\ \qquad\qquad\qquad \text{Output}(x) = f(\text{Input}(x), \theta_1, \theta_2) \\ \text{COMPONENT}(c) \end{array}$$

could be replaced by:

$$\begin{array}{l} \text{COMPONENT}(x) \wedge \text{PARAMETER1}(y) \wedge \text{PARAMETER2}(z) \wedge \\ \quad \neg\text{AB}(x) \wedge \neg\text{AB}(y) \wedge \neg\text{AB}(z) \Rightarrow \text{Output}(x) = f(\text{Input}(x), y, z) \\ \text{PARAMETER1}(\theta_1), \text{PARAMETER2}(\theta_2), \text{COMPONENT}(c) \end{array}$$

The component-based DX approach can hence be generalized by allowing the set *COMPS* to include not only components (including sensors and actuators), but also parameters. This framework is adopted in the following, *COMPS* standing for the set of *generalized components*, in one-to-one correspondence with FDI faults.

III.4. ARRs vs. R-conflicts

In the two approaches, diagnosis is triggered when discrepancies occur between the modeled (correct) behavior and the observations (*OBS*). As seen in section II.2, in DX, diagnoses are generated from the identification of R-conflicts,

[3] F can be assumed to be written in first-order predicate calculus, even if in practice a constraint logic programming framework is frequently used, the truth value of F being thus evaluated with respect to a given semantics of the constraints in a given domain.

[4] with reference to their influence on the state variable vector in a state space model.

where an R-conflict is a set of components the correctness of which supports a discrepancy. In the ARR framework, discrepancies come from ARRs, which are not satisfied by *OBS*.

The fundamental correspondence between ARRs and R-conflicts is now established using the following definitions and property.

Definition 3.1: The *support* of an analytical redundancy relation ARR_i is the set of components (columns of the signature matrix) with a non zero element[5] in the row corresponding to this ARR_i.

Definition 3.2: The *scope* of a component c_j is the set of ARRs (rows of the signature matrix) with a non zero element in the column corresponding to c_j.

In II.1.3, ARRs have been defined with respect to a syntactic property (observed variables), and sets of ARRs are supposed to be (in some cases, proven to be) complete, in the sense that they are sensitive to relevant faults. Note that proving this property in the general case amounts to prove a general diagnosability property of faults. We will take it as an assumption, to be proven for particular systems under consideration, and moreover make a distinction between the standard view of completeness in FDI and a view taking ARR supports into account.

ARR-d-completeness Property: A set E of ARRs is said to be d-complete if:

- E is finite;
- for any OBS, if $SM \cup OBS \models \perp$, then $\exists\ ARR_i \in E$ such that $\{ARR_i\} \cup OBS \models \perp$.

ARR-i-completeness Property: A set E of ARRs is said to be i-complete if:

- E is finite;
- for any set C of components, $C \subseteq COMPS$, and for any OBS, if $SM(C) \cup OBS \models \perp$, then $\exists\ ARR_i \in E$ such that the support of ARR_i is included in C and $\{ARR_i\} \cup OBS \models \perp$.

It will be clear from the comparison that d-completeness guarantees detectability, and i-completeness refers to isolation.

Proposition 3.1: Assuming the SRE property, let *OBS* be a set of observations for a system modeled by *SM* (or *SD*).

1) Given an analytical redundancy relation ARR_i violated by *OBS*, the support of ARR_i is an R-conflict;
2) If E is a d-complete set of ARRs, then if there exists an R-conflict for (*SD, COMPS, OBS*), there exists an analytical redundancy relation $ARR_i \in E$ violated by *OBS*;
3) If E is i-complete, then given an R-conflict C for (*SD, COMPS, OBS*), there exists an analytical redundancy relation $ARR_i \in E$ violated by *OBS* whose support is included in C.

Proof:

1) By hypothesis, $\{ARR_i\} \cup OBS \models \perp$; since, if C is the support of ARR_i, ARR_i is a consequence of SM(C), it follows that SM(C) $\cup$ $OBS \models \perp$, i.e. C is an R-conflict.
2) Suppose now that an R-conflict has been detected and that E is d-complete. Since an R-conflict exists, $SM \cup OBS \models \perp$, and d-completeness gives an $ARR_i \in E$ such that $\{ARR_i\} \cup OBS \models \perp$.
3) Last, let C be an R-conflict and suppose that E is i-complete. By definition of R-conflicts, one has SM(C) $\cup$ $OBS \models \perp$, and i-completeness gives the result.

In consequence, the support of an ARR can be defined as a *potential R-conflict* (cf. the related concept of possible conflict in [29]).

Corollary 3.1: If both the SRE property holds and the ARR-i-completeness holds, the set of minimal R-conflicts for OBS and the set of minimal supports of ARRs (taken in any i-complete set of ARRs) violated by OBS are identical.

Example (polybox continued):

The potential R-conflicts are: C_1 = {A1, M1, M2} (support of ARR1), C_2 = {A2, M2, M3} (support of *ARR2*) and C_3 = {A1, A2, M1, M3} (support of *ARR3*).

With $f = 10$ and $g = 12$, *ARR1* and *ARR3* are not satisfied, which gives rise to the minimal R-conflicts C_1 and C_3.

With $f = 10$ and $g = 10$, *ARR1* and *ARR2* are not satisfied, which gives rise to the minimal R-conflicts C_1 and C_2.

With $f = 10$ and $g = 14$, *ARR1*, *ARR2* and *ARR3* are not satisfied, which gives rise to the minimal R-conflicts C_1, C_2 and C_3.

Given *SM*, *COMPS*, *OBS*, the equivalence between really computed minimal R-conflicts for that *OBS* on the one hand and minimal supports of those really computed ARRs which are falsified by *OBS* on the other hand, depends both on the existence of a complete problem solver for DX (computation of prime implicates) and of a computable i-complete set of ARRs. Proposition and corollary 3.1 state the conditions under which a formal equivalence holds. This is a key point of the comparison between the FDI and DX approaches. Notice that corollary 3.1 was stated in [11] as proposition 4.1, omitting the condition of i-completeness. This statement was thus exact only in the cases where an i-complete set of ARRs exists. [29] suggested rightly that some conditions were needed, but gave only a sufficient condition of effective computability without any characterization and did not point out any concept similar to i-completeness. This is the case for example for linear algebraic equations, but it has not been proven in general. The completeness properties will be discussed more deeply in V.1.

III.5. The matrix framework

The FDI approach uses the signature matrix crossing ARRs in rows and sets of components in columns. It was shown in II.1 that, given an observation *OBS*, diagnosis is achieved by identifying those columns, which are identical (or closest with respect to a distance function) to the observed signature.

In the DX approach, it has been seen in II.2 that (minimal) diagnoses are obtained as (minimal) hitting sets of the collection of (*OBS*-) R-conflicts. From III.4 above, under the assumption of i-completeness, such R-conflicts can be viewed as the supports of those ARRs which are not satisfied by *OBS*, i.e. looking at the corresponding set of rows I. A

[5] It will be seen later that an extension can be done so that the elements of the FS matrix can take a value different from 1, when not equal to 0.

(minimal) hitting set of the collection of R-conflicts can thus be viewed as a (minimal) set J of singleton columns (i.e. columns corresponding to one single component) such that each of the rows of I intersects at least one column of J (i.e. has a non zero element in this column).

It is thus quite natural to adopt this matrix framework as a formal basis on which to compare the two approaches.

Let $SARR = \{ARR_i \mid i = 1\ldots n\}$ be a set, assumed to be i-complete, of ARRs and $COMPS = \{c_j \mid j = 1\ldots m\}$ be the set of components of the system. $FS = [s_{ij}]_{i = 1\ldots n,\ j = 1\ldots m}$ is the signature matrix. The j^{th} column of FS is the signature of a fault on c_j and is noted FS_j.

Definition 3.3: Any observation OBS splits the set $SARR$ into two subsets. The subset of ARRs which are violated, i.e. $\{ARR_i \equiv (r_i = 0) \mid \text{val}(r_i, OBS) \neq 0\}$, is defined as R_{false}. The subset of ARRs which are satisfied, i.e. $\{ARR_i \equiv (r_i = 0) \mid \text{val}(r_i, OBS) = 0\}$, is defined as R_{true}. Obviously, one has $R_{true} = SARR \setminus R_{false}$.

OBS is thus described through its signature OS, which is the binary column vector defined by: for all $i = 1\ldots n$, $OS_i = 1$ if $ARR_i \in R_{false}$ and $OS_i = 0$ if $ARR_i \in R_{true}$. Note that this is equivalent to: $OS_i = \text{Fa}_{\text{OBS}}(ARR_i)$, where Fa_{OBS} stands for "not satisfied" and denotes the *falsity* value of the relation ARR_i with respect to OBS.

The FDI theory compares the observed signature to the fault signatures whereas DX considers each line corresponding to an ARR in R_{false} separately, isolating R-conflicts before searching for a common explanation. In the following, these approaches are called *column view* and *line view* respectively.

III.6. Multiple faults

In the matrix framework proposed in III.5, the DX approach deals with multiple faults by implicitly considering sets of singleton columns. By default, there is no limitation on the number of possible simultaneous faults: minimal diagnoses are built as minimal hitting sets of the collection of minimal R-conflicts and are not limited in size. Single and multiple faults are thus dealt with in exactly the same framework.

In the FDI approach, as seen in II.1.5, dealing with multiple faults requires adding new columns to FS, corresponding to the considered multiple faults (a maximum of $2^{|COMPS|} - |COMPS| - 1$ if all possible multiple faults are considered). Let us call *MF property*, the constraint which has to be satisfied by the new columns.

For $J = \{j1,\ldots,jk\} \subseteq \{1,\ldots,m\}$, let us note C_J the subset $\{c_j \,/\, j \in J\}$[6], and s_{iJ} the matrix element of FS at row i and column FS_J (meaning the column added for C_J representing a multiple fault). Then, for any row i, we have:

$$s_{iJ} \neq 0 \text{ if and only if } \exists \mu\ 1 \leq \mu \leq k\ s_{ij_\mu} \neq 0 \qquad \text{(MF property)}$$

The correspondance between the DX and the FDI perspectives in the case of multiple faults can now be checked. Let the set of singleton columns $\{FS_{j_1}, \ldots, FS_{j_k}\}$ be one hitting set of a rows set I. $\{FS_{j_1}, \ldots, FS_{j_k}\}$ is viewed as a new column FS_J corresponding to $C_J = \{c_{j_1}, \ldots, c_{j_k}\}$. It results from the hitting set definition that each row of I must intersect the column FS_J if and only if it intersects at least one of the FS_{j_μ} columns. The column FS_J must have thus a non zero element in a given row i of I if and only if at least one of the FS_{j_μ} columns has a non zero element in row i, i.e. FS_J exactly verifies the MF-property.

As the extended matrix is computed for any possible set I of rows, the MF property has to hold for each row i and extended column FS_J. Consequently, the correspondence between the DX and FDI approaches is shown to be well-stated in the matrix framework.

The MF property expresses the intuitive idea that a multiple fault may affect an ARR if and only if at least one of the single faults it is made up of may affect this ARR. This means that the scope of a multiple fault is the union of the scopes of its single fault constituents.

The MF property implies an assumption on the way multiple faults manifest themselves in relation with the manifestation of single faults (for instance, no compensation or MF-exo assumption). Section IV discusses this point and shows that the MF property has to be adapted with respect with the assumptions that are made about the combination of the effects of the single faults.

IV. Comparing DX and FDI approaches assumptions and results

This section makes an intensive comparison of the DX and FDI approaches. It is shown that every approach adopts different diagnosis exoneration assumptions by default. Under the same assumptions, in particular with no exoneration at all, it is shown that the results provided by both approaches are identical and the theoretical proofs are included.

For explicitness purpose, the formulas corresponding to the different assumption cases used in the comparison are labeled as explained: C/LV: Column/Line View, S/MF: Single/Multiple Fault, (no-)exo: (no) ARR-based exoneration.

IV.1. Exoneration assumptions for the comparison

The originality and the power of both the FDI and DX approaches result from the fact that they are based only on the correct behavior of the components: no model of faulty behavior is needed. Nevertheless, different assumptions are adopted by default by each approach, leading to different computations of the diagnoses. These assumptions concern the manifestations of the faults through observations.

The DX approach makes absolutely no assumption about how a component may behave when it is faulty. This is because this approach is only based on a *reductio ad absurdum* principle: any discrepancy between the correct model and the observations necessarily implies that a component is faulty. This ensures the fundamental property of the DX approach, i.e. its logical soundness. In the matrix framework, this means that, for any given OBS, only those rows (ARRs) which are not satisfied by OBS are considered: for each one, its support constitutes the associated R-conflict. Possible diagnoses (sets of faulty components) are built from

[6] Component C_j is here straightforwardly identified to $C_{\{j\}}$.

these R-conflicts. However, the DX approach allows one to state an explicit exoneration assumption at the level of every component: assume any component, the model of which is satisfied in a given context, correct in this context. Beyond the default assumption of DX (nothing assumed about faulty behavior), this exoneration assumption is equivalent to state that the occurrence of any fault always manifests in the sense that a faulty component does not behave according to its corresponding model. This hypothesis is commonly expressed explicitly in *SD* by modeling components with biconditionals (relating the explicit correctness assumption and the functioning law). Note that, as conditions of proposition 2.1 are no more satisfied in this case, only minimal diagnoses are still characterized in terms of R-conflicts, a superset of a diagnosis being not in general a diagnosis. We do refer to this assumption as to the component-based exoneration (COMP-exo) assumption.

Definition 4.1 (COMP-exo assumption): If the correct behavioral model of a component is satisfied in a given context (given observation *OBS* and assumption of correct behavior of some given components), then this component is assumed to be correct in this context.

Conversely, the FDI approach is based on a direct reasoning about the effects of a fault (column), viewed as a non satisfaction of the correct behavioral model of the corresponding component, on the ARRs (rows). In addition to the obvious fact that a fault cannot affect an ARR which it is not in its scope, which is the direct reasoning used in DX, the idea is that a fault necessarily manifests itself by affecting the ARRs in its scope, causing them not to be satisfied by any given OBS. Hence, given *OBS*, not only, as in DX, is any component in the support of a non satisfied ARR a fault candidate, but also any component in the support of a satisfied ARR is implicitly exonerated (satisfied rows are thus also used in the reasoning). In fact this result is not sound but rests on an ARR-based exoneration (ARR-exo) assumption, which is implicitly made in the FDI approach and has to be considered explicitly in order to compare the FDI approach with the DX approach.

Definition 4.2 (ARR-exo assumption): A set of faulty components necessarily shows its faulty behavior, i.e. causes any ARR in its scope not to be satisfied by any given OBS. Or, equivalently, given *OBS*, each component of the support of a satisfied ARR is exonerated, i.e. is considered as functioning properly.

In the following, the comparison between DX and FDI approaches is made only in the case of no-exoneration at all, i.e. no COMP-exo in DX (which is the default case) and no ARR-exo in FDI (which is not the default case). The comparison of the FDI ARR-exo assumption and the DX COMP-exo assumption has been made, relying on the concept of alibi [30] but is out of the scope of this paper and will be published apart

IV.2. The no-exoneration case

In this subsection, under the SRE property, the no-exoneration case is now given a formal account in the matrix framework previously introduced, in order to specify formally which (sets of) components have to be considered as diagnoses in each case.

From the matrix viewpoint, the fact that ARR_i, if satisfied by *OBS*, exonerates c_j appears (cf. II.1.4) in FS as $s_{ij} = 1$. In order to release the default ARR-exo assumption in the FDI approach, it is necessary to express that a faulty component may or may not affect the ARRs in its scope. To make the difference with the previous case, the symbol X can be used instead of 1 for this purpose. We can now represent the fact that c_j belongs to the support of ARR_i but is not necessarily exonerated when $ARRi$ is satisfied by *OBS*, by s_{ij} = X. The semantics of s_{ij} = X is thus: a fault in c_j can explain why ARR_i is not satisfied by *OBS*, but ARR_i may happen to be satisfied by *OBS* even when c_j is faulty (to be compared with the semantics of $s_{ij} = 1$: a fault in c_j implies that ARR_i cannot be satisfied by any *OBS*).

The generalized use of an exoneration assumption for each component of the support of each ARR is called the *exoneration case* and corresponds to the assumption by default of the FDI approach (elements of FS take their values in {0, 1}). As said above, in the present comparison, we consider only the total lack of exoneration, called the *no-exoneration case* (elements of FS take their values in {0,X}). In this later case, definitions 3.1 and 3.2 translate to: the *support* of an ARR_i is the set $\{c_j \mid s_{ij} = \text{X}\}$; the *scope* of a component c_j is the set $\{ARR_i \mid s_{ij} = \text{X}\}$.

IV.2.1. The single fault no-exoneration case (SF-no-exo case)

The column associated with the faulty component must have X in non satisfied rows and 0 or X in satisfied rows. In this column view, the matching of the observed signature with a fault signature is thus based on the fact that an X in the fault signature is consistent with either a 0 or a 1 in the observed signature. So, it is just like using only non satisfied rows: the faulty component must have X in each such row.

So acceptable diagnoses are those $\{c_j\}$ verifying:

$$R_{false} \subseteq \text{Scope}(c_j) \qquad \text{(CV-SF-no-exo)}$$

In the line view, $\{c_j\}$ is an acceptable diagnosis if it hits all the supports of not satisfied ARRs, that is to say:

$$\forall i \ (ARR_i \in R_{false} \Rightarrow c_j \in \text{Support}(ARR_i)) \quad \text{(LV-SF-no-exo)}$$

(LV-SF-no-exo) and (CV-SF-no-exo) are straightforwardly equivalent, because each one is equivalent to: $\forall i$ ($\text{Fa}_{\text{OBS}}(ARR_i)$ $= 1 \Rightarrow s_{ij} = \text{X}$).

We have thus the result:

Theorem 4.1: Under the assumption of i-completeness, FDI single fault diagnoses in the ARR-no-exoneration case are identical to non empty DX single fault diagnoses.

Example (polybox continued) Releasing the exoneration assumption in the polybox example leads to the following single fault signature matrix:

TABLE V
POLYBOX SINGLE FAULTS SIGNATURES WITHOUT EXONERATION

	F_{A1}	F_{A2}	F_{M1}	F_{M2}	F_{M3}
ARR1	X	0	X	X	0
ARR2	0	X	0	X	X
ARR3	X	X	X	0	X

The following results are then obtained:

With outputs $f = 10$ and $g = 12$, i.e. observed signature (1,0,1), there are 2 single fault diagnoses {A1} and {M1}.

With outputs $f = 10$ and $g = 10$, i.e. observed signature (1,1,0), there is only one single fault diagnosis {M2}.

With outputs $f = 10$ and $g = 14$, i.e. observed signature (1,1,1), there is no single fault diagnosis.

With outputs $f = 12$ and $g = 12$, i.e. observed signature (0,0,0), there are 5 single fault diagnoses.

These results obtained by FDI are identical to those obtained by DX (cf. II.2.3.2).

Let us remark also that, except in the case of normal observation (null observed signature), these results are the same as under the default exo assumption (cf. II.1.6). This is because, as each one of the ARRs can be derived from the other two, the observed signatures (1,0,0), (0,1,0) and (0,0,1) are physically impossible. But this would not be the case in general. For instance, it is not the case here for the normal observation $f = 12$, $g = 12$, i.e. observed signature (0,0,0): in the exo case (cf. II.1.6), no single fault diagnosis exists, when in the no-exo case, five single-fault diagnoses corresponding to the five components are proposed.

IV.2.2. The multiple fault no-exoneration case (MF-no-exo case)

In this case, (CV-SF-no-exo) can be straightforwardly extended to: C_J is a possible diagnosis iff

$$R_{false} \subseteq \text{Scope}(C_J) \qquad \text{(CV-MF-no-exo)}$$

No COMP-exo and multiple faults is the default case in DX. The way the line view selects a set of column vectors (cf III.6) to build the equivalent extended matrix column interprets as follows: a multiple fault can explain that a given ARR is not satisfied if and only if at least one of its faults can explain it, i.e. several faults never produce more / less than the combination of their separate effects. On the other hand, it is admitted that a faulty component does not necessarily affect an ARR in its scope (no-exo) and that several faults may compensate each other (compensation), resulting in a satisfied ARR.

With the help of the ordering 0<X, the no-exoneration fault interaction law can be stated very simply:

$$s_{iJ} = \sup_{j \in J} \{s_{ij}\} \qquad \text{(MF-no-exo)}$$

Thus in the line view the diagnoses are the sets C_J such that:

$$\forall i\ (ARR_i \in R_{false} \Rightarrow \exists j \in J,\ C_j \in \text{Support}(ARR_i)) \quad \text{(LV-MF-no-exo)}$$

This, due to (MF-no-exo), translates to:

$$\forall i\ (ARR_i \in R_{false} \Rightarrow C_J \in \text{Support}(ARR_i))$$

that in turn is the same as $R_{false} \subseteq \text{Scope}(C_J)$, i.e. (CV-MF-no-exo).

Theorem 4.2: Under the assumption of i-completeness, FDI diagnoses in the ARR no-exoneration case are identical to non empty DX diagnoses.

Example (polybox continued): For the polybox example, the following extended signature matrix is obtained from the usual one (see II.1.5) by replacing each 1 by X (all signatures of at least triple faults are identical to (X,X,X)):

TABLE VI
POLYBOX EXTENDED SIGNATURE MATRIX WITHOUT EXONERATION

	F_{A1}	F_{A2}	F_{M1}	F_{M2}	F_{M3}	$F_{\substack{A1\\A2}}$	$F_{\substack{A1\\M1}}$	$F_{\substack{A1\\M2}}$	$F_{\substack{A1\\M3}}$	$F_{\substack{A2\\M1}}$	$F_{\substack{A2\\M2}}$	$F_{\substack{A2\\M3}}$	$F_{\substack{M1\\M2}}$	$F_{\substack{M1\\M3}}$	$F_{\substack{M2\\M3}}$
ARR1	X	0	X	X	0	X	X	X	X	X	X	0	X	X	X
ARR2	0	X	0	X	X	X	0	X	X	X	X	X	X	X	X
ARR3	X	X	X	0	X	X	X	X	X	X	X	X	X	X	X

The following results are then obtained:

With outputs $f = 10$ and $g = 12$, i.e. observed signature (1,0,1), there are 4 minimal diagnoses: the 2 single fault diagnoses {A1} and {M1} and the 2 double fault diagnoses {A2, M2} and {M2, M3}, and 22 superset diagnoses.

With outputs $f = 10$ and $g = 10$, i.e. observed signature (1,1,0), there are 5 minimal diagnoses: the single fault diagnosis {M2} and the 4 double fault diagnoses {A1, A2}, {A1, M3}, {A2, M1} and {M1, M3}, and 20 superset diagnoses.

With outputs $f = 10$ and $g = 14$, i.e. observed signature (1,1,1), there are 8 minimal double fault diagnoses: {A1, A2}, {A1, M2}, {A1, M3}, {A2, M1}, {A2, M2}, {M1, M2}, {M1, M3} and {M2, M3}, and 16 superset diagnoses.

These results obtained by FDI are identical to those obtained by DX (cf. II.2.3.2). In the case where $f = 12$ and $g = 12$, i.e. observed signature (0,0,0), any non empty subset of components is a diagnosis: there are 5 minimal single fault diagnoses and 26 superset diagnoses. The only difference between FDI and DX is that, the "no-fault" column of signature (0,0,0) is left implicit in the signature matrix.

It can be noticed that, except in the $f = 10$ and $g = 14$ case (where anyhow, no exoneration can apply as no ARR is satisfied), these results are different from those obtained under the default exo assumption (II.1.6).

V. BENEFITS AND PERSPECTIVES ARISING FROM THE UNIFIED FRAMEWORK

V.1. The SRE and ARR-completeness properties

The SRE property is required to perform a sound comparison. Indeed, it imposes that the models *SM* and *SD* are isomorphic both from a semantic and a syntactic point of view.

The ARR-d-completeness property (cf. the definition in III.4) is a minimum requirement.

Let $M(\boldsymbol{x}, \boldsymbol{o})$ be the equation set that represents *SM*, where $\boldsymbol{x}$ and $\boldsymbol{o}$ denote the vectors of variables contained in X and O respectively. Elimination theory allows one to go from $M(\boldsymbol{x}, \boldsymbol{o})$ to a set of ARRs, $E(\boldsymbol{o})$. When equivalence is preserved, $\forall \boldsymbol{o}$ ($\exists \boldsymbol{x}\, M(\boldsymbol{x}, \boldsymbol{o}) \Leftrightarrow E(\boldsymbol{o})$), d-completeness is satisfied.

Notice that in [24] such an equivalence is included in the definition of an ARR, i.e. only d-complete ARRs are considered.

An old result of algebraic geometry [21] states that the equivalence holds for *polynomial algebraic* equations. This result only ensures existence and is not constructive, i.e. cannot be used in practice to build *E*. Now, recent computer algebra techniques, such as *Gröbner bases, Ritt's algorithm* are a step in this direction [34].

The i-completeness property is a novel concept since it requires to take into account the ARRs' supports. This is not common in the FDI community. The problem is related to the fact that having a basis of ARRs does not guarantee that all the potential minimal R-conflicts are represented by the ARRs' supports.

Example (polybox continued): The polybox example, presented in II.1.3, illustrates the above issues. {*ARR1*, *ARR2*} is ARR-d-complete but not ARR-i-complete. Indeed, let us consider C = {A1, A2, M1, M3} and $OBS = \{a = 2, b = 2, c = 3, d = 3, e = 2, f = 10, g = 12\}$, then SM($C$) ∪ OBS $\models \perp$, but neither *ARR1* or *ARR2* have a support included in C. It is only when adding *ARR3*, that can be obtained by combining *ARR1* and *ARR2*, and whose support is {A1, A2, M1, M3} that ARR-i-completeness is obtained.

The ARR-i-completeness problem is even thornier, as illustrated by the following example:

Example (the inverted polybox):

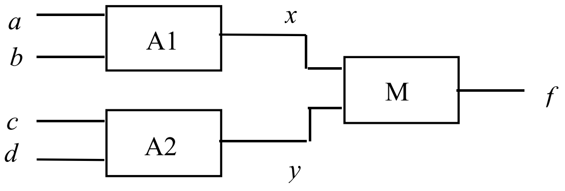


Fig. 2. The inverted polybox

Here *COMPS* = {A1, A2, M}, where A1 and A2 are adders and M is a multiplier, with models as in section II. We assume that $O = \{a, b, c, d, f\}$ and $X = \{x, y\}$.

The unique ARR is given by:

ARR1: $f - (a + b) \cdot (c + d) = 0$, with support {A1, A2, M}.

Let us consider the following observations: $OBS = \{a = -1, b = c = d = f = 1\}$ and C={A1, M}, then SM(C) ∪ OBS $\models \perp$, indeed $x = 0$ due to SM(A1) and $f = 0$ due to SM(M) and the absorbant property of 0 for multiplication. However, the support of *ARR1* is not included in C. This proves the non ARR-i-completeness.

Notice that the DX approach captures the {A1, M} R-conflict for *OBS* because:

$SD \cup \{\neg\mathrm{AB(A1)}\} \cup \{a = -1, b = 1\} \models x = 0$

$SD \cup \{\neg\mathrm{AB(M)}\} \cup \{x = 0\} \models f = 0$

and this conflict does not appeal to the behavior of A2. Thus DX single fault diagnoses are {A1} and {M}, different from FDI ones which are {A1}, {A2} and {M} and come from the violation of *ARR1* by OBS.

Since ARR-i-completeness is not satisfied, theorem 4.1. does not apply which explains that FDI and DX diagnoses are different.

The problem arises when particular values of some variables, appearing as input in a component's model SM(C), determine the component output independently of remaining inputs.

The ARR-i-completeness issue is naturally linked to the redundancy and minimality issues. It is known in DX that only minimal (for subset inclusion) R-conflicts are relevant, the non minimal ones being redundant. On the other hand, it is common in FDI to derive additional ARRs by combination. Although combined ARRs are redundant when considered just as equations, they must be considered jointly with their associated support to decide whether they are needed to obtain i-completeness or can just be ignored. The following proposition states under which conditions a combined ARR is redundant with respect to a set of ARR_is:

Proposition 5.1: The necessary and sufficient condition for a given ARR_j to be redundant with respect to a set of ARR_is, $i \in I, j \notin I$, is: $\exists\, I' \subseteq I$ such that

1) for any observation *OBS*, if all ARR_is, $i \in I'$, are satisfied by *OBS*, then ARR_j is satisfied by *OBS* (or, equivalently, if ARR_j is not satisfied by *OBS*, necessarily at least one of the ARR_is is not satisfied by *OBS*): $\bigwedge_{i \in I'} \mathrm{ARR}_i[OBS] \Rightarrow \mathrm{ARR}_j[OBS]$ is valid.

2) the support of ARR_j contains the support of each ARR_i, $i \in I'$:

Supp(ARR_j) $\supseteq \cup_{i \in I'}$ Supp(ARR_i).

ARR[*OBS*] designates the ground formula obtained from *ARR* by substituting each observed variable by its value in *OBS*: if $OBS = \{X_j = \mathrm{v}_j\}$ then ARR[OBS] = ARR[X_j/v_j].

The proof of proposition 5.1 is not given due to space limitations.

V.2. Off-line vs. on-line computation of R-conflicts

From the computational point of view, the main difference between the FDI and DX approaches is that in FDI most of the computational work is done off-line. Using just the knowledge of which variables are observed, i.e. sensor locations, modeling knowledge is compiled: ARRs are obtained by combining model equations or constraints, and eliminating unobserved variables. The only thing that has to be done on-line, i.e. when a given *OBS* is acquired, is to compute the truth value (with respect to *OBS*) of each ARR and to compare the obtained observed signature with the fault theoretical signatures (columns of the signature matrix). In terms of R-conflicts, this means that potential R-conflicts are compiled and that, for a given *OBS*, R-conflicts are exactly those potential R-conflicts which are supports of those ARRs which are not satisfied by *OBS*.

Conversely, in DX, the computational task starts as soon as OBS is known, nothing being compiled off-line.

The idea, coming from FDI, of compiling ARRs can be used as so in the DX framework for obtaining potential R-conflicts. This has indeed already been proposed in the DX framework: [25] proposes to compute in advance all possible linear combinations of models eliminating all occurrences of unobserved variables, i.e. all possible ARRs, for the monostable circuit, an analog electronic circuit. When this is possible, this is a way to get the best from each approach:

- modeling knowledge is compiled (under ARRs form) according to sensor locations before any observation has been made, which is the main advantage of the FDI approach;
- thanks to explicit correctness assumptions, potential R-conflicts (supports of ARRs) are computed at the same time to give rise, given an *OBS*, to R-conflicts;
- R-conflicts are used to generate the diagnoses, which is the main advantage of the DX approach.

V.3. Logical soundness, decision and robustness

As seen in section 4, the DX logical diagnosis theory does not make any kind of assumption about the faults *a priori*, which guarantees logically sound results. In the most general case, single as well as multiple faults are considered: a fault may be observable or not at the symptom level and multiple faults may as well compensate, i.e. being themselves not observable. When the application domain suggests specific assumptions, these are explicitly stated as additional axioms, for example the exoneration assumptions as defined in section IV.1

Conversely, the FDI approach implicitly adopts assumptions to restrict the number of diagnosis candidates, e.g. exoneration and single fault assumptions. These assumptions are justified in statistical terms.

It has been mentioned in II.1.6 that, when the observed signature fits no fault signature, some FDI applications accept the closest fault signatures using a similarity-based consistency criterion, e.g. with respect to some distance. The reason for accepting an approximate matching is that it is a way to cope with model uncertainties, e.g. unknown disturbances or model errors. Another issue is to guarantee some kind of robustness in the decision procedure which assesses whether a residual is zero or not. Viewing this operation as hypothesizing a whole set of possible observed signatures, the formal proposed framework relating observed and fault signatures still holds.

Another way to deal with uncertainties in FDI is to make use of as many ARRs as can be derived, even though these may be redundant from a detection and localization point of view. It can be argued that additional signature bits ensure more robust detection in the presence of noise and disturbances. Although a definition of *logically redundant* ARRs is provided in section V.1, the redundancy properties of ARRs in noisy environments must be stated in statistical terms and are not studied in this paper.

The robustness issue arises from the type of models being used, which are essentially numeric with uncertainties represented either by unknown disturbances or by stochastically characterized signals. There are two families of methods: those which act at the residual generation step (unknown input observers [2], disturbance optimal decoupling [8]) and those which act at the residual interpretation step (statistical decision methods [5], fuzzy interpretation [6]). These methods have no equivalent in DX.

DX manages uncertainty by focusing on the use of high level of abstraction models, which are qualitative or symbolic. Also widely used in DX, interval models (also known as semi-qualitative models) are based on the assumption that uncertainties are bounded [3], [25]. These have been investigated for several years in the DX community as realizing a perfect compromise between precision and robustness; more recently, interval models have been considered in pure FDI approaches [1], [28].

VI. Conclusion

The first goal of FDI was historically fault detection and associated decision procedures. Its main interest was to offer sophisticated techniques, such as observers and filters, so as to interpret observations to produce a set of symptoms (residuals). Nevertheless, the residuals can be designed in such a way that they are also informative from the fault localization point of view. DX approached the diagnosis problem the other way around, focusing on fault localization by pointing out the subsets of the system description that conflict with the observations. Our study proves that a significant part of the two theories fits into a common framework which allows a precise comparison. When they adopt the same hypotheses with respect to how faults manifest themselves and how many faults can occur simultaneously, FDI and DX views agree on diagnoses. This opens the possibility of a fruitful cooperation between these two diagnostic approaches.

Some points have been left out of this comparison. There is presently no equivalent in DX of the notion of unknown disturbance or noise. Conversely, DX makes a systematic use of fault models, whose counterpart in FDI can be found in assumptions about the additive or multiplicative disturbances that model the faults but always with respect to a correct behavior model. Fault models have been left out of the framework of the present paper. Temporal aspects of diagnosis, which are crucial in the state tracking problem, have not been approached neither. Further studies are needed to integrate these aspects, which would be beneficial to both communities.

VII. References

[1] O. Adrot, D. Maquin and J. Ragot "Fault detection with model parameter structured uncertainties", *European Control Conference ECC'99,* Karlsruhe, Germany, CD-ROM BA.5, F210.pdf, 1999.

[2] E. Alcorta-Garcia, P.M. Frank "Deterministic non-linear observer-based approaches to fault diagnosis: a survey", *Control Engineering Practice* 5(5), p. 663-670, 1997.

[3] J. Armengol, J. Vehi, L. Travé-Massuyès and .M.A. Sainz "Application of multiple sliding time windows to fault detection based on interval

models", *12th International Workshop on Principles of Diagnosis DX'01,* Via Lattea, Italy, p. 9-16, 2001.

[4] M. Basseville, I. Nikiforov "*Detection of abrupt changes – Theory and Applications"*, Information and System Sciences Serie, Prentice Hall, Englewood Cliffs, 1993.

[5] V. Brusoni, L. Console, P. Terenziani and D. Theseider Dupré "A spectrum of definitions for temporal model-based diagnosis", *Artificial Intelligence* 102(1), p. 39-79, 1998.

[6] J.-Ph. Cassar , M. Staroswiecki "Advanced Design of the Decision Procedure in Failure Detection and Isolation Systems, *IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS'94*, Espoo, Finland, p. 380-385, 1994.

[7] S. Cauvin, M.O. Cordier, C. Dousson, P. Laborie, F. Lévy, J. Montmain M. Porcheron, I. Servet and L. Travé-Massuyès "Monitoring and alarm interpretation in industrial environments", *AI Communications*, p. 139-173, 1998.

[8] J. Chen, R.J. Patton and H.Y. Zhang "A multi-criteria optimization approach to the design of robust fault detection algorithms", *International Conference on Fault Diagnosis Tooldiag'93*, Toulouse, France, 1993.

[9] CEP "Control Engineering Practice", *Special volume on Supervision, fault detection, and diagnosis of technical systems*, Vol. 5(5), 1997.

[10] M.O. Cordier, P. Dague, M. Dumas, F. Lévy, J. Montmain, M. Staroswiecki and L. Travé-Massuyès "AI and Automatic control approaches of model-based diagnosis: links and underlying hypotheses", 4th *IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS 2000,* Budapest, Hungary, p. 274-279, 2000.

[11] M.O. Cordier, P. Dague, M. Dumas, F. Lévy, J. Montmain, M. Staroswiecki and L. Travé-Massuyès "A comparative analysis of AI and control theory approaches to model-based diagnosis", *14th European Conference on Artificial Intelligence ECAI'00*, Berlin, 2000, p. 136-140.

[12] R. Davis "Diagnostic Reasoning based on structure and behavior", *Artificial Intelligence* 24, p. 347-410, 1984.

[13] J. De Kleer, A. Mackworth and R. Reiter "Characterizing diagnoses and systems", *Artificial Intelligence* 56(2-3), p. 197-222, 1992.

[14] J. De Kleer, O. Raiman and M. Shirley "One step lookahead is pretty good", *2nd International Workshop on Principles of Diagnosis DX'91*, Milan, p. 136-142. Also in *Readings in Model-Based Diagnosis*, Hamscher W., Console L., de Kleer J. (eds.), Morgan Kaufmann, 1992, p. 138-142, 1991.

[15] J. De Kleer, B.C. Williams "Diagnosing multiple faults", *Artificial Intelligence* 32(1), p. 97-130, 1987.

[17] P.M. Frank "Analytical and qualitative model-based fault diagnosis – A survey and some new results", *European Journal of Control,* Vol. 2, p. 6-28, 1996.

[16] *Pattern Recognition*, Second Edition by Sergios Theodoridis, Konstantinos Koutroumbas Publisher: Academic Press; 2nd edition February 2003, ISBN: 0126858756.

[18] J.J. Gertler "Analytical redundancy methods in fault detection and isolation", IFAC Fault Detection, Supervision and Safety for Technical Processes, pages 9-21, Baden-Baden, Germany, 1991.

[19] J. Gertler "*Fault detection and diagnosis in engineering systems"*, Marcel Dekker Inc., 1998.

[20] W. Hamscher, L. Console and J. de Kleer (eds.) "*Readings in Model-Based Diagnosi"s*, Morgan Kaufmann, San Mateo, CA, 1992.

[21] W.V.D. Hodge, D. Pedoe "*Methods of Algebraic Geometry"*, Cambridge University Press, 1952.

[22] R. Isermann "Supervision, fault detection and fault-diagnosis methods – An introduction", *Control Engineering Practice*, Vol. 5(5), p. 639-652, 1997.

[23] R. Isermann, "Process fault diagnosis based on process knowledge", *IFAC-AIPAC'89, Advanced Information Processing in Automatic Control*, volume II, pages 23-27, Nancy, 1989.

[24] M. Krysander, M. Nyberg "Structural analysis utilizing MSS sets with application to a paper plant", *13th International Workshop on Principles of Diagnosis DX'02,* Semmering, Austria, p. 51-57, 2002.

[25] E. Loiez, P. Taillibert "Polynomial Temporal Band Sequences for Analog Diagnosis", *International Joint Conference on Artificial Intelligence (IJCAI 97)*, Nagoya, Japon, 23-29 aug. 1997.

[26] R.J. Patton, J. Chen "A review of parity space approaches to fault diagnosis", 4th *IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS'91*, Baden-Baden, Germany, p. 239-255, 1991.

[27] Y. Peng, J. Reggia "*Abductive inference models for diagnostic problem solving"*, Springer-Verlag, 1990.

[28] S. Ploix, O. Adrot and J. Ragot "Bounding approach to the diagnosis of a class of uncertain static systems", *4th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS'2000,* Budapest, Hungary, p. 149-154, 2000.

[29] B. Pulido, C. Alonso "Possible conflicts, ARRs, and conflicts", *13th International Workshop on Principles of Diagnosis DX'02,* Semmering, Austria, p. 122-128, 2002.

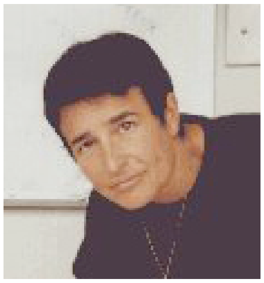
[30] O. Raiman "The alibi principle, *Readings in Model-Based Diagnosis"*, Hamscher W., Console L., de Kleer J. (eds.), Morgan Kaufmann, San Mateo, CA, p. 66-70, 1992.

[39] R. Reiter "A theory of diagnosis from first principles", *Artificial Intelligence* 32(1), p. 57-96, 1987.

[32] M. Staroswiecki, J.P. Cassar and V. Cocquempot "Generation of optimal structured residuals in the parity space", *12th IFAC Word Congress*, Sydney, Australia, Vol. 5, p. 535-542, 1993.

[33] M. Staroswiecki "Quantitative and qualitative models for faults detection and isolation, International Journal of Mechanical Systems and Signal Processing, 14, 3, p. 301-325, 2000.

[34] M. Staroswiecki, G. Comtet-Varga "Analytical redundancy relations for fault detection and isolation in algebraic dynamic systems", *Automatic*a, 37(5), p. 687-699, 2001.



Marie-Odile Cordier was born in Paris (France) in 1950. She received a Ph.D. degree in Computer Science in 1979 and an « Habilitation à Diriger des Recherches » in 1996, both from the University of Paris Sud/Orsay, France.

She was Associated Professor at University of Paris Sud/Orsay, France, from 1973 and became full Professor at University of Rennes in 1988, performing her research activity at IRISA-INRIA. She is currently the scientific leader of the DREAM Team (Diagnostic, Reasoning and Modeling). Her main research interests are in Artificial Intelligence, focusing in Model-Based Diagnosis, on-line monitoring, model acquisition, using model checking techniques and Inductive Logic Programming, and temporal abductive reasoning. She has been responsible for several industrial projects, published numerous papers in international conference proceedings and scientific journals, and served as Program Committee member and area chair of several conferences.

Prof. Cordier current responsibilities include co-leader of the French Imalaia Group and member of the European Network of Excellence *MONET*, both studying the links between AI and Control Theory methods in the field of monitoring and diagnosis. She is an ECCAI fellow since 2001. Her e-mail address is cordier@irisa.fr.

Philippe Dague received a "DEA" in Mathematics from University Paris 7 in 1971, in Theoretical Physics from University Paris 6 in 1972, and in Computer Science from university Paris 6 in 1983 ; Engineering degree from "Ecole Centrale de Paris" in 1972. He received his PhD in Theoretical Physics in 1976 and the "Habilitation à Diriger des Recherches" in Computer

Sciences, all from University Paris 6. He was assistant in Mathematics starting at University of Poitiers, then at University Paris 6 from 1976 to 1983 and research engineer in Computer Science at IBM Paris Scientific Center from 1983 to 1992. He is currently professor at the University Paris 13 since 1992, working at Galilée Institute from 1998. From 1992, he is member of LIPN-UMR 7030, the laboratory of Computer Science of the University Paris 13 and Mixed Research Unit of CNRS, responsible of the ADAge (Symbolic and Numeric Machine Learning, Diagnosis and Agents) group. His research interests in AI are Model-Based Diagnosis and Qualitative Reasoning, and active in establishing a bridge between FDI and AI MBD communities. He has been involved in many collaborations and industrial projects, at the National and the European level and is author of about 60 papers in international or national conferences and journals. His e-mail address is Philippe.Dague@lipn.univ-paris13.fr.

François Lévy received a French 'Agrégation' in math's, Phd and 'Habilitation' in computer science at Paris North University.

He has been a math teacher during18 years, Assistant Professor in 1989 and Full Professor since 1993 at the University of Paris North. His research interests are in Logic (default logic, temporal reasoning), Diagnosis (monitoring of telecommunication networks), Cognitive Modeling (Natural Language Understanding, causality). Some of his relevant publications are

F. Lévy, “A Survey of Belief Revision and Updating in Classical Logic” in : Revision and Updating in Knowledge Bases, Lea Sombe ed., (J. WIley & sons, 1994), pp 29-59 ; F. Lévy, J. Joachim Quantz, "Representing Beliefs in a Situated Event Calculus", Ecai 98, Brighton, 25-28 Aug 1998 pp 537-541 ; N. Chaignaud, F. Lévy : "Common Sense Reasonning: Experiments and Implementation" Ecai 96, Budapest, 14-16 Aug 1996 pp 604-608.

Jacky Montmain was born in Lyon, France. He graduated from the Ecole Nationale Supérieure d’Ingénieurs Electroniciens de grenobe in 1987, and received his Doctorate in 1992 from National Polytechnic Institute in Grenoble, both in control theory.

He is a research engineer at the French Atomic Commission, temporarily attached to the Ecole des Mines in Alès. His research interests include process control and supervision, fuzzy approaches to decision-making, fault detection, and model-based diagnosis. Recent publication include J. Montmain, S. Gentil, Dynamic causal model diagnostic reasoning for online technical process supervision, Automatica 36, (2000) 1137-1152 ; J. Montmain: Supervision Applied to Nuclear Fuel Reprocessing. AI Commun. 13(2): 61-82 (2000).

Marcel Staroswiecki was born in Melitopol (Ukraina) in 1945. He obtained the Engineering Degree from the Ecole Nationale Supérieure d’Ingéniers des Arts et Métiers (silver medal), in 1968. He then obtained a PhD in Automatic Control in 1970, and the DSc in Physical Sciences in 1979, both at University of Lille, France.

He joined the University of Lille as an Assistant Professor in 1969 and he became a full Professor in 1983. He is currently teaching Automatic Control at Ecole Polytechnique Universitaire de Lille. He has been heading the Laboratoire d’Automatique et d’Informatique Industrielle de Lille (LAIL-CNRS), the French national network on Fault Detection and Isolation, and was a Chargé de mission at the French ministery of research. He is involved in several European projects and networks, and acts as an expert for the European Commission and the French ministery of Research. He was the IPC chair of the 2003 IFAC Safeprocess Symposium in Washington DC. He is currently working on Fault Detection and isolation, and on Fault Tolerant Control, and he co-authored two books : *Actionneurs intelligents* (smart actuators), Dunod, Paris, 1994 and *Diagnosis and Fault Tolerant Control*, Springer-Verlag, Berlin Heidelberg, 2003.

Prof. Staroswiecki is member of two IFAC Technical Committes, Intelligent Components and Instruments, and Safeprocess.

Louise Travé-Massuyès received a Ph.D. degree in control in 1984 and an Engineering Degree specialized in control, electronics and computer science in 1982, both from the *Institut National des Sciences Appliquées (INSA)* in Toulouse, France; Award from the *Union des Groupements d'Ingenieurs de la Region Midi-Pyrénées*. She received an « Habilitation à Diriger des Recherches » from Paul Sabatier University in 1998.

She is currently Research Director of *Centre National de Recherche Scientifique (CNRS),* working at *LAAS*, Toulouse, France, in which she is the scientific leader the “Qualitative Diagnosis, Supervision and Control” Group for several years. Her main research interests are in Qualitative and Model-Based Reasoning and applications to dynamic systems Supervision and Diagnosis. She has been particularly active in bridging the AI and Control Engineering Model-Based Diagnosis communities, as leader of the BRIDGE Task Group of the MONET European Network. She has been responsible from several industrial and european projects and published more than 100 papers in international conference proceedings and scientific journals.

Dr. Travé-Massuyès current responsibilities include; member of the *IFAC Safeprocess* Technical Committee; member of the European Network of Excellence *MONET* Steering Committee; member of the French CNRS Network *RTP 20* on "Diagnosis, Reliability and Safety" Steering Committee; co-leader of the French Imalaia Group. She is a Senior Member of the *IEEE* Computer Society. Her e-mail address is louise@laas.fr. HYPERLINK