8th IFAC Symposium on Fault Detection,
Supervision and Safety of Technical Processes (SAFEPROCESS)
August 29-31, 2012. Mexico City, Mexico

IFAC

Meta-Diagnosis in FDI: Reasoning About False Analytical Redundancy Relations

Nuno Belard *,**,*** **Yannick Pencolé** **,***
Michel Combacau **,***

* *Airbus France; 316 route de Bayonne; 31060 Toulouse, France.*
** *LAAS-CNRS; 7 Av. du Colonel Roche; F-31077 Toulouse, France.*
*** *Université de Toulouse; UPS, INSA, INP, ISAE; LAAS; 7 Av. du Colonel Roche, F-31077 Toulouse, France.*
{ *nuno.belard@airbus.com* *ypencole@laas.fr* *combacau@laas.fr* }

Abstract: Fault Detection and Isolation (FDI) analytical-redundancy-based systems rely on a model of a real-world system and on some observations obtained from sensor readings to determine what faults are present in that same system at a given time. In this framework, it is sometimes assumed that the models used are a true representation of the artefact under study. Unfortunately, in real-world applications this is not always the case and errors in models may entail false diagnoses with huge economic consequences. Call the problem of detecting and identifying faults in models a problem of meta-diagnosis; an unsolved issue in the FDI community and a very difficult problem to address especially in the case of complex systems. In this paper, we contribute by providing this community with a method of meta-diagnosis making use of the link between the FDI analytical redundancy approach and the DX consistency-based logical approach; and illustrate such contribution with a DC motor example. Finally, the meta-diagnosis is generalised for detecting and identifying errors in observations and algorithms.

Keywords: Maintenance, Meta-diagnosis, FDI and DX approaches, Complex systems

1. INTRODUCTION

Just like nuclear power plants or satellites, aircraft are complex systems with tens of interacting subsystems and thousands of underlying parts; some of which may be complex systems on their own. As so, the paradigm behind the maintenance of such systems has suffered tremendous changes and the time where mechanics knew exactly what components to change based on their own experience is long gone. In this context, Fault Detection and Isolation (FDI) analytical-redundancy-based systems are slowly invading the industrial world and Airbus is not an exception.

In order to fit model-based diagnosis needs, [Airbus] engineers build subsystem models considering a compromise between the correctness of what is encoded in the model, the complexity of that same model and its future usage. Thus, models are typically built not to represent every possible real-world phenomena, but instead to approximate every predicted situation the system will be confronted to. Then, they are tested through a finite set of experiments and its validity is eventually admitted. If on one hand, this is perfectly acceptable from an industrial point of view; on the other hand one must be conscientious that:

- models representing artefacts of the real-world can never be proved true due to the problem of induction in the scientific method Hume (1978); Popper (1963);
- it is impossible in practice to perfectly represent, in a model, all the situations a system will be confronted to and, even if this was possible, the model would be too complex for practical real-life applications (qualification problem McCarthy (1977)); and
- human errors are ubiquitous in the real-world and models are sometimes badly encoded.

In a nutshell, models will always be subjected to errors (hereafter referred to as "model faults") which may entail the computation of false diagnoses and, consequently, huge economic losses in practical applications. Now, at this point, one can follow two distinct paths. The first one is to develop methods for reasoning in the presence of possibly faulty - or uncertain, or failing to correctly represent noise, and so on - models. This is what several works in the FDI community have been aiming at. Chen and Patton (1999), Erik Frisk and Lars Nielsen (2006) and Adrot et al. (1999) are just some of the many examples. However, even such fields make some hypotheses about the unknown, about how approximate the model is to reality or about how the reality is structured; this among many others hypotheses described, for instance, by Patton et al. (2000). Straightforwardly, these assumptions are sometimes confirmed to be false, which brings us directly to the second path. Concerning this second option, to our knowledge the FDI community is chronically missing of a method in its arsenal to detect and isolate faults in diagnostic models; a tool which would greatly help engineers to repair diagnostic models of complex systems.

Motivated by the discussion above, we contribute in this paper by providing FDI analytical redundancy approach with a method for detecting and isolating faults in diagnostic models of complex systems: meta-diagnosis.

As for the organisation, Section 2 starts by providing the reader with an example of how model faults can entail errors in a DC motor fault diagnosis. Then, in Section 3, the *Bridge* (cf. Cordier et al. (2000b) and Cordier et al. (2000a)) between the FDI analytical redundancy approach and the DX (a community focused on the principles of model-based diagnosis) consistency-based logical approach will be used to transform our problem into Reiter (1987) framework; a precondition enabling solution of our problem through the computation of meta-diagnoses using a "theory of meta-diagnosis" by Belard et al. (2011a), in Section 4. We will then return to the FDI community's world and explain and generalise our results. This will be the subject of Section 5. Finally, the main contributions and a few generalisations as well as some perspectives are provided in Section 6. A final remark: the reader is assumed to be familiar with the model-theoretic notions of structure, substructure, isomorphism or satisfiability. If not, Hodges (1993) provides the material needed.

2. WHEN MODELS HURT MODEL-BASED DIAGNOSIS: A DC MOTOR TRIAL

The present section is devoted to illustrating how faults in diagnostic models can affect the computed diagnoses and to providing a basis example for the remainder of this paper. We will start by building a diagnostic model of a DC motor and move to exposing how a diagnostic system based on this model fails to compute a valid diagnosis.

2.1 Modelling a DC motor

Imagine a team of engineers working in a diagnostic system of a DC motor through an FDI analytical redundancy approach. Moreover, suppose these engineers perceive the motor as depicted in Figure 1. Straightforwardly, from the

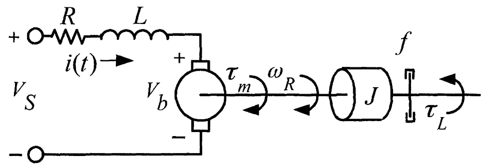

Fig. 1. A graphical representation of a DC motor.

electrical part of this circuit they can extract the following equations of interest:

$$V_L = V_S - R\cdot\ i - K_v \cdot \omega_R$$

$$\frac{d(i)}{d(t)} = -\frac{R}{L}\cdot\ i - \frac{K_v}{L}\cdot\omega_R + \frac{V_S}{L}$$

where R is the resistance of the resistor R, L is the inductance of the inductor L, V_L is the voltage across L, V_S is the drive voltage, i is value of the current in the rotor loop, ω_R is the shaft angular velocity and K_v is the velocity constant determined by the flux density, the reluctance of the core, and the number of turns of the armature winding. Moreover, admit R and L are modelled as constants.

As for the mechanical part of the circuit, suppose it is represented by the following equation:

$$\frac{d(\omega_R)}{d(t)} = \frac{K_T}{J}\cdot\ i - \frac{f}{J}\cdot\omega_R - \frac{\tau_L}{J}$$

where J is the rotor moment of inertial, f is the coefficient of viscous friction, τ_L is the load and K_T is the torque constant determined by the flux density, the reluctance of the core, and the number of turns of the armature winding. Again, suppose J and f are modelled as constants.

Now, hypothesizing no electromagnetic losses entails that electrical and magnetic powers equal each other and $K_v = K_T = K$. From all this information, the following state-space representation can be built:

$$\begin{pmatrix} \dot{i} \\ \dot{\omega}_R \end{pmatrix} = \begin{pmatrix} -\frac{R}{L} & -\frac{K}{L} \\ \frac{K}{J} & -\frac{f}{J} \end{pmatrix} \cdot \begin{pmatrix} i \\ \omega_R \end{pmatrix} + \begin{pmatrix} \frac{1}{L} & 0 \\ 0 & -\frac{1}{J} \end{pmatrix} \cdot \begin{pmatrix} V_S \\ \tau_L \end{pmatrix}$$

Suppose also that $R = 1\Omega$, $L = 0.5$H, $K = 0.01$N.m/A, $J = 0.01$kg.m^2/s^2, $f = 0.1$N.m.s and $\tau_L = 0$.

Moreover, imagine there is a set of sensors placed for diagnosis purposes so that the voltage across L, the angular velocity of the shaft ω_R and the current i can be measured. Furthermore, admit that engineers predict possible faults in the rotor causing changes in the resistance and inductance of the rotor loop, in the rotor moment of inertia, in the coefficient of viscous friction or in the torque constant/velocity constant, resp. $\mathrm{f_R}$, $\mathrm{f_L}$, $\mathrm{f_J}$, $\mathrm{f_f}$, and $\mathrm{f_K}$.

The next step into building a diagnostic system using an analytical redundancy approach is the computation of the Analytical Redundancy Relations (ARR) of interest. Imagine that these engineers determined the following ARR's, where the first three correspond to a motor with a positive drive voltage and the fourth one corresponds to a motor in an open circuit:

$$\text{ARR1 } r_1 \text{ where } r_1 = V_{L_{obs}} - [V_s - R\cdot\ i(t) - K\cdot\omega_R(t)]$$
$$\text{ARR2 } r_2 \text{ where } r_2 = i_{obs} - i(t)$$
$$\text{ARR3 } r_3 \text{ where } r_3 = \omega_{R_{obs}} - \omega_R(t)$$
$$\text{ARR4 } r_4 \text{ where } r_4 = \omega_{R_{obs}} - \omega_R(t)$$

As for their structures, ARR1 has a structure $\{\mathrm{f_L,f_R,f_K}\}$, ARR2 has a structure $\{\mathrm{f_L,f_R,f_K,f_J,f_f}\}$, ARR3 has a structure $\{\mathrm{f_L,f_R,f_K,f_J,f_f}\}$ and ARR4 has a structure $\{\mathrm{f_K,f_J,f_f}\}$

In this case, when rejecting both the exoneration and the no-compensation assumptions Cordier et al. (2000b) the fault signature matrix was determined to be:

	$\mathrm{f_L}$	$\mathrm{f_R}$	$\mathrm{f_K}$	$\mathrm{f_J}$	$\mathrm{f_f}$
ARR1	x	x	x	0	0
ARR2	x	x	x	x	x
ARR3	x	x	x	x	x
ARR4	0	0	x	x	x

This is the model that will be used for diagnosis.

2.2 Diagnosing the DC motor

All the necessary elements to tackle the problem are now gathered. Suppose the DC motor is installed in an helicopter. Suppose that, in a given mission, the DC motor was started with a drive voltage of 1V. Then, 1s after its start V_L, i and ω_R were measured to be, respectively, 0.0822V, 0.7348A and 0.0713rad/s. Some minutes later, admit the DC motor was put in an open circuit and that 0.2s after this, ω_R was measured to be 0.0135rad/s. Figure 2 illustrates these four measurements (crosses) and depicts the predicted values of those same variables. Based on such observations, the vector of residuals was computed to be $[1,1,1,0]^T$. The consequent diagnosis, assuming the

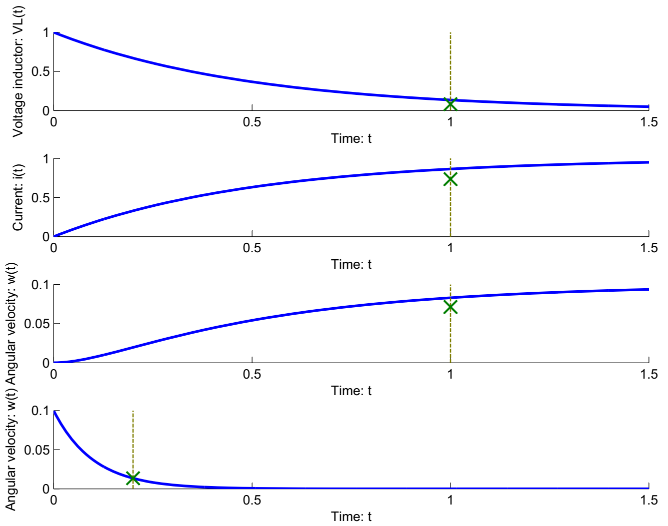

Fig. 2. The first three figures correspond to the model of $V_L(t)$, $i(t)$ and $\omega_R(t)$ in the first situation plotted against the observed values of these variables. The fourth figure corresponds to the second situation.

existence of no faults or repairs between observations, is: $\{f_L\}$, $\{f_R\}$, $\{f_K\}$ or any multiple fault resulting as a combination of at least these faults.

Finally, suppose that when the helicopter went back to the hangar for maintenance purposes maintenance operators checked for these three faults, but everything was found to be normal. They are now facing a meta-diagnosis problem that needs to be solved: why does the diagnostic system determines some faults to be present while their own perception is that every component is normal?

3. FROM AN FDI APPROACH TO A DX ONE: TRANSFORMING THE PROBLEM

A precondition for computing meta-diagnoses through the work of Belard et al. (2011a) is having a problem formalised in a DX consistency-based logical framework. As so, the present section is devoted to transforming the FDI analytical-redundancy-based framework of the previous section into a DX consistency-based logical one. This will be done with the help of *Bridge* (cf. Cordier et al. (2000b) and Cordier et al. (2000a)).

3.1 Characterising diagnoses in a consistency-based logical approach

Before going any further, let us state that the DX consistency-based logical approach sees model-based diagnosis as a reasoning problem that aims at retrieving abnormalities in a real-world artefact given a representation of that artefact (the so-called SD) and a set of observations OBS. Now, reality is accessible to every one of us, observer-participants, through *information*: a mathematical structure (of the cognitive-world), hereafter noted Ψ, underlying every observer-participant's percepts. As for percepts, they are the result of perception, a process through which observer-participants interpret a substructure of information corresponding to their sensed part of reality. This discussion will prove to be useful in the remainder of this paper. For now, let us move one step further and introduce the concept of believed system:

Definition 1. (Believed system). A believed system S is a pair (SD,COMPS) where:

(1) SD, the believed system description, is a set of first-order sentences. Semantically, typically represents the systems' structure, behaviour, functions and goals.
(2) COMPS, the believed system components, is a finite set of constants. Semantically, represents the systems' components whose faults one wants to diagnose.

As for SD-sentences, they rely on the Abnormality predicate to represent the behaviour of components:

Definition 2. (Abnormality predicate). Ab(.), the abnormality predicate, is a unary predicate. If $c \in$ COMPS is a believed system component, then Ab(c) and $\neg$Ab(c) are well-formed-formulas. Semantically, Ab(.) represents the abnormality of a component; and $\neg$Ab(.) its normality.

Observe that, in the same way as ARR's may contain errors, sentences in believed system descriptions may suffer from the same problem. This point will be highlighted in the following section. Apart from believed systems, consistency-based approaches rely on observations:

Definition 3. (Observations). The set of observations, OBS, is a set of first-order sentences. Semantically, they represent the perception of some observables in the believed system description.

Believed systems and observations are, thus, the DX consistency-based approach equivalent of the ARR's and measurements of the FDI analytical-redundancy-based approach. Now, the concept of diagnostic problem arises from the notions of believed system and observations:

Definition 4. (Diagnostic problem). A diagnostic problem DP is a tuple (SD,COMPS,OBS).

As for its solution, it is a set of *health states* respecting some conditions. Intuitively, health states determine the normality or abnormality of every element in COMPS:

Definition 5. (Health state). $\sigma(\Delta,\mathrm{COMPS}\backslash\Delta)$, a health - state, is $[\bigwedge_{c\in\,\Delta}\mathrm{Ab}(c)]\wedge[\bigwedge_{c\in(\mathrm{COMPS}\backslash\Delta)}\neg\mathrm{Ab}(c)]$.

Finally, the notion of consistency-based diagnosis appears, complemented with the concepts of partial and kernel diagnosis enabling an easier representation of diagnoses. Courtesy of Reiter (1987) and de Kleer et al. (1992):

Definition 6. (Diagnosis). Let $\Delta \subseteq$ COMPS. A diagnosis, D, for the diagnostic problem (SD,COMPS,OBS) is the set of all diagnosis hypotheses $\sigma(\Delta,\mathrm{COMPS}\backslash\Delta)$ such that:

$$\mathrm{SD}\cup\mathrm{OBS}\cup\sigma(\Delta,\mathrm{COMPS}\backslash\Delta)$$

is satisfiable.

Definition 7. (Partial diagnosis). A partial diagnosis for a diagnostic problem DP is a satisfiable conjunction Q of Ab-literals such that for every satisfiable conjunction R of Ab-literals covered by Q [1], $\mathrm{SD}\cup\mathrm{OBS}_{\mathrm{cons}}\cup\mathrm{R}$ is satisfiable and $\mathrm{SD}\cup\mathrm{OBS}_{\mathrm{cons}}\cup\mathrm{R} \models \mathrm{OBS}_{\mathrm{exp}}$.

Definition 8. (Kernel diagnosis). A kernel diagnosis is a partial diagnosis whose only partial diagnosis covering it is itself.

[1] A conjunction of Ab-literals is said to cover another conjunction of Ab-literals if every literal of the former is present in the later.

3.2 Re-diagnosing a DC motor

One is now ready to analyse the DC motor problem from a DX consistency-based perspective; and transform ARR's and measurements into believed systems and observations. First, the set COMPS may be defined as {L,R,K,J,f}, semantically representing the inductor, resistor, rotor moment of inertia, torque/velocity constant and coefficient of viscous friction "components". Moreover, SD can be designed by adapting the structures and the residuals in the analytical redundancy relations. The following first-order sentences extended with the appropriate axioms for arithmetic, the equations describing $i(t)$ and $\omega_R(t)$ and so on determine, SD.

$$\mathrm{S}_1: \quad Closed(circuit) \wedge \neg \mathrm{Ab(L)} \wedge \neg \mathrm{Ab(R)} \wedge \neg \mathrm{Ab(K)} \Rightarrow \Rightarrow (V_{L_{obs}} - [V_s - R \cdot i(t) - K \cdot \omega_R(t)] = 0)$$

$$\mathrm{S}_2: \quad Closed(circuit) \wedge \neg \mathrm{Ab(L)} \wedge \neg \mathrm{Ab(R)} \wedge \neg \mathrm{Ab(K)} \wedge \wedge \neg \mathrm{Ab(J)} \wedge \neg \mathrm{Ab(f)} \Rightarrow (i_{obs} - i(t) = 0)$$

$$\mathrm{S}_3: \quad Closed(circuit) \wedge \neg \mathrm{Ab(L)} \wedge \neg \mathrm{Ab(R)} \wedge \neg \mathrm{Ab(K)} \wedge \wedge \neg \mathrm{Ab(J)} \wedge \neg \mathrm{Ab(f)} \Rightarrow (\omega_{R_{obs}} - \omega_R(t) = 0)$$

$$\mathrm{S}_4: \quad Open(circuit) \wedge \neg \mathrm{Ab(K)} \wedge \neg \mathrm{Ab(J)} \wedge \neg \mathrm{Ab(F)} \Rightarrow \Rightarrow (\omega_{R_{obs}} - \omega_R(t) = 0)$$

The reader may notice that sentence S_1 corresponds to $ARR1$, and so on. This will be exploited later in this paper. As for the observations, they are two sets associated to the two situations mentioned before:

$$\mathrm{OBS}_1 = \{Closed(circuit)\ ,\ t = 1\ ,\ V_s = 1\ ,\ V_{L_{obs}} = 0.0822\ ,\ i_{obs} = 0.7348\ ,\ \omega_{R_{obs}} = 0.0713\}$$

$$\mathrm{OBS}_2 = \{Open(circuit)\ ,\ t = 0.2\ ,\ \omega_{R_{obs}} = 0.0135\}$$

Kernel diagnoses can then be computed for both the diagnostic problems DP_1 = (SD,COMPS,OBS_1) and DP_2 = (SD,COMPS,OBS_2), using the same assumptions as in the previous section, i.e. rejecting both the exoneration and no-compensation hypotheses (a default in consistency-based approaches). These are KD_1 = {{Ab(L)},{Ab(R)} ,{Ab(K)}} and $\mathrm{KD}_2 = \{\emptyset\}$. Finally, assuming that no faults or repairs occurred between both situations, the final kernel diagnosis is: KD = {{Ab(L)},{Ab(R)},{Ab(K)}}; the exact same result as the one obtained in FDI's analytical redundancy approach. This comes with no surprise since as affirmed in Cordier et al. (2000b), when no fault models are present - and they never are in the present FDI analytical redundancy approach - and when "releasing the exoneration and the no-compensation assumptions", then "FDI and DX views agree on diagnoses".

4. META-DIAGNOSIS

At this point, let us stop for a moment and re-gain a high-level vision of this paper. In Section 1 the need for a method to detect and isolate faults in FDI analytical redundancy approaches was introduced. This need was then illustrated with a DC motor example in Section 2. When this stage was reached, we stated that a precondition for computing meta-diagnoses through the work of Belard et al. (2011a) was having a problem formalised in a DX consistency-based logical framework. As so, the previous section was devoted to transforming the FDI analytical-redundancy-based depart framework into a DX consistency-based logical one. It is now finally the time to tackle the issue by using the tools formalized and developed by Belard et al. (2011a) and Belard et al. (2011b).

4.1 Characterising meta-diagnoses

First of all, the atoms in meta-diagnosis are *meta-components*, i.e. the parts of the model whose normality/abnormality one wants to assess. Meta-components behaviour and interactions are described in a *meta-system description* thanks to the unary predicate M-Ab(·); semantically representing meta-component's abnormality.

Definition 9. (Meta-system). A meta-system is a pair (M-SD,M-COMPS) where:

(1) M-SD, the meta-system description, is a set of first-order sentences. Semantically, it typically represents the model structure and behaviour.
(2) M-COMPS, the meta-system components, is a finite set of constants. Semantically, it is a representation of the model meta-components to be meta-diagnosed.

As stated by Belard et al. (2011a), the choice of meta-components depends on the users' goals and underlying hypotheses, that is, meta-components are defined according to what the user considers to be a fault in the model. For instance, to determine if each sentence in the believed system description describes the artefact to be diagnosed in a correct manner; one can associate a meta-component to every sentence in this description. We encourage the reader to keep this in mind; for it will be useful for understanding the work presented in the next subsection. From meta-systems, one can move the definition of *meta-observations.*

Definition 10. (Meta-observations). The set of meta - observations, M-OBS, is a finite set of first-order sentences. Semantically, it represents the perception of some parameter of the diagnostic system or the artefact itself.

In the same way as believed systems and observations were reunited to form diagnostic problems, meta-systems and meta-observations form meta-diagnostic problems:

Definition 11. (Meta-diagnostic problem). A meta-dia - gnostic problem M-DP is (M-SD,M-COMPS,M-OBS).

Solving a diagnostic problem consists in determining the normality/abnormality of meta-components. This is why, before focusing on a definition of meta-diagnoses, the concept of meta-health state needs to be apprehended:

Definition 12. (Meta-health state). Let $\Phi \subseteq$ M-COMPS be a set of meta-components. The meta-health state π(Φ,M-COMPS\Φ) is the conjunction:

$$[\bigwedge_{\mathrm{mc} \in \Phi} \mathrm{M\text{-}Ab(mc)}] \wedge [\bigwedge_{\mathrm{mc} \in (\mathrm{M\text{-}COMPS} \setminus \Phi)} \neg \mathrm{M\text{-}Ab(mc)}]$$

Definition 13. (Meta-diagnosis). Let $\Phi \subseteq$ M-COMPS. A meta-diagnosis, M-D, for the meta-diagnostic problem (M-SD,M-COMPS,M-OBS) is a set of all M-D hypotheses π(Φ,M-COMPS\Φ) such that:

$$\mathrm{M\text{-}SD} \cup \mathrm{M\text{-}OBS} \cup \pi(\Phi, \mathrm{M\text{-}COMPS} \setminus \Phi)$$

is satisfiable.

Before ending this subsection, the reader may notice that Definitions 6 and 13 are syntactically equivalent. Thus, a meta-diagnostic problem can be seen as a diagnostic problem where the artefact being diagnosed is a diagnostic system; and a DX consistency-based logical approach can be used to tackle meta-diagnostic problem. This is proved in Belard et al. (2011a) and Belard et al. (2011b).

4.2 Meta-diagnosing a DC motor model

With a formal theory of meta-diagnosis introduced, the DC motor diagnostic model can finally be meta-diagnosed. According to the theory, meta-diagnosis starts by defining the meta-components. Now, in the previous subsection it was stated that: "meta-components are defined according to what the user considers to be a fault the model". Now, intuitively, one is looking for the reason behind the invalid diagnoses computed. But how can this help the choice of meta-components?

In order to answer this question and definitely solve the problem, four definitions and a theorem are now borrowed from Belard et al. (2011a) [2]:

Definition 14. (Ontological truth of a theory). Given:

- T, a logical theory,
- Ω, the set of all structures,
- Mod(T), the set of all structures models of T, and
- $\Psi \in \Omega$, the structure of information (Subsection 3.1),

the theory T is said to be an ontological truth iff:

$$\exists_{\mathrm{s}\in\mathrm{Mod(T)}}\ \exists_{\mathrm{t}\in\Omega}\ (\mathrm{s}\subseteq\mathrm{t})\wedge(\mathrm{t}\rightleftharpoons\Psi)$$

where s$\subseteq$t means that s is a substructure of t and t$\rightleftharpoons\Psi$ means that t is isomorphic to Ψ. If T is an ontologically true theory, then so are every sentences in T.

Definition 15. (True health state). σ_{T}, the true health state, is the only ontologically true health state (cf. Definitions 5 and 14) at a given time.

Definition 16. (Validity of a diagnosis). Let σ_{T} the true health state. A diagnosis, D, is said to be valid iff $\sigma_{\mathrm{T}} \in$ D; and invalid otherwise.

Definition 17. (Ontological truth of B.S. description). A believed system description is an ontological truth iff, for all ontological true theories OBS, SD$\cup$OBS is an ontologically true theory (cf. Definition 14).

Theorem 1. If (SD,COMPS) is an ontologically true believed system, then for every diagnostic problem (SD, COMPS,OBS) with ontologically true observations, every diagnosis D is valid.

Will all this material, it can now be formally stated that one is looking for the reason behind invalid diagnosis (cf. Definition 16). Moreover, Theorem 1 indicates that this reason can be the ontological falsehood of the believed system description (cf. Definition 17). This being so, meta-components can be chosen as a representation of the sentences in SD, i.e. M-COMPS = $\{\mathrm{S}_1,\mathrm{S}_2,\mathrm{S}_3,\mathrm{S}_4\}$. Moreover, since either our meta-component is abnormal or the sentence is ontologically true, the following meta-system description, M-SD, can be built:

$$\neg\text{M-Ab}(\mathrm{S}_1) \Rightarrow \big[Closed(circuit)\wedge\neg\mathrm{Ab(L)}\wedge\neg\mathrm{Ab(R)}\wedge \wedge\neg\mathrm{Ab(K)} \Rightarrow (V_{L_{obs}} - [V_s - R\cdot\ i(t) - K\cdot\omega_R(t)] = 0)\big]$$

$$\neg\text{M-Ab}(\mathrm{S}_2) \Rightarrow \big[Closed(circuit)\wedge\neg\mathrm{Ab(L)}\wedge\neg\mathrm{Ab(R)}\wedge \wedge\neg\mathrm{Ab(K)}\wedge\neg\mathrm{Ab(J)}\wedge\neg\mathrm{Ab(f)} \Rightarrow (i_{obs} - i(t) = 0)\big]$$

$$\neg\text{M-Ab}(\mathrm{S}_3) \Rightarrow \big[Closed(circuit)\wedge\neg\mathrm{Ab(L)}\wedge\neg\mathrm{Ab(R)}\wedge \wedge\neg\mathrm{Ab(K)}\wedge\neg\mathrm{Ab(J)}\wedge\neg\mathrm{Ab(f)} \Rightarrow (\omega_{R_{obs}} - \omega_R(t) = 0)\big]$$

$$\neg\text{M-Ab}(\mathrm{S}_4) \Rightarrow \big[Open(circuit)\wedge\neg\mathrm{Ab(K)}\wedge\neg\mathrm{Ab(J)}\wedge \wedge\neg\mathrm{Ab(f} \Rightarrow (\omega_{R_{obs}} - \omega_R(t) = 0)\big]$$

[2] The reader is encouraged to refer to Belard et al. (2010), Belard et al. (2011a) and Belard et al. (2011b) for more details.

extended with the appropriate axioms for arithmetic, the equations describing $i(t)$ and $\omega_R(t)$ and so on. As for meta-observations, they contain all the observations at system-level. Moreover, this theory is extended by taking into account the unobserved abnormality in the three components analysed by maintenance operators: L, R and K. As so,

$$\text{M-OBS}_1 = \{\neg\mathrm{Ab(L)}\ ,\ \neg\mathrm{Ab(R)}\ ,\ \neg\mathrm{Ab(K)}\ ,\ Closed(circuit)\ ,\ t = 1\ ,\ V_s = 1\ ,\ V_{L_{obs}} = 0.0822\ ,\ i_{obs} = 0.7348\ ,\ \omega_{R_{obs}} = 0.0713\}$$

$$\text{M-OBS}_2 = \{\neg\mathrm{Ab(L)}\ ,\ \neg\mathrm{Ab(R)}\ ,\ \neg\mathrm{Ab(K)}\ ,\ Open(circuit)\ ,\ t = 0.2\ ,\ \omega_{R_{obs}} = 0.0135\}$$

With a meta-system and some meta-observations the kernel meta-diagnoses for both meta-diagnostic problems M-DP$_1$ = (M-SD,M-COMPS,M-OBS$_1$)) and M-DP$_2$ = (M-SD,M-COMPS,M-OBS$_2$)) can now be computed. These are M-KD$_1$ = {M-Ab(S$_1$)} and M-KD$_2$ = $\{\emptyset\}$. Finally, since it is assumed, as in Section 1 that no changes in the model occurred between both situations, then the final kernel diagnosis is: M-KD = {M-Ab(S$_1$)}. In a nutshell, the meta-diagnosis tells us that the first sentence in the believed system description must be repaired.

5. BACK TO AN FDI APPROACH

In the previous section a method was provided for computing meta-diagnoses using a DX consistency-based logical approach. However, what does that means in terms of the elements in FDI's analytical redundancy framework: ARR's and measurements?

The answer to such question relies on the fact that, under some assumptions, believed system description sentences are perfectly equivalent to ARR's residuals and their structure. We recall, once again, the words of Cordier et al. (2000b): "Releasing the exoneration and the no-compensation assumptions (...) FDI and DX views agree on diagnoses". As also stated in that work, this is true whenever one is not interested in fault models; for the present analytical redundancy approach cannot handle these objects. Now, since these were exactly the assumptions made all along the present paper (and these assumptions can always be made when one wants to meta-diagnose FDI analytical redundancy approach models) then the result from the previous section tells us that ARR1 needs to be repaired, either in terms of its structure or in terms of its residual definition.

Now, using this result, engineers inspected the residual in ARR1, i.e. $r_1 = V_{L_{obs}} - [V_s - R\cdot\ i(t) - K\cdot\omega_R(t)$ and found out the problem. In fact, our DC motor was always tested in low temperatures. Unfortunately, when the helicopter flew to a region where high temperatures were registered the resistance of our resistor changed. This happens because being made of iron, the resistor has an electrical resistance temperature coefficient of, by and large, 0.0062K^{-1}. When changing the residual equation of ARR1 to $r_1 = V_{L_{obs}} - [V_s - R_0\cdot(1+\alpha\cdot(T-T_0))\cdot\ i(t) - K\cdot\omega_R(t)$ with $R_0 = 1\Omega$, the test temperature $T_0 = 5$C and the temperature registered in the region where the flight took place $T = 45$C; engineers verified that, in fact, the diagnostic system provided the correct answers. The model was then assumed as being correct once again. Figure 3 extends Figure 2 with the new model to provide a visual representation of what has been stated.

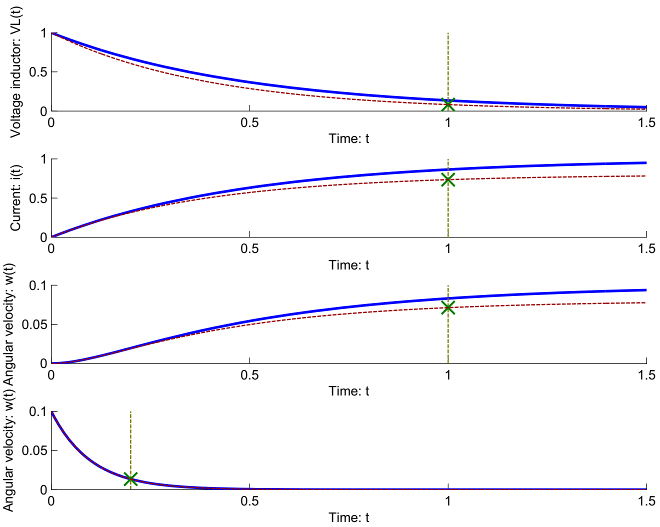

Fig. 3. An extension of Figure 2 with the repaired model of the DC motor depicted in a dashed red line.

6. DISCUSSION AND PERSPECTIVES

In this paper, we provided FDI community with a theory of meta-diagnosis, that is, a framework for detecting and isolating faults in diagnostic models. We have motivated the need for such contribution by the ubiquity of diagnostic model's faults and its associated economic impact. Furthermore, we have illustrated our point trough a DC motor example that, although not conveying the complexity of Airbus models with thousands of ARR's, enables the reader to understand our framework. The presented framework also illustrates how the cooperation between the FDI analytical redundancy approach and the DX consistency-based approach can enrich both communities.

Let us now open up with a two-fold remark. First, although in this paper we focused on detecting and isolating ontologically false ARR, this property is far from being the only one affecting these sentences. In fact, diagnosability, completeness and many other properties whose absence is considered as a fault in the diagnostic model, can be meta-diagnosed thanks to this work. In Belard et al. (2010) we provide a detailed logical account on models' properties and their relations with diagnoses' properties. Finally, despite our interest in diagnostic models, they are also far from being the only possible faulty part of diagnostic systems. In fact, both observations and diagnostic algorithms can be assigned some properties whose absence is considered abnormal. The framework presented covers such cases as proved by Belard et al. (2011a).

Finally, our perspectives are related to the question: do we really have to go to a DX consistency-based logical approach and come back every-time we meta-diagnose models in an FDI analytical redundancy framework?

Since Definition 6 and Definition 13 are syntactically equivalent and Cordier et al. (2000b) tells us that under some assumptions "FDI and DX views agree on diagnoses", then one could certainly try to investigate the possibility of directly computing meta-diagnoses through an FDI analytical redundancy approach whenever no meta-fault-models are used.

In this perspective, one could take, for instance, the first M-SD-sentence in the DC motor problem:

$$\neg\text{M-Ab}(\text{S}_1) \Rightarrow \big[Closed(circuit) \wedge \neg\text{Ab}(\text{L}) \wedge \neg\text{Ab}(\text{R}) \wedge \\ \wedge \neg\text{Ab}(\text{K}) \Rightarrow (V_{L_{obs}} - [V_s - R\cdot\ i(t) - K\cdot\omega_R(t)] = 0)\big]$$

and use the analogy between both worlds to formulate a meta-ARR with a meta-residual:

$$m - r_1 = (r_1 - r_{1_{obs}})$$

with $r_1 = 0$ if $V_{L_{obs}} - [V_s - R\cdot\ i(t) - K\cdot\omega_R(t)]) = 0$ and 1 otherwise; and $r_{1_{obs}} = 0$ if no component in the structure of ARR1 is meta-observed as faulty and 1 otherwise. Moreover, rejecting the exoneration and no-compensation hypotheses a $m - r_1 < 0$ could be interpreted as a possible misdetection and not as a fault in the model. Finally, the M-ARR would have the structure {ARR1}.

This is the subject of our present investigations.

REFERENCES

Adrot, O., Maquin, D., and Ragot, J. (1999). Fault detection with model parameter structured uncertainties. In *Proceedings of the 5th European Control Conference (ECC-99)*.

Belard, N., Pencolé, Y., and Combacau, M. (2010). Defining and exploring properties in diagnostic systems. In *Proceedings of the 21th International Workshop on Principles of Diagnosis (DX-10)*.

Belard, N., Pencolé, Y., and Combacau, M. (2011a). A Theory of Meta-diagnosis: Reasoning about Diagnostic Systems. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-11)*.

Belard, N., Pencolé, Y., and Combacau, M. (2011b). MEDITO: a logic-based meta-diagnosis tool. In *Proceedings of the 23rd International Conference on Tools with Artificial Intelligence (ICTAI-11)*.

Chen, J. and Patton, R. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers.

Cordier, M.O., Dague, P., Dumas, M., Lvy, F., Montmain, J., Staroswiecki, M., and Travé-Massuys, L. (2000a). A Comparative Analysis of AI and Control Theory Approaches to Model-based Diagnosis. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI-00)*.

Cordier, M.O., Dague, P., Dumas, M., Lvy, F., Montmain, J., Staroswiecki, M., and Travé-Massuys, L. (2000b). AI and Automatic Control Theory approaches of model-based diagnosis : links and underlying hypotheses. In *Proceedings of the 4th IFAC International Symposium on Fault Detection, Supervision and Safety for Technical Processes (Safeprocess'2000)*.

de Kleer, J., Mackworth, A.K., and Reiter, R. (1992). Characterizing Diagnoses and Systems. *Artificial Intelligence*, 56.

Erik Frisk and Lars Nielsen (2006). Robust residual generation for diagnosis including a reference model for residual behavior. *Automatica*, 42.

Hodges, W. (1993). *Model Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press.

Hume, D. (1978). *A Treatise of Human Nature*. Edited by L.A. Selby-Bigge and P.H. Nidditch. Oxford University Press.

McCarthy, J. (1977). Epistemological Problems of Artificial Intelligence. In *Proceeding of the 5th International Joint Conference on Artificial Intelligence (IJCAI-77)*.

Patton, R., Frank, P., and Clark, R. (2000). *Issues of Fault Diagnosis for Dynamic Systems*. Springer.

Popper, K. (1963). Science as Falsification. In *Conjectures and Refutations*. Routledge and Kegan Paul.

Reiter, R. (1987). A Theory of Diagnosis from First Principles. *Artificial Intelligence*, 32.