MONET2

Project Full Title: Network of Excellence on Model Based Systems and Qualitative Reasoning.

Contract: Concerted Action / Thematic Network

Contract No: IST-33540

Deliverable MBD2:

Common Diagnostic Framework Specification

Date: 30th January 2003
Version: 1.2
Status: Release

MONET Project Office
Department of Computer Science
University of Wales
Aberystwyth
Ceredigion
SY23 3DB

Contents

Authors

Louise Travé-Massuyès	LAAS-CNRF
Bridge Task Group	MONET Project

1 Introduction

1.1 Purpose of this Document

This deliverable proposes a Common Diagnosis Framework which is the result of a comparative analysis of MBD technologies developed in control theory (FDI) and AI (DX). It is aimed at helping identify for participants from different backgrounds the requirements of diagnostic processes, and relating the existing theories, methods, techniques, and systems to these requirements. The analysis compares and characterizes the presumptions and scope of applicability of MBD technologies developed in FDI and DX. It identifies competing and complementary methods using the DX MBD technology alongside the FDI MBD technology as well as it identifies and defines useful “bridges” for beneficial combinations of techniques.

2 Review of the MBD technologies in the AI and Control fields

(by P. Dague and L. Travé-Massuyès)

In DX as well as in FDI, the basic principle of model-based diagnosis is to compare the actual behaviour of the system, as it is observed, and its expected behaviour, as it can be predicted from its model and observations.

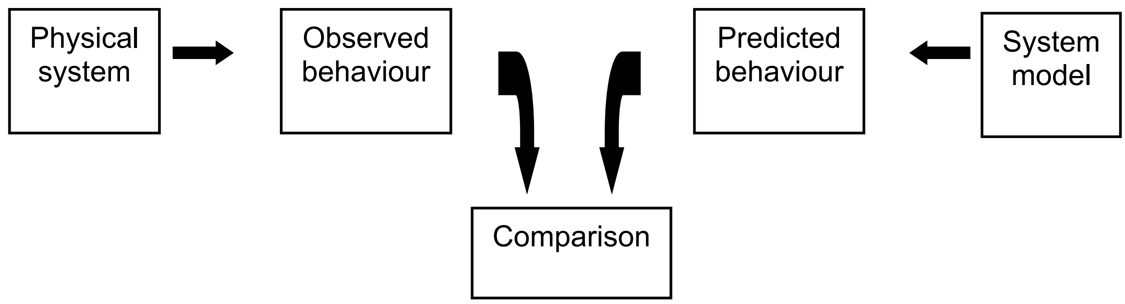

Figure 1: Principle of model-based diagnosis

2.2 MBD technology in Artificial Intelligence

We will briefly review in this section the state of the art of the logical theory of model-based diagnosis developed in AI (see Dubuisson, 2001a, 1st chapter, for a more exhaustive and detailed account and bibliography). We will point out in particular how it deals with the three subtasks of the diagnosis task:

- fault detection,
- fault localization, possibly followed by fault identification,
- fault recovery (repair, reconfiguration, etc.).

2.2.1 Principles and advantages

The method called *model-based diagnosis* or *diagnosis from first principles* was born in mid 70s and formalized at the beginning of 80s. An increasing number of works has been carried out since and it became a full research field in AI, as proved by the annual international workshop on Principles of Diagnosis which holds since 1989 (it is interesting to notice that it just followed the creation in 1987 of the annual international workshop on Qualitative

Reasoning, another field of AI, and that the two communities have a strong interaction, in particular in what concerns the modelling task). The term “model” is used here in contrast with associational knowledge, in general empirical, used in traditional early AI diagnostic techniques, such as expert systems, but also with black box models as used for example in pattern recognition, neural networks techniques or case-based reasoning (one speaks also of deep knowledge in contrast with shallow knowledge).

A model is thus an explicit white box model describing the functioning (or dysfunctioning) of the system, analogous to the use of this term in Control Theory (see next section). More precisely, in a component-centred approach, which is the one commonly used for technical systems (more generally a process-centred view can be considered), a model includes two types of knowledge: the first one describes the *structure* of the system to be diagnosed, i.e. the connections between the components, the second one describes the *behaviour* of each component. The whole is called the *system description* (SD). Applying compositional modelling allows one to obtain the global behavioural model of the system from SD. Only structural knowledge and components parameters are specific to the given system. Behavioural knowledge comes more or less directly from physical laws and is thus generic, only depending on the physical field in question (electronics, thermodynamics, hydraulics, mechanics, etc.), and can thus be reused, on the form of a library of models of generic components. What is called *components* (COMPS) in this description has to be understood in terms of diagnostic purpose and the granularity of the description is fixed by the level of diagnosis required: elementary components are thus the physical supports of the finest primary faults that have to be localized, i.e. for which fault recovery can apply (e.g. repairable, replaceable or reconfigurable parts of the system). Finally the system description is completed by the knowledge of *observations* (OBS) coming in general from sensors measurements (such observations may be inputs, such as commands, or outputs or intermediate variables values).

First approaches to diagnosis were based on *abductive* reasoning. It consists in starting from the effects (observed symptoms) and in trying to infer their possible causes (primary defects), as currently done in medicine for example. In this context, diagnosis requires the use of fault models and rests on the search of corroboration between the observed behaviour and some modelled faulty behaviour. The main drawback is that fault detection rests on an exhaustive description of symptoms, and fault identification on an exhaustive description of causes and their relationships to symptoms. Lack of completeness of fault models implies that corroboration, and thus abductive reasoning, is not logically sound: it can happen that faults are not detected as it can happen that identified faults are not the real ones.

The fundamental idea was then to base diagnosis on a refutation principle instead of a corroboration principle. Not only this guarantees logical soundness but this allows also diagnosis to be performed from the modelling of the single correct behaviour of the system: there is no need of any prior knowledge about faults or failures. Indeed, any discrepancy between the observed behaviour and the modelled correct behaviour of the system refutes the fact that the system behaves properly and thus signs necessarily the presence of fault(s). In other words, finding that the system description together with the observations is inconsistent is a sound proof of fault detection. The idea to localize the detected fault(s) is to stop assuming that all components are correct and to check consistency of the system description, containing now some hypothesized faulty component(s), with the observations. Notice that, as only correct behaviour of components is modelled, hypothesizing a faulty component boils down to simply removing its model without replacing it by anything else (the model of the system is thus less constrained). If inconsistency still exists, it means that hypothesized fault(s) is (are) refuted also. Any possible combination of hypothesized fault(s), which restores the consistency with observations, is thus a possible diagnosis. Technically, each correct behavioural model M of a component c is expressed in the system description by recording explicitly the assumption under which this model is valid, i.e. the correctness of c, which is used to be written as ¬AB(c), where the unary predicate AB(.) means Abnormal: $\neg AB(c) \in M$. So, at the beginning, all components are supposed to be correct by default: $\{\neg AB(c) \mid c \in COMPS\}$. Fault detection occurs when $SD \cup OBS \cup \{\neg AB(c) \mid c \in COMPS\}$ is proved to be inconsistent. Hypothesizing that the component c is faulty means switching from ¬AB(c) to AB(c). A *diagnosis* is thus a subset Δ of COMPS such that: $SD \cup OBS \cup \{AB(c) \mid c \in \Delta\} \cup \{\neg AB(c) \mid c \in COMPS - \Delta\}$ is consistent. This model-based diagnosis technique, which, by contrast with abductive diagnosis, consists in checking for inconsistency and revising the assumptions of components correctness in order to restore consistency is called *consistency-based diagnosis*, and can be viewed as a logical theory of diagnosis.

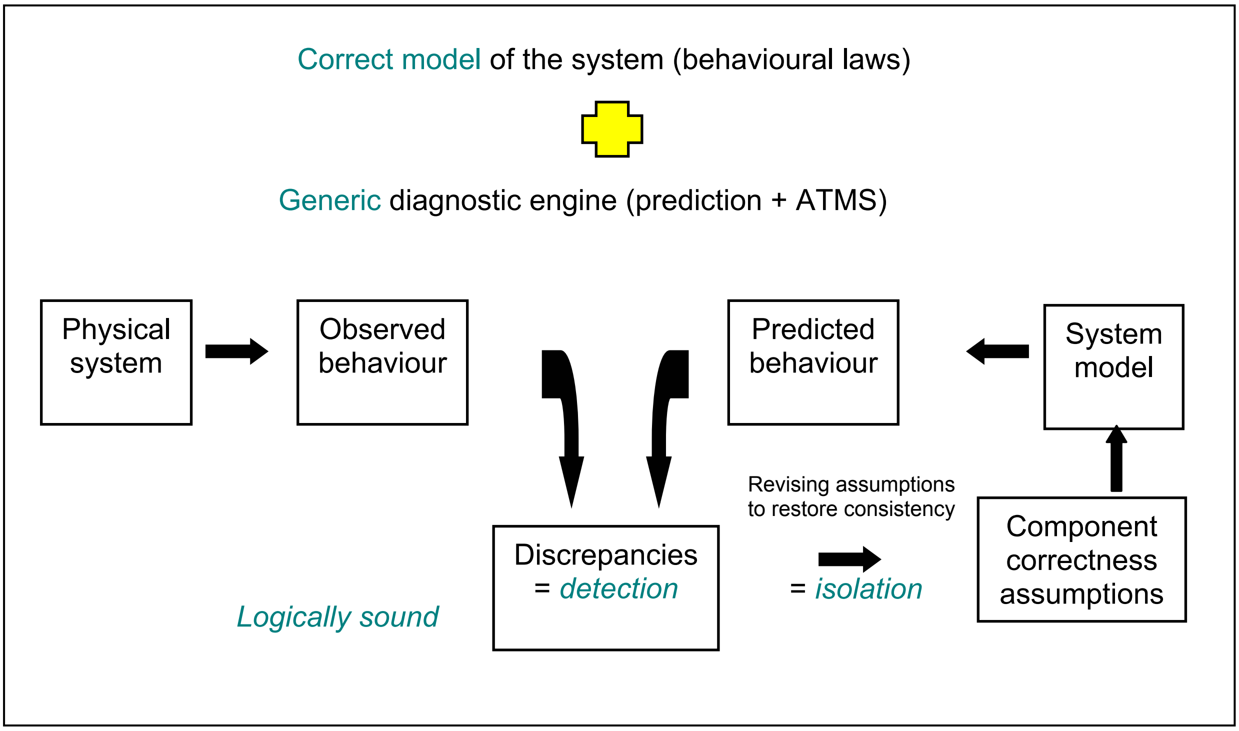

Figure 3: Consistency-based diagnosis principle

It does not develop any specific technique for fault detection, which is simply considered as proving the inconsistency of the correct functioning system description with observations: as such, any consistency checking method currently used in logics may be chosen, the system description being generally considered as expressed in first order predicate logic with equality. E.g., a contradiction between an observed and a predicted value for a same variable just boils down to check that both values are not equal. No threshold analysis to trigger detection is proposed. The usual way to deal with this is to express system description in a constraint-based language, e.g. with interval-based constraints, and to express explicitly models uncertainty, parameters dispersion or sensors precision in the system description. The theory mainly develops fault localization techniques. Indeed, the search space for all possible diagnoses is exponential in the number of components and smarter techniques than just guessing hypothesized faults before checking them can be developed, as it will be explained in the next sections. The theory deals naturally with multiple faults, which are difficult to handle in other frameworks, and indeed several diagnoses of various cardinalities generally exist. As it will be seen, preference criteria are then defined to keep only some of them or to rank them. This initial version of the theory using only correct modelling of the components, the identification task cannot be performed in general and no explanation is given as a faulty component has an arbitrary not defined behaviour. But it will be seen that the theory can be extended by using fault models of components, allowing thus to perform abductive diagnosis in the framework of consistency-based diagnosis. Finally some works exist about repair (or reconfiguration) and how to interleave it with fault localization (or identification): using a model-based framework is here also an advantage as the same model is shared for diagnostic reasoning and repair reasoning.

Before describing in more detail these different topics, it is important to precise the implicit properties that the models need to have in order that the logical soundness of the method is ensured. In order to conclude to the presence of a defect, one has to be sure that the models well reflect the correct behaviour of the components. More precisely, it is assumed that the models are *necessary conditions of correctness* of components. This means that, if a component is correct, then its behavioural model is satisfied or, equivalently, that if it happens that its model is proved to be violated then it is necessarily faulty (we saw it is precisely the way it is written in SD: ¬AB(c) ⇒ M). In these conditions, and only in these, detecting inconsistency between model and observations is a sound proof that a component is faulty. Ideally, it should be wanted that models fit exactly the reality and are thus also sufficient conditions of correctness of the components, i.e. that satisfaction of a component's model implies the correctness of this component. But both properties are impossible to be satisfied in practice by a same model: the first property is ensured by using some coarse model as the second one would require very precise models. When the last one is not true, it means that some inconsistencies, that should be found, are not and thus that some defects may not be detected: in practice what happens most often is that the localization is less precise and the ambiguity between different possible diagnoses is greater, which requires to replace or reconfigure a larger set of components than would be needed. In the following, we will consider only, as usual in the theory, models of the first type, ensuring logical soundness of fault detection.

To sum up, the main advantages, often decisive, of the consistency-based diagnosis are the following:

- Logical soundness of diagnostic process is guaranteed, as it is based on consistency checking;
- Diagnosis rests on design knowledge more than on exploitation knowledge: in particular, the system does not need to be operational;
- Faults and symptoms do not need to be anticipated: only correct behavioural models are needed for fault detection and localization; but if some fault models exist, they can be efficiently taken into account in the same framework for fault identification and symptom explanation;
- Multiple faults are automatically taken into account in this framework;
- Knowledge about the system (models) is separated from knowledge about the diagnostic task itself (generic diagnostic engine with its strategies);
- Models provide an adequate framework for explanation (the structure and the behaviour of the system are explicitly represented);
- Modelling is constructive and modular: the model of the system is derived from a library of generic models of the elementary components, making easier models reusability;
- Same model is used for diagnosis and reconfiguration or recovery;
- Development, validation and maintenance costs are decreased for all these reasons.

The strong point of the theory, i.e. being model-based, turns out to be often its main drawback for real applications. From one side, the fact that the theoretical framework is independent from the modelling formalism (nothing is imposed to SD) and from the algorithms used for checking consistency between models and observations makes this theory universal, i.e. applicable a priori for any type of model and any chosen theorem prover. From the other side, when working on real applications, one sees that the most difficult task is indeed the modelling itself.

The present limitations to the scope of this approach concern dealing with time, for which more research has to be done. Nowhere time is mentioned in the theory, now most of systems are dynamic ones. In this case, time will appear in the modelling, i.e. in the system description, either explicitly (e.g. as differential equations) or implicitly (e.g. as state changes in finite state machines). But the repercussion on the diagnostic process itself, i.e. diagnosis across time, has not been well formalized as the rest of the theory (even if a lot of works and real experiments exist). The theory can be easily extended to the case of successive diagnostic snapshots at different time slices. This so-called *state-based diagnosis* works fine if observations at each instant allow one to reconstruct the state of the system at this instant. If it is not the case, the model has to be simulated along time while be checked against observations: this *simulation-based diagnosis* requires much more computation, which can prevent it to be used on-line, depending of the dynamics of the system. A more important difficulty occurs for *time-varying* systems, i.e. systems whose parameters evolve along time or more generally whose faults evolve (e.g. appear) along time. Research for applying model-based diagnosis, and in particular consistency-based diagnosis, to such time-varying systems has progressed and indeed several applications to real world systems exist or are currently developed. But there is not yet a unifying consistency-based diagnosis theory for time-varying systems, as it exists for not time-varying ones, and this is still a research issue.

2.2.2 The different steps of the basic diagnostic process

The steps for achieving fault detection and localization in the pure consistency-based approach (i.e. only with models that are necessary conditions of correctness of components) for a system in a static state of faults (which excludes transient or intermittent faults) are described in the following.

2.2.2.1 Conflict detection

Fault detection is achieved by proving inconsistency of the modelled correct behaviour with the observations. In general, the problem solver used to do this is able to report, more or less precisely, the possible causes of the detected inconsistency, i.e. the components for which the correctness assumption is involved in this inconsistency. This can be done for example by keeping track of these assumptions when the corresponding correct model has been used in a prediction that led to a contradiction (with an observation or another prediction). This gives information for fault localization. Actually, if a contradiction is supported by the correctness assumptions of a subset **C** of components, this means that these components cannot be all correct and thus that at least one of them is necessarily faulty. Such a subset **C** of components is called a conflict set or a Reiter-conflict, here in short a *conflict*. Thus a conflict is a subset **C** of COMPS such that: $SD \cup OBS \cup \{\neg AB(c) \mid c \in \mathbf{C}\}$ is inconsistent. The more observations accumulate, the more contradictions have chance to be obtained and the more the collection of conflicts increases (obviously only minimal conflicts for set inclusion have to be kept), allowing by intersecting them to refine the fault(s) localization (this remains true for a sequence of sets of observations at different time instants, as the faults are assumed to be permanent: the collection of conflicts for the whole sequence is the union of the sets of conflicts detected for each time stamped set of observations).

2.2.2.2 Generation and test of assumptions

The second step consists in computing the diagnoses from the conflicts, i.e. in generating the assumptions of abnormal behaviour of components c, switching their mode from ¬AB(c) to AB(c), in such a way that consistency is restored with observations, i.e. that all conflicts disappear.

The idea is that for making each conflict disappear, it is needed to switch the behavioural mode of at least one component of each conflict. Expressed in set theory, it means that a diagnosis has to intersect each (minimal) conflict. With the modelling assumptions made, switching a component from ¬AB to AB cannot bring inconsistency if no one existed and thus any superset of a diagnosis is still a diagnosis, which implies that the converse of the property above is true, i.e. any subset of components that intersects all the (minimal) conflicts is a diagnosis. For a given set E and a given collection **F** of subsets of E, let us define a *hitting set* of **F** as any subset of E that intersects each member of **F**. The following result then holds: *the diagnoses are exactly the hitting sets of the collection of minimal conflicts*. Diagnoses of size 1 correspond to single faults, diagnoses of size greater than 1 to multiple faults. As any superset of a diagnosis is still a diagnosis, all diagnoses are thus characterized by the minimal (for set inclusion) diagnoses, and attention is focused only on these *minimal diagnoses*. They are computed as the minimal hitting sets of the collection of minimal conflicts. A hitting set algorithm has been proposed by (Reiter, 1987; Greiner *et al*., 1989). It is exponential in the size of the collection, this problem being actually NP-complete.

2.2.2.3 Typical ATMS-based algorithm

The typical diagnostic algorithm was proposed by (de Kleer, Williams, 1987) as GDE (General Diagnosis Engine). It uses an ATMS (Assumption-based Truth Maintenance System), invented the year before by de Kleer, coupled to the solver. The role of the ATMS is to store the inferences done by the solver in the form of *justifications*, to update the labels of the facts from these justifications and the assumptions, and to keep up to date the collection of *nogoods*, which are exactly, in this context of diagnosis, the minimal conflicts. Minimal diagnoses are then generated from these conflicts by a hitting set algorithm. Almost all existing diagnostic tools are inspired by GDE and use the de Kleer's ATMS or one of its variants (such as an incremental non-monotonic TMS with focusing abilities).

2.2.2.4 Discrimination between diagnoses

The third step (which does nt necessary exist) consists in gaining additional information allowing the remaining diagnoses to be discriminated at best. There are mainly two ways of obtaining new information: making new measurements on the system if possible (this is basically the case of off-line sequential diagnosis for a static or periodic dynamical system, e.g. new probing done by an expert diagnosing an electronic device, but it could be applied to on-line diagnosability at the design stage) or changing the input of the system if possible in order to put it in another context (this is the case of active diagnosis).

In the first case, techniques have been developed to choose the best next measurement. In GDE, for example (de Kleer, Williams, 1987), the chosen measurement is the one that minimizes the entropy of the probabilities of the diagnoses that could result from this measurement. These techniques make use of the initial probabilities of component failure and assume the independence of faults. If these probabilities are unavailable or unknown, the techniques can be adapted (de Kleer, 1990) to qualitative probabilities (e.g. if it is just known that components fail with close and very small probabilities).

The second case is a matter for the automatic generation of discriminating tests, for which much work remains to be done (cf. (Shirley, Davis, 1983; Singh, 1986) in the framework of logical circuits and (McIlraith, Reiter, 1992; Struss, 1994) for a general formalization).

2.2.2.5 The diagnostic cycle

The diagnostic process progresses in a cyclic form: once new measurement to be done has been selected in the third step, one comes back to the first step. It is essential for the diagnostic algorithm to be incremental with respect to the increasing set of observations. The hitting set algorithm can easily be made incremental with respect to new conflicts. For the ATMS, the updating process of labels and nogoods is incremental with respect to new justifications.

There is no general stopping condition to the diagnostic algorithm, except that all possible measurements have been made. It is a matter of trade off between the computation time required by an additional cycle and the expected gain.

2.2.3 Extension with fault models

The first idea of consistency-based diagnosis, which is also one of its strengths, is to do without any knowledge about faults or failures. Nevertheless, such knowledge, when available, allows one to go further by identifying faults that have been previously localized and also by ranking diagnoses, e.g. according to their probability or their explanatory power (in this last case, this allows one to integrate in a sound way abductive reasoning into consistency-based reasoning). This is why extensions of the basic theory have been proposed, which allow fault models to be used in addition to correct models: GDE^+ (Struss, Dressler, 1989) et Sherlock (de Kleer, Williams, 1989). The important feature is that, in contrast with fault-based associational and empirical approaches, the logical validity of the obtained diagnosis is guaranteed.

2.2.3.1 Using fault models

Instead of modelling as previously only the correct behavioural mode ¬AB(c) of each component c, its faulty behavioural mode Ab(c) is also modelled. In general, several mutually exclusive fault modes, the most typical ones (as open circuit and short circuit in electricity), are represented. But, in order to deal with the unavoidable lack of exhaustiveness of these fault modes, that make the classical abductive approach unsound, an unknown mode with no behavioural model is explicitly added, which represents all dysfunctioning cases whose behaviours are unknown or not listed above. If we denote the correct mode G, the explicitly modelled fault modes F_i and the unknown, not modelled, mode U, we have thus the equivalences:

$$\neg Ab(c) \Leftrightarrow G(c)$$
$$Ab(c) \Leftrightarrow F_1(c) \vee \ldots \vee F_n(c) \vee U(c)$$

Each component c has now a finite number of mutually exclusive *behavioural modes* {G(c), F_1(c), …, F_n(c), U(c)}, the last one being not modelled. These are these modes that play now the role of assumptions. In this framework, a conflict becomes a mode assignment to some components, which is inconsistent with the observations. A diagnosis is then a mode assignment to all components, which restores consistency with observations.

The general diagnostic process and the algorithms seen above can be extended to this general framework. In particular, the ATMS works now directly with the new mode assumptions and has not to be modified and the hitting set algorithm can be adapted to this case. But obviously the combinatorial complexity is greater. Focusing techniques based on probabilities or explanation power (cf. de Kleer, 1991) are needed to avoid combinatorial explosion. In practice, correct models are used first to detect and localize the faults, and then only those fault models of the faulty components involved in some diagnosis in order to identify the faults.

2.2.4 Diagnosis and repair / reconfiguration

One has not to forget that the real goal of the diagnostic process is to restore the functionality of the system (at best). Moreover, all the diagnostic process has to take into account what recovery actions are available: fault localization precision must hence be consistent with repair actions granularity and fault identification is not necessary in general for actions like replacement or reconfiguration, but is required for maintaining on-line the proper functioning of the system (fault accommodation).

Without even considering general recovery actions, it is only recently (from 1991) that works tackled in its whole the diagnostic / repair (or reconfiguration) process viewed as alternating successive stages of diagnosis and of repair (Friedrich *et al*., 1992, 1994). Taking into account actions leads naturally to *planning*: (Friedrich and Nejdl, 1992) and (Sun and Weld, 1993) have introduced independently the notion of repair plans. (Provan, 1999) uses the same diagnostic software to carry out at once diagnosis and reconfiguration by swapping the status of component behavioural modes (hypotheses to be established for diagnosis, established facts for reconfiguration) with the status of actions (known for diagnosis, hypotheses to be selected for reconfiguration). One will refer to (McIlraith, 1994; Baroni *et al*., 2000) for an attempt of an extended theory of diagnosis, test and repair, integrating the notion of diagnostic plan and repair plan, expressed in an actions-based language.

2.2.5 Modelling

Modelling is the real key issue of MBD. Even modelling the correct behaviour is not easy, as it is not limited to nominal behaviour: indeed the presence of a fault somewhere in the system may cause a lot of components to adapt themselves to an environment changed by the effects of the fault (this is particularly true for looped systems) and, while remaining correct, to behave differently that in the nominal case. This implies to model the largest range of correct behaviour of each component for some environment, i.e. inputs, or other. In this case, precise parameter values of numerical models as those used in design or in the engineering world are not available in general (Struss, 1991).

Qualitative modelling can be a resort to this situation. Models with a higher level of abstraction have in general a larger scope of validity with respect to the functioning context of the modelled components. Moreover any modelling cannot have a universal scope and has to be suited to the task aimed at. Diagnostic models do not require anyway the precision of design models. The reader is referred to (Weld, de Kleer, 1990; Travé-Massuyès *et al*., 1997) for more detailed information and for bibliography about qualitative modelling and reasoning.

2.2.5.1 Modelling lines

Classically and somewhat roughly, two modelling lines have to be considered: *discrete versus continuous* (i.e. domains of values of modelled physical quantities are either finite, numeric or symbolic, or isomorphic to a real interval) and *static versus dynamic* (i.e. modelled physical laws are either time independent, e.g. algebraic equations, or not, e.g. differential equations). Obviously, categorization has in fact to be more precise, in particular in what concerns time: how the time variable is represented and what are the data influenced by time. Time can be explicitly represented and thus, as any other physical quantity, either discrete (numeric or symbolic) or continuous. It can also be implicit, represented by events.

In DX, difficulties of modelling, of designing predictive and consistency checking algorithms, and finally of the whole diagnostic process, grow when going from static discrete case to time-varying dynamic continuous case. Complex systems are in general heterogeneous and require models of different types to be used. For example, most of industrial plants to be supervised are dynamic continuous processes. But they have their own control-command system, naturally modelled as an event-driven discrete system. It is the prototype of a *hybrid* system, mixing continuous and discrete features (for physical quantities and for time).

For complex systems, the diagnosis cannot be tackled directly at the level of elementary components and their behavioural models. A new modelling line is thus a *hierarchy* in abstraction. Two sorts of hierarchy can be considered: *structural* (Chittaro and Ranon, 1999) (Darwiche, 1998, 1999) and *functional* (Chittaro, 1995; Chittaro and Ranon, 1996; Chittaro *et al*. 1998) (Provan, 1998, 2001).

2.2.5.1.1 Causality

Another important concept, which has a relationship with time, is *causality*. It has been the subject of a lot of works, first in econometrics, and more recently in AI, in particular in qualitative reasoning (de Kleer and Brown, 1986)(Iwasaki and Simon, 1986).

Taking into account causality may offer much potentiality to the diagnosis, in particular with regard to the explanation task. Different model-based diagnosis approaches can thus be distinguished according to this new modelling line (Travé-Massuyès *et al*., 1993): based on non-causal models (such as algebraic equations), based on causal models (most often, influence graphs between variables) or based on hybrid models (mixing non-causal and causal models, as the CA~EN system (Bousson et al., 1994; Travé-Massuyès and Pons, 1997) used in the TIGER project (Travé-Massuyès and Milne, 1997) for gas turbine monitoring).

In the framework of abductive diagnosis, in most cases, a diagnosis is required to explain only the non-temporal part of observations, and only consistency is used for temporal aspects. This is in this way that the abductive approach of (Console *et al*., 1989) is extended in (Console and Torasso, 1991b) to the computation of explanations that are temporally consistent with respect to quantitative temporal constraints. A similar approach is used in the DIAPO system of EDF (Porcheron *et al*., 1994; Porcheron, 2001) for diagnosing pumps of the primary cooling circuit in pressurized water nuclear plants.

2.2.6 On-line diagnosis and supervision

The basic idea of on-line model-based diagnosis is the same as for off-line model-based diagnosis: compare the predicted (from modelling) behaviour and the observed behaviour, any inconsistency between them indicating at least one fault, that one will try to localize by investigating the deductive path which led to the inconsistency. In addition to real time problems (response time of the diagnostic software) which can be crucial for processes with fast dynamics or at the time of abrupt failures and can require off-line compilation techniques, on-line diagnosis introduces a new *temporal* part since prediction and observation times have now to be synchronized. Indeed, as soon as the response time of the own dynamics of the system cannot be considered as negligible, it is generally impossible to carry out the diagnosis on snapshots of the system, i.e. as a simple sequence of static diagnoses, and the diagnostic tool must integrate the dynamics of the system as well as the dynamics of the faults. This raises problems much more difficult to solve than for the off-line diagnosis: temporal tracking and simulation of the evolving process, emergence of new dysfunctioning modes, which identify themselves either to abrupt failures or to time evolving failures (degraded modes) inducing slow drifts to be distinguished from those due to the unavoidable approximation of the prediction, etc.

In spite of these difficulties, a lot of works have been carried out from ten years about diagnosis of dynamical and, to a certain extent, time-varying systems and, more generally, about supervision (we refer to (Dubuisson, 2001a, 1st chapter) for more information and bibliography). In the framework of consistency-based diagnosis, one can quote, between 1989 and 1996, the following academic systems: MIMIC, DATMI, MIDAS, QDIAG, DIAMON, IDUN, DYANA, QDOCS, most of them being based on a qualitative modelling and using the qualitative simulator QSIM or its extensions. In the framework of abductive diagnosis, an approach to diagnosis of time-varying systems is proposed in (Console *et al.*, 1994) that considers the whole spectrum of choices for the logical notion of explanation but limits itself to the only consistency for the temporal aspects. It is based on the behavioural modes of the components and on a model of the temporal evolution of these modes, specifying the possible transitions between defects. Other works (Gamper and Nejdl, 1997) consider at once a temporal behaviour described by qualitative temporal constraints based on Allen's algebra and a time-varying behaviour described by transition graphs between modes, and require that some qualitative abstractions of the temporal information about observations be implied by the constraints involved in the part of the model used in explanation.

Though some of works above gave rise to applications, a complete theory of model-based supervision, which would extent the one presented here, in particular in integrating explicitly the temporal part, is not yet operational: (Raiman *et al.*, 1991) exploits only an axiom of non-intermittency of faults; (Friedrich and Lackinger, 1991) is a preliminary attempt to extend the theory by taking into account intermittent failures or, more generally, failures with temporal properties (a probabilistic approach based on Markov processes allows one to determine plausibility values of the temporal diagnoses from failures and repairs probabilities).

It is interesting to notice that a series of works (Dressler and Freitag, 1994; Dressler, 1996; Malik and Struss, 1996; Struss *et al.*, 1997) have shown experimentally that it is sometimes possible to diagnose a dynamical system and supervise a time-varying system (modelled in general qualitatively, for example with QSIM or with qualitative deviations) by considering only the intrastate constraints (at the observation instants) and ignoring the interstate constraints (concerning the evolution between successive observation instants) and thus the temporal problems. (Struss, 1997) has stated theoretically the conditions under which it is thus possible to limit oneself to a *state-based diagnosis* without resorting to a *simulation-based diagnosis*, obviously more costly in terms of computation.

Let us mention an approach, which uses only a one-step simulation (time being just a partially ordered sequence of states), and rests on a modelling of the system with a *transition system*. Tracking the system consists in determining, from the observations, what transitions occurred (in practice, several states are tracked, namely the most likely with respect to the given prior probabilities of the transitions). This is the approach adopted by Livingstone software (Williams and Nayak, 1996) for diagnosis and self-reconfiguration of *autonomous systems* dedicated to spatial exploration. This software implements the *mode identification* MI and *mode reconfiguration* MR modules, which, with a planning and scheduling module and a reactive execution module, constitute the architecture of the *Remote Agent* developed by the NASA in the framework of the *New Millennium Program* (Muscettola *et al.*, 1998).
The same sort of discrete modelling is used for *alarms* monitoring and interpretation in industrial environments (Cauvin *et al.*, 1998). Modelling is generally done by *discrete-event systems*, most often automata, and a lot of works (Cordier and Thiébaux, 1994; Sampath *et al.*, 1995, 1996, 1998; El Fattah and Provan, 1997; Baroni *et al.*, 1999, 2000; Lamperti and Zanella, 1999, 2000a, 2000b) (Pencolé, 2000; Lamperti and Zanella, 2001) (Cordier and Largouët, 2001). Note also the on-line chronicles recognition techniques, where a *chronicle* is a sequence of events linked by temporal constraints (generally generated off-line by machine learning techniques from the log files). These works are applied to the supervision of power distribution networks (AUSTRAL project from EDF (Laborie and Krivine, 1997a,

1997b) and project from Cise (Lamperti and Pogliano, 1997)) and telecommunication networks (GASPAR project from France Telecom (Bibas *et al*., 1996a, 1996b; Dousson, 1996; Rozé, 1997) and MAGDA project from Telecommunication Research French Network (Pencolé, 2000; Aghasaryan and Dousson, 2001; Pencolé *et al*., 2001)).

Supervision introduces still many additional features. One refers to (Montmain, 2000) for an exhaustive survey of the problems posed by the supervision and the various models for supervision, from both AI and control theory perspectives.

2.3 MBD technology in Control

The application of control " modern " theories (Rosenbrock, 1970) to fault detection and diagnosis started in the 60's. Nowadays, this topic defines a well-identified research line in the control field. Several sessions are devoted to this topic in all major conferences of the field (Decision and Control Conference, IFAC World Congress on Automatic Control, etc.) and a dedicated IFAC symposium, untitled SAFEPROCESS, is held every three years.

This section presents the main ideas underlying the model-based approach and the different methods, without entering the technical details. For more details, the reader is referred to the various references given in the text, in particular (Gertler, 1993, 1998; Iserman, 1997; Patton, 1989; CEP, 1997) for survey papers.

The control approach is based on the use of *numeric analytical models* and tackles the diagnosis problem from the fault detection point of view, stating the first step as the one of generating fault indicators, so-called *residuals*. The works focus on dynamic systems, dynamic aspects introducing the main source of complexity. The theory is well established for linear models, although more recent results have been obtained for non-linear models (cf. for example (Frank, 1987; Staroswiecki *et al.,* 1989). In consequence, we do limit our presentation to linear models in the following of the section.

The surveillance task includes two steps:
- Residual generation;
- Residual evaluation.

A residual is the expression of a redundancy introduced by the instrumentation system. For instance, if two sensors are available to measure the same variable, then a residual is obtained by comparing the value provided by each of the sensors. More generally, if there are several computation paths to provide the estimation of a given variable, then there are as many residuals. Evidently, this raises the essential problem of the instrumentation system design, in other words the problem of determining the sensor number and placement so as to achieve diagnosability and a given redundancy degree (Carpentier *et al*., 1997; Gissinger *et al.,* 2000; Van der Velde and Carignan, 1984; Basseville *et al.,* 1987; Travé-Massuyès *et al.,* 2001).

The value of a residual is ideally zero in the nominal case and non-zero in a faulty situation. The residual expressions are obtained a priori, as a function of observable variables (residual generation step). It is then evaluated (possibly on-line) using the measured values (residual evaluation step). A decision procedure, which takes into account the statistical characteristics of measurement noises and modelling errors, is necessary in general to evaluate residuals, i.e. to decide whether their value is zero or not. A spectrum of statistical approaches is available for this purpose (Basseville and Nikoforov, 1993).

Two main streams are known for generating residuals. One of them stands on monitoring the variables that are relevant for describing the system, i.e. the evolutional quantities accounting for the system state across time. The output estimation and the parity space approaches belong to this stream. The other one stands on monitoring the structural parameters of the system, whose values should ideally remain constant, and can be obtained by parameter estimation techniques (Escobet and Travé-Massuyès, 2001). In the linear case, it has been proved that the three methods are equivalent in the sense that one can always find a mathematical transformation to turn the residuals obtain by one method into the residuals obtained by the other. Residual generation also depends on the type of model that is used (state space model in the form of a set of differential and algebraic equations, input-output model given a set of transfer functions, graphical model (structural, bond-graph)).

2.3.1 Robustness

It is generally required that the residuals are robust to system's disturbances, modelling errors and noises (Gertler and Kunwer, 1993). *Decoupling techniques* are used for this purpose. These techniques position the residuals in subspaces, which are not affected by elements to be rejected. In the two first cases, it is possible to perform a perfect decoupling if there are few disturbances or if the modelling errors concern a restricted number of parameters. In the contrary, optimisation techniques are used to find a partial decoupling. Two main methods can be mentioned: the reduction of the rank of the matrix concerning the disturbances and/or modelling errors, the minimisation of a quadratic performance index under equality constraints. For the case of noises, which are generally assumed to have a zero mean and a high frequency band, it is preferred to filter the residual signals with low-pass band filters or to base the test with respect to a threshold around zero.

2.3.2 The Localisation Task

The localization task is accounted for at the residual generation step itself: the idea is to generate " intelligent " residuals, in the sense that they should include information about the components that are responsible for their alarming state (transition to a non zero value). The design of " intelligent " residuals complexity depends on the approach and the type of model. However, the basic principle is always the same: projection operators allow one to project the signals onto well-chosen sub-spaces. Two approaches are available: the *structured residuals* and the *directional residuals* approaches. In both cases, the method is applied to a set of residuals, the number of available residuals defining the fault discrimination degree.

Structured residuals are designed so that each residual is sensitive to a subset of faults and insensitive to the other faults. Hence, when a fault occurs, a subset of residuals transition to a non-zero value whereas the others remain zero: this pattern is known as the *fault signature*. Given a set of anticipated faults, the set of predefined fault signatures forms the so-called *signature matrix*. The minimal requirement is that the residuals are weakly isolating, i.e. that every fault signature (under no disturbance assumption) is different and non-zero. Then, the observed signature is compared to the predefined fault signatures and the corresponding fault is obtained from matches. Note that residuals are defined as dynamic quantities. Their evaluation may be dependent on input exiting conditions (transitory states versus equilibrium states).

Directional residuals are so designed that, in response to a particular fault, the residual vector is confined to a fault-specific straight line, at all times including transients. With directional residuals, fault isolation amounts to determining to which pre-defined fault direction the observed signature lies the closest.

A set of residuals in which every residual responds to one and only one fault is structured and directional at the same time. Such a residual set, so-called diagonal, is important for multiple faults. Indeed, under the common assumption that multiple fault effects do not interact, a different signature is directly obtained for every multiple fault. This limits however the number of faults to the number of independent outputs. Alternatively, a different signature can be defined for every anticipated multiple fault but this requires a greater number of residuals.

2.3.3 Sensitivity to faults, robustness w.r.t. disturbances

Let us note that in practice it is desired to structure or direct the residuals while minimising/maximising the residuals sensitivity with respect to specified fault subsets. In addition, let us note that the disturbance decoupling problem comes back to turn the residual vector insensitive to disturbances. This latter problem is hence a particular case of residual structuring. Hence, in practice, the whole problem is formulated as an optimisation problem whose goals are to generate a set of residuals sensitive to faults in a structured way, and robust with respect to disturbances and model uncertainties.

2.3.4 Parameter estimation approach

The idea is to monitor the system's structural parameters. The mathematical relations linking the variables are characterized by parameters whose nominal values are known and should remain constant in large operational domains; estimation techniques hence allow one to follow the evolution of these parameters and to detect significant variations (Basseville and Nikiforov, 1993; Isermann, 1989; Escobet and Travé-Massuyès, 2001).

The system model in discrete form is given as follows:

$$x_{t+1} = f(x_t, u_t, \theta s_t, v_t)$$
$$y_t = g(x_t, \theta c_t, \varepsilon_t)$$

where x_t and u_t and y_t are the state, input and output vectors at instant t, respectively, v_t and ε_t are input and output disturbances and θs_t and θc_t are the system and sensor parameter vectors respectively. Parameters are estimated as a function h of input and output variables at present and past instants:

$$\begin{pmatrix} \hat{\theta} s_t \\ \hat{\theta} c_t \end{pmatrix} = h(y_t, y_{t-1}, \ldots, u_t, u_{t-1}, \ldots)$$

In this approach, fault localization and identification appear as by-products of fault detection if the reference is taken as the *physical* model of the system. Indeed, a physical model is obtained from physics' principles and makes explicit the physical parameters contributions. On the contrary, if a black-box model, obtained by identification techniques, is used, then the diagnosis information critically depends on the knowledge of the relation Φ between identified parameters and physical parameters.

In each situation, the residuals are given by:

1) *Physical parameters residuals*

$$\Delta\varphi_t = \begin{pmatrix} \varphi s_t \\ \varphi c_t \end{pmatrix} - \begin{pmatrix} \Phi^{-1}(\hat{\theta} s_t) \\ \Phi^{-1}(\hat{\theta} c_t) \end{pmatrix}$$

2) *Model parameters residuals*

$$\Delta\theta_t = \begin{pmatrix} \hat{\theta}s_t \\ \hat{\theta}c_t \end{pmatrix} - \begin{pmatrix} \theta s_t \\ \theta c_t \end{pmatrix}$$

Let us mention that a necessary condition to perform good parameter estimation is that the input signals reach a reasonable excitation level.

2.3.5 Output estimation, observers

The idea is to monitor the system's variables, as quantities being representative of the system's state. Let us consider the system model in a state space form:

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$

where *x(t)* is the state vector of dimension *n*, *u(t)* the input vector of dimension *m* and *y(t)* the output vector of dimension *p*. A, B and C are the dynamic, input and output matrices of appropriate dimensions.

The state of a (dynamic) system is known at a given instant *t* if the values of all the so-called *state variables x(t)* are known. Unfortunately, it is generally not possible to measure all the state variables and there exists a relation, given by the matrix *Y*, between the output (or measured variables) *y(t)* and the state variables *x(t*). Some state variables hence remain as *unknown variables*. The idea underlying this approach is to use an observer *to estimate the unknown variables* from measured input and output signals.

The structure of such an observer is given as:

$$\frac{d\hat{x}}{dt} = A\hat{x}(t) + Bu(t) + He(t)$$
$$e(t) = y(t) - C\hat{x}(t)$$

e(t) is the output error. The state error $\tilde{x}(t)$ (difference between the real and estimated state) is then given by:

$$\frac{d\tilde{x}(t)}{dt} = \frac{dx(t)}{dt} - \frac{d\hat{x}(t)}{dt} = (A - HC)\tilde{x}(t)$$

When the observer is stable, which is obtained by an appropriate choice of the observer feedback gain matrix *H*, the state error tends asymptotically towards zero.

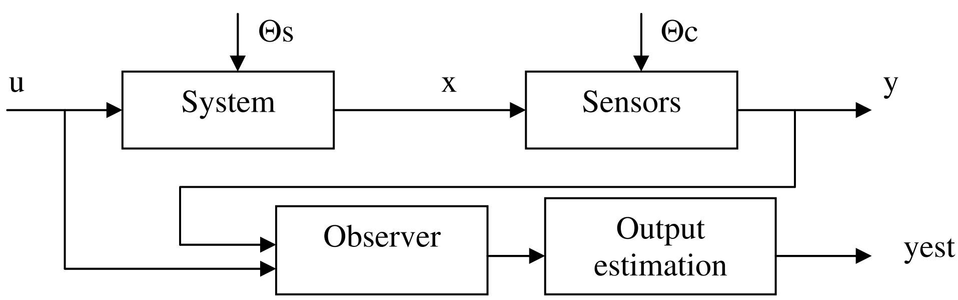

Figure 5: *Observer*

Two types of faults are considered:

1) *Additive faults* influence *x(t)* and *e(t)* through summation. They are considered as unknown inputs acting on the system, which are normally zero and which, when present, cause a change in the system outputs independent of the known inputs. They are represented in the model in the same way as additive disturbances as follows:

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) + F\nu(t) + Lf_L(t)$$
$$y(t) = Cx(t) + N\vartheta(t) + Mf_M(t)$$

$\nu(t)$: non-measurable input disturbances;
$\vartheta(t)$: non-measurable output disturbances;
$f_L(t)$: input faults, acting on *x(t)* through *L* (e.g. actuator fault);
$f_M(t)$: : output faults, acting on *x(t)* through *M* (e.g. sensor fault).

Let us notice that the knowledge about faults is put into matrices *L* and *M*. Moreover, let us outline that a fault has exactly the same status as a disturbance.

Additive faults act both on the state error $\tilde{x}(t)$ and the output error *e(t)*. Moreover, their action is different for $f_L(t)$ and $f_M(t)$ as shown below:

$$\begin{cases} e(t) = C\tilde{x}(t) + Mf_M(t) \\ \dfrac{d\tilde{x}(t)}{dt} = (A - HC)\tilde{x}(t) + Lf_L(t) - HMf_M(t) \end{cases}$$

These two errors can hence be taken as residuals.

2) *Multiplicative faults* are changes (abrupt or gradual) in some system parameters. They cause changes in the system outputs which depend also of the know inputs. They are represented in the model as variations ΔA, ΔB and ΔC of the *A*, *B* and *C* matrices parameters as follows:

$$\frac{dx(t)}{dt} = (A + \Delta A)x(t) + (B + \Delta B)u(t)$$
$$y(t) = (C + \Delta C)x(t)$$

The action of multiplicative faults on $\tilde{x}(t)$ and *e*(t) is not as obvious as the one of additive faults as shown below:

$$\begin{cases} e(t) = C\tilde{x}(t) + \Delta Cx(t) \\ \dfrac{d\tilde{x}(t)}{dt} = (A - HC)\tilde{x}(t) + (\Delta A - H\Delta C)x(t) + \Delta Bu(t) \end{cases}$$

This limits the scope of this approach.

Several classical architectures based on observers have been proposed for multivariable systems:

- Observer driven one output: only one system output is used to estimate all the remaining ones, which are then compared to their measure. This is suited to unique sensor fault detection (Clark, 1978a).
- Bank of observers, each one driven by one different system output: this generalises the precedent principle and allows one to detect multiple sensor faults (this architecture is known as DOS which stands for *Dedicated Observer Schema*) (Clark, 1978b).
- Bank of observers, each one driven by all the system outputs but one, whore corresponding sensor is monitored (Frank, 1987).

The observer feedback gain matrix H must be so designed that the global system (the plant and its observer) is stable. The remaining freedom can be used together with a filter W such as:

$$r(t) = We(t)$$

to minimise residual sensitivity to disturbances (decoupling with respect to environmental disturbances and modelling errors) (Frank, 1990; Patton and Chen, 1991) and to perform a transformation which structures or orients the residual set.

If signals are stochastic, Kalman filters are used instead of observers but Kalman filters not provide additional freedom degrees, as they are all already allocated to the optimal estimation of the system state. In this case, the evaluation of the residuals *e(t)* (filter innovation process) requires a statistical hypothesis test to determine whether statistical parameters of *e(t)* have deviate from normal.

2.3.6 Parity space approach

The simplest way to generate residuals is to use directly an input-output model of the system in the following form:

$y(t)=S(z)u(t)$ or equivalently $H(z)y(t)=G(z)u(t)$

The input and output vectors are *u(t)* and *y(t), S(z)* is the transfer, whose elements are rational functions of the *z* variable, whereas matrices *H(z)* and *G(z)* are polynomial matrices, *H(z)* being diagonal. Let us consider input additive faults, *p(*t), and output additive faults, *q(t),* as well as multiplicative faults (parameter variations) we obtain:

$$H(z)y(t)=G(z)u(t)+H(z)q(t)+F(z)p(t)+E(t)$$

Reorganising the set of equations of the model so that they give zero under nominal conditions, we obtain the so-called *primary parity equations*:

$$\begin{aligned} r(t) &= H(z)y(t)-G(z)u(t) \\ &= H(z)q(t)+F(z)p(t) \end{aligned}$$

The first line is known as the residual *computation form*, which expresses the residuals as a function of observable variables, whereas the second line is the *internal form*, which expresses the residuals as a function of faults.

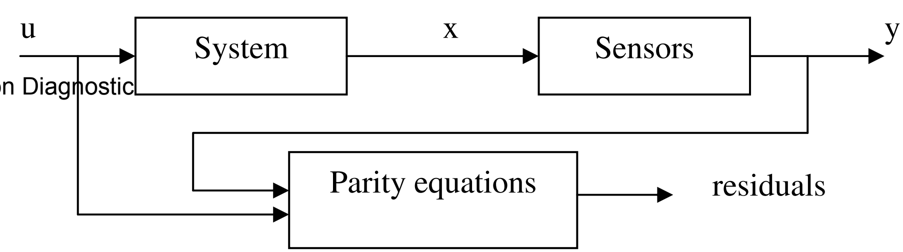

Figure 6: *Parity space input-output model*

When multiplied by numbers, polynomial or rational functions in z, parity equations (also called *analytical redundancy relations*, or ARR for short) remain parity equations. An appropriated linear transformation can hence be used to provide the residuals with interesting properties (structuring, decoupling, etc.):

r'(t)=W(.)r(t)

where *W(.)* is a constant polynomial or rational matrix *W(z)* if we deal with additive faults and disturbances or a time varying numeric matrix *W(t)* if we deal with multiplicative faults (Staroswiecki *et al.,* 1993).

Let us notice that parity equations can also be obtained from a state space model of the system by eliminating the state variables *x(t)*, which makes explicit the principle of this method, which is *to eliminate the unknown variables*. This can be performed by a substitution method or by a projection method. The initial model is given by:

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$

where *x(t)* is the state vector of dimension *n*, *u(t)* the input vector of dimension *m* and *y(t)* the output vector of dimension *p*. *A, B* and *C* are the dynamic, input and output matrices of appropriate dimensions.

According to a projection method, let us derive *n* times the second equation and substitute the expression of $\frac{dx(t)}{dt}$, which provides:

$$\begin{pmatrix} y \\ \dot{y} \\ \ddot{y} \\ \vdots \\ y^{n} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^n \end{pmatrix} x + \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ CB & \ddots & & & \vdots \\ CAB & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ CA^{n-1}B & \cdots & CAB & CB & 0 \end{pmatrix} \begin{pmatrix} u \\ \dot{u} \\ \ddot{u} \\ \vdots \\ u^{n} \end{pmatrix}$$

which can be written:

$$Y = OBS \times x + M \times U$$

by defining *Y* and *U* as the vectors composed from *y* and *u* and their n derivatives. *x* can be eliminated if we find a matrix *P* such that $P \times OBS = 0$. Then, we have:

PY=PMU

And the residuals are given by:

r(t)=PY-PMU

A survey of the parity space approach is available in (Patton and Chen, 1991).

2.3.7 More recent results

Following the success of techniques known as “ soft computing ” techniques which gather fuzzy set theory, neural and evolutional optimisation (e.g. genetic algorithms) based techniques, there has been a considerable amount of work using these techniques. The “ knowledge based ” approach has been intensively developed and the use of qualitative models and qualitative simulation (as used in AI) has been proposed. However, this works remain based on the above-presented three approaches (parameter estimation, observers and parity space). A survey of these recent developments is available in (Frank, 1996).

Fuzzy sets provide an interesting alternative to residuals evaluation, as they allow one to obtain a smooth evaluation. For instance, architectures integrating fuzzy models and fuzzy observers *(“ knowledge observer ”)* have been proposed. A survey of the use of fuzzy concepts is available in (Isermann, 1995).

Qualitative models have been used to obtain qualitative observers, implementing QSIM-like qualitative simulation techniques. DOS architectures have been proposed (Frank, 1996).

Neural networks, which are indeed universal classifiers, are well known in the pattern recognition diagnosis area but they can as well be used in model-based architectures, in the residual generation step as well as in the residual evaluation step. Several contributions have presented how neural networks can be used as an alternative model of the system, which can then handle non-linear aspects, or as the basis of a decision module, which switches linear model according to the system operating state. The fuzzy and neural approaches can also cooperate (Ayoubi, 1996).

3 Comparative analysis

The DX and FDI communities are two distinct and parallel research communities concerned with the diagnosis of technical systems.

Both DX and FDI recognise themselves as model-based diagnosis approaches and they indeed stand on similar principles, in particular the use of a model to represent the physical system. However, the FDI community has evolved from the Engineering world and uses techniques from control theory and statistical process control. The DX community is more recent. Its foundations have been derived from the fields of Computer Science and Artificial Intelligence. Each community has developed its own tools, techniques and approaches. Each also has its own set of conferences, publications and terminology. There are however growing numbers of researchers, in both research communities, who are trying to understand and incorporate the approaches of their parallel research field.

Pioneering work along the comparison of approaches on both sides has been performed by the French IMALAIA[1] Group, which has recently proved the equivalence of several concepts (Cordier *et al*., 2001a, 2001b). At the European level, the BRIDGE Group within the MONET European Network of Excellence (http://monet.aber.ac.uk) also works in favour of exchanges and collaborations between both communities.

Although the basic principles are similar, the reference backgrounds of the DX and the FDI communities are at the origin of several differences between the corresponding approaches. These are discussed in the following sections.

3.4 Detection versus localization point of view

The first key point to be outlined is that:

FDI tackles the diagnosis problem from the fault detection point of view, formulated as the generation of a set of residuals (more precisely, of their formal expression), which should have the required fault sensitivity and robustness with respect to all types of anticipated disturbances. In doing so, properties that are useful for fault localization are also considered (structured or directional residuals).

DX takes the opposite point of view, tackling the diagnosis problem from the fault localization algorithmic problem and considering that fault detection relies on a theorem prover, which detects inconsistencies and is not part of the diagnosis theory. Historically, the modelling issues have also been considered separately, in the Qualitative Reasoning research area. This dichotomy has reduced since modelling was recognised as an essential step.

This different starting angle has numerous consequences, which are presented in the following.

3.5 Disturbances, modelling errors, noise

Essentially concerned by fault detection, the FDI community uses precise numeric models, which have better sensitivity to variations, and takes explicitly into account disturbances (additive unknown inputs or outputs), modelling errors (stated as multiplicative disturbances) and noises. These are accounted for in the design of robust residuals based on decoupling techniques and signal filtering techniques.

In DX, the problems related to occurrence of disturbances, modelling errors and noise are in general “ ignored ”, as being naturally rejected by the level of abstraction of the models (symbolic or qualitative). Semi-qualitative models (e .g. interval models), which are closer to numeric models, require special care about these problems. So far, the working assumption (sometimes implicit) has been that interval bounds are chosen so that disturbances, modelling errors and noise are absorbed.

3.6 ARR versus conflicts

Let us analyse more precisely the correspondence between the concepts used on each side. In both sides, limiting the analysis to the use of normal behaviour models, diagnosis is

[1] IMALAIA *(“ Intégration de Modèles Alliant Automatique et IA ”)* is a group which is supported by CNRS, as a project of *GDR Automatique* and *GDR I3.*

triggered by the inconsistencies between the modelled (correct) behaviour and the observations. As presented in section 3.2, in the logical theory, the generation of diagnosis candidates is driven by the identification of R-conflicts. In the FDI parity space approach, the inconsistencies arise from the ARR that are not satisfied by observations (hence leading to non zero residuals). The *support* of an ARR is naturally defined as the set of components whose model is involved in the ARR formal expression (Cordier *et al.,* 2000). If the components of the support are not faulty, the ARR is satisfied by any observations. On the other hand, the violation of an ARR by a set of observations incriminates at least one of the components of its support. The ARR supports can hence be identified as *potential conflicts*, in the sense that they do become R-conflicts whenever the corresponding ARR is violated by a set of observations. The notion of (ARR, support) redundant pair, in the sense that this pair is a logical consequence of analogous pairs, can be precisely defined, giving rise to the concept of *minimal potential conflict*, although to be distinguished from the minimal R-conflict concept (Cordier *et al.,* 2000).

3.7 Off-line processing approach versus on-line processing approach

In the FDI parity space approach, the ARRs are determined off-line, which is equivalent to compile potential conflicts. Nevertheless, the component correct behaviour assumptions are not represented explicitly in the models. Hence, the determination of the ARR supports is either performed by hand or by graph theory techniques (Cassar & Staroswiecki, 1997). Then, on-line, the observed signature is determined (this is the detection step, which decides for every ARR whether it is satisfied or not by the observations) and the diagnosis candidates (localization step) arise from a simple comparison of the observed signature and the fault signatures. On the other hand, the DX community performs on-line propagation instantiated by the observations in the models and R-conflicts are automatically deduced, during the propagation, from the explicit statement of the assumptions supporting the validity of the component models. This is generally performed by means of an ATMS coupled with the inference engine. The minimal diagnosis sets are obtained as the R-conflicts hitting sets.

In summary, the FDI parity space approach performs a sophisticated off-line processing and only consistency checking tests remain for the on-line phase, which uses the observations. On the contrary, the DX approach is deductive in the sense that it uses a diagnosis engine which, totally decoupled from the models, the system structure and observations, implements the on-line diagnosis reasoning.

Let us notice that the two approaches can be combined to merge mutual advantages. Recent contributions in the DX community are based on this idea (Loiez, 1997; Pulido, 2001; Washio *et al.,* 1998; Katsilis and Chantler, 1997).

3.8 System and instrumentation changes

A deductive approach, such as the DX approach, is robust to system changes (e.g. multiple operational mode systems such as hybrid systems) and instrumentation changes (addition of one sensor, additional tests performed to refine the diagnosis results). The works concerning the choice of the best next measurement have no equivalent in the FDI community. On the other hand, in such situations, a compiled approach such as the FDI parity space approach would require determining again the new ARR set. An observer approach would require redesigning the observer in order to achieve the desired properties for the residuals. The same stands for the parameter estimation approach, whose algorithm parameters need to be tuned according to the parameters to be estimated and the available input-output signals. The FDI approaches would require anticipating all the possible configurations of the system and its instrumentation.

3.9 Models granularity

The FDI approaches consider the faults just as disturbances, acting either on model parameters (multiplicative faults) or on input and output signals (additive faults). Faults are hence generally defined with respect to a nominal behaviour model and the use of " real " fault models remains infrequent. A fault is rather an abstract concept, which may not have a direct correspondence with physical components, in particular for identified models. In DX, considering the consistency-based approach, only a correct behaviour model is used, a fault being detected when observations are inconsistent with the model. In this case, no assumption is made about the faults. In the extension including fault models, explicit faulty behavioural models are used and their consistency is tested against the observations following the same principle. This allows one to discard some faults. The faults generally considered in DX are abrupt, changing radically a component's behaviour. Indeed, this is easily represented in high abstraction level formalisms as uses in IA.

Let us outline that *semi-qualitative models* provide an interesting compromise between numeric models and qualitative models. They indeed keep numeric aspects while allowing one to represent imprecision.

For non-linear models, the FDI analytic approach requires powerful mathematical tools, which may be difficult to use, this being a consequence of working with precise models. As a matter of fact, the set of results now well established is concerned with linear models. This is in accordance with the fact that the FDI community generally formulates the surveillance problem as the surveillance of a nominal operating mode for which a linearized model can be obtained. This strengthens the idea that the FDI approaches are better suited to degradation faults, which manifest as deviations with respect to the nominal behaviour. In DX, there are no limitations in the theory itself but the modelling formalisms (symbolic, qualitative) lead to the inverse conclusion.

3.10 Compositional aspect of the modelling

In the DX field, compositional modelling has always been considered as essential, for obvious reusability and maintenance reasons. A given application domain is characterised by basic components (e.g. resistance, capacity, transistor, etc. for the electronic domain) for which a library of models is built. The global model of a system, for which the structure is given, is then assembled from the model " fragments " in the library. This step is made automatic. The same library can hence be used to build the model of any system. Let us note however that compositional modelling is submitted to the condition that the component behaviours are independent of the context. This condition may be violated when considering precise models, as used in FDI, because the connection of two components can as well induce interacting effects. These should hence be preferred for the surveillance of individual equipments or for the surveillance of high-level functions modelled as a whole.

The compositional modelling aspect has a key importance when dealing with hybrid systems. Indeed, it allows one to automatically build the model after an operational mode transition or a reconfiguration.

3.11 Fault assumptions

The DX logical diagnosis theory does not include any assumption about the faults, which provides logically sound diagnosis conclusions. In the default case, multiple faults are handled as single faults, a fault may as well be non observable (i.e. have no associated symptom), and several faults may as well compensate, producing no observable effects. When the application domain suggests some assumptions, these are explicitly taken into account by additional axioms (i.e. the *exoneration* axiom). On the contrary, the FDI approach generally adopts the single fault assumption with observable effects. If multiple faults are considered, it is generally assumed that they have no interacting effects (implicit non-compensation assumption). When comparing DX and FDI approaches, a special attention must hence be put on clarifying the assumptions, as they are different by default.

The approaches can also be combined from this angle : the FDI approach can indeed be enhanced by explicitly stating fault assumptions and by the techniques for generating diagnosis from conflict sets.

3.12 Models explanation capacity

Numeric models have very little explanation power whereas qualitative models have been thought to capture intuitive features. In particular, the causality links between variables in a model are often taken into account in abductive reasoning schemas. These can be represented by a causal structure, which supports causal qualitative models (Travé-Massuyès and Pons, 1997). Similar structures are used in FDI (Staroswiecki *et al*., 1989) for the design of residuals, that is at the off-line processing stage.

The compositional aspect of modelling as well as the explicit formulation of working assumptions is also important for explanation purposes.

3.13 Dynamic systems

Dynamic systems are at the core of the control community topics. They have been intensively studied, providing a well-established theory. Given this background, the FDI approaches naturally take into account dynamic aspects. Residuals are considered as dynamic quantities (temporal signal) and time is explicitly stated in the models. In DX, historically, the basic theory does not include time in the diagnosis process. However, the practical problems of evaluating a temporal residual exist in a similar way (i.e. state-based diagnosis versus simulation-based diagnosis (Struss, 1997)).

3.14 Discrimination among several candidates versus diagnosability conditions

The problem of discriminating among several candidates is not formulated in the FDI community. The set of faults to be considered is anticipated and fault probabilities are taken into account at this level (i.e. a fault with very low probability is not included in the set). Then, residuals are so designed that they are at least weakly isolating, i.e. that every fault has a different signature. If this is not possible, the problem is formulated as determining the sensor(s) to be added to obtain the desired discrimination level. This is known as the diagnosability problem related to the design of the instrumentation system (Travé-Massuyès *et al*., 2001; Carpentier *et al.,* 1997; Gissinger *et al,* 2000). Let us notice that, on the inverse, this later problem has been rarely considered in the DX community. Now, in DX, given a poor diagnosability level, the theory is based on generating all the alternative diagnosis candidates and proposes tools to give probability or preference degrees. Hence, in practice, it is possible to generate only the most plausible according to some criteria (fault probability, explanation degree, etc.).

4 Comparative analysis summary

FDI	DX
Looks for one diagnostic	Looks for a set of diagnostics

Centered on detection	Centered on diagnosis
Fault and variable centered	Component centered
Must enumerate faults and their effects	Enumeration of possible faults and their effects not required
Very good fault detection based on model based predictor	Efficient algorithms for fault isolation
Faults are considered as deviations w.r.t. normal behavior (generally additive deviation faults)	All kind of faults are considered within the same framework: deviations w.r.t the normal behavior (unknown fault mode) and explicitely represented fault models
No representation of an “unknown fault” mode, so runs into a problem in case of a non anticipated fault.	Closed World Assumption fulfilled by the systematic account of an “unknown fault” mode for every component.
Implicit (model and fault) assumptions	Explicit (model and fault) assumptions
Model parameter estimation methods	Model compositionality
Models are compiled into residuals, detection filters...	Models support the diagnosis reasoning process
Interval models are used on both sides	
One operating mode systems	Multiple operating mode systems
Robustness issues: noise, disturbances, modeling errors	High level of abstraction models (qualitative)
FDI pays great attention on detection, then goes directly to fault identification.	DX put its focus on fautl isolation
Systems with complex dynamics	Systems complexity is on the number of components
The diagnosis reasoning is embedded within the model: residuals generation, DOS (Dedicated Observers Sheme), ...	The diagnosis reasoning is automated: “dependency-recording” techniques
The approaches are corroboration-based, hence not logically sound.	The reasoning is refutation based (in its consistency framework) hence logically sound
Generally only single fault. If multiple faults are considered, their effects must be anticipated.	Deals with multiple faults within the same framework
Off-line structural and symbolic computation (at least in the linear case) from the model itself before any observation is known	On-line propagation techniques (with observations) on-line propagation techniques (with observations)
Explanation is not accounted for in FDI methods which uses some mathematical transformation where intuition can be lost.)	Providing an explanation understandable by human operators is important in DX (comes from the cognitive part of AI). This leads DX to use sometimes causal models etc, nearer from explanation.
Time is naturally taken into account.	Dealing with time in DX is more recent.
In both approaches, a lot is done by hand and not still automated. (models in DX, residuals or observors in FDI…)	
Principles are similar. Some concepts can be bridged, e.g. ARR supports correspond to possible R-conflicts.	

5 Characterization of a diagnosis problem: table of dimensions

This section derives a series of dimensions that can be used to characterize a diagnosis problem. The table indicates in the two right hand side columns the field (DX or FDI) which deals more typically with the topic (note that this does not mean that the other field does not approach the topic).

	DX	FDI
TASK		
Monitoring - Fault detection	-	X
Fault localisation	X	-
Fault identification – State Assessment	X	X
Preventive/Corrective actions advice	X	X
Objective reconfiguration and planning	-	X
System's structure reconfiguration	-	X
Fault accommodation	X	-
Alarm filtering	X	-
MODEL : PHYSICAL SYSTEM REPRESENTATION		
Analytical quantitative models	-	X
Constraint-based qualitative models	X	-
Interval models	X	X
Causal influence models	X	X
Discreet Events models	X	X
Chronicles	X	-
Hybrid models	X	-
Cause to symptoms causal graphs	X	-
MODEL: TIME REPRESENTATION		
Continuous	-	X
Sampled	-	X
Asynchronous (event-triggered)	X	-
TEMPORAL FEATURES OF THE SYSTEM		
System internal dynamics	-	X
Multiple operating modes	X	-
Multiple configurations	X	-
Wear	-	X
SYSTEM'S OBSERVABLES		
Continuous signals	-	X
Events	X	X
Sensors or tests (predefined or variable number and location)	X	-
TYPE OF FAULTS		
Structural fault (abrupt)	X	-
Drift fault (incipient)	-	X
Configuration fault	-	-
ASSUMPTIONS ABOUT FAULTS		
Single fault assumption	-	X
Multiple faults assumption	X	-
Single fault exoneration (no masking)	-	X
Multiple faults exoneration (no compensation)	-	X

	DX	FDI
MODELS USE		
Prediction	-	X
Consistency check	X	X
Explanation (cause retrieval)	X	-

6 Common Diagnosis Framework

A hierarchical decomposition of the system in four levels seems relevant for the montoring and diagnosis using model-based techniques. DX and FDI model-based diagnosis approaches integrate nicely within this framework:

1. Level 1: signal processing level,
2. Level 2: equipments level, taken individually (component variables monitoring),
3. Level 3: sub-system (functional) level (use of redundancies),
4. Level 4: System level (state of each equipment, consistency of commands and observations with respect to the current system mode).

Each level uses the information from the lower levels. Low levels mostly perform detection whereas higher levels include more sophisticated algorithms to process information to localize the faults.

Level 1 essentially includes signal-processing techniques such as statistical analysis, filtering, etc. Level 2 deals with every component individually, monitoring physical quantities (e.g. temperature of a transistor) and behavioural quantities. The behavioural aspect can as well be tackled with model-based techniques, mainly numeric such as Kalman filters or other kind of state observer, e.g. interval model-based. These two levels mainly include fault detection algorithms, the fault isolation problem being not relevant. Recovery actions are generally simple and can be hard coded. The techniques that seem relevant for these two levels better come from FDI. If hybrid aspects exist for some equipment, the number of behavioural modes should remain low so that compiled approaches (residual generation) are applicable. Qualitative and semi-qualitative models can be interesting for handling behavioural aspects.

The two upper levels are the ones that may require more "intelligence" for dealing with unanticipated situations. From our previous comparative analysis, we do believe that DX techniques are essential at these two levels. Indeed, hybrid aspects that are typical of these levels are better handled by DX techniques.

Hybrid modelling formalisms may provide a significant advantage to unify the modelling through all the levels. A model based diagnosis and recovery module can then in a unified way handle the equipment aspects together with the functional ones and get a global view to perform better state identification. For example, at the two first levels, an alarm on some equipment can as well be the consequence of a fault on another equipment (e.g., cascade effects). The relational analysis leading to the localization of the source cause can only be performed at the functional or even system level.

Hybrid model based diagnosis systems can take advantage of FDI and DX approaches whose foundations are consistent and which can enhance mutually (compilation approach of FDI, logical soundness of DX). It makes the modelling better suited to the different equipment and sub-system to be considered, as it allows one to range the granularity of the models over a spectrum, including pure numerical models, semi-qualitative (interval) models, qualitative and symbolic.

The use of interval models is particularly interesting, as preserving numerical aspects while allowing one to account for uncertainty. These models can advantageously used to generate adaptive thresholds for fault detection (Armengol *et al.,* 1999, 2000) and are given attention in both the FDI and the DX communities.

Although some of the recovery actions can be hard coded, for obvious efficiency and safety purposes, we do believe that a class of actions could be decided in an automated way. DX approaches the problem of structure reconfiguration involving planning of the sequence of recovery actions, wheareas fault accomodation can take advantage of FDI control methods.

7 A few items for transferring the MBD technology into industry

The main gaps and issues between the technical state of the art for model based diagnosis as it exists in the research community and the need of industry to use these technologies were identified in (Travé-Massuyès and Milne, 1998) and this helped find ways to close them.

As a matter of fact, it has been demonstrated that the technologies are capable of helping with many complex problems, and have reached the maturity to be fully applied in industry. Several real world products and advanced systems are now available and many advanced prototypes have been demonstrated (MONET ILC, 1998)(Cauvin et al., 1998).

In spite of their reasonably mature state, the use of MBD techniques by industry has not reached its full potential yet. The problems with transferring this technology into industry are not with the diagnosis mechanisms but rather with the models to be built.

Although model based methods have many advantages, systems modelling, at any level, is inherently difficult. Even with qualitative approaches a high level modelling intelligence is required which does not exist « in the machine ».

The real challenge posed by the modelling is the automation of its acquisition and update and its sharing along the whole life cycle of a system: design, manufacture, test, diagnosis, maintenance. Only some academic works exist about model acquisition, without industrial application today (Amsterdam, 1993; Xia *et al*., 1993; Nayak, 1995; Schut and Bredeweg, 1996; Smith, 1998; Stolle and Bradley, 1998), and preliminary works about sensor placement for diagnosability (Scarl, 1994; Travé-Massuyès *et al*., 2001). Projects have been lauched in 2000 (as the European IDD project, *Integrated Design process for on-board Diagnosis*, in the automotive industry) which consist in studying diagnosticability of a system from its design stage and aim at coupling CAD tools to model-based diagnosis tools and updating automatically diagnostic models, and thus diagnostic facilities, when design models evolve. It would be highly desirable to build models from data that already exist, ideally in an electronic form. This would save the need to re-input the data and possibly to have to redefine the information. Also, engineering models as used in FDI are generally context dependent and they should be bridged to more conceptual and higher abstraction models, as used in DX, by making explicite the underlying hypotheses and performing the adequate abstraction. Although researchers are active on this topic (e;g. the IDD European project and associated research), as a matter of fact, there is no such automatic modelling tool today and the modelling task is a costly and time consuming exercise.

This difficulty has to be balanced with the benefits of the MBD approach and the models. Before launching a model-based diagnosis industrial project, it should be demonstrated that the investment has a low risk and high return. The analysis should evaluate the cost of building an actual standard diagnosis system against the estimated modelling cost. It should also compare the following issues :

- Reusability and genericity: in the model based approach, the diagnosis algorithms are generic and several component models are generally reusable from one application to another. This rises the fundamental problem of the domain specific nature of modelling tasks. Certain components are relatively straight forward to model, store and reuse. The solid state nature of electrical components for example lend themselves fairly well because a discrete numerical description of their function is possible which is in stark comparison to some mechanical components whose functional description can vary continuously with their physical dimensions and overall design. The number of reusable component models should be clearly identified.
- Maintenance: proper planning should be put in place to update the model based system as the primary system evolves. What is ideally required is a method to automatically update the models periodically both as the characteristics of the system change and the components of the system change.
- By-product services: once built, models can be useful in many ways and this benefit should be evaluated. A well-established data library of models could be highly valuable as a knowledge management tool. Standardisation is here one of the major keys. This calls for the choice of a well-established modelling language and of a dedicated friendly user interface. Let us outline that, used independently of the diagnoser, the model can act as a standard simulator and can be useful in many ways. For instance, sensor placement or other design features impacting on fault detection and identification can be simulated and analysed. The model can also be applied to dedicated model-based methods concerned with the diagnosability degree of the system and the automatic optimisation of on-board sensor placements (Travé-Massuyès *et al.,* 2001; Carpentier et al., 1997; Gissinger *et al*., 2000). It should also be noted that the model-based diagnosis techniques themselves can be used to debug the models, hence providing valuable aid to the modeller. The diagnosis engine is then used against well labelled data (normal and faulty). Inconsistencies and resulting diagnosis indicate the component models that need to be revised in order to fit the data.

A good integration of model based techniques with existing software environments is also very important. These issues are not unique to model based systems but are standard problems of any integration of additional software into existing environments.

The success of the modelling stage is related to the team that is created for the purpose. This team should involve design experts, i.e. system's equipment and software design engineers, operation experts, and model-based diagnosis experts. Careful management of the team is still needed to ensure success.

Let us outline that, used independently of the diagnoser, the model can act as a standard simulator and can be useful in many ways. For instance, sensor placement or other design features impacting on fault detection and identification can be simulated and analysed. The model can also be applied to dedicated model-based methods concerned with the diagnosability degree of the system and the automatic optimisation of on-board sensor placements (Travé-Massuyès *et al.,* 2001; Carpentier et al., 1997; Gissinger *et al*., 2000).

It should also be noted that the model-based diagnosis techniques themselves can be used to debug the models, hence providing valuable aid to the modeller. The diagnosis engine is then used against well labelled data (normal and faulty). Inconsistencies and resulting diagnosis indicate the component models that need to be revised in order to fit the data.

8 References

Model-Based Diagnosis in Artificial Intelligence

Aghasaryan A., Dousson C., Mixing chronicle and Petri net approaches in evolution monitoring problems, working notes of the 12th International Workshop on Principles of Diagnosis DX-01, Sansicario, Italy, 2001.

Allen J. F., Maintaining Knowledge about Temporal Intervals, Comm. ACM 26(6), p. 832-843, 1983.

Amsterdam J., Automated qualitative modeling of dynamical physical systems, PhD dissertation, AI Lab, MIT, 1993.

Autio K., Structural abstraction in model-based diagnosis, Proceedings of ECAI-98, Brighton, p. 269-273, 1998.

Baroni P., Lamperti G., Pogliano P., Zanella M., Diagnosis of large active systems, Artificial Intelligence 110(1), p. 135-183, 1999.

Baroni P., Lamperti G., Pogliano P., Zanella M., Diagnosis of a class of distributed discrete-event systems, IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans, 30(6), 2000.
Benazera E., Travé-Massuyès L., Un formalisme discret de modélisation pour le dagnostic réactif, Workshop sur la Modélisation des Systèmes Réactifs MSR'01, Toulouse, France, October 2001.

Benazera E., Travé-Massuyès L., Dague P., Hybrid model based diagnosis for Autonomous spacecrafts, ESA Workshop on On-Board Autonomy, ESTEC, Noordwijk, The Netherlands, October 2001.

Bernard D. E., Dorais G. A., Fry C., Gamble Jr.E. B., Kanefsky B., Kurien J., Millar W., Muscettola N., Nayak P. P., Pell B., Rajan K., Rouquette N., Smith B., Williams B. C., Design of the Remote Agent Experiment for Spacecraft Autonomy, IEEE Aerospace 1998.

Bibas S., Cordier M.-O., Dague P., Dousson C., Lévy F. , Rozé L., Alarm driven supervision of telecommunication networks: I-Off-line scenario generation, Annales des télécommunications 51 (9-10), p. 493-500, 1996a.

Bibas S., Cordier M.-O., Dague P., Lévy F., Rozé L., GASPAR: A model-based system for monitoring telecommunication networks, Proceedings of the IMACS IEEE/SMC Multiconference of Computational Engineering in Systems Applications CESA-96/MBS, Lille, 1996b.

Böttcher C., Dressler O., A framework for controlling model-based diagnosis systems with multiple actions, Annals of Mathematics and Artificial Intelligence, special issue on Model-Based Diagnosis, 11(1-4), J.C. Baltzer AG, p. 241-261, 1994.

Bousson K., Travé-Massuyès L, Putting more Numbers in the Qualitative Simulator CA-EN, Int. Conf., on Intelligent Systems Engineering, Hamburg-Harburg, Germany, 62-69, 1994.

Bousson K., Zimmer L., Travé-Massuyès L., Causal Model-Based Diagnosis of Dynamic Systems, 5th Int. Workshop on Principles of Diagnosis, New Paltz, (NY), USA, 1994.

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

Cascio F., Console L., Guagliumi M., Osella M., Panati A., Sottano S., Theseider Dupré D., Generating on-board diagnostics of dynamic automotive systems based on qualitative decisions, Ai Communications 12(1), p. 33-44, IOS Press, 1999.

Cauvin S., Cordier M.-O., Dousson C., Laborie P., Lévy F., Montmain J., Porcheron M., Servet I., Travé-Massuyès L., Monitoring and alarm interpretation in industrial environments, AI Communications Journal 11, p.139-173, IOS Press, 1998.

Cheung J. T.-Y., Stephanopoulos G., Representation of process trends - Part II. The problem of scale and qualitative scaling, Computers Chem. Engng., 14, 4/5, p. 511-539, 1990.

Chittaro L., Functional diagnosis and prescription of measurements using effort and flow variables, IEEE Control Theory and Applications 142(5), p. 420-432, 1995.

Chittaro L., Guida G., Tasso C., Toppano E., Functional and teleological knowledge in the multi-modeling approach for reasoning about physical systems: a case study in diagnosis, IEEE Transactions on Systems, Man and Cybernetics 23(6), p. 1718-1751, 1993.

Chittaro L., Ranon R., Augmenting the diagnostic power of flow-based approaches to functional reasoning, Proceedings AAAI-96, Portland, OR, p. 1010-1015, 1996.

Chittaro L., Ranon R., Automatic derivation of hierarchical representation for flow-based functional models, working notes of the 10th International Workshop on Principles of Diagnosis, Loch Awe, Scotland, 1999.

Chittaro L., Ranon R., Soldati A., Introducing deviations and multiple abstraction levels in the functional diagnosis of fluid transfer systems, Artificial Intelligence in Engineering 12(4), p. 355-373, 1998.
Clancy D., Kurien J., Model-based diagnosis: Real-world applications and the engineering process, Proceedings of ISAIRAS 2001.

Console L., Dressler O., Model-based diagnosis in the real world: Lessons learned and challenges remaining, Proceedings IJCAI-99, Stockholm, p. 1393-1400, 1999.

Console L., Friedrich G. (eds.), Golumbic M.C. (ed.-in-chief), Annals of Mathematics and Artificial Intelligence, special issue on Model-Based Diagnosis, 11(1-4), J.C. Baltzer AG, 1994.

Console L., Picardi C., Ribaudo M., Diagnosis and diagnosability analysis using PEPA, Proceedings ECAI-2000, Berlin, p. 131-135, 2000.

Console L., Portinale L., Theseider Dupré D., Torasso P., Diagnosing time-varying misbehavior: an approach based on model decomposition, Annals of Mathematics and Artificial Intelligence, special issue on Model-Based Diagnosis, 11(1-4), J.C. Baltzer AG, p. 381-398, 1994.

Console L., Theseider Dupré D., Torasso P., A theory of diagnosis for incomplete causal models, Proceedings IJCAI-89, Detroit, MI, p. 1311-1317, 1989.

Console L., Torasso P., Integrating models of the correct behavior into abductive diagnosis, Proceedings ECAI-90, Stockholm, p. 160-166, 1990.

Console L., Torasso P., A spectrum of logical definitions of model-based diagnosis, Computational Intelligence, 7(3), p. 133-141, 1991a.

Console L., Torasso P., On the co-operation between abductive and temporal reasoning in medical diagnosis, Artificial Intelligence in Medicine 3(6), p. 291-311, 1991b.

Cordier M.-O., Largouët C., Using model-checking techniques for diagnosing discrete-event systems, working notes of the 12th International Workshop on Principles of Diagnosis DX-01, Sansicario, Italy, 2001.

Cordier M.-O., Thiébaux S., Event-based diagnosis for evolutive systems, working notes of the 5th International Workshop on Principles of Diagnosis DX-94, New Paltz, NY, 1994.

Dague P., Model-based diagnosis of analog electronic circuits, Annals of Mathematics and Artificial Intelligence, special issue on Model-Based Diagnosis, 11(1-4), J.C. Baltzer AG, p. 439-492, 1994.

Dague P., Jehl O., Taillibert P., An interval propagation and conflict recognition engine for diagnosing continuous dynamic systems, International Workshop on Expert Systems in Engineering, Vienne, in Lecture Notes in Artificial Intelligence 462, GottlobG., Nejdl W. (eds.), Springer-Verlag, p. 16-31, 1990.

Darwiche A., Model-based diagnosis using structured system descriptions, Journal of Artificial Intelligence Research 8, p. 165-222, 1998.

Darwiche A., Utilizing device behavior in structure-based diagnosis, Proceedings of IJCAI-99, Stockholm, p. 1096-1101, 1999.

Davis R., Diagnostic reasoning based on structure and behavior, Artificial Intelligence 24(1), p. 347-410, 1984.

Debouk R., Lafortune S., Teneketzis D.C., Coordinated decentralized protocols for failure diagnosis of discrete-event systems, Journal of Discrete Event Dynamical Systems: Theory and Applications 10(1-2), p. 33-86, 2000.

De Coste D., Dynamic across time measurement interpretation, Proceedings AAAI-90, Boston, MA, p. 373-379, July-August 1990.

de Kleer J., Williams B. C., Diagnosing multiple faults, Artificial Intelligence 32(1), p. 97-130, 1987.

de Kleer J., Williams B. C., Diagnosis with behavioral modes, Proceedings IJCAI-89, Detroit, MI, p. 1324-1330, 1989.

de Kleer J., Using crude probability estimates to guide diagnosis, Artificial Intelligence 45(3), p. 381-391, 1990.

de Kleer J., Focusing on probable diagnoses, Proceedings AAAI-91, Anaheim, CA, p. 842-848, 1991.

de Kleer J., Mackworth A., Reiter R., Characterizing diagnoses and systems, Artificial Intelligence 56(2-3), p. 197-222, 1992.

de Kleer J., Raiman O., Shirley M., One step lookahead is pretty good, working notes of the 2nd International Workshop on Principles of Diagnosis DX'91, p. 136-142, Milan, 1991.

Dousson C., Alarm driven supervision for telecommunication networks: II-On-line chronicle recognition, Annales des Télécommunications 51(9-10), p. 501-508, 1996.

Dressler O., On-line diagnosis and monitoring of dynamic systems based on qualitative models and dependency-recording diagnosis engines, Proceedings ECAI-96, Budapest, p. 481-485, 1996.

Dressler O., Freitag H., Prediction sharing across time and contexts, Proceedings AAAI-94, Seattle, WA, p. 1136-1141, 1994.

Dubuisson B. (ed.) Diagnostic, intelligence artificielle et reconnaissance des formes, traité IC2, Hermes, 2001a (in particular the chapter 1: Dague P., Théorie logique du diagnostic à base de modèles, p. 17-105).

El Fattah Y., Provan G., Modeling temporal behavior in the model-based diagnosis of discrete-event systems (a preliminary note), working notes of the 8th International Workshop on Principles of Diagnosis DX'97, Mont Saint Michel, France, 1997.

Friedrich G., Gottlob G., Nejdl W., Formalizing the repair process, Proceedings ECAI-92, Vienne, p. 709-713, 1992.

Friedrich G., Gottlob G., Nejdl W., Formalizing the repair process – extended report, Annals of Mathematics and Artificial Intelligence, special issue on Model-Based Diagnosis, 11(1-4), J.C. Baltzer AG, p. 187-201, 1994.

Friedrich G., Lackinger F., Diagnosing temporal misbehaviour, Proceedings IJCAI-91, Sydney, p. 1116-1122, 1991.

Friedrich G., Nejdl W., Choosing observations and actions in model-based diagnosis/repair systems, Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning KR-92, Cambridge, MA, p. 489-498, 1992.

Gamper J., Nejdl W., Abstract temporal diagnosis in medical domains, Artificial Intelligence in Medicine 10(3), p. 209-234, 1997.

Goldstone D. J., Controlling inequality reasoning in a TMS-based analog diagnosis system, Proceedings AAAI-91, Anaheim, CA, p. 512-517, 1991.

Greiner R., Smith B. A., Wilkerson R. W., A correction to the algorithm in Reiter's theory of diagnosis, Artificial Intelligence 41(1), p. 79-88, 1989.

Hamscher W., Modeling digital circuits for troubleshooting, Artificial Intelligence 51(1-3), p. 223-271, 1991a.

Hamscher W., ACP: Reason maintenance and inference control for constraint propagation over intervals, Proceedings AAAI-91, Anaheim, CA, p. 506-511, 1991b.

Hamscher W., Console L., de Kleer J. (eds.) Readings in Model-Based Diagnosis, Morgan Kaufmann, San Mateo, CA, 1992.

Holtzblatt L.J., Neiberg M.J., Piazza R.L., Vilain M.B., Temporal methods: Multi-dimensional modeling of sequential circuits, Annals of Mathematics and Artificial Intelligence, special issue on Model-Based Diagnosis, 11(1-4), J.C. Baltzer AG, p. 399-413, 1994.

Hou A., A theory of measurement in diagnosis from first principles, Artificial Intelligence 65(2), p. 281-328, 1994.

Hunt J., Pugh D., Price C., Failure mode effects analysis: A practical application of functional modeling, Journal of Applied Artificial Intelligence 9(1), p. 33-44, 1995.

International Workshop on Principles of Diagnosis, 0 to 12th : DX'89, Paris, DX'90, Stanford, CA, DX'91, Milan, DX'92, Rosario, WA, DX'93, Aberystwyth, Wales, DX'94, New Paltz, NY, DX'95, Goslar, DX'96, Val Morin, Québec, DX'97, Mont-Saint-Michel, DX'98, Cap Code, MA, DX’99, Loch Awe, Scotland, DX’00, Morelia, Mexico, DX’01, Sansicario, Italy, proceedings, 1989-2001.

Katsillis G., Chantler M.J., Can dependency-based diagnosis cope with simultaneous equations?, working notes of the 8th International Workshop on Principles of Diagnosis DX’97, Mont Saint Michel, France, 1997.

Kay H., Rinner B., Kuipers B., Semi-quantitative system identification, Artificial Intelligence 119(1-2), p. 103-140, 2000.

Kurien J., Nayak P., Back to the future for consistency-based trajectory tracking, Proceedings AAAI-2000, Austin, TX, p. 370-377, 2000a.

Kurien J., Nayak P., Model-based Monitoring, Diagnosis and Control, Thesis Proposal, Brown University Department of Computer Science, 2000b.

Laborie P., Krivine J.-P., GEMO: A model-based approach for an alarm processing function in power distribution networks, Proceedings of International Conference on Intelligent System Application to Power Systems ISAP-97, Séoul, 1997a.

Laborie P., Krivine J.-P., Automatic generation of chronicles and its application to alarm processing in power distribution systems, working notes of the 8th International Workshop on Principles of Diagnosis DX’97, Mont Saint Michel, France, 1997b.

Lamperti G., Pogliano P., Event-based reasoning for short circuit diagnosis in power transmission networks, Proceedings IJCAI-97, Nagoya, p. 446-451, 1997.

Lamperti G., Zanella M., Diagnosis of discrete-event systems integrating synchronous and asynchronous behavior, working notes of the 10th International Workshop on Principles of Diagnosis DX-99, Loch Awe, Scotland, 1999.

Lamperti G., Zanella M., Uncertain temporal observations in diagnosis, Proceedings ECAI-2000, Berlin, p. 151-155, 2000a.

Lamperti G., Zanella M., Generation of diagnostic knowledge by discrete-event model compilation, Proceedings KR-2000, Breckenridge, CO, p. 333-344, 2000b.

Lamperti G., Zanella M., Principles of distributed diagnosis of discrete-event systems, working notes of the 12th International Workshop on Principles of Diagnosis DX-01, Sansicario, Italy, 2001.

Loiez P., Contribution au diagnostic de systèmes analogiques, Thèse de l'Université des Sciences et Technologies de Lille, 3 Mars 1997.

Loiez E., Taillibert P., Polynomial Temporal Band Sequences for Analog Diagnosis, Proceedings of IJCAI-97, Nagoya, p. 474-479, August 1997.

Malik A., Struss P., Diagnosis of dynamic systems does not necessarily require simulation, working notes of the 7th International Workshop on Principles of Diagnosis DX-96, Val Morin, Canada, 1996.

McIlraith S., Towards a theory of diagnosis, testing and repair, working notes of the 5th International Workshop on Principles of Diagnosis DX-94, New Paltz, NY, 1994.

McIlraith S., Biswas G., Clancy D. and Gupta V., Toward Diagnosing Hybrid Systems, in Proceedings of DX-00, 2000.

McIlraith S., Reiter R., On tests for hypothetical reasoning, Readings in Model-Based Diagnosis, Morgan Kaufmann, p. 89-96, 1992.

Montmain J., Des modèles pour la supervision, mémoire d'habilitation à diriger des recherches, ENSIEG, Grenoble, 2000.

P.J. Mosterman, G. Biswas, A comprehensive methodology for building hybrid models of physical systems, Artificial Intelligence 121:171-209, 2000.

Mozetic I., Hierarchical model-based diagnosis, International Journal of man-Machine Studies 35(3), p. 329-362, 1991.

Nayak P.P., Automated modeling of physical systems, PhD dissertation, Springer, LNAI 1003, 1995.
Nayak P. P., Williams B. C., Fast Context Switching in Real-time Propositional Reasoning. In Proceedings of AAAI-97 (best paper prize winner), Providence, Rhode Island, p. 50-56, July 1997.

Nejdl W., Bachmayer J., Diagnosis and repair iteration planning versus n-step look ahead planning, working notes of the 4th International Workshop on Principles of Diagnosis DX-93, Aberystwyth, Wales, 1993.

Panati A., Theseider Dupré D., State-based vs simulation-based diagnosis of dynamic systems, Proceedings ECAI-2000, Berlin, p. 176-180, 2000.

Pell B., Bernard D. E., Chien S. A., Gat E., Muscettola N., Nayak P. P., Wagner M. D., Williams B. C., A Remote Agent Prototype for Spacecraft Autonomy, Proceedings of the SPIE Conference on Optical Science, Engineering, and Instrumentation, 1996.

Pencolé Y., Decentralized diagnoser approach: application to telecommunication networks, working notes of the 11th International Workshop on Principles of Diagnosis DX-00, Morelia, Mexico, 2000.

Pencolé Y., Cordier M.-O., Rozé L., Incremental decentralized diagnosis approach for the supervision of a telecommunication network, working notes of the 12th International Workshop on Principles of Diagnosis DX-01, Sansicario, Italy, 2001.

Porcheron M., Ricard B., Busquet J.-L., Parent P., DIAPO: A case study in applying advanced AI techniques to the diagnosis of a complex system, Proceedings ECAI-94, Stockholm, p. 43-47, 1994.

Porcheron M., Techniques d'intelligence artificielle pour le diagnostic automatique de systèmes industriels, mémoire d'habilitation à diriger des recherches, université Paris 13, 2001.

Provan G., Modeling and diagnosing rocket engines: A multimodeling approach, working notes of the 9th International Workshop on Principles of Diagnosis DX-98, Cap Code, MA, 1998.

Provan G., Model-Based Diagnostics and Recovery: an Integrated Approach, working notes of the 10th International Workshop on Principles of Diagnosis DX-99, Loch Awe, Scotland, 1999.

Provan G., Hierarchical model-based diagnosis, working notes of the 12th International Workshop on Principles of Diagnosis DX-01, Sansicario, Italy, 2001.

Raiman O., de Kleer J., Saraswat V., Shirley M., Characterizing non-intermittent faults, Proceedings AAAI-91, Anaheim, CA, p. 849-854, 1991.

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

Rinner B., From tracking continuous mode hypotheses to diagnosing technical systems, working notes of the 12th International Workshop on Principles of Diagnosis DX-01, Sansicario, Italy, 2001.

Rinner B., Kuipers B., Monitoring piecewise continuous behaviors by refining semi-quantitative trackers, Proceedings IJCAI-99, Stockholm, p. 1080-1086, 1999.

Rouquette N., Quantitative diagnosis of continuous-valued, steady-state systems, working notes of the 6th International Workshop on Principles of Diagnosis DX-95, Goslar, Germany, 1995.

Rozé L., Supervision of telecommunication network: A diagnoser approach, working notes of the 8th International Workshop on Principles of Diagnosis DX'97, Mont Saint Michel, France, 1997.

Sachenbacher M., Malik A., Struss P., From electrics to emissions: Experiences in applying model-based diagnosis to real problems in cars, working notes of the 9th International Workshop on Principles of Diagnosis DX-98, Cape Code, MA, 1998.

Sachenbacher M., Struss P., AQUA: A framework for automated qualitative abstraction, working notes of the 15th International Workshop on Qualitative Reasoning QR-01, San Antonio, TX, 2001.

Scarl E., Sensor placement for diagnosability, Annals of Mathematics and Artificial Intelligence, special issue on Model-Based Diagnosis, 11(1-4), J.C. Baltzer AG, p. 493-510, 1994.

Schut C., Bredeweg B., An overview of approaches to qualitative model construction, Knowledge Engineering Review 11(1), p. 1-25, 1996.

Shirley M., Davis R., Generating distinguishing tests based on hierarchical models and symptom information, Proceedings of the International Conference on Computer Design: VLSI in Computers, Silver Springs, MD, 1983.

Singh N., SATURN: An automatic test generation system for digital circuits, Proceedings of AAAI-86, Philadelphia, PA, p. 778-783, 1986.

Smith N., A new architecture for automated modelling, Proceedings of AAAI-98, Madison, WI, p. 225-231, 1998.

Stolle R., Bradley E., Multimodal reasoning for automatic model construction, Proceedings of AAAI-98, Madison, WI, p. 181-188, 1998.

Struss P., Diagnosis as a process, working notes of the International Workshop on Model-based Diagnosis, Paris, 1989. Revised version in Readings in Model-Based Diagnosis, Morgan Kaufmann, p. 408-418, 1992.

Struss P., What's in SD? Towards a theory of modeling for diagnosis, working notes of the 2nd International Workshop on Principles of Diagnosis DX'91, Milan, 1991. Also in Readings in Model-Based Diagnosis, Morgan Kaufmann, p. 419-448, 1992.

Struss P., Testing for discrimination of diagnoses, working notes of the 5th International Workshop on Principles of Diagnosis DX'94, New Paltz, NY, 1994.

Struss P., Fundamentals of model-based diagnosis of dynamic systems, Proceedings IJCAI-97, Nagoya, p. 480-485, 1997.

Struss P., Dressler O., Physical negation: Integrating fault models into the general diagnostic engine, Proceedings IJCAI-89, Detroit, MI, p. 1318-1323, 1989.

Struss P., Sachenbacher M., Carlen C., Insights from building a prototype for model-based on-board diagnosis of automotive systems, working notes of the 11th International Workshop on Principles of Diagnosis DX-00, Morelia, Mexico, 2000.

Struss P., Sachenbacher M., Dummert F., Diagnosing a dynamic system with (almost) no observations, working notes of the 11th International Workshop on Qualitative Reasoning QR-97, Cortona, Italie, 1997.

Sun Y., Weld D., A framework for model-based repair, Proceedings AAAI-93, Washington, DC, p.182-187, 1993.

Travé-Massuyès L., Bousson K., Evrard J.-M., Guerrin F., Lucas B., Missier A., Tomasena M., Zimmer L., Non causal vs. causal qualitative modelling and simulation, Intelligent Systems Engineering 2(3), p. 159-182, 1993.

Travé-Massuyès L., Dague P., Guerrin F., Le raisonnement qualitatif pour les sciences de l'ingénieur, Hermès, 1997.

Travé-Massuyès L., Milne R., Application-oriented qualitative reasoning, The Knowledge Engineering Review 10(2), 1995.

Travé-Massuyès L., Milne R., TIGERTM: Gas turbines condition monitoring using qualitative model based diagnosis, IEEE Expert Intelligent Systems & Applications 12(3), p. 21-31, 1997.

L. Travé-Massuyès, Milne R., Gaps between research and industry related to model based and qualitative reasoning, In the proceedings of the ECAI'98 workshop on "Model Based Systems and Qualitative Reasoning", Brighton (UK, August 23-28, 1998.

Travé-Massuyès L., Pons R., Causal ordering for multiple mode systems, 11th Int. Workshop on Qualitative Reasoning about Physical Systems, Cortona, Italy, 1997.

Vilain M., Kautz H., Constraint Propagation Algorithms for Temporal Reasoning, Proceedings AAAI-86, Philadelphia, PA, p. 377-382, 1986.

Vilain M., Kautz H., Van Beek P., Constraint Propagation Algorithms for Temporal Reasoning: A Revised Report, Readings in Qualitative Reasoning About Physical Systems, Morgan Kaufmann, p. 373-381, 1990.

Washio T., Hotoda H., Discovering admissible simultaneous equations of large scale systems, in Proc. of the 15th National Conference on Artificial Intelligence AAAI-98, Madison, WI, p. 189-196, 1998.

Weld D., de Kleer J. (eds.), Readings in Qualitative Reasoning About Physical Systems, Morgan Kaufmann, 1990.

Williams B. C., Doing Time: Putting Qualitative Reasoning on Firmer Ground, Proceedings AAAI-86, Philadelphia, PA, p. 105-112, 1986.

Williams B. C., Nayak P. P., Immobile Robots: AI in the New Millennium, AI Magazine, Vol. 17, 3, p. 16-35, Fall 1996.

Williams B. C., Nayak P. P., A Model-based Approach to Reactive Self-Configuring Systems, Proceedings of AAAI-96, Portland, Oregon, p. 971-978, August 1996.

Williams B. C., Millar W., Automated Decomposition of Model-based Learning Problems, Proceedings of the Tenth International Workshop on Qualitative ReasoningQR-96, Fallen Leaf Lake, California, AAAI Technical Report WS-96-01, p. 265-273, 1996.

Williams B. C., Nayak P. P., A Reactive Planner for a Model-based Executive, Proceedings of IJCAI-97, Nagoya, p. 1178-1185, August 1997.

Williams B. C., Millar W., Decompositional, Model-based Learning and its Analogy to Diagnosis, Proceedings of AAAI-98, Madison, Wisconsin, p. 197-204, July 1998.

Williams B.C., Gupta V., Unifying Model-based and Reactive Programming within a Model-based Executive, Proceedings of the Tenth International Workshop on Principles of Diagnosis DX-99, 1999.

Williams B.C., J. Kurien, Back to the future for Consistency-based Trajectory Tracking, Proceedings of AAAI-2000, Austin, Texas, 2000.

Williams B.C., S.H. Chung and J.M.V. Eepoel, Improving Model-Based Mode Estimation through off-line Compilation, Proceedings of ISAIRAS-2001.

Xia S., Linkens D.A., Bennett S., Automatic modeling and analysis of dynamical physical systems using qualitative reasoning and bond graphs, Intelligent Systems Engineering Journal, p. 201-212, 1993.

Model-Based Diagnosis in Control

Aghasaryan A., Fabre E., Benveniste A., Boubour R., Jard C., Fault detection and diagnosis in distributed systems: an approach by partially stochastic Petri nets, Journal of Discrete Event Dynamical Systems, special issue on hybrid systems, 8(2), p. 203-231, 1998.

Armengol J., Vehi J., Travé-Massuyès L., Sainz M.A., Generation of error-bounded envelopes by means of Modal Interval Analysis, Workshop on Applications of Interval Analysis to Systems and Control with special emphasis on recent advances in Modal Interval Analysis (MISC 1999), p. 251-261, Girona, Catalonia, Spain, 1999.

Armengol J., Vehi J., Travé-Massuyès L., Sainz M.A., Interval model-based fault detection using multiple sliding time windows, 4th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS 2000), p. 168 -173. Budapest, Hungary, 2000.

Ayoubi M., Fuzzy systems design based on hybrid neural structure and application to the fault diagnosis of technical processes, Control Engineering Practice, Vol. 4 (1), 35-42, 1996.

Azor R., I.Y. Bar-Itzhack, R.R. Harman, Satellite Angular Rate Estimation from Vector Measurements, Journal of Guidance, Control, and Dynamics, Vol. 21, 3, p. 450-457, May-June 1998.

Basseville M., Benveniste A., Moustadikes G., Rougée A., Optimal sensor location for detecting changes in dynamical behavior, IEEE Transactions on Automatic Control, AC 32(12), p. 1067-1075, December 1987.

Basseville M., Nikiforov , I.V., Detection of abrupt changes - Theory and application, Prentice Hall, Information and System Sciences Series, 1993.

Carpentier T., Litwak R., Cassar, J-P., Criteria for the evaluation of F.D.I. systems application to sensors location. In Proceedings IFAC Workshop SAFEPROCESS Hull, UK, 1997.

Cassa, J.P., Staroswiecki M., A structural approach for the design of failure detection and identification systems, Proceedings IFAC, IFIP, IMACS Conference on Control of Industrial Systems, p. 329-334, Belfort, 1997.

Chantler M.J., Daus S., Vikatos T., Coghill G.M., The use of quantitative dynamic models and dependency recording for diagnosis, working notes of the 7th International Workshop on Principles of Diagnosis DX-96, Val Morin, Canada, 1996.

Clark R .N, A simplified instrument detection scheme, IEEE Trans. Aerospace Electron. Syst. 14(3), p. 558-563, 1978.

Clark R .N, Instrument fault detection, IEEE Trans. Aerospace Electron. Syst. 14(3), p. 456-465, 1978.

Control Engineering Practice (CEP), Special volume on Supervision, fault detection, and diagnosis of technical systems 5(5), 1997.

Dubuisson B. (ed.) Automatique et statistiques pour le diagnostic, traité IC2, Hermes, 2001b.

Escobet T., Travé-Massuyès E., Parameter estimation methods for fault detection and isolation, Bridge Workshop Notes, Sansicario (I), March 5-9 2001.

Fabre E., Benveniste A., Jard C., Ricker L., Smith M., Distributed state reconstruction for discrete event systems, Proceedings of the 2000 IEEE Control and Decision Conference (CDC), Sydney, 2000.

Frank P.M., Advanced fault detection and isolation schemes using nonlinear and robust observers, 10th IFAC Congress, Munchen, Vol. 3, 63-68, 1987.

Frank P.M., Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy, Automatica, Vol. 26, 459-474, 1990.

Frank P.M., Analytical and qualitative model-based fault diagnosis - A survey and some new results, European Journal of Control, Vol. 2, 6-28, 1996.

Gertler J.J., Analytical redundancy methods in fault detection and isolation. Proc. of TOOLDIAG Conference, Toulouse, 1993.

Gertler J.J., Fault Detection and Diagnosis in Engineering Systems, Marcel Dekker, 1998.

Gertler J.J., Fault detection and diagnosis in engineering systems, Marcel Dekker Inc., ISBN :0-8247-9427-3, 1998.

Gissinger G.L, Loung M., Reynaund H.-F., Failure Detection and Isolation - Optimal design of instrumentation system, Proceedings of the IFAC Workshop SAFEPROCESS 2000.

Isermann R., Process fault diagnosis based on process model knowledge, Proc. of Advanced Information Processing in Automatic Control, AIPAC, Nancy, France, 1989.

Isermann R., Supervision, fault detection and fault-diagnosis methods- An introduction, Control Engineering Practice 5(5), 639-652, 1997.

Isermann, R. On fuzzy logic application for automatic control, supervision and fault detection, Third European Congress on Fuzzy and Intelligent Technologies EUFIT'95, Aachen, Vol. 2, 738-753, 1995.

Karnopp D.C., Rosenberg R.C., Analysis and simulation of multiport systems, MIT Press, 1968.

Lunze J., Schröder J., Supavatanakul P., Diagnosis of discrete-event systems: the method and an example, working notes of the 12th International Workshop on Principles of Diagnosis DX-01, Sansicario, Italy, 2001.

Patton R.J., Frank P.M., Clark R.N., (Eds.) Fault detection in dynamic systems - Theory and applications, Prentice Hall International, 1989.

Patton R.J., Chen J., A review of parity space approaches to fault diagnosis, IFAC SAFEPROCESS Symposium, Baden-Baden, 1991.

Ploix S., Adrot O., Ragot J., "Bounding approach to the diagnosis of a class of uncertain static systems". IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS'2000, Budapest, Hungary, June 14-16, 2000.

Rosembrock H.H., State-space and multivariable theory, John Wiley and Sons,1970.

Rosenberg R.C., Karnopp D.C., Introduction to physical systems dynamics, McGraw-Hill, 1983.

Sampath M., Lafortune S., Teneketzis D.C., Active diagnosis of discrete-event systems, IEEE Trans. Autom. Control 43(7), p. 908-929, 1998.

Sampath M., Sengupta R., Lafortune S., Sinnamohideen K., Teneketzis D.C., Diagnosability of discrete-event systems, IEEE Trans. Autom. Control 40(9), p. 1555-1575, 1995.

Sampath M., Sengupta R., Lafortune S., Sinnamohideen K., Teneketzis D.C., Failure diagnosis using discrete-event models, IEEE Trans. Control Systems Technology 4(2), p. 105-124, 1996.

Staroswiecki, M. and Ph. Declerck, Analytical Redundancy in Non-linear Interconnected Systems by Means of Structural Analysis, in Proc. of the IFAC Symposium on Advanced Information Processing in Automatic Control, (AIPAC'89), Nancy, II, 23-27, 1989.

Staroswiecki, M., J. Ph. Cassar and V. Cocquempot, Generation of Optimal Structured Residuals in the Parity Space, Proceedings of the 12th IFAC World Congress, Sydney, 1993.

Travé-Massuyès L. Escobet T., Milne R., Model based Diagnosability and sensor placement, Proceedings of IJCAI-01, Seattle, WA, 2001.

Van der Velde W.E.,C.R. Carignan C. R., Number and placement of control system components considering possible failures, Journal of Guidance, Control, Dynamics, 7(6), p. 703-709, 1984.

Washio T., Sakuma M., Kitamura M., A new approach to quantitative and credible diagnosis for multiple faults of components and sensors, Artificial Intelligence 91(1), p. 103-110, 1997.

BRIDGE model-based diagnosis publications

Basseville M., Cordier M.O., Surveillance et diagnostic de systèmes dynamiques: approches complémentaires du traitement du signal et de l'intelligence artificielle, Publication IRISA n° 1004, 1996.

Cordier M.-0., Dague P., Dumas M., Levy F., Montmain J., Staroswiecki M., Travé-Massuyès L., AI and Automatic Control Approaches of Model-Based Diagnosis : Links and Underlying Hypotheses, IFAC Symposium SAFEPROCESS 2000, Budapest, Hungary, pp. 274-279, June 2000.

Cordier M.-0., Dague P., Dumas M., Levy F., Montmain J., Staroswiecki M., Travé-Massuyès L., A comparative analysis of AI and control theory approches to model-based diagnosis, International Workshop on Principles of Diagnosis DX'00, Morelia, Mexico, pp. 33-40, June 8-10, 2000.

Cordier M.-0., Dague P., Dumas M., Levy F., Montmain J., Staroswiecki M., Travé-Massuyès L., A comparative analysis of AI and control theory approches to model-based diagnosis, ECAI '00, Berlin, Germany, August 20-25.

Krysander M., Nyberg M. (2002), Structural analysis utilizing MSS sets with application to a paper plant, 13th International Workshop on Principles of Diagnosis (DX'02), Semmering (Austria), 2-4 May 2002, pp.51-57.

Pulido Junquera B., Alonso Gonzalez C. (2002), Possible conflicts, ARRs, and conflicts, 13th International Workshop on Principles of Diagnosis (DX'02), Semmering (Austria), 2-4 May 2002, pp.51-57.

9 Document History

Version	*Date*	*Changes made to document*	*Changed by*
1.1	30th January 2003	Task Group produced the Document Edited by Louise Travé-Massuyès	Bridge TG
1.2	2nd October 2003	Updated to standard format. No change to text thus date not changed	RIR